GPCR Oligomerisation Modulation by Conformational State and Lipid Interactions Revealed by MD Simulations and Markov Models

Wanling Song, Anna L. Duncan, & Mark S.P. Sansom *

Department of Biochemistry, University of Oxford, UK

*Corresponding author mark.sansom@bioch.ox.ac.uk

Abstract

G protein coupled receptors (GPCRs) play key roles in cellular signalling pathways. GPCRs are suggested to form dimers and higher order oligomers in response to activation and other stimuli. However, a lack of structural details of the involvement of GPCR oligomerisation in receptor activation has held back our understanding of GPCR signalling in an *in vivo* context. Here, we combined extensive molecular dynamics simulations with Markov state modelling in order to understand the oligomerisation process of the Adenosine A2a receptor in response to receptor activation. The Markov state models reveal that activation leads to an increased propensity for oligomerisation, accompanied by a more interconnected oligomerisation network, when compared to the receptor in the inactive state. Coupling of the active state to a mini Gs protein shifted the oligomerisation pattern. Oligomerisation is also sensitive to changes in local membrane environment, such as the packing density of GPCRs in a membrane patch, which enhanced oligomerisation, and removal of PIP_2 from the lipid bilayer membrane which disfavoured oligomerisation. These results support combinatory allosteric modulation of GPCR oligomerisation whereby the receptor may generate a specific oligomerisation profile in response to its local environment in order to favour the formation of a specific supra-molecular signalling complex.

Introduction

G Protein-Coupled Receptors (GPCRs) are the largest superfamily in the human genome. They are major drug targets due to involvement in many physiological processes. The canonical model of GPCR activation involves the formation of a ternary structure of a receptor, an agonist and a signalling partner [1]. However, this paradigm may require revision given increasing evidence of the existence of GPCR oligomerization and its biological importance [2]. For example, single-molecule imaging has revealed that a range of GPCRs (e.g. the dopamine D_2 receptor and the M_2 muscarinic receptor) can form metastable oligomers in living cells with lifetimes ranging from sub-milliseconds to seconds [3,4] . The degree of oligomerization varies among receptors. For example, at moderate expression levels, β_2 adrenergic receptors form dimers, whilst cannabinoid CB_1 receptor and opsin form tetramers [5]. The degree of oligomerization is positively correlated with receptor density and with the presence of an agonist for some GPCRs, including the serotonin 5-HT_{2c} [6], dopamine D_2 [7] and secretin receptors [8]. The biological importance of GPCR oligomerization may lie in the possibility of functional communication between adjacent protomers (i.e. individual GPCR molecules) within these assemblies. Such communication is suggested e.g. by bivalent drugs that simultaneously bind to two receptor molecules to generate diverse pharmacological profiles [9,10], and by rescue of functionally deficient luteinizing hormone receptor via intramolecular interactions [11]. More recently, super-resolution microscopy revealed that distinct clusters of GLP-1R (a Class B GPCR) form nanodomains to initiate endocytosis [12], and that G protein coupling and β-arrestin mediated internalization are associated with the formation of CXCR4 oligomers [13]. A recent single-molecule study links formation of μ-opioid receptor dimers with specific agonists and their downstream signals [14].

The detection of G protein oligomers has pointed towards a scenario in which a signalling complex may consist of an oligomeric assembly of both GPCRs and G proteins. Single particle analyses showed that a heterotetramer of A_1-A_{2a} receptors exists as a dimer of dimers which simultaneously couples to both a Gi and a Gs protein [15], and that addition of agonist resulted in a supramolecular complex of four receptors and four G proteins [16]. Similar to GPCR oligomerization, that of G proteins is dynamic and subject to modulation by such factors as palmitoylation [17] and the presence of ligands[18].

A variety of GPCR dimer interfaces have been revealed by crystal structures: *e.g.* the symmetric interfaces of TM1,2,8/TM1,2,8 in the β_2 adrenergic [19] and μ-opioid receptors [20,21], and of TM5,6/TM5,6 as in the CXCR4 [22] and $P2Y_{12}$ receptors [23]. Asymmetric dimer interfaces have also been observed, e.g. TM1-3/TM5,6 in A_{2a} [24], and TM4,5/TM7 in the $P2Y_1$ receptors [25]. Lipids, *e.g.* cholesterol molecules and some fatty acid chains, are frequently seen at such dimer interfaces in *e.g.*

$P1Y_{12}$, A2aR and β_2AR and others [26]. Indeed, cholesterol was proposed to have an impact on GPCR oligomerization either via alteration of bilayer properties, *e.g.* thickness, fluidity, acyl tail order, and/or by modulating the potential transmembrane interactions surface of the receptors by binding at specific lipid sites [27].

Molecular dynamics (MD) simulations, in particular coarse-grained (CG) simulations, provide a useful tool to investigate protein:protein interfaces and the impact of membrane lipids on GPCR oligomerization. Rhodopsin was found to form interfaces with different association affinities corresponding to TM5/TM5 and TM1,2,8/TM1,2,8 interactions in CG simulations in POPC membranes [28]. Addition of cholesterol molecules to pure phospholipid membranes shifted the dominant dimer interface of CXCR4 from TM1/TM5-7 to TM3,4/TM3,4 in CG simulations [29]. Increasing the concentration of cholesterol in POPC membranes correlates with enhanced plasticity and flexibility of dimeric association of the serotonin$_{1A}$ receptor, a modulatory effect exerted via both direct cholesterol binding and indirect membrane effects in CG simulations [30]. Kinetics of the homodimerization of opioid receptors have also been explored via CG simulations in a POPC/CHOL membrane and Markov state models, revealing that dimerization of μOR is dynamic with kinetically distinct dimeric assemblies [31,32].

Despite these efforts, our current understanding of GPCR oligomerization still remains very limited. In particular, we remain unclear how GPCR oligomerization responds to receptor activation, and how lipid interactions in more complex membranes containing multiple lipid species may modulate such responses. These missing pieces of the puzzle hold us back from a fuller understanding of GPCR signalling in an *in vivo* context, taking place within a crowded and complex membrane environment. Here, we present a more complete view of GPCR activation in which receptor oligomerization, G protein coupling and lipid interactions are interconnected. We have employed Markov State Models (MSMs) to investigate the oligomerization behaviour of a Class A GPCR (the A2aR). By combining extensive ensembles of unbiased MD simulation data (over 2.6 ms of CG-MD in total) with MSMs we demonstrate that receptor activation results in an increased propensity to form oligomers when compared to the inactive state (in which the receptor remains predominantly in a monomeric form with occasional dimerization). Our results suggest potential combinatory allosteric modulation of GPCR signalling, in which the receptor may respond to the surrounding membrane environment to generate a unique oligomerization profile favouring the formation of specific supramolecular complexes. This helps us to understand the cell membrane-level context of the structural details of GPCR signalling complexity, presenting new possibilities to manipulate GPCR function.

Results

MD simulations

We based initial simulations on large-scale membrane systems using a mixture of lipids to form an *in vivo* mimetic model of the plasma membrane environment within which multiple copies of the receptor can freely move, enabling both association and dissociation events to occur. We have focused on oligomerization of the Adenosine A2a receptor (A2aR) as this receptor has been demonstrated experimentally to form dimers and higher order oligomers [33,34]. Thus we placed 9 copies (positioned and oriented randomly in membrane plane relative to one another) of the A2aR in a 45 x 45 nm^2 membrane with 10 lipid species present (Fig 1A) to simulate the oligomerization process (i.e. the 9-copy systems). To study how the oligomerization may change in response to receptor activation, we generated 3 ensembles of simulations for the 9-copy systems, one each for the receptor in the inactive state (PDB id 3EML) , the active state (PDB id 5G53 receptor only) and the active state in complex with a mini Gs protein (PDB id 5G53 receptor and mini Gs). For each conformational state 10 repeat simulations of 50 µs duration were performed. We also performed simulations of the three conformational states with a higher protein density (~4% by area) by including of 16 copies of the receptor in a membrane of the same dimensions (termed 16-copy), as a number of studies (e.g. [5]) have shown that altering receptor density can alter GPCR oligomerization. Given recent studies, both computational and experimental [35-37], demonstrating the interaction of phosphatidylinositol 4,5-bisphosphate (PIP_2) with Class A GPCRs, we also performed simulations with setups identical to the 9-copy system but with a lipid bilayer devoid of PIP_2 (9-copy NoPIP_2) to see whether binding of PIP_2 could modulate GPCRs via altering oligomerization.

Overall, 60 MD trajectories with a cumulative simulation time of 2.75 ms were generated (Table 1). The time evolution of the minimum distance between pairs of receptors illustrated that multiple association and dissociation events occur within the timescales simulated, and protein-protein interactions of different durations were sampled (Fig. 2A and SI Fig. S1). On average, each trajectory sampled over 20 association or dissociation events in the 9-copy systems and over 60 in the 16-copy systems (Fig. 2B). The total number of association events in the 16-copy systems was approximately $(16/9)^2$ times higher than in the 9-copy system, indicating a second order reaction. Ensemble averages of the number of monomers, dimers, and higher order oligomers present suggested that all nine simulation systems approached an equilibrium towards the end of simulations, the characteristics of which were dependent on the membrane environment and the receptor conformational state (Fig. 3 and SI Fig. S2). For example, comparing the 9-copy simulations of the Inactive vs. Active states, it can be seen that after 25 to 30 µs, the number of monomers drops to ~5 for the Inactive state compared to 2 or 3 for the Active state. The effect of omitting PIP_2 from the simulation of the Active state of

the receptor seems to decrease the receptor tendency to oligomerise such that the number of monomers remained at ~5 (Fig. 3C). Overall, given these initial observations, we decided to analyse the kinetics of association and dissociation observed in the simulations in more detail using Markov State Models (MSMs) [38].

