Hidden GPCR structural transitions addressed by multiple walker supervised molecular dynamics (mwSuMD)

Giuseppe Deganutti[1], Ludovico Pipitò[1], Roxana M. Rujan[1], Tal Weizmann[1], Peter Griffin[1], Antonella Ciancetta[2], Stefano Moro[3], and Christopher A. Reynolds[1]**

[1] Centre for sport, Exercise and Life Sciences, Faculty of Health and Life Sciences, Coventry University, Alison Gingell Building, Coventry, CV1 5FB, UK

[2] Dipartimento di Scienze Chimiche, Farmaceutiche ed Agrarie, University of Ferrara, Via Luigi Borsari 46, 44121 Ferrara, Italy

[3] Molecular Modeling Section (MMS), Dipartimento di Scienze del Farmaco, University of Padua via Marzolo 5, 35131, Padova, Italy

*Corresponding Authors:
Giuseppe Deganutti - ad5288@coventry.ac.uk
Christopher A. Reynolds - ad5291@coventry.ac.uk

Keywords:
G protein-coupled receptors; binding; protein activation; molecular dynamics; supervised molecular dynamics

Abstract

G protein-coupled receptors (GPCRs) are the most abundant membrane proteins and the target of about 35% of approved drugs, but the structural basis of GPCR pharmacology is still a matter of intense study. Here, we present an unbiased molecular dynamics adaptive sampling algorithm, namely multiple walker supervised molecular dynamics (mwSuMD), that performs well on different hidden transitions involving GPCRs. Molecular dynamics (MD) simulations aim at expanding the knowledge of GPCR dynamics by building upon the recent advances in structural biology. However, the timescale limitations of classic MD hinder its applicability to numerous structural processes happening in time scales longer than microseconds (hidden structural transitions), limiting the overall MD impact on the study of GPCRs, hence our new algorithm. By increasing the complexity of the simulated process, we report the binding and unbinding of the vasopressin peptide from its receptor V_2, the inactive-to-active transition of the glucagon-like peptide-1 receptor (GLP-1R), and the stimulatory (G_s) and inhibitory (G_i) G proteins binding to the adrenoreceptor β_2 (β_2 AR) and the adenosine 1 receptor (A_1R), respectively. Finally, we report on the heterodimerization between the adenosine receptor A_2 (A_{2A}R) and the dopamine receptor D_2 (D_2R). We demonstrate mwSuMD usefulness for studying atomic-level GPCR transitions that are challenging to address with classic MD simulations.

Introduction

Supervised molecular dynamics[1,2] (SuMD) is a powerful technique for studying ligand-receptor binding and unbinding pathways; here we present a significant enhancement to the method, namely multiple walker supervised molecular dynamics (mwSuMD) that permits a wider range of conformational transitions relevant to drug design to be studied. We validated the method

by applying it to G protein-coupled receptors (GPCRs), as these are both fundamental drug targets and well-validated test systems. GPCRs are the most abundant family of membrane receptors in eukaryotes[3] and the target for more than one-third of drugs approved for human use[4]. Vertebrate GPCRs are subdivided into five subfamilies (Rhodopsin or class A, Secretin or class B, Glutamate or class C, Adhesion, and Frizzled/Taste2) according to function and sequence[5,6]. Common features of all GPCRs are seven transmembrane (TM) helices connected by three extracellular loops (ECLs) and three intracellular loops (ICLs), while an extended and structured N-terminus extracellular domains (ECD) is found in all subtypes, but class A. The primary function of GPCRs is transducing extracellular chemical signals into the cytosol by binding and activating four G protein families ($G_{s/olf}$, $G_{i/o}$, $G_{12/13}$ and $G_{q/11}$) responsible for decreasing ($G_{i/o}$) or increasing ($G_{s/olf}$) the cyclic adenosine-3',5'-monophosphate (cAMP), and generating inositol-1,4,5-triphosphate (IP_3) and diacylglycerol (DAG) to increase Ca^{2+} intracellular levels (G_q)[7].

GPCR structures have been solved by X-ray and cryo-electron microscopy (cryo-EM) at an increasing pace since the first X-ray structures in 2000[8] and 2007[9]. However, many aspects of their pharmacology remain elusive. For example, the structural determinants of the selectivity displayed towards specific G proteins or the ability of certain agonists to drive a preferred intracellular signaling pathway over the others (i.e. functional selectivity or bias)[10]. What makes GPCRs challenging proteins to characterize with standard techniques is their inherent flexibility and the transitory nature of the complexes formed with extracellular and intracellular effectors. One of the possible approaches to integrate or sometimes overcome the limits of experimental conditions is performing molecular dynamics (MD) simulations. MD is a computational methodology that predicts the movement and interactions of (bio)molecules in systems of variable complexity, at atomic detail, enabling useful working hypotheses and rationalization of experimental data. However, standard MD sampling is limited to the microsecond or, in the best conditions, the millisecond time scale[11,12]. For this reason, different algorithms have been designed to speed up the simulation of rare events such as ligand (un)binding and conformational transitions. Amongst the most popular and effective ones, there are metadynamics[13], accelerated MD (aMD)[14], and Gaussian-accelerated MD (GaMD)[15]. Such methods, which introduce an energy potential to overcome the energy barriers preventing the complete exploration of the free energy surface, thus *de facto* biasing the simulation, have been used to propose activation mechanisms of GPCRs[16,17]. Energetically unbiased MD protocols, on the other hand, comprise the weighted ensemble MD (weMD)[18] and SuMD[1,19]. SuMD has been successfully applied to the (un)binding mechanism of both small molecules, peptides, and small proteins[1,19–23]. Since SuMD is optimized only for (un)bindings, we have designed a new version of the software, namely multiple walker SuMD (mwSuMD), that extends the applicability of the method to conformational transitions and protein:protein binding.

We tested mwSuMD on a series of increasingly complex hidden structural transitions involving both class A and class B1 GPCRs. Firstly, we validated the method on the nonapeptide arginine vasopressin (AVP) by simulating binding (dynamic docking) and unbinding paths from the vasopressin 2 receptor (V_2R). AVP is an endogenous hormone (**Figure S1a)** that mediates antidiuretic effects on the kidney by signaling through three class A GPCR subtypes: V_{1a} and V_{1b} receptors activate phospholipases via $G_{q/11}$ protein, while the V_2 receptor (V_2R) activates adenylyl cyclase by interacting with G_s protein [24] and is a therapeutic target for hyponatremia, hypertension, and incontinence[25]. Dynamic docking, although more computationally demanding than standard molecular docking, provides insights into the binding mode of ligands in a fully hydrated and flexible environment. Moreover, it

informs about binding paths and the complete mechanism of formation leading to an intermolecular complex, delivering in the context of binding kinetics[26] and structure-kinetics relationship (SKR) studies[27].

We then show that mwSuMD can be employed to simulate the receptor activation of the class B1 GPCR glucagon-like peptide-1 receptor (GLP-1R) upon binding of the small molecule PF06882961. GLP-1R is a validated target in type 2 diabetes and probably the best-characterized class B1 GPCR from a structural perspective. GLP-1R is the only class B1 receptor with structurally characterized non-peptidic orthosteric agonists, which makes it a model system for studying the druggability of the entire B1 subfamily.

The further case studies we report are the G_s and G_i proteins binding to the adrenoreceptor β_2 (β_2 AR) and the adenosine 1 receptor (A_1R), starting from different conditions. GPCRs preferentially couple to very few G proteins out of 23 possible counterparts[28,29]. More importantly, agonists can modify the receptor selectivity profile by imprinting unique intracellular conformations from the orthosteric binding site. The mechanism behind these phenomena is one of the outstanding questions in the GPCR field[28]. It is increasingly accepted that dynamic and transient interactions determine whether the encounter between a GPCR and a G protein results in productive or unproductive coupling[30]. MD simulations are considered a useful orthogonal tool for providing working hypotheses and rationalizing existing data on G protein selectivity. However, so far, it has not delivered as expected. Attempts so far have employed energetically biased simulations or have been confined to the $G\alpha$ subunit[16,17].

The last GPCR key process simulated through mwSuMD is the heterodimerization in the membrane between the adenosine receptor A_2 ($A_{2A}R$) and the dopamine receptor D_2 (D_2R). The $A_{2A}R$:D_2R heterodimer[31] is a therapeutic target for neurodegenerative diseases, Parkinson's disease, and schizophrenia[32–34] due to the reciprocal antagonistic allosteric effect between monomers[35]. $A_{2A}R$ activation reduces the binding affinity of D_2R agonists, while $A_{2A}R$ antagonists enhance the dopaminergic tone by decreasing the adenosine negative allosteric modulation on D_2R. Heterobivalent ligands able to inhibit $A_{2A}R$ and activate D_2R represent a valuable pharmacological tool[36] and, in principle, therapeutic options for conditions characterized by reduction of dopaminergic signaling in the central nervous system. The successive dynamic docking of the heterobivalent ligand compound 26[37] to the heterodimer suggested by mwSuMD produced a ternary complex stabilized by lipids.

Results and Discussion

Short mwSuMD time windows improve the AVP dynamic docking prediction

AVP has an amphipathic nature and interacts with both polar and hydrophobic V_2R residues located on both TM helices and ECLs (**Figure S1b**). Although AVP presents an intramolecular C1-C6 disulfide bond that limits the overall conformational flexibility of the backbone, it has a high number of rotatable bonds, making dynamic docking complicated[38]. We assessed the performance of mwSuMD and the original version of SuMD in reconstructing the experimental V_2R:AVP complex using different settings, simulating a total of 92 binding events (**Table S1**). As a reference, the AVP RMSD during a classic (unsupervised) equilibrium MD simulation of the AVP:V_2R complex was 3.80 ± 0.52 Å (**Figure S2**). SuMD[1,19] produced a minimum root mean square deviation (RMSD) to the cryo-EM complex of 4.28 Å, with most of the replicas (distribution mode) close to 10 Å (**Figure 1a**). MwSuMD, with the same settings (**Figure 1b, Table S1**) in terms of time window duration (600 ps), metric supervised (the distance between AVP and V_2R), and acceptance method (slope) produced slightly more precise results (distribution mode RMSD = 7.90 Å) but similar accuracy (minimum RMSD = 4.60). Supervising the AVP RMSD to the experimental complex rather than the distance (**Figure 1c**) and using the SMscore (**Equation 1**) as the acceptance method (**Figure 1d**) worsened the prediction. Supervising distance and RMSD at the same time (**Figure 1e**), employing the DMscore (**Equation 2**), recovered accuracy (minimum RMSD = 4.60 Å) but not precision (distribution mode RMSD = 12.40 Å). Interestingly, decreasing the time window duration from 600 ps to 100 ps impaired the SuMD ability to predict the experimental complex (**Figure 2a**), but enhanced mwSuMD accuracy and precision (**Figure 2b-d**). The combination of RMSD as the supervised metric and SMscore produced the best results in terms of minimum RMSD and distribution mode RMSD, 3.85 Å and 4.40 Å, respectively (**Figure 2d, Video S1**), in agreement with the AVP deviations in the equilibrium MD simulation of the AVP:V_2R complex.

These results suggest that short time windows can dramatically improve the dynamic docking performance of mwSuMD. However, it is necessary to know the final bound state to employ the RMSD, while the distance as the supervised metric is required to dynamically dock ligands with unknown bound conformation. Both distance and RMSD-based simulations delivered insights into the binding path and the residues involved along the recognition route. For example, mwSuMD suggested V_2R residues E184^{ECL2}, P298^{ECL3}, and E303^{ECL3} (**Figure S3a**) as involved during AVP binding, although not in contact with the ligand in the orthosteric complex.

Further to binding, a SuMD approach was previously employed to reconstruct the unbinding path of ligands from several GPCRs [1,2,39]. We assessed mwSuMD capability to simulate AVP unbinding from V_2R. Five mwSuMD and five SuMD replicas were collected using 100 ps time windows (**Table 1**). Overall, mwSuMD outperformed SuMD in terms of time required to complete a dissociation (**Figure S4, Video S2**), producing dissociation paths almost 10-fold faster than SuMD**.** Such rapidity in dissociating inherently produces a limited sampling of metastable states along the pathway, which can be compensated by seeding classic (unsupervised) MD simulations from configurations extracted from the unbinding route[40,41]. Here, the set of V_2R residues involved during the dissociation was comparable to the binding (**Figure S3b**), though ECL2 and ECL3 were slightly more involved during the association than the dissociation, in analogy with other class A and B GPCRs[21,40].

