Activation of G-protein coupled receptors is thermodynamically linked to lipid solvation

Alison N. Leonard^a and Edward Lyman^{a,b,†}

^a Department of Physics and Astronomy, University of Delaware, Newark, Delaware

^b Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware

[†] Corresponding author

Abstract

Preferential lipid solvation of the G-protein coupled A_{2A} adenosine receptor (A_{2A}R) is evaluated from 35 μ sec of all-atom molecular dynamics simulation. A coarse-grained transition matrix algorithm is developed to overcome slow equilibration of the first solvation shell, obtaining statistically robust estimates of the free energy of solvation by different lipids for the receptor in different activation states. Results indicate preference for solvation by unsaturated chains, which favors the active receptor. A model for lipid-dependent GPCR activity is proposed in which the chemical potential of lipids in the bulk membrane modulates receptor activity. The enthalpy and entropy associated with moving saturated vs. unsaturated lipids from bulk to A_{2A}R's first solvation shell are compared. In the simulated mixture, saturated chains are disordered (i.e., obtain a favorable entropic contribution) when partitioning to the receptor surface, but this is outweighed by a favorable enthalpic contribution for unsaturated chains to occupy the first solvation shell.

Introduction

Lipid composition is highly differentiated among different cell types (Řezanka et al., 2018; Pradas et al., 2018), between the membranes within a single cell (Doralicia et al., 2019), and even between the leaflets of a single membrane (Lorent et al., 2020). These distinct lipid environments are known to influence the function of membrane proteins such as G-protein coupled receptors (GPCRs) (Periole, 2017) — transmembrane proteins that comprise the largest class of proteins in the human genome and are important drug targets (Katritch et al., 2013). Understanding lipid-dependent activity of GPCRs might therefore yield new therapeutic strategies based on tissue- or context-specific pharmacological targeting.

Tight, site-specific interactions between lipids and GPCRs have been detected using x-ray diffraction (Yeagle, 2014), electron microscopy (Gonen et al., 2005), and high resolution mass spectrometry (Bolla et al., 2019; Montenegro et al., 2017), and so mechanisms have tended to focus on specific cholesterol (Cang et al., 2013; Hanson et al., 2008; Lee et al., 2013; McGraw et al., 2019; Rouviere et al., 2017; Gutierrez et al., 2016; O'Malley et al., 2011; Lee & Lyman, 2013) and lipid (Javanainen et al., 2019; Grossfield et al., 2006; Neale et al., 2015; Song et al., 2019; Naranjo et al., 2016) interactions. In contrast, resolving lipid solvation of a transmembrane protein (i.e., weaker but still preferential interactions) presents a challenge for experiment, requiring labeled protein (Loura et al., 2010; Soubias & Gawrisch, 2005) or lipid probes (Marsh et al., 1982; Marsh, 1990). The best studied member of the family is rhodopsin, for which NMR (Soubias & Gawrisch, 2005; Soubias & Gawrisch, 2012; Soubias et al., 2008; Feller et al., 2003), simulation (Grossfield et al., 2006), and theoretical (Botelho et al., 2002; Huber et al., 2004) work suggest that the negative intrinsic curvature of polyunsaturated fatty acid (PUFA)-containing lipids shifts the receptor toward active conformations. Javanainen, et al. found in simulations of A_{2A}R that PUFA chains in the first solvation shell mediate partitioning of the receptor into ordered lipid phases, following experimental partitioning measurements in phase separated GUVs reported by Gutierrez and Malmstadt (Javanainen et al., 2019;

Gutierrez et al., 2019). Yang and Lyman also reported preferential solvation of A_{2A}R by the monounsaturated chains of dioleoyl phospholipids relative to saturated lipids (Yang & Lyman, 2019).

But is the preference for acyl chain unsaturation coupled with receptor activity? In the simplest model for GPCR activation, the receptor is in a conformational equilibrium between an inactive “R” state and an active “R*” state. In this model, anything which shifts the R/R* equilibrium will modulate activity (Leff, 1995). Thus if the R and R* states recruit different lipid environments, changes in lipid solvation will influence the R/R* equilibrium and modulate activity. In more complex models of receptor function, the receptor can adopt multiple states which couple to different pathways (e.g., different G-proteins or arrestins) (Nygaard et al., 2013), and the extent to which the lipid environment shifts the balance of these multiple states will modulate signaling output.

The present work addresses this question by all-atom simulations of the inactive and partially-active A_{2A}R in a mixture of dipalmitoylphosphatidylcholine (DPPC), cholesterol, and dioleoylphosphatidylcholine (DOPC). Statistically significant differences in lipid solvation between the inactive and partially-active (agonist bound, but without G-protein) receptors are found. The effect of these differences is to shift the R $\leftrightarrow$ R* equilibrium, with the active receptor recruiting on average three more unsaturated lipids to its first solvation shell than the inactive receptor. Thermodynamic analysis of the first solvation shell indicates this represents a difference in free energy of about -0.32 kcal/mol. The consequences of state-dependent lipid solvation are analyzed within the context of a thermodynamic model that incorporates ligand binding, the R $\leftrightarrow$ R* conformational equilibrium, and lipid solvation.

Fluctuations in the lipid composition of the first solvation shell are significant even during 15- μ sec all-atom simulations (Yang & Lyman, 2019), but statistically reliable estimates of the lipid solvation free energy require determining the equilibrium composition of the first shell with high precision. To address this problem, a new coarse-grained transition matrix (CGTM) algorithm is developed specifically to equilibrate the lipid environment around integral membrane proteins. The CGTM method permits a rigorous assessment of statistical uncertainty, and can readily be applied to simulations of other membrane proteins.

Results

Equilibrium distribution of the first lipid solvation shell of A_{2A}R by the CGTM approach

The 15 μ sec inactive and partially-active trajectories in the present work were reported previously (Yang & Lyman, 2019). It was found that the first solvation shell of A_{2A}R is enriched in unsaturated DOPC relative to bulk concentration (see Figure 1A). The previous report included a histogram of first-shell lifetimes, which ranged from 5 nsec to the length of the 15- μ sec simulation in the case of three cholesterol. The average first-shell lifetime for all three lipid types was reported to be $\sim$ 100 nsec.

