Article

Insights into Membrane Protein−Lipid Interactions from Free Energy Calculations

Robin A. Corey,[†] Owen N. Vickery,[†] Mark S. P. Sansom,[ID] and Phillip J. Stansfeld*[ID]

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Ⓢ *Supporting Information*

ABSTRACT: Integral membrane proteins are regulated by specific interactions with lipids from the surrounding bilayer. The structures of protein−lipid complexes can be determined through a combination of experimental and computational approaches, but the energetic basis of these interactions is difficult to resolve. Molecular dynamics simulations provide the primary computational technique to estimate the free energies of these interactions. We demonstrate that the energetics of protein−lipid interactions may be reliably and reproducibly calculated using three simulation-based approaches: potential of mean force calculations, alchemical free energy perturbation, and well-tempered metadynamics. We employ these techniques within the framework of a coarse-grained force field and apply them to both bacterial and mammalian membrane protein−lipid systems. We demonstrate good agreement between the different techniques, providing a robust framework for their automated implementation within a pipeline for annotation of newly determined membrane protein structures.
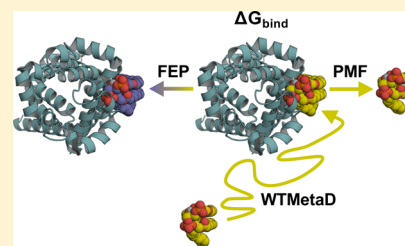
INTRODUCTION

Integral transmembrane proteins have diverse functions within cells, and as such are key targets for many drugs, ranging from antibiotics to anticancer agents. Structurally, they are unified by the presence of a hydrophobic span of residues that both anchors the protein within the core of a lipid bilayer membrane and presents the flanking residues to the surrounding polar lipid head groups. The resulting protein−lipid interactions are important for function, with many membrane proteins, including, for example ion channels, transporters, and receptors, regulated by specific lipid interactions.[1] Lipid-binding sites thus provide potential druggable allosteric sites on many biologically important membrane proteins.

Structural studies of membrane proteins often rely on their extraction from their native bilayer environment through use of detergents. As a consequence of this, lipids which bind to the protein are often lost before structural (X-ray diffraction or cryoelectron microscopy) data are gathered. Although there are cases where X-ray or electron scattering density may be observed for lipids bound to membrane proteins (for examples, see refs [2−4]), the often modest resolution of such data presents challenges to the unambiguous assignment of the molecular identity of the bound lipid species.

Molecular simulations provide high resolution insights into the interactions of lipids with membrane proteins. They can both predict the location of lipid-binding sites in advance of structural studies[5−7] and can extend structural observations on the lipid interactions of a given membrane protein to other members of a protein family.[8] In addition to identification of potential lipid interaction sites, for example, from estimates of lipid−protein "fingerprints",[9] molecular simulations can provide estimates of the residence times of lipids at binding

sites on a membrane protein[10] and of free energies of interaction of specific lipids.[11,12]

Validation of computational predictions of specific lipid interactions can be achieved via a number of biophysical approaches, including, for example, native mass spectrometry (nMS)[13] which can be employed in tandem with molecular simulation.[14] The relatively slow throughput of these techniques, however, means that only a tiny fraction of the possible interactions has so far been identified. Moreover, experimental quantification of the strength and specificity of protein−lipid interactions remains more challenging, with notable recent attempts using nMS[15] and surface plasmon resonance (SPR)-based methods.[16]