Markov State Models

MSMs can be used to model oligomerisation in terms of a network of transitions between oligomeric states, based on information extracted from MD simulations [38]. The MD simulations were analysed in terms of oligomer formation (using a distance cut-off of 0.75 nm and hierarchical clustering)in order to build MSMs for each of the nine simulation ensembles (Fig. 1B; see Methods for details). In order to characterise the oligomeric states, clustering was performed of the oligomer structures occurring in the MD simulations; this was implemented separately within each oligomeric order if the order was lower than or equal to 5, whilst oligomers of order higher than 5 (i.e. hexamers and above) were distinguished only by their oligomeric orders. An MSM was estimated for each set of simulations using pyEMMA[39], which reweights the transitions such that the equilibrium kinetics and stationary distributions can be recovered. After statistical validation, the MSMs were used to compute metastable conformations, equilibrium probabilities, kinetics, and oligomerization networks (see Methods for details). The following analyses (of equilibrium distributions, oligomerization kinetics and pathways) are based on the information obtained from the MSMs.

Receptor activation and packing density led to unique oligomerization fingerprints

The equilibrium distributions calculated from MSMs (see Table S1) support the tentative conclusions drawn from examination of the raw data (see above). Thus, it is clear that the active state favours oligomerization compared to the inactive state (e.g. in the 9-copy systems, for the active state, 25% of the population was in oligomeric form (including dimers and higher orders) compared to 7% for the inactive state). Furthermore, the active state is also associated with a shift in favour of higher order oligomers. Not surprisingly, comparison of the 9-copy and 16-copy simulations indicates that a higher receptor density in the membrane favours oligomerization. The presence of the mini-Gs protein bound to the receptor also noticeably shifted the oligomerization profiles in favour of higher order oligomers. Conversely, omission of PIP_2 from the bilayer disfavours oligomerization, especially to form oligomers of order higher than 3 (albeit in shorter simulations). Taken together, the oligomerization fingerprints of equilibrium distribution suggest that A2aR activation may induce higher order oligomerization in a PIP_2-dependent fashion. The positive correlation of oligomerization

with both receptor activation and protein density is in agreement with observations on a number of GPCRs[6-8].

Activation increased oligomeric plasticity

To better understand GPCR oligomerization in response to receptor activation, we characterized the oligomeric assemblies and compared the distributions of these assemblies between different conformational states. For the 9-copy systems, we calculated the distributions of oligomeric assemblies of order lower than pentamers (corresponding to 99.97%, 99.93% and 91.68% of the total populations of the inactive, active, and active + mini Gs states respectively). To describe these oligomers, we used the following metrics (Fig. 4): (i) for dimers, the relative orientation angle pair (θ_1, θ_2)[40], in which θ_1 describes the orientation of the given monomer with respect to a reference monomer and θ_2 describes the rotation of the second monomer around its *z* axis to match the reference monomer (Fig 4A); (ii) for trimers, a bending angle φ defined by the centres of mass of the three monomers (Fig 4B); (iii) for tetramers and pentamers, the projected lengths D_1 and D_2 of their 1st and 2nd principal axes onto the *x* and *y* axes (Fig 4C).

For dimers (Fig. 4D), the inactive state had major populations at (100˚,50˚) and (100˚,170˚) corresponding to the interfaces ICL2,TM5//H8,TM1 and ICL2,TM4,TM5//ICL2,TM4,TM5 (where ICL = intracellular loop, TM = transmembrane helix and A//B indicates an interface between surface A and surface B). In contrast the active state exhibited a wider range of sampled angles with enhanced populations at (100˚,50˚), (25˚,75˚) and (115˚,125˚), corresponding to the interfaces ICL2,TM5//H8,TM1, ICL2,TM4//H8,TM1 and ICL3,TM5,TM6//ICL3,TM5,TM6. For trimers (Fig 4E), the inactive state showed a preference towards more linear assemblies with $\varphi > 150$˚. This preference was weakened in the active state with a much wider range of φ values. For pentamers and tetramers (Fig 4FG), a comparable situation occurred, with the active state exhibiting a broader distribution of D_1 and D_2 values, indicating that an amalgam of diverse assemblies were present, whereas the inactive state showed a narrower distribution.

Thus, overall comparisons of the distributions of these three metrics for oligomeric assemblies from different simulations suggested that when in the active state the receptor showed higher oligomeric plasticity than when in the inactive state. This was the case whether or not the mini-Gs was bound to the active state receptor.

Comparable trends were observed in the 16-copy simulations, *i.e.* higher oligomerization plasticity was shown in the active state in all oligomeric orders (SI Fig. S3). This consistency suggests that the increased plasticity may be a consequence of the conformational changes during receptor activation,

which involve a prominent outward tilt of TM6 in the intracellular leaflet of the membrane. This would be expected to enlarge the potential oligomerisation surface of the receptor thus accommodating more extensive protein-protein associations in more diverse combinations. In the absence of PIP_2 (see SI Fig. S4) the oligomeric distributions are altered relative to those from the simulations in the presence of PIP_2, and in particular the absence of PIP_2 seems to be associated with a degree of reduction in overall oligomeric plasticity.

The active state receptor exhibits a more diverse combination of association interfaces

Characterization of these oligomeric assemblies allowed us to determine the association interfaces and their relative frequencies from the MSMs. Eight association interfaces were identified in the 9-copy systems (Fig. 5; also SI Fig. S5 for the 16-copy systems), some of which have been seen in crystal structures, including ICL3,TM5,TM6//H8,ICL1TM1 in the A2aR (PDB id 4EIY), ICL2/TM4,TM5//ICL2,TM4,TM5 in the CXCR4 (PDB id 3ODU), ICL3,TM5,TM6//ICL3,TM5,TM6 in the P2Y12R (PDB id 4NTJ), and H8,TM1,TM7//H8,TM1,TM7 in the β_2R (PDB id 2RH1) and μ-opioidR (PDB id 5C1M). As anticipated from the preceding analysis of conformational plasticity, the association landscape changed between different conformational states of the receptor. Thus, the active state showed the most heterogeneous population of association interfaces. The oligomerization of the inactive state were dominated by interfaces of ICL2,TM5//H8,TM1 and ICL2,TM4,TM5//ICL2,TM4,TM5. In comparison, that of the active state showed increased occurrence of ICL3,TM5,TM6//ICL3,TM5,TM6 and ICL2,TM4,TM5//ICL3,TM5,TM6, both of which involved TM5 and TM6 that have tilted outward after receptor activation. This tilt also led to a unique interface of ECL1,TM2,TM3//ICL2,TM5 in the active state, which bordered on the enlarged surface of the intracellular side of ICL2 to TM5. The association interfaces in the oligomerization of the active + mini Gs were, to a large extent, governed by the presence of mini Gs. Interfaces such as *e.g.* ICL2,TM4,TM5//ICL2,TM4,TM5 and ICL2,TM5//H8,TM1, occurred with much lower probability due to steric hindrance of the two mini Gs, whereas interfaces ICL3,TM5,TM6//H8,ICL1,TM1 and ICL3,TM5,TM6//ICL3,TM5,TM6 occurred more frequently due to the ample space provided for the two mini Gs by the tilting of helices TM6 and TM5.

Enhanced PIP_2 interactions led to more stable protein-protein associations

Visual inspection of the MD simulation trajectories showed that PIP_2 molecules were seen at the association interfaces mediating the protein-protein interactions (Fig 6A). As it is difficult to incorporate lipid interactions in MSMs, we calculated the PIP_2 interactions directly from simulation

trajectories. Six PIP_2 binding sites, all of which formed the core of the association interfaces, were observed (Fig. 6B; Table 2). Two of these lipid binding sites (TM6_TM7 and ICL3_TM5_TM6) saw a significant increase in PIP_2 residence time in the active state compared to the inactive state, reflecting the greater accessibility to PIP_2 of the tilted TM6 helix in the active state receptor. This paralleled the higher frequency of occurrence of the association interfaces (termed the dominant interfaces) in the active state that involved these two lipid binding sites, *i.e.* ICL3,TM5,TM6//ICL3, TM5,TM6; ICL2,TM4,TM5//ICL3,TM5,TM6 and ECL1,TM2,TM3//ICL2,TM5.

The active + mini Gs state attracted more PIP_2 molecules due to the associated mini Gs providing a larger potential binding site. However, the most notable changes in both residence time and the number of bound lipids, compared to the other two states, occurred at the binding sites defined by TM3_TM5, H8_TM1 and by ICL1_TM4, all of which contributed to the core of the dominant interfaces for this state, *i.e.* ICL3,TM5,TM6//ICL3,TM5,TM6 and ICL3,TM5,TM6//H8,ICL1,TM1.

We also examined the association interfaces in the 9-copy $NoPIP_2$ systems from which the influence of PIP_2 interactions was absent (SI Fig S6). Two prominent changes showed up when comparing the oligomerization interfaces in these two membrane environments: firstly the multiplicity of the interfaces decreased in the 9-copy $NoPIP_2$ systems, indicating that the binding of PIP_2 altered the association energy landscape to allow more patterns of association; secondly, the dominant interfaces shifted from those centring around the PIP_2 binding site of ICL3_TM5_TM6 to those involving H8, TM1 and TM7 in $NoPIP_2$ systems, indicating the lipid-mediated stabilization of this interaction. Interestingly, the association interface H8,TM1,TM7//H8,TM1,TM7, which saw a significant increase of occurrence in the 9-copy $NoPIP_2$ systems (Fig 6C), has been seen to bind to cholesterol both in MD simulations[29,35] and in crystal structures[19,41]. Thus the stability of this interface would be dependent on cholesterol-mediated interactions.