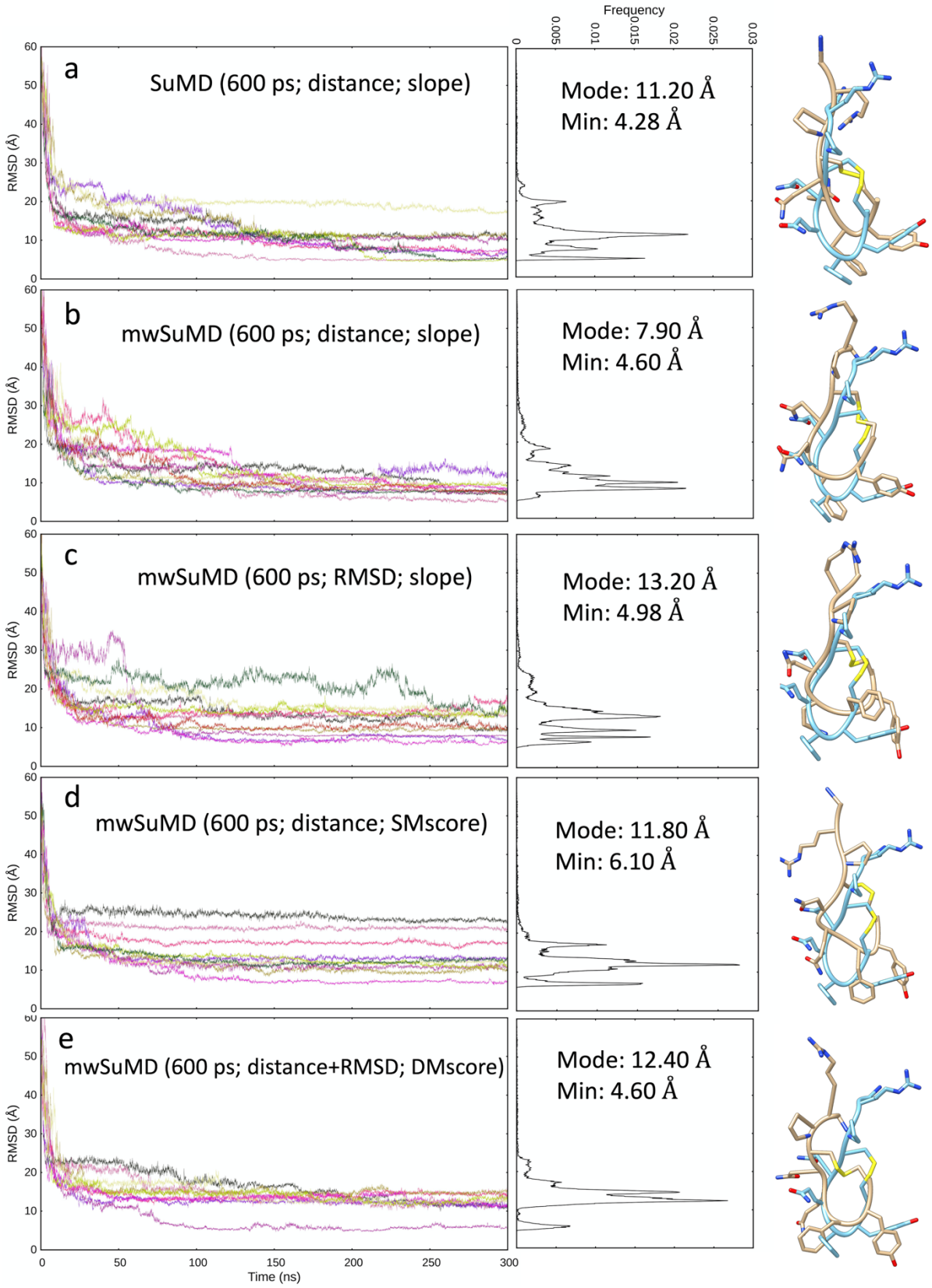

Figure 1. AVP SuMD and mwSuMD binding simulations to V_2R (600 ps time windows). For each set of settings **(a-e),** the RMSD of AVP Cα atoms to the cryo-EM structure 7DW9 is reported during the time course of each SuMD (**a**) or mwSuMD (**b-e**) replica alongside the RMSD values distribution and the snapshot corresponding to the lowest RMSD values (AVP from the cryo-EM structure 7DW9 in cyan

stick representation, while AVP from simulations in a tan stick). A complete description of the simulation settings is reported in Table 1 and the Methods section.

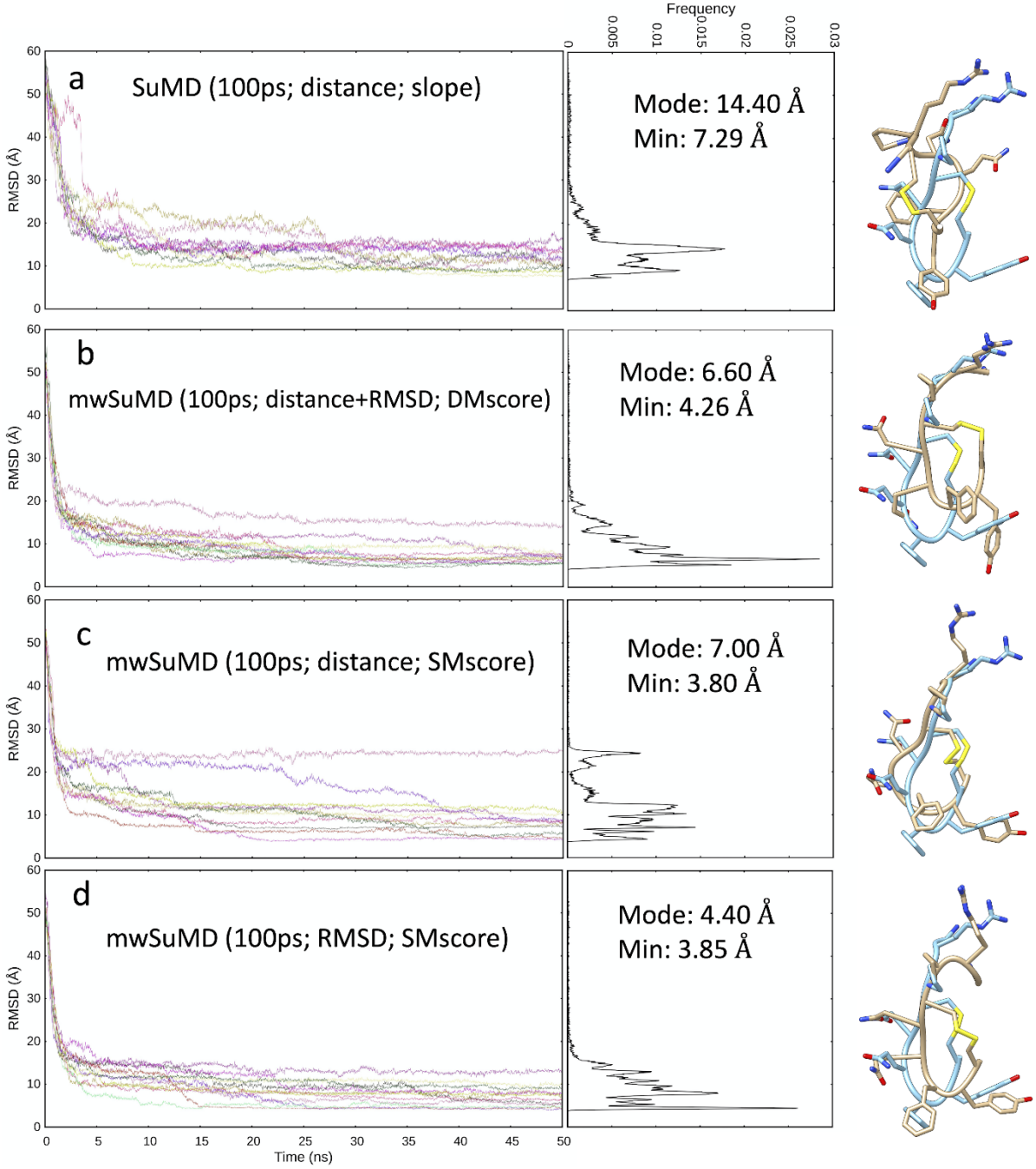

Figure 2. AVP SuMD and mwSuMD binding simulations to V_2R (100 ps time windows). For each set of settings **(a-d)** the RMSD of AVP Cα atoms to the cryo-EM structure 7DW9 is reported during the time course of each SuMD (**a**) or mwSuMD (**b-d**) replica alongside the RMSD values distribution and the snapshot corresponding to the lowest RMSD values (AVP from the cryo-EM structure 7DW9 in cyan stick representation, while AVP from simulations in a tan stick). A complete description of the simulation settings is reported in Table 1 and the Methods section.

PF06882961 binding and GLP-1R activation

The GLP-1R has been captured by cryo-EM in both the inactive apo (ligand-free) and the active (G_s-bound) conformations, and in complex with either peptides or non-peptide agonists[42–47]. In the inactive apo GLP-1R, residues forming the binding site for the non-peptide agonist PF06882961 are dislocated and scattered due to the structural reorganization of the transmembrane domain (TMD) and extracellular domain (ECD) (**Figure S5**) that occurs on activation. Moreover, GLP-1R in complex with GLP-1 or different agonists present distinct structural features, even amongst structurally related ligands (**Figure S6**). This complicates the scenario and suggests divergent recognition mechanisms amongst different agonists. We simulated the binding of PF06882961 using multistep supervision on different metrics of the system (**Figure 3**) to model the structural hallmark of GLP-1R activation (**Video S3**, **Video S4**).

Several metrics were supervised in a consecutive fashion. Firstly, the distance between PF06882961 and the TMD as well as the RMSD of the ECD to the active state (stage 1); secondly, the RMSD of ECD and ECL1 to the active state (stage 2); thirdly, the RMSD of PF06882961 and ECL3 to the active state (stage 3); lastly, only the RMSD of TM6 (residues I345-F367, Cα atoms) to the active state (stage 4). The combination of these supervisions produced a conformational transition of GLP-1R towards the active state (**Figure 3, Video S4**). Noteworthy, the sequence of these supervisions was arbitrary and does not necessarily reflect the right order of the steps involved in GLP-1R activation. This kind of planned multistep approach is feasible when the end-point receptor inactive and active structures are available, and the inherent flexibility of different domains is known. In class B GPCRs, the ECD is the most dynamic sub-structure, followed by the ECL1 and ECL3 which display high plasticity during ligand binding[21,48]. For this reason, we first supervised these elements of GLP-1R, leaving the bottleneck of activation, TM6 outward movement, as the last step. However, the protocol employed can be tweaked to study how each conformational transition takes place and influences the receptor domains. Structural elements not directly supervised, such as TM1 or TM7, displayed an RMSD reduction to the active state because they were influenced by the movement of supervised helixes or loops. For example, the supervision of ECL3 (stage 3) and TM6 (stage 4) facilitated the spontaneous rearrangement of the ECD to an active-like conformation after the ECD had previously experienced transient high flexibility during stages 2 and 3 (**Figure 3**).

During the supervision of ECL3 and PF06882961 (stage 3), we observed a loosening of the intracellular polar interactions that stabilize GLP-1R TM6 in the inactive state. As a result, the subsequent supervision of TM6 (residues I345-F367, Cα atoms) rapidly produced the outward movement towards the active state, in the last step of the mwSuMD simulation (stage 4). Taken together, these results suggest a concerted conformational transition for ECD and ECL1 during the binding of PF06882961 and an allosteric effect between ECL3 and the bottom of TM6. Interestingly, while the intracellular polar interactions were destabilized by the ECL3 transition to an active-like conformation (stages 2 and 3), the outward movement of TM6 (stage 4) did not favor the closure of ECL3 towards PF06882961, which appear to be driven by direct interactions between the ligand and R310$^{5.40}$ or R380$^{7.35}$. Since we were interested in reconstructing the binding of PF06882961 to GLP-1R and the successive receptor structural transitions to prepare the intracellular G protein binding site, our mwSuMD simulation did not include G_s. Therefore, any allosteric effect triggered by the binding of the effector could have

been overlooked, as well as the complete stabilization of TM6 in the active conformation, which is known to be achieved only when the intracellular effector is bound[49].