The scale of fluctuations in Figure 1A suggests significant uncertainty in the average first shell lipid composition. This motivated the construction of a CGTM to compute the equilibrium first shell occupancy. Coarse-grained states were defined by binning the first shell occupancies in increments of 5 mol % of each lipid type (roughly equal to two lipids); the partially-active receptor occupied 63 and the inactive receptor 58 distinct states in the respective 15- μ sec simulations. The need for improved sampling is apparent in Figure 1B, which plots the cumulative probabilities of the five most populated states in the partially-active trajectory, together accounting for nearly 50% of the total probability. Equilibration of these states is not achieved by the end of the 15 μ sec simulation — note the nonzero slopes at the end of the simulation, indicating that the occupancy of these states is still changing in time. The fact that one-half of the total probability is accounted for by only five states also raises the question of whether the sampling is broad enough in the 15 μ sec trajectories.

To address sampling limitations, the CGTM approach was implemented as described in Methods. A total of 144 trajectories per receptor state, each 50 nsec in duration (72 taken from non-

overlapping segments of the 15- μ sec trajectory and 72 seeded from under-sampled states) were used to estimate the transition matrix. The state populations predicted by this transition matrix (denoted π^{Eq}) are reported in Figure 2A (top panel) for the partially active receptor. For comparison, the average occupancies of the same states were obtained directly from the 15- μ sec trajectory (denoted $\pi^{15\mu\text{s}}$) and are shown in Figure 2A (bottom panel). The same comparison is made for the inactive receptor in Fig. 2B. Both states preferentially recruit unsaturated DOPC over saturated DPPC despite the fourfold elevation of DPPC in bulk, as previously reported (Yang & Lyman, 2019). See Materials and Methods for a complete description of molecular dynamics simulations.

Comparison of the top and bottom panels of Figures 2A and 2B shows that the distributions obtained from the CGTM are broader, indicating improved sampling of states that were under-sampled in the 15- μ sec trajectories. However, the average composition (that is, the average ratio of DPPC/DOPC/cholesterol) of the first shell obtained from π^{Eq} and $\pi^{15\mu\text{s}}$ (Table 1) are strikingly similar. The fact that this is the case for simulations of both receptor states suggests that this is not a coincidence, and indicates that if the "brute force" simulations were run longer, A_{2A}R would sample a broader distribution of solvation states, but that the overall composition of the first shell would not change.

The free energy difference to bring a lipid from bulk into the first shell can also be estimated directly from π^{Eq} of each lipid type in the first shell vs. bulk:

$$\Delta G = -k_{\text{B}}T \ln \frac{p_{1\text{st}}}{p_{\text{bulk}}}, \text{ Eq. 1}$$

where $p_{1\text{st}}$ and p_{bulk} are the fractional occupancies of each lipid in the 1st shell or in bulk, and ΔG is computed for DOPC and DPPC separately. The preference of the receptor for unsaturated lipids can then be quantified by the difference in ΔG for the two types of phospholipids:

$$\Delta\Delta G = \Delta G_{\text{unsat}} - \Delta G_{\text{sat}}. \text{ Eq. 2.}$$

A negative $\Delta\Delta G$ represents preference for unsaturated DOPC and a positive $\Delta\Delta G$ preference for saturated DPPC.

$\Delta\Delta G$ for both receptor states are given in Table 1. The partially-active receptor recruits a first solvation shell that is more enriched in DOPC than the inactive receptor. Both receptor states recruit a solvation shell that is depleted in cholesterol relative to bulk. These observations can be made with high confidence thanks to the convergence of first shell occupancies afforded by the CGTM (See Materials and Methods, Supplemental Figures 6-1 and 6-2). This suggests a link between lipid solvation and protein function, which is revisited in the Discussion.

Well-converged lipid solvation was obtained for this system with only 2.4 μ sec of total simulation time: (number of initial states) $\times$ (number of trajectories per state) $\times$ (duration of each trajectory) = $16 \times 5 \times 30 \text{ ns} = 2.4 \mu\text{sec}$ (See Supplemental Figures 6-1 and 6-2). This is less than 1/5 the duration of the 15 μ sec simulations which never reached clear convergence. Moreover, the CGTM approach is naively parallel — all data used to obtain the results presented here were obtained in a single day on a commodity supercomputing resource.

The CGTM model can be visualized as the exchange of probability between different solvation states, drawing the transition matrix as a network of states, with the transition probabilities as the links between states as shown in Figure 3. The network reveals a limited number of "bridge" states that connect states with elevated DPPC to states with elevated DOPC. Figure 2 show that both 15 μ sec simulations were initialized in relatively underpopulated states. The limited number of "bridge" states suggests why equilibration of the first solvation shell requires such a long brute force simulation time

— the system can cycle among elevated DPPC states, for example, for some time before crossing the bridge to sample elevated DOPC states. The network representation also shows how a similar first shell composition is obtained by the long single trajectories and the CGTM method, despite distinct underlying state distributions. Once the long trajectory crosses into the elevated DOPC states (lower left of the panels in Figure 3), it is likely to remain there for an extended period of time, suggesting that the correspondence between the long trajectories and the CGTM model is specific to the topology of the transition matrix, which in turn is likely specific to the membrane composition and perhaps also the protein. It seems impossible to predict this outcome in advance, underlining the value of the CGTM approach.

Thermodynamic modeling of acyl chain solvation

An independent assessment of lipid solvation was determined by estimating the (Gibbs) free energy associated with bringing a single lipid of either type (DOPC or DPPC) from bulk to the receptor's first solvation shell, based on separate calculations of the enthalpic (ΔH) and entropic (ΔS) contributions:

$$\Delta G = \Delta H - T\Delta S. \text{ Eq. 3}$$

In estimating the ΔH term in Eq. 3, only dihedral and non-bonded energies of the acyl chains are considered, since the headgroups of DOPC and DPPC are identical. Details of the calculation are given in Materials and Methods.

Differences in the conformational entropy of the hydrocarbon chains are expected to be an important component of the lipid solvation free energy. The chain conformational entropy ΔS is estimated from the Gibbs entropy:

$$\Delta S = -k_B \sum_i p_i \ln p_i \text{ Eq. 4}$$

where k_B is the Boltzmann constant and p_i is the observed probability of configuration i of the chain. Following Grossfield, et al. (Grossfield et al., 2006), the configurations are enumerated using a rotational isomeric states model in which each dihedral is assigned to a discrete state based on symmetry. For example, a C-C-C-C dihedral could be (t)rans, (g⁺)auche, or (g⁻)auche. A configuration is defined by the sequence of dihedral states, 13 for a palmitoyl chain and 15 for an oleoyl chain, differing from the previous work which grouped sets of three neighboring dihedrals into a single state to reduce the size of the configuration space. Despite the large state space, good convergence is seen for the resulting estimates of S (See Materials and Methods, Supplemental Figure 4-3).

