SCIENTIFIC REPORTS

OPEN

Molecular details of dimerization kinetics reveal negligible populations of transient μ-opioid receptor homodimers at physiological concentrations

Derya Meral[1], Davide Provasi[1], Diego Prada-Gracia[1], Jan Möller[2,3], Kristen Marino[1], Martin J. Lohse[2,3] & Marta Filizola [1]

Various experimental and computational techniques have been employed over the past decade to provide structural and thermodynamic insights into G Protein-Coupled Receptor (GPCR) dimerization. Here, we use multiple microsecond-long, coarse-grained, biased and unbiased molecular dynamics simulations (a total of ~4 milliseconds) combined with multi-ensemble Markov state models to elucidate the kinetics of homodimerization of a prototypic GPCR, the μ-opioid receptor (MOR), embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC)/cholesterol lipid bilayer. Analysis of these computations identifies kinetically distinct macrostates comprising several different short-lived dimeric configurations of either inactive or activated MOR. Calculated kinetic rates and fractions of dimers at different MOR concentrations suggest a negligible population of MOR homodimers at physiological concentrations, which is supported by acceptor photobleaching fluorescence resonance energy transfer (FRET) experiments. This study provides a rigorous, quantitative explanation for some conflicting experimental data on GPCR oligomerization.

Methodological limitations and inadequate structural information have thus far hampered a full delineation of the thermodynamics, kinetics, and stoichiometry of G Protein-Coupled Receptor (GPCR) interactions in living cells, including the role of di-/oligomerization in GPCR function. The majority of published studies examining di-/oligomerization report on effects derived from activation of one receptor in the presence of another (e.g., see[1]) or modulation of the activity of one GPCR using ligands targeting another one (e.g., see[2,3]). Although there is clear evidence of signaling from a dimeric unit in the case of GPCR members of the glutamate family[4,5], differential signaling of a receptor complex is still a matter of debate for rhodopsin family GPCRs[6–10] in spite of unambiguous evidence of physical interaction for many of them[11,12], most recently provided by single-molecule imaging[13–15]. Notably, these studies suggest transient formation of receptor dimers, with only a small fraction of the receptors existing as dimers at any given time.

Although very powerful and informative, the experimental technologies employed so far do not provide the level of molecular detail that is required to eventually understand the role of di-/oligomerization in GPCR function. To this end, molecular dynamics (MD) simulations provide particularly useful information, and we and others have used various MD-based techniques to derive structural and thermodynamic information about GPCR dimers/oligomers (e.g.[16,17]). Using the opioid receptors as model systems, we recently carried out extensive MD simulations of coarse-grained (CG) representations of arrays of 16 inactive μ-opioid receptors (MOR), δ-opioid receptors (DOR), or κ-opioid receptors (KOR) to study their self-assembly in a membrane-mimetic environment[16,18]. While umbrella sampling simulations of putative dimer interfaces of inactive MOR or KOR inferred by their crystal structures (i.e., the symmetric interfaces involving transmembrane helices (TM) 5 and 6 (TM5,6/TM5,6), and/or transmembrane helices 1, 2 and helix 8 (TM1,2,H8/TM1,2,H8)) confirmed their thermodynamic

[1]Department of Pharmacological Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [2]Max Delbrück Center for Molecular Medicine, Berlin, Germany. [3]Institute of Pharmacology and Toxicology, Würzburg, Germany. Correspondence and requests for materials should be addressed to M.F. (email: marta.filizola@mssm.edu)

stability in a 1-palmitoyl,2-oleoyl-sn-glycero-3-phosphocholine (POPC)/10% cholesterol (CHOL) membrane model, the TM5,6/TM5,6 interface did not form during unbiased simulations of receptor self-association, leading us to conclude that this interface was kinetically unable to form on the microsecond timescale of those simulations[16].

Here, we take advantage of state-of-the-art algorithms using biased and unbiased MD trajectories to investigate longer timescales, and provide, for the first time, molecular details of the kinetics of the GPCR dimerization process through Markov state model (MSM) analysis of multiple microsecond-long, coarse-grained simulations of pairs of inactive or activated MOR crystal structures embedded in a POPC/CHOL membrane model. The results of this analysis are supported by acceptor photobleaching fluorescence resonance energy transfer (FRET) experiments.

Results

With the intention of sampling the association/dissociation between inactive or activated structures of MOR, 352 independent systems with randomly-oriented protomer pairs for each receptor conformation were simulated for at least 5 μs each using the MARTINI force field[19–22], adding up to ~2 milliseconds (ms) of total simulation time for each system (see Supplementary Table 1 for a summary of all simulations presented here). The resulting simulation data were recast in terms of three collective variables (CVs), specifically: (i) the distance d between the centers-of-mass (COM) of the protomers, (ii) the angle α_1 formed between the COM of transmembrane helix 1 of one protomer, its COM, and the COM of the second protomer, and (iii) the angle α_2 obtained by swapping the two protomers. The resulting distributions of the relative position of the protomers sampled during unbiased simulations of inactive or activated MOR structures can be seen in Supplementary Fig. 1, where the three aforementioned CVs, [d, α_1, α_2], have been projected onto two dimensions to obtain the probability distibution of the relative position of the protomers' centers of mass, i.e. x = $d \cos\alpha$ and $y = d \sin\alpha$ (see figure legend for additional details). A preliminary MSM was built after k-means clustering of the CVs to help select regions that are relevant to receptor dimerization using transition path theory (TPT). Configurations within these regions were selected for further exploration by umbrella sampling in order to aid convergence (see Methods section for details). Subsequently, the unbiased and biased simulations were analyzed in tandem using the transition-based reweighting method (TRAM)[23]. In the following sections, we present the structural and kinetic information obtained from the multi-ensemble MSM produced by the TRAM analysis.

Preferred orientations of inactive or activated MOR structures in putative dimeric configurations.

Figure 1 shows the free energies of the relative positions of the MOR protomer pairs obtained from the TRAM analysis. Two highly populated regions, clustered around $\alpha_1, \alpha_2 \approx 0°$ and $\alpha_1, \alpha_2 \approx \pi$ and below distances of 4.5 nm for either inactive or activated MOR conformations suggest a preference for MOR homodimerization through TM4,5,6,7 or TM1,2,H8 contacts, respectively. Notably, the region around $\alpha_1, \alpha_2 \approx \pi$ presents a broader distribution in the inactive MOR simulations (see larger dark red area in Fig. 1a) compared to those of the activated MOR (Fig. 1b), with the region near TM4 (orange helix in the insets of Fig. 1) of the activated MOR structure less populated compared to that of the inactive MOR.

In order to further elucidate the preferred protomer orientations within putative dimeric configurations of MOR, the microstates of the MSM produced by the TRAM analysis were clustered using the PCCA+ algorithm. Specifically, this method allows for the separation of long- and short-lived microstates and distinguishes the unbound (i.e., monomeric) microstates from the longer-lived dimeric or dimer-like microstates (for details, see Methods section).

Removing all microstates of the monomeric cluster from the transition matrix revealed four main kinetically disconnected regions for both the inactive and activated MOR systems (see Figs 2 and 3, respectively). These regions, herein referred to as components, encompass spatially and kinetically linked PCCA clusters of microstates between which there are no direct transitions; instead, the progression from one component to another requires passage through the monomeric PCCA cluster. These four main components represent four possible protomer-protomer orientations, two asymmetric – i.e., ($\alpha_1 \approx 0, \alpha_2 \approx \pi$) and ($\alpha_1 \approx \pi, \alpha_2 \approx 0$) – and two symmetric – i.e., ($\alpha_1 \approx 0, \alpha_2 \approx 0$) and ($\alpha_1 \approx \pi, \alpha_2 \approx \pi$) – referred to as $C_{0\pi}$, $C_{\pi 0}$, C_{00}, and $C_{\pi\pi}$, respectively. The two asymmetric components, $C_{0\pi}$ and $C_{\pi 0}$, whose structures can be grouped together given that protomers are identical, encompass a much larger population than the symmetric components (see red nodes in Figs 2 and 3), forming 87% and 89% of the populations of the bound-like PCCA clusters for the inactive and activated MOR systems, respectively. The symmetric component C_{00} (purple nodes in Figs 2 and 3) forms 2% and 9% of the total bound-like population in the inactive and activated MOR systems, respectively, while $C_{\pi\pi}$ (yellow nodes in Figs 2 and 3) represents 11% and 2% of the population in the same systems. While these four components encompass all microstates of the inactive MOR system with the exception of the unbound states (Fig. 2), two small additional clusters of microstates (0.1% and 0.06% of the stationary occupancy; blue and green circles in Fig. 3) are identified in the activated MOR system. Within each component, the MOR protomers sample a broad range of orientations, which can be further grouped according to the PCCA clusters within those components. Specifically, by imposing symmetry over α_1 and α_2, 16 and 15 groups of receptor configurations, referred to as macrostates, were obtained for the inactive and activated MOR systems, respectively (see Methods section for further details).