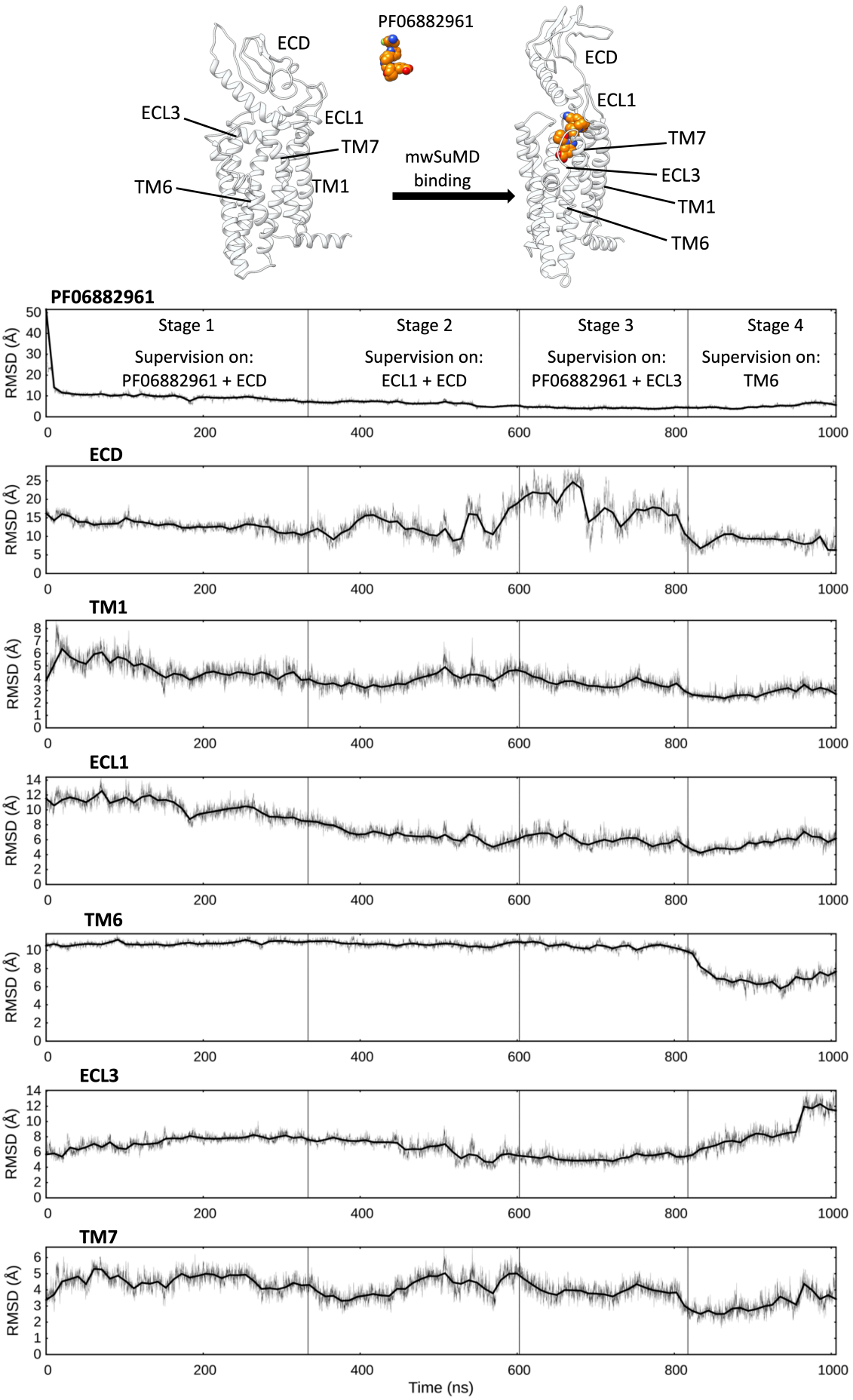
PF06882961
ECD
ECL3
ECL1
TM7
TM6
TM1
mwSuMD
binding
ECD
ECL1
TM7
ECL3
TM1
TM6
PF06882961
Stage 1
Supervision on:
PF06882961 + ECD
Stage 2
Supervision on:
ECL1 + ECD
Stage 3
Supervision on:
PF06882961 + ECL3
Stage 4
Supervision on:
TM6
ECD
TM1
ECL1
TM6
ECL3
TM7
RMSD (Å)
Time (ns)

Figure 3. MwSuMD simulation of PF06882961 binding to GLP-1R and receptor activation. Each panel reports the root mean square deviation (RMSD) to a GLP-1R structural element or the position of the ligand in the active state (top panel), over the time course (all but ECL3 converging to the active state). ECD: extracellular domain; TM: transmembrane helix; ECL: extracellular loop. The mwSuMD simulation was performed with four different settings over 1 microsecond in total.

G proteins — class A GPCR binding simulations

We tested the ability of mwSuMD to simulate the binding between the prototypical class A receptor, the β_2 adrenoreceptor (β_2 AR), and the stimulatory G protein (G_s), *without energy input*. mwSuMD simulations started from the intermediate, agonist-bound conformation of β_2 AR and the inactive G_s to resemble pre-coupling conditions. Three mwSuMD replicas were performed by supervising the distance between G_s helix 5 (H5) and β_2 AR as well as the RMSD of the intracellular end of TM6 to the fully-active state of the receptor (**Table S1**). To monitor the progression of the simulations, we computed the RMSD of the Cα atoms of the Gα and Gβ subunits to the experimental complex[50] (**Video S5, Figure 4ab**). During two out of three replicas, both Gα and Gβ reached values close to 5 Å (minimum RMSD = 3.94 Å and 3.96 Å respectively), in good agreement with the reference (the β_2 AR:G_s complex, PDB 3SN6, **Figure 4c).** The flexibility of $G_s\beta$ is backed by both MD and cryo-EM data suggesting G protein rocking motions around $G_s\alpha$:receptor interactions[21,51].

According to the model of G protein activation, the binding to the receptor allosterically stabilizes the orthosteric agonist, adrenaline in our simulations, and destabilizes the guanosine 5'-diphosphate (GDP) within Gα, resulting in the exchange with the ribonucleoside guanosine 5'triiphosphate (GTP) upon opening of the G protein alpha-helical domain (AHD). triggering the subsequent dissociation of Gα from G$\beta\gamma$. In our simulations, adrenaline was not further stabilized in the timescale of the simulations (**Figure 4d),** probably because the simulations sampled intermediate states, therefore, suboptimal β_2 AR:G_s interactions that were unable to allosterically stabilize the agonist. Upon receptor activation by the orthosteric agonist, TM6 undergoes an outward movement to accommodate the G protein that is accompanied by an anticlockwise rotation. We did not observe this rotation, which suggests that mwSuMD did not sample the complete G_s coupling. One of the β_2 AR residues undergoing rotation upon receptor activation is E268$^{6.30}$, involved in the conserved salt bridge (named ionic lock) with R131$^{3.50}$ that stabilizes the inactive state. Interestingly, during simulations, E268$^{6.30}$ formed hydrogen bonds with the G_s residues R385$^{H5.17}$, and R389$^{H5.21}$, both conserved across G protein subfamilies G_s, $G_{i/o}$, and $G_{q/11}$ (**Table S2**). We speculate that these interactions, not observed in any GPCR active state cryo-EM or X-ray structure, stabiles the early stage of G_s binding and that the TM6 full rotation occurs at a late stage of the coupling as a rate-limiting step of the process. GDP, instead, was slightly destabilized by G_s binding to β_2 AR (**Figure 4e**), although a complete dissociation requires the opening of the AHD, the first step for GDP release, which requires timescales longer than our simulations[52].

Usually, ICL3 of the GPCR and the G protein loop hgh4 are masked out from deposited cryo-EM structures due to their high flexibility and therefore low resolution. During our simulations, these two loops formed polar intermolecular interactions through R239^{ICL3}, R260^{ICL3}, K235^{ICL3},

and E322$^{hgh4.12}$, D323$^{hgh4.13}$. Further transient interactions not visible in the experiential structures, involved a mix of conserved and unique residues forming hydrogen bonds (**Table S2**): R63^{ICL1}-E392$^{H5.24}$, K232$^{5.71}$-D378$^{H5.10}$, K235$^{5.74}$- D378$^{H5.10}$, K235^{ICL3}-D343$^{H4.13}$, K267$^{6.29}$-L394^{c}, R239^{ICL3}-E314$^{hgh4.04}$, and S137$^{3.56}$-D381$^{H5.13}$. None of the interactions reported in **Table S2** is evident from the experimental β_2 AR:G_s complex, implying that mwSuMD can deliver useful working hypotheses for mutagenesis and spectroscopic experiments from out-of-equilibrium simulations. Results also suggest that the G_s binding is driven by a combination of conserved and unique transitory interactions with β_2 AR, possibly contributing to G protein selectivity. The conserved interactions would be necessary for the binding regardless of the receptor:G protein couple involved, while the transitory interactions should produce an effective engagement of the G protein.

