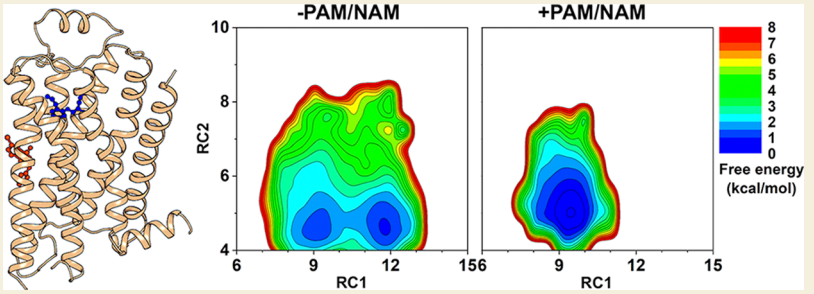
JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau
Article

Deep Learning Dynamic Allostery of G-Protein-Coupled Receptors

Hung N. Do, Jinan Wang, and Yinglong Miao*

Cite This: JACS Au 2023, 3, 3165−3180
Read Online
ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: G-protein-coupled receptors (GPCRs) make up the largest superfamily of human membrane proteins and represent primary targets of ∼1/3 of currently marketed drugs. Allosteric modulators have emerged as more selective drug candidates compared with orthosteric agonists and antagonists. However, many X-ray and cryo-EM structures of GPCRs resolved so far exhibit negligible differences upon the binding of positive and negative allosteric modulators (PAMs and NAMs). The mechanism of dynamic allosteric modulation in GPCRs remains unclear. In this work, we have systematically mapped dynamic changes in free energy landscapes of GPCRs upon binding of allosteric modulators using the Gaussian accelerated molecular dynamics (GaMD), deep learning (DL), and free energy prOfiling Workflow (GLOW). GaMD simulations were performed for a total of 66 μs on 44 GPCR systems in the presence and absence of the modulator. DL and free energy calculations revealed significantly reduced dynamic fluctuations and conformational space of GPCRs upon modulator binding. While the modulator-free GPCRs often sampled multiple low-energy conformational states, the NAMs and PAMs confined the inactive and active agonist-G-protein-bound GPCRs, respectively, to mostly only one specific conformation for signaling. Such cooperative effects were significantly reduced for binding of the selective modulators to "non-cognate" receptor subtypes. Therefore, GPCR allostery exhibits a dynamic "conformational selection" mechanism. In the absence of available modulator-bound structures as for most current GPCRs, it is critical to use a structural ensemble of representative GPCR conformations rather than a single structure for compound docking ("ensemble docking"), which will potentially improve structure-based design of novel allosteric drugs of GPCRs.

KEYWORDS: *G-protein-coupled receptors (GPCRs), allostery, Gaussian accelerated Molecular Dynamics (GaMD), deep learning, conformational selection, drug design*

INTRODUCTION

G-protein-coupled receptors (GPCRs) are the largest superfamily of human membrane proteins with >800 members. GPCRs play key roles in cellular signaling and mediate various physiological activities, including vision, olfaction, taste, neurotransmission, endocrine, and immune responses.[1] They represent primary targets of ∼1/3 of currently marketed drugs.[2] GPCRs can be classified into six different classes, including class A (Rhodopsin-like), B (secretin receptors), C (metabotropic glutamate receptors (mGluRs)), D (fungal mating pheromone receptors), E (cyclic AMP receptors), and class F (frizzled/TAS2 receptors).[3,4] GPCRs share a characteristic structural fold of seven transmembrane (TM) helices (TM1-TM7) connected by three extracellular loops (ECL1-ECL3) and three intracellular loops (ICL1-ICL3). For decades, the primary endogenous agonist-binding ("orthosteric") site has been targeted for drug design of GPCR agonists, antagonists, and inverse agonists.[5] However, the orthosteric site is usually highly conserved in different subtypes of GPCRs. An orthosteric drug often binds and activates/deactivates multiple GPCRs simultaneously with poor selectivity, thereby causing toxic side effects.[6]

Alternatively, allosteric modulators have been discovered to bind topographically distant ("allosteric") sites of GPCRs with advantages.[7−12] They are able to modulate the binding affinity and signaling of orthosteric ligands, including positive and negative allosteric modulators (PAMs and NAMs).[13] The allosteric effect has been shown to depend on the orthosteric

Received: August 26, 2023
Revised: October 10, 2023
Accepted: October 18, 2023
Published: November 2, 2023

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

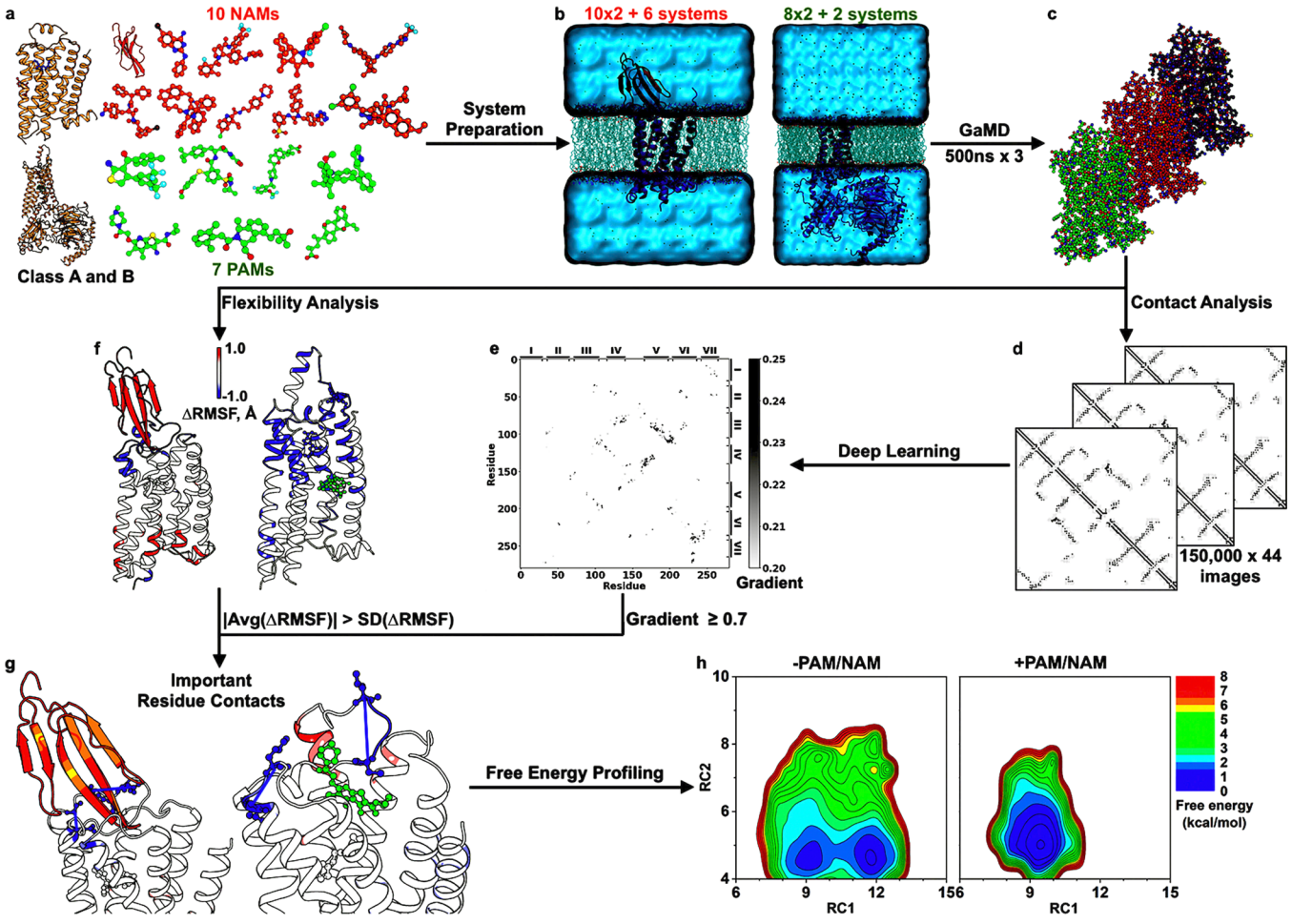


Figure 1. Workflow of DL dynamic allostery of GPCRs. Starting from 10 NAMs, seven PAMs, 18 different experimental structures, and eight computational models of class A and B GPCR-NAM/PAM complexes (a), 2 × 10 structural +6 model simulation systems of inactive antagonist-bound GPCRs in the presence/absence of NAM and 2 × 8 structural +2 model simulation systems of active agonist-bound GPCRs in the presence/absence of PAM were built (b). Three independent 500 ns GaMD simulations were performed on each system (c). Residue contact maps were calculated for 150,000 × 44 GaMD simulation frames (d) and analyzed by Deep Learning, yielding saliency (attention) maps of residue contact gradients (e). Changes in root-mean-square fluctuations (ΔRMSFs) upon NAM/PAM binding in GPCRs were calculated from the GaMD simulations (f). If the absolute average ΔRMSF calculated from three simulations of a residue was smaller than the standard deviation of ΔRMSF, the flexibility change for that residue was considered not significant and related residue pairs were neglected for further analysis. The characteristic residue contacts selected were those with ≥0.7 gradients and significant flexibility changes upon modulator binding (g). They served as reaction coordinates for free energy profiling of dynamic allostery of GPCRs (h).

probe,[14] with a "ceiling level" determined by the magnitude and direction of cooperativity between the orthosteric and allosteric ligands. Because the allosteric site is usually more divergent in residue sequences and conformations, allosteric modulators offer higher receptor selectivity in comparison to the orthosteric ligands. They serve as important chemical probes and promising selective therapeutics of GPCRs.

Important insights have been obtained using X-ray crystallography and cryo-electron microscopy (cryo-EM) about structural changes induced by allosteric modulator binding in certain GPCRs.[7,9,10,15] For class A GPCRs, binding of the LY2119620 PAM in the M_2 muscarinic receptor (M_2R) led to side chain rotation of residue W7.35 and slight contraction of the receptor extracellular pocket, which was preformed in the active agonist-bound structure.[16,17] GPCR residues are numbered according to the Ballesteros-Weinstein scheme[18] for class A and Wootten scheme[19] for class B GPCRs. Binding of the muscarinic toxin MT7 NAM to the antagonist bound M_1 muscarinic receptor (M_1R) resulted in conformational changes in the ECL2, TM1, TM2, TM6 and TM7 extracellular domains, as well as the TM2 and TM6 intracellular domains.[20] In the free fatty acid receptor GPR40 (FFAR1), AgoPAM binding in a lipid-facing pocket formed by TM3-TM4-ICL2 induced conformational changes in the ICL2, TM4 and TM5 of the active receptor.[21] The ICL2 adopted a short helical conformation and the TM5 was shifted along its helical axis toward the extracellular side relative to the TM4.[21] A similar allosteric site was identified for binding of the NDT9513727 and Avacopan NAMs between TM3-TM4-TM5 on the lipid-exposed surface of the $C5a_1$ receptor (C5AR1).[22] For class B GPCRs, the LSN3160440 PAM was found to bind between the extracellular domains of TM1 and TM2 of the GLP-1 receptor (GLP1R).[23] In the glucagon receptor (GLR), NAM binding outside of the 7TM bundle between TM6-TM7 restricted the outward movement of the TM6 intracellular domain required for activation and G-protein coupling of the receptor.[24] The ECL2 stretched to the central axis of the TM helical bundle, allowing for interactions from TM3 to TM6 and TM7 in the inactive class B GPCRs.[24] Despite remarkable

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

advances, the X-ray and cryo-EM structures represent static snapshots of GPCRs. Many GPCR structures exhibit small/negligible differences in the absence and presence of allosteric modulators, notably for the A_1AR,[25] M_2R,[16] M_4R,[26,27] β_2-adrenoceptor (β_2AR),[28−30] C5AR1,[31] CB_1 cannabinoid receptor (CB_1),[32] chemokine receptor CCR2,[33] dopamine receptor 1 (D_1R),[34] GPBA receptor (GPBAR),[35] and GLP1R.[23,36,37] A dynamic review has been suggested for allosteric modulation of GPCRs.[38] However, the dynamic mechanism of GPCR allostery remains unclear.