In order to discriminate between compact dimeric complexes of MOR and loosely bound protomer pairs, we enforced a definition of dimeric states similar to that used in references[16] and[18]. Specifically, bound MOR protomer pairs were considered to be fully formed dimers whenever each protomer had at least 10 inter-protomer contacts within a distance of 0.8 nm. A dimeric microstate was defined as any microstate for which the probability to observe a dimeric configuration is higher than 90%. Using this definition, we calculated contact maps for all dimeric microstates within each macrostate, and report the different dimeric interactions in each macrostate in Supplementary Figs 2 and 3 for inactive and activated MOR, respectively. Corresponding structures for each
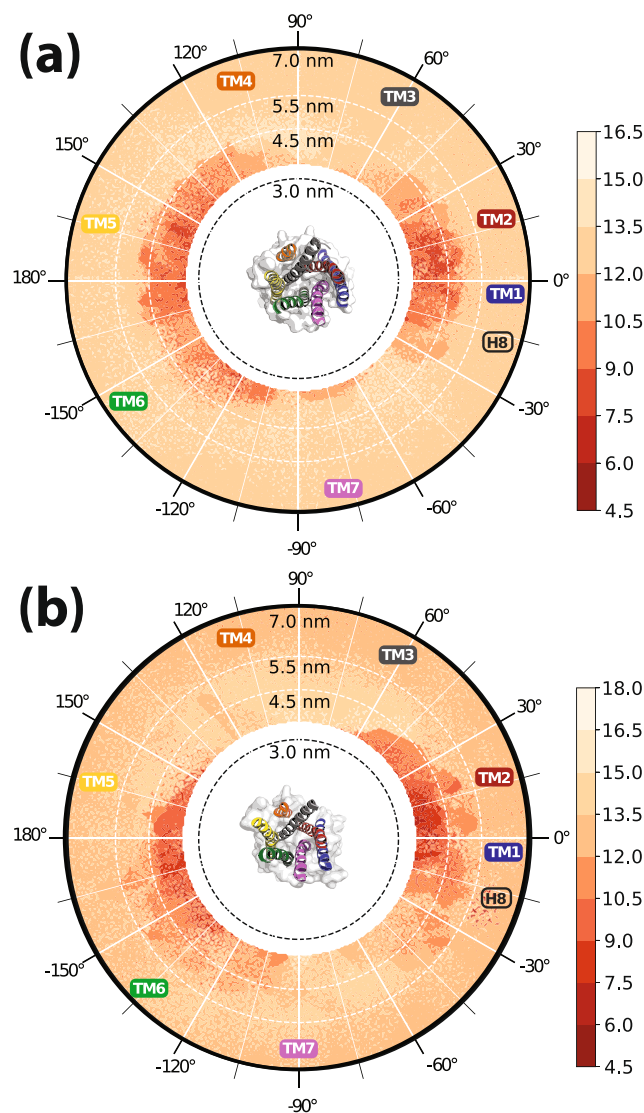
Figure 1. Free energy landscapes of the (**a**) inactive and (**b**) activated μ-opioid receptor (MOR) simulations derived using the transition-based reweighting method (TRAM). In each figure, the CVs, [d, α_1, α_2], have been projected onto the d and α dimensions, taking into account the interchangeability of the two protomers by duplicating the trajectories through swapping of the two angular CVs prior to the free energy calculation. Dark red/light orange color coding is matched to lowest/highest free-energy values. The insets in the center refer to surface representations of the inactive and activated MOR structures as seen from the extracellular side, with the transmembrane helix 1 (TM1) depicted in blue TM2 in red, TM3 in dark gray, TM4 in orange, TM5 in yellow, TM6 in green, and TM7 in magenta.

interface are depicted in Supplementary Figs 4 and 5. Receptor domains participating in the interface identified at this stage were used to label their corresponding macrostates in the kinetic networks presented in Figs 2 and 3. Specifically, any receptor region with more than three contacts is listed as part of the interface name and regions with less than three contacts are presented in parentheses.

For both active and inactive MOR structures, we observe that each component (the asymmetric components $C_{0\pi}$ and $C_{\pi0}$, and the two symmetric components, C_{00} and $C_{\pi\pi}$) is populated with dimeric microstates. While some of the dimeric interfaces are common to the active and the inactive receptors, others are specific to one conformation only, as can be seen in Figs 2 and 3. Within the asymmetric components (red clusters in Figs 2 and 3), both MOR structures form TM5,6/TM1,2,H8 and TM5/TM1,2,H8 interfaces whereas interfaces involving TM4 (e.g., TM4,5/TM1,2,H8) appear to be unique to the inactive MOR. Out of the two symmetric components, the $C_{\pi\pi}$ component (yellow clusters in Figs 2 and 3) yields distinct interfaces for the inactive and activated MOR structures. While the active MOR structure is only able to form dimeric TM5/TM5 interfaces, the two dimeric macrostates that are found in the $C_{\pi\pi}$ component of inactive MOR exhibit TM5,6/TM5,6 interfaces that are similar to those seen in the MOR inactive crystal structure (see Supplementary Figs 4 and 5). Finally, simulations of
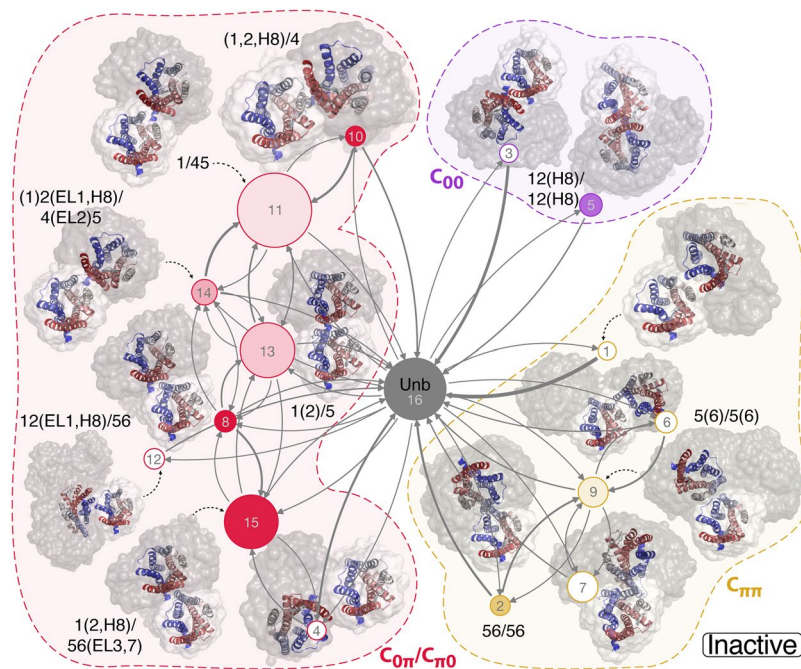
Figure 2. Kinetic networks of inactive MOR macrostates within asymmetric and symmetric components. Asymmetric components $C_{0\pi}$ ($\alpha_1 \approx 0$, $\alpha_2 \approx \pi$) and $C_{\pi 0}$ ($\alpha_1 \approx \pi$, $\alpha_2 \approx 0$) are grouped together and shown in red while symmetric components $C_{\pi\pi}$ ($\alpha_1 \approx \pi$, $\alpha_2 \approx \pi$) and C_{00} ($\alpha_1 \approx 0$, $\alpha_2 \approx 0$) are shown in yellow and purple, respectively. Each node represents a macrostate consisting of a group of PCCA clusters that are spatially and kinetically linked, and whose area is proportional to the probability of the state. The nodes are shaded according to the fraction of dimeric microstates (i.e., those characterized by a dimer population larger than 90%), with a solid color representing a fraction of 100%, and white signifying a fraction of 0%. These macrostates are further labeled according to the interfaces obtained from the contact map analysis of the dimeric microstates. The arrow thickness is proportional to the transition probability between the connected macrostates, ranging between $\sim 3 \times 10^{-5}$ (thinnest line) and 0.04 (thickest line). Representative configurations of the dimeric microstates are shown in cartoon models with the inter-protomeric variability within each macrostate captured by 100 randomly selected frames from that macrostate shown as gray clouds.

both the inactive and activated MOR systems yield dimers exhibiting the same TM1,2,H8/TM1,2,H8 interface (C_{00} component, purple clusters in Figs 2 and 3) seen in the inactive crystal structures of MOR and KOR.

To elucidate which residues are the most relevant to the dimerization process, we calculated the average number of contacts formed by each residue over the complete set of microstates. The residues involved in an average number of contacts larger than 0.2 are listed in Supplementary Table 2 and their spatial distribution across the inactive or active MOR structures is shown in Supplementary Fig. 6. In the inactive MOR system, these residues belong to TM1, TM2, TM5, and H8, as well as IL2, IL3, EL1, and EL2, whereas in the active MOR system additional TM6 residues, but not residues in IL2 and EL2, are among those with the highest average numbers of contacts.