Active state showed a more interconnected oligomerization network

A simulated 10 ms Monte Carlo MSM trajectory was generated for the 9-copy ensemble (see e.g. Fig 7A and 7B). The resultant trajectories revealed a dynamic process of oligomerization in which A2aR switched between oligomeric states at sub-millisecond timescales and switched to more stable oligomeric assemblies when activated and when coupled to mini Gs. The trajectories showed a higher occurrence of dimers and higher order oligomers in the active state comparing to the inactive state. We then calculated the lifetime and on-/off-rate based on the mean first passage time (MFPT). The computed lifetimes of monomer and dimer were 388 µs and 36 µs respectively for the inactive state compared with 359 µs and 380 µs respectively for the active state (SI Table S2), indicating a 10x stabilization of the dimer upon activation. The reaction rates (SI Tables S3 and S4) also revealed a

more interconnected oligomerization network in the active state, with more frequent transitions seen between higher order oligomer and dimers, due to the higher stability of the active dimers (Fig. 7C). In contrast, for the inactive state the high order oligomers tended to directly dissociate to monomers,. Comparable trends were seen for the 16-copy simulation systems, *i.e.* a more interconnected oligomerization network for the active state with frequent transitions seen between different orders of oligomers and longer lifetimes for the dominant oligomeric states (SI Fig S7).

Discussion

By combining extensive MD simulations with Markov State models, we have explored the oligomerization process of the A2a receptor as a function of receptor activation state and density of receptors in a bilayer. This was made possible by the use of large scale CG simulation systems containing multiple copies of the receptors together with unbiased simulations on multi-microsecond timescales. One of the key findings from this study is that the oligomerization plasticity and dynamics are significantly increased when the A2aR is in the active state compared to when it is in the inactive state. Whilst the inactive receptor remained predominantly in the monomeric form whilst occasionally dimerising (with a large k_{off}), the activated receptor showed increased population and longer lifetimes of dimers, alongside a more interconnected oligomerization network in which transitions between different orders of oligomers occurred more frequently. Comparing to the inactive state, the activated A2aR also exhibited a more diverse ensemble of oligomeric assemblies, resulting from an increased range of combinations of association interfaces which ultimately resulted from the increased intracellular surface around the outward tilted TM6. This was underscored by the shift towards TM6-associated interfaces in the active state oligomers. Given that this outward tilt of TM6 is a common feature in Class A GPCR activation, we suggest that the increase of oligomerization plasticity and dynamics might be an element of a shared mechanism for receptor activation. Our results also suggest an explanation of the difficulty in observing dimers or higher order oligomers in crystals.

The interactions of membrane lipids, *e.g.* cholesterols and polyunsaturated fatty acids, have been reported to affect GPCR oligomerization[42,43], via altering the receptor surface and intercalating between protomers[27]. Here we report the effect of PIP_2, which has previously been shown to interact with a number of Class A GPCRs [35-37], on A2aR oligomerization. We saw from the simulations that PIP_2 bound to the conserved PIP_2 binding sites at the intracellular side of A2aR intercalating between protomers and enhancing protein-protein interactions. A direct correlation between the probability of association interfaces and the probability of PIP_2 binding at these locations was revealed by comparing to the PIP_2 interactions in our previous simulations of single copy of the receptor[35]. This

was further confirmed by comparing to systems devoid of PIP_2. PIP_2 binding also contributed to the shift towards the TM6-associated interfaces of oligomerization in the active state. Given that our previous simulation studies have shown that the interactions of PIP_2 are largely conserved across Class A GPCRs[36], it is tempting to suggest that the effect of PIP_2 on GPCR oligomerisation may also be shared. Such interactions could be important in determining GPCR oligomerisation *e.g.* in functionally important microdomains with a high local concentration of PIP_2[44]. Interestingly, mutation of intracellular cationic residues, particularly those on ICL3, have been shown to play a critical role in formation of quaternary structures related to the functions of GPCR oligomers[45].

Recent single-molecule imaging techniques have revealed that G proteins recruited by monomers and dimers of the β_2AR elicited different signalling pathways [46]. The block of recruitment of G protein to β_2AR dimers by the inverse agonist isoproterenol suggested that the dimers in this case were responsible for the basal activity of β_2AR. The specific effect of FTY720-P, the active form of the pro-drug fingolimod, on S1PR1 was demonstrated to be dependent on receptor oligomerization causing a short-lived Gαi-mediated intracellular response that differed from that of the endogenous agonist S1P[47]. These studies suggest that the oligomerization of GPCRs may initiate signalling pathways different from those activated by the corresponding monomers. Signalling outcomes may also be dependent on the specific arrangement within oligomeric assemblies. Through a peptide-interfacing approach, A2aR-D2R heterotetramers have been shown to have different oligomeric interfaces when coupled to AC5 in the presence or absence of agonist[48].

A number of simulation studies using the MARTINI force field have provided valuable insights into the association of membrane proteins [28,32,49-53] and into protein-lipid interactions [54-56]. However, the current MARTINI model may also have some limitations [57,58] . In particular, the lack of size-specific ε for Lennard-Jones (LJ) potentials and the weak bond constants may sometimes lead to an overestimation of direct protein-protein interactions in the aqueous phase. Whilst the Inactive and Active simulation ensembles showed reasonable association-dissociation behaviour, in part due to the involvement of lipids in the protein-protein interface, the Active + mini Gs ensembles displayed a lower frequency of dissociation resulting from direct interactions of mini Gs in the aqueous phase. This low sampling of dissociations in turn resulted in large uncertainties for properties calculated from MSMs of the Active + mini Gs state. Future studies of GPCR oligomerisation using MARTINI 3 (http://cgmartini.nl/index.php/martini3beta) which includes size-dependent LJ potentials may provide further information on *e.g.* the impact of G protein coupling on GPCR oligomerisation.

In the current study we have employed relatively simple models of the lipid mixture present in mammalian plasma membranes, as in our previous studies of GPCRs [35,59]. There have been recent advances in computational studies of more complex models of the lipid bilayer composition of cell

membranes [60,61] alongside experimental studies of the asymmetric lipid composition of plasma membranes [62] which will enable such models to be used in GPCR simulation studies in the future.

Conclusions

Reflecting upon our simulation results in the context of experimental studies [46-48], we suggest that GPCR oligomerization may provide greater functional flexibility for the receptor signalling array via generating different supramolecular complexes that could initiate different signalling outcomes (Fig 8). From our MSM, we observed enhanced oligomerization plasticity facilitating the transition to oligomeric assemblies of the G protein-coupled state due to a larger sampling of oligomeric interfaces and faster transition dynamics between assemblies in the active state. In addition, distinctive oligomerization fingerprints, *i.e.* the level of oligomerization and the probabilities of various oligomeric assemblies, were elicited when the active A2aR was in different membrane conditions, which suggested that the oligomerization energy landscape was sensitive to the membrane environment. A number of studies have revealed that the coupling mechanism between GPCRs and G proteins is highly efficient, involving direct collisions with no intermediate[63,64]. The array of functional supramolecular complexes of GPCRs coupled with signalling partners including G proteins, therefore, would be dependent on the array of oligomeric assemblies presented by the activated receptor. Our simulations have revealed an additional level of complexity of GPCR allosteric modulation by receptor oligomerization, whereby the receptor can generate a specific array of oligomers according to the environment that would lead to specific signalling complexes and hence functional outcomes. This hypothesis provides a structural explanation for the observation of multiple pharmacological profiles of GPCRs [65], and potentially expands the druggable space to the protein-protein association interface and protein-membrane interface. Understanding the mechanisms of combinatory modulation of GPCR oligomerization and the role therein of lipid interactions, may present new opportunities for therapeutic targeting of GPCRs.

Methods

MD simulation set-up. Simulations used the MARTINI coarse-grained forcefield [66] as previously applied in simulations of monomeric A2aR [35]. Nine or sixteen copies of the receptor in a given conformational state were individually rotated through a random angle at the centre of a box of size of 45 x 45 x 25 nm^3 and then randomly translated in the *xy* plane. The resultant configuration was embedded in a membrane bilayer with the specified complex lipid composition (see Table 1) using the *insane.py* script[67]. Electrolyte solution corresponding to ~0.15 M NaCl was added and additional

sodium ions were added to neutralise the system. All the simulations were performed using GROMACS 5.1[68]. The CG simulations parameters were taken from ref [35]. A summary of simulations performed is provided in Table 1, amounting to a total of ~2.6 ms of simulation data collected.

Oligomeric assembly characterisation. The minimum distance between each pair of proteins was monitored as a function of time (SI Fig. S1), from which the density of such distance distribution was estimated. The density estimation showed a first peak at ~0.55 nm and a first trough at ~0.7 nm. We found that a cut-off of 0.75 nm could best discriminate effective associations. Using this cut-off and hierarchical clustering, the number of oligomers of different orders were counted and each copy of the receptor was assigned to those identified oligomers separately for each frame. The identified oligomeric assemblies with the same oligomeric order were grouped together into an 'oligomer pool', and characterised, *i.e.* clustered, within each pool if the order is lower than 6. Since calculation of root mean square deviation (RMSD) based on structural coordinates is sensitive to the order of the receptor copies, we set out to assign an order of copies to the structures. We took a reference structures randomly from each oligomer pool and calculated the coordinate RMSDs for the rest of the structures in that pool, testing all possible orders of the receptor copies. The order giving the lowest RMSD was assigned to the structure. The oligomeric structures in each oligomer pool were then clustered using the KMeans method based on their reordered coordinates. The number of clusters were manually tested for each oligomer pool to ensure a homogeneous distribution of structures in each cluster. On average, 20-60 clusters were generated for each oligomer pool, with 100 to >50,000 structures in each cluster. These clusters were individually labelled as *A.b* where *A* was the oligomeric order and *b* the cluster identifier in their oligomer pool. Such a characterisation procedure was repeated for each ensemble of simulations. A repository of all the characterised oligomeric structures in both atomistic coarse-grained coordinates can be found here: http://doi.org/10.5281/zenodo.3896363.