Molecular dynamics (MD) is a powerful computational technique for simulating biomolecular dynamics on an atomistic level.[39] For GPCRs, MD has been applied to simulate binding of both orthosteric and allosteric ligands.[40−45] Binding of known NAMs to the M_2R was observed in conventional MD (cMD) simulations using the specialized supercomputer Anton.[46] The modulators formed cation-π interactions with aromatic residues in the receptor extracellular vestibule, which was confirmed by mutation experiments and later by X-ray structure of M_2R recognized by a PAM.[16] Microsecond-time scale cMD simulations revealed mechanistic insights into allosteric modulation by Na^+ in dopamine and opioid receptors.[47,48] Accelerated MD (aMD) simulations also captured Na^+ binding to the highly conserved D2.50 allosteric site, which stabilized a muscarinic GPCR in the inactive state.[49] Recently, spontaneous binding of prototypical PAMs to the putative ECL2 allosteric site of the A_1AR was captured in Gaussian accelerated molecular dynamics (GaMD) simulations.[50] Moreover, metadynamics simulations captured binding of the BMS-986187 PAM to the δ-opioid receptor.[51] Metadynamics and GaMD enhanced sampling simulations revealed positive binding cooperativity between allosteric and orthosteric ligands of the CCR2.[52] Despite these exciting advances, MD simulations of allosteric modulation have been limited to mostly few selected class A GPCRs.[45]

Recently, we have developed the GaMD, deep learning (DL), and free energy profiling workflow (GLOW) to predict molecular determinants and map free energy landscapes of biomolecules.[53] GaMD is an unconstrained enhanced sampling technique that works by applying a harmonic boost potential to smooth the biomolecular potential energy surface.[54] Since this boost potential usually exhibits a near Gaussian distribution, cumulant expansion to the second order ("Gaussian approximation") can be applied to achieve proper energy reweighting.[55] GaMD allows for simultaneous unconstrained enhanced sampling and free energy calculations of large biomolecules.[54] GaMD has been successfully demonstrated on enhanced sampling of ligand binding, protein folding, protein conformational changes, protein−membrane/peptide/protein/nucleic acid/carbohydrate interactions.[56] In GLOW, DL of image-transformed residue contact maps calculated from GaMD simulation frames allows us to identify important residue contacts by classic gradient-based pixel attribution in the saliency (attention) maps.[53,57] Finally, free energy profiles of these residue contacts are calculated through reweighting of GaMD simulations to characterize the biomolecular systems of interest.[53]

In this work, we applied GLOW to systematically map dynamic changes in free energy landscapes of GPCRs upon binding of allosteric modulators (Figure 1). A total of 18 different high-resolution experimental structures of class A and B GPCRs are collected for modeling and 8 computational models were generated by changing target receptors of the modulators to different subtypes. Our comprehensive DL analysis of extensive GaMD simulations has provided important mechanistic insights into the dynamic allostery of GPCRs.

METHODS

Setup of GPCR Simulation Systems

A total of 18 unique experimental structures and eight computational models of allosteric modulator-bound class A and B GPCRs were prepared for simulations (Figure 1a and Table S1). The GPCR structures bound by NAMs included the MT7-bound M_1R (PDB: 6WJC),[20] Cmpd-15-bound β_2AR (PDB: 5X7D),[28] AS408-bound β_2AR (PDB: 6OBA),[30] NDT9513727-bound C5AR1 (PDB: 6C1Q),[31] Avacopan-bound C5AR1 (PDB: 6C1R),[31] ORG27569-bound CB_1 receptor (PDB: 6KQI),[32] GTPL9431-bound CCR2 (PDB: 5T1A),[33] NNC0640-bound GLP1R (PDB: 5VEX),[36] PF-06372222-bound GLP1R (PDB: 6LN2),[37] and MK-0893-bound GLR (PDB: 5EE7).[24] The GPCR structures bound by only antagonists without NAMs included the 5D5B[58] (β_2AR without Cmpd-15), 6PRZ[59] (β_2AR without AS408), and 7V3Z[60] (CB_1 without ORG27569) PDB structure. Other GPCR structures bound by only antagonists were obtained by removing the NAMs from the corresponding NAM-bound GPCR structures. Six computational models of NAM-bound GPCRs included the MT7-bound M_2R and M_4R, which were built by aligning the 6WJC PDB structure of M_1R to the 5ZK3[61] and 5DSG[62] PDB structures of M_2R and M_4R, respectively, and copying atomic coordinates of the atropine antagonist and MT7 NAM, as well as the Cmpd-15-bound α_{1B}-adrenoceptor (α_{1B}AR), α_{2A}-adrenoceptor (α_{2A}AR), α_{2C}-adrenoceptor (α_{2C}AR), and β_1-adrenoceptor (β_1AR), which were built by aligning the 5X7D PDB structure of β_2AR to the 7B6W,[63] 6KUX,[64] 6KUW,[65] and 7BVQ[66] PDB structures of α_{1B}AR, α_{2A}AR, α_{2C}AR, and β_1AR, respectively, and copying atomic coordinates of the carazolol antagonist and Cmpd-15 NAM. The GPCR structures bound by PAMs included the MIPS521-bound A_1AR (PDB: 7LD3),[25] LY2119620-bound M_2R (PDB: 6OIK),[17] LY2119620-bound M_4R (PDB: 7V68),[26] Cmpd-6FA-bound β_2AR (PDB: 6N48),[29] LY3154207-bound D_1R (PDB: 7LJC),[34] AgoPAM-bound FFAR1 (PDB: 5TZY),[21] INT777-bound GPBAR (PDB: 7CFN),[35] and LSN3160440-bound GLP1R (PDB: 6VCB).[23] The GPCR structures bound by only agonists without PAM included the 7LD4[25] (A_1AR without MIPS521), 7TRK[27] (M_4R without LY2119620), 6E67[67] (β_2AR without Cmpd-6FA), 7JV5[68] (D_1R without LY3154207), and 5TZR[21] (FFAR1 without AgoPAM). Other GPCR structures bound by only agonists were obtained by removing the PAMs from the corresponding PAM-bound GPCR structures. Two computational models of PAM-bound GPCRs included LY2119620-bound M_1R, which was built by aligning the 6OIK PDB structure of M_2R to the 6OIJ PDB structure of M_1R[17] and copying atomic coordinates of the LY2119620 PAM, and LY32154207-bound D_2 receptor (D_2R), which was built by aligning the 7LJC PDB structure of D_1R to the 7JVR PDB structure of D_2R[68] and copying atomic coordinates of the SKF-81297 agonist and LY3154207 PAM.

SWISS-MODEL[69] homology modeling was applied to restore missing residues in the GPCR structures and models, particularly in the ECL2, ICL2, and ECL3. Charges of the ligands are listed in Table S1. All water and heteroatom molecules except the ligands and receptor-bound ions (including the sodium ion in the 6C1R PDB structure of C5AR1 and zinc ion in the 6LN2 PDB structure of GLP-1 receptor) were removed from the structures. The GPCR complexes were embedded in POPC membrane lipid bilayers and solvated in 0.15 M NaCl (Figure 1b). The AMBER[70] force field parameter sets were used for our GaMD simulations, specifically ff19SB[71] for proteins, GAFF2[72] for ligands using the AM1-BCC[73] charging method, LIPID17 for lipids, and TIP3P[74] for water, except for the A_1AR and iperoxo-bound M_4R simulations, which were obtained from previous studies[25,27] where the CHARMM36m[75] force field and CGenFF[76,77] parameter set was used. The ligand force field parameters were included in Supporting Data 2. Previous studies comparing CHARMM and AMBER force field parameter sets showed minor differences in the

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

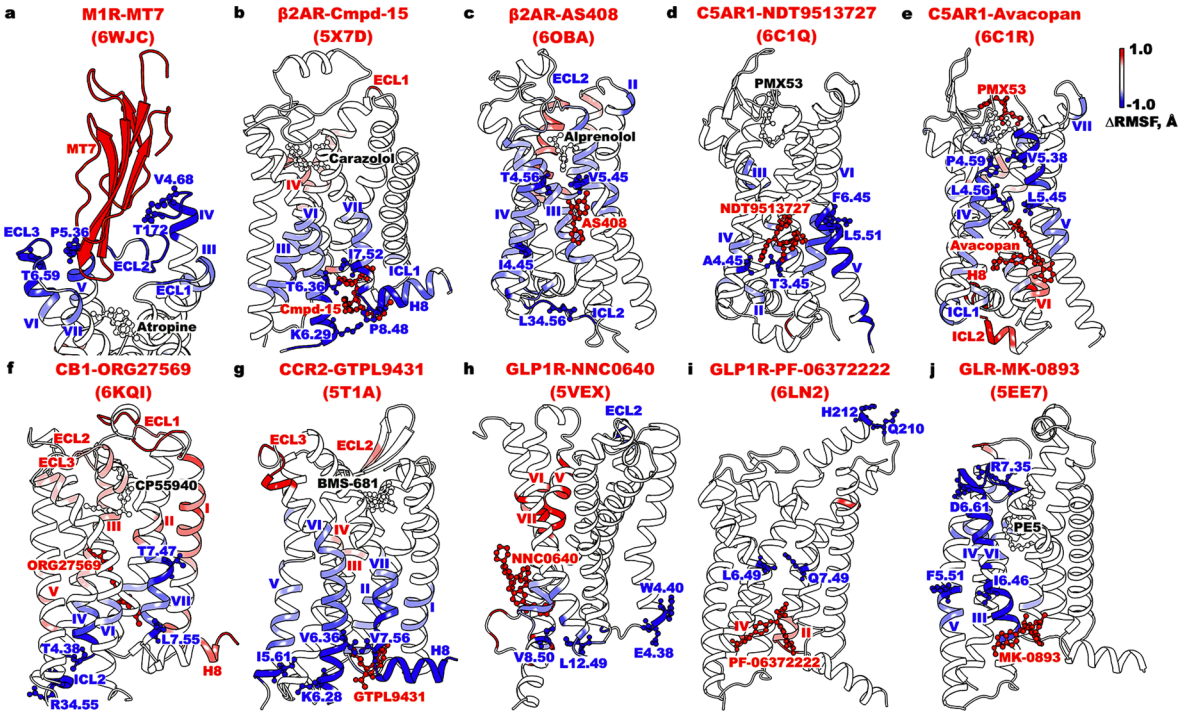


Figure 2. Characteristic residue contacts in the negative allosteric modulation of class A and B GPCRs calculated from GaMD simulations of the MT7-bound M_1R (PDB: 6WJC) (a), Cmpd-15-bound β_2AR (PDB: 5X7D) (b), AS408-bound β_2AR (PDB: 6OBA) (c), NDT9513727-bound C5AR1 (PDB: 6C1Q) (d), Avacopan-bound C5AR1 (PDB: 6C1R) (e), ORG27569-bound CB_1 (PDB: 6KQI) (f), GTPL9431-bound CCR2 (PDB: 5T1A) (g), NNC0640-bound GLP1R (PDB: 5VEX) (h), PF-06372222-bound GLP1R (PDB: 6LN2) (i), and MK-0893-bound GLR (PDB: 5EE7) (j). The seven TM helices are labeled I−VII, H8 for helix 8, ECL1-ECL3 for extracellular loops 1−3, and ICL1-ICL3 for intracellular loops 1−3. A color scale of −1.0 (blue) to 0 (white) to 1.0 (red) is used to show the ΔRMSF upon NAM binding, and NAMs are colored orange.

dynamic behaviors of the simulation systems between the two force fields, and both CHARMM and AMBER force fields could reproduce experimental data.[78−81]

Simulation Protocols

All-atom dual-boost GaMD simulations[54] were performed on the GPCR structures and models with and without allosteric modulators (Figure 1c and Table S1). The simulations of the A_1AR with and without the MIPS521 PAM were obtained from a previous study.[25] GaMD simulations of the other systems followed a similar protocol. Periodic boundary conditions were applied to the simulation systems. Bonds containing hydrogen atoms were restrained with the SHAKE[82] algorithm, and a 2 fs time step was used. The temperature was kept at 310 K using the Langevin thermostat[83,84] with a friction coefficient of 1.0 ps^{-1}. The pressure was kept constant at 1.0 bar using the Berendsen barostat[85] with semi-isotropic coupling with enabled surface tension in the $X-Y$ plane. The Berendsen coupling constant was set to 0.5 ps. The electrostatic interactions were calculated using the particle mesh Ewald (PME) summation[86] with a cutoff of 9.0 Å for long-range interactions. The systems were energetically minimized for 5000 steps using the steepest-descent algorithm and equilibrated with the constant number, volume, and temperature (NVT) ensemble at 310 K. They were further equilibrated for 375 ps at 310 K with the constant number, pressure, and temperature (NPT) ensemble. The cMD simulations were then performed for 10 ns using the NPT ensemble with constant surface tension at 1 atm of pressure and 310 K temperature. GaMD implemented in GPU version of AMBER 20[54,87,88] was applied to simulate the GPCR systems. The simulations involved an initial short cMD of 4−10 ns to calculate GaMD acceleration parameters and GaMD equilibration of added boost potential for 16−40 ns. Three independent 500 ns GaMD production simulations with randomized initial atomic velocities were performed for each system at the "dual-boost" level, with one boost potential applied to the dihedral energetic term and the other to the total potential energetic term. The reference energy was set to the lower bound $E = V_{max}$, and the upper limit of the boost potential standard deviation, σ_0, was set to 6.0 kcal/mol for both the dihedral and total potential energetic terms. The GaMD simulations are summarized in Table S1.