Molecular simulations can also be used to quantify the strength of protein−lipid interactions, via free energy calculations (Figure 1A). Several free energy techniques have been developed for the calculation of binding free energies between ligands and (water soluble) proteins,[17] and these can be modified for analysis of protein−lipid interactions. Membrane proteins and lipids pose particular challenges of sampling and convergence for accurate free energy estimation,[18] arising from the relatively slow rates of lipid diffusion and from the diversity of lipid species present in complex biological membranes.[19] To date, most studies[5,11,18,20,21] have combined umbrella sampling with a potential of mean force (PMF) calculation along a one-dimensional reaction coordinate connecting the binding site with the surrounding membrane[18] (Figure 1B). Convergence of such calculations (i.e. the point at which additional sampling via additional simulation does not substantially change the outcome) is often
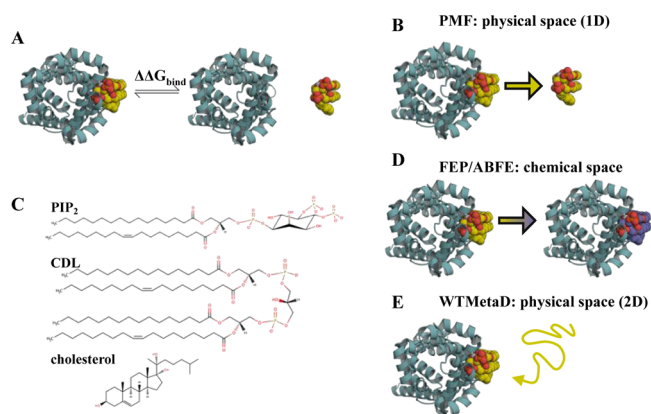
Figure 1. Introduction to free energy calculations. (A) Overview of free energy calculations. A membrane protein, as viewed from above the membrane, is shown in cyan cartoon, and a lipid in yellow, orange and red spheres. Two states are modelled: the left state is the protein bound to the target lipid, the right is the protein bound to a generic lipid (not shown), with the target lipid unbound. Free energy calculations aim to compute the difference in free energy between these states ($\Delta\Delta G_{bind}$). (B) PMF calculations create a reaction coordinate in physical space by pulling the lipid away from or towards the binding site. This coordinate can then be sampled, for example with umbrella sampling, to provide a 1D energetic landscape, allowing calculation of $\Delta\Delta G_{bind}$ between the target and a generic lipid. (C) Chemical structures of PIP$_2$, CDL and cholesterol. (D) FEP and ABFE calculations build alchemical pathways which either change the bound lipid into a different species, in this study to that of the bulk membrane, or fully remove the lipid from the simulation box. This provides the binding free energy difference between the target lipid and a generic lipid, $\Delta\Delta G_{bind}$. (E) WTMetaD biases the diffusion of a target lipid around the protein through addition of a time-dependent Gaussian of energy to the CV. These Gaussians can then be reconstructed into a full 2D energy landscape, with comparison of binding regions and the bulk membrane giving $\Delta\Delta G_{bind}$.

achieved through use of a coarse-grained (CG) biomolecular force field, such as Martini,[22,23] which allows for efficient sampling of molecular systems. While a powerful technique, the difficulty in demonstrating convergence makes this approach challenging to implement in a high throughout fashion. Furthermore, it is computationally demanding, requiring >50 μs of simulation per protein−lipid interaction, currently equivalent to ∼1−2 weeks on a typical GPU-node. It is therefore important that we explore additional approaches in order to extend the reach and to evaluate the robustness of PMF-based estimates of free energies of protein−lipid interactions.

Here, we present an analysis of the determination of protein−lipid binding interactions using PMFs alongside two other powerful free energy approaches, adapted here for investigation of protein−lipid interactions. These are free energy perturbation (FEP[24]) or absolute binding free energy (ABFE[25]) calculations, whereby a molecule is partially or fully perturbed via non-natural (i.e., alchemical) chemical space (Figure 1D), and well-tempered metadynamics (WTMetaD[26]), where a history-dependent bias is added to a free energy surface (FES) (Figure 1E) to reduce simulation time spent sampling local energy minima. We compare PMF, FEP/ABFE, and WTMetaD in terms of ease of accuracy and computational cost. We use all three methods on a panel of experimentally well-characterized proteins which are representative of bacterial, mitochondrial, and mammalian cell

membranes. Through comparison of the methods applied here we outline a mechanism for producing robust and reproducible estimates of protein−lipid interactions from molecular simulations.

■ METHODS

Equilibrium CG Simulations. Simulations were run using the CG Martini v2 biomolecular force field.[22,23] In this forcefield, molecules are coarse grained through the representation of approximately 4 heavy atoms and associated hydrogens as a single bead or particle. While this simplification provides the force field with reduced resolution,[27] it has repeatedly been shown as highly proficient in the identification of specific interactions between proteins and membrane lipids, including cardiolipin (CDL),[11,21,28,29] phosphatidylinositol (4,5) bisphosphate (PIP$_2$),[5,14,20] cholesterol[20,31] (Figure 1C), and others. The Martini force field also restricts internal protein dynamics, reducing concerns about potential conformational changes caused by, for example, ion binding or residue protonation.

We follow the MemProtMD protocol for setting up CG simulations of integral membrane proteins in an equilibrated bilayer.[32,33] Briefly, the input proteins are aligned accordingly on the xy plane, and lipids (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; POPC or 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine; POPE) are placed randomly around the transmembrane region of protein, in a z range of 8 nm. The starting protein coordinates were used as follows: chicken (*Gallus gallus*) Kir2.2 ion channel, PDB code 3SPC;[34] bovine (*Bos taurus*) AAC transport protein, PDB code 1OKC;[4] bacterial (*Aquifex aeolicus*) leucine transporter (LeuT) transport protein, PDB code 2A65[35] (simulated here as a dimer); and human (*Homo sapiens*) GPCR adenosine 2A receptor (A$_{2A}$R), PDB code 5IU4[36] with the bRIL subunit removed. In each case, nonprotein atoms were removed and any missing loops were added using MODELLER[37] or SWISS-MODEL[38] prior to CG conversion. Phospholipids were modeled with palmitoyl (4-beads) and oleoyl (5-beads) tails. Cholesterol was modeled using the virtual-site description.[39] Full details of the simulations are given in the Supporting Methods.

