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,[1,4]**

[1] Centre for 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 Fossato di Mortara 17/19, 44121 Ferrara, Italy

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

[4] School of Life Sciences, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ

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

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

Abstract

The structural basis for the pharmacology of G protein-coupled receptors (GPCRs), the most abundant membrane proteins and the target of about 35% of approved drugs, is still a matter of intense study. What makes GPCRs challenging to study is the inherent flexibility and the metastable nature of interaction with extra- and intracellular partners that drive their effects. Here, we present a molecular dynamics (MD) adaptive sampling algorithm, namely multiple walker supervised molecular dynamics (mwSuMD), to address complex structural transitions involving GPCRs without energy input. We first report the binding and unbinding of the vasopressin peptide from its receptor V_2. Successively, we present the complete transition of the glucagon-like peptide-1 receptor (GLP-1R) from inactive to active, agonist and G_s-bound state, and the GDP release from G_s. To our knowledge, this is the first time the whole sequence of events leading from an inactive GPCR to the GDP release is simulated without any energy bias. We demonstrate that mwSuMD can address complex binding processes intrinsically linked to protein dynamics out of reach of classic MD.

Introduction

Supervised molecular dynamics[1,2] (SuMD) is an efficient technique for studying ligand-receptor binding and unbinding pathways; here, we present the multiple walker enhancement (mwSuMD) to study a significantly wider range of structural transitions relevant to the drug design. We validated the method by applying it to G protein-coupled receptors (GPCRs), since

their inherent flexibility is essential to their function and because these 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]. X-ray and cryo-electron microscopy (cryo-EM) show that GPCRs possess seven transmembrane (TM) helices connected by three extracellular loops (ECLs) and three intracellular loops (ICLs), with an extended and structured N-terminal extracellular domain (ECD) 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 ($G_{s/olf}$) 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].

Many aspects of GPCR 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 others (*i.e.* functional selectivity or bias)[8]. GPCRs are challenging proteins to characterize experimentally due to their inherent flexibility and the transitory nature of the complexes formed with extracellular and intracellular effectors. Importantly, agonists can allosterically 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[9].

Molecular dynamics (MD) is a powerful computational methodology that predicts the movement and interactions of (bio)molecules in systems of variable complexity, at atomic detail. However, classic MD sampling is limited to the microsecond or, in the best conditions, the millisecond time scale[10,11]. 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, are metadynamics[12], accelerated MD (aMD)[13], and Gaussian-accelerated MD (GaMD)[14], which introduce an energy potential to overcome the energy barriers preventing the complete exploration of the free energy surface, thus *de facto* bias the simulation. Energetically unbiased MD protocols, on the other hand, comprise weighted ensemble MD (weMD)[15], swarms approach[16], AdaptiveGoal[17], and SuMD[1,18], which have largely been applied to (unbinding) small molecules, peptides, and small proteins[1,18–22]. Since SuMD is optimized for (un)bindings, we have designed mwSuMD to address more complex conformational transitions and protein-protein associations. GPCRs preferentially couple to very few G proteins out of 23 possible counterparts[9,23]. 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[24]. 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 have usually employed energetically biased simulations, have been confined to the $G\alpha$ subunit, or considered a pre-formed GPCR:G protein complex[23,25–27].

Firstly, we validated mwSuMD on the nonapeptide arginine vasopressin (AVP) by simulating binding (dynamic docking) and unbinding paths from the vasopressin 2 receptor (V_2R). 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[28] and structure-kinetics relationship (SKR) studies[29]. We then studied the class B1 GPCR glucagon-like peptide-1 receptor (GLP-1R) activation by 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. After GLP-1R agonist binding and activation, the coupling of G_s and the release of GDP, the rate-limiting step of the G protein activation, was simulated for the first time using an energy-unbiased method.

These results demonstrate the usefulness of mwSuMD for illuminating the molecular events involved in GPCR function.

Results and Discussion

Short mwSuMD time windows improve the AVP dynamic docking prediction

Arginine vasopressin (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 $G_{q/11}$, while the V_2 receptor (V_2R) activates adenylyl cyclase by interacting with G_s [30] and is a therapeutic target for hyponatremia, hypertension, and incontinence[31]. AVP is amphipathic and in the bound state interacts with both polar and hydrophobic V_2R residues located on TM helices and ECLs (**Figure S1b**). Although AVP presents an intramolecular C1-C6 disulfide bond that limits the backbone's overall conformational flexibility, it has many rotatable bonds, making dynamic docking complicated[32]. We compared the performance of mwSuMD to the parent algorithm 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 X-ray AVP:V_2R complex was 3.80 ± 0.52 Å (**Figure S2**). SuMD[1,18] produced a minimum root mean square deviation (RMSD) to the cryo-EM complex of 4.28 Å, with most of the replicas (*i.e.,* distribution's mode) having an RMSD close to 10 Å (**Figure S3a**). mwSuMD, with the same settings (**Figure S3b, 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 (i.e., distribution's 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 S3c**) and using the SMscore (**Equation 1**) as the acceptance method (**Figure S3d**) worsened the prediction. Supervising distance and RMSD at the same time (**Figure S3e**), 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 1a**), but enhanced mwSuMD accuracy and precision (**Figure 1b-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 1d, Video S1**), in agreement with the AVP deviations in the equilibrium MD simulation of the X-ray AVP:V_2R complex (**Figure S2**)**.** These results confirm the inherent complexity of reproducing the AVP:V_2R complex via dynamic docking and suggest that short time windows can improve mwSuMD performance on this system. However, it is necessary to know the final bound state to employ the RMSD as the supervised metric, while the distance is required to dynamically dock ligands with unknown bound conformation as previously reported[1,23]. Both distance and RMSD-based simulations delivered insights into the binding path and the residues involved along the recognition route. For example, mwSuMD pinpointed V_2R residues $E184^{ECL2}$, $P298^{ECL3}$, and $E303^{ECL3}$ (**Figure S4a**) as involved during AVP binding, although not in contact with the ligand in the orthosteric complex. None of them are yet characterized through mutagenesis studies according to the GPCRdb[33].

Further to binding, a SuMD approach was previously employed to reconstruct the unbinding path of ligands from several GPCRs[2,34]. We assessed mwSuMD's capability to simulate AVP unbinding from V_2R. Five mwSuMD and five SuMD replicas were collected using 100 ps time windows (**Table S1**). Overall, mwSuMD outperformed SuMD in terms of time required to complete a dissociation (**Figure S5, Video S2**), producing dissociation paths

almost 10 times 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 pathway[35,36]. Here, the V_2R residues involved during the dissociation were comparable to the binding (**Figure S4b**), although ECL2 and ECL3 were slightly more involved during the association than the dissociation, in analogy with other class A and B GPCRs[20,35].