Table 2 gives ΔH and $-T\Delta S$ for each lipid type. They are defined so that negative values indicate a favorable thermodynamic contribution for that lipid to transfer from the bulk to the receptor surface. The entropic contribution is favorable for both lipids, but more so for DPPC. In the simulated mixture, the bulk lipids are more ordered than those at the protein surface, and this is more pronounced for DPPC. However, all 3 enthalpic contributions favor solvation by DOPC (DOPC ΔH are either more negative or are positive but smaller in magnitude than for DPPC). Taken together, the thermodynamics of the hydrocarbon chains yield a favorable free energy for DOPC to partition to the receptor surface in both activation states, but more so for the partially-active state. The thermodynamic analysis is consistent with the partitioning free energies that are obtained directly from relative probabilities for each lipid to be in bulk vs. at the receptor surface, as reported in Table 1. In that case, solvation by DOPC is again favorable, and more so for the partially-active receptor. Absolute numerical agreement is not expected, given that thermodynamic analysis only includes contributions from the hydrocarbon chains. Details of the calculations of ΔH and S are given in Materials and Methods.

Figure 4 illustrates some typical acyl chain conformations at the receptor surface. In some cases, a lipid has one acyl chain in contact with the protein surface and one in contact with other lipids.

The chain that is in contact with the receptor shows increased gauche dihedrals, corresponding to a higher conformational entropy compared to acyl chains in the bulk, which are more ordered on average.

Discussion

Previous analyses of all-atom simulations have characterized localized and specific lipid-GPCR interactions — sites on the protein surface which bind lipids at high affinity compared to the typical lipid-protein interaction (Song et al., 2019; Neale et al., 2015; Cang et al., 2013; Yen et al., 2018; Lee & Lyman, 2013; Lee et al., 2013; Rouviere et al., 2017; Yang & Lyman, 2019). This mode of interaction is simple to detect in simulations, since it yields long-lived lipid binding. It is straightforward to imagine how such interactions may have functional consequences (e.g., via allosteric coupling to ligand or G-protein binding), since they are similar to other small molecule-protein interactions.

Among MD studies which focus on cholesterol, Cang, et al. used density mapping around the β_2 -adrenergic receptor to show competitive and cooperative associations between 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) and cholesterol in 5- μ s all-atom simulations, and observed spontaneous binding of cholesterol at the cholesterol consensus motif (Cang et al., 2013). Lyman and coworkers (Rouviere et al., 2017; McGraw et al., 2019) reported a series of 6- μ s all-atom simulations of A_{2A}R bound to either agonist or antagonist, also observing cholesterol binding to the cholesterol consensus motif, and reporting two new putative cholesterol binding sites. Several studies have characterized anionic lipid interactions with GPCRs at the inner leaflet, including close interactions with PIP₂ that are elevated in the active, G-protein bound state and enhance receptor/G-protein interactions in some cases (Neale et al., 2015; Song et al., 2019; Yen et al., 2018). Relevant to the present work, Grossfield, Feller, and Pitman examined the role of direct interaction of polyunsaturated fatty acids (PUFA) with Rhodopsin, a class-A cousin of A_{2A}R, using a series of short (100 ns) simulations to sample the lipid-protein interactions (Grossfield et al., 2006).

Conceptually contrasting with these previous reports, the present work addresses the functional role of preferential lipid solvation — lower affinity but still selective interactions between lipids and membrane proteins, that lead the protein to recruit a first shell composition that differs from the bulk. A recent paper by Yang and Lyman suggested enrichment of DOPC over DPPC at the surface of A_{2A}R, but a lack of clear equilibration of the A_{2A}R's first lipid solvation shell hampered robust functional inferences (Yang & Lyman, 2019). These observations were based on long (15 μ sec) all-atom simulations, which emphasized the need for an approach that obtains rigorous convergence of the lipid environment around the protein.

This work presents such an approach based on a coarse-grained transition matrix (CGTM) populated by a series of short simulations. The equilibrium distribution predicted by the CGTM yields the composition of the first solvation shell. Because the method is naively parallel, large scale computing resources can be deployed, and results obtained in a matter of days, without requiring specialized hardware or long, contiguous trajectories.

Analysis of the composition of the first solvation shell indicates that the receptor recruits DOPC over DPPC in both the inactive and partially-active states. More importantly, the free energy for DOPC to partition to the first solvation shell is significantly *more favorable* for the partially-active than for the inactive receptor, indicating a mechanism coupling receptor function and membrane composition that is elaborated below.

The state-dependent difference in receptor solvation suggests a functional mechanism based on lipid-dependent shifting of the $R \leftrightarrow R^*$ equilibrium, depicted schematically in Figure 5. The apo receptor fluctuates spontaneously between the inactive R state and the active R* state. (These spontaneous, thermally excited fluctuations underlie the observed constitutive activity of A_{2A}R

(McGraw et al., 2019), and many other receptors.) The fraction of receptors that are in R vs. R* is governed by the free energy difference between these two states:

$$\frac{[R^*]}{[R]} = e^{-\Delta G_{RR^*}/RT} \quad \text{Eq. 5}$$

where $\Delta G_{RR^*} = G_{R^*} - G_R$ is the free energy difference between the R and R* states and $[\]$ denotes the fraction of receptors in the R or R* state. Since it is the R* state that catalyzes nucleotide exchange at the G-protein, this ratio determines the signaling output.

The results presented here indicate that ΔG_{RR^*} is lipid dependent, such that the $R \leftrightarrow R^*$ equilibrium is modulated by the chemical potential of lipids in the bulk. Consider a change in lipid composition which increases the chemical potential of unsaturated relative to saturated chains. This will favor receptor activation, as it drives unsaturated chains to the receptor surface. Because the $R \leftrightarrow R^*$ equilibrium directly impacts experimental observables, the proposed model is experimentally testable. Lipid mixtures that favor R* should increase constitutive signaling and apparent affinity of agonists, while mixtures that shift the equilibrium toward R will suppress constitutive activity and increase the affinity of inverse agonists.

Although the results presented herein are specific to the particular lipid mixture that was simulated, they establish a new conceptual framework for thinking about lipid-dependent membrane protein function. It is not possible to determine how specific effects on activity might change for a different lipid environment without performing simulations in that mixture — this is because the free energy to partition to the receptor surface depends on the chemical potential for each species in the bulk, and lipid chemical potentials are collective properties, determined by the overall composition. However, this challenge is readily overcome using the CGTM approach, which permits rapid sampling of the lipid environment around the receptor. Thus, the results can be easily extended to more realistic mixtures and to other integral membrane proteins. Comparison across systems, especially GPCRs (Sejdiu & Tieleman, 2020), should help to determine generic aspects of lipid-dependent GPCR function.