For the active MOR system, the three residues with the highest average contact numbers belong to H8, and they are: F350, C351, and I352. While F350 and C351 have significant, albeit slightly lower, average contact numbers in the inactive receptor, I352 has the highest average contact number in this receptor state. In TM1, the residues with the highest average contacts formed are A73$^{1.37}$, S76$^{1.40}$, and I69$^{1.33}$ for the inactive system, and I69$^{1.33}$, V80$^{1.44}$, and M72$^{1.36}$ for the active system (amino acid numbers correspond to the mouse MOR sequence while superscripts refer to the Ballesteros-Weinstein numbering scheme[24]). In TM2, V126$^{2.62}$, L129$^{2.65}$, and M130$^{2.66}$ present the largest average number of contacts formed in both the inactive and active structures. In TM5, the residues most likely to form contacts are I238$^{5.44}$, M243$^{5.49}$, and Y227$^{5.33}$ in the inactive system, or L257$^{5.63}$, R258$^{5.64}$, and Y227$^{5.33}$ in the active receptor. In IL3, S261 and K260 can form more than 0.2 contacts on average in the active system, while in the inactive system only K260 presents an average contact number above this threshold. In the active MOR system, P134 and F135 in EL1 exhibit high contact numbers, while in the inactive receptor, only P134 is above the cut-off of 0.2 contacts. Unique to the TM6 of the active MOR system, residues I298$^{6.53}$, I302$^{6.57}$, and L305$^{6.60}$ are most likely to form inter-protomer contacts with averages above 0.2 contacts. Exclusive to the inactive system, K174 and H171 in IL2, and H223 and P224 in EL2 form above 0.2 contacts on average.

Since the average number of inter-protomer contacts formed by each of the aforementioned residues does not provide information about which two residues are in contact, we also calculated the probability of occurrence for each contact over all microstates. Supplementary Table 3 lists all inter-protomer contacts with probabilities larger than 0.1 calculated for the inactive and active MOR systems. For the active system, the most likely contacts are formed between H8 and TM5 or IL3 with probabilities over 0.25, followed by contacts between TM6 and TM1 or TM2 with probabilities between 0.13 and 0.22. Both groups of contacts fall within the asymmetric component, $C_{0\pi}$ (red clusters in Fig. 3), which is the most populated component. Specifically, the contacts in dimers of
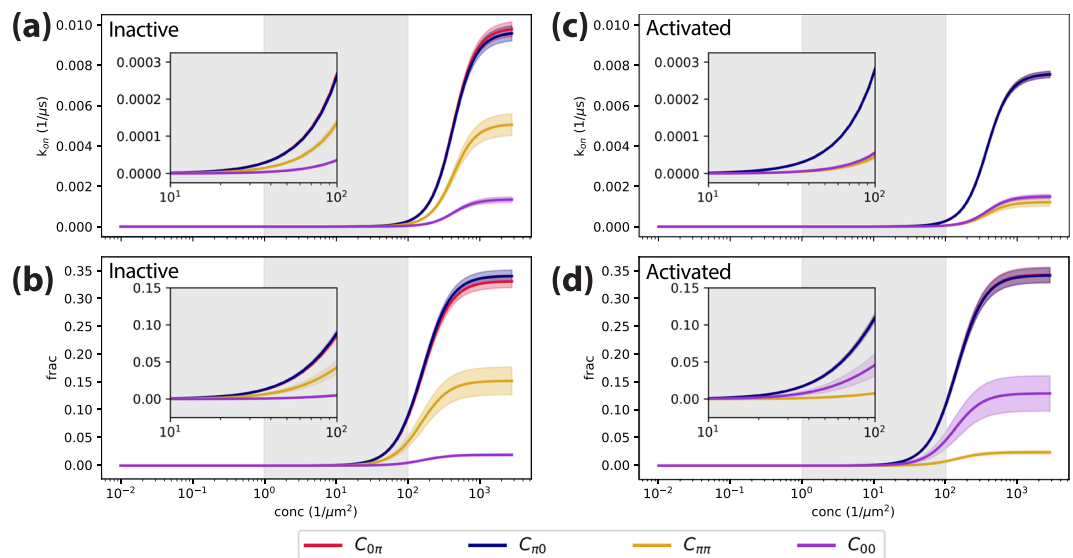
Figure 3. Kinetic networks of active MOR macrostates within asymmetric and symmetric components. Asymmetric components $C_{0\pi}$ ($\alpha_1 \approx 0$, $\alpha_2 \approx \pi$) and $C_{\pi 0}$ ($\alpha_1 \approx \pi$, $\alpha_2 \approx 0$) are grouped together and shown in red while symmetric components $C_{\pi\pi}$ ($\alpha_1 \approx \pi$, $\alpha_2 \approx \pi$) and C_{00} ($\alpha_1 \approx 0$, $\alpha_2 \approx 0$) are shown in yellow and purple, respectively. Two additional macrostates that do not belong to any component are shown as green and blue circles, respectively. Each node represents a macrostate consisting of a group of PCCA clusters that are spatially and kinetically linked, and whose area is proportional to the probability of the state. The nodes are shaded according to the fraction of dimeric microstates (i.e., those characterized by a dimer population larger than 90%), with a solid color representing a fraction of 100% and white a fraction of 0%. These macrostates are further labeled according to the interfaces obtained from the contact map analysis of the dimeric microstates. The arrow thickness is proportional to the transition probability between the connected macrostates, ranging between 3.7×10^{-5} (thinnest line) and ~ 0.08 (thickest line). Representative configurations of the dimeric microstates are shown in cartoon models with the inter-protomeric variability within each macrostate captured by 100 randomly selected frames from that macrostate shown as gray clouds.

active MOR with the highest probabilities are between IL3 and H8 (i.e., S261(IL3)-I352/F350/C351/E349(H8), and K260(IL3)-I352/C351/F350(H8)), and between TM5 and H8 (i.e., L257[5.63]-C351/F350/I352(H8) and R258[5.64]-F350/C351(H8)). For the inactive system, the most likely contacts also belong to the asymmetric component, $C_{0\pi}$ (red clusters in Fig. 2), and those with highest probability are: K174(IL2)-I352(H8), M130[2.66]-Y227[5.33], and A73[1.37]-I238[5.44].

Kinetic models of dimerization of inactive or activated MOR receptors.

A k-means discretization with ~ 1500 centers and a lag time of 50 ns was used to build a MSM from TRAM. In order to validate this MSM, we checked for the convergence of the implied timescales at lag times of 20 ns, 50 ns, and 100 ns (see Supplementary Fig. 4). Specifically, timescale distributions at each of these lag times were calculated by bootstrap sampling over the unbiased simulations only. Additionally, the Chapman-Kolmogorov test, which compares the evolution of the probability calculated with the TRAM kinetic model with the one from simulations, was applied at the macrostate level for all transitions from the unbound macrostate to the bound macrostates (Supplementary Fig. 8) to further assess the reliability of the Markovian assumption.

To better differentiate the MOR dimer interfaces (shaded nodes in Figs 2 and 3), we chose to study the kinetics of the system at the macrostate instead of the microstate level with each macrostate representing groups of spatially and kinetically linked PCCA clusters of microstates. A reduced transition matrix was built for this simplified grouping using the validated approach described in[25] (see Methods section).

The kinetic networks of the macrostates for the inactive and activated MOR systems are shown in Figs 2 and 3, respectively. The macrostates have all been numbered and those containing dimers have been assigned a dimeric interface label that depends on the number of contacts formed. Transition path analysis was then used to obtain the partitioning of the total flux between the monomeric state of MOR and macrostates that contain dimeric interfaces. As shown by this partitioning (see Supplementary Tables 4 and 5 for inactive and active MOR, respectively), while some dimeric interfaces form directly from the unbound state (macrostates #16 and #15 for inactive and active MOR, respectively), others transition through multiple intermediates, suggesting a different dimerization mechanism depending on the interface formed. For instance, in the inactive system, the TM1,2,(H8)/TM1,2,(H8) interface forms directly from the monomeric state with a probability of 94% (see Supplementary Table 4) and only 5% through a loosely bound intermediate (indicated as cluster #3 in Fig. 2). On the other hand, interfaces involving helices TM4, TM5, or TM6, form through metastable intermediates with a higher probability. For instance, while the TM5,6/TM5,6 complex (C_{00}) forms directly with a probability of 64% (see Supplementary

Figure 4. Concentration dependence of the k_{on} values (panels a and c for the inactive and active MOR, respectively) and dimeric fractions (panels b and d for the inactive and active MOR, respectively). The gray regions indicate the typical concentration ranges commonly used in experiments. The insets refer to concentration values between 10 and 100 μm^{-2}.

Table 4), 25% of the flux goes through the TM5,(6)/TM5,(6) state and 7% visits both the TM5,(6)/TM5,(6) state and the loosely bound state #7 in Fig. 2. Notably, states in the asymmetric components have the most heterogeneous dynamics with interfaces between TM1,2 and TM4,5, or TM5,6 freely interchanging in inactive MOR (see Supplementary Table 4).