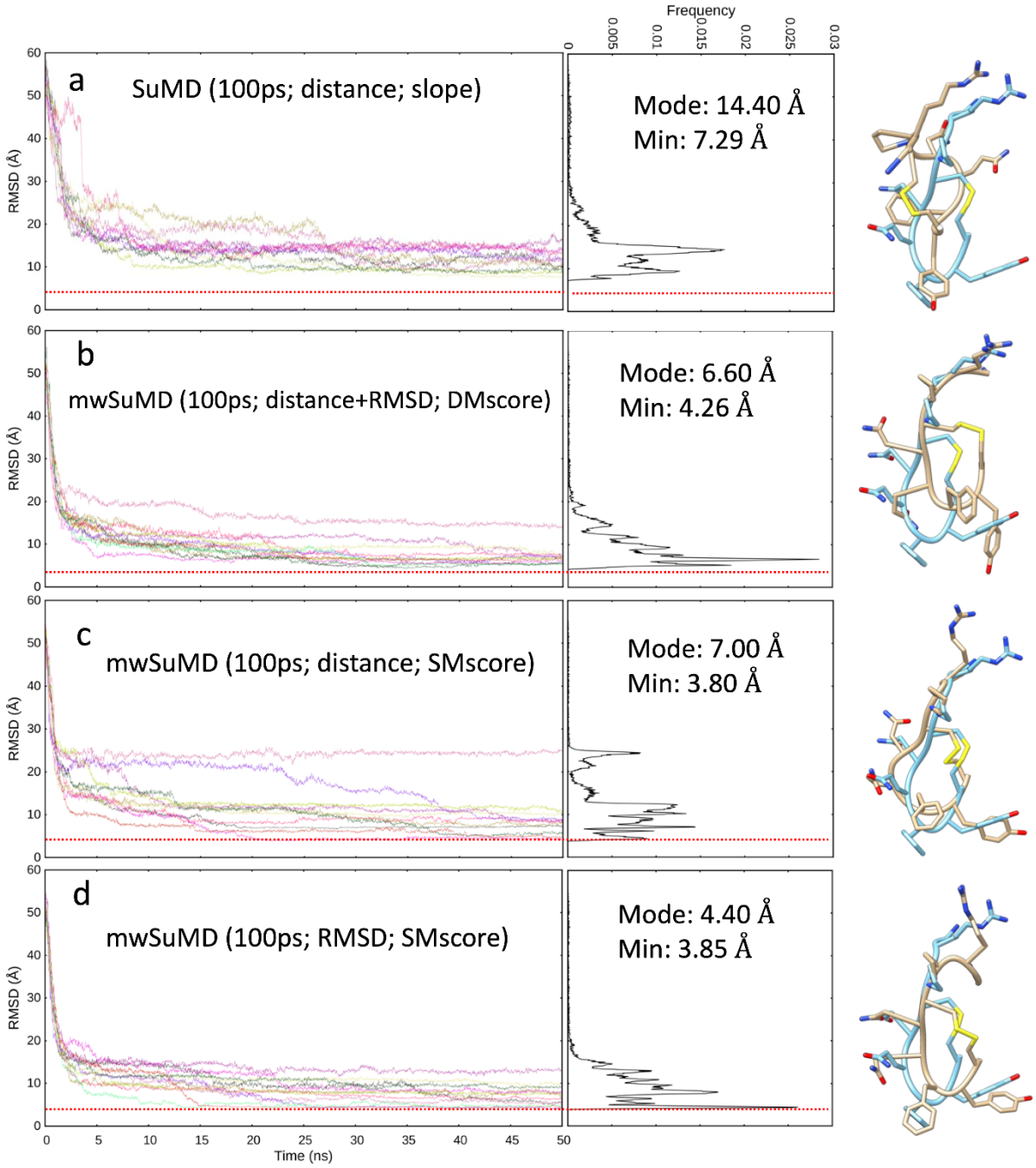

Figure 1. AVP SuMD and mwSuMD binding simulations to V_2R (100 ps time windows). For each set of settings **(a-d)** the RMSD of AVP Cα atoms to the cryo-EM structure 7DW9 is reported during the time course of each SuMD (**a**) or mwSuMD (**b-d**) replica alongside the RMSD values distribution and

the snapshot corresponding to the lowest RMSD values (AVP from the cryo-EM structure 7DW9 is in a cyan stick representation, while AVP from simulations is in a tan stick representation). A complete description of the simulation settings is reported in Table S1 and the Methods section. The dashed red line indicates the AVP RMSD during a classic (unsupervised) equilibrium MD simulation of the X-ray AVP:V_2R complex (**Figure S2**).

PF06882961 binding and GLP-1R activation

The GLP-1R has been captured by cryo-electron microscopy (cryo-EM) in both the inactive apo (ligand-free) and the active (G_s-bound) conformations and in complex with either peptide or non-peptide agonists[37–42]. 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 S6**) 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 S7**). This complicates the scenario and suggests divergent recognition mechanisms amongst different agonists. We simulated the binding of PF06882961 using multistep supervision on different system metrics (**Figure 2**) to model the structural hallmark of GLP-1R activation (**Video S5**, **Video S6**).

Several metrics were supervised consecutively. 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 2, Video S6**). Noteworthy, mwSuMD like any other CV-based technique requires some knowledge of the simulated system. 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[20,43]. 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, were influenced by the movement of supervised helixes or loops and therefore displayed an RMSD reduction to the active state. 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 2**). 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. 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 appears to be driven by direct interactions between the ligand and R310$^{5.40}$ or R380$^{7.35}$. Interestingly, the mwSuMD simulation during Stage 4 (TM6 supervision) sampled a

counterclockwise helix rotation (**Figure S8a**) consistent with the GLP-1R cryo-EM structures in the active, G_s-coupled state[40,44].

It is worth noting that 6LN2, the only inactive GLP-1R structure available complete with the ECD, was stabilized and solved with an antibody bound to the ECD. This strategy might have forced the ECD into a closed conformation that engages the EC *vestibule* of GLP-1R and possibly restrained the whole TMD in an altered conformation that deviates from the physiological conditions. This might explain why the RMSD of the TMD elements monitored during the simulation rarely reach values lower than 3 or 4 Å.