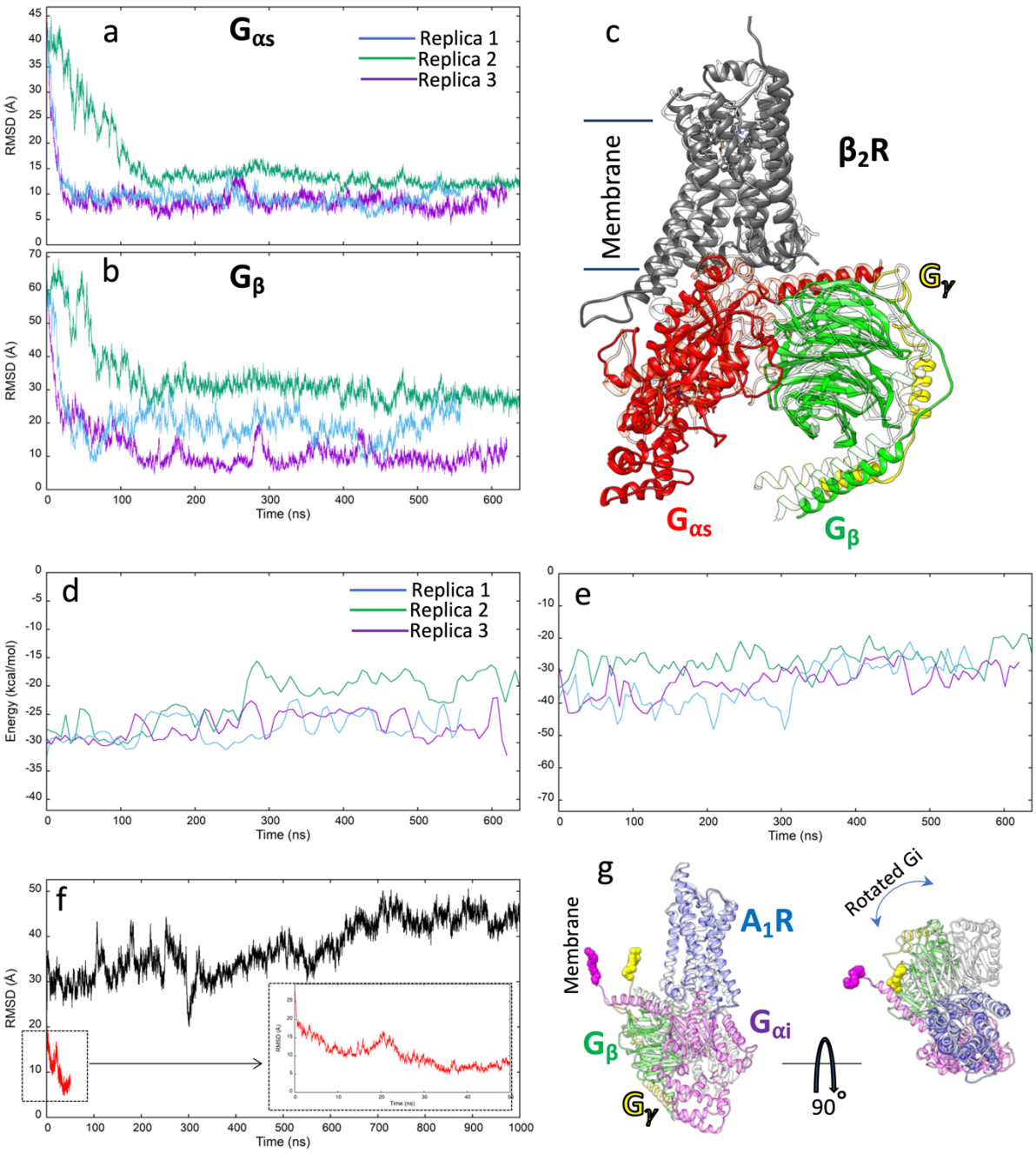

Figure 4. G protein binding simulations to β_2AR and A_1R. a) RMSD of $G_s\alpha$ to the experimental complex (PDB 3NS6) during three mwSuMD replicas; **b)** RMSD of $G_s\beta$ to the experimental complex (PDB 3NS6) during three mwSuMD replicas; **c)** superposition of the experimental G_s: **β_2** AR complex (transparent ribbon) and the MD frame with the lowest $G_s\alpha$ RMSD **(**3.94 Å**); d**) adrenaline MM-GBSA binding energy during three mwSuMD replicas; **e**) GDP MM-GBSA binding energy during three mwSuMD replicas; **f**) RMSD of $G_i\alpha$ (residues 243-355) to the experimental complex (PDB 6D9H) during a mwSuMD simulation (red, magnified in the box) and a 1000-ns long classic MD simulation (black); **g**) two-view superposition of the experimental G_i:**A_1** R complex (transparent ribbon) and the MD frame with the lowest $G_i\alpha$ RMSD **(**4.82 Å**).**

A possible pitfall of the above-reported G_s:β_2 AR mwSuMD binding simulation is that G proteins bear potential palmitoylation and myristoylation sites that can anchor the inactive trimer to the plasma membrane[53,54], de facto restraining possible binding paths to the receptor. To address this point and test the possible system dependency of mwSuMD, we prepared a different class A GPCR, the adenosine A1 receptor (A_1R), and its principal effector, the inhibitory G protein (G_i) considering $G_{i\alpha}$ residue C3 and G_γ residue C65 as palmitoylated and geranylgeranylated respectively and hence inserted in the membrane. Both classic (unsupervised) and mwSuMD simulations were performed on this system (**Video S6, Figure 4f**). In about 50 ns of mwSuMD, the $G_{i\alpha}$ subunit engaged its intracellular binding site on A_1R and formed a complex in close agreement with the cryo-EM structure (PDB 6D9H, RMSD ≈ 5 Å). The membrane anchoring affected the overall G_i binding and the final complex, which was rotated compared to the experimental structure due to the lipidation of $G_{i\alpha}$ and G_γ (**Figure 4g**). This suggests that future, more comprehensive studies of G protein binding and activation should consider several G protein orientations around the receptor as the starting points for mwSuMD simulations, to evaluate as many binding paths as possible. For comparison, 1 μs of cMD did not produce a productive engagement as the $G_{i\alpha}$ remained at RMSD values > 40 Å, suggesting the effectiveness of mwSuMD in sampling G protein binding rare events without the input of energy. Recently, the G_i binding to A_1R was simulated by combining the biased methods aMD with SuMD[55] but without taking into account the role played by membrane-anchoring post-translational modifications on the G_i binding pathway.

The heterodimerization between A_{2A} and D_2R, and binding simulations of the heterobivalent ligand compound 26.

The current structural model of the A_{2A}R:D_2R heterodimer is that TM4 and TM5 from both the two receptors contribute to form the primary interface of the dimer, although the involvement of TM7 is not ruled out[56]. Following this interaction model, we first dynamically docked A_{2A}R and D_2R in an explicit 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) membrane model, then simulated the binding of the heterobivalent compound 26[37] (CP26) to the preformed A_{2A}R:D_2R heterodimer (**Video S7**). Since membrane proteins are characterized by slow lateral diffusion[57], we favored the encounter between A_{2A}R and D_2R by input energy as metadynamics and adiabatic MD, during mwSuMD (hybrid metadynamics/aMD/mwSuMD), followed by 1.5 μs of classic MD (cMD) to relax the system and check the stability of the A_{2A}R:D_2R interactions.

During the first 200 ns of simulation with energy bias (**Figure 5a,c** and **Figure S7a**), $A_{2A}R$ and D_2R rapidly moved close to each other and reached a distance of about 30 Å (computed between centroids), before stabilizing at around 40 Å (**Figure 5a**). The computed molecular mechanics combined with the Poisson–Boltzmann and surface area continuum solvation (MM-PBSA) binding energy suggested two energy minima (**Figure 5c**) at these distances. The successive cMD simulation did not produce remarkable changes in the distance between receptors (**Figure 4b**), although the energy fluctuated before reaching about -10 kcal/mol, at the end of the simulation (**Figure 5d**). The sharp energy minima after 25 and 150 ns were due to the high number of direct contacts between $A_{2A}R$ and D_2R (**Figure S7**), favored by the energy added to the system. When the input of energy bias was stopped (**Figure 5b,d**) the POPC residues re-equilibrated at the interface between proteins and mediated intracellular polar interactions between $R150^{4.40\ D2R}$, $Y146^{4.36\ D2}$ and $R199^{5.60\ A2A}$, $Y103^{3.51\ A2A}$ as well as extracellular polar interactions between the top of $TM4^{D2}$, $TM5^{D2}$ and $TM5^{A2A}$, $TM6^{A2A}$ (**Figure 5f**), suggesting that the $A_{2A}R$:D_2R heterodimerization relies on lipids to mediate short-range interactions between receptors.

The dynamic docking of the herobivalent ligand C26 further stabilized the $A_{2A}R$:D_2R dimer (**Figure 5e),** in line with experimental data[37]. C26 reached the bound state rapidly inserting the agonist pharmacophore within the D_2R orthosteric site (**Figure S8, Video S7**), while the pyrazole-triazole-pyrimidine scaffold remained in metastable complex with $A_{2A}R$, before completely binding the orthosteric site at the end of the simulation (**Figure S9, Video S7**). In the final state, the long linker between pharmacophores extended over the top of the interface formed by $A_{2A}R$ and D_2R at the level of the receptors' ECL2 (**Figure 5g**). A network of polar interactions between POPC, $Y179^{A2A}$, and $Y192^{D2}$ contributed to stabilizing this ternary complex. Interestingly, the latter residues were pinpointed as important for $A_{2A}R$:D_2R interactions[56]. From a binding energy perspective, C26 reached the most stable configurations between 80 and 100 ns (**Figure S10**), before the pyrazole-triazole-pyrimidine component of the ligand completed the binding to $A_{2A}R$. This suggests some contribution of the linker to the overall stability of the ternary complex with $A_{2A}R$ and D_2R. Two out of four mwSuMD replicas produced $A_{2A}R$:D_2R:C26 ternary complexes with C26 engaged both by the orthosteric site of $A_{2A}R$ and D_2R, while in the remaining two replicas the A_2AR pharmacophore remained stacked on the extracellular vestibule of the receptor, although in the proximity of the binding site (**Figure S8**).

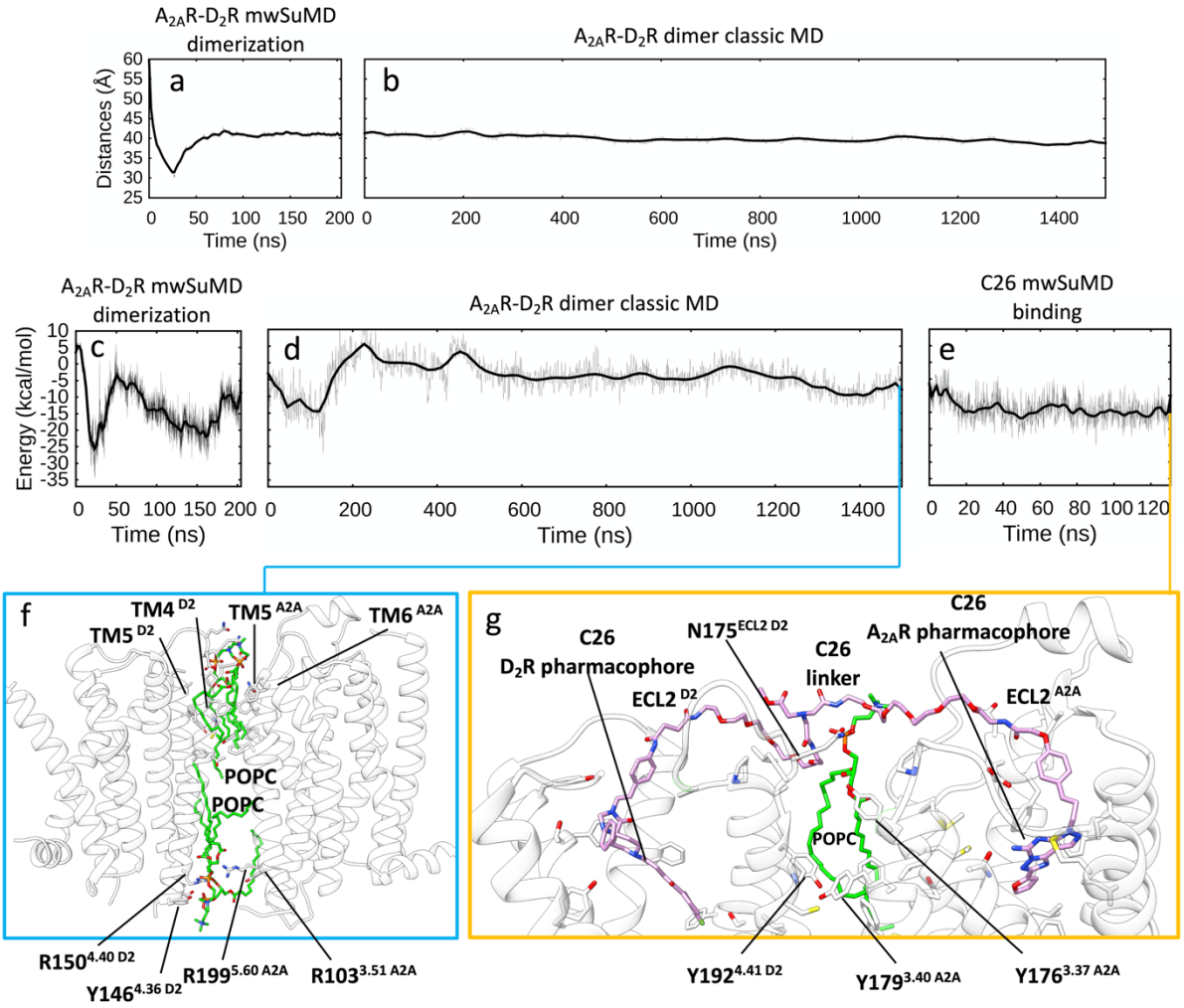

Figure 5. $A_{2A}R$:D_2R heterodimerization and formation of the ternary complex with C26. a) Distance between the centroids of $A_{2A}R$ and D_2R during the hybrid metadynamics/aMD/mwSuMD simulation; **b**) distance between the centroids of A2AR and D2R during the successive cMD simulation; **c**) MM-PBSA binding energy between A2AR and D2R during the hybrid metadynamics/aMD/mwSuMD simulation; **d**) MM-PBSA binding energy between A2AR and D2R during the successive cMD simulation; **e**) MM-PBSA binding energy between A2AR and D2R during the mwSUMD binding of C26. **f**) $A_{2A}R$:D_2R heterodimer (white ribbon) after 1.5 μs of cMD; POPC residues (green stick) were involved in polar and hydrophobic interactions; **g**) extracellular view of the $A_{2A}R$:D_2R:C26 ternary complex (D_2R TM2 and TM3 removed for clarity).