MSM and validation. To calculate the transition matrix *P* from the identified oligomeric assemblies, we needed to restore the oligomerisation process using the clustering results as input. We therefore constructed an 'oligomerisation trajectory' for each copy of the receptors in a trajectory by tracing which cluster this copy of receptor belonged to at each time step. Thus the 'oligomerisation trajectory' looked like this: e.g. *1, … , 2.3, … , 4.7, … , 3.10, … , 1, 1, 1*, where '*1*' denoted the monomeric state and *2.3, 4.7* etc. were the cluster labels defined above. Maximum likelihood estimation was used to generate a transition matrix P_{ij} which mapped out the probability of transitioning from state *I* at time *t* to state *j* at time $t + \tau$, where τ is the lag time of the model [69]. The transition matrix was calculated using pyEMMA [39]. The Markovian behaviour of the process can be checked by plotting the implied timescales (k) from eigenvalues μ of the transition matrix *P* at various lag times τ based on the relationship $k = -\tau/ln(\mu)$. The smallest τ, at which the implied timescales started to converge to an

unchanged rate, was taken to build the MSM. A list of the implied timescale plots can be found in SI Fig S8.

MSM Analysis. To assist understanding of the oligomerisation mechanism, the microstates resulted from Kmeans clustering were lumped into macrostates of different oligomeric orders S_I where I = *1,2,..,9* in the 9-copy systems, and *1,2,…,16* in the 16-copy systems. The macrostate transition probabilities were obtained from the microstate probabilities with:

$$\tilde{P}_{IJ} = \frac{1}{\pi_I} \sum_{i \in I} \sum_{i \in J} \pi_i p_{ij}$$

where $\pi_I = \sum_{i \in I} \pi_i$ is the macrostate equilibrium probability. The mean first passage time (mfpt) between the macrostates was calculated as described in ref [70] via the functionality in pyEMMA. The lifetime of a macrostate S_I was calculated as mfpt(I,J), where $J \in \Theta$ and $J \notin I$.

Lipid Interaction Analysis. Lipid binding sites, interaction counts and residence times were calculated using PyLipID (https://github.com/wlsong/PyLipID). A continuous lipid contact was defined using a dual cut-off scheme such that the interaction started when lipid headgroup moved to a given residue closer than 0.55 nm and ended when the headgroup moved farther than 1.0 nm. Lipid binding sites were derived from the lipid contact information using graph theory and community analysis. Graphs were constructed using residues as nodes and the frequency with which a given pair of residues interacted with the same lipid molecule as edges. The best partition function of the community library (https://github.com/taynaud/python-louvain) was used to detect community structures. Each calculated community was considered a binding site.

Lipid residence time of a binding site was calculated from the normalised survival time-correlation function $\sigma(t)$:

$$\sigma(t) = \frac{1}{N_j} \frac{1}{T-t} \sum_{j=1}^{N_j} \sum_{v=0}^{T} \tilde{n}_j(v, v+t)$$

where T is the total simulation time, N_j the total number of lipids of a given type, and $\tilde{n}_j(v, v+t)$ is a function that takes the value 1 if a lipid occupies the given binding site for a duration of t after forming the contact at time v, and takes the value 0 if otherwise.

Acknowledgements

Research in the M.S.P.S. group is supported by Wellcome (208361/Z/17/Z), BBSRC (BB/R00126X/1), and PRACE (Partnership for Advanced Computing in Europe; 2016163984). W.S. acknowledges support from the Newton International Fellowship. This project made use of time on ARCHER via the HECBioSim, supported by EPSRC (EP/L000253/1).

References

1 Glukhova, A. *et al.* Rules of engagement: GPCRs and G proteins. *ACS Pharmacol. Transl. Sci.* **1**, 73-83 (2018).

2 Milligan, G., Ward, R. J. & Marsango, S. GPCR homo-oligomerization. *Curr. Opin. Cell Biol.* **57**, 40-47 (2019).

3 Guo, W. *et al.* Dopamine D2 receptors form higher order oligomers at physiological expression levels. *EMBO J.* **27**, 2293-2304 (2008).

4 Pisterzi, L. F. *et al.* Oligomeric size of the M2 muscarinic receptor in live cells as determined by quantitative fluorescence resonance energy transfer. *J. Biol. Chem.* **285**, 16723-16738 (2010).

5 Walsh, S. M. *et al.* Single proteoliposome high-content analysis reveals differences in the homo-oligomerization of GPCRs. *Biophys. J.* **115**, 300-312 (2018).

6 Ward, R. J., Pediani, J. D., Godin, A. G. & Milligan, G. Regulation of oligomeric organization of the serotonin 5-hydroxytryptamine 2C (5-HT2C) receptor observed by spatial intensity distribution analysis. *J Biol Chem* **290**, 12844-12857 (2015).

7 Tabor, A. *et al.* Visualization and ligand-induced modulation of dopamine receptor dimerization at the single molecule level. *Sci. Reports* **6**, 33233 (2016).

8 Ward, R. J., Pediani, J. D., Harikumar, K. G., Miller, L. J. & Milligan, G. Spatial intensity distribution analysis quantifies the extent and regulation of homodimerization of the secretin receptor. *Biochem J* **474**, 1879-1895 (2017).

9 Hiller, C., Kühhorn, J. & Gmeiner, P. Class A g-protein-coupled receptor (GPCR) dimers and bivalent ligands. *J. Med. Chem.* **56**, 6542-6559 (2013).

10 Shonberg, J., Scammells, P. J. & Capuano, B. Design Strategies for Bivalent Ligands Targeting GPCRs. *ChemMedChem* **6**, 963-974 (2011).

11 Rivero-Muller, A. *et al.* Rescue of defective G protein-coupled receptor function in vivo by intermolecular cooperation. *Proc Natl Acad Sci U S A* **107**, 2319-2324 (2010).

12 Buenaventura, T. *et al.* Agonist-induced membrane nanodomain clustering drives GLP-1 receptor responses in pancreatic beta cells. *PLOS Biology* **17**, e3000097 (2019).

13 Ge, B. *et al.* Single-molecule imaging reveals dimerization/oligomerization of CXCR4 on plasma membrane closely related to its function. *Sci Rep* **7**, 16873 (2017).

14 Möller, J. *et al.* Single-molecule analysis reveals agonist-specific dimer formation of μ-opioid receptors. *Nature Chem. Biol.* **DOI 10.1038/s41589-020-0566-1** (2020).

15 Navarro, G. *et al.* Quaternary structure of a G-protein-coupled receptor heterotetramer in complex with Gi and Gs. *BMC Biology* **14**, 26 (2016).

16 Shivnaraine, R. V. *et al.* Single-molecule analysis of the supramolecular organization of the M2 muscarinic receptor and the Gαi1 protein. *J. Amer. Chem. Soc.* **138**, 11583-11598 (2016).

17 Yang, H., Qu, L., Ni, J., Wang, M. & Huang, Y. Palmitoylation participates in G protein coupled signal transduction by affecting its oligomerization. *Molec. Membrane Biol.* **25**, 58-71 (2008).

18 Li, Y., Shivnaraine, R. V., Huang, F., Wells, J. W. & Gradinaru, C. C. Ligand-induced coupling between oligomers of the M2 Receptor and the Gi1 protein in live cells. *Biophys. J.* **115**, 881-895 (2018).

19 Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β2-adrenergic G protein–coupled receptor. *Science* **318**, 1258-1265 (2007).

20 Manglik, A. *et al.* Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321-326 (2012).

21 Huang, W. *et al.* Structural insights into m-opioid receptor activation. *Nature* **524**, 315-321 (2015).

22 Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066-1071 (2010).

23 Zhang, K. *et al.* Structure of the human P2Y12 receptor in complex with an antithrombotic drug. *Nature* **509**, 115 (2014).
24 Liu, W. *et al.* Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**, 232-236 (2012).
25 Zhang, D. *et al.* Two disparate ligand-binding sites in the human P2Y1 receptor. *Nature* **520**, 317-321 (2015).
26 Duncan, A. L., Song, W. & Sansom, M. S. P. Lipid-Dependent Regulation of Ion Channels and G Protein–Coupled Receptors: Insights from Structures and Simulations. *Annual Review of Pharmacology and Toxicology* (2019).
27 Gahbauer, S. & Böckmann, R. A. Membrane-mediated oligomerization of G protein coupled receptors and its implications for GPCR function. *Frontiers Physiol.* **7** (2016).
28 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. Amer. Chem. Soc.* **134**, 10959-10965 (2012).
29 Pluhackova, K., Gahbauer, S., Kranz, F., Wassenaar, T. A. & Bockmann, R. A. Dynamic cholesterol-conditioned dimerization of the G protein coupled chemokine receptor type 4. *PLoS Comp. Biol.* **12**, e1005169 (2016).
30 Prasanna, X., Sengupta, D. & Chattopadhyay, A. Cholesterol-dependent conformational plasticity in GPCR dimers. *Sci. Reports* **6**, 31858 (2016).
31 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.* **11** (2015).
32 Meral, D. *et al.* Molecular details of dimerization kinetics reveal negligible populations of transient mu-opioid receptor homodimers at physiological concentrations. *Sci. Reports* **8**, 7705 (2018).
33 Vidi, P. A., Chemel, B. R., Hu, C. D. & Watts, V. J. Ligand-dependent oligomerization of dopamine D2 and adenosine A(2A) receptors in living neuronal cells. *Molec. Pharmacol.* **74**, 544-551 (2008).
34 Vidi, P. A., Chen, J. J., Irudayaraj, J. M. K. & Watts, V. J. Adenosine A(2A) receptors assemble into higher-order oligomers at the plasma membrane. *FEBS Lett.* **582**, 3985-3990 (2008).
35 Song, W. L., Yen, H. Y., Robinson, C. V. & Sansom, M. S. P. State-dependent lipid interactions with the A2a receptor revealed by MD simulations using in vivo-mimetic membranes. *Structure* **27**, 392-403 (2019).
36 Yen, H. Y. *et al.* PtdIns(4,5)P-2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **559**, 424-427 (2018).
37 Huang, W. *et al.* Structure of the neurotensin receptor 1 in complex with β-arrestin 1. *Nature* **579**, 303–308 (2020).
38 Husic, B. E. & Pande, V. S. Markov state models: from an art to a science. *J. Amer. Chem. Soc.* **140**, 2386-2396 (2018).
39 Scherer, M. K. *et al.* PyEMMA 2: A software package for estimation, validation, and analysis of Markov models. *J. Chem. Theor. Comput.* **11**, 5525-5542 (2015).
40 Gahbauer, S., Pluhackova, K. & Bockmann, R. A. Closely related, yet unique: Distinct homo- and heterodimerization patterns of G protein coupled chemokine receptors and their fine-tuning by cholesterol. *PLoS Comput Biol* **14**, e1006062 (2018).
41 Hanson, M. A. *et al.* A specific cholesterol binding site is established by the 2.8 Å structure of the human β2-adrenergic receptor. *Structure* **16**, 897-905 (2008).
42 Wang, J., He, L., Combs, C. A., Roderiquez, G. & Norcross, M. A. Dimerization of CXCR4 in living malignant cells: control of cell migration by a synthetic peptide that reduces homologous CXCR4 interactions. *Molec. Cancer Therapeut.* **5**, 2474-2483 (2006).
43 Guixà-González, R. *et al.* Membrane omega-3 fatty acids modulate the oligomerisation kinetics of adenosine A2A and dopamine D2 receptors. *Scientific Reports* **6**, 19839 (2016).