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 of TM6 towards the active state, in the last step of the mwSuMD simulation (stage 4). A more detailed analysis revealed that the central polar network, which is pivotal for mediating GLP-1 signalling[45], and the residues at the TM6 kink level adopted active-like conformations during the final stage of the simulation (**Figure 3a**). In particular, the central polar network (E364$^{6.53}$, H363$^{6.52}$, and Q394$^{7.49}$) experienced side chain rearrangements (**Figure 3b**) while extensive TM6 kink dislocation occurred at L360$^{6.49}$, P358$^{6.48}$, L357$^{6.47}$ and I357$^{6.46}$ (**Figure 3c**).

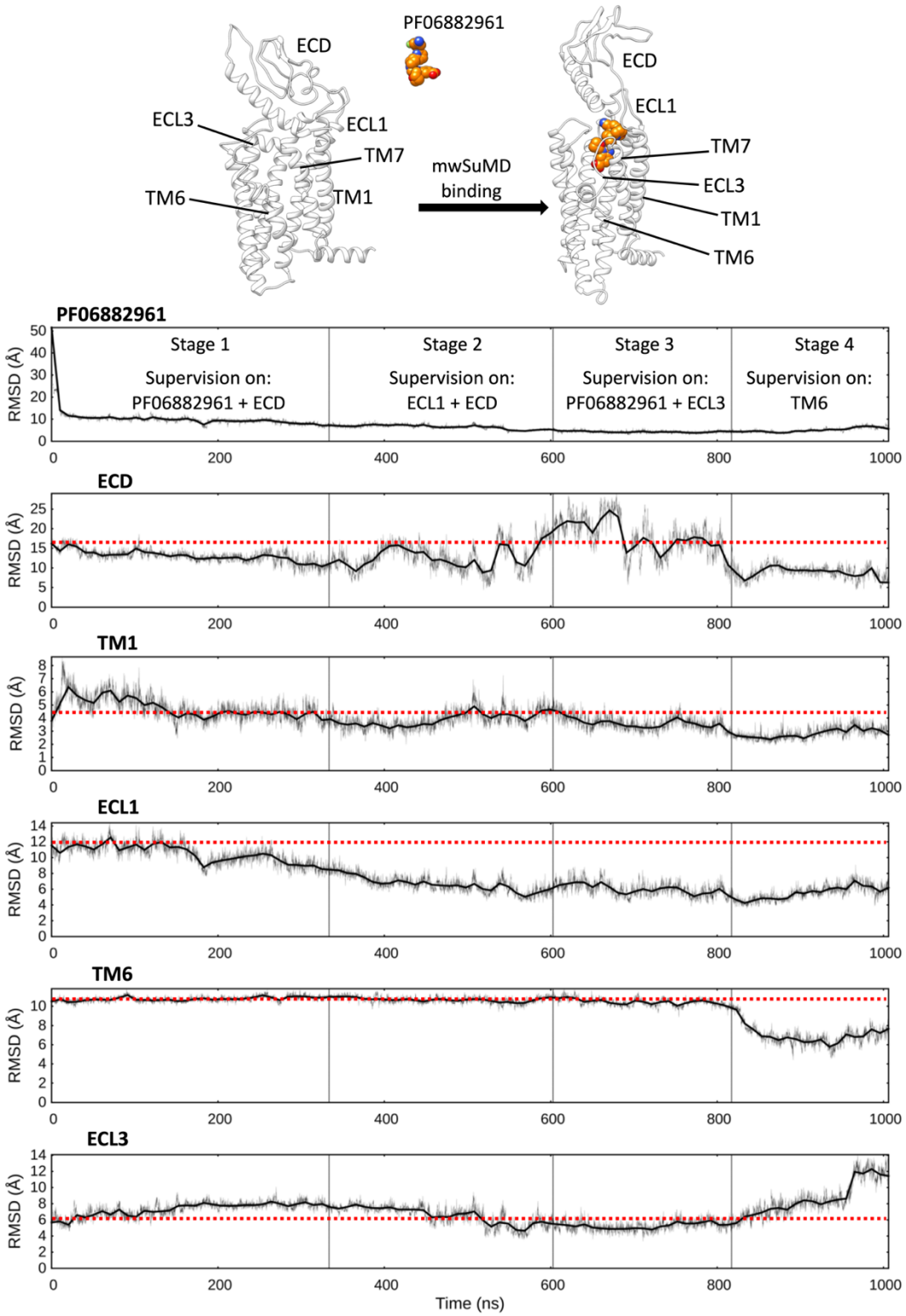

Figure 2. mwSuMD simulation of PF06882961 binding to GLP-1R and receptor activation. Each panel reports the root mean square deviation (RMSD) to the position of the ligand in the active state (top panel) or a GLP-1R structural element, over the time course (all but ECL3 converging to the active state). ECD: extracellular domain; TM: transmembrane helix; ECL: extracellular loop. The mwSuMD simulation was performed with four different settings over 1 microsecond in total. The red dashed lines show the initial RMSD value for reference.