Materials and Methods

Determining the lipid solvation of A_{2A}R

Lipids were assigned to a solvation shell around A_{2A}R by a Voronoi tessellation based on centers of mass of the lipids and α -carbon atoms of the protein, following Beaven, et al (Beaven et al., 2017). Each lipid was assigned a shell number according to whether it shares a boundary with the receptor (first solvation shell), with a first-shell lipid (second solvation shell), and so on. Lipids in the fifth shell were considered "bulk" because this shell was furthest from the receptor and also contained ratios of lipid types nearest to the whole system averages (within 1%). Because site-specific binding was not considered, both leaflets were averaged together. Figure 6A shows an example of Voronoi boundaries for a single leaflet in a single frame.

The lipid composition of the first shell was characterized by the percent of each lipid type, rounded to the nearest 5% and constrained to total 100%. For example, a first-shell solvation state could be DPPC/DOPC/Cholesterol in a ratio of 0.35/0.45/0.20.

Because this binning does not consider the lipids' spatial distribution, the resulting states are coarse grained. The procedure followed is identical to constructing a Markov State Model, but with coarse graining; hence the name Coarse-Grained Transition Matrix (CGTM) modeling. To compute the equilibrium distribution of states (π^{Eq}), a CGTM was populated by a series of short (50 nsec) trajectories initialized in states selected to span the configuration space. For this analysis, 16 initial states were selected, and replicas were run in each of the 16 states for each receptor state (partially active and inactive). MD simulation details for these short simulations are described in Materials and Methods. Initial configurations were taken from the 15- μ sec simulations for under-sampled states. The

more populated states were adequately represented in the original 15- μ sec simulations; therefore, nonoverlapping segments of the original trajectories were extracted.

Figure 6B illustrates how transitions between states were counted and the CGTM populated. Assuming the system can occupy a state s_i existing in $s = [s_1, s_2 \dots s_N]$, discrete sampling of the system results in a probability distribution $\pi = [\pi_1, \pi_2 \dots \pi_N]$ in which π_i is the sampled probability of the system occupying s_i . Transition probabilities between states are computed from transition events c_{ij} between s_i and s_j :

$$T_{ij} = \frac{c_{ij}}{\sum_j c_{ij}}. \quad \text{Eq. 6}$$

A CGTM $\mathbb{T}$ containing these probabilities describes the evolution of state populations π_i over time:

$$\mathbb{T} = \begin{bmatrix} T_{11} & T_{12} & \dots \\ T_{21} & T_{22} & \\ \vdots & & \end{bmatrix}; \quad \pi_{t+1} = \pi_t \mathbb{T}. \quad \text{Eq. 7}$$

When there is no net population flow among states, the system is in equilibrium. π^{Eq} can be computed by solving the matrix equation:

$$\pi^{\text{Eq}} = \pi^{\text{Eq}} \mathbb{T}. \quad \text{Eq. 8}$$

If there is no net probability flow among states, Eq. 8 holds and the system is in equilibrium. This is guaranteed to be true if microscopic detailed balance is satisfied:

$$\pi_i T_{ij} = \pi_j T_{ji}. \quad \text{Eq. 9}$$

Detailed balance is therefore a sufficient condition for a system to exhibit equilibrium. Detailed balance implies “coarse balance”—balanced flux between every pair of arbitrarily defined set of microstates (Zuckerman, 2010). While coarse graining configuration space results in loss of the Markov property, the coarse balance property implies the reliability of equilibrium properties estimated using trajectories that sample the equilibrium probability distribution. Further verification of the computed π^{Eq} is therefore provided by checking the system for detailed balance and showing the robustness of π^{Eq} as a function of lagtime (time between discrete sampling events).

Convergence of lipid distribution obtained by the CGTM approach

To test convergence of π^{Eq} computed from the CGTM, the number of replicas of short simulations in a given initial state (N_{tr}) was varied from 5 – 9 and the trajectory length (t_{tr}) was varied from 10 – 50 ns. The change in the cumulative root mean squared deviation of the state probabilities were computed:

$$\sigma_{\pi} = \frac{\sum_i \left((\pi_{i,t_{\text{tr},n}}^{\text{Eq}} - \pi_{i,t_{\text{tr},m}}^{\text{Eq}})^2 \right)^{\frac{1}{2}}}{t_{\text{tr},n} - t_{\text{tr},m}}. \quad \text{Eq. 10}$$

Here, $\pi_{i,t_{\text{tr},n}}^{\text{Eq}}$ is the equilibrium probability of state i estimated from trajectory length $t_{\text{tr},n}$ in ns, with $t_{\text{tr},n} > t_{\text{tr},m}$. Convergence of σ_{π} was compared for various values of N_{tr} .

Supplemental Figure 6-1 shows that the state equilibrium probabilities are well converged using the CGTM approach. By the time each trajectory has reached 30 nsec the convergence metric σ_{π} shows that the cumulative state probabilities are changing by about 1%, and including more data (whether in the form of longer individual trajectories or as additional trajectories) does not improve the statistics. Supplemental Figure 6-1 shows that well-converged lipid solvation can be obtained for this system with a total of only 2.4 μ sec of total simulation time: (number of initial states) * $N_{\text{tr}} * t_{\text{tr}} = 16 * 5 * 30 \text{ ns} = 2.4 \mu\text{s}$. This is significantly less than the duration of the initial 15 μ sec simulations which never reach clear convergence. Moreover, the CGTM approach is naively parallel — all data used to produce Figure 2 can easily be obtained in a single day on a commodity supercomputing resource.

While 5 trajectories/initial configuration shows good convergence of π^{Eq} , 9 total replicas were run for each initial configuration. Supplemental Figure 6-2A shows the error in π^{Eq} computed among

blocks of 3 trajectories each; actual error is likely considerably lower, since better convergence is obtained when all 9 trajectories are used to compute π^{Eq} . Additionally, T and the resulting π^{Eq} computed for the agonist- and antagonist-bound receptors were checked for detailed balance. Deviation from Eq. 9 can be computed for each $\pi_i^{\text{Eq}}T_{ij}/\pi_j^{\text{Eq}}T_{ji}$ pair: $\text{abs}(\pi_i^{\text{Eq}}T_{ij} - \pi_j^{\text{Eq}}T_{ji})/(0.5\pi_i^{\text{Eq}}T_{ij} + 0.5\pi_j^{\text{Eq}}T_{ji})$. On average, observed transitions deviated from detailed balance by 5%. Detailed balance was robustly observed for high frequency transitions, with larger deviations seen for infrequent transitions.