Deep Learning and Free Energy Profiling of GaMD Simulations with the GLOW Workflow

GLOW[53] was applied to systematically analyze GPCR allostery (Figure 1d,e). Residue contact maps of 6,600,000 GaMD simulation frames obtained from a total of 66 μs GaMD simulations, which were inclusive of three replicas of each GPCR simulation system, were calculated and transformed into images for DL (Figure 1d). A contact definition of ≤4.5 Å between any heavy atoms in the two protein residues was used. For DL, 80% of the residue contact map images were randomly assigned to the training set, while the remaining 20% was put in the validation set for each GPCR. The residue contact map images of each GPCR system were separated into four different classes for DL analysis based on the absence and presence of NAMs and PAMs, including NAM-free ("Antagonist"), NAM-bound ("AntagonistNAM"), PAM-free ("Agonist"), and PAM-bound ("AgonistPAM"). DL models of two-dimensional (2D) convolutional neural networks (CNNs)[53] were built to classify the frame images with and without the allosteric modulator bound for each GPCR subfamily for 15 epochs (Figures S7 and S8). The best-fit 2D-CNN architecture consisted of four convolutional layers of 3 × 3 kernel sizes, with 32, 32, 64, and 64 filters, respectively, followed by three dense layers, the first two of which included 512 and 128 filters with a dropout rate of 0.5 each.[53] The final dense layer was the classification layer.[53] "ReLu" activation was used for all layers in the 2D-CNN, except the final dense layer, where "softmax" activation was used.[53] A maximum pooling layer of 2 × 2 kernel size was added after each convolutional layer.[53] Saliency (attention) maps of residue contact gradients were calculated through backpropagation by vanilla gradient-based pixel attribution[57] using the residue contact map of the most populated structural cluster of each GPCR system (Figure 1e). The hierarchical agglomerative clustering algorithm was used to

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

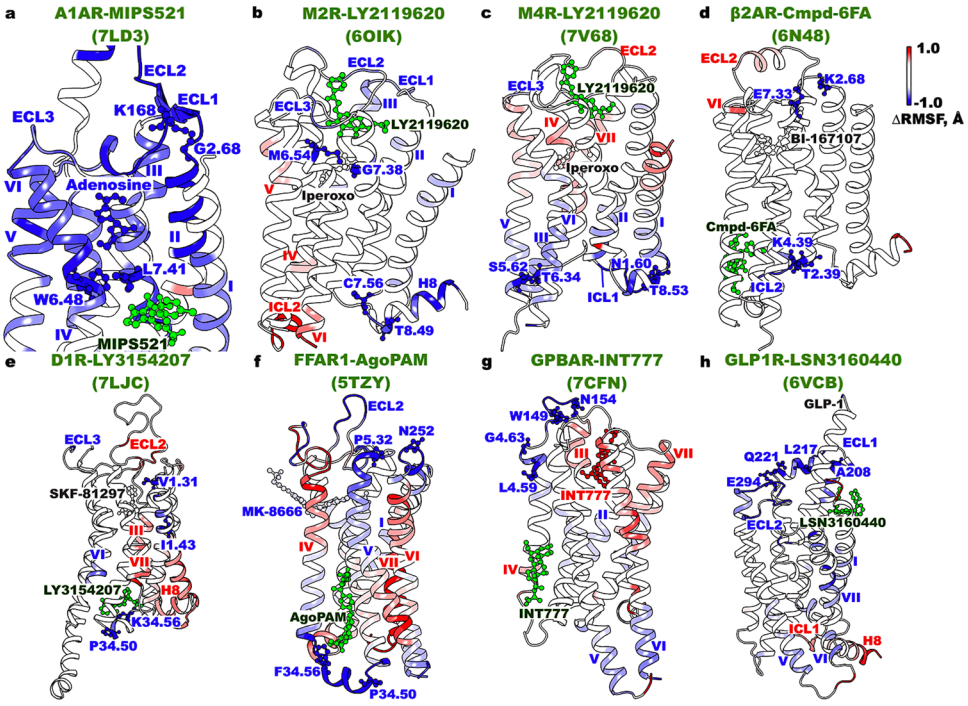


Figure 3. Characteristic residue contacts in the positive allosteric modulation of class A and B GPCRs calculated from GaMD simulations of the MIPS521-bound A_1AR (PDB: 7LD3) (a), LY2119620-bound M_2R (PDB: 6OIK) (b), LY2119620-bound M_4R (PDB: 7V68) (c), Cmpd-6FA-bound β_2AR (PDB: 6N48) (d), LY3154207-bound D_1R (PDB: 7LJC) (e), AgoPAM-bound FFAR1 (PDB: 5TZY) (f), INT777-bound GPBAR (PDB: 7CFN) (g), and LSN3160440-bound GLP1R (h). The seven TM helices are labeled I−VII, H8 for helix 8, ECL1-ECL3 for extracellular loops 1−3, and ICL1-ICL3 for intracellular loops 1−3. A color scale of −1.0 (blue) to 0 (white) to 1.0 (red) is used to show the ΔRMSF upon PAM binding, and PAMs are colored green.

cluster snapshots of receptor conformations with all GaMD production simulations combined for each simulation system.[53]

Furthermore, root-mean-square fluctuations (RMSFs) of the receptors and orthosteric ligands within the GPCR complexes were calculated by averaging the RMSFs calculated from individual GaMD simulations of each GPCR system. Changes in the RMSFs (ΔRMSF) upon binding of allosteric modulators were calculated by subtracting the RMSFs of GPCRs without modulators from those with modulators bound (Figure 1f). If the absolute average of ΔRMSF calculated from three simulations of a residue was smaller than the corresponding standard deviation, the flexibility change for that residue was considered not significant, and related residue pairs were neglected for further analysis. Important residue contacts were selected with the highest contact gradients (≥0.7) in the attention maps from DL and significant changes in the GPCR residue flexibility upon modulator binding (Figure 1g,h). They were finally used as reaction coordinates (RCs) to calculate free energy profiles by reweighting the GaMD simulations using the *PyReweighting* toolkit,[53−55] with bin sizes of 0.5−1.0 Å and cutoff of 100−500 frames in one bin. The hierarchical agglomerative clustering algorithm was also used to cluster snapshots of GPCR conformations with all GaMD production simulations combined for each system and obtain representative conformations of the low-energy conformational states that did not overlap with the PDB structures.

RESULTS

GaMD Simulations on Effects of Allosteric Modulator Binding to GPCRs

All-atom GaMD simulations were obtained for systems of the A_1AR, M_1R, M_2R, M_4R, α_{1B}AR, α_{2A}AR, α_{2C}AR, β_1AR, β_2AR, C5AR1, CB_1, CCR2, D_1R, D_2R, FFAR1, GPBAR, GLP1R, and GLR. GaMD simulations performed in this study recorded overall similar averages (∼13−16 kcal/mol) and standard deviations (∼4−5 kcal/mol) of boost potentials across the different GPCR systems, except for the A_1AR simulations that were obtained from a previous study[25] using a different force field parameter set[75] (Table S1). We first examined the structural dynamics of the GPCR orthosteric and allosteric ligands. Time courses of the orthosteric and allosteric ligand RMSDs relative to the simulation starting structures were plotted in Figures S1−S3. In most of the GPCR systems, the orthosteric ligands showed similar RMSDs in the absence and presence of allosteric modulators during the GaMD simulations (Figures S1−S3). This was consistent with previous findings that modulator binding mostly does not cause large changes in the X-ray and cryo-EM structures of the GPCRs.[23,25−37] However, a number of orthosteric ligands, including adenosine in A_1AR and PE5 in GLR, exhibited significantly smaller structural deviations in the presence of the MIPS521 PAM and MK-0893 NAM, respectively (Figures S2a and S3d).

In general, binding of allosteric modulators reduced fluctuations of the orthosteric ligands and GPCRs as shown in Figures 2, 3, S4 and S5. NAM binding primarily stabilized the allosteric binding pockets, with additional reduced flexibility observed in the extracellular or intracellular domains of the inactive receptors to prevent GPCR activation (Figures 2 and S4). In particular, binding of MT7 significantly reduced fluctuations of the extracellular mouth between ECL2 and ECL3 in M_1R (Figure 2a), which was consistent with the structural data[20] (Table S2). Furthermore, NAM binding to the extracellular pocket directly above the orthosteric pocket prevented the dissociation of the atropine antagonist, which explained their experimental cooperativity value (logα) of 0.75 ± 0.04[20] (Table S3). Binding of Cmpd-15 in β_2AR and GTPL9431 in CCR2 reduced fluctuations of the allosteric pocket formed by TM6, TM7, ICL1, and H8 Figure 2b,g). The finding for the Cmpd-15-bound β_2AR was in good agreement

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

with previous experimental data that the Cmpd-15 NAM stabilized ICL1, H8, as well as TM6 and TM7 intracellular end[28] (Table S2). Notably, AS408, NDT9513727, and Avacopan bound to a similar TM3-TM4-TM5 region on the lipid-facing surfaces of the β_2AR and C5AR1. They reduced fluctuations in the intracellular domains of TM3, TM4, and TM5 (Figure 2c−e), as well as ICL2 in β_2AR (Figure 2c). This was consistent with a previous finding that the AS408 NAM stabilized TM5 in an inactive conformation[30] (Table S2), which in turn stabilized but did not block the alprenolol antagonist, explaining the relatively low binding cooperativity of -0.7 ± 0.5[30] (Table S3). Moreover, the NDT9513727 and Avacopan NAMs were shown to stabilize TM5 through the hydrophobic stacking between TM4 and TM5[31,89] (Table S2). Binding of NNC0640 to GLP1R and MK-0893 to GLR stabilized the lipid-facing pocket on the intracellular domains of TM7 and TM6, respectively (Figures 2h,j and S4h,j), to restrict the movement of the TMs required for GPCR activation[24,36] (Table S2). Notably, binding of the MK-0893 NAM to GLR also stabilized the PE5 antagonist significantly and even prevented the dissociation of PE5 from GLR, which was observed in Sim2 of the PE5-bound GLR without the NAM (Figure S3d). In particular, the PE5 antagonist moved from the orthosteric pocket to the large gap between TM1 and TM2 extracellular ends, then interacted with ECL2, and dissociated to the bulk solvent in Sim2 (Figure S6). This dissociation pathway was similar to previous findings of class A GPCRs.[44,90] While PF-06372222 also bound to a similar region in GLP1R, no significant flexibility change was observed in the receptor, likely because the modulator-free receptor was already stable (Figures 2i and S4i). Lastly, binding of ORG27569 to CB_1 reduced fluctuations of the TM2, TM4, TM6, TM7 intracellular domains and ICL2 (Figure 2f), being consistent with the experimental finding that the ORG27569 NAM disfavored TM6 movement by stabilizing TM2 and TM4[32] (Table S2).

Binding of PAMs to GPCRs generally reduced fluctuations of the extracellular domains, orthosteric agonist-binding pocket, and intracellular G-protein coupling domains to enhance GPCR activation and signaling (Figures 3 and S5). Specifically, binding of MIPS521 to the A_1AR significantly reduced fluctuations of the adenosine agonist, the orthosteric pocket constituted by extracellular domains of TM2, ECL1, ECL2, TM3, TM5, TM6, ECL3, and TM7, and the intracellular domains of TM5, TM7, and ICL2 (Figures 3a and S5a), which was consistent with the previous observation that the PAM stabilized the adenosine-A_1AR-G-protein complex through effects on TM6 and TM7[25] (Table S2). LY2119620 binding to the M_2R reduced flexibility of TM2, TM7, ECL1-ECL3, and H8 (Figures 3b and S5b), being consistent with the experimental finding of TM7 conformational change and slight contraction of the receptor extracellular pocket upon PAM binding in this receptor[16,17] (Table S2). LY2119620 binding to the extracellular pocket helped prevent the dissociation of the iperoxo agonist from the M_2R, explaining their experimental binding cooperativity of 1.40 ± 0.09[16,17] (Table S3). In the M_4R, LY2119620 binding significantly reduced flexibility of the G-protein coupling domains of ICL1, TM3, TM5, TM6, and H8, while stabilizing ECL3 to a lesser extent (Figures 3c and S5c). These findings were consistent with the structural data that LY2119620 binding to the M_4R caused slight contraction of the ECL3 and conformational changes at the G-protein binding interfaces[26] (Table S2). In the β_2AR, Cmpd-6FA binding reduced fluctuations of orthosteric residues in the ECL1 and TM7 extracellular end as well as the intracellular domains of ICL1, TM3, and ICL2 (Figures 3d and S5d), being consistent with the structural data that minor conformational differences were observed at the ICL2 upon PAM binding[29] (Table S2). Binding of LY3154207 to the D_1R reduced fluctuations of the TM1 extracellular domain and ECL3, as well as the TM6 intracellular domain and ICL2 (Figures 3e and S5e), being consistent with the structural data that LY3154207 binding might stabilize the ICL2 helix for receptor activation and G-protein coupling[34] (Table S2). Similar to the 6N48 PDB structure of β_2AR, the 5TZY PDB structure of AgoPAM-bound FFAR1 did not have a G-protein bound. Binding of AgoPAM reduced flexibility of the ECL1, ECL2, ECL3 and TM1, TM2, TM3, and TM5 extracellular ends as well as the ICL2, and TM3 and TM4 intracellular ends, which constituted the G-protein binding pocket (Figures 3f and S5f). The stabilization of ICL2 observed in the AgoPAM-bound FFAR1 was in good agreement with the structural data of the PAM-bound 5TZY and PAM-free 5TZR PDB structures.[21] In particular, the ICL2 adopted a short helical conformation in the PAM-bound 5TZY PDB and was completely missing in the PAM-free 5TZR PDB structure of FFAR1[21] (Table S2). Two molecules of INT777 with −1 charge served as the orthosteric and allosteric ligands of GPBAR, which potentially caused electrostatic repulsion and destabilized orthosteric INT777 and most of the orthosteric residues (Figures 3g and S5g). In fact, electrostatic repulsion between orthosteric and allosteric ligands could weaken the binding of one in the presence of the other.[46] Even so, binding of INT777 to the allosteric site reduced fluctuations of ECL2 and the TM4 extracellular end, as well as the TM5 and TM6 intracellular ends (Figures 3g and S5g), being consistent with the experimental finding that the TM6 intracellular end is more contracted in the activated GPBAR[35] (Table S2). Finally, the LSN3160440 PAM binding significantly reduced fluctuations of the large orthosteric pocket of GLP1R formed by TM2, ECL1, ECL2, TM4, and ECL3, and the G-protein coupling domains in TM5 and TM6 (Figures 3h and S5h).