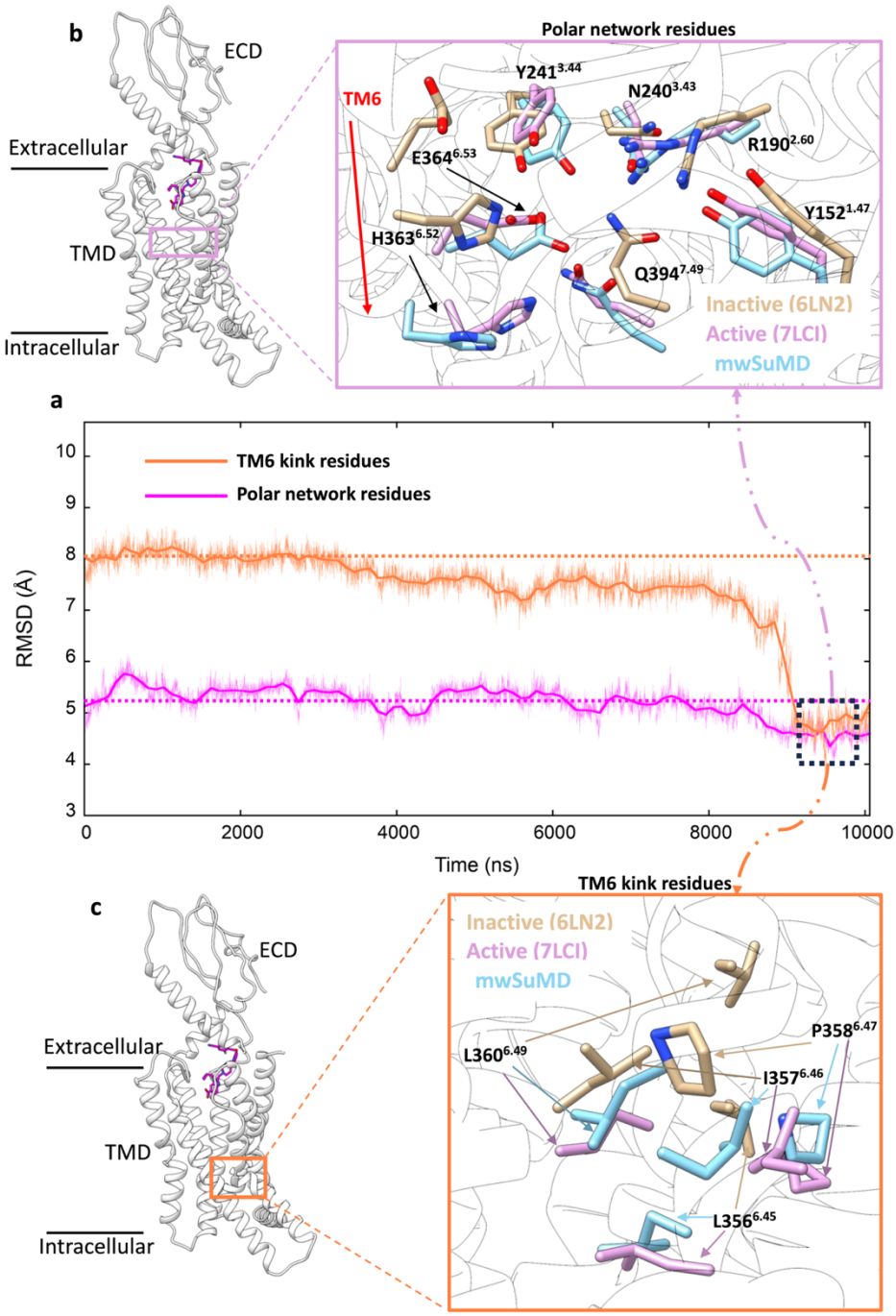

Figure 3. GLP-1R key structural motifs during mwSuMD GLP-1R activation. a) RMSD to the active state GLP-1R (7LCI) of the residues forming the central polar network (magenta) and TM6 kink (orange) during mwSuMD of receptor activation; at the end of the simulations minimum values were reached (dashed square). **b)** The position of the polar network within the core of TMD (left-hand panel) and comparison between inactive, active and mwSuMD final states for the side chains of the residues forming of the polar network; **c)** The position of the TM6 kink (right-hand panel) and comparison between inactive, active and mwSuMD final states for the side chains of the residues forming the TM6 kink.

G_s protein binding to GLP-1R and GDP release

We then focused on simulating the G_s binding to GLP1-R, after activation, *without energy input*. For this purpose, we first tested the binding between the prototypical class A receptor β_2 adrenoreceptor (β_2 AR), and the stimulatory G protein (G_s) (**Video S3, Figure S9a,b**) by supervising the distance between G_s helix 5 (α5) and β_2AR as well as the RMSD of the intracellular end of TM6 to the fully active state of the receptor (see Supplementary Methods). During two out of three replicas, both Gα and Gβ achieved distance 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 S9c).** A possible pitfall is that G proteins bear potential palmitoylation and myristoylation sites that anchor the inactive trimer to the plasma membrane[46,47], *de facto* restraining possible binding paths to the receptor. To address this point and test different conditions, we prepared the adenosine A1 receptor (A_1R), and its principal effector, the inhibitory G protein (G_i) considering the $G_{i\alpha}$ residue C3 and G_γ residue C65 as palmitoylated and geranylgeranylated respectively and hence inserted in the membrane. Recently, the G_i binding to A_1R was simulated by combining the biased methods GaMD with SuMD[48] but without considering membrane-anchoring post-translational modifications. Both classic (unsupervised) and mwSuMD simulations were performed on this system for comparison (**Video S4, Figure S9d)**. In about 50 ns of mwSuMD, the $G_{i\alpha}$ subunit engaged its intracellular binding site on A_1R and formed a complex in good agreement with the cryo-EM structure (PDB 6D9H, RMSD $\approx$ 5 Å). For comparison, 1 μs of cMD did not produce a productive engagement as the $G_{i\alpha}$ remained at RMSD values > 40 Å (**Figure S9d**), suggesting the effectiveness of mwSuMD in sampling G protein binding rare events without the input of energy. 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 S9e)**.

Encouraged by results obtained on G_s and G_i binding to β_2 AR and A_1R (**Figure S9**), we extracted the GLP-1R active conformation described above and simulated the G_s binding to its intracellular side. Starting from the inactive, membrane-anchored G_s, we performed three independent mwSuMD replicas by supervising the distance between G_s helix α5 and GLP-1R residues located at the intracellular binding interface (**Figure 4a,e**). All three mwSuMD replicas showed the G_s approaching GLP-1R, with two out of three reaching an RMSD of the $G_{s\alpha}$ subunit close to or less than 10 Å, compared to the experimental complex 7LCI (**Figure 4e**). Replica 2, in particular, well reproduced the cryo-EM GLP-1R:G_s complex as suggested by RMSDs to 7LCI of 7.59 $\pm$ 1.58 Å, 12.15 $\pm$ 2.13 Å, and 13.73 $\pm$ 2.24 Å for G_α, G_β and G_γ respectively. Such values are not far from the RMSDs measured in our previous simulations of GLP-1R in complex with G_s and GLP-1[49] (G_α = 6.18 $\pm$ 2.40 Å; G_β = 7.22 $\pm$ 3.12 Å; G_γ = 9.30 $\pm$ 3.65 Å), which indicates overall higher flexibility of G_β and G_γ compared to G_α, which acts as a sort of fulcrum bound to GLP-1R.

According to the model of G protein activation, the G protein binding has the effect of allosterically stabilizing the orthosteric agonist in complex with the receptor[49] and destabilizing the guanosine-diphosphate (GDP) bound to Gα, triggering its release and exchange with guanosine-triphosphate (GTP)[50], upon opening of the G protein alpha-helical domain (AHD). Following this model, PF06882961 and GDP were respectively stabilized and destabilized during the simulated G_s association (**Figure S8b,c**). The analysis of atomic contacts along the

binding path of G_s to GLP-1R highlights a few persistent interactions not observed in the equilibrium MD simulations of GLP-1R:G_s cryo-EM complexes[49]; for example, we propose the hidden interaction between D344$^{6.33}$ and R385$^{\alpha5}$ to be important for G_s coupling (**Table S2**), which would explain the GLP-1 EC_{50} reduction upon mutation of position 344 to Ala[51].