Conclusion

Classic MD simulations sample the phase space with an efficiency that depends on the energy barrier between neighboring minima. Processes like (un)binding and protein activation require the system to overcome numerous energy barriers, some of which create a bottleneck that slows the transition down to the millisecond, or second, time scale. To overcome some of these limits, we have developed an energetically-unbiased adaptive sampling algorithm,

namely multiple walker mwSuMD, which is based on traditional SuMD, while drawing on parallel multiple replica methods[58,59], and tested it on complex structural events characterizing GPCRs.

MwSuMD performed similarly to SuMD for the dynamic docking of AVP to V2R when time windows of 600 ps were employed. Time windows of 100 ps remarkably improved mwSuMD. Usually, dynamic docking is performed to predict the geometry of complexes or sample the binding path of an already known intermolecular complex, or both. The RMSD of AVP to the experimental coordinates as the supervised metric produced the best results. Consequently, the RMSD should be the metric of choice to study the binding path of well-known intermolecular complexes. The distance, on the other hand, is necessary when limited structural information about the binding mode is available. In the absence of structural information regarding the final bound state, it is possible to sample numerous binding events employing mwSuMD and evaluate the final bound states rank by applying end-point free energy binding methods like the molecular mechanics energies combined with the Poisson–Boltzmann or generalized Born and surface area continuum solvation (MM/PBSA and MM/GBSA) models. Our simulations suggested a remarkable predictivity of distance-driven mwSuMD, as demonstrated by the lowest deviation from the experimental AVP:V2R complex. Remarkably, the dissociation of AVP from V2R was simulated much more rapidly by mwSuMD than by SuMD, suggesting it is an efficient tool for studying the dissociation of ligands from GPCRs.

We increased the complexity of binding simulations by considering GLP-1R and the non-peptide agonist PF06882961. Using mwSuMD, we obtained a binding of the ligand in good agreement with the cryo-EM structure, followed by an active-like conformational transition of GLP-1R. The choice of the metrics supervised was driven by structural data available[45] and extensive preparatory MD simulations, however, alternative binding routes are possible from either the bulk solvent or the membrane[40,60,61]. Future studies on GLP-1R and other class B1 GPCR should consider different starting points for the ligand and alternative apo receptor conformations to improve the sampling.

MwSuMD was further tested on the G_s and G_i binding to β_2 AR and A_1R, respectively. MwSuMD produced G protein:GPCR complexes in remarkable agreement with experimental structural data without the input of energy in a few hundred nanoseconds when starting from inactive G_s and the intermediate active β_2 AR, or a few tens of nanoseconds when considering the active-state A_1R and G_i was anchored to the plasma membrane through the palmitoylation and the geranylgeranylation of $G\alpha\gamma$[53,54,62].

The final case study was the dimerization process between A_{2A}R and D_2R in a membrane model. To speed up the encounter between receptors, we introduced an energy bias in the form of abMD and MetaD. Although mwSuMD is an unbiased adaptive sampling method, it can be easily coupled to many forms of bias to favor the simulation of energy-requiring processes. Our results suggest a fundamental contribution of the phospholipids on the stabilization of the heterodimer, in agreement with experiments[63,64] and in disagreement with X-ray or protein-protein molecular docking results frequently predicting extended interfaces between monomers[65]. MwSuMD was able to dynamically dock the heterobivalent ligand CP26, supporting a stabilizing effect on the A_{2A}R:D_2R heterodimer. A complete characterization of the possible interfaces between GPCR monomers, which falls beyond the goal of the present work, should be achieved by preparing different initial unbound states

characterized by divergent relative orientations between monomers to dynamically dock in an explicit membrane.

In summary, we showcased the extended applicability domain of mwSuMD to key aspects of GPCRs structural biology. However, given the generality and simplicity of its implementation, we anticipate that mwSuMD can be employed to study a wide range of phenomena characterizing membrane and cytosolic proteins.

Methods

Force field, ligands parameters, and general systems preparation

The CHARMM36[66,67]/CGenFF 3.0.1[68–70] force field combination was employed in this work. Initial ligand force field, topology and parameter files were obtained from the ParamChem webserver[68]. Restrained electrostatic potential (RESP)[71] partial charges were assigned to all the non-peptidic small molecules but adrenaline and guanosine-5'-diphosphate (GDP) using Gaussian09 (HF/6-31G* level of theory) and AmberTools20.

Six systems were prepared for MD (**Table S1**). Hydrogen atoms were added using the pdb2pqr[72] and propka[73] software (considering a simulated pH of 7.0); the protonation of titratable side chains was checked by visual inspection. The resulting receptors were separately inserted in a 1-palmitoyl-2-oleyl-sn-glycerol-3-phosphocholine (POPC) bilayer (previously built by using the VMD Membrane Builder plugin 1.1, Membrane Plugin, Version 1.1. at: http://www.ks.uiuc.edu/Research/vmd/plugins/membrane/), through an insertion method[74]. Receptor orientation was obtained by superposing the coordinates on the corresponding structure retrieved from the OPM database[75]. Lipids overlapping the receptor transmembrane helical bundle were removed and TIP3P water molecules[76] were added to the simulation box by means of the VMD Solvate plugin 1.5 (Solvate Plugin, Version 1.5. at <http://www.ks.uiuc.edu/Research/vmd/plugins/solvate/). Finally, overall charge neutrality was reached by adding Na^+/Cl^- counter ions up to the final concentration of 0.150 M), using the VMD Autoionize plugin 1.3 (Autoionize Plugin, Version 1.3. at <http://www.ks.uiuc.edu/Research/vmd/plugins/autoionize/).

System equilibration and general MD settings

The MD engine ACEMD 3[77] was employed for both the equilibration and productive simulations. The equilibration was achieved in isothermal-isobaric conditions (NPT) using the Berendsen barostat[78] (target pressure 1 atm) and the Langevin thermostat[79] (target temperature 300 K) with low damping of 1 ps^{-1}. For the equilibration (integration time step of 2 fs): first, clashes between protein and lipid atoms were reduced through 1500 conjugate-gradient minimization steps, then a positional constraint of 1 kcal mol^{-1} $Å^{-2}$ on all heavy atoms was gradually released over different time windows: 2 ns for lipid phosphorus atoms, 60 ns for protein atoms other than alpha carbon atoms, 80 ns for alpha carbon atoms; a further 20 ns of equilibration was performed without any positional constraints.

Productive trajectories (**Table S1**) were computed with an integration time step of 4 fs in the canonical ensemble (NVT). The target temperature was set at 300 K, using a thermostat damping of 0.1 ps^{-1}; the M-SHAKE algorithm[80,81] was employed to constrain the bond lengths involving hydrogen atoms. The cut-off distance for electrostatic interactions was set at 9 Å, with a switching function applied beyond 7.5 Å. Long-range Coulomb interactions were handled using the particle mesh Ewald summation method (PME)[82] by setting the mesh spacing to 1.0 Å.

Vasopressin binding simulations

The vasopressin 2 receptor (V_2R) in complex with vasopressin (AVP) and the G_s protein[83] was retrieved from the Protein Data Bank[84] (PDB ID 7DW9). The G_s was removed from the system

and the missing residues on ECL2 (G185-G189) were modeled from scratch using Modeller 9.19[85]. AVP was placed away from V_2R in the extracellular bulk and the resulting system was prepared for MD simulations and equilibrated as reported above.

During SuMD simulations, the distance between the centroids of AVP residues C1-Q4 and V_2R residues Q96, Q174, Q291, and L312 (Cα atoms only) was supervised over time windows of 600 ps or 100 ps (**Table S1**). MwSuMD simulations considered the same distance, the RMSD of AVP residues C1-Q4 to the experimental bound complex or the combination of the two during time windows of 600 ps (3 walkers) or 100 ps (10 walkers) (**Table S1**). Slope, SMscore, or DMscore (see Methods section **MwSuMD protocol**) was used in the different mwSuMD replicas performed (**Table S1**). Simulations were stopped after 300 ns (time window duration = 600 ps) or 50 ns (time window duration = 100 ps) of total SuMD or mwSuMD simulation time.

Vasopressin unbinding simulations

The V_2R:AVP complex was prepared for MD simulations and equilibrated as reported above. During both SuMD and mwSuMD simulations (**Table S1**), the distance between the centroids of AVP residues C1-Q4 and V_2R residues Q96, Q174, Q291, and L312 (Cα atoms only) was supervised over time windows of 100 ps (10 walkers seeded for mwSuMD simulations). Replicas were stopped when the AVP-V_2R distance reached 40 Å.

GLP-1R:PF06882961 binding simulations

The inactive, ligand-free glucagon-like peptide receptor (GLP-1R) was retrieved from the Protein Data Bank[84] (PDB ID 6LN2)[86]. Missing residues in the stalk and ICL2 were modeled with Modeller 9.29. The PF06882961 initial conformation was extracted from the complex with the fully active GLP-1R[87] (PDB ID 7LCJ) and placed away from GLP-1R in the extracellular bulk. The resulting system was prepared for MD simulations and equilibrated as reported above. CGenFF dihedral force field parameters of PF06882961 with the highest penalties (dihedrals NG2R51-CG321-CG3C41-CG3C41 (penalty=143.5) and NG2R51-CG321-CG3C41-OG3C51 (penalty=152.4)) were optimized (**Figure S11**) employing Gaussian09 (geometric optimization and dihedral scan at HF/6-31g(d) level of theory) and the VMD force field toolkit plugin[88].

Four classic MD replicas, for a total of 8 μs, were performed on the inactive, ligand-free receptor (prepared for MD simulations and equilibrated as reported above) to assess the possible binding path to the receptor TMD and therefore decide the initial position of PF06882961 in the extracellular bulk of the simulation box. A visual inspection of the trajectories suggested three major conformational changes that could allow ligand access to the TMD (**Figure S12**). Transitory openings of the ECD (distance $Q47^{ECD}$ - $S310^{ECL2}$), TM6-TM7 (distance $H363^{6.52}$ - $F390^{7.45}$), and TM1-ECL1 (distance $E138^{1.33}$ and $W214^{ECL1}$) were observed. Since the opening of TM1-ECL1 was observed in two replicas out of four, we placed the ligand in a favorable position for crossing that region of GLP-1R.

MwSuMD simulations (**Table S1**) were performed stepwise to dock the ligand within GLP-1R first and then relax the receptor towards the active state. The PF06882961 binding was obtained by supervising at the same time the distance between the ligand and GLP-1R TM7 residues L379-F381, which are part of the orthosteric site (Cα atoms only), and the RMSD of the ECD (residues W33-W120, Cα atoms only) to the active state (PDB ID 7LCJ) until the former distance reached 4 Å. In the second phase of mwSuMD, the RMSD of the ECD (residues W33-W120, Cα atoms only) and the ECL1 to the active state (PDB ID 7LCJ)

Cα atoms of residues M204-L224) were supervised until the latter reached less than 4 Å. During the third phase, the RMSD of PF06882961, as well as the RMSD of ECL3 (residues A368-T378, Cα atoms), were supervised until the former reached values lower than 3 Å. In the last mwSuMD step, only the RMSD of TM6 (residues I345-F367, Cα atoms) to the active state (PDB ID 7LCJ) was supervised until less than 5 Å.