Overall, NAM binding stabilized the allosteric and antagonist-binding sites, with additional reduced flexibility observed in the extracellular and intracellular domains of the inactive receptors to prevent GPCR activation. Therefore, NAMs can be designed to bound locations on inactive GPCRs that inhibit GPCR activation by blocking TM5 and especially TM6 movement or slowing antagonist dissociation. Meanwhile, PAM binding stabilized the receptor extracellular domains, orthosteric agonist-binding pocket, and G-protein coupling regions of the active receptors to enhance GPCR activation and signaling. As a result, PAMs can be designed to bind locations on active GPCRs that enhance GPCR activation and signaling by stabilizing agonist binding, opening the ECL2-ECL3 pocket to recruit another agonist to the orthosteric pocket, or stabilizing G-protein binding.

Deep Learning Important Residue Contacts Underlying Allosteric Modulation of GPCRs

Classification of GPCRs bound by “Antagonist”, “Antagonist-NAM”, “Agonist”, “AgonistPAM” was carried out with high accuracies on both the training and validation sets for all GPCRs (Figures S7 and S8). The saliency (attention) residue contact maps of gradients are plotted in Figure S9 for NAM-bound GPCRs and Figure S10 for PAM-bound GPCRs. The important residue contacts for allosteric modulation of GPCRs identified from DL (Table S2) and structural flexibility analysis were

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

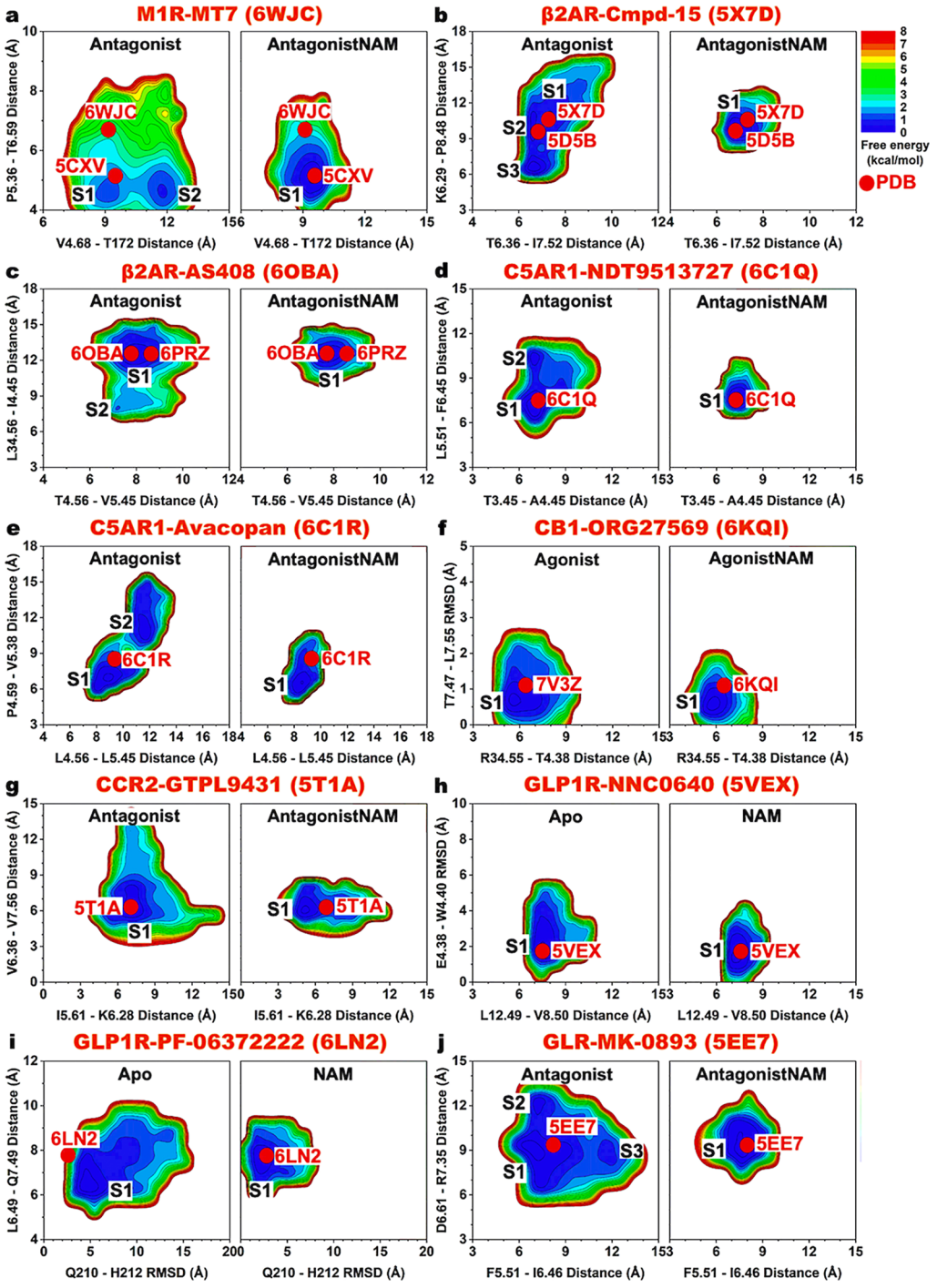


Figure 4. 2D free energy profiles of characteristic residue contacts in the allosteric modulation of class A and B GPCRs bound by NAMs. (a) C_α-atom distances between V4.68-T172^{ECL2} and P5.36-T6.39 in the M_1R without and with the MT7 NAM. The inactive antagonist-bound GPCR without and with NAM are denoted "Antagonist" and "AntagonistNAM", respectively. (b) C_α-atom distances between T6.36-I7.52 and K6.29-P8.48 in the β_2AR without and with the Cmpd-15 NAM. (c) C_α-atom distances between T4.56-V5.45 and L34.56^{ICL2}-I4.45 in the β_2AR without and with the AS408 NAM. (d)C_α-atom distances between T3.45-A4.45 and L5.51-F6.45 in the C5AR1 without and with the NDT9513727 NAM. (e) C_α-atom distances between L4.56-L5.45 and P4.59-V5.38 in the C5AR1 without and with the Avacopan NAM. (f) C_α-atom distance between R34.55^{ICL2}-T4.38 and backbone RMSD of T7.47-L7.55 relative to the 6KQI PDB structure in the CB_1 receptor without and with the ORG27569 NAM. The agonist bound GPCR without and with NAM are denoted "Agonist" and "AgonistNAM", respectively. (g) C_α-atom distances between I5.61-K6.28 and V6.36-V7.56 in the CCR2 without and with the GTPL9431 NAM. (h) C_α-atom distance between L12.49^{ICL1}-V8.50 and backbone RMSD of E4.38-W4.40 relative to the 5VEX PDB structure in the GLP1R without and with the NNC0640 NAM. The apo GPCR without and with NAM are denoted "Apo" and "NAM", respectively. (i) Backbone RMSD of Q210^{ECL1}-H212^{ECL1} relative to the 6LN2 PDB structure and C_α-atom distance between L6.49-Q7.49 in the GLP1R without and with the PF-06372222 NAM. (j) C_α-atom distances between F5.51-I6.46 and D6.61-R7.35 in the GLR without and with the MK-0893 NAM. The PDB structures are mapped to the free energy surfaces as red dots. The RMSDs of PDB structures were calculated by averaging from the corresponding 10 ns short cMD simulations.

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

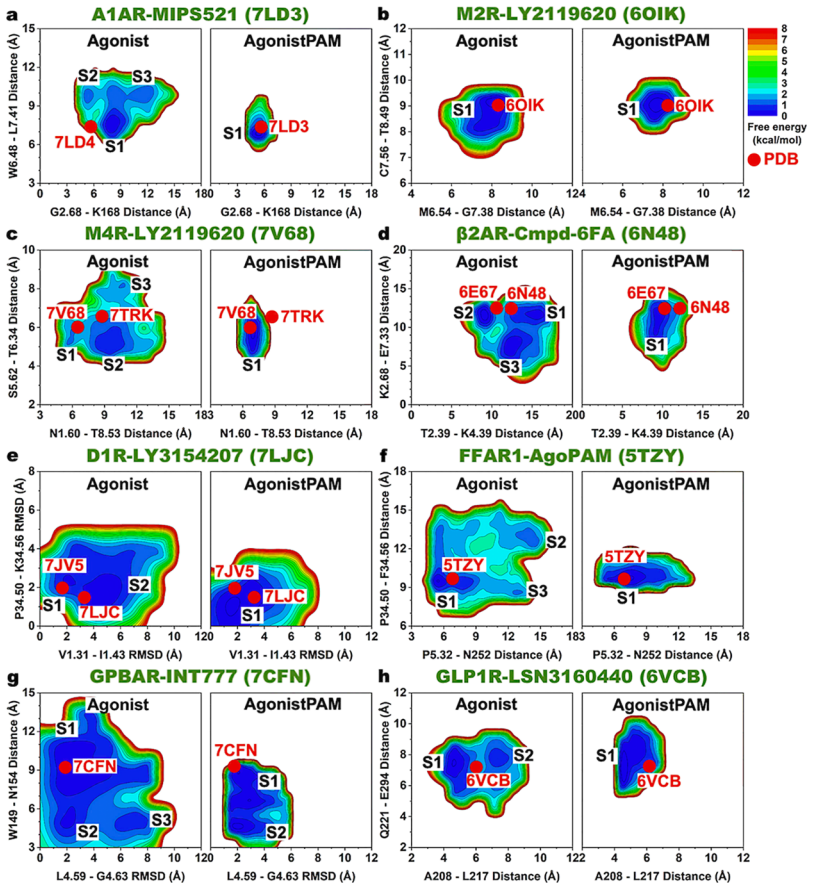


Figure 5. 2D free energy profiles of characteristic residue contacts in the allosteric modulation of class A and B GPCRs bound by PAMs. (a) C_α-atom distances between G2.68-K168^{ECL2} and W6.48-L7.41 in the A_1AR without and with the MIPS521 PAM. The active agonist bound GPCR without and with PAM are denoted "Agonist" and "AgonistPAM", respectively. (b) C_α-atom distances between M6.54-G7.38 and C7.56-T8.49 in the M_2R without and with the LY2119620 PAM. (c) C_α-atom distances between N1.60-T8.53 and S5.62-T6.34 in the M_4R without and with the LY2119620 PAM. (d) C_α-atom distances between T2.39-K4.39 and K2.68-E7.33 in the β_2AR without and with the Cmpd-6FA PAM. (e) Backbone RMSDs of V1.31-I1.43 and P34.50^{ICL2}-K34.56^{ICL2} relative to the 7LJC PDB structure in the D_1R without and with the LY3154207 PAM. (f) C_α-atom distances between P5.32-N252^{ECL3} and P34.50^{ICL2}-F34.56^{ICL2} in the FFAR1 without and with the AgoPAM PAM. (g) Backbone RMSD of L4.59-G4.63 and C_α-atom distance between W149^{ECL2}-N154^{ECL2} in the GPBAR without and with the INT777 PAM. (h) C_α-atom distances between A208^{ECL1}-L217^{ECL1} and Q221^{ECL1}-E294^{ECL2} in the GLP1R without and with the LSN3160440 PAM. The PDB structures are mapped to the free energy surfaces as red dots. The RMSDs of PDB structures were calculated by averaging from the corresponding 10 ns short cMD simulations.

included in Figures 2 and 3. Notably, most of these reaction coordinates involve residues located in the allosteric modulator-binding sites and structural domains that have been identified to be important for allosteric modulation of GPCRs in previous experimental structural studies (Table S2).

The important residue contacts in NAM-bound GPCRs were mostly located near allosteric binding sites (Figure 2). In the MT7-bound M_1R, residue contacts V4.68-T172^{ECL2} and P5.36-T6.59 involved the TM4, TM5, and TM6 extracellular domains and ECL2 (Figure 2a). In the Cmpd-15-bound β_2AR, residue contacts K6.29-P8.48 and T6.36-I7.52 involved the TM6 and TM7 intracellular domains and H8 (Figure 2b). Previous structural data showed that residue T6.36 played a crucial role in the negative allosteric modulation of Cmpd-15 in the β_2AR[28] (Table S2). In the AS408-bound β_2AR, residue contact T4.56-V5.45 connected the middle of TM4 and TM5, while residue contact L34.56^{ICL2}-I4.45 connected the TM4 intracellular domain and ICL2 (Figure 2c). Residue V5.45 was shown to control the negative allosteric modulation of AS408 in the β_2AR[30] (Table S2). In the NDT9513727-bound C5AR1, residue contact T3.45-A4.45 connected the TM3 and TM4 intracellular ends, while residue contact L5.51-F6.45 connected the middle of TM5 and TM6 (Figure 2d). In the Avacopan-bound C5AR1, residue contacts L4.56-L5.45 and P4.59-V5.38 connected the TM4 and TM5 extracellular ends (Figure 2e). In the ORG27569-bound CB_1, residue contact R34.55^{ICL2}-T4.38 connected the TM4 intracellular domain and ICL2, while residues T7.47-L7.55 involved the TM7 intracellular end (Figure 2f). In the GTPL9431-bound CCR2, residue contacts I5.61-K6.28 and V6.36-V7.56 connected the TM5, TM6, and TM7 intracellular domains (Figure 2g). In the NNC0640-bound GLP1R, residue contact L12.49^{ICL1}-V8.50 connected ICL1 and H8, and residues E4.38-W4.40 involved the TM4 intracellular domain (Figure 2h). In the PF-06372222-bound GLP1R, residues Q210^{ECL1}-H212^{ECL1} involved the ECL1, while residue contact L6.49-Q7.49 connected the middle of TM6 and TM7 (Figure 2i). Lastly, in the MK-0893-bound GLR, residue contact F5.51-I6.46 connected the middle of TM5 and TM6, while residue contact D6.61-R7.35 connected the extracellular domains of TM6 and TM7 (Figure 2j).