While certain interfaces form similarly (directly or indirectly) in the inactive and active MOR systems, others exhibit a different dynamic behavior in the two systems. For instance, while the TM1,2,(H8)/TM1,2,(H8) dimer assembles directly in the inactive MOR, it forms through an intermediate state (#11 in Fig. 3) with a probability of 99% (Supplementary Table 5) in the active MOR. On the other hand, interfaces involving helix TM5 (in the symmetric $C_{\pi\pi}$ and the asymmetric components) mostly occur through metastable intermediates in both inactive and active MOR systems (see Supplementary Table 5).

Concentration dependence of the kinetic rates and fractions of MOR homodimers. Since the concentrations used in the simulations reported in this work are much larger than most experimental concentrations (specifically, $1.37 \times 10^4 \mu m^{-2}$ compared to the typical experimental range of $1–100 \mu m^{-2}$), we calculated the protein concentration dependence of MOR dimerization kinetics. Briefly (see *Methods* for further details), we describe the kinetics as a two-step reaction:

$$A + B \underset{k_{+1}}{\overset{k_{-1}}{\rightleftharpoons}} A \cdot B \underset{k_{ij}}{\overset{k_{ji}}{\rightleftharpoons}} AB$$
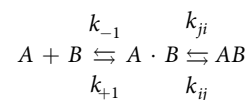
where $A + B$ is the fully dissociated protomeric state, $A \bullet B$ is the encounter complex, AB is the proper complex, and k_{ji} and k_{ij} stand for the rates obtained from the TRAM transition matrix. k_{-1} and k_{+1} are second- and first-order diffusive rates which depend on the protomers' relative 2D diffusion constant D and the 2D concentration c as described in[26]. The overall dimerization rate is $k_{on} = \gamma k_{+1}$ where the capture probability γ is the probability to form an associated state AB from the encounter complex, instead of dissociating to $A + B$[27,28].

In order to compare the kinetics of the main components, the TRAM transition matrix was simplified, further aggregating the macrostates into the C_{00}, $C_{\pi\pi}$, $C_{0\pi}$ and $C_{\pi0}$ components. The reduced transition matrix was validated by comparing its implied timescales to the timescales of the full transition matrix of the TRAM model (see Supplementary Fig. 9). The rates k_{ji} and k_{ij} were calculated for the four components from the reduced transition matrix.

The concentration dependence of the association rates and fractions of homodimers of activated or inactive MOR structures, obtained using a value of the diffusion constant of 0.1 $\mu m^2/s$ (see Discussion), are shown in Fig. 4. As shown in this figure, for concentrations lower than $10^2 \mu m^{-2}$, $k_{on} \sim 0$, suggesting that MOR dimerization is negligible at these concentrations. However, at concentrations above $10^3 \mu m^{-2}$, which include the concentrations used for the simulations presented herein, the conversion of the encounter complex $A \bullet B$ to the dimeric configurations becomes the rate-limiting step, i.e., the dimerization becomes reaction-limited, and the k_{on} values converge to different values depending on the dimeric component of the inactive or activated MOR structure. Specifically, the k_{on} values of the asymmetric components ($C_{0\pi}$ and $C_{\pi0}$) converge to $\sim 10 \pm 0.4\,ms^{-1}$ and $\sim 8.0 \pm 0.2\,ms^{-1}$ for the inactive and activated MOR structures, respectively. These values are significantly

Kinetic Components	Inactive MOR		Active MOR	
	k_{on} ($\mu s^{-1} \mu m^{-2}$)	T_{off} (μs)	k_{on} ($\mu s^{-1} \mu m^{-2}$)	T_{off} (μs)
$C_{0\pi}$	$3 \times 10^{-6} \pm 6 \times 10^{-8}$	210 ± 20	$3 \times 10^{-6} \pm 5 \times 10^{-8}$	280 ± 20
$C_{\pi 0}$	$3 \times 10^{-6} \pm 6 \times 10^{-8}$	220 ± 20	$3 \times 10^{-6} \pm 5 \times 10^{-8}$	280 ± 20
$C_{\pi\pi}$	$1 \times 10^{-6} \pm 1 \times 10^{-7}$	130 ± 30	$5 \times 10^{-7} \pm 8 \times 10^{-8}$	100 ± 20
C_{00}	$4 \times 10^{-7} \pm 3 \times 10^{-8}$	90 ± 10	$6 \times 10^{-7} \pm 4 \times 10^{-8}$	380 ± 60

Table 1. The k_{on} values (calculated for protein densities of $10^2 \mu m^{-2}$) and dimer lifetimes of the four kinetic components, $C_{0\pi}$, $C_{\pi 0}$, $C_{\pi\pi}$, and C_{00}, derived for the inactive and active MOR systems.
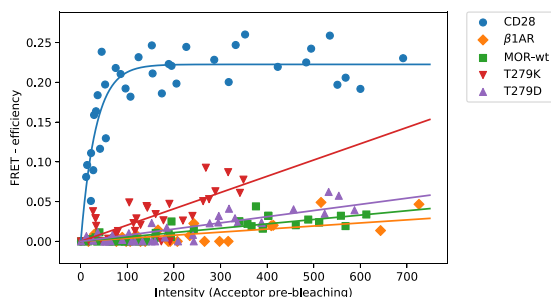
Figure 5. FRET-efficiencies of SNAP-labeled MOR constructs plotted as a function of acceptor intensity (i.e., expression level) before photobleaching. CD28 (blue color) is used as reference dimer whereas β1AR (orange color) is used as reference monomer. Wild-type MOR (green color), as well as inactive MOR mutant T279D (purple color) and constitutively active MOR mutant T279K (red color), all show an unspecific increase in FRET efficiencies.

larger than the k_{on} values of the symmetric components, $C_{\pi\pi}$ and C_{00}, which converge to ~5.0 ± 0.5 ms^{-1} and ~1.0 ± 0.2 ms^{-1}, respectively, for the inactive MOR system and ~1.0 ± 0.2 ms^{-1} and ~1.0 ± 0.1 ms^{-1}, respectively, for the activated MOR structures.

The bound states of both the inactive and activated MOR systems reach a total fraction of ~0.85 ± 0.04 considering the sum of all four components. Notably, the faster rate at which asymmetric complexes contributes to the larger fraction of dimers in the two $C_{0\pi}$ and $C_{\pi 0}$ components. In the inactive MOR, these components reach a cumulative fraction of ~0.7 ± 0.02, while the symmetric $C_{\pi\pi}$ and C_{00} components are at ~0.15 ± 0.02 and ~0.02 ± 0.002, respectively, as shown in Fig. 4a. Similar to the inactive MOR system, the asymmetric components have a cumulative fraction of ~0.7 ± 0.02 in the activated receptor whereas opposite values (0.02 ± 0.003 for $C_{\pi\pi}$ and ~0.13 ± 0.03 for C_{00}) are obtained for the symmetric components (see Fig. 4b).

As the concentration decreases to the upper limit of the typical experimental range at $10^2 \mu m^{-2}$, the k_{on} values reach ~0.3 ms^{-1} for the asymmetric components ($C_{0\pi}$ or $C_{0\pi}$) of both the inactive and activated MOR structures with dimer fractions of ~9% and ~12%, respectively. The $C_{\pi\pi}$ component of the inactive MOR dimers reaches a k_{on} value of 0.1 ms^{-1} with a dimer fraction of ~0.05, while the C_{00} conformations become negligibly small. For the activated MOR dimers, k_{on} values for both symmetric components are ~0.05 ms^{-1} at this concentration reaching dimer fractions below 0.05. The k_{on} calculated for receptor concentrations of $10^2 \mu m^{-2}$ and dimer lifetime values (see Methods section for details) are presented in Table 1 for each main component of the inactive and active MOR systems. The lifetimes of the asymmetric components are ~220 μs and ~280 μs for the inactive and active MOR systems, respectively. Among the symmetric components, $C_{\pi\pi}$ has a lifetime of 130 μs and 100 μs for the inactive and active MOR systems, respectively, while C_{00} exhibits dimer lifetimes of 90 μs and 380 μs for the same systems.