Finally, the computed π^{Eq} is independent of lagtime, a necessary but not sufficient condition of a system in equilibrium (Swope et al., 2004). Intuitively, this can be understood as the system exhibiting predictable flow from one state to another despite passing through intermediate states for which the Markov property holds; meaning transitions into and out of those states are also well sampled. While π^{Eq} shown in Figure 2 are computed with a lagtime of 240 ps (the time between subsequent trajectory frames), increasing the lagtime to 960 ps shows no noticeable change in π^{Eq} ; the distributions overlap (Supplemental Figure 6-2B).

Computation and convergence of configurational entropy of acyl chains

A free energy scan about each type of dihedral in the carbon backbone of DOPC and DPPC acyl chains is shown in Supplemental Figure 4-1. DOPC chains contain 3 types of rotatable dihedrals: C-C-C-C, C=C-C-C/C-C-C=C, and C-C=C-C. DPPC acyl chains have one type: C-C-C-C. Relative energy maxima from these plots were taken as boundaries for binning into dihedral states.

Supplemental Figure 4-2 compares the resulting acyl backbone dihedral S of 1st shell and bulk lipids. For both lipid types, configurational entropy of acyl chains is greater at the protein surface than in bulk. Increase in S indicates it is entropically favorable for the acyl chains to neighbor the protein surface because they sample a larger dihedral configuration space. The acyl chains of saturated DPPC, which in the cholesterol-rich bulk are relatively ordered, exhibit the largest change in S moving from bulk to the protein surface. Therefore, while S contributes favorably to the solvation of the protein by both lipid types, it is entropically more favorable for saturated DPPC to solvate the protein.

Each DOPC acyl chain backbone contains 15 dihedrals, and each DPPC backbone 13. The dependence of S on the number of dihedrals considered to be coupled when defining a microstate is plotted in Supplemental Figure 4-3A. Accounting for correlations between dihedral states in the same lipid tail results in a decrease in the estimate of S , as expected — correlation reduces entropy. Notably, estimates of S for the DPPC backbone do not decrease when coupling beyond 1 or 2 neighboring dihedrals ($N = 2$ or $N = 3$) is considered. However, because the DOPC acyl backbone contains multiple types of dihedrals which are each assigned to distinct states and can occupy only specific sites, it is necessary to consider the entire tail to be coupled to get an accurate estimate of S for DOPC.

Supplemental Figure 4-3B shows the convergence of S with increasing sample size. Each point represents the average S computed from different sized non-overlapping trajectory blocks. Good convergence is seen for both lipid types. Convergence is reached more rapidly if fewer neighboring dihedrals are considered to be coupled because this reduces the number of possible states. An upper limit of S can therefore be computed, but given the good convergence seen in Figure 4-3B, the present results consider the entire tail to be coupled.

Molecular dynamics simulations

Initial coordinates for the inactive state of the receptor were based on the 1.8 Å resolution structure bound to ZM241385, Protein Data Bank (PDB) entry [4E1Y](#) (Liu et al., 2012). The partially active state was based on the UK432097-bound structure in which intracellular loop 3 is replaced with a BRIL fusion construct, PDB entry [3QAK](#) (Xu et al., 2011); the BRIL construct was excised, and the

missing residues were rebuilt as described previously (Lee et al., 2013). The protein was embedded in a symmetric lipid bilayer containing 504 DPPC, 273 cholesterol, and 132 DOPC molecules and solvated with approximately 55000 modified TIP3P waters (Jorgensen et al., 1983; Durell et al., 1994), one Na⁺ as observed in the crystal structure, and 10 Cl⁻ to neutralize the system. The lipid mixture (0.55:0.15:0.30 DPPC:DOPC:Chol) is in the liquid–liquid coexistence region of the phase diagram. The protein and lipids were modeled with the CHARMM36 force field (Klauda et al., 2010; Huang et al., 2017; Watts et al., 2018), and the ligand was modeled with the CHARMM general force field (Vanommeslaeghe et al., 2010) with atoms typed by the ParamChem server (Vanommeslaeghe & MacKerell, 2012).

Production simulations were run on Anton2. Integration was performed under constant pressure (1 atm), temperature (295 K), and particle number with the multigrator (Lippert et al., 2013) method with a 2.5 fsec time step. Temperature was controlled by a Nose-Hoover (Nosé, 1984) thermostat coupled every 24 time steps and pressure controlled by the Martyna–Tobias–Klein barostat (semi-isotropic) coupled every 480 time steps (Martyna et al., 1994). Electrostatics were computed using the *k*-space Gaussian split Ewald method (Shan et al., 2005) with long-range interactions computed every third time step and hydrogens constrained by M-SHAKE (Kräutler et al., 2001). Lennard-Jones interactions were cut off at ≥ 11 Å. Production simulations were run for 15 μ sec, and configurations were stored every 240 psec.

Additionally, to verify the equilibrium lipid solvation of A_{2A}R, a series of 50-nsec simulations ("short simulations") were run using GROMACS 2019.4 simulation package (Abraham et al., 2015). The simulations were performed in the NPT ensemble at 1 bar and 295 K with the Nosé–Hoover thermostat (Hoover, 1985; Nosé, 1984; Nosé, 1984) and Parrinello–Rahman barostat (Parrinello & Rahman, 1980; Parrinello & Rahman, 1981). A coupling constant of $\tau_p = 5.0$ psec was used for pressure control, and for temperature $\tau_t = 1$ psec. The protein, lipid bilayer, and water were separately coupled to the heat bath, and the semi-isotropic pressure coupling was applied separately in the *xy*-direction (bilayer plane) and the *z*-direction (bilayer normal). Electrostatics were computed with the particle-mesh Ewald method with a real-space cutoff of 9 Å, Lennard-Jones interactions were truncated at 11 Å, and configurations were stored every 240 psec.

Images of lipids and receptors were generated and values of acyl chain dihedrals determined with the Visual Molecular Dynamics program (Humphrey et al., 1996). Internal energies and nonbonded interaction energies were computed using the CHARMM program (Brooks et al., 2009).

Acknowledgements

The authors thank Alex Sodt and Dan Zuckerman for extensive discussion regarding implementation of the CGTM, and Alex for assistance with Voronoi analysis. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) supported by National Science Foundation Grant ACI-1548562. Anton2 computer time was provided by the Pittsburgh Supercomputing Center (PSC) through Grant R01GM116961 from the National Institutes of Health. The Anton2 machine was made available by D. E. Shaw Research

References

- Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S. r.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, 1-2, 19.
- Beaven, A. H.; Maer, A. M.; Sodt, A. J.; Rui, H.; Pastor, R. W.; Andersen, O. S.; Im, W. Gramicidin A Channel Formation Induces Local Lipid Redistribution I: Experiment and Simulation. *Biophys. J.* **2017**, 112 (6), 1185.