44 van den Bogaart, G. *et al.* Membrane protein sequestering by ionic protein-lipid interactions. *Nature* **479**, 552–555 (2011).
45 Navarro, G. *et al.* Interactions between intracellular domains as key determinants of the quaternary structure and function of receptor heteromers. *J. Biol. Chem.* **285**, 27346-27359 (2010).
46 Kasai, R. S., Fujiwara, T. K. & Kusumi, A. Metastable GPCR dimers trigger the basal signal by recruiting G-proteins. *BioRxiv*, 2020.2002.2010.929588 (2020).
47 Patrone, M. *et al.* Combinatorial allosteric modulation of agonist response in a self-interacting G-protein coupled receptor. *Comms. Biol.* **3**, 27 (2020).
48 Navarro, G. *et al.* Evidence for functional pre-coupled complexes of receptor heteromers and adenylyl cyclase. *Nature Comms.* **9**, 1242 (2018).
49 Periole, X., Huber, T., Marrink, S. J. & Sakmar, T. P. G protein-coupled receptors self-assemble in dynamics simulations of model bilayers. *J. Amer. Chem. Soc.* **129**, 10126-10132 (2007).
50 Chavent, M., Chetwynd, A. P., Stansfeld, P. J. & Sansom, M. S. Dimerization of the EphA1 receptor tyrosine kinase transmembrane domain: Insights into the mechanism of receptor activation. *Biochem.* **53**, 6641-6652 (2014).
51 Domanski, J., Sansom, M. S. P., Stansfeld, P. J. & Best, R. B. Balancing force field protein-lipid interactions to capture transmembrane helix-helix association. *J. Chem. Theor. Comput.* **14**, 1706-1715 (2018).
52 Gahbauer, S. & Bockmann, R. A. Comprehensive characterization of lipid-guided G protein-coupled receptor dimerization. *J. Phys. Chem. B* **124**, 2823-2834 (2020).
53 Periole, X., Zeppelin, T. & Schiott, B. Dimer interface of the human serotonin transporter and effect of the membrane composition. *Sci. Reports* **8**, 5080 (2018).
54 Hedger, G. & Sansom, M. S. P. Lipid interaction sites on channels, transporters and receptors: recent insights from molecular dynamics simulations. *Biochim. Biophys. Acta* **1858**, 2390–2400 (2016).
55 Corradi, V. *et al.* Lipid–protein interactions are unique fingerprints for membrane proteins. *ACS Central Sci.* **4**, 709-717 (2018).
56 Duncan, A. L., Corey, R. A. & Sansom, M. S. P. Defining how multiple lipid species interact with inward rectifier potassium (Kir2) channels. *Proc. Natl. Acad. Sci. USA* **117**, 7803-7813 (2020).
57 Javanainen, M., Martinez-Seara, H. & Vattulainen, I. Excessive aggregation of membrane proteins in the Martini model. *PLoS ONE* **12**, e0187936 (2017).
58 Alessandri, R. *et al.* Pitfalls of the Martini model. *J. Chem. Theor. Comput.* **15**, 5448-5460 (2019).
59 Koldsø, H. & Sansom, M. S. P. Organization and dynamics of receptor proteins in a plasma membrane. *J. Amer. Chem. Soc.* **137**, 14694−14704 (2015).
60 Ingolfsson, H. I. *et al.* Computational lipidomics of the neuronal plasma membrane. *Biophys. J.* **113**, 2271-2280 (2017).
61 Marrink, S. J. *et al.* Computational modeling of realistic cell membranes. *Chem. Rev.* **119**, 6184-6226 (2019).
62 Lorent, J. H. *et al.* Plasma membranes are asymmetric in lipid unsaturation, packing and protein shape. *Nature Chem. Biol.* **16**, 644–652 (2020).
63 Hein, P., Frank, M., Hoffmann, C., Lohse, M. J. & Bünemann, M. Dynamics of receptor/G protein coupling in living cells. *EMBO J.* **24**, 4106-4114 (2005).
64 Falkenburger, B. H., Jensen, J. B. & Hille, B. Kinetics of M1 muscarinic receptor and G protein signaling to phospholipase C in living cells. *J. Gen. Physiol.* **135**, 81-97 (2010).
65 Ferre, S. *et al.* G protein-coupled receptor oligomerization revisited: Functional and pharmacological perspectives. *Pharmacol. Rev.* **66**, 413-434 (2014).
66 de Jong, D. H. *et al.* Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theor. Comput.* **9**, 687-697 (2013).

67 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. *J. Chem. Theor. Comput.* **11**, 2144-2155 (2015).
68 Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19-25 (2015).
69 Prinz, J.-H. *et al.* Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **134**, 174105 (2011).
70 Singhal, N. & Pande, V. S. Error analysis and efficient sampling in Markovian state models for molecular dynamics. *J. Chem. Phys.* **123**, 204909 (2005).

Tables

Table 1. Overview of Simulations Performed

System	A2aR State	Number of copies of protein	Duration (μs) x repeats	Lipid bilayer composition
9-copy	Inactive	9	50 x 10	*Extracellular side:* POPC(20%), DOPC(20%), POPE(5%), DOPE(5%), Sph(15%), GM3(10%), CHOL(25%) *Intracellular side:* POPC(5%), DOPC(5%), POPE(20%), DOPE(20%), POPS(8%), DOPS(7%), PIP2(10%), CHOL(25%)
	Active		50 x 10	
	Active + mini Gs		50 x 10	
16-copy	Inactive	16	50 x 5	
	Active		50 x 5	
	Active + mini Gs		50 x 5	
9-copy NoPIP2	Inactive	9	25 x 5	*Extracellular side:* POPC(20%), DOPC(20%), POPE(5%), DOPE(5%), Sph(15%), GM3(10%), CHOL(25%) *Intracellular side:* POPC(5%), DOPC(5%), POPE(25%), DOPE(25%), POPS(8%), DOPS(7%), CHOL(25%)
	Active		25 x 5	
	Active + mini Gs		25 x 5	

All simulation boxes were 45 x 45 x 25 nm^3.

Table 2. PIP_2 Binding Sites in 9-Copy Simulations

Binding site§	**TM6_TM7**			**TM3_TM5**			**ICL2_TM3_TM4**			**ICL3_TM5_TM6**			**H8_TM1**			**ICL1_TM4**		
A2aR state*	I	A	AmG	I	A	AmG	I	A	AmG	I	A	AmG	I	A	AmG	I	A	AmG
k_{off} (μs^{-1})	0.34	0.08	0.03	0.11	0.16	0.04	0.23	0.25	0.03	0.91	0.12	0.09	0.14	0.29	0.04	0.22	0.28	0.04
Residence time (µs)	2.8	11.8	28.5	8.9	6.2	22.1	4.2	3.9	26.4	1.1	8.1	10.1	6.8	3.4	21.7	4.5	3.5	23.4
Number of lipids	3.4	2.8	2.5	2.4	2.4	3.7	2.2	2.3	2.1	1.1	3.1	3.1	2.3	1.2	3.5	1.4	1.3	2.3

§ The binding sites are illustrated in Fig. 6B, to which the colour coding used above corresponds.
* I = inactive; A = active; AmG = active + mini Gs.