In PAM-bound GPCRs, important residue contacts were mostly located in the extracellular domains, orthosteric agonist-binding pocket, and intracellular G protein-binding regions (Figure 3). In the MIPS521-bound A_1AR, residue contact

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

G2.68-K168^{ECL2} involved the TM2 extracellular end and ECL2, while residue contact W6.48-L7.41 connected the middle of TM6 and TM7, interacting with the adenosine agonist (Figure 3a). In the LY2119620-bound M_2R, residue contact M6.54-G7.38 were found interacting with the iperoxo agonist near the TM6 and TM7 extracellular ends, while residue contact C7.56-T8.49 connected the TM7 intracellular domain and H8 (Figure 3b). In the LY2119620-bound M_4R, residue contacts N1.60-T8.53 and S5.62-T6.34 involved the TM1, TM5, and TM6 intracellular ends and H8 (Figure 3c). In the Cmpd-6FA-bound β_2AR, residue contact T2.39-K4.39 connected the TM2 and TM4 intracellular ends, while residue contact K2.68-E7.33 connected the TM2 and TM7 extracellular domains (Figure 3d). In the LY3154207-bound D_1R, residues V1.31-I1.43 and P34.50^{ICL2}-K34.56^{ICL2} involved the TM1 extracellular end and ICL2, respectively (Figure 3e). In the AgoPAM-bound FFAR1, residue contact P5.32-N252^{ECL3} connected the TM5 extracellular domain and ECL3, while residue contact P34.50^{ICL2}-F34.56^{ICL2} formed the short helix of ICL2 (Figure 3f). In the INT777-bound GPBAR, residues L4.59-G4.63 and residue contact W149^{ECL2}-N154^{ECL2} involved the TM4 extracellular domain and ECL2, respectively (Figure 3g). In the LSN3160440-bound GLP1R, residue contacts A208^{ECL1}-L217^{ECL1} and Q221^{ECL1}-E294^{ECL2} connected ECL1 and ECL2 (Figure 3h).

Free Energy Profiling of Important Residue Contacts in GPCR Allostery

2D free energy profiles were calculated from GaMD simulations for the important residue contacts identified from DL and structural flexibility analyses of GPCRs (Figures 4 and 5). Overall, binding of NAMs and PAMs reduced the conformational space of the inactive antagonist-bound and active agonist-bound GPCRs, respectively. Moreover, in case the modulator-free GPCRs were able to sample multiple low-energy conformational states, binding of the NAMs and PAMs confined the GPCR residue contacts to fewer low-energy states (only 1 for most GPCRs) (Figures 4 and 5).

Binding of the MT7 NAM to M_1R confined the TM4 extracellular domain and ECL2 from two conformational states ("S1" and "S2") to only the "S1" state, in which the extracellular mouth adopted the more closed conformation (Figure 4a), being consistent with the previous structural data[20] (Table S2). Cmpd-15 binding to the β_2AR reduced the conformational space from three states ("S1″-"S3") to only the "S1" state, in which the intracellular pocket formed by TM6, TM7, and H8 adopted the more open conformation to accommodate the NAM and restrict TM6 movement required for GPCR activation[28] (Figures 4b and 2b). Similarly, AS408 binding to the β_2AR confined the TM4 intracellular end and ICL2 from two conformational states ("S1" and "S2") to only state "S1″, in which the intracellular pocket formed by the TM4 intracellular end and ICL2 adopted a more open conformation for stable NAM binding and inactive conformation[30] (Figures 4c and 2c). Binding of NDT9513727 to the C5AR1 reduced the number of low-energy conformational states from two to one (the "S1" state), where the allosteric pocket located between TM3 and TM4 intracellular ends as well as the middle of TM5 and TM6 adopted a more open conformation to accommodate the NAM and restrict the TM movement required for GPCR activation[59] (Figures 4d and 2d). Avacopan-binding to the C5AR1 confined the TM4 and TM5 extracellular domains from two conformational states to only state "S1″, where the TM4 and TM5 extracellular ends adopted the more closed conformation to stabilize NAM binding and inactive conformation[59] (Figures 4e and 2e). The hydrophobic stacking found between TM4 and TM5 extracellular ends was again in good agreement with previous findings.[31,89] Binding of MK-0893 to the GLR confined the conformational space from three states ("S1″-S3") to only state "S1″, in which the middle of TM5 and TM6 as well as TM6 and TM7 extracellular ends adopted the more closed conformations to restrict the required TM6 movement for receptor activation[36] (Figure 4j). In the cases of ORG27569-bound CB_1, GTPL9431-bound CCR2, NNC0640-bound GLP1R, and PF-06372222-bound GLP1R, the modulator-free GPCRs already sampled only one low-energy conformational state (Figure 4f–i).

In PAM-bound GPCRs, binding of MIPS521 to the A_1AR confined the TM2 extracellular domain and ECL2 as well as the middle of TM6 and TM7 from three conformational states ("S1″-"S3") to only the "S1" state, in which the extracellular mouth and orthosteric agonist-binding pocket adopted the more closed conformation to stabilize agonist binding and GPCR signaling (Figures 5a and 3a). This finding was highly consistent with previous experimental and computational studies of the A_1AR allosteric modulation,[25,53,91,92] where the PAM stabilized the adenosine-A_1AR-G-protein complex through effects on TM6 and TM7[25] (Table S2). Moreover, the extracellular mouth was shown to play a crucial role in the dissociation of orthosteric ligands from class A GPCRs.[44,90] Closure of this extracellular mouth would prevent the dissociation of adenosine and further stabilize agonist binding and GPCR signaling. The modulator-free GPCR in the case of LY2119620-bound M_2R sampled only one low-energy conformational state (Figure 5b). LY2119620-binding to the M_4R reduced the conformational space of the TM1 intracellular end and H8 from three states ("S1″-"S3") to only state "S1″, in which the G-protein-binding region adopted the more closed conformation to stabilize G-protein binding and the receptor active conformation[26] (Figure 5c). Binding of Cmpd-6FA to the β_2AR confined the TM2 and TM7 extracellular ends as well as TM2 and TM4 intracellular ends from three states ("S1″-S3") to the "S1" state, in which the G protein-binding domain adopted an open conformation for G-protein coupling (Figure 5d and Table S2). Binding of the LY3154207 PAM to the D_1R confined the TM1 extracellular end and ICL2 from two conformational states ("S1" and "S2") to only state "S1″, stabilizing G-protein binding and receptor active conformation[34] (Figures 5e and 3e). AgoPAM-binding to the FFAR1 reduced the conformational space of the TM5 extracellular end, ECL3, and ICL2 from two states ("S1" and "S2") to state "S1″, where the extracellular mouth between TM5 and ECL3 as well as the ICL2 adopted a closed and short helix conformation, respectively (Figures 5f and 3f). Here, the finding that the ICL2 preferred a short helix conformation in the AgoPAM-bound FFAR1 resembled the structural data of 5TZY and 5TZR PDB structures well[21] (Table S2). Binding of INT777 to the allosteric site of GPBAR reduced the number of low-energy conformational states from three ("S1" to "S3") to two ("S1" and "S2") (Figure 5g). Lastly, LSN3160440-binding to the GLP1R confined ECL1 and ECL2 from two conformational states ("S1" and "S2") to only the "S1" state, in which the extracellular mouth between ECL1 and ECL2 adopted the closed conformation to stabilize the peptide agonist and receptor active conformation for signaling (Figures 5h and 3h).

Since most reaction coordinates involved residues located at or near the allosteric binding pockets of the NAM- and PAM-

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

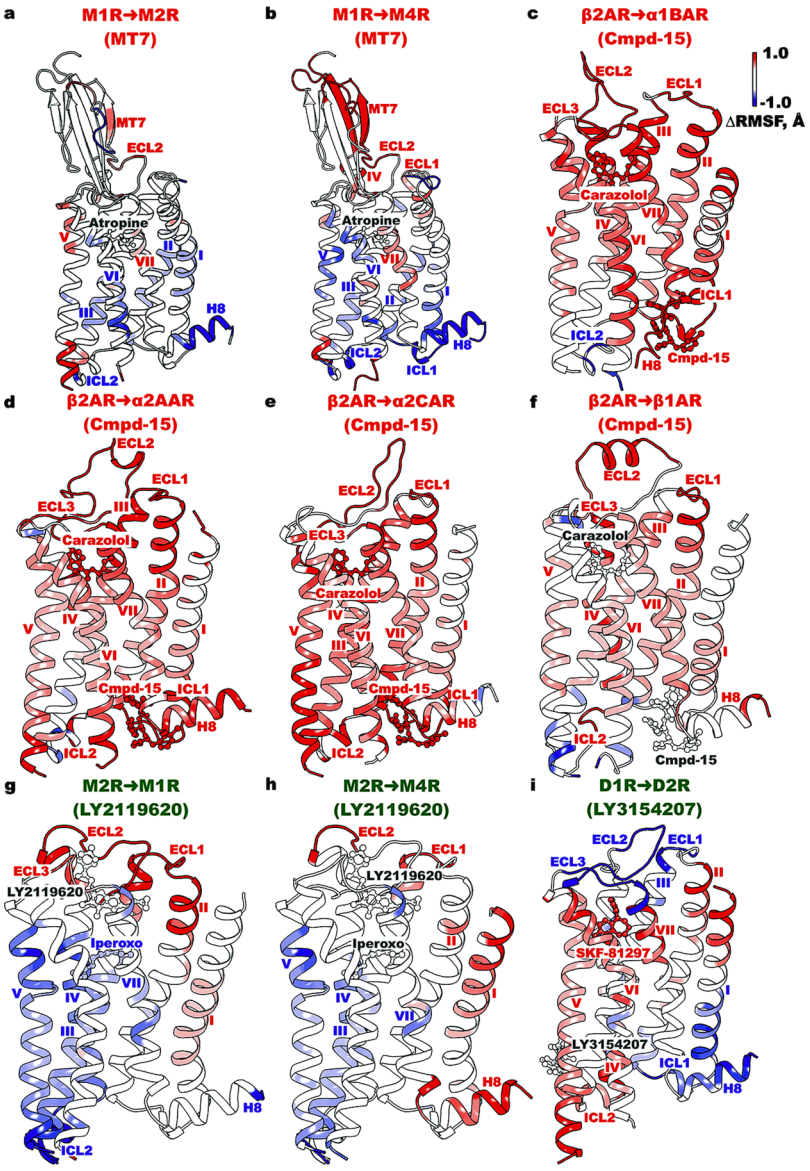


Figure 6. Increased system fluctuations were observed for binding of NAMs (MT7 and Cmpd-15) and PAMs (LY2119620 and LY3154207) to "non-cognate" GPCRs in GaMD simulations. Changes in root-mean-square fluctuations (ΔRMSFs) of the receptor, orthosteric and allosteric ligands in the M_2R (a) and M_4R (b) compared to the M_1R bound by the MT7 NAM (PDB: 6WJC), the $\alpha_{1B}AR$ (c), $\alpha_{2A}AR$ (d), $\alpha_{2C}AR$ (e), and β_1AR (f) compared to the β_2AR bound by the Cmpd-15 NAM (PDB: 5X7D), the M_1R (g) and M_4R (PDB: 7V68) (h) compared to the M_2R bound by the LY2119620 PAM (PDB: 6OIK), and the D_2R (i) compared to the D_1R bound by the LY3154207 PAM (PDB: 7LJC). A color scale of −1.0 (blue) to 0 (white) to 1.0 (red) is used to show the ΔRMSF.

bound GPCRs (Figures 2 and 3), the free energy profiles in Figures 4 and 5 also illustrated the conformational changes of the allosteric modulator binding sites. Because conformations of the allosteric modulator-bound pockets were already sampled in the modulator-free GPCRs, binding of allosteric modulators to GPCRs involved a "conformational selection" mechanism. We selected new reaction coordinates based on experimental observations, especially for the M_2R and M_4R that did not have reaction coordinates identified for residues in the modulator-binding sites yet (Figure 3) and calculated their free energy profiles in Figure S11. The reaction coordinates included the $\chi 2$ dihedral angle of residue W7.35 and the C_α-atom distance between residues $Y177^{ECL2}$-W7.35 in the M_2R (Figure S11a,b) and the C_α-atom distance between residues N6.58-W7.35 in the M_4R (Figure S11c,d). Two different low-energy conformational states were observed in the free energy profiles of the GPCRs without and with PAM bound for both M_2R and M_4R. However, while the GPCRs without PAM favored the "S2" states (Figure S11a,c), the "S1" states were preferred in the GPCRs with PAM bound, given the lower free energy observed for the "S1" states in these systems (Figure S11b,d). We also calculated the time courses of ICL2 secondary structures in the FFAR1 without and with AgoPAM (Figure S12). While the ICL2 could not be resolved in the FFAR1 structure without AgoPAM bound,[21] our simulations showed that the ICL2 α-helix was formed even in the FFAR1 without AgoPAM (Figures S12a and 5f). AgoPAM binding to FFAR1 made the ICL2 α-helix more stable (Figures S12b and 5f), which was consistent with the experimental structures of FFAR1.[21] These findings further supported our conclusion that binding of modulator to GPCRs involved a "conformational selection" mechanism.