- Bolla, J. R.; Agasid, M. T.; Mehmood, S.; Robinson, C. V. Membrane Protein–Lipid Interactions Probed Using Mass Spectrometry. *Annu. Rev. Biochem.* **2019**, *88* (1), 85.
- Botelho, A. V.; Gibson, N. J.; Thurmond, R. L.; Wang, Y.; Brown, M. F. Conformational energetics of rhodopsin modulated by nonlamellar-forming lipids. *Biochemistry* **2002**, *41* (20), 6354.
- Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30* (10), 1545.
- Cang, X.; Du, Y.; Mao, Y.; Wang, Y.; Yang, H.; Jiang, H. Mapping the functional binding sites of cholesterol in β 2-adrenergic receptor by long-time molecular dynamics simulations. *J. Phys. Chem. B* **2013**, *117* (4), 1085.
- Doralicia, C.; Pablo, V. E.; Catalina Ana, R. In *Int. J. Mol. Sci.*, 2019; Vol. 20.
- Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-Induced Forces between Two Hydrophilic Groups. *The Journal of Physical Chemistry* **1994**, *98* (8), 2198.
- Feller, S. E.; Gawrisch, K.; Woolf, T. B. Rhodopsin exhibits a preference for solvation by polyunsaturated docosahexaenoic acid. *J. Am. Chem. Soc.* **2003**, *125* (15), 4434.
- Gonen, T.; Cheng, Y.; Sliz, P.; Hiroaki, Y.; Fujiyoshi, Y.; Harrison, S. C.; Walz, T. Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* **2005**, *438* (7068), 633.
- Grossfield, A.; Feller, S. E.; Pitman, M. C. Contribution of omega-3 fatty acids to the thermodynamics of membrane protein solvation. *J. Phys. Chem. B* **2006**, *110* (18), 8907.
- Grossfield, A.; Feller, S. E.; Pitman, M. C. A Role for Direct Interactions in the Modulation of Rhodopsin by ω -3 Polyunsaturated Lipids. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (13), 4888.
- Gutierrez, M. G.; Deyell, J.; White, K. L.; Dalle Ore, L. C.; Cherezov, V.; Stevens, R. C.; Malmstadt, N. The lipid phase preference of the adenosine A2A receptor depends on its ligand binding state. *Chemical communications (Cambridge, England)* **2019**, *55* (40), 5724.
- Gutierrez, M. G.; Mansfield, K. S.; Malmstadt, N. The Functional Activity of the Human Serotonin 5-HT1A Receptor Is Controlled by Lipid Bilayer Composition. *Biophys. J.* **2016**, *110* (11), 2486.
- Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Stevens, R. C.; Kuhn, P. A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β 2-Adrenergic Receptor. *Structure* **2008**, *16* (6), 897.
- Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985**, *31* (3), 1695.
- Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14* (1), 71.
- Huber, T.; Botelho, A. V.; Beyer, K.; Brown, M. F. Membrane model for the G-protein-coupled receptor rhodopsin: hydrophobic interface and dynamical structure. *Biophys. J.* **2004**, *86* (4), 2078.
- Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33.
- Javanainen, M.; Enkavi, G.; Guixà-González, R.; Kulig, W.; Martinez-Seara, H.; Levental, I.; Vattulainen, I. In *PLoS Comp. Biol.*, 2019; Vol. 15.
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79* (2), 926.
- Katritch, V.; Cherezov, V.; Stevens, R. C. Structure-Function of the G Protein–Coupled Receptor Superfamily. *Annu. Rev. Pharmacol. Toxicol.* **2013**, *53* (1), 531.
- Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **2010**, *114* (23), 7830.

- Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **2001**, 22 (5), 501.
- Lee, J. Y.; Lyman, E. Predictions for Cholesterol Interaction Sites on the A2A Adenosine Receptor. *J. Am. Chem. Soc.* **2013**, 134 (40), 16512
- Lee, J. Y.; Patel, R.; Lyman, E. Ligand-dependent cholesterol interactions with the human A2A adenosine receptor. *Chem. Phys. Lipids* **2013**, 169, 39.
- Leff, P. The two-state model of receptor activation. *Trends Pharmacol. Sci.* **1995**, 16 (3), 89.
- Lippert, R. A.; Predescu, C.; Ierardi, D. J.; Mackenzie, K. M.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure. *The Journal of chemical physics* **2013**, 139 (16), 164106.
- Liu, W.; Chun, E.; Thompson, A. A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G. W.; Heitman, L. H.; Ijzerman, A. P.; Cherezov, V. et al. *Crystal structure of the chimeric protein of A2aAR-BRIL in complex with ZM241385 at 1.8 Å resolution*; Protein Data Bank, Rutgers University, 2012.
- Lorent, J. H.; Levental, K. R.; Ganesan, L.; Rivera-Longsworth, G.; Sezgin, E.; Doktorova, M.; Lyman, E.; Levental, I. Plasma membranes are asymmetric in lipid unsaturation, packing and protein shape. *Nat. Chem. Biol.* **2020**, 16 (6), 644.
- Loura, L. M.; Prieto, M.; Fernandes, F. Quantification of protein-lipid selectivity using FRET. *Eur. Biophys. J.* **2010**, 39 (4), 565.
- Marsh, D. Lipid-protein interactions in membranes. *FEBS Lett.* **1990**, 268 (2), 371.
- Marsh, D.; Watts, A.; Pates, R. D.; Uhl, R.; Knowles, P. F.; Esmann, M. ESR spin-label studies of lipid-protein interactions in membranes. *Biophys. J.* **1982**, 37 (1), 265.
- Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994**, 101 (5), 4177.
- McGraw, C.; Yang, L.; Levental, I.; Lyman, E.; Robinson, A. S. Membrane cholesterol depletion reduces downstream signaling activity of the adenosine A2A receptor. *Biochim. Biophys. Acta Biomembr.* **2019**, 1861 (4), 760.
- Montenegro, F. A.; Cantero, J. R.; Barrera, N. P. Combining Mass Spectrometry and X-Ray Crystallography for Analyzing Native-Like Membrane Protein Lipid Complexes. *Front. Physiol.* **2017**, 8, 892.
- Naranjo, A. N.; McNeely, P. M.; Katsaras, J.; Robinson, A. S. Impact of purification conditions and history on A2A adenosine receptor activity: The role of CHAPS and lipids. *Protein Expression Purif.* **2016**, 124, 62.
- Neale, C.; Herce, Henry D.; Pomès, R. g.; García, Angel E. Can Specific Protein-Lipid Interactions Stabilize an Active State of the Beta 2 Adrenergic Receptor? *Biophys. J.* **2015**, 109 (8), 1652.
- Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984**, 81 (1), 511.
- Nosé, S. i. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics* **1984**, 52 (2), 255.
- Nygaard, R.; Zou, Y.; Dror, R. O.; Mildorf, T. J.; Arlow, D. H.; Manglik, A.; Pan, A. C.; Liu, C. W.; Fung, J. J.; Bokoch, M. P. et al. The dynamic process of $\beta(2)$ -adrenergic receptor activation. *Cell* **2013**, 152 (3), 532.
- O'Malley, M. A.; Helgeson, M. E.; Wagner, N. J.; Robinson, A. S. The morphology and composition of cholesterol-rich micellar nanostructures determine transmembrane protein (GPCR) activity. *Biophys. J.* **2011**, 100 (2), 11.
- Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters* **1980**, 45 (14), 1196.
- Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **1981**, 52 (12), 7182.