Membrane-anchored Gi protein:A_1R simulations

Since the full-length structure of the inactive human Gi protein has not been yet resolved by X-ray or cryo-EM, it was modeled by superimposing the AlphaFold2[89] models of the Gαi (P63096-F1), Gβ (Q9HAV0-F1), and Gγ (P50151-F1) subunits to the PDB file 6EG8 (a Gs heterotrimer). The resulting homotrimer (without GDP) was processed through Charmm-GUI[90] to palmitoylate residue $C3^{G\alpha i}$ and geranylgeranylate residue $C65^{G\gamma}$ [53,91]. The side chains of these two lipidated residues were manually inserted into a 120 x 120 Å POPC membrane and the resulting system was (previously built by using the VMD Membrane Builder plugin 1.1, Membrane Plugin, Version 1.1. at: http://www.ks.uiuc.edu/Research/vmd/plugins/membrane/). Lipids overlapping the palmitoyl and geranylgeranyl groups were removed and TIP3P water molecules[76] were added to the simulation box by means of the VMD Solvate plugin 1.5 (Solvate Plugin, Version 1.5. at <http://www.ks.uiuc.edu/Research/vmd/plugins/solvate/). Finally, overall charge neutrality was reached by adding Na^+/Cl^- counter ions up to the final concentration of 0.150 M), using the VMD Autoionize plugin 1.3 (Autoionize Plugin, Version 1.3. at <http://www.ks.uiuc.edu/Research/vmd/plugins/autoionize/). The first stage of equilibration was performed as reported above (Methods section **System equilibration and general MD settings**) for 120 ns, followed by a second stage in the NVT ensemble for a further 1 μs without any restraints to allow the membrane-anchored heterotrimeric Gi protein to stabilize within the intracellular side of the simulation box. After this two-stage, long equilibration, the active state A_1R in complex with adenosine (PDB 6D9H) was manually inserted into the equilibrated membrane above the Gi protein using the corresponding structure retrieved from the OPM database as a reference, and the system further equilibrated for 120 ns as reported above (Methods section **System equilibration and general MD settings**). The A_1R-Gi system was then subjected to both a 1 μs-long classic MD simulation and a mwSuMD simulation (**Table S1**). During the mwSuMD simulation, the RMSD of helix 5 (H5) G_αs residues 329-354 to the PDB 6D9H was supervised, seeding three walkers of 100 ps each until the productive simulation time reached 50 ns (total simulation time 150 ns).

A2A:D2R heterodimerization

The inactive state A_{2A}R and D_2R were retrieved from the Protein Data Bank[84] (PDB ID 5NM4 and 6LUQ, respectively)[92,93]. Antagonists bound to the orthosteric site were removed and no modeling of the missing IC loops was attempted. A_{2A}R and D_2R were manually placed roughly 40 Å away from each other, on the plane of the membrane, orienting the two receptors to favor the dimerization through the interface formed by TM5 and TM6, as suggested by Borroto-Esquela D. O. *et al*.[56] The resulting system was prepared for MD simulations and equilibrated as reported above.

The heterodimerization between A_{2A}R and D_2R was simulated with mwSuMD, seeding batches of three walkers with a duration of 100 ps each (**Table S1).** During each walker, the distance between TM5 of A_{2A}R and D_2R was supervised. At the same time, the distance between the centroids of A_{2A}R and D_2R was used as a collective variable for adiabatic MD[94]

(abMD) and well-tempered metadynamics[95,96] (wtMetaD) performed with Plumed 2.6[97]. For abMD, a distance target of 30 Å and a force constant of 10000 kJ*mol^{-1}*Å^{-1}) was used, while mwMetaD was performed by seeding gaussian functions every 1 ps (sigma=1 Å; height=0.837 kJ/mol; T=310K) with a bias factor of 30. When the $A_{2A}R$ - D_2R distance reached values lower than 40 Å and the first contacts between proteins were formed, the abMD was stopped and wtMetaD continued with an harmonic energy wall at 30 Å to avoid artificial crushing between the receptors due to the added energy bias. When the distance between $A_{2A}R$ and D_2R was stable at about 30 Å, the collective variable biased by wtMetaD was set as the number of atomic contacts between $A_{2A}R$ and D_2R, until reaching 200 ns of simulation. Finally, to relax the system and challenge the stability of the heterodimer formed during the biased mwSuMD simulation, a 1.5 μs classic MD simulation was performed.

$A_{2A}R$-D_2R heterobitopic ligand binding simulations

The A_2AR-D_2R heterobivalent ligand compound 26[37] was parameterized as reported above and placed in the bulk solvent of the A_2AR:D_2R complex from the classic MD. Four mwSuMD replicas were collected supervising at the same time the distance between the A_{2A} antagonist pyrazole-triazole-pyrimidine scaffold and the centroid of $A_{2A}R$ residues F168, N253, and A277 (Cα atoms) as well as the distance between the D_2 antagonist 4-fluorobenzyl scaffold and the centroids of the Cα of D_2R residues C118, F198, and V115 (Cα atoms). Ten walkers of 100 ps were simulated for every mwSuMD batch of replicas.

Gs protein:β_2 AR binding simulations

The model of the adrenergic β_2 receptor (β_2 AR) in an intermediate active state was downloaded from GPCRdb (https://gpcrdb.org/). The full agonist adrenaline (ALE) was inserted in the orthosteric site by superposition with the PDB ID 4LDO (fully-active β_2 AR)[98]. The structure of the inactive, GDP bound G_s protein[99] was retrieved from the Protein Data Bank[84] (PDB ID 6EG8) and placed in the intracellular bulk. The resulting system (G_s > 50 Å away from (β_2 AR) was prepared for MD simulations and equilibrated as reported above. The PDB ID 3SN6 (fully-active β_2 AR in complex with G_s[50]) was used as the reference for RMSD computations. Three mwSuMD replicas (**Table S1**) were performed supervising at the same time the distance between the helix 5 (H5) G_αs residues R385-L395 and the β_2 AR residues V31-P330 as well as the RMSD of β_2 AR TM6 residues C265-I278 (Cα atoms only) to the fully active state, during 100 ps time windows (5 walkers).

Multiple walker SuMD (mwSuMD) protocol

The supervised MD (SuMD) is an adaptive sampling method[100] for speeding up the simulation of binding events between small molecules (or peptides[101,102]) and proteins[1,19] without the introduction of any energetic bias. Briefly, during the SuMD a series of short unbiased MD simulations are performed, and after each simulation, the distances between the centers of mass (or the geometrical centers) of the ligand and the predicted binding site (collected at regular time intervals) are fitted to a linear function. If the resulting slope is negative (showing progress towards the target) the next simulation step starts from the last set of coordinates and velocities produced, otherwise, the simulation is restarted by randomly assigning the atomic velocities.

In the implementation for AceMD, mwSuMD needs as input the initial coordinates of the system as a pdb file, the coordinates, and the atomic velocities of the system from the equilibration stage, the topology file of the system, and all the necessary force filed parameters. The user can decide to supervise one (X) or two metrics (X', X'') of the simulated system over short simulations seeded in batches, called walkers. In the former case, either the slope of the linear function interpolating the metric values or a score can be adopted to decide whether to continue the mwSuMD simulation. When the user decides to supervise two metrics, then a specific score is used. In the present work, distances between centroids, RMSDs, or the number of atomic contacts between two selections were supervised (**Table S1**). The choice of the metrics is system and problem dependent, as the RMSD should be most useful when the final state is known, while the distance is required when the target state is unknown; details on the scores are given below. The decision to restart or continue mwSuMD after any short simulation is postponed until all the walkers of a batch are collected. The best short simulation is selected and extended by seeding the same number of walkers, with the same duration as the step before.

For each walker, the score for the supervision of a single metric (SMscore) is computed as the square root of the product between the metric value in the last frame ($X_{last\ frame}$) and the average metric value over the short simulation ($\bar{X}$):

$$SMscore = \sqrt{X_{last\ frame} * \bar{X}} \quad (1)$$

If the metric is set to decrease (e.g. binding or dimerization) the walker with the lowest SMscore is continued, otherwise (e.g. unbinding or outwards opening of domains), it is the walker with the highest score to be extended. Using the SMscore rather than the slope should give more weight to the final state of each short simulation, as it is the starting point for the successive batch of simulations. Considering the average of the metric should favor short simulations consistently evolving in the desired direction along the metric.

If both X' and X'' are set to increase during the mwSuMD simulations, the score for the supervision of two metrics (DMscore) on each walker is computed as follows:

$$DMscore = \left(\left(\frac{X'_{last\ frame}}{\bar{X}'_{batch\ walkers}} - 1\right) + \left(\frac{X''_{last\ frame}}{\bar{X}''_{batch\ walkers}} - 1\right)\right) * 100 \quad (2)$$

Where $X'_{last\ frame}$ and $X''_{last\ frame}$ are the metrics values in the last frame, while $\bar{X}'_{batch\ walkers}$ and $\bar{X}''_{batch\ walkers}$ represent the average value of the two metrics over all the walkers in the batch. Subtracting the value 1 to the metric ratio ensures that if one of the two metrics from the last frame ($X'_{last\ frame}$ or $X''_{last\ frame}$) is equal to the average ($\bar{X}''_{batch\ walkers}$ or $\bar{X}''_{batch\ walkers}$) then that metric addend is null and DMscore depends only on the remaining metric. If any of the two metrics is set to decrease, then the corresponding component in Equation 2 is multiplied by -1 to maintain a positive score. Considering the average value of the two metrics over all the walkers rather than only over the considered walker should be more representative of the system evolution along the defined metric. In other words, the information about the metric is taken from all the walkers to better describe the evolution of the system.

The DMScore is designed to preserve some degree of independence between the two metrics supervised. Indeed, if the variation of one of them slows down and gets close to zero, the other metric is still able to drive the system's evolution. It should be noted that DMScore

works at its best if the two metrics have similar variations over time, as it is in the case of distance and RMSD (both of which are distance-based). Notably, when a walker is extended by seeding a new batch of short simulations and the remaining walkers are stopped, the atomic velocities are not reassigned. This allows the simulations to be as short as a few picoseconds if desired, without introducing artifacts due to the thermostat latency to reach the target temperature (usually up to 10-20 ps when a simulation is restarted reassigning the velocities of the atoms).

The current implementation of mwSuMD is for python3 and exploits MDAnalysis[103] and MDTRaj[104] modules.

MD Analysis

Interatomic distances were computed through MDAnalysis[103]; root mean square deviations (RMSD) were computed using VMD[105] and MDAnalysis[103].

Interatomic contacts and ligand-protein hydrogen bonds were detected using the GetContacts scripts tool (https://getcontacts.github.io), setting a hydrogen bond donor-acceptor distance of 3.3 Å and an angle value of 120° as geometrical cut-offs. Contacts and hydrogen bond persistency are quantified as the percentage of frames (over all the frames obtained by merging the different replicas) in which protein residues formed contacts or hydrogen bonds with the ligand.

The MMPBSA.py[106] script, from the AmberTools20 suite (The Amber Molecular Dynamics Package, at http://ambermd.org/), was used to compute molecular mechanics energies combined with the generalized Born and surface area continuum solvation (MM/GBSA) method or the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) approach, after transforming the CHARMM psf topology files to an Amber prmtop format using ParmEd (documentation at <http://parmed.github.io/ParmEd/html/index.html).

Supplementary Videos were produced employing VMD and avconv (at https://libav.org/avconv.html). Molecular graphics images were produced using the UCSF Chimera[107] (v1.14).

Numbering system

Throughout the manuscript, the Ballesteros-Weinstein residues numbering system for class A GPCRs[108] and the Wootten residues numbering system for class B GPCRs[109] are adopted.

ASSOCIATED CONTENT

Supporting Information (PDF)

Supporting Videos S1-S7 (mpg).

All the MD trajectories (stripped of POPC, water molecules, and ions) and topology files (psf and pdb) are available here:

https://zenodo.org/record/7351548

The mwSuMD software used in this study is available at:

https://github.com/pipitoludovico/mwSuMD

Author Contributions

GD and CAR supervised the project; GD conceived the software and planned the simulations; GD, LP, RMR, TW, and PG carried out the simulations; GD analyzed the data; GD, AC, SM, and CAR interpreted the results, GD wrote the manuscript with input from CAR, AC, and SM; all the authors edited and reviewed the final version of the manuscript.

Acknowledgments
GD is a member of the GPCRs-focused European COST action ERNEST. CAR is grateful for a Royal Society Industry Fellowship.

Competing Interest
All authors declare that they have no conflicts of interest.

Bibliography

(1) Cuzzolin, A.; Sturlese, M.; Deganutti, G.; Salmaso, V.; Sabbadin, D.; Ciancetta, A.; Moro, S. Deciphering the Complexity of Ligand-Protein Recognition Pathways Using Supervised Molecular Dynamics (SuMD) Simulations. *J. Chem. Inf. Model.* **2016**, *56*, 687–705.

(2) Deganutti, G.; Moro, S.; Reynolds, C. A. A Supervised Molecular Dynamics Approach to Unbiased Ligand-Protein Unbinding. *J. Chem. Inf. Model.* **2020**, *60*, 1804–1817.