Extending Replica 2, we further investigated the G_s activation mechanism by supervising the opening of the G_s alpha-helical domain (AHD), which is considered a necessary step to allow GDP release from the Ras-like domain[52]. We first easily obtained the opening of AHD (**Figure 4b**) and successively supervised the GDP unbinding in a further three replicas, seeded after the AHD opening. In one of these three mwSuMD simulations, the nucleotide dissociated from G_s (**Figure 4d**). **Video S7** shows the full G_s binding**,** AHD opening, and GDP release. mwSuMD suggested several structural changes as implicated in GDP dissociation (**Figure 4f-h**): *i*) the AHD opening; *ii*) the conformational change of αG-α4 loop (residues A303$^{G\alpha}$-P332$^{G\alpha}$), in concert with a loosening of interactions between D323$^{G\alpha}$ and K342$^{6.31}$, and *iii*) the rupture of the hydrogen bond between GDP and D295$^{G\alpha}$ triggered by the movement of αG away from the GDP binding site (**Figure 4f**). The involvement of αG through the αG-α4 during the release of GDP from G_s is supported by hydrogen/deuterium exchange experiments[53], while the role of D275$^{G\alpha}$ has been probed with functional assays on the G_i isoform mutant D272$^{G\alpha i}$A[54]. Moreover, a similar αG behavior was recently suggested by MD simulations of the G_s binding pathway to β_2AR[55]. We also note that the αG-α4 loop length and aminoacidic composition diverge among G protein isoforms, further suggesting a role in G protein selectivity consistent with the hypothesis that, in different G proteins, distinct domains of the Gα subunit could be responsible for receptor selectivity[56]. However, a more subtle allosteric communication through internal structural elements like the β2-β3 strands, prompted by α5 tilting, could have weakened GDP phosphate binding as previously suggested by other groups[57–59]. Interestingly, no significant conformational changes of the β6-α5 loop happened before or during the GDP dissociation, suggesting that its conformational change as captured in the nucleotide-free GLP-1R:G_s complex (**Figure S10**) occurs after the GDP release as a result of the loss of binding stabilization, rather than being an initiator of the GDP dissociation.

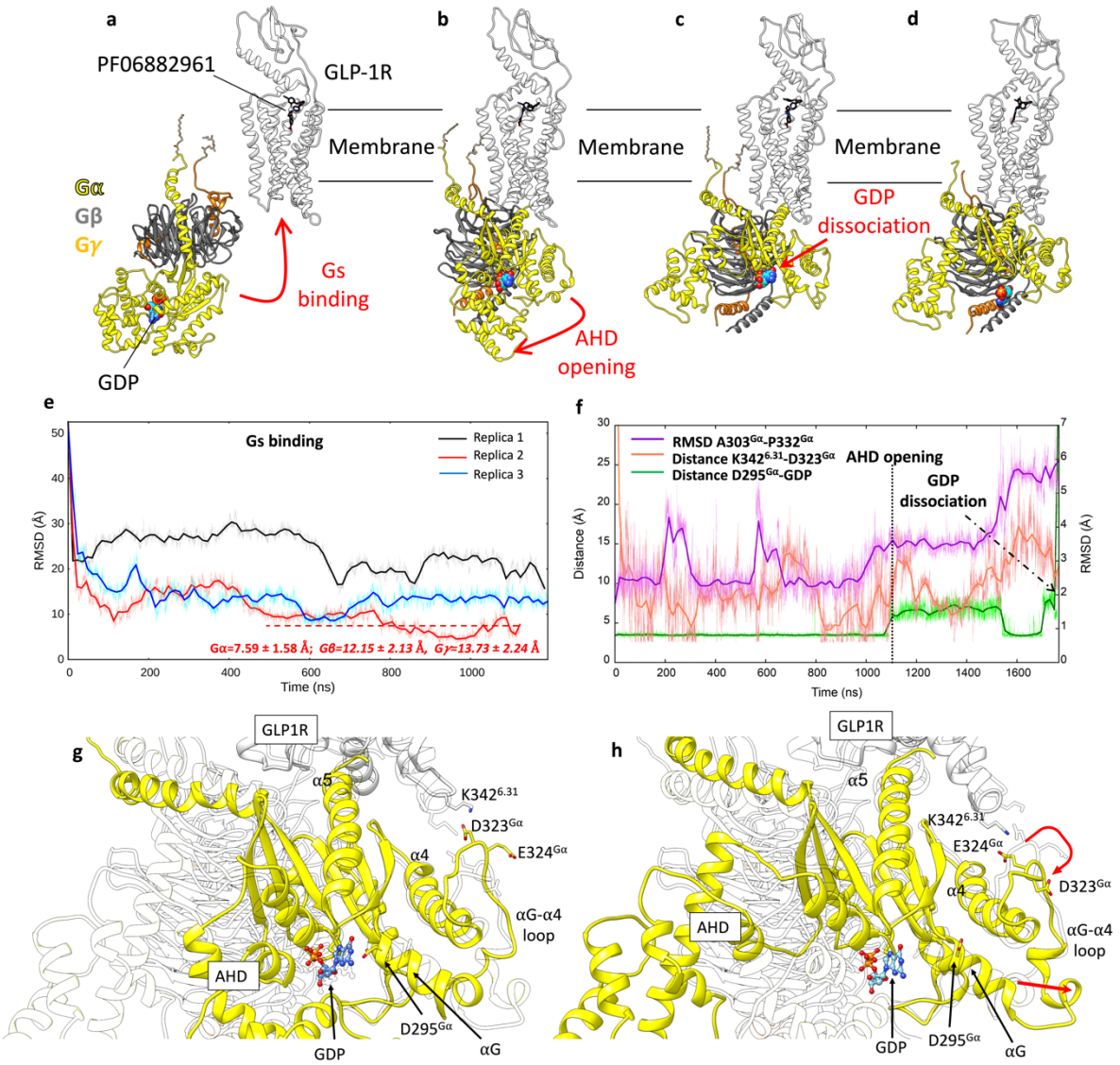

Figure 4. GLP-1R activation and G_s binding. a-d) sequence of simulated events during the mwSuMD G_s:GLP-1R simulations. **e)** RMSD of $G_s\alpha$ to the experimental GLP-1R:G_s complex (PDB 7LCJ) during three mwSuMD replicas; the RMSD to the experiential bound conformation (7LCJ) during the second part of Replica2 (red dashed line) is reported for each G_s subunit. RMSDs were computed on G_α residues 11 to 43 and 205 to 394 to the experimental structure 7LCI after superimposition on GLP-1R residues 140-240 Cα atoms. **f)** RMSD of the αG-α4 loop (purple), the distance between $K342^{6.31}$ and $D323^{G\alpha}$ (salmon), and the distance between GDP and $D295^{G\alpha}$ (green) during G_s binding, AHD opening and GDP dissociation; **g)** and **h)** comparison between states extracted from before and after AHD opening. Before AHD opening (**a**), GLP-1R ICL3 interacted with $D323^{G\alpha}$ and $D295^{G\alpha}$ interacted with GDP; after AHD opening (**b**) and αG-α4 loop reorganization (curved red arrow), αG and $D295^{G\alpha}$ moved away from GDP (straight red arrow), destabilizing its binding to G_s.