- Periole, X. Interplay of G Protein-Coupled Receptors with the Membrane: Insights from Supra-Atomic Coarse Grain Molecular Dynamics Simulations. *Chem. Rev.* **2017**, *117* (1), 156.
- Pradas, I.; Huynh, K.; Cabré, R.; Ayala, V.; Meikle, P. J.; Jové, M.; Pamplona, R. Lipidomics Reveals a Tissue-Specific Fingerprint. *Front. Physiol.* **2018**, *9*, 1165.
- Řezanka, T.; Kolouchová, I.; Gharwalová, L.; Palyzová, A.; Sigler, K. Lipidomic Analysis: From Archaea to Mammals. *Lipids* **2018**, *53* (1), 5.
- Rouviere, E.; Arnarez, C.; Yang, L.; Lyman, E. Identification of Two New Cholesterol Interaction Sites on the A2A Adenosine Receptor. *Biophys. J.* **2017**, *113* (11), 2415.
- Sejdiu, B. I.; Tieleman, D. P. Lipid-Protein Interactions Are a Unique Property and Defining Feature of G Protein-Coupled Receptors. *Biophys. J.* **2020**, *118* (8), 1887.
- Shan, Y.; Klepeis, J. L.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *The Journal of chemical physics* **2005**, *122* (5), 54101.
- Song, W.; 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* **2019**, *27* (2), 392.
- Soubias, O.; Gawrisch, K. Probing specific lipid-protein interaction by saturation transfer difference NMR spectroscopy. *J. Am. Chem. Soc.* **2005**, *127* (38), 13110(2).
- Soubias, O.; Gawrisch, K. The role of the lipid matrix for structure and function of the GPCR rhodopsin. *Biochim. Biophys. Acta Biomembr.* **2012**, *1818* (2), 234.
- Soubias, O.; Niu, S. L.; Mitchell, D. C.; Gawrisch, K. Lipid-rhodopsin hydrophobic mismatch alters rhodopsin helical content. *J. Am. Chem. Soc.* **2008**, *130* (37), 12465.
- Swope, W. C.; Pitera, J. W.; Suits, F. Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 1. Theory †. *The Journal of Physical Chemistry B* **2004**, *108* (21), 6571.
- Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I. et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31* (4), 671.
- Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *Journal of chemical information and modeling* **2012**, *52* (12), 3144.
- Watts, C. R.; Gregory, A.; Frisbie, C.; Lovas, S. n. Effects of force fields on the conformational and dynamic properties of amyloid β (1-40) dimer explored by replica exchange molecular dynamics simulations. *Proteins* **2018**, *86* (3), 279.
- Xu, F.; Wu, H.; Katritch, V.; Han, G. W.; Cherezov, V.; Stevens, R.; Network, G. *Agonist bound structure of the human adenosine A2a receptor*; Protein Data Bank, Rutgers University, 2011.
- Yang, L.; Lyman, E. Local Enrichment of Unsaturated Chains around the A2A Adenosine Receptor. *Biochemistry* **2019**, *58* (39), 4096.
- Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochim. Biophys. Acta Biomembr.* **2014**, *1838* (6), 1548.
- Yen, H. Y.; Hoi, K. K.; Liko, I.; Hedger, G.; Horrell, M. R.; Song, W.; Wu, D.; Heine, P.; Warne, T.; Lee, Y. et al. PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **2018**, *559* (7714), 423.
- Zuckerman, D. M. *Statistical physics of biomolecules : an introduction*; CRC Press/Taylor & Francis: Boca Raton, FL, 2010.

Figure 1

Figure 1. First lipid solvation shell of receptor does not converge over a 15 μ sec all-atom simulation. Panel A: Composition of the first lipid solvation shell; agonist bound receptor (similar results are obtained for the inactive receptor). The number of each lipid type in the first solvation shell is obtained by the Voronoi analysis described in Methods. Red, DPPC; blue, DOPC; green, cholesterol. Horizontal dashed lines show the average over entire simulation of the first shell, solid lines show the bulk composition. Both leaflets are included. Panel B: Probabilities of the 5 most populated states over time; agonist-bound receptor. Probability is integrated over time with 125 evenly-spaced sampling events. The following states are plotted with DPPC/DOPC/Chol compositions given as percentages: 1st 35/50/15; 2nd 30/55/15; 3rd 40/45/15; 4th 40/40/20; 5th 35/45/20.

Figure 2

Figure 2. Comparison of occupation probabilities π^{Eq} computed from CGTM, and raw occupancies $\pi^{15\mu s}$ computed from 15- μ sec trajectory. Blue: DOPC; red: DPPC; yellow: cholesterol. Each state is a unique ratio of DPPC/DOPC/Cholesterol, binned in 5 mol % increments. Panels A (top and bottom), partially active receptor. Panels B (top and bottom), inactive receptor.

Figure 3

Figure 4

Figure 4. Lipid tails conforming to the receptor surface. Red, DPPC; blue, DOPC; white, cholesterol. Note the conformation of the acyl chains, many of which are contorted as they conform to the receptor surface. Only a subset of first shell lipids are shown for clarity. Partially-active receptor.