Figures

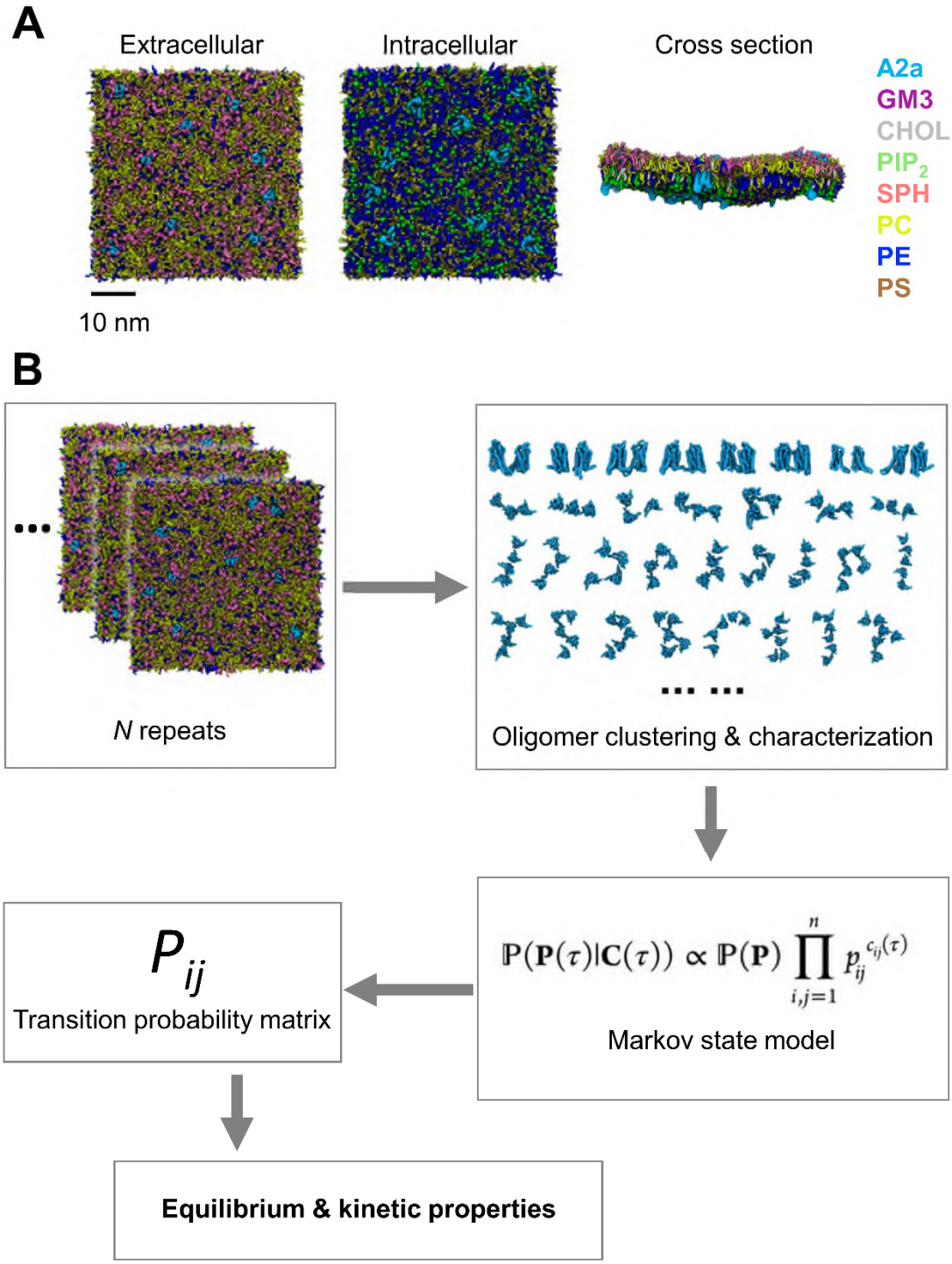

Figure 1. Simulations and Markov State Models of A2aR oligomerization. (**A**) System setup of A2aR oligomerization simulations. The selected number of receptor molecules (pale blue) were randomly inserted into a mixed lipid membrane of area 45 x 45 nm^2. Views of the systems from the extracellular and intracellular surfaces and in cross section are shown, with the lipid species present colour coded. Details of the different simulation systems setups are listed in Table 1. (**B**) Schematic workflow for constriction of the Markov State Model. The oligomers from an ensemble of simulation repeats were characterised and a transition matrix between the receptor monomers, dimers and these oligomers was calculated to generate a Markov state model using pyEMMA[39].

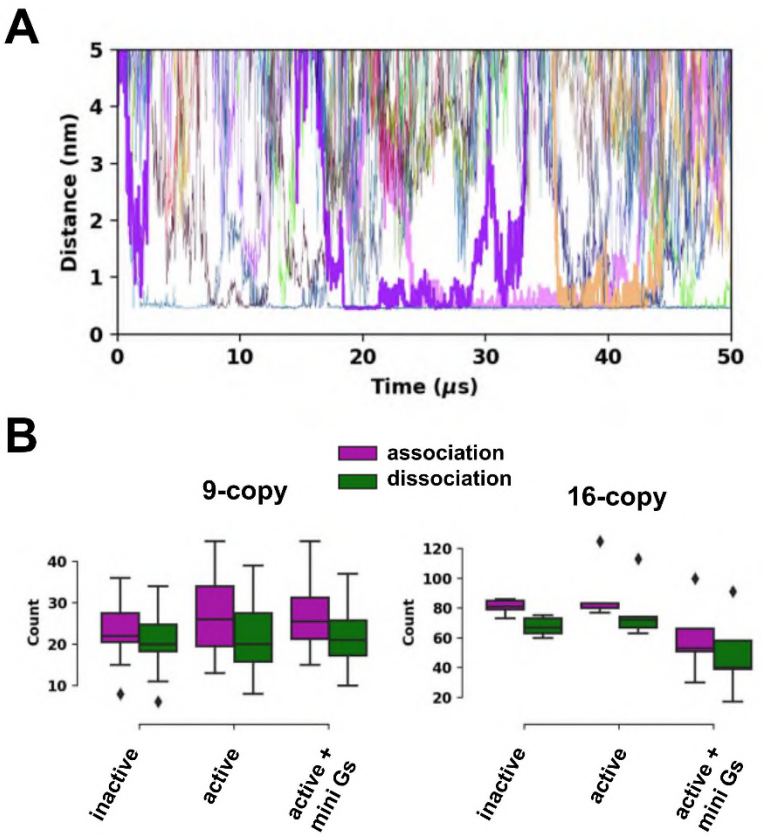

Figure 2. A2aR oligomerisation sampled by MD simulations. (A) the time evolution of minimum distance between pairs of receptors (from a simulations trajectory in the inactive state 9-copy simulation ensemble), illustrating that both association and dissociation events occur within the timescale simulated (selected events are highlighted by bold traces). (**B)** The number of association (defined as a pair of proteins coming closer than 0.75 nm) and dissociation (defined as a pair of proteins separating to further than 0.75 nm) events sampled in each trajectory for each of the 9-copy and 16-copy system ensembles. **(C)** The associations and dissociations have led to a dynamic equilibrium as illustrated by the time course of oligomer formation for the inactive state 9-copy, active state 9-copy, and active state 9-copy $NoPIP_2$ simulations. Data from all the trajectories in the same simulation ensemble were plotted out to illustrate the ensemble trend.

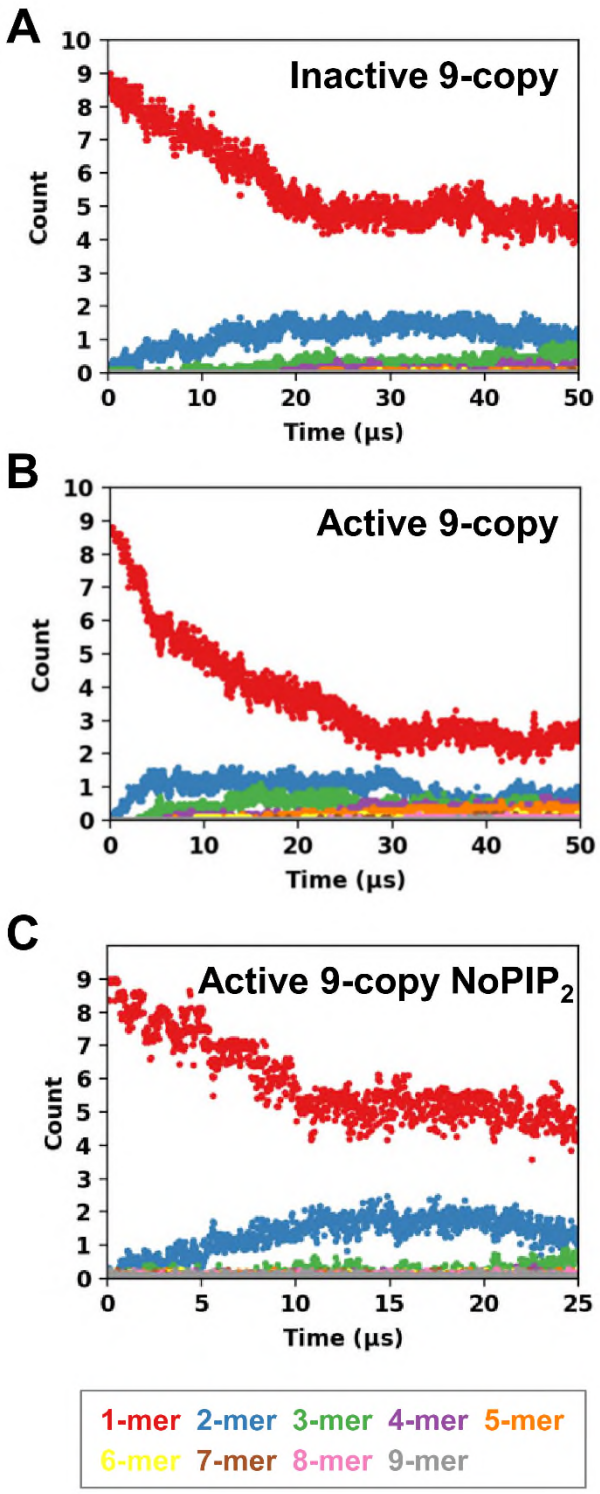

Figure 3. Time courses of oligomer formation for the **A** inactive state 9-copy, **B** active state 9-copy, and **C** active state 9-copy $NoPIP_2$ simulations. The number of monomers, dimers, and higher order oligomers present as a function of time. Data from trajectories in the same set of simulations were merged and are displayed together.