Discussion

Classic MD simulations sample 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 creating bottlenecks that slow the transition down to the millisecond, or second, time scale. To reduce some of these limits, we have developed, and tested on complex structural events characterizing GPCRs, an energetically unbiased adaptive sampling algorithm, namely multiple walker SuMD, which is based on traditional SuMD, while drawing on parallel multiple replica methods[60,61],

Our simulations propose that remarkable predictivity can be obtained with distance-driven mwSuMD, as demonstrated by the lowest deviation from the experimental AVP:V2R complex. 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. This is due to the more extensive sampling obtainable by seeding multiple parallel short simulations instead of a single simulation for batch.

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 either 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[62]) models.

We increased the complexity of binding simulations by considering GLP-1R and the non-peptide agonist PF06882961. Using mwSuMD, we obtained a binding of PF06882961in 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 the structural data available[40] and extensive preparatory MD simulations. However, binding routes are possible from either the bulk solvent or the membrane[35,63,64]. These results show the power of the mwSuMD method, indicating that future applications could include ligand binding from the membrane, or alternative apo receptor conformations to improve the sampling for more difficult receptors.

mwSuMD enabled us to simulate the G_s binding to the active GLP-1R and the subsequent GDP release. Our results suggest a concerted effect on GDP binding produced by AHD opening, αG-α4 loop rearrangement, and αG shift away from the GDP site. The full rotation and elongation of α5 as in cryo-EM structures would occur after the GDP release, supporting the role of hidden, metastable interactions as the driving force of G protein coupling and selectivity, as per recent work on GLP-1R[51].

We stress that a complete understanding of a complex molecular event like G protein coupling requires the collection of numerous G protein binding paths and GDP dissociation events. mwSuMD is designed to yield a mechanistic description of structural events. For this reason, it integrates well with mutagenesis and kinetics experiments. We note that mwSuMD trajectories, since they describe the sequential states along a transition pathway, can

represent a precious backbone for further MD sampling aimed at quantifying or predicting the kinetics of the transition. Approaches such as path collective variable metadynamics[65], Markov State Models[66], or machine learning models[67,68] can be informed by mwSuMD. We will address these novel opportunities created by mwSuMD in future work.

In summary, we showcased the applicability domain of mwSuMD to key, but scarcely understood, aspects of GPCR structural biology, pharmacology, and drug design hitherto unaddressed by unbiased simulations. Given the generality and simplicity of its implementation, we anticipate that mwSuMD can be employed to study a wide range of structural phenomena characterizing both membrane and cytosolic proteins. mwSuMD is an ongoing project updated on https://github.com/pipitoludovico/mwSuMD_multiEngine and in the recent implementation can exploit ACEMD[69], NAMD[70], GROMACS[71] and OPENMM[72] as graphic processing units (GPU)-based MD engines while offering the option to run also on CPUs with GROMACS.

Methods

Multiple walker SuMD (mwSuMD) protocol

The supervised MD (SuMD) is an adaptive sampling method[73] for speeding up the simulation of binding events between small molecules (or peptides[74,75]) and proteins[1,18] without the introduction of any energetic bias. Briefly, during 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, otherwise, the simulation is restarted by randomly assigning the atomic velocities.

mwSuMD is designed to increase the sampling from a specific configuration by seeding user-decided parallel replicas (walkers) rather than one short simulation as per SuMD. Since one replica for each batch of walkers is always considered productive, mwSuMD gives more control than SuMD on the total wall-clock time used for a simulation. On the flip side, to maximise mwSuMD it is optimal to assign one walker per GPU, requiring multiple GPUs to be effective. However, modern multi-threaded GPUs can still employ mwSuMD with a smaller cost in GPU performance. In the implementation for ACEMD used in this work, 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 field 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 is 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), the walker with the highest score is continued. 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:

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

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 in the case of distance and RMSD (both of which are distance-based). Notably, differently from SuMD, 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[76] and MDTRaj[77] modules.

Force field, ligands parameters, and general systems preparation

The CHARMM36[78,79]/CGenFF 3.0.1[80–82] force field combination was employed in this work. Initial ligand force field, topology and parameter files were obtained from the ParamChem webserver[80]. Restrained electrostatic potential (RESP)[83] 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[84] and propka[85] 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[86]. Receptor orientation was obtained by superposing the coordinates on the corresponding structure retrieved from the OPM database[87]. Lipids overlapping the receptor transmembrane helical bundle were removed and TIP3P water molecules[88] were added to the simulation box employing 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 ACEMD3[69] was employed for both the equilibration and productive simulations. The equilibration was achieved in isothermal-isobaric conditions (NPT) using the Berendsen barostat[89] (target pressure 1 atm) and the Langevin thermostat[90] (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[91,92] 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)[93] 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[94] was retrieved from the Protein Data Bank[95] (PDB 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[96], considering the solution with the lowest DOPE score out of ten conformations produced. 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 (backbone and side chains), anticipated to bind deep into V_2R, and the V_2R residues lining the peptide binding site 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 (backbone and side chains) 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[95] (PDB 6LN2)[97]. Missing residues in the stalk (129-134) and ICL2 (256 to 263) were modeled with Modeller 9.19, considering the solutions with the lowest DOPE score out of ten conformations produced. The PF06882961 initial conformation was extracted from the complex with the fully active GLP-1R[44] (PDB 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[98].

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's heavy atoms centroid and the centroid of GLP-1R TM7 residues $L379^{7.34}$-$F381^{7.36}$ (Cα atoms only), which are part of the orthosteric site, and the RMSD of the ECD (residues $W33^{ECD}$-$W120^{ECD}$, Cα atoms only) to the active state (PDB 7LCJ) until the former distance reached 4 Å. In the second phase of mwSuMD, the RMSD of the ECD (residues $W33^{ECD}$-$W120^{ECD}$, Cα atoms only) and the ECL1 to the active state (PDB 7LCJ, Cα atoms of residues $M204^{2.74}$-$L224^{3.27}$) 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 $A368^{6.57}$-$T378^{7.33}$, Cα atoms), were supervised until the former reached values lower than 3 Å. In the last mwSuMD step, only the RMSD of TM6 (residues $I345^{6.34}$-$F367^{6.56}$, Cα atoms) to the active state (PDB 7LCJ) was supervised until less than 5 Å. RMSDs were computed after superimposition on TM2, ECL1, and TM3 residues 170-240 (Cα atoms), which is the GLP-1R less flexible part[49].