Figure 5

Figure 5. Activation energy depends on lipid solvation. In membrane composition “A,” the $R \rightarrow R^*$ equilibrium (left to right arrow; inactive receptor in dark blue, active receptor in light blue) is determined by the free energy difference between R and R^* , ΔG_{RR}^A , which differs in composition “B.” The vertical arrows indicate the free energy associated with a change in receptor environment in each state (e.g., as listed in Table 1). Differences in the R/R^* equilibrium are observable as differences in constitutive activity and ligand affinities.

Figure 6

Figure 6. Method for constructing CGTM. Panel A: Example of Voronoi tessellation of the on lea left. Voronoi cells for DOPC (blue), DPPC (red), and cholesterol (yellow). Agonist-bound receptor. Panel B: Schematic for definition of transitions used to populate the CGTM. Short simulations were initialized in states selected to span configuration space. Here, "state 13" is the initial state. Transitions were tallied with inclusion of redundancy. To test convergence, the number of replicas in each initial state (N_{tr}) was varied, as well as the length of the trajectories (t_{tr}).

Supplemental Figure 6-1

Supplemental Figure 6-1. Convergence of π^{Eq} is achieved with modest sampling. The convergence metric (σ_{π} , Eq. 10) demonstrates that the state probabilities are well converged after each individual trajectory reaches 30 nsec, using 5 independent trajectories per initial state. Blue: 5 trajectories per initial state. Red: 9 trajectories per initial state. Agonist-bound receptor.

Supplemental Figure 6-2

Supplemental Figure 6-2. Statistical reliability of π^{Eq} . Panel A: Error in π^{Eq} computed by block averaging. Standard error in π^{Eq} computed from 3 sets of 3 trajectories each per initial configuration for the partially active state (top) and inactive state (bottom). Panel B: Equivalence of π^{Eq} computed using different lagtimes. Only slight variation in state probabilities are seen if π^{Eq} is computed with a lagtime of 240, 480, 720, or 960 ps. π^{Eq} is independent of lagtime in equilibrated systems. Agonist receptor; computed with overlapping trajectory windows. Similar equivalence was exhibited by antagonist-bound receptor.

Supplemental Figure 4-1

Supplemental Figure 4-1. DOPC and DPPC free energy landscapes for various dihedral types.

Dihedral types in acyl chains; example minimum energy configurations shown with dihedrals highlighted in black. (left) DOPC dihedrals include (top) C-C-C-C, maxima at 0° and $\pm 130^\circ$; (middle) C-C-C=C, maxima at 180° and $\pm 45^\circ$; and (bottom) C-C=C-C, maxima at $\pm 100^\circ$. (right) DPPC includes C-C-C-C, maxima at 0° and $\pm 130^\circ$.

Supplemental Figure 4-2

Supplemental Figure 4-2. Acyl chain dihedral entropy (S) of first shell and bulk lipids. Error computed from block averaging is on the order of 1%; therefore, error bars are not shown. For both receptors, entropy favors DPPC partitioning from bulk to the first shell significantly more than DOPC.

Supplemental Figure 4-3

Supplemental Figure 4-3. Convergence of acyl backbone dihedral entropy (S). Panel A:

Dependence of S on the number of neighboring dihedrals (N) considered to be coupled. Sampled from simulation of agonist-bound receptor, *sn*-2 chain. Error computed from block averaging is on the order of 1%; therefore error bars are not shown. Panel B: Convergence of S as a function of trajectory length. Plotted as a function of length of trajectory used in block averaging. Two data sets are included for DOPC, one with the entire tail considered to be coupled, and one with 13 of the 15 dihedrals considered to be coupled; $N = 13$ or 15 respectively. Sampled from simulation of agonist-bound receptor, *sn*-1 chain. Error computed from block averaging is on the order of 1%; therefore error bars are not shown.

Table 1

		DPPC	DOPC	Chol	
Agonist	$\pi^{15\mu s}$	36.9%	45.9%	17.2%	
	π^{Eq}	35.9±0.4%	46.7±0.5%	17.4±0.2%	
	π^{Bulk}	55.5%	14.3%	30.2%	$\Delta\Delta G$ (Eqs. 1 & 2) DOPC – DPPC
	ΔG , Eq. 1 [kcal/mol]	0.24±0.01	-0.68±0.01	0.33±0.01	-0.92±0.01
Antagonist	$\pi^{15\mu s}$	49.8%	37.3%	12.9%	
	π^{Eq}	46.7±0.5%	38.8±0.4%	14.5±0.1%	
	π^{Bulk}	55.3%	14.8%	30.0%	$\Delta\Delta G$ (Eqs. 1 & 2) DOPC – DPPC
	ΔG , Eq. 1 [kcal/mol]	0.06±0.01	-0.54±0.01	0.49±0.01	-0.60±0.01

Table 1. Fraction of each lipid type in first solvation shell vs. bulk, and free energies to transfer from bulk to first shell. Raw first-shell occupancies from 15- μ sec simulations ($\pi^{15\mu s}$) and equilibrium distribution computed by the CGTM approach (π^{Eq}), compared with bulk distribution (π^{Bulk}). Free energies obtained from π^{Eq} values (Eqs. 1 and 2). Errors in π^{Eq} utilize convergence of probabilities to within 1% (See Materials and Methods, Supplemental Figure 6-1).

Table 2

	Bulk → First Shell Contribution [kcal/mol]	DOPC	DPPC	
Agonist	ΔH, dihedrals	0.055 ± 0.02	0.187 ± 0.02	
	ΔH, Lennard-Jones	14.85 ± 0.06	15.24 ± 0.07	
	ΔH, Coulomb	-6.738 ± 0.05	-4.851 ± 0.06	
	$-T\Delta S$, dihedrals	<u>$+ -0.543 \pm 0.005$</u>	<u>$+ -2.525 \pm 0.007$</u>	$\Delta\Delta G$ (Eqs. 3 & 2) DOPC - DPPC
	ΔG, sum, Eq. 3	7.62 ± 0.08	8.05 ± 0.09	-0.43 ± 0.11
Antagonist	ΔH, dihedrals	0.026 ± 0.02	0.115 ± 0.01	
	ΔH, Lennard-Jones	13.80 ± 0.04	14.97 ± 0.05	
	ΔH, Coulomb	-5.682 ± 0.07	-4.827 ± 0.06	
	$-T\Delta S$, dihedrals	<u>$+ -0.313 \pm 0.004$</u>	<u>$+ -2.162 \pm 0.006$</u>	$\Delta\Delta G$ (Eqs. 3 & 2) DOPC - DPPC
	ΔG, sum, Eq. 3	7.83 ± 0.08	8.10 ± 0.08	-0.27 ± 0.10

Table 2. Contributions to $\Delta\Delta G$ to bring a lipid molecule from bulk to first solvation shell.