(3) de Mendoza, A.; Sebé-Pedrós, A.; Ruiz-Trillo, I. The Evolution of the GPCR Signaling System in Eukaryotes: Modularity, Conservation, and the Transition to Metazoan Multicellularity. *Genome Biol. Evol.* **2014**, *6*, 606–619.

(4) Sriram, K.; Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **2018**, *93*, 251–258.

(5) Schiöth, H. B.; Fredriksson, R. The GRAFS Classification System of G-Protein Coupled Receptors in Comparative Perspective. *Gen. Comp. Endocrinol.* **2005**, *142*, 94–101.

(6) G protein-coupled receptors | G protein-coupled receptors | IUPHAR/BPS Guide to PHARMACOLOGY https://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=694 (accessed May 18, 2022).

(7) Syrovatkina, V.; Alegre, K. O.; Dey, R.; Huang, X.-Y. Regulation, Signaling, and Physiological Functions of G-Proteins. *J. Mol. Biol.* **2016**, *428*, 3850–3868.

(8) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science* **2000**, *289*, 739–745.

(9) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. High-Resolution Crystal Structure of an Engineered Human Beta2-Adrenergic G Protein-Coupled Receptor. *Science* **2007**, *318*, 1258–1265.

(10) Tan, L.; Yan, W.; McCorvy, J. D.; Cheng, J. Biased Ligands of G Protein-Coupled Receptors (GPCRs): Structure-Functional Selectivity Relationships (SFSRs) and Therapeutic Potential. *J. Med. Chem.* **2018**, *61*, 9841–9878.

(11) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129–1143.

(12) Durrant, J. D.; McCammon, J. A. Molecular Dynamics Simulations and Drug Discovery. *BMC Biol.* **2011**, *9*, 71.

(13) Bussi, G.; Laio, A. Using Metadynamics to Explore Complex Free-Energy Landscapes. *Nat. Rev. Phys.* **2020**.

(14) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919–11929.

(15) 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.

(16) Mattedi, G.; Acosta-Gutiérrez, S.; Clark, T.; Gervasio, F. L. A Combined Activation Mechanism for the Glucagon Receptor. *Proc. Natl. Acad. Sci. USA* **2020**.

(17) Miao, Y.; McCammon, J. A. Mechanism of the G-Protein Mimetic Nanobody

Binding to a Muscarinic G-Protein-Coupled Receptor. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3036–3041.
(18) Zuckerman, D. M.; Chong, L. T. Weighted Ensemble Simulation: Review of Methodology, Applications, and Software. *Annu. Rev. Biophys.* **2017**, *46*, 43–57.
(19) Sabbadin, D.; Moro, S. Supervised Molecular Dynamics (SuMD) as a Helpful Tool to Depict GPCR-Ligand Recognition Pathway in a Nanosecond Time Scale. *J. Chem. Inf. Model.* **2014**, *54*, 372–376.
(20) Deganutti, G.; Atanasio, S.; Rujan, R.-M.; Sexton, P. M.; Wootten, D.; Reynolds, C. A. Exploring Ligand Binding to Calcitonin Gene-Related Peptide Receptors. *Front. Mol. Biosci.* **2021**, *8*, 720561.
(21) Dong, M.; Deganutti, G.; Piper, S. J.; Liang, Y.-L.; Khoshouei, M.; Belousoff, M. J.; Harikumar, K. G.; Reynolds, C. A.; Glukhova, A.; Furness, S. G. B.; Christopoulos, A.; Danev, R.; Wootten, D.; Sexton, P. M.; Miller, L. J. Structure and Dynamics of the Active Gs-Coupled Human Secretin Receptor. *Nat. Commun.* **2020**, *11*, 4137.
(22) Deganutti, G.; Prischi, F.; Reynolds, C. A. Supervised Molecular Dynamics for Exploring the Druggability of the SARS-CoV-2 Spike Protein. *J. Comput. Aided Mol. Des.* **2021**, *35*, 195–207.
(23) Cary, B. P.; Deganutti, G.; Zhao, P.; Truong, T. T.; Piper, S. J.; Liu, X.; Belousoff, M. J.; Danev, R.; Sexton, P. M.; Wootten, D.; Gellman, S. H. Structural and Functional Diversity among Agonist-Bound States of the GLP-1 Receptor. *Nat. Chem. Biol.* **2021**.
(24) Birnbaumer, M. Vasopressin Receptors. *Trends Endocrinol. Metab.* **2000**, *11*, 406–410.
(25) Ball, S. G. Vasopressin and Disorders of Water Balance: The Physiology and Pathophysiology of Vasopressin. *Ann Clin Biochem* **2007**, *44*, 417–431.
(26) Guo, D.; Heitman, L. H.; IJzerman, A. P. Kinetic Aspects of the Interaction between Ligand and G Protein-Coupled Receptor: The Case of the Adenosine Receptors. *Chem. Rev.* **2017**, *117*, 38–66.
(27) Guo, D.; Heitman, L. H.; IJzerman, A. P. The Role of Target Binding Kinetics in Drug Discovery. *ChemMedChem* **2015**, *10*, 1793–1796.
(28) 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*, 317–322.
(29) Wall, M. J.; Hill, E.; Huckstepp, R.; Barkan, K.; Deganutti, G.; Leuenberger, M.; Preti, B.; Winfield, I.; Carvalho, S.; Suchankova, A.; Wei, H.; Safitri, D.; Huang, X.; Imlach, W.; La Mache, C.; Dean, E.; Hume, C.; Hayward, S.; Oliver, J.; Zhao, F.-Y.; Spanswick, D.; Reynolds, C. A.; Lochner, M.; Ladds, G.; Frenguelli, B. G. Selective Activation of Gαob by an Adenosine A1 Receptor Agonist Elicits Analgesia without Cardiorespiratory Depression. *Nat. Commun.* **2022**, *13*, 4150.
(30) Culhane, K. J.; Gupte, T. M.; Madhugiri, I.; Gadgil, C. J.; Sivaramakrishnan, S. Kinetic Model of GPCR-G Protein Interactions Reveals Allokairic Modulation of Signaling. *Nat. Commun.* **2022**, *13*, 1202.
(31) Canals, M.; Marcellino, D.; Fanelli, F.; Ciruela, F.; de Benedetti, P.; Goldberg, S. R.; Neve, K.; Fuxe, K.; Agnati, L. F.; Woods, A. S.; Ferré, S.; Lluis, C.; Bouvier, M.; Franco, R. Adenosine A2A-Dopamine D2 Receptor-Receptor Heteromerization: Qualitative and Quantitative Assessment by Fluorescence and Bioluminescence Energy Transfer. *J. Biol. Chem.* **2003**, *278*, 46741–46749.
(32) Fuxe, K.; Guidolin, D.; Agnati, L. F.; Borroto-Escuela, D. O. Dopamine Heteroreceptor Complexes as Therapeutic Targets in Parkinson's Disease. *Expert Opin. Ther. Targets* **2015**, *19*, 377–398.
(33) Rivera-Oliver, M.; Díaz-Ríos, M. Using Caffeine and Other Adenosine Receptor

Antagonists and Agonists as Therapeutic Tools against Neurodegenerative Diseases: A Review. *Life Sci.* **2014**, *101*, 1–9.
(34) Richardson, P. J.; Kase, H.; Jenner, P. G. Adenosine A2A Receptor Antagonists as New Agents for the Treatment of Parkinson's Disease. *Trends Pharmacol. Sci.* **1997**, *18*, 338–344.
(35) Fuxe, K.; Borroto-Escuela, D. O. Heteroreceptor Complexes and Their Allosteric Receptor-Receptor Interactions as a Novel Biological Principle for Integration of Communication in the CNS: Targets for Drug Development. *Neuropsychopharmacology* **2016**, *41*, 380–382.
(36) Soriano, A.; Ventura, R.; Molero, A.; Hoen, R.; Casadó, V.; Cortés, A.; Fanelli, F.; Albericio, F.; Lluís, C.; Franco, R.; Royo, M. Adenosine A2A Receptor-Antagonist/Dopamine D2 Receptor-Agonist Bivalent Ligands as Pharmacological Tools to Detect A2A-D2 Receptor Heteromers. *J. Med. Chem.* **2009**, *52*, 5590–5602.
(37) Pulido, D.; Casadó-Anguera, V.; Gómez-Autet, M.; Llopart, N.; Moreno, E.; Casajuana-Martin, N.; Ferré, S.; Pardo, L.; Casadó, V.; Royo, M. Heterobivalent Ligand for the Adenosine A2A-Dopamine D2 Receptor Heteromer. *J. Med. Chem.* **2022**.
(38) Gioia, D.; Bertazzo, M.; Recanatini, M.; Masetti, M.; Cavalli, A. Dynamic Docking: A Paradigm Shift in Computational Drug Discovery. *Molecules* **2017**, *22*.
(39) Atanasio, S.; Deganutti, G.; Reynolds, C. A. Addressing Free Fatty Acid Receptor 1 (FFAR1) Activation Using Supervised Molecular Dynamics. *J. Comput. Aided Mol. Des.* **2020**, *34*, 1181–1193.
(40) Deganutti, G.; Barkan, K.; Preti, B.; Leuenberger, M.; Wall, M.; Frenguelli, B. G.; Lochner, M.; Ladds, G.; Reynolds, C. A. Deciphering the Agonist Binding Mechanism to the Adenosine A1 Receptor. *ACS Pharmacol. Transl. Sci.* **2021**, *4*, 314–326.
(41) Deganutti, G.; Barkan, K.; Ladds, G.; Reynolds, C. A. A Multisite Model of Allosterism for the Adenosine A1 Receptor. *BioRxiv* **2020**.
(42) Zhao, P.; Liang, Y.-L.; Belousoff, M. J.; Deganutti, G.; Fletcher, M. M.; Willard, F. S.; Bell, M. G.; Christe, M. E.; Sloop, K. W.; Inoue, A.; Truong, T. T.; Clydesdale, L.; Furness, S. G. B.; Christopoulos, A.; Wang, M.-W.; Miller, L. J.; Reynolds, C. A.; Danev, R.; Sexton, P. M.; Wootten, D. Activation of the GLP-1 Receptor by a Non-Peptidic Agonist. *Nature* **2020**, *577*, 432–436.
(43) Kawai, T.; Sun, B.; Yoshino, H.; Feng, D.; Suzuki, Y.; Fukazawa, M.; Nagao, S.; Wainscott, D. B.; Showalter, A. D.; Droz, B. A.; Kobilka, T. S.; Coghlan, M. P.; Willard, F. S.; Kawabe, Y.; Kobilka, B. K.; Sloop, K. W. Structural Basis for GLP-1 Receptor Activation by LY3502970, an Orally Active Nonpeptide Agonist. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 29959–29967.
(44) Ma, H.; Huang, W.; Wang, X.; Zhao, L.; Jiang, Y.; Liu, F.; Guo, W.; Sun, X.; Zhong, W.; Yuan, D.; Xu, H. E. Structural Insights into the Activation of GLP-1R by a Small Molecule Agonist. *Cell Res.* **2020**, *30*, 1140–1142.
(45) Zhang, X.; Belousoff, M. J.; Zhao, P.; Kooistra, A. J.; Truong, T. T.; Ang, S. Y.; Underwood, C. R.; Egebjerg, T.; Šenel, P.; Stewart, G. D.; Liang, Y.-L.; Glukhova, A.; Venugopal, H.; Christopoulos, A.; Furness, S. G. B.; Miller, L. J.; Reedtz-Runge, S.; Langmead, C. J.; Gloriam, D. E.; Danev, R.; Sexton, P. M.; Wootten, D. Differential GLP-1R Binding and Activation by Peptide and Non-Peptide Agonists. *Mol. Cell* **2020**, *80*, 485–500.e7.
(46) Cong, Z.; Chen, L.-N.; Ma, H.; Zhou, Q.; Zou, X.; Ye, C.; Dai, A.; Liu, Q.; Huang, W.; Sun, X.; Wang, X.; Xu, P.; Zhao, L.; Xia, T.; Zhong, W.; Yang, D.; Eric Xu, H.; Zhang, Y.; Wang, M.-W. Molecular Insights into Ago-Allosteric Modulation of the