Membrane-anchored G_s protein:GLP-1R simulations and GDP dissociation.

The PDB 6EG8 was processed through Charmm-GUI[99] to palmitoylate residue $C3^{G\alpha i}$ and geranylgeranylate residue $C65^{G\gamma}$. The resulting system was inserted into a 120 x 120 Å POPC membrane and 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[88] 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 G_s protein to stabilize within the intracellular side of the simulation box. After this two-stage long equilibration, GLP-1R from the final frame of the activation simulation (in complex with PF06882961) was manually inserted into the equilibrated membrane above the G_s 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 GLP-1R-G_s system was then subjected to three simulations (**Table S1**). Each mwSuMD replica was interrupted by 500 ns of classic MD twice, to relax the system during the transition. In the supervised stages, the distance between residues M386-L394 $^{G\alpha s}$ (all-atoms centroid) of helix 5 (α5) and the GLP-1R intracellular residues R176$^{2.46}$, R348$^{6.37}$, S352$^{6.41}$, and N405$^{7.60}$ (Cα atoms only) was monitored, seeding three walkers of 200 ps each.

The AHD opening was simulated starting from the GLP-1R:G_s binding mwSuMD replica with the final lowest G_s RMSD, the lowest PF06882961 binding energy and the highest GDP binding energy (Replica 2 in **Figure 4e**) by supervising the distance between AHD residues G70-R199 $^{G\alpha s}$ and K300-L394$^{G\alpha s}$ (all-atoms centroids) during three walkers of 100 ps each. 300 ns of classic MD was performed to relax the system. Finally, the GDP unbinding was supervised as the distance between GDP (all-atoms centroid) and residues E50$^{G\alpha s}$, K52$^{G\alpha s}$, T55$^{G\alpha s}$, K293$^{G\alpha s}$, and V367$^{G\alpha s}$ (Cα atoms only) of $G_{\alpha s}$; five walkers were used in a 50 ps long mwSuMD simulations.

MD Analysis

Interatomic distances were computed through MDAnalysis[76]; root mean square deviations (RMSD) were computed using VMD[100] and MDAnalysis[76]. 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[101] 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[102] (v1.14).

Numbering system

Throughout the manuscript, the Ballesteros-Weinstein residues numbering system for class A[103] and the Wootten residues numbering system for class B GPCRs[104] 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/7944479

The mwSuMD software version used in this study is available at:

https://github.com/pipitoludovico/mwSuMD

Author Contributions

GD and CAR conceived and supervised the project; GD designed and implemented 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. GD and CAR are grateful for support from the BBSRC (BB/W016974/1) and Diabetes UK (BDA 20/0006307).

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

References

(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) 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.

(9) 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.

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

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

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

(13) 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.

(14) 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.

(15) Zuckerman, D. M.; Chong, L. T. Weighted Ensemble Simulation: Review of Methodology, Applications, and Software. *Annu. Rev. Biophys.* **2017**, *46*, 43–57.

(16) Fleetwood, O.; Matricon, P.; Carlsson, J.; Delemotte, L. Energy Landscapes Reveal Agonist Control of G Protein-Coupled Receptor Activation via Microswitches. *Biochemistry* **2020**, *59*, 880–891.

(17) Lovera, S.; Cuzzolin, A.; Kelm, S.; De Fabritiis, G.; Sands, Z. A. Reconstruction of Apo A2A Receptor Activation Pathways Reveal Ligand-Competent Intermediates and State-Dependent Cholesterol Hotspots. *Sci. Rep.* **2019**, *9*, 14199.

(18) 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.
(19) 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.
(20) 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.
(21) 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.
(22) 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.* **2022**, *18*, 256–263.
(23) 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.
(24) 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.
(25) Mattedi, G.; Acosta-Gutiérrez, S.; Clark, T.; Gervasio, F. L. A Combined Activation Mechanism for the Glucagon Receptor. *Proc. Natl. Acad. Sci. USA* **2020**.
(26) 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.
(27) Mafi, A.; Kim, S.-K.; Goddard, W. A. The Dynamics of Agonist-B2-Adrenergic Receptor Activation Induced by Binding of GDP-Bound Gs Protein. *Nat. Chem.* **2023**, *15*, 1127–1137.
(28) 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.
(29) Guo, D.; Heitman, L. H.; IJzerman, A. P. The Role of Target Binding Kinetics in Drug Discovery. *ChemMedChem* **2015**, *10*, 1793–1796.
(30) Birnbaumer, M. Vasopressin Receptors. *Trends Endocrinol. Metab.* **2000**, *11*, 406–410.
(31) Ball, S. G. Vasopressin and Disorders of Water Balance: The Physiology and Pathophysiology of Vasopressin. *Ann Clin Biochem* **2007**, *44*, 417–431.
(32) Gioia, D.; Bertazzo, M.; Recanatini, M.; Masetti, M.; Cavalli, A. Dynamic Docking: A Paradigm Shift in Computational Drug Discovery. *Molecules* **2017**, *22*.
(33) 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*, D356-64.
(34) 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.
(35) 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.

(36) Deganutti, G.; Barkan, K.; Ladds, G.; Reynolds, C. A. A Multisite Model of Allosterism for the Adenosine A1 Receptor. *BioRxiv* **2020**.

(37) 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.

(38) 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.

(39) 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.

(40) 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.

(41) 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.

(42) 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**.

(43) 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.

(44) 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.

(45) Wootten, D.; Reynolds, C. A.; Koole, C.; Smith, K. J.; Mobarec, J. C.; Simms, J.; Quon, T.; Coudrat, T.; Furness, S. G. B.; Miller, L. J.; Christopoulos, A.; Sexton, P. M. A Hydrogen-Bonded Polar Network in the Core of the Glucagon-Like Peptide-1 Receptor Is a Fulcrum for Biased Agonism: Lessons from Class B Crystal Structures. *Mol. Pharmacol.* **2016**, *89*, 335–347.

(46) 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.

(47) 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.