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

To examine the simulation convergence, 2D free energy profiles of characteristic residue contacts were calculated from the individual GaMD simulations of each GPCR system (Figures 4 and 5). The standard deviations of the 2D free energy surfaces of NAM- and PAM-bound GPCRs were then calculated and plotted in Figures S13 and S14, respectively. Overall, for the NAM-bound GPCRs, most of the low-energy conformational states showed near zero standard deviations of 2D free energy across the GaMD simulations, demonstrating that GaMD simulations converged in the low-energy regions (Figure S13). Exceptions included the "S2" states of the M_1R without MT7 (Figure S13a), β_2AR without Cmpd-15 (Figure S13b), β_2AR without AS408 (Figure S13c), C5AR1 without NDT9513727 (Figure S13d), C5AR1 without Avacopan (Figure S13e), and the "S2" and "S3" state of the GLR without MK-0893 (Figure S13j). However, the standard deviations of these low-energy states were also relatively low, at ∼3−5 kcal/mol. For the PAM-bound GPCRs, more low-energy states exhibited ∼3−5-kcal/mol standard deviations of free energy, given relatively larger sizes of the simulation systems (Figure S14). Nevertheless, the GaMD simulations still showed almost zero energy standard deviations in one or more low-energy conformational states of the PAM-bound GPCR systems. These states included the "S1" states of the A_1AR with MIPS521 (Figure S14a), M_2R with LY2119620 (Figure S14b), "S1" and "S2" state of the M_4R with and without LY2119620, respectively, (Figure S14c), "S1" and "S3" state of the β_2AR with and without Cmpd-6FA, respectively, (Figure S14d), "S1" and "S2" states of the D_1R with and without LY3154207 (Figure S14e), "S1" states of the FFAR1 with and without AgoPAM (Figure S14f), "S1" and "S2" state of the GPBAR with the INT777 PAM (Figure S14g), and "S1" states of the GLP1R with and without LSN3160440 (Figure S14h). Moreover, the time courses of characteristic residue contacts in the allosteric modulation of class A and B GPCRs were plotted in Figures S15−S18, which showed similar dynamic behavior in each independent GaMD simulation of most GPCR systems without and with allosteric modulators. Finally, we calculated representative 2D free energy profiles of characteristic residue contacts of one class A GPCR without and with NAM bound (the M_1R without and with MT7), one class A GPCR without and with PAM bound (the M_4R without and with LY2119620), and one class B GPCR (the GLP1R without and with LSN3160440) using only the last 250 ns of their GaMD simulations and compared the results with the 2D free energy profiles calculated from the whole 500 ns of GaMD simulations (Figure S19). Since the 2D free energy profiles calculated from the last 250 ns and full 500 ns of GaMD simulations were highly similar in terms of the conformational spaces and low-energy conformational states sampled for all six systems (Figure S19), it was reasonable for us to use the whole 500 ns of GaMD production simulations for analysis. Our systems were properly equilibrated as extensive GaMD equilibration simulations were performed on each of them prior to the production simulations.

Representative conformations of the low-energy conformational states of GPCRs uncovered during the allosteric modulation were shown in Figures S20−S23 and included in Supporting Data 1. In the "S2" state of the M_1R, the TM4 extracellular domain became distorted, illustrating a more flexible extracellular domain in the absence of the MT7 NAM (Figure S20a). The "S2" and "S3" states of A_1AR were observed in the absence of the MIPS521 PAM (Figure 5a). In these two states, the β-sheet between ECL1 and ECL2 was completely distorted, pointing to a more flexible extracellular mouth and less stable binding of the adenosine agonist in the absence of the PAM (Figure S21a,b). The "S2" and "S3" states of the FFAR1 saw a complete distortion of the α-helix secondary structures in the TM5 extracellular domain, ECL3, and ICL2 in the absence of AgoPAM (Figure S22a,b). Notably, in the "S2" and "S3" states of the GLR without the MK-0893 NAM, the PE5 antagonist moved downward toward a pocket formed between TM5 and TM6 due to outward movement of TM6 (Figure S23a,b).

Selectivity of GPCR Allosteric Modulators

Additional GaMD simulations were performed on artificially generated computational models to examine the binding selectivity of the MT7 and Cmpd-15 NAMs to the muscarinic and adrenergic receptors, respectively, as well as the LY2119620 and LY3154207 PAMs to the muscarinic and dopamine receptors, respectively. Flexibility changes were calculated by subtracting RMSFs of the cognate from the "non-cognate" GPCRs of the modulators (Figure 6). Furthermore, 2D free energy profiles of the heavy-atom RMSDs of orthosteric and allosteric ligands relative to their respective starting structures were calculated and are shown in Figure S24. Overall, modulator binding in the "non-cognate" GPCRs resulted in higher complex fluctuations and mostly larger conformational space (i.e., reduced cooperative effects) compared to the cognate GPCRs, demonstrating the binding preference and selectivity of allosteric modulators toward their cognate subtypes.

Significantly higher fluctuations were observed for binding of allosteric modulators to "non-cognate" GPCRs, especially in the allosteric pockets and various receptor domains, compared to their binding to the cognate GPCRs (Figure 6). Compared to the MT7-bound M_1R, the NAM showed moderately increased to much higher fluctuations in the model M_2R and M_4R, respectively. Furthermore, NAM binding increased fluctuations in ECL2 of the M_2R (Figure 6a) and the TM4 extracellular end, ECL1, and ECL2 in the M_4R (Figure 6b). Compared to the Cmpd-15-bound β_2AR, the NAM showed much higher fluctuations in the "non-cognate" subtypes of the α_{1B}AR, α_{2A}AR, and α_{2C}AR, and significantly increased fluctuations of these three GPCR-antagonist complexes (Figure 6c−e). The flexibility increase was smaller in the Cmpd-15-bound β_1AR, likely due to the receptor similarity in its sequence and structure to the β_2AR. Even so, binding of Cmpd-15 to the β_1AR significantly increased fluctuations in the TM2 extracellular end, ECL1, ECL2, TM4, ICL2, and H8 (Figure 6f). The binding preference of the LY2119620 PAM reduced from the M_2R to the M_4R and then M_1R.[93] Here, binding of LY2119620 to the M_1R significantly increased fluctuations in the TM2, TM3 and TM4 extracellular ends, ECL1, ECL2, and ECL3 compared to the LY2119620-bound M_2R (Figure 6g). In the LY2119620-bound M_4R, PAM binding only slightly increased fluctuations in the TM2 and TM3 extracellular ends, ECL1, and ECL2 compared to the M_2R (Figure 6h). Our simulation results were thus consistent with the previous experimental finding.[93] Finally, binding of the LY3154207 PAM to the D_2R increased fluctuations mostly in the TM1 and TM2 extracellular ends, TM4 intracellular end, ICL2, TM5, TM6, TM7, and the SKF-81297 agonist compared to the cognate D_1R (Figure 6i).

Binding of allosteric modulators to "non-cognate" GPCRs mostly increased conformational space of the orthosteric and allosteric ligands with higher RMSDs. Moreover, most of the modulator-bound "non-cognate" GPCRs sampled more low-

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

energy conformational states compared to the cognate GPCRs (Figure S24). In particular, the MT7-bound M_4R sampled four low-energy conformational states ("S1″-"S4") of the atropine antagonist and MT7 NAM compared to only two conformations ("S1" and "S2") of the MT7-bound M_1R (Figure S24a,c). While the MT7-bound M_2R sampled the same number of conformational states as the M_1R, the "S3" state of the MT7-bound M_2R showed higher RMSD of the atropine antagonist at ∼3.8 Å compared to the ∼1.0 Å atropine RMSD in the "S2" state of M_1R (Figure S24a,b). The Cmpd-15-bound α_{1B}AR and α_{2C}AR sampled three low-energy conformational states ("S2″-"S4" and "S1″,"S3″-"S4″, respectively) compared to only one state ("S1") of the Cmpd-15-bound β_2AR (Figure S24d,e,g). While the Cmpd-15-bound α_{2A}AR sampled the same number of low-energy conformational states as the cognate β_2AR, the "S3" state of the α_{2A}AR showed higher RMSDs of both the carazolol antagonist and Cmpd-15 NAM at ∼2.3 and ∼6.0 Å, respectively, compared to the ∼1.0 Å and ∼2.0 Å carazolol and Cmpd-15 RMSDs of state "S1" of the β_2AR (Figure S24d,f). The LY2119620-bound M_1R sampled only one low-energy conformational state ("S2") compared to two states in the M_2R ("S1" and "S2") and M_4R ("S3" and "S4") (Figure S24i−k). Nevertheless, in the "S2" state, the LY2119620 PAM adopted the ∼6.0 Å relatively higher RMSD conformation. Finally, the LY3154207-bound D_2R sampled two low-energy conformational states ("S2" and "S3"), both of which exhibited significantly higher agonist and PAM RMSDs, compared to the one "S1" state of LY3154207-bound D_1R (Figure S24l,m).

The standard deviations of 2D free energy surfaces obtained from individual GaMD simulations of the cognate and "non-cognate" GPCRs bound by allosteric modulators were plotted in Figure S25. Generally, the GaMD simulations of the cognate and "non-cognate" GPCRs converged well (i.e., almost zero standard deviations of free energy) in one or more low-energy states. In particular, these states included the "S1" and "S2" states of the MT7-bound M_1R (Figure S25a), "S1" state of the MT7-bound M_2R (Figure S25b), "S2" and "S4" state of the MT7-bound M_4R (Figure S25c), "S1" state of the Cmpd-15-bound β_2AR (Figure S25d), "S2" state of the Cmpd-15-bound α_{1B}AR (Figure S25e), "S3" states of the Cmpd-15-bound α_{2A}AR and α_{2C}AR (Figure S25f,g), "S1" state of the Cmpd-15-bound β_1AR (Figure S25h), "S1" state of the LY2119620-bound M_2R (Figure S25i), "S3" state of the LY2119620-bound M_4R (Figure S25j), "S2" state of the LY2119620-bound M_1R (Figure S25k), "S1" state of the LY3154207-bound D_1R (Figure S25l), and "S2" state of the LY3154207-bound D_2R (Figure S25m). Moreover, the time courses of the RMSDs of orthosteric and allosteric ligands relative to the respective starting structures and models of cognate and "non-cognate" GPCRs were plotted in Figures S1−S3 and S26. The ligand RMSDs were mostly converged across their respective 500 ns GaMD production simulations.

Representative conformations of the low-energy states of the cognate and "non-cognate" GPCRs bound by allosteric modulators were shown in Figures S27−S30 and included in the Supporting Data 1. Particularly, in the new "S5" state of the model MT7-bound M_4R, the NAM clearly tilted compared to its starting pose, demonstrating the unstable binding of MT7 in the "non-cognate" M_4R (Figure S27d). In the "S2" state of the LY2119620-bound M_2R, a more vertical pose of the LY2119620 PAM was observed in the extracellular mouth of the M_2R, compared to the 6OIK PDB structure (Figure S29a). Finally, the "S2" and "S3" states of the model LY3154207-bound D_2R showed partial dissociation of the PAM from the "non-cognate" receptor. The LY3154207 PAM moved downward on the intracellular side from the "S2" to "S3" state, compared to its original pose located in an intracellular pocket formed by TM3, TM4, and ICL2 (Figure S30).

DISCUSSION

Allosteric modulators have emerged as more selective drug candidates than orthosteric agonist and antagonist ligands. However, many X-ray and cryo-EM structures of GPCRs resolved so far exhibit negligible differences upon the binding of allosteric modulators. Consequently, the mechanism of dynamic allosteric modulation in GPCRs remains unclear, despite their critical importance. It is important to note that as membrane proteins, GPCRs are well known to be extremely flexible. Conventional MD simulations often suffer from insufficient conformational sampling of the GPCRs. Moreover, it is very difficult to define relevant collective variables to characterize protein allosteric modulation by its nature for popular enhanced sampling methods. To address these challenges, we applied GaMD for unconstrained enhanced sampling of GPCRs without predefinition of the collective variables. In addition, previous studies of GPCR allostery have been mostly limited to one or a few receptor systems. In this work, we have carried out extensive enhanced sampling MD simulations (66 μs GaMD) on 44 GPCR systems, which include all available X-ray and cryo-EM experimental structures of allosteric modulator-bound GPCRs and eight additional computational models to examine modulator selectivity. We have integrated GaMD and DL in GLOW to map dynamic changes in free energy landscapes of GPCRs upon the binding of allosteric modulators. By intersecting DL-predicted residue contacts with the highest gradient and residues with the largest flexibility changes, we selected characteristic residue contacts for free energy profiling to decipher the effects of allosteric modulator binding on GPCRs.

The PAM and NAM binding primarily reduced the dynamic fluctuations of the GPCR complexes. NAMs stabilized the allosteric and antagonist-binding sites to prevent GPCR activation. PAMs stabilized the receptor extracellular domains, orthosteric agonist-binding pocket, and G-protein coupling regions to enhance GPCR activation and signaling. Furthermore, the conformational space of the GPCRs was significantly reduced upon modulator binding. The NAMs and PAMs confined the GPCRs to mainly one specific conformation for signaling. These effects transcended across classes A and B GPCRs. We also proved that the binding of allosteric modulators to GPCRs involved a "conformational selection" mechanism. It is preferable to use high-resolution structures of modulator-bound GPCRs for drug design, although this will likely yield ligands similar to those of the bound modulators. In the absence of available modulator-bound structures as for most current GPCRs, it is critical to use a structural ensemble of representative GPCR conformations rather than a single structure for compound docking ("ensemble docking"),[94] which will potentially improve structure-based design of novel allosteric drugs of GPCRs. NAM and PAM binding were also found selective toward their cognate receptor subtypes. Significantly higher fluctuations were observed for modulator binding to "non-cognate" GPCR subtypes, for which the orthosteric and allosteric ligands exhibited larger RMSDs.