MOR dimerization assessed by FRET acceptor photobleaching experiments. To experimentally assess the degree of MOR dimerization in a concentration-dependent manner, we conducted acceptor photobleaching experiments to determine FRET efficiencies between mouse MOR pairs (see Supplementary Fig. 10). Specifically, SNAP-tagged MOR constructs were selectively labeled with an equimolar combination of SNAP-Surface 549 (SS549) and SNAP-Surface Alexa Fluor 647 (SS647) on the outer cell membrane. This FRET pair has a sufficient spectral overlap, which leads to an estimated Förster-distance[29] of 74 Å. The donor (SS549) is independently photostable and is not significantly affected during acceptor photobleaching (Supplementary Fig. 11). A construct coding for a dimeric cell-surface receptor[30], CD28, was used as a reference dimer. As expected and shown in Fig. 5, this construct shows expression dependence, fast increase, and saturation of FRET efficiencies, with a $k_d = 22.08 \pm 6.2$ relative fluorescent units (RFU), and a FRET efficiency max = $22.25 \pm 0.6\%$. Based on previously published single-molecule experiments[15] suggesting a low degree of di-/oligomerization (monomers or dimers) for the β1-adrenergic receptor (β1AR), we used this system as a reference monomer in the experiments reported here. As expected, β1AR showed a quite low linear (unspecific) increase of FRET efficiencies (Fig. 5) with the highest obtained FRET efficiency of 4.7% at an expression level of 726 RFUs (slope = 0.004%/RFU). Experiments with MOR do not deviate significantly from the β1AR example of a monomer (Fig. 5). The system does not saturate and its highest observed FRET efficiency is 3.4% at 613 RFUs (slope = 0.005%/RFU).

To further assess the dependence of the active or inactive state of MOR on its di-/oligomerization, we performed experiments with two MOR mutants: T279K and T279D (Fig. 5). In the inactive MOR crystal structure, the T279[6.34] residue forms a hydrogen bond with R165[3.50], which replaces the classical R[3.50]-D/E[6.30] salt bridge that stabilizes the inactive state of rhodopsin-like GPCRs. Previous studies[31] show that the T279K mutant is a constitutively active receptor, presumably due to an electrostatic repulsion with R165[3.50], whereas the T279D mutant is an inactive receptor possibly due to its ability to form a salt-bridge with R165[3.50]. For both mutants, the FRET efficiency also increased linearly (unspecifically) with expression levels as shown in Fig. 5. The inactive T279D mutant exhibited a slightly increased slope (0.007%/RFU) in comparison to the wild-type MOR whereas the constitutively active T279K mutant showed a more significant, but still unspecific, increase in FRET efficiencies with a slope of 0.02%/RFU. For a comparison of the significance values see Supplementary Table 6.

Discussion

While it is well established that the MOR is functional in its monomeric form[32], a conclusion has not been reached as to whether it is mostly present as monomers, dimers, or higher order oligomers under physiological conditions *in vivo*. In this work, we report, for the first time, a multi-ensemble Markov model describing the kinetics underlying MOR association and dissociation, and providing predictions of MOR homodimerization rates and fractions at physiological concentrations.

A few cautionary points should be kept in mind while considering the results of this study. Firstly, unlike what happens under physiological conditions, the coarse-grained MARTINI force field used here to simulate MOR prevents any major conformational rearrangement during receptor association. Notwithstanding this limitation, the computational approach presented here provides inferences of MOR dimerization that are verified experimentally through measurements of FRET efficiency. Secondly, since an unfeasible amount of simulation time would be required to observe relevant events such as MOR dimerization at physiological concentrations, the simulations presented here were performed at much higher receptor densities (i.e., $\sim 1.37 \times 10^4 \mu m^{-2}$) than those afforded under physiological conditions[33,34] or commonly used in fluorescence resonance energy transfer (FRET) or spatial brightness experiments (e.g., $0.1 \mu m^{-2}–100 \mu m^{-2}$; see for instance[35–37]). Furthermore, the natural cellular environment of MOR is far more crowded and complex than its counterpart in our simulations, which is composed of POPC and CHOL molecules only. Thus, it is not surprising that the diffusion coefficients obtained from the simulations presented here are several orders of magnitude larger (i.e., $\sim 10 \mu m^2/s$) than those observed in experiments (i.e., $0.02 \mu m^2/s$ to $0.2 \mu m^2/s$)[13,15], where the MOR movement within the membrane is impaired. It is also important to point out that the diffusion coefficients obtained from the simulations reported here are calculated using portions of the trajectories for which the distance between protomers range between 8 nm to 11 nm. While these values are appropriate given the simulation box size, the diffusive behavior of the protomers is expected to be different at lower concentrations where the distance between them is much larger.

The above considerations prompted us to use monomer diffusion coefficients obtained from experiments rather than from simulations in the diffusive step of the reaction-diffusion model employed to study the concentration-dependence of kinetic rates and fractions of MOR homodimers at physiological concentrations. Using a physiologically relevant diffusion coefficient of $0.1 \mu m^2/s$, we obtain negligible dimeric fractions of MOR at most experimental concentrations, including physiological concentrations. These homodimers are characterized by relatively short lifetimes that depend on the interface and do not exceed ~ 0.3 ms. Notably, at the lower end of a typical experimental concentration range (i.e., $0.1–10/\mu m^2$), MOR dimer populations become negligible even for much larger diffusion coefficients such as those obtained from simulations. This was confirmed by linear increases in FRET efficiencies by acceptor photobleaching, which argued in favor of a very low degree of homomerization for MOR.

Compared to other GPCRs whose dimerization has been studied experimentally, our results suggest smaller populations of MOR homodimers with much shorter lifetimes. For instance, unlike the β_1-adrenergic receptor, which had been reported to exhibit a low degree of di-/oligomerization by single-molecule experiments[15], the $\beta_2 AR$ had been shown to form 60% of dimer and higher order oligomer populations, at concentrations of $\sim 0.25/\mu m^2$ and with lifetimes on the order of a few seconds[15]. Similarly, the M_1 muscarinic receptor had been shown to form $\sim 30\%$ dimers at concentrations of $\sim 1.8/\mu m^2$ with lifetimes on the second time-scale[13]. Our computations suggest that MOR dimer populations are much lower at these concentrations, suggesting a weaker propensity for MOR dimerization compared to $\beta_2 AR$ and M_1 muscarinic receptor. Additionally, in the case of the $\beta_2 AR$, agonist binding showed no significant effect on dimerization[15] while our results suggest a small increase in dimer fractions upon activation at concentrations near $10^2/\mu m^2$. However, this difference becomes irrelevant at smaller and more physiological concentrations, as confirmed by FRET acceptor photobleaching measurements of the inactive T279D and constitutively active T279K mutants of MOR.

Notably, the differences observed in the dimerization kinetics of the inactive and activated MOR systems are accompanied by distinct dimer interfaces. The available inactive and active MOR crystal structures suggest that while dimerization can occur at the symmetric interface of TM1,2,H8/TM1,2,H8 for both systems, the outward movement of the TM6 in the MOR activated crystal structure would prevent the TM5,6/TM5,6 interface from occurring[38]. The interfaces formed in the simulations presented here corroborate these conjectures and furthermore suggest that for the active MOR structure a TM5/TM5 interface might also be possible. Unlike our earlier simulation results in which we found the TM4/TM4 interface to form with high probability between active MOR protomers, we only observed here the TM5/TM5 interface[16]. In contrast, the TM1,2,H8/TM1,2,H8 interface, which has been also inferred by the inactive MOR and KOR crystal structures[38,39], also formed in our previous simulations, albeit with much lower probability.

Compared to the symmetric interfaces, our results suggest that asymmetric interfaces of MOR dimers have significantly higher propensities to form, in line with our earlier results[16]. While asymmetric dimers of the inactive

MOR involving the TM1,2,H8/TM4, TM1,2,H8/TM4,5 and TM1,2,H8/TM5 interfaces were heavily favored in both earlier and present simulations, the TM1,2,H8/TM5,6 interface was observed only in the simulations reported here. Furthermore, the small population of dimer interfaces involving TM7 observed in the earlier study are not observed in the current simulations, nor is the TM4/TM5 interface. On the other hand, the active MOR structures also have high propensities for the asymmetric dimer interfaces of TM1,2,H8/TM5,6 and TM1,2,H8/TM5. The TM1,H8/TM6,7 interface observed in CXCR4[40] and the TM4/TM7 interface of CCR5[41] are absent in the simulations of both inactive and active MOR despite having been reported to form ~8% of the dimer population in our earlier results.

Predictions of residues involved in dimeric contacts suggest several potential mutations that can significantly affect the dimerization process. For both the inactive and active systems, we found that H8 includes some of the residues with the largest average contact numbers as well as being involved in some of the most likely contacts. For instance, I352 in H8 forms an average of 1.48 and 1.3 contacts in the active and inactive systems, respectively, while participating in the top 3 most likely contacts that appear in both systems, between IL3 and H8 in the active MOR system and between IL2 and H8 in inactive MOR. Similar to I352, F350 and C351 in H8 are also involved in putative dimeric interfaces of MOR, in addition to playing a significant role in TM5 and H8 contacts in the active MOR system. This analysis also draws attention to $M130^{2.66}$ as a potential point of mutation, since it plays an important role in TM5-TM2 and TM2-EL2 contacts in the inactive MOR system, in addition to TM6-TM2 contacts in the active MOR. Notably, $L129^{2.65}$ is likely to have an effect similar to $M130^{2.66}$. Among the TM1 residues involved in dimerization, $A73^{1.37}$ plays an important role in TM5-TM1 contacts in the inactive MOR system, while $I69^{1.33}$ participates in prominent TM1-TM6 contacts. Combined, the appropriate mutations at these sites might enhance the dimerization propensities of the MOR system by modulating contacts that are crucial to the formation of asymmetric interfaces, which are suggested to be the most likely candidates for dimerization according to our results.