Human Glucagon-like Peptide-1 Receptor. *Nat. Commun.* **2021**, *12*, 3763.
(47) Cong, Z.; Zhou, Q.; Li, Y.; Chen, L.-N.; Zhang, Z.-C.; Liang, A.; Liu, Q.; Wu, X.; Dai, A.; Xia, T.; Wu, W.; Zhang, Y.; Yang, D.; Wang, M.-W. Structural Basis of Peptidomimetic Agonism Revealed by Small Molecule GLP-1R Agonists Boc5 and WB4-24. *BioRxiv* **2022**.
(48) Cong, Z.; Liang, Y.-L.; Zhou, Q.; Darbalaei, S.; Zhao, F.; Feng, W.; Zhao, L.; Xu, H. E.; Yang, D.; Wang, M.-W. Structural Perspective of Class B1 GPCR Signaling. *Trends Pharmacol. Sci.* **2022**, *43*, 321–334.
(49) Hilger, D.; Kumar, K. K.; Hu, H.; Pedersen, M. F.; O'Brien, E. S.; Giehm, L.; Jennings, C.; Eskici, G.; Inoue, A.; Lerch, M.; Mathiesen, J. M.; Skiniotis, G.; Kobilka, B. K. Structural Insights into Differences in G Protein Activation by Family A and Family B GPCRs. *Science* **2020**, *369*.
(50) Rasmussen, S. G. F.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal Structure of the B2 Adrenergic Receptor-Gs Protein Complex. *Nature* **2011**, *477*, 549–555.
(51) Liang, Y.-L.; Belousoff, M. J.; Fletcher, M. M.; Zhang, X.; Khoshouei, M.; Deganutti, G.; Koole, C.; Furness, S. G. B.; Miller, L. J.; Hay, D. L.; Christopoulos, A.; Reynolds, C. A.; Danev, R.; Wootten, D.; Sexton, P. M. Structure and Dynamics of Adrenomedullin Receptors AM1 and AM2 Reveal Key Mechanisms in the Control of Receptor Phenotype by Receptor Activity-Modifying Proteins. *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 263–284.
(52) 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.; Hubbell, W. L.; Kobilka, B. K.; Sunahara, R. K.; Shaw, D. E. SIGNAL TRANSDUCTION. Structural Basis for Nucleotide Exchange in Heterotrimeric G Proteins. *Science* **2015**, *348*, 1361–1365.
(53) Linder, M. E.; Middleton, P.; Hepler, J. R.; Taussig, R.; Gilman, A. G.; Mumby, S. M. Lipid Modifications of G Proteins: Alpha Subunits Are Palmitoylated. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3675–3679.
(54) Zhang, Z.; Melia, T. J.; He, F.; Yuan, C.; McGough, A.; Schmid, M. F.; Wensel, T. G. How a G Protein Binds a Membrane. *J. Biol. Chem.* **2004**, *279*, 33937–33945.
(55) Li, Y.; Sun, J.; Li, D.; Lin, J. The Full Activation Mechanism of the Adenosine A1 Receptor Revealed by GaMD and Su-GaMD Simulations. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2203702119.
(56) Borroto-Escuela, D. O.; Rodriguez, D.; Romero-Fernandez, W.; Kapla, J.; Jaiteh, M.; Ranganathan, A.; Lazarova, T.; Fuxe, K.; Carlsson, J. Mapping the Interface of a GPCR Dimer: A Structural Model of the A2A Adenosine and D2 Dopamine Receptor Heteromer. *Front. Pharmacol.* **2018**, *9*, 829.
(57) Ramadurai, S.; Holt, A.; Krasnikov, V.; van den Bogaart, G.; Killian, J. A.; Poolman, B. Lateral Diffusion of Membrane Proteins. *J. Am. Chem. Soc.* **2009**, *131*, 12650–12656.
(58) Dickson, A.; Brooks, C. L. WExplore: Hierarchical Exploration of High-Dimensional Spaces Using the Weighted Ensemble Algorithm. *J. Phys. Chem. B* **2014**, *118*, 3532–3542.
(59) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
(60) Stanley, N.; Pardo, L.; Fabritiis, G. D. The Pathway of Ligand Entry from the Membrane Bilayer to a Lipid G Protein-Coupled Receptor. *Sci. Rep.* **2016**, *6*, 22639.
(61) Bokoch, M. P.; Jo, H.; Valcourt, J. R.; Srinivasan, Y.; Pan, A. C.; Capponi, S.;

Grabe, M.; Dror, R. O.; Shaw, D. E.; DeGrado, W. F.; Coughlin, S. R. Entry from the Lipid Bilayer: A Possible Pathway for Inhibition of a Peptide G Protein-Coupled Receptor by a Lipophilic Small Molecule. *Biochemistry* **2018**.

(62) Vögler, O.; Barceló, J. M.; Ribas, C.; Escribá, P. V. Membrane Interactions of G Proteins and Other Related Proteins. *Biochim. Biophys. Acta* **2008**, *1778*, 1640–1652.

(63) Goddard, A. D.; Dijkman, P. M.; Adamson, R. J.; Watts, A. Lipid-Dependent GPCR Dimerization. *Methods Cell Biol.* **2013**, *117*, 341–357.

(64) Prasanna, X.; Sengupta, D.; Chattopadhyay, A. Cholesterol-Dependent Conformational Plasticity in GPCR Dimers. *Sci. Rep.* **2016**, *6*, 31858.

(65) Meng, X.-Y.; Mezei, M.; Cui, M. Computational Approaches for Modeling GPCR Dimerization. *Curr. Pharm. Biotechnol.* **2014**, *15*, 996–1006.

(66) Huang, J.; MacKerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.

(67) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71–73.

(68) Vanommeslaeghe, K.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.

(69) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. 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.

(70) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell, A. D. Extension of the CHARMM General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.

(71) Woods, R. J.; Chappelle, R. Restrained Electrostatic Potential Atomic Partial Charges for Condensed-Phase Simulations of Carbohydrates. *Theochem* **2000**, *527*, 149–156.

(72) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32*, W665-7.

(73) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PK Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.

(74) Sommer, B. Membrane Packing Problems: A Short Review on Computational Membrane Modeling Methods and Tools. *Comput Struct Biotechnol J* **2013**, *5*, e201302014.

(75) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: Orientations of Proteins in Membranes Database. *Bioinformatics* **2006**, *22*, 623–625.

(76) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926.

(77) Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* **2009**, *5*, 1632–1639.

(78) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684.

(79) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides: The

Frictional Dependence of Isomerization Rates of N-Acetylalanyl-N'-Methylamide. *Biopolymers* **1992**, *32*, 523–535.
(80) Forester, T. R.; Smith, W. SHAKE, Rattle, and Roll: Efficient Constraint Algorithms for Linked Rigid Bodies. *J. Comput. Chem.* **1998**.
(81) Kr□utler, V.; van Gunsteren, W. F.; H□nenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* **2001**, *22*, 501–508.
(82) 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*, 8577.
(83) Zhou, F.; Ye, C.; Ma, X.; Yin, W.; Croll, T. I.; Zhou, Q.; He, X.; Zhang, X.; Yang, D.; Wang, P.; Xu, H. E.; Wang, M.-W.; Jiang, Y. Molecular Basis of Ligand Recognition and Activation of Human V2 Vasopressin Receptor. *Cell Res.* **2021**, *31*, 929–931.
(84) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
(85) Fiser, A.; Sali, A. Modeller: Generation and Refinement of Homology-Based Protein Structure Models. *Meth. Enzymol.* **2003**, *374*, 461–491.
(86) Wu, F.; Yang, L.; Hang, K.; Laursen, M.; Wu, L.; Han, G. W.; Ren, Q.; Roed, N. K.; Lin, G.; Hanson, M. A.; Jiang, H.; Wang, M.-W.; Reedtz-Runge, S.; Song, G.; Stevens, R. C. Full-Length Human GLP-1 Receptor Structure without Orthosteric Ligands. *Nat. Commun.* **2020**, *11*, 1272.
(87) Zhang, X.; Johnson, R. M.; Drulyte, I.; Yu, L.; Kotecha, A.; Danev, R.; Wootten, D.; Sexton, P. M.; Belousoff, M. J. Evolving Cryo-EM Structural Approaches for GPCR Drug Discovery. *Structure* **2021**, *29*, 963–974.e6.
(88) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid Parameterization of Small Molecules Using the Force Field Toolkit. *J. Comput. Chem.* **2013**, *34*, 2757–2770.
(89) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589.
(90) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865.
(91) Mystek, P.; Rysiewicz, B.; Gregrowicz, J.; Dziedzicka-Wasylewska, M.; Polit, A. Gγ and Gα Identity Dictate a G-Protein Heterotrimer Plasma Membrane Targeting. *Cells* **2019**, *8*.
(92) Weinert, T.; Olieric, N.; Cheng, R.; Brünle, S.; James, D.; Ozerov, D.; Gashi, D.; Vera, L.; Marsh, M.; Jaeger, K.; Dworkowski, F.; Panepucci, E.; Basu, S.; Skopintsev, P.; Doré, A. S.; Geng, T.; Cooke, R. M.; Liang, M.; Prota, A. E.; Panneels, V.; Nogly, P.; Ermler, U.; Schertler, G.; Hennig, M.; Steinmetz, M. O.; Wang, M.; Standfuss, J. Serial Millisecond Crystallography for Routine Room-Temperature Structure Determination at Synchrotrons. *Nat. Commun.* **2017**, *8*, 542.
(93) Fan, L.; Tan, L.; Chen, Z.; Qi, J.; Nie, F.; Luo, Z.; Cheng, J.; Wang, S. Haloperidol Bound D2 Dopamine Receptor Structure Inspired the Discovery of Subtype Selective Ligands. *Nat. Commun.* **2020**, *11*, 1074.
(94) Marchi, M.; Ballone, P. Adiabatic Bias Molecular Dynamics: A Method to Navigate

the Conformational Space of Complex Molecular Systems. *J. Chem. Phys.* **1999**, *110*, 3697.
(95) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562–12566.
(96) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603.
(97) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput Phys Commun* **2014**, *185*, 604–613.
(98) Ring, A. M.; Manglik, A.; Kruse, A. C.; Enos, M. D.; Weis, W. I.; Garcia, K. C.; Kobilka, B. K. Adrenaline-Activated Structure of B2-Adrenoceptor Stabilized by an Engineered Nanobody. *Nature* **2013**, *502*, 575–579.
(99) Liu, X.; Xu, X.; Hilger, D.; Aschauer, P.; Tiemann, J. K. S.; Du, Y.; Liu, H.; Hirata, K.; Sun, X.; Guixà-González, R.; Mathiesen, J. M.; Hildebrand, P. W.; Kobilka, B. K. Structural Insights into the Process of GPCR-G Protein Complex Formation. *Cell* **2019**, *177*, 1243–1251.e12.
(100) Deganutti, G.; Moro, S. Estimation of Kinetic and Thermodynamic Ligand-Binding Parameters Using Computational Strategies. *Future Med. Chem.* **2017**, *9*, 507–523.
(101) Salmaso, V.; Sturlese, M.; Cuzzolin, A.; Moro, S. Exploring Protein-Peptide Recognition Pathways Using a Supervised Molecular Dynamics Approach. *Structure* **2017**, *25*, 655–662.e2.
(102) Bower, R. L.; Yule, L.; Rees, T. A.; Deganutti, G.; Hendrikse, E. R.; Harris, P. W. R.; Kowalczyk, R.; Ridgway, Z.; Wong, A. G.; Swierkula, K.; Raleigh, D. P.; Pioszak, A. A.; Brimble, M. A.; Reynolds, C. A.; Walker, C. S.; Hay, D. L. Molecular Signature for Receptor Engagement in the Metabolic Peptide Hormone Amylin. *ACS Pharmacol. Transl. Sci.* **2018**, *1*, 32–49.
(103) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
(104) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. Mdtraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109*, 1528–1532.
(105) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J Mol Graph* **1996**, *14*, 33–38, 27.
(106) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.
(107) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
(108) Ballesteros, J. A.; Weinstein, H. [19] Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. In *Receptor Molecular Biology*; Methods in Neurosciences; Elsevier, 1995; Vol. 25, pp. 366–428.
(109) 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. USA* **2013**, *110*, 5211–5216.