(48) 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.
(49) Deganutti, G.; Liang, Y.-L.; Zhang, X.; Khoshouei, M.; Clydesdale, L.; Belousoff, M. J.; Venugopal, H.; Truong, T. T.; Glukhova, A.; Keller, A. N.; Gregory, K. J.; Leach, K.; Christopoulos, A.; Danev, R.; Reynolds, C. A.; Zhao, P.; Sexton, P. M.; Wootten, D. Dynamics of GLP-1R Peptide Agonist Engagement Are Correlated with Kinetics of G Protein Activation. *Nat. Commun.* **2022**, *13*, 92.
(50) Gregorio, G. G.; Masureel, M.; Hilger, D.; Terry, D. S.; Juette, M.; Zhao, H.; Zhou, Z.; Perez-Aguilar, J. M.; Hauge, M.; Mathiasen, S.; Javitch, J. A.; Weinstein, H.; Kobilka, B. K.; Blanchard, S. C. Single-Molecule Analysis of Ligand Efficacy in B2AR-G-Protein Activation. *Nature* **2017**, *547*, 68–73.
(51) Yuan, S.; Xia, L.; Wang, C.; Wu, F.; Zhang, B.; Pan, C.; Fan, Z.; Lei, X.; Stevens, R. C.; Sali, A.; Sun, L.; Shui, W. Conformational Dynamics of the Activated GLP-1 Receptor-G_s Complex Revealed by Cross-Linking Mass Spectrometry and Integrative Structure Modeling. *ACS Cent. Sci.* **2023**.
(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) Du, Y.; Duc, N. M.; Rasmussen, S. G. F.; Hilger, D.; Kubiak, X.; Wang, L.; Bohon, J.; Kim, H. R.; Wegrecki, M.; Asuru, A.; Jeong, K. M.; Lee, J.; Chance, M. R.; Lodowski, D. T.; Kobilka, B. K.; Chung, K. Y. Assembly of a GPCR-G Protein Complex. *Cell* **2019**, *177*, 1232–1242.e11.
(54) Ham, D.; Ahn, D.; Ashim, J.; Cho, Y.; Kim, H. R.; Yu, W.; Chung, K. Y. Conformational Switch That Induces GDP Release from Gi. *J. Struct. Biol.* **2021**, *213*, 107694.
(55) Batebi, H.; Pérez-Hernández, G.; Rahman, S. N.; Lan, B.; Kamprad, A.; Shi, M.; Speck, D.; Tiemann, J. K. S.; Guixà-González, R.; Reinhardt, F.; Stadler, P. F.; Papasergi-Scott, M. M.; Skiniotis, G.; Scheerer, P.; Kobilka, B. K.; Mathiesen, J. M.; Liu, X.; Hildebrand, P. W. Mechanistic Insights into G-Protein Coupling with an Agonist-Bound G-Protein-Coupled Receptor. *Nat. Struct. Mol. Biol.* **2024**.
(56) Glukhova, A.; Draper-Joyce, C. J.; Sunahara, R. K.; Christopoulos, A.; Wootten, D.; Sexton, P. M. Rules of Engagement: Gpcrs and G Proteins. *ACS Pharmacol. Transl. Sci.* **2018**, *1*, 73–83.
(57) Sun, X.; Singh, S.; Blumer, K. J.; Bowman, G. R. Simulation of Spontaneous G Protein Activation Reveals a New Intermediate Driving GDP Unbinding. *Elife* **2018**, *7*.
(58) Kaya, A. I.; Lokits, A. D.; Gilbert, J. A.; Iverson, T. M.; Meiler, J.; Hamm, H. E. A Conserved Hydrophobic Core in Gαi1 Regulates G Protein Activation and Release from Activated Receptor. *J. Biol. Chem.* **2016**, *291*, 19674–19686.
(59) Flock, T.; Ravarani, C. N. J.; Sun, D.; Venkatakrishnan, A. J.; Kayikci, M.; Tate, C. G.; Veprintsev, D. B.; Babu, M. M. Universal Allosteric Mechanism for Gα Activation by GPCRs. *Nature* **2015**, *524*, 173–179.
(60) 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.
(61) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
(62) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478–9508.

(63) 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.
(64) 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**.
(65) Pérez de Alba Ortíz, A.; Vreede, J.; Ensing, B. The Adaptive Path Collective Variable: A Versatile Biasing Approach to Compute the Average Transition Path and Free Energy of Molecular Transitions. *Methods Mol. Biol.* **2019**, *2022*, 255–290.
(66) Husic, B. E.; Pande, V. S. Markov State Models: From an Art to a Science. *J. Am. Chem. Soc.* **2018**, *140*, 2386–2396.
(67) Kokh, D. B.; Kaufmann, T.; Kister, B.; Wade, R. C. Machine Learning Analysis of TRAMD Trajectories to Decipher Molecular Determinants of Drug-Target Residence Times. *Front. Mol. Biosci.* **2019**, *6*, 36.
(68) Pérez, A.; Martínez-Rosell, G.; De Fabritiis, G. Simulations Meet Machine Learning in Structural Biology. *Curr. Opin. Struct. Biol.* **2018**, *49*, 139–144.
(69) 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.
(70) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
(71) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput Phys Commun* **1995**, *91*, 43–56.
(72) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput. Biol.* **2017**, *13*, e1005659.
(73) Deganutti, G.; Moro, S. Estimation of Kinetic and Thermodynamic Ligand-Binding Parameters Using Computational Strategies. *Future Med. Chem.* **2017**, *9*, 507–523.
(74) 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.
(75) 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.
(76) 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.
(77) 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.
(78) 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.
(79) 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.
(80) 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.
(81) 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.
(82) 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.
(83) Woods, R. J.; Chappelle, R. Restrained Electrostatic Potential Atomic Partial Charges for Condensed-Phase Simulations of Carbohydrates. *Theochem* **2000**, *527*, 149–156.
(84) 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.
(85) 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.
(86) Sommer, B. Membrane Packing Problems: A Short Review on Computational Membrane Modeling Methods and Tools. *Comput Struct Biotechnol J* **2013**, *5*, e201302014.
(87) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. OPM: Orientations of Proteins in Membranes Database. *Bioinformatics* **2006**, *22*, 623–625.
(88) 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.
(89) 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.
(90) 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.
(91) Forester, T. R.; Smith, W. SHAKE, Rattle, and Roll: Efficient Constraint Algorithms for Linked Rigid Bodies. *J. Comput. Chem.* **1998**.
(92) Krutler, V.; van Gunsteren, W. F.; Hanenberger, 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.
(93) 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.
(94) 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.
(95) 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.
(96) Fiser, A.; Sali, A. Modeller: Generation and Refinement of Homology-Based Protein Structure Models. *Meth. Enzymol.* **2003**, *374*, 461–491.

(97) 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.

(98) 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.

(99) 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.

(100) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J Mol Graph* **1996**, *14*, 33–38, 27.

(101) 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.

(102) 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.

(103) 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.

(104) 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.