In conclusion, our study presents a computational framework that elucidates, for the first time, the dimerization kinetics of MOR by combining unbiased and biased coarse-grained simulations coupled with multi-ensemble Markov modeling and a reaction-diffusion model, which provides details both at the molecular and the ensemble scale. Our computational results suggest that the MOR does not dimerize at concentrations used in most experimental methods in contrast to other prototypic GPCRs, such as the β_2-adrenergic receptors and the M_1 muscarinic receptor. These results are supported by FRET acceptor photobleaching experiments, which confirm that MOR does not exhibit concentration-dependent specific dimerization over a wide range of expression levels. In striking contrast to CD28, MOR exhibits a linear (unspecific) increase in FRET efficiencies rather similar to the largely monomeric $\beta1AR$[15]. Moreover, FRET measurements performed on both the inactive T279D mutant receptor and the constitutively active T279K mutant suggested that MOR dimeric fractions might be independent on a specific conformational state of the receptor. Notably, the slope of FRET efficiencies was significantly higher for the constitutively active T279K mutant compared to the inactive T279D mutant and wild-type MOR. This might be due to either an enhanced propensity of this mutant receptor to form true complexes from encounter complexes or to its enhanced internalization, which leads to non-specific aggregation through clustering in clathrin coated pits and vesicles. In summary, our study provides a quantitative explanation for some conflicting experimental results on GPCR oligomerization.

Methods

System set-up.
All non-receptor atoms of the crystal structures of inactive (PDB ID: 4DKL[38]) and activated (PDB: 5C1M[42]) MOR were removed, including co-crystallized ligands, the T4L fusion protein in the inactive structure, and the G protein mimetic nanobody in the activated structure. The missing intracellular loop 3 (ICL3) in the inactive MOR structure was built by homology modeling using MODELLER[43] and the high-resolution crystal structure of DOR[44] as a template. The N-terminal segment of the activated structure was removed and the remaining residues of H8 added so that each structure consisted of sequence residues 65 to 352. The *martinize* python script was used to coarse grain the structures within the MARTINI v2.2 force field[19–22]. A modified version of the elastic network was applied to maintain the receptor tertiary structure[45,46]. Specifically, a harmonic force was applied to all backbone (BB) bead pairs with relative distance less than 0.9 nm. A force constant of 1000 kJ mol^{-1} nm^{-2} was used when both beads belonged to regions with helical secondary structure while 250 kJ mol^{-1} nm^{-2} was used when either bead was in a loop region. Two inactive or activated MOR protomers were randomly rotated around the z-axis and placed in a $15 \times 15 \times 11$ nm^3 box along the diagonal such that the COM of the receptors were approximately 6.25 nm apart. The receptors were then embedded in a POPC/10% CHOL membrane constructed using the *insane* python script[47]. Each system contained approximately 576 POPC and 64 CHOL molecules. A total of 704 dimer/membrane complexes were generated: 352 for the inactive MOR and 352 for the activated receptor. Each complex was solvated and water and ions were added to neutralize the charge.

Simulations in the unbiased ensemble.
Up to 10,000 steps of energy minimization were performed on each system followed by 500 ns of equilibration keeping position restraints on the BB beads of the receptor. The position restraints were slowly removed by performing four 50 ns runs with gradually decreasing force constants (k = 500, 100, 50, 10 kJ/mol/nm^2). Each of the 704 receptor/membrane complexes was then run for at least 5 μs. All simulations were performed using Gromacs 5.0[48] and in the NPT ensemble at 310 K and 1 bar. Constant temperature and pressure were maintained using the velocity-rescale[49] and Berendsen[50] algorithms, respectively. The Coulombic interactions decayed smoothly to zero between 0 and 1.2 nm while the van der Waals interactions decayed to zero between 0.9 and 1.2 nm.

Collective variables and preliminary kinetic model.
The results were first analyzed by describing the system with a three-dimensional order parameter comprising (i) the distance *d* between the protomers' COM, (ii) the angle α_1 formed by the COM of one protomer, the COM of the second protomer, and the COM of its transmembrane (TM) helix 1, and (iii) the corresponding angle α_2 obtained swapping the two protomers. To efficiently

account for angle periodicity, the four-dimensional order-parameter (d$sin\alpha_1$, d$cos\alpha_1$, d$sin\alpha_2$, d$cos\alpha_2$) was used in the calculations. An initial kinetic model was set up using this order parameter and discretizing the trajectories into 6000 microstates using k-means clustering for all d < ½ L, and using a lag-time of 50 ns. Configurations with the protomers at a distance d ≥ ½ L were assigned to a separate microstate. To improve statistics and account for the interchangeability of the protomers in a homomeric system, each unbiased trajectory was duplicated with swapped values of the angles α_1 and α_2, corresponding to exchanging the two proteins. The transition matrix was estimated by maximum likelihood from the count matrix for both the active and inactive systems, and subsequently coarse-grained using PCCA+ to obtain 37 clusters for both the inactive or active systems, respectively. The number of clusters was chosen as the smallest value with at least one macrostate including only microstates with d > 4.2 nm and with a flat angular distribution, which corresponded to the unbound state. One unique cluster with these properties was identified for both active and inactive systems. Removing this macrostate from this preliminary transition probability matrix resulted in five main disconnected sets of clusters in the case of the inactive MOR system and six sets for the activated MOR. Once the system enters in one of such disconnected clusters, it cannot transition into another one without first passing through the unbound cluster. K-means clustering, transition matrix estimation and spectral clustering were performed using pyemma (version 2.3)[51] python functions.

Simulation details for umbrella sampling runs. To aid convergence of the free-energy and kinetic estimation, umbrella sampling runs were performed restraining the system around selected k-mean centers. Specifically, to select the centers for biased sampling, Transition Path Theory (TPT) from the microstate with d ≥ ½L to the most probable microstate in each of the disconnected components including bound structures, was applied to the preliminary kinetic model described above, and the k-mean centers of microstates with a committor probability p > 10% were selected as umbrella positions. Overall, 314 and 323 umbrella centers were selected for the active and inactive receptor systems, respectively. The position of the centers of umbrella sampling runs is displayed in Supplementary Fig. 12. Umbrella sampling simulations were run for 0.3 μs, using the same parameters of the unbiased simulations, and applying harmonic restraints with elastic constants of $k_d = 250$ kJ/mol/nm^2 and $k_\alpha = 100$ kJ/mol, respectively, for the distance and angle components of the order parameter. For a small set of runs the elastic constants were reduced to $k_d = 100$ kJ/mol/nm^2 and $k_\alpha = 80$ kJ/mol, respectively. Restraints were applied using the Plumed 2.1 plugin[52].

Markov state modeling and multi-ensemble Markov models. The data from the unbiased and biased simulations were combined using the transition-based reweighting analysis method (TRAM)[23]. To reduce the number of microstates in the calculation, microstates in the unbound PCCA+ component of the preliminary kinetic clustering that did not show transitions out of the component in the 50 ns-lag count matrix were resampled, reducing their number from 5214 to 703 and from 5412 to 911 for the inactive and active MOR systems, respectively. The resulting k-means clustering comprised 1490 microstates in total for both the inactive and active receptor systems. The final multi-ensemble model was estimated using the TRAM algorithm and a convergence threshold based on the maximal free energy change of 10^{-9}. Further decreasing the convergence threshold to 10^{-12} resulted in transition matrices without appreciable differences from those with relaxed convergence criteria. The implied timescales from the transition matrix for the unbiased thermodynamic state were calculated for multiple lag-times and their convergence checked. Errors on timescales and rates were obtained by bootstrap. Specifically, 10 bootstrap samples were extracted from the data by sampling of the unbiased trajectories and the TRAM estimator was used to obtain a distribution of transfer matrices. Best estimate and confidence intervals are reported using the mean and the 5% and 95% percentiles for the microstate free-energies and timescales.

Coarse-graining of the transfer matrix. To obtain kinetic information at a more coarse-grained level, the final TRAM transition matrix was clustered using the PCCA+ algorithm and 51 or 35 clusters for the inactive and activated MOR structures, respectively. The number of clusters to be used was decided using the same criteria applied in the case of the unbiased MSM. Furthermore, the approach used to find disconnected components in the unbiased MSM was repeated for the TRAM transition matrices, resulting in four main disconnected components for both systems and two additional components of negligible populations for the active system. The four main components correspond to the four possible combinations of the two heavily populated regions of the free energy landscapes centered around values of the angles α_1, $\alpha_2 \approx 0$ and α_1, $\alpha_2 \approx \pi$, respectively. We refer to these groups as the C_{00} component for the TM1,2,H8/TM1,2,H8 region (centered around α_1, $\alpha_2 \approx 0$), the $C_{\pi\pi}$ for the TM4, 5, 6/TM4, 5, 6 region (centered around α_1, $\alpha_2 \approx \pi$), and the $C_{0\pi}$ component for the TM1, 2,H8/TM4, 5, 6 region (centered around $\alpha_1 \approx 0$, $\alpha_2 \approx \pi$ and $\alpha_1 \approx \pi$, $\alpha_2 \approx 0$).