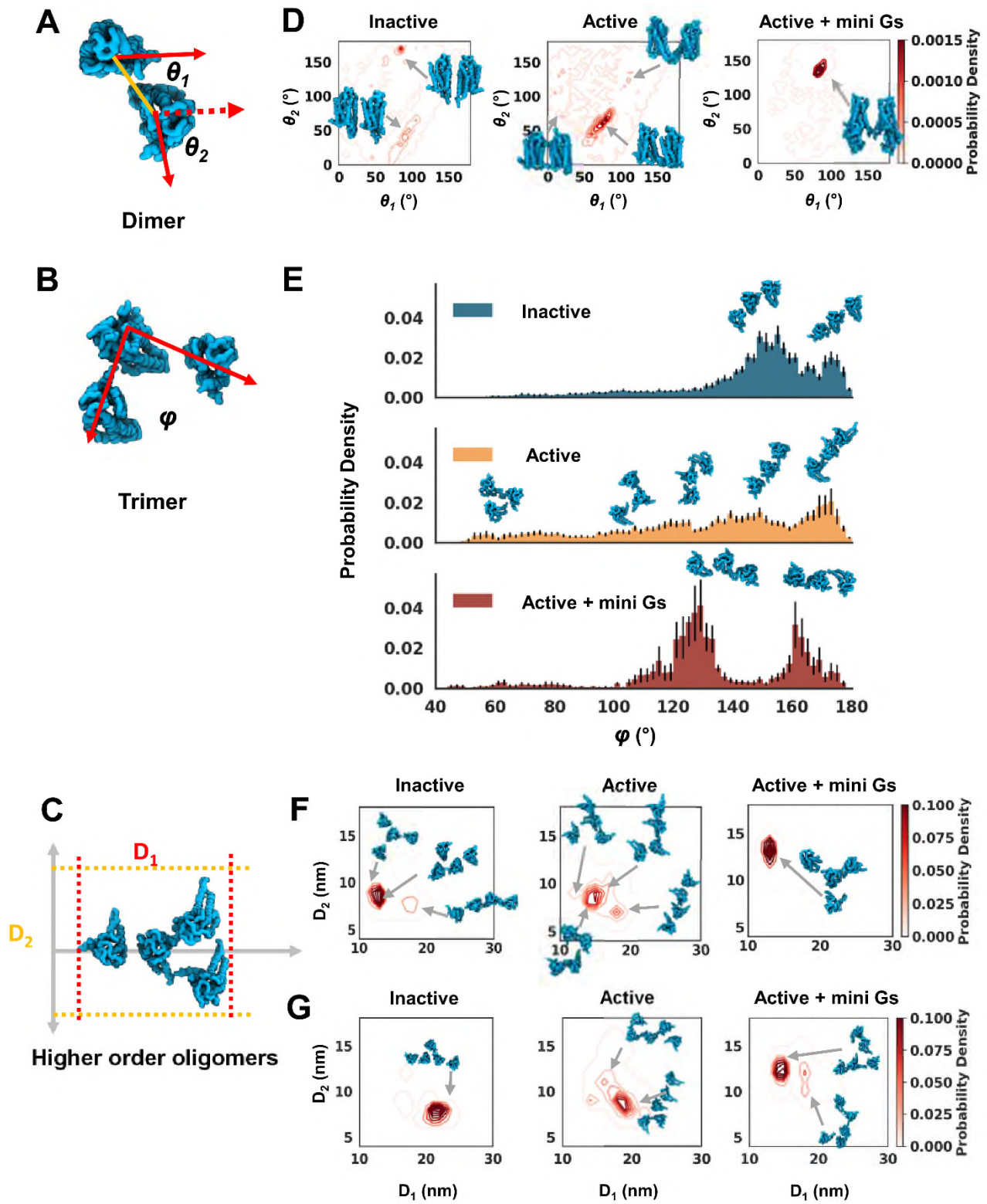

Figure 4. Distribution of assemblies within each oligomeric order in the 9-copy ensembles. The equilibrium distributions are calculated from the MSMs. (**A**) Two relative orientation angles (θ_1, θ_2) were used to describe dimers. (**B**) The bending angle θ defined by the centre of mass of the three monomers was used to describe trimers. (**C**) The projected lengths D_1 and D_2 of their 1st and 2nd principal axes onto the *x* and *y* axes was used to describe tetramers and pentamers. The distributions of these oligomeric assemblies from different conformational states are shown in **D - G**. Data for 16-copy and 9-copy NoPIP2 ensembles are shown in SI Fig. S3 and S4.

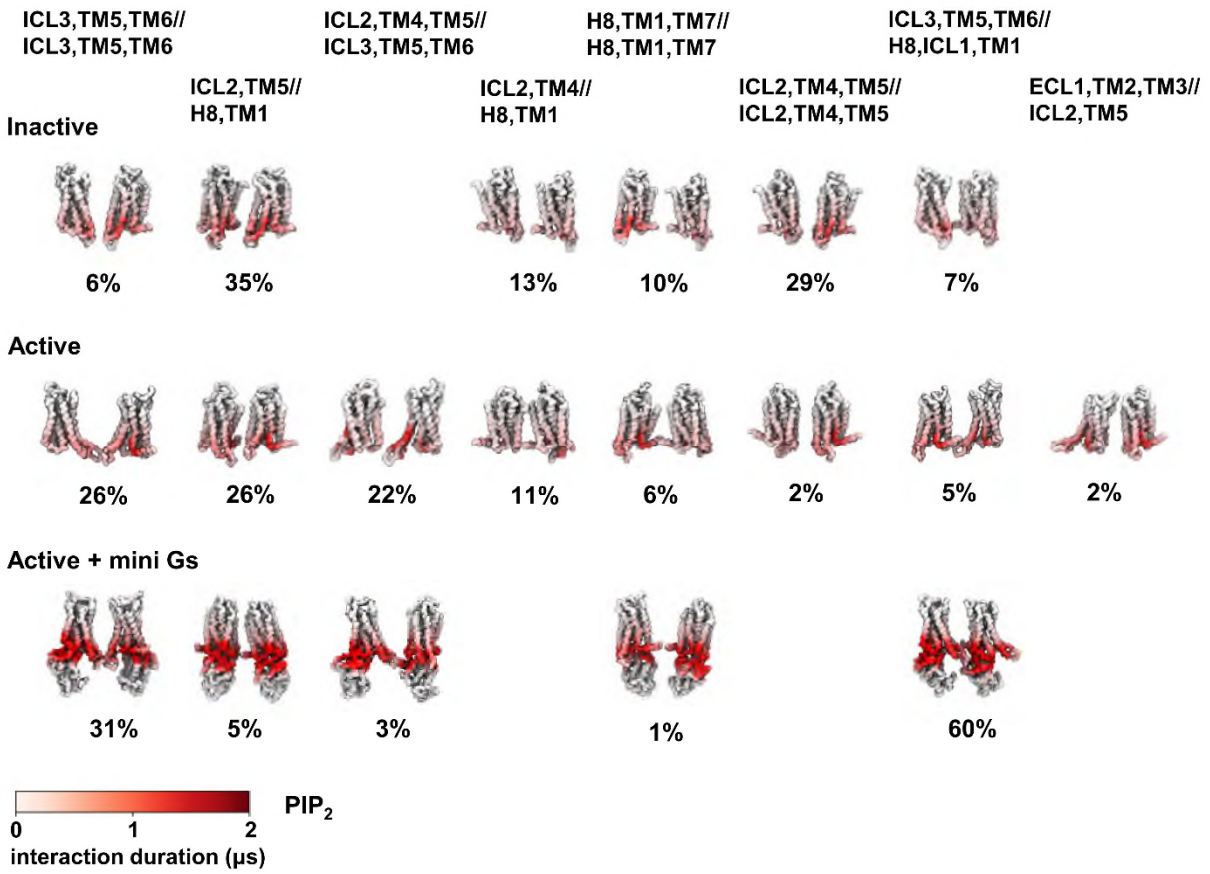

Figure 5. Association interfaces in the 9-copy ensembles. Eight protein-protein association interfaces are compared for the three conformational states of A2aR. The interface distributions (listed below each structure image) are calculated from MSMs equilibrium distribution, via counting the probability-weighted occurrence in all oligomeric assemblies then normalising the total frequency to 1. PIP_2 residence times from their respective simulations are mapped onto the receptor structures with a white-to-red colourmap. Data for 16-copy and 9-copy NoPIP2 ensembles are shown in SI Fig. S5 and S6.