Overall, the effects of NAM and PAM binding on the receptor dynamics were consistent across different GPCRs. Two of the studied structures were class B GPCRs that were bound by only

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180

negative allosteric modulators (without any orthosteric ligands bound). They were the GLP1R bound by NNC0640 (PDB: 5VEX)[36] and GLP1R bound by PF-06372222 (PDB: 6LN2).[37] Although the effect of PF-06372222 binding was not so clear, the effect of NNC0640 binding to GLP1R was consistent with the overall trend of NAM binding to other GPCRs. In fact, exceptions were observed only in two cases: the GPBAR bound by the INT777 PAM and the GLP1R bound by the PF-06372222 NAM. In the GPBAR, the fact that two molecules of the same charged ligand bound to the receptor could potentially create electrostatic repulsion, leading to increased fluctuations in the orthosteric ligand and other parts of the receptor[46] (Figure 3g). In addition, potential inaccuracies in, especially, the ligand force field parameters could contribute to the inconsistencies observed in these two cases. Further ligand parameter optimization could be helpful to achieve more consistent results in these GPCR systems. The allosteric effect of GPCRs has been shown to depend on the orthosteric probe.[14] We will carry out further studies in the future to identify characteristic features of different PAMs and NAMs for a given orthosteric ligand and to understand their functional mechanisms. In summary, we have deciphered the mechanism of dynamic allostery in class A and B GPCRs through the DL of extensive GaMD simulations. Our findings are expected to facilitate the rational design of selective GPCR allosteric drugs.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.3c00503.

Example input files for GaMD equilibration and production simulations of GPCRs, Figures S1−S30, and Tables S1−S3 (PDF)

PDB files of the representative low-energy conformational states in Figures S20−S23 and S27−S30 (ZIP)

Ligand parameter files used in the GaMD simulations of GPCRs (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Yinglong Miao − *Computational Biology Program and Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66047, United States;* Present Address: Computational Medicine Program and Department of Pharmacology, University of North Carolina − Chapel Hill, Chapel Hill, North Carolina, USA 27599; orcid.org/0000-0003-3714-1395; Email: Yinglong_Miao@med.unc.edu

Authors

Hung N. Do − *Computational Biology Program and Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66047, United States;* Present Address: Theoretical Biology and Biophysics Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States; orcid.org/0000-0002-6497-4096

Jinan Wang − *Computational Biology Program and Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66047, United States;* orcid.org/0000-0003-0162-212X

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.3c00503

Author Contributions

CRediT: **Hung Nguyen Do** data curation, formal analysis, investigation, methodology, software, validation, visualization, writing-original draft, writing-review & editing; **Jinan Wang** project administration, supervision; **Yinglong Miao** conceptualization, project administration, resources, supervision, validation, writing-original draft, writing-review & editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Miao Lab members for valuable discussions. This work used supercomputing resources with allocation award TG-MCB180049 through the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296, and project M2874 through the National Energy Research Scientific Computing Center (NERSC), which is a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231, and the Research Computing Cluster and BigJay Cluster funded through NSF Grant MRI-2117449 at the University of Kansas. This work was supported by the National Institutes of Health (R01GM132572).

■ REFERENCES

(1) 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), 185−194.

(2) Hauser, A. S.; Chavali, S.; Masuho, I.; Jahn, L. J.; Martemyanov, K. A.; Gloriam, D. E.; Babu, M. M. Pharmacogenomics of GPCR Drug Targets. *Cell* **2018**, *172* (1−2), 41−54.

(3) Stevens, R. C.; Cherezov, V.; Katritch, V.; Abagyan, R.; Kuhn, P.; Rosen, H.; Wüthrich, K. The GPCR Network: a large-scale collaboration to determine human GPCR structure and function. *Nat. Rev. Drug Discovery* **2013**, *12* (1), 25−34.

(4) Isberg, V.; Mordalski, S.; Munk, C.; Rataj, K.; Harpsøe, K.; Hauser, A. S.; Vroling, B.; Bojarski, A. J.; Vriend, G.; Gloriam, D. E. GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res.* **2016**, *44* (D1), D356−364, DOI: 10.1093/nar/gkv1178.

(5) Neubig, R. R.; Spedding, M.; Kenakin, T.; Christopoulos, A.; International Union of Pharmacology Committee on Receptor, N.; Drug, C.. International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. *Pharmacol Rev.* **2003**, *55* (4), 597−606.

(6) Allen, J. A.; Roth, B. L. Strategies to discover unexpected targets for drugs active at G protein-coupled receptors. *Annu. Rev. Pharmacol. Toxicol.* **2011**, *51*, 117−144.

(7) Christopoulos, A. Allosteric binding sites on cell-surface receptors: Novel targets for drug discovery. *Nat. Rev. Drug Discovery* **2002**, *1* (3), 198−210.

(8) May, L. T.; Leach, K.; Sexton, P. M.; Christopoulos, A. Allosteric modulation of G protein-coupled receptors. *Annu. Rev. Pharmacol.* **2007**, *47*, 1−51.

(9) Jeffrey Conn, P.; Christopoulos, A.; Lindsley, C. W. Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nat. Rev. Drug Discovery* **2009**, *8* (1), 41−54.

(10) Keov, P.; Sexton, P.; Christopoulos, A. Allosteric modulation of G protein-coupled receptors: A pharmacological perspective. *Neuropharmacology* **2011**, *60* (1), 24−35.

(11) Goblyos, A.; Ijzerman, A. P. Allosteric modulation of adenosine receptors. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808* (5), 1309−1318.
(12) Nawaratne, V.; Leach, K.; Felder, C. C.; Sexton, P. M.; Christopoulos, A. Structural determinants of allosteric agonism and modulation at the M4 muscarinic acetylcholine receptor: identification of ligand-specific and global activation mechanisms. *J. Biol. Chem.* **2010**, *285* (25), 19012−19021.
(13) Wakefield, A. E.; Bajusz, D.; Kozakov, D.; Keseru, G. M.; Vajda, S. Conservation of Allosteric Ligand Binding Sites in G-Protein Coupled Receptors. *J. Chem. Inf. Model.* **2022**, *62* (20), 4937−4954.
(14) Valant, C.; Felder, C. C.; Sexton, P. M.; Christopoulos, A. Probe Dependence in the Allosteric Modulation of a G Protein-Coupled Receptor: Implications for Detection and Validation of Allosteric Ligand Effects. *Mol. Pharmacol.* **2012**, *81* (1), 41−52.
(15) Thal, D. M.; Glukhova, A.; Sexton, P. M.; Christopoulos, A. Structural insights into G-protein-coupled receptor allostery. *Nature* **2018**, *559* (7712), 45−53.
(16) Kruse, A. C.; Ring, A. M.; Manglik, A.; Hu, J.; Hu, K.; Eitel, K.; Hübner, H.; Pardon, E.; Valant, C.; Sexton, P. M.; Christopoulos, A.; Felder, C. C.; Gmeiner, P.; Steyaert, J.; Weis, W. I.; Garcia, K. C.; Wess, J.; Kobilka, B. K. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **2013**, *504* (7478), 101−106.
(17) Maeda, S.; Qu, Q.; Robertson, M. J.; Skiniotis, G.; Kobilka, B. K. Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes. *Science* **2019**, *364* (6440), 552−557.
(18) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in Neurosciences*; Stuart, C. S. Ed.; Academic Press: 1995; Vol. *25*, pp 366−428.
(19) Wootten, D.; Simms, J.; Miller, L. J.; Christopoulos, A.; Sexton, P. M. Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 5211−5216.
(20) Maeda, S.; Xu, J.; Fm, N. K.; Clark, M. J.; Zhao, J.; Tsutsumi, N.; Aoki, J.; Sunahara, R. K.; Inoue, A.; Garcia, K. C.; Kobilka, B. K. Structure and selectivity engineering of the M1 muscarinic receptor toxin complex. *Science* **2020**, *369* (6500), 161−167.
(21) Lu, J.; Byrne, N.; Wang, J.; Bricogne, G.; Brown, F. K.; Chobanian, H. R.; Colletti, S. L.; Di Salvo, J.; Thomas-Fowlkes, B.; Guo, Y.; Hall, D. L.; Hadix, J.; Hastings, N. B.; Hermes, J. D.; Ho, T.; Howard, A. D.; Josien, H.; Kornienko, M.; Lumb, K. J.; Miller, M. W.; Patel, S. B.; Pio, B.; Plummer, C. W.; Sherborne, B. S.; Sheth, P.; Souza, S.; Tummala, S.; Vonrhein, C.; Webb, M.; Allen, S. J.; Johnston, J. M.; Weinglass, A. B.; Sharma, S.; Soisson, S. M. Structural basis for the cooperative allosteric activation of the free fatty acid receptor GPR40. *Nat. Struct. Mol. Biol.* **2017**, *24* (7), 570−577.
(22) Liu, H.; Kim, H. R.; Deepak, R.; Wang, L.; Chung, K. Y.; Fan, H.; Wei, Z.; Zhang, C. Orthosteric and allosteric action of the C5a receptor antagonists. *Nat. Struct Mol. Biol.* **2018**, *25* (6), 472−481.
(23) Bueno, A. B.; Sun, B.; Willard, F. S.; Feng, D.; Ho, J. D.; Wainscott, D. B.; Showalter, A. D.; Vieth, M.; Chen, Q.; Stutsman, C.; et al. Structural insights into probe-dependent positive allosterism of the GLP-1 receptor. *Nat. Chem. Biol.* **2020**, *16* (10), 1105−1110.
(24) Jazayeri, A.; Doré, A. S.; Lamb, D.; Krishnamurthy, H.; Southall, S. M.; Baig, A. H.; Bortolato, A.; Koglin, M.; Robertson, N. J.; Errey, J. C.; Andrews, S. P.; Teobald, I.; Brown, A. J. H.; Cooke, R. M.; Weir, M.; Marshall, F. H. Extra-helical binding site of a glucagon receptor antagonist. *Nature* **2016**, *533*, 274−277.
(25) Draper-Joyce, C.; Bhola, R.; Wang, J.; Bhattarai, A.; Nguyen, A.; Cowie-Kent, I.; O'Sullivan, K.; Venugopal, H.; Valant, C.; Thal, D.; et al. Positive allosteric mechanisms of adenosine A1 receptor-mediated analgesia. *Nature* **2021**, *597*, 571−576.
(26) Wang, J.; Wu, M.; Chen, Z.; Wu, L.; Wang, T.; Cao, D.; Wang, H.; Liu, S.; Xu, Y.; Li, F.; et al. The unconventional activation of the muscarinic acetylcholine receptor M4R by diverse ligands. *Nat. Commun.* **2022**, *13*, 2855 DOI: 10.1038/s41467-022-30595-y.
(27) Vuckovic, Z.; Wang, J.; Pham, V.; Mobbs, J. I.; Belousoff, M. J.; Bhattarai, A.; Burger, W. A. C.; Thompson, G.; Yeasmin, M.; Leach, K.; et al. Pharmacological hallmarks of allostery at the M4 muscarinic receptor elucidated through structure and dynamics. *eLife* **2023**, *12*, No. e83477.
(28) Liu, X.; Ahn, S.; Kahsai, A. W.; Meng, K. C.; Latorraca, N. R.; Pani, B.; Venkatakrishnan, A. J.; Masoudi, A.; Weis, W. I.; Dror, R. O.; et al. Mechanism of intracellular allosteric beta2AR antagonist revealed by X-ray crystal structure. *Nature* **2017**, *548*, 480−484.
(29) Liu, X.; Masoudi, A.; Kahsai, A. W.; Huang, L. Y.; Pani, B.; Staus, D. P.; Shim, P. J.; Hirata, K.; Simhal, R. K.; Schwalb, A. M.; et al. Mechanism of beta2AR regulation by an intracellular positive allosteric modulator. *Science* **2019**, *364*, 1283−1287.
(30) Liu, X.; Kaindl, J.; Korczynska, M.; Stossel, A.; Dengler, D.; Stanek, M.; Hubner, H.; Clark, M. J.; Mahoney, J.; Matt, R. A.; et al. An allosteric modulator binds to a conformational hub in the beta2adrenergic receptor. *Nat. Chem. Biol.* **2020**, *16*, 749−755.
(31) Liu, H.; Kim, H. R.; Deepak, R.; Wang, L.; Chung, K. Y.; Fan, H.; Wei, Z.; Zhang, C. Orthosteric and allosteric action of the C5a receptor antagonists. *Nat. Struct. Mol. Biol.* **2018**, *25* (6), 472−481.
(32) Shao, Z.; Yan, W.; Chapman, K.; Ramesh, K.; Ferrell, A. J.; Yin, J.; Wang, X.; Xu, Q.; Rosenbaum, D. M. Structure of an allosteric modulator bound to the CB1 cannabinoid receptor. *Nat. Chem. Biol.* **2019**, *15*, 1199−1205.
(33) Zheng, Y.; Qin, L.; Zacarias, N. V.; de Vries, H.; Han, G. W.; Gustavsson, M.; Dabros, M.; Zhao, C.; Cherney, R. J.; Carter, P.; et al. Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. *Nature* **2016**, *540*, 458−461.
(34) Zhuang, Y.; Krumm, B.; Zhang, H.; Zhou, X. E.; Wang, Y.; Huang, X. P.; Liu, Y.; Cheng, X.; Jiang, Y.; Jiang, H.; et al. Mechanism of dopamine binding and allosteric modulation of the human D1 dopamine receptor. *Cell Res.* **2021**, *31*, 593−596.
(35) Yang, F.; Mao, C.; Guo, L.; Lin, J.; Ming, Q.; Xiao, P.; Wu, X.; Shen, Q.; Guo, S.; Shen, D. D.; et al. Structural basis of GPBAR activation and bile acid recognition. *Nature* **2020**, *587*, 499−504.
(36) Song, G.; Yang, D.; Wang, Y.; de Graaf, C.; Zhou, Q.; Jiang, S.; Liu, K.; Cai, X.; Dai, A.; Lin, G.; et al. Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators. *Nature* **2017**, *546*, 312−315.
(37) Wu, F.; Yang, L.; Hang, K.; Laursen, M.; Wu, L.; Han, G. W.; Ren, Q.; Roed, N. K.; Lin, G.; Hanson, M. A.; et al. Full-length human GLP-1 receptor structure without orthosteric ligands. *Nat. Commun.* **2020**, *11*, 1272 DOI: 10.1038/s41467-020-14934-5.
(38) Changeux, J.-P.; Christopoulos, A. Allosteric Modulation as a Unifying Mechanism for Receptor Function and Regulation. *Cell* **2016**, *166* (5), 1084−1102.
(39) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9* (9), 646−652.
(40) Grossfield, A. Recent progress in the study of G protein-coupled receptors with molecular dynamics computer simulations. *Biochim. Biophys. Acta, Biomembr.* **2011**, *1808* (7), 1868−1878.
(41) Johnston, J. M.; Filizola, M. Showcasing modern molecular dynamics simulations of membrane proteins through G protein-coupled receptors. *Curr. Opin. Struct. Biol.* **2011**, *21* (4), 552−558.
(42) Vanni, S.; Rothlisberger, U. A Closer Look into G Protein Coupled Receptor Activation: X-Ray Crystallography and Long-Scale Molecular Dynamics Simulations. *Curr. Med. Chem.* **2012**, *19* (8), 1135−1145.
(43) Latorraca, N. R.; Venkatakrishnan, A.; Dror, R. O. GPCR dynamics: structures in motion. *Chem. Rev.* **2017**, *117* (1), 139−155.
(44) Miao, Y.; McCammon, J. A. G-protein coupled receptors: advances in simulation and drug discovery. *Curr. Opin. Struct. Biol.* **2016**, *41*, 83−89.
(45) Lamim Ribeiro, J. M.; Filizola, M. Allostery in G protein-coupled receptors investigated by molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2019**, *55*, 121−128.
(46) Dror, R. O.; Green, H. F.; Valant, C.; Borhani, D. W.; Valcourt, J. R.; Pan, A. C.; Arlow, D. H.; Canals, M.; Lane, J. R.; Rahmani, R.; et al.

Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature* **2013**, *503* (7475), 295−299.
(47) Shang, Y.; LeRouzic, V.; Schneider, S.; Bisignano, P.; Pasternak, G. W.; Filizola, M. Mechanistic Insights into the Allosteric Modulation of Opioid Receptors by Sodium Ions. *Biochemistry* **2014**, *53* (31), 5140−5149.
(48) Selent, J.; Sanz, F.; Pastor, M.; De Fabritiis, G. Induced effects of sodium ions on dopaminergic G-protein coupled receptors. *PLoS Comput. Biol.* **2010**, *6* (8), No. e1000884.
(49) Miao, Y.; Caliman, A. D.; McCammon, J. A. Allosteric Effects of Sodium Ion Binding on Activation of the M3Muscarinic G-Protein Coupled Receptor. *Biophys. J.* **2015**, *108* (7), 1796−1806.
(50) 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 DOI: 10.1038/s41598-018-35266-x.
(51) Shang, Y.; Yeatman, H. R.; Provasi, D.; Alt, A.; Christopoulos, A.; Canals, M.; Filizola, M. Proposed Mode of Binding and Action of Positive Allosteric Modulators at Opioid Receptors. *ACS Chem. Biol.* **2016**, *11* (5), 1220−1229.
(52) An, X.; Bai, Q.; Bing, Z.; Liu, H.; Zhang, Q.; Liu, H.; Yao, X. Revealing the Positive Binding Cooperativity Mechanism between the Orthosteric and the Allosteric Antagonists of CCR2 by Metadynamics and Gaussian Accelerated Molecular Dynamics Simulations. *ACS Chem. Neurosci.* **2020**, *11* (4), 628−637.
(53) Do, H. N.; Wang, J.; Bhattarai, A.; Miao, Y. GLOW: A Workflow Integrating Gaussian-Accelerated Molecular Dynamics and Deep Learning for Free Energy Profiling. *J. Chem. Theory Comput.* **2022**, *18* (3), 1423−1436.
(54) 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*, 3584−3595.
(55) 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*, 2677−2689.
(56) Wang, J.; Arantes, P.; Bhattarai, A.; Hsu, R.; Pawnikar, S.; Huang, Y.-M.; Palermo, G.; Miao, Y. Gaussian accelerated molecular dynamics: principles and applications. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11*, No. e1521, DOI: 10.1002/wcms.1521.
(57) Keras-Vis; GitHub: 2017. https://github.com/raghakot/keras-vis.
(58) Huang, C. Y.; Olieric, V.; Ma, P.; Howe, N.; Vogeley, L.; Liu, X.; Warshamange, R.; Weinert, T.; Panepucci, E.; Kobilka, B.; et al. In meso in situ serial X-ray crystallography of soluble and membrane proteins at cryogenic temperatures. *Acta Crystallogr., Sect. D: Struct. Biol.* **2016**, *72*, 93−112.
(59) Ishchenko, A.; Stauch, B.; Han, G. W.; Batyuk, A.; Shiriaeva, A.; Li, C.; Zatsepin, N.; Weierstall, U.; Liu, W.; Nango, E.; et al. Toward G protein-coupled receptor structure-based drug design using X-ray lasers. *IUCrJ* **2019**, *6*, 1106−1119.
(60) Wang, X.; Liu, D.; Shen, L.; Li, F.; Li, Y.; Yang, L.; Xu, T.; Tao, H.; Yao, D.; Wu, L.; et al. A Genetically Encoded F-19 NMR Probe Reveals the Allosteric Modulation Mechanism of Cannabinoid Receptor 1. *J. Am. Chem. Soc.* **2021**, *143*, 16320−16325.
(61) Suno, R.; Lee, S.; Maeda, S.; Yasuda, S.; Yamashita, K.; Hirata, K.; Horita, S.; Tawaramoto, M. S.; Tsujimoto, H.; Murata, T.; et al. Structural insights into the subtype-selective antagonist binding to the M2 muscarinic receptor. *Nat. Chem. Biol.* **2018**, *14*, 1150−1158.
(62) Thal, D. M.; Sun, B.; Feng, D.; Nawaratne, V.; Leach, K.; Felder, C. C.; Bures, M. G.; Evans, D. A.; Weis, W. I.; Bachhawat, P.; et al. Crystal structures of the M1 and M4 muscarinic acetylcholine receptors. *Nature* **2016**, *531*, 335−340.
(63) Deluigi, M.; Morstein, L.; Schuster, M.; Klenk, C.; Merklinger, L.; Cridge, R. R.; de Zhang, L. A.; Klipp, A.; Vacca, S.; Vaid, T. M.; et al. Crystal structure of the alpha1B-adrenergic receptor reveals molecular determinants of selective ligand recognition. *Nat. Commun.* **2022**, *13*, 382 DOI: 10.1038/s41467-021-27911-3.
(64) Qu, L.; Zhou, Q. T.; Wu, D.; Zhao, S. W. Crystal structures of the alpha2A adrenergic receptor in complex with an antagonist RSC. *To be published* 2019.
(65) Chen, X. Y.; Wu, D.; Wu, L. J.; Han, G. W.; Guo, Y.; Zhong, G. S. Crystal structure of human alpha2C adrenergic G-protein-coupled receptor. *To be published*, 2019.
(66) Xu, X.; Kaindl, J.; Clark, M. J.; Hubner, H.; Hirata, K.; Sunahara, R. K.; Gmeiner, P.; Kobilka, B. K.; Liu, X. Binding pathway determines norepinephrine selectivity for the human beta1AR over beta2AR. *Cell Res.* **2021**, *31*, 569−579.
(67) Liu, X.; Xu, X.; Hilger, D.; Aschauer, P.; Tiemann, J. K. S.; Du, Y.; Liu, H.; Hirata, K.; Sun, X.; Guixa-Gonzalez, R.; et al. Structural Insights into the Process of GPCR-G Protein Complex Formation. *Cell* **2019**, *177*, 1243−1251.e1212.
(68) Zhuang, Y.; Xu, P.; Mao, C.; Wang, L.; Krumm, B.; Zhou, X. E.; Huang, S.; Liu, H.; Cheng, X.; Huang, X. P.; et al. Structural insights into the human D1 and D2 dopamine receptor signaling complexes. *Cell* **2021**, *184*, 931−942.e918.
(69) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.; de Beer, T.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46* (W1), W296−W303.
(70) Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **2003**, *66*, 27−85.
(71) Tian, C.; Kasavajhala, K.; Belfon, K.; Raguette, L.; Huang, H.; Migues, A.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16* (1), 528−552.
(72) Wang, J.; Wolf, R.; Caldwell, J.; Kollman, P.; Case, D. Development and testing of a general AMBER force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(73) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23* (16), 1623−1641.
(74) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(75) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmuller, H.; MacKerell, A. D., Jr CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71−73.
(76) 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*, 3144−3154.
(77) 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*, 3155−3168.
(78) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Systematic Validation of Protein Force Fields against Experimental Data. *PLoS One* **2012**, *7* (2), No. e32131.
(79) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* **2011**, *100* (9), L47−L49.
(80) Chen, W.; Shi, C.; MacKerell, A. D., Jr; Shen, J. Conformational Dynamics of Two Natively Unfolded Fragment Peptides: Comparison of the AMBER and CHARMM Force Fields. *J. Phys. Chem. B* **2015**, *119* (25), 7902−7910.
(81) Sandoval-Perez, A.; Pluhackova, K.; Bockmann, R. A. Critical Comparison of Biomembrane Force Fields: Protein−Lipid Interactions at the Membrane Interface. *J. Chem. Theory Comput.* **2017**, *13* (5), 2310−2321.
(82) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.
(83) Evans, D. J. Computer "experiment" for nonlinear thermodynamics of Couette flow. *J. Chem. Phys.* **1983**, *78*, 3297.

(84) Hoover, W. G.; Ladd, A. J. C.; Moran, B. High-Strain-Rate Plastic Flow Studied via Nonequilibrium Molecular Dynamics. *Phys. Rev. Lett.* **1982**, *48*, 1818 DOI: 10.1103/PhysRevLett.48.1818.
(85) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684−3690.
(86) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577−8593.
(87) Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routined microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(88) *Amber 2021*; University of California: San Francisco, 2021.
(89) Xiaoli, A.; Yuzhen, N.; Qiong, Y.; Yang, L.; Yao, X.; Bing, Z. Investigating the Dynamic Binding Behavior of PMX53 Cooperating with Allosteric Antagonist NDT9513727 to C5a Anaphylatoxin Chemotactic Receptor 1 through Gaussian Accelerated Molecular Dynamics and Free-Energy Perturbation Simulations. *ACS Chem. Neurosci.* **2022**, *13*, 3502−3511.
(90) Do, H. N.; Akhter, S.; Miao, Y. Pathways and Mechanism of Caffeine Binding to Human Adenosine A2A Receptor. *Front. Mol. Biosci.* **2021**, *8*, No. 673170, DOI: 10.3389/fmolb.2021.673170.
(91) Nguyen, A.; Vecchio, E.; Thomas, T.; Nguyen, T.; Aurelio, L.; Scammells, P.; White, P.; Sexton, P.; Gregory, K.; May, L.; Christopoulos, A. Role of the Second Extracellular Loop of the Adenosine A1 Receptor on Allosteric Modulator Binding, Signaling, and Cooperativity. *Mol. Pharmacol.* **2016**, *90* (6), 715−725.
(92) Nguyen, A.; Baltos, J.-A.; Thomas, T.; Nguyen, T.; Muñoz, L.; Gregory, K.; White, P.; Sexton, P.; Christopoulos, A.; May, L. Extracellular Loop 2 of the Adenosine A1 Receptor Has a Key Role in Orthosteric Ligand Affinity and Agonist Efficacy. *Mol. Pharmacol.* **2016**, *90* (6), 703−714.
(93) Schober, D. A.; Croy, C. H.; Xiao, H.; Christopoulos, A.; Felder, C. C. Development of a radioligand, [(3)H]LY2119620, to probe the human M(2) and M(4) muscarinic receptor allosteric binding sites. *Mol. Pharmacol.* **2014**, *86* (1), 116−123.
(94) Amaro, R. E.; Baudry, J.; Chodera, J. D.; Demir, O.; McCammon, J. A.; Miao, Y.; Smith, J. C. Ensemble Docking in Drug Discovery. *Biophys. J.* **2018**, *114* (10), 2271−2278.

https://doi.org/10.1021/jacsau.3c00503
JACS Au 2023, 3, 3165−3180