The coarse-grained rates connecting the PCCA components were calculated using the approach described in ref.[25]. Briefly, reduced dynamic models defined on the PCCA states were obtained by imposing the same cross-relaxation times observed in the full microscopic model. The coarse-grained transfer matrix T_{IJ}^{CG} was obtained by solving

$$\sum_{i \in I, \, j \in J} A_{ij}^{-1} \pi_j^{micro} = B_{IJ}^{-1} \pi_J^{CG}$$

(1)

with

$$\begin{cases} A_{ij} = \delta_{ij} + \pi_i^{micro} - T_{ij}^{micro} \\ B_{IJ} = \delta_{IJ} + \pi_I^{CG} - T_{IJ}^{CG} \end{cases}$$

(2)

where π^{micro} and T^{micro} are the microscopic stationary probability and transfer matrix from the TRAM estimation, respectively, and $\pi_I^{CG} = \sum_{i \in I} \pi_i^{micro}$ are the coarse-grained probabilities on the PCCA+ components. This procedure results in a coarse-grained Markov model that best emulates the dynamics of the full model at both short and long times. To validate the resulting approximation, the Chapman-Kolmogorov test was performed comparing the time-behavior of the occupation probability of each of the PCCA components from the unbiased simulations to the same quantity calculated using the full transition matrix and the coarse-grained transition matrix.

Definition of dimeric states and contact map calculation. For each microstate k, the probability $p_k^{(Tot)}(n)$ of observing at least n inter-protomer contacts (defined as backbone beads below 0.8 nm) and the probability $p_k^{(Micro)}(i, j)$ of each individual contact being formed were calculated as averages over the frames of the unbiased trajectories. Microstates with a high probability of having at least 10 inter-protomer contacts, i.e. $p_k^{(Tot)}(n = 10) \geq 0.9$, were defined as dimeric microstates. Contact maps for each macrostate K were obtained by averaging $p_K^{(PCCA,dimeric)}(i, j) = \sum_{k \in K}' \pi_k p_k^{(Micro)}(i, j)$ where π_k is the stationary probability for microstate k from the converged TRAM model and the prime on the sum indicates it is over all dimeric microstates assigned to the macrostate.

Reaction-diffusion model. In order to extract from the simulations inferences for different receptor concentrations, we employ the simple kinetic model

$$A + B \underset{k_{+1}}{\overset{k_{-1}}{\rightleftarrows}} A \cdot B \underset{k_{ij}}{\overset{k_{ji}}{\rightleftarrows}} AB \tag{3}$$

where $A \cdot B$ represents an encounter complex and AB a proper complex, k_{+1} and k_{-1} are a second- and first- order diffusive rates that depend on the proteins' relative 2D diffusion constant D and 2D concentration c, and k_{ij} are the rates obtained from the coarse-graining of the TRAM transition matrix. The effective on-rate for the dimerization reaction is given by $k_{on} = \gamma k_{+1}$ (e.g., see[27]) where γ is the capture probability $\gamma = \sum_j k_{ij} / \left(k_{-1} + \sum_j k_{ij} \right)$, i.e. the probability to form an associated state AB rather than dissociating to A + B, starting from the encounter complex $A \cdot B$. To model the diffusive second-order rate k_{+1}, we use the approach in ref.[26] for finite protein concentration, and obtain

$$k_{+1} = \frac{4\pi D}{\log \frac{\lambda}{r_0}} \tag{4}$$

where $\lambda = (\pi c)^{-1/2}$ is the mean-free path at concentration c and r_0 is the protein distance in the encounter complex. Here, we model the encounter complex with the dissociated state from the PCCA+ analysis obtained from the simulations, and use the mean distance between protomers in this state $\sum_{i \in U} d_i \pi_i$, where π_i are the TRAM steady-state probabilities. For the inactive and active system, we obtain $r_0 = 6.2$ nm and 6.3 nm, respectively. The relative diffusion constant D of the proteins was estimated by fitting the time-dependence of the variance of protein-protein distance (d) distribution at increasing lag times (up to 10 ns) for trajectories starting at d > 80 nm. We obtained similar values D $= 10^{-8} \mu m^2$/ns for the active and inactive receptors, in line with estimates from previous publications with the same force field[53]. Note that this value is not rescaled with the common factor 4 for the Martini forcefield. We calculate the capture probability following the Northrup-Allison-McCammon protocol[28] as:

$$\gamma = \frac{f_{ass}(r_1)}{1 - f_{esc}(r_1) p_{ass}(r_1)} \tag{5}$$

where f_{ass} is the fraction of trajectories that start with the monomers at distance r_0 and enter an associated state (i.e. any PCCA cluster except the dissociated one) before reaching an "escape" distance r_1, and $p_{ass} = 1 - (\log \lambda / r_1)/(\log \lambda / r_0)$ is a factor accounting for the fact that only a fraction of the trajectories that reaches r_1 will eventually escape rather than return to r_0 Values for $f_{ass,esc}$ were obtained from the unbiased simulations for $r_1 = 7$ nm.

To model the concentration dependence of k_{on} we used the approach in ref.[23] to calculate the rates from the combined dissociated states (A + B and $A \cdot B$) to the associated states AB in the kinetic model; following this approach, k_{-1} is obtained from the definition of γ. The k_{off} values are reported as the rates from the associated states to the encounter complex $A \cdot B$. Final times and rates were scaled using a factor 4 to account for faster dynamics in the Martini CG force-field.

Cell-culture. All experiments were performed with transiently transfected HEK 293 cells (Sigma, Munich, Germany). Cells were cultured in DMEM (Dulbecco's modified Eagle's medium) (PAN Biotech, Aidenbach, Germany), supplemented with 4,5 g/L glucose, 2 mM L-glutamin 10% FCS (Biochrome), 100 units/mL penicillin and 0.1 mg/mL streptomycin, at 37 °C, 5% CO2. To split cells, growth medium was removed by aspiration and cells were washed once with 10 ml of phosphate-buffered saline (Sigma), followed by trypsinization for 1–2 minutes in 3-ml of Trypsin 0.05%/EDTA 0.02% (PAN biotech) solution. Cells were routinely tested for mycoplasma infection using PCR mycoplasma test kit (Applichem GmbH, Darmstadt, Germany).

Microscope sample preparation and SNAP-labeling. Cells seeded on 25 mm coverslips were transfected with SNAP-tagged receptor-constructs using Effectene Transfection Reagent (QIAGEN, Hilden, Germany)

according to manufacturer's protocol. 24–36 hours after transfection cells were labeled with SNAP-Surface Dyes 549 and Alexa Fluor 647 (New England Biolabs, Frankfurt am Main, Germany) according to the manufacturer's instructions. Briefly, 24 hours after transfection, the culture medium was exchanged with labeling medium (5 μM SNAP-dye in DMEM) followed by a 30 minute incubation at 37 °C. After incubation cells were washed three times with DMEM and immediately taken for imaging in HBSS-Buffer.

FRET-Acceptor photoBleaching (AB) in confocal microscopy. Prepared samples were placed into an Attofluor™ Cell Chamber (Fisher Scientific GmbH, Schwerte, Germany). Coverslips were washed once with HBSS (Hank's Balanced Salt Solution) and the chamber was filled with 500 mL HBSS. The chamber was mounted onto a Leica SP8 confocal laser-scanning microscope. Cells were imaged using the Leica FRET-AB wizard with a HC PL APO CS2 40x/1.3 numerical aperture oil immersion objective. A 1.5 mW white light laser was set to 85% and a 560 nm laser line was used at 5% power for the donor imaging. For the acceptor imaging a 652 nm laser line at 2% power was used and for the bleaching step increased to 50% over 10 frames. 512 × 512 pixel images of the bottom cell-membrane expressing SNAP-tagged receptor-constructs were acquired with a hybrid detector in standard mode. Emission of donor channel was recorded within 575–640 nm and acceptor channel was acquired between 658 nm–776 nm. The zoom factor was set to 5.5 x resulting in a pixel-size of 0.103 μm and the laser scanning speed was set to 400 Hz. There were at maximum 2 cells taken for analysis per image. FRET efficiencies were calculated with the manufacturer's Wizard tool based on the provided formula (see Supplementary Fig. 10). Potential vesicles close to the cell surface were excluded by the drawing of the region of interest.

Data availability. The data that support the findings of this study are available from the corresponding author upon request.