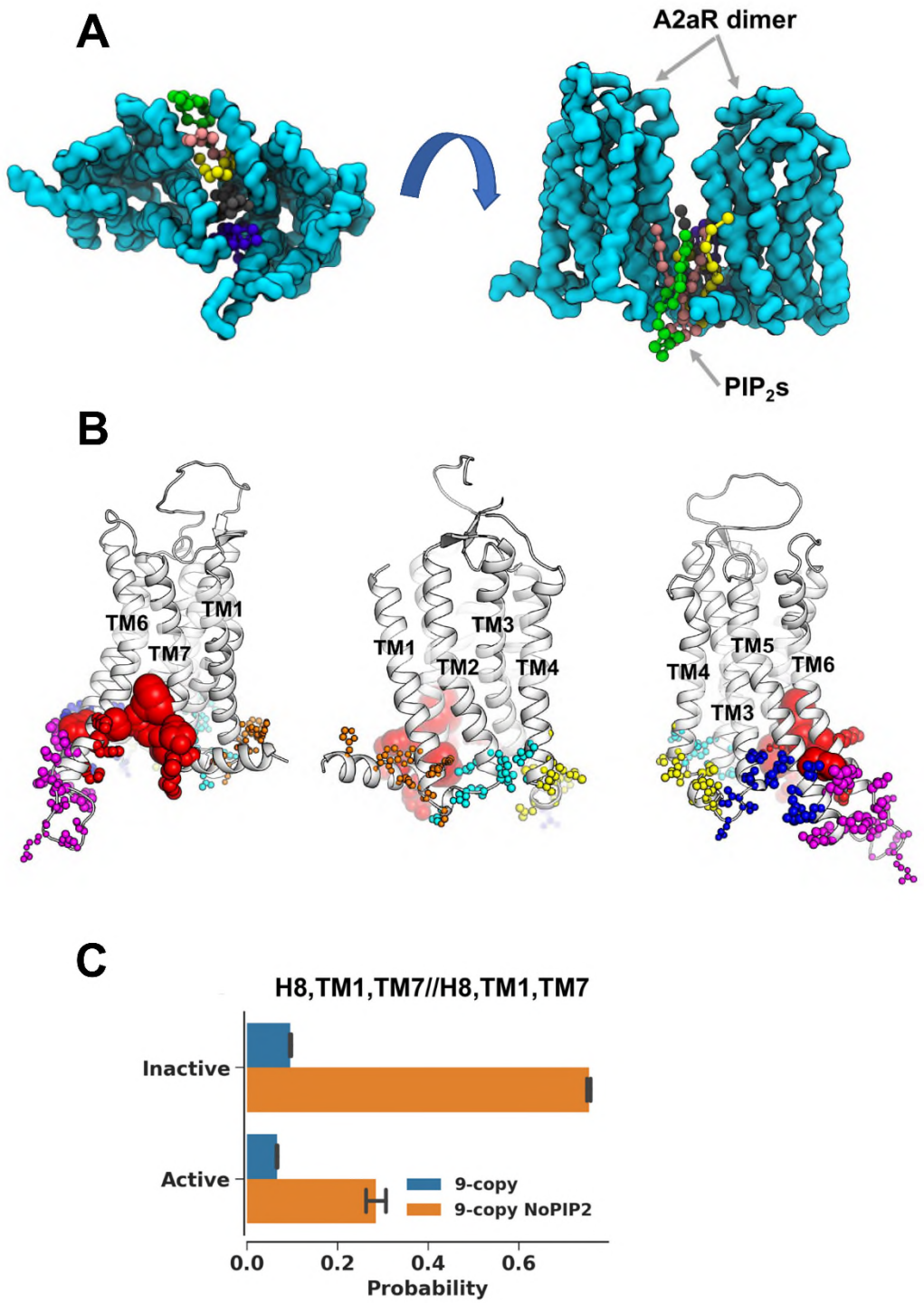

Figure 6. PIP_2-mediated protein-protein interactions. (**A**) Snapshot from the 9-copy active simulations of the interaction interface ICL2,TM5//H8,TM1 with mediating PIP_2 molecules shown in ball and stick. The receptor backbone beads were shown in cyan. (**B**) PIP_2 binding sites on the active A2aR structure. The 6 binding sites were calculated using PyLipID (https://github.com/wlsong/PyLipID), and shown using different colours. The residues in each binding site are shown as atomic spheres with sphere scale proportional to their PIP_2 residence times for the active state 9-copy simulations. (**C**) Comparison of the relative frequency of the H8,TM1,TM7//H8,TM1,TM7 interaction interface in the 9-copy systems and 9-copy $NoPIP_2$ systems.

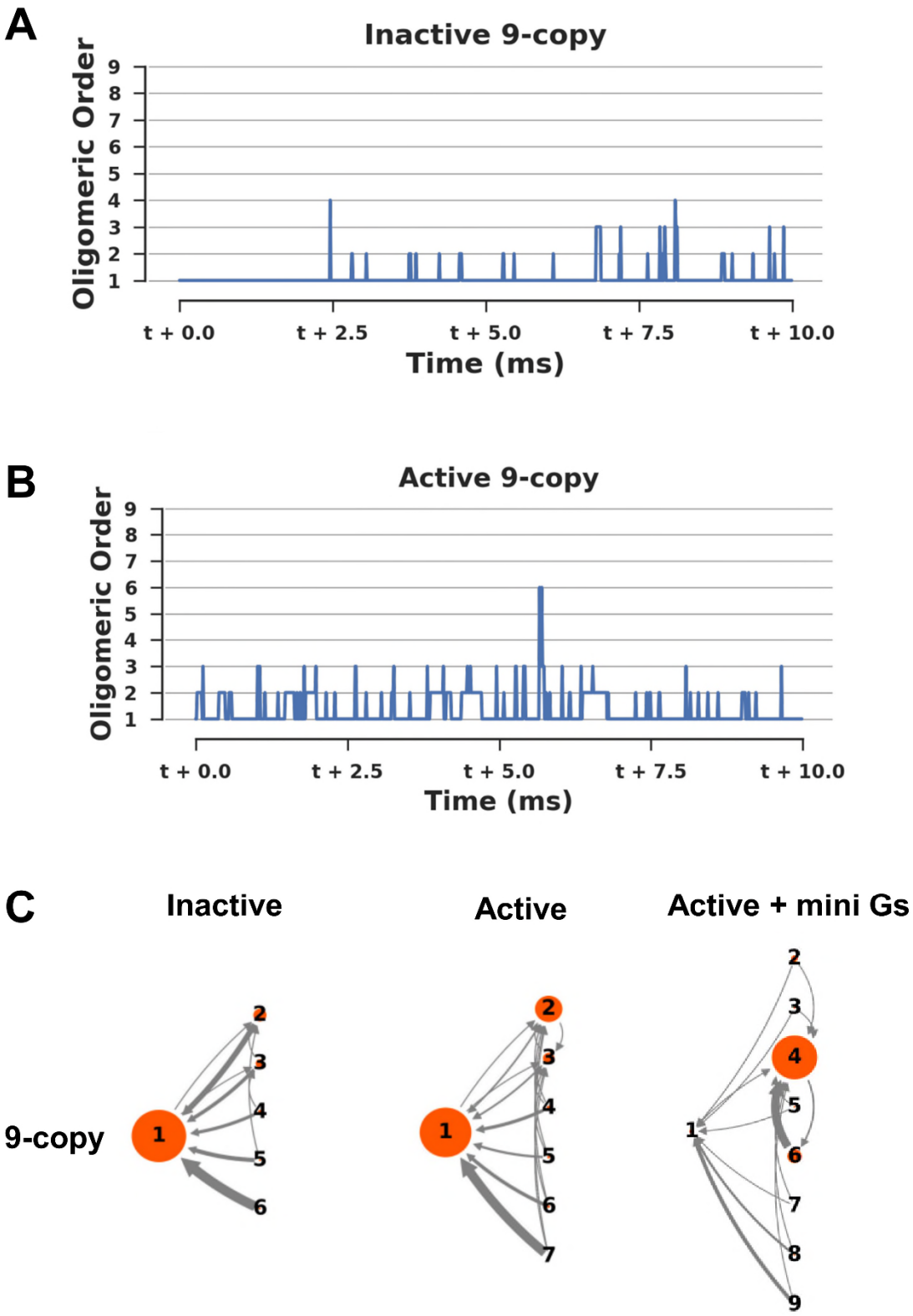

Figure 7. Oligomerization kinetics of the 9-copy ensemble. A 10 ms trajectory was generated for the inactive state **(A)** and active state **(B)** using the Monte Carlo algorithm and MSM transition probability. (**C)** The oligomerization networks for each of the three conformational states are shown. The thickness of the arrows is proportional to the corresponding reaction rate (only reaction rates >50 s^{-1} are shown) and the size of the circles to the equilibrium distributions. Reaction rates were calculated as the reciprocal of the corresponding mean first passage time.

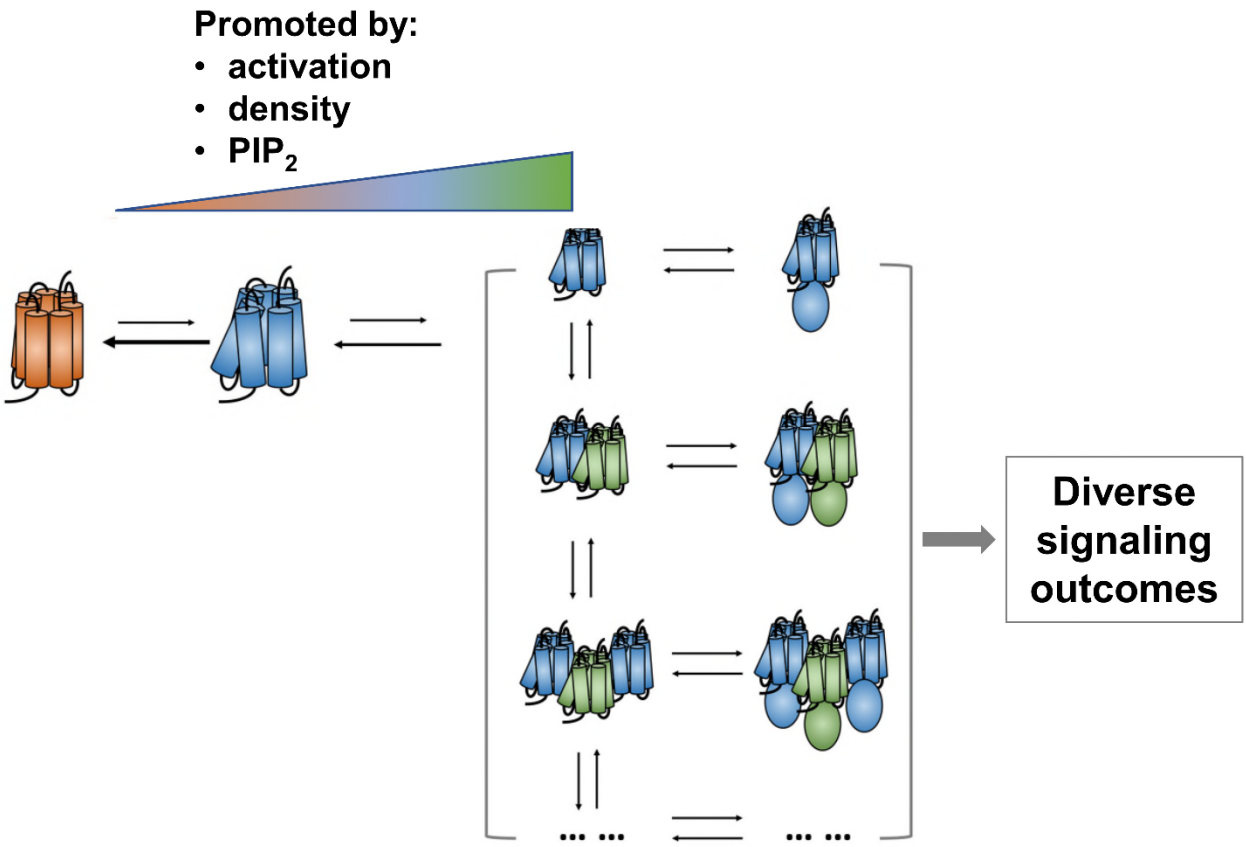

Figure 8. Allosteric modulation and GPCR oligomerization. The activated GPCR generates an array of oligomeric assemblies which couples to signalling partners, generating an array of supramolecular signalling complexes capable of initiating different signalling pathways. This mechanism of combinatory modulation of GPCR is responsive to the lipid bilayer environment and the conformational state of the receptor. Receptor activation, PIP_2 and/or mini-Gs binding, and an increase in receptor density in the bilayer all promote oligomerization.