References
1. Sartania, N., Appelbe, S., Pediani, J. D. & Milligan, G. Agonist occupancy of a single monomeric element is sufficient to cause internalization of the dimeric beta2-adrenoceptor. *Cell Signal* **19**, 1928–1938 (2007).
2. Parenty, G., Appelbe, S. & Milligan, G. CXCR2 chemokine receptor antagonism enhances DOP opioid receptor function via allosteric regulation of the CXCR2-DOP receptor heterodimer. *Biochem J* **412**, 245–256 (2008).
3. Vilardaga, J. P. *et al*. Conformational cross-talk between alpha2A-adrenergic and mu-opioid receptors controls cell signaling. *Nat Chem Biol* **4**, 126–131 (2008).
4. Pin, J. P., Galvez, T. & Prezeau, L. Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. *Pharmacol Ther* **98**, 325–354 (2003).
5. Brock, C. *et al*. Activation of a dimeric metabotropic glutamate receptor by intersubunit rearrangement. *J Biol Chem* **282**, 33000–33008 (2007).
6. Salahpour, A. & Masri, B. Experimental challenge to a 'rigorous' BRET analysis of GPCR oligomerization. *Nat Methods* **4**, 599–600, author reply 601 (2007).
7. Lohse, M. J. G protein-coupled receptors: too many dimers? *Nat Methods* **3**, 972–973 (2006).
8. Chabre, M. & le Maire, M. Monomeric G-protein-coupled receptor as a functional unit. *Biochemistry* **44**, 9395–9403 (2005).
9. James, J. R., Oliveira, M. I., Carmo, A. M., Iaboni, A. & Davis, S. J. A rigorous experimental framework for detecting protein oligomerization using bioluminescence resonance energy transfer. *Nat Methods* **3**, 1001–1006 (2006).
10. Bouvier, M., Heveker, N., Jockers, R., Marullo, S. & Milligan, G. BRET analysis of GPCR oligomerization: newer does not mean better. *Nat Methods* 4, 3–4, author reply 4 (2007).
11. Milligan, G. A day in the life of a G protein-coupled receptor: the contribution to function of G protein-coupled receptor dimerization. *Br J Pharmacol* **153**(Suppl 1), S216–229 (2008).
12. Milligan, G. G protein-coupled receptor hetero-dimerization: contribution to pharmacology and function. *Br J Pharmacol* **158**, 5–14 (2009).
13. Hern, J. A. *et al*. Formation and dissociation of M-1 muscarinic receptor dimers seen by total internal reflection fluorescence imaging of single molecules. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 2693–2698 (2010).
14. Kasai, R. S. *et al*. Full characterization of GPCR monomer-dimer dynamic equilibrium by single molecule imaging. *J Cell Biol* **192**, 463–480 (2011).
15. Calebiro, D. *et al*. Single-molecule analysis of fluorescently labeled G-protein-coupled receptors reveals complexes with distinct dynamics and organization. *Proc Natl Acad Sci USA* **110**, 743–748 (2013).
16. Marino, K., Prada-Gracia, D., Provasi, D. & Filizola, M. Impact of Lipid Composition and Receptor Conformation on the Spatio-Temporal Organization of mu-Opioid Receptors in a Multi-component Plasma Membrane Model. *PLOS Computational Biology* **12** (2016).
17. Periole, X., Knepp, A. M., Sakmar, T. P., Marrink, S. J. & Huber, T. Structural determinants of the supramolecular organization of G protein-coupled receptors in bilayers. *J Am Chem Soc* **134**, 10959–10965 (2012).
18. Provasi, D., Boz, M. B., Johnston, J. M. & Filizola, M. Preferred Supramolecular Organization and Dimer Interfaces of Opioid Receptors from Simulated Self-Association *PLOS Comp. Biol.*, accepted (2015).
19. de Jong, D. H. *et al*. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *Journal of Chemical Theory and Computation* **9**, 687–697 (2013).
20. Marrink, S. J., de Vries, A. H. & Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *Journal of Physical Chemistry B* **108**, 750–760 (2004).
21. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *Journal of Physical Chemistry B* **111**, 7812–7824 (2007).
22. Monticelli, L. *et al*. The MARTINI coarse-grained force field: Extension to proteins. *Journal of Chemical Theory and Computation* **4**, 819–834 (2008).
23. Wu, H., Paul, F., Wehmeyer, C. & Noé, F. Multiensemble Markov models of molecular thermodynamics and kinetics. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E3221–E3230 (2016).
24. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods in Neuroscience* **25**, 366–428 (1995).
25. Hummer, G. & Szabo, A. Optimal Dimensionality Reduction of Multistate Kinetic and Markov-State Models. *The Journal of Physical Chemistry B* **119**, 9029–9037 (2015).
26. Hardt, S. L. Rates of diffusion controlled reactions in one, two and three dimensions. *Biophys Chem* **10**, 239–243 (1979).
27. Shoup, D. & Szabo, A. Role of diffusion in ligand binding to macromolecules and cell-bound receptors. *Biophys J* **40**, 33–39 (1982).
28. Northrup, S. H., Allison, S. A. & McCammon, A. Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *The Journal of Chemical Physics* **80**, 1517 (1984).

29. Spence, M. T. & Johnson I. D. The molecular probes handbook: a guide to fluorescent probes and labeling technologies. *Live Technologies Corporation* 38–39 (2010).
30. Tabor, A. *et al.* Visualization and ligand-induced modulation of dopamine receptor dimerization at the single molecule level. *Sci Rep* **6**, 33233 (2016).
31. Huang, P. *et al.* Functional role of a conserved motif in TM6 of the rat mu opioid receptor: constitutively active and inactive receptors result from substitutions of Thr6.34(279) with Lys and Asp. *Biochemistry* **40**, 13501–13509 (2001).
32. Kuszak, A. J. *et al.* Purification and functional reconstitution of monomeric mu-opioid receptors: allosteric modulation of agonist binding by Gi2. *J Biol Chem* **284**, 26732–26741 (2009).
33. Ko, M. C. *et al.* Studies of mu-, kappa-, and delta-opioid receptor density and G protein activation in the cortex and thalamus of monkeys. *J Pharmacol Exp Ther* **306**, 179–186 (2003).
34. Wolfe, L. S., Morgan, I. G. & Gombos, G. Isolation of plasma membranes from rat brain. *Biochim Biophys Acta* **241**, 737–751 (1971).
35. Massotte, D. G protein-coupled receptor overexpression with the baculovirus-insect cell system: a tool for structural and functional studies. *Biochim Biophys Acta* **1610**, 77–89 (2003).
36. Massotte, D. *et al.* Characterization of delta, kappa, and mu human opioid receptors overexpressed in baculovirus-infected insect cells. *J Biol Chem* **272**, 19987–19992 (1997).
37. Stanasila, L., Pattus, F. & Massotte, D. Heterologous expression of G-protein-coupled receptors: human opioid receptors under scrutiny. *Biochimie* **80**, 563–571 (1998).
38. Manglik, A. *et al.* Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321–326 (2012).
39. Wu, H. *et al.* Structure of the human kappa-opioid receptor in complex with JDTic. *Nature* **485**, 327–332 (2012).
40. Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071 (2010).
41. Tan, Q. *et al.* Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science* **341**, 1387–1390 (2013).
42. Huang, W. *et al.* Structural insights into mu-opioid receptor activation. *Nature* **524**, 315–321 (2015).
43. Fiser, A., Do, R. K. G. & Sali, A. Modeling of loops in protein structures. *Protein Science* **9**, 1753–1773 (2000).
44. Fenalti, G. *et al.* Molecular control of delta-opioid receptor signalling. *Nature* **506**, 191–196 (2014).
45. Periole, X., Cavalli, M., Marrink, S. J. & Ceruso, M. A. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *Journal of Chemical Theory and Computation* **5**, 2531–2543 (2009).
46. Provasi, D., Johnston, J. M. & Filizola, M. Lessons from Free Energy Simulations of delta-Opioid Receptor Homodimers Involving the Fourth Transmembrane Helix. *Biochemistry* **49**, 6771–6776 (2010).
47. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *Journal of Chemical Theory and Computation* **11**, 2144–2155 (2015).
48. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
49. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **126**, (2007).
50. 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. *The Journal of Chemical Physics* **81**, 3684–3690 (1984).
51. Scherer, M. K. *et al.* PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *Journal of Chemical Theory and Computation* **11**, 5525–5542 (2015).
52. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED2: New feathers for an old bird. *Comp. Phys. Comm.* **185** (2014).
53. Goose, J. E. & Sansom, M. S. Reduced lateral mobility of lipids and proteins in crowded membranes. *PLoS Comput Biol* **9**, e1003033 (2013).

Acknowledgements

Author Contributions

M.F. and D.P. conceived and designed the study; D.P.-G. and K.M. run the simulations; D.M., D.P.-G., D.P., and M.F. analyzed the computational results; J.M. performed the experiments under M.J.L. supervision; D.M., D.P. and M.F. wrote the manuscript with contributions by J.M. and M.J.L.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-26070-8.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.