Potential of Mean Force Calculations. PMF calculations were set up and run as described previously.[18] Calculations start from the complex formed between the protein and target lipid, which is then inserted into a generic membrane (i.e., POPE or POPC). The lipid is then removed from the protein through application of steered MD, in which a distance-dependent pulling force is applied between the lipid head group and the protein. This trajectory then forms the collective variable (CV) to be analyzed. The specific CVs used here are outlined in the Supporting Methods.

Snapshots were taken of the system in which the lipid is at specific window along the CV (using a 0.05 nm spacing for optimal histogram overlap, see e.g. Supporting Figure 1), with each snapshot used to seed an independent simulation. For these, an umbrella potential with a force of 1000 kJ mol^{-1} nm^{-2} was used to keep the lipid in place along the reaction coordinate, with 100 kJ mol^{-1} nm^{-2} xy positional restraints on 3−4 protein backbone beads to prevent the protein from rotating during the simulations—details provided in the Supporting Methods.

Simulations were run for 1 μs to allow convergence (e.g., Supporting Figure 1). The first 200 ns were removed from each simulation as equilibration, and the final 800 ns were

combined into a 1D energy profile using the weighted histogram analysis method,[40] as implemented in *gmx wham*,[41] and employing 200 rounds of Bayesian bootstrapping to report on statistical accuracy. When plotting, the bulk region of the membrane is considered to have a free energy of 0 kJ mol^{-1}, and the binding energy well is set to 0 nm on the x-axis.

Note that, as the lipid binding site will either be occupied by the target lipid or a generic bulk lipid, this analysis will not provide us with ΔG_{bind} of our target lipid to the site, but instead $\Delta\Delta G_{bind}$ between the target and a generic lipid. This is the biologically appropriate term, as protein–lipid binding will always occur in competition with other lipids, and the effective affinity of the interaction will be dependent on the nature of the other lipids present. Accordingly, if you carry out PMF calculations of, for example, POPC in a POPC membrane, the value reported should be 0 kJ mol^{-1} (e.g. see ref 21).

Free Energy Perturbation. The bound PMF systems were used as the input for the FEP calculations. The target lipid was alchemically transformed into the generic lipid, along a coordinate in chemical space, termed λ. Additional simulations were run perturbing the target lipid in the bulk membrane ("free"), that is, with no protein visible to the lipid. $\Delta\Delta G_{bind}$ can then be calculated as described in Figure 2. This value should be equivalent to that obtained using PMF calculations.

Charges and Lennard Jones interactions were turned off separately, with a softcore parameter (sigma = 0.3) used for the Lennard Jones interactions.[42,43] Bonded interactions were not perturbed as these cancel out between the bound and free states. For each system, we used a single topology method, where molecule A is converted into molecule B in a single transformation. Details of the transformation are provided in the Supporting Methods. In all cases, we used 10 windows to perturb the Columbic interactions, and either 10 (PIP$_2$) or 20 (CDL) windows were used to perturb the Lennard-Jones interactions.

Each λ window was minimized using the steepest descent method, followed by 5 ns of *NPT* equilibrium. Five independent production simulations of 250 ns were then run with randomized initial velocities, using a leap-frog stochastic dynamics integrator.

The free energy pathway was then constructed from the individual λ windows using the Alchemical Analysis package.[44] Energy values were calculated on the final 225 ns of simulation data, with this simulation length showing good conversion (e.g., Supporting Figure 2). Analysis was run using the MBAR method,[45] although we observed good agreement between multiple analysis methods (e.g., Supporting Figure 2). Data from 5 repeats were averaged, and the standard deviations calculated.

Absolute Binding Free Energy Calculations. A$_{2A}$R-cholesterol simulations followed a similar setup as described above, but with the target lipid fully, rather than partially, decoupled from the simulation box (Supporting Figure 3). To keep the cholesterol molecule in the binding site and within the correct plane and orientation of the membrane at high values of λ, we followed a cholesterol-restraining scheme described previously,[46] and outlined in the Supporting Methods. $\Delta\Delta G_{bind}$ was calculated from the energy required to decouple the cholesterol molecule in the bound and free state, in each case perturbing the Lennard-Jones interactions over 29 λ windows (with 0.05 spacing from 0 to 0.6, and 0.025 spacing from 0.6 to 1). The energetic input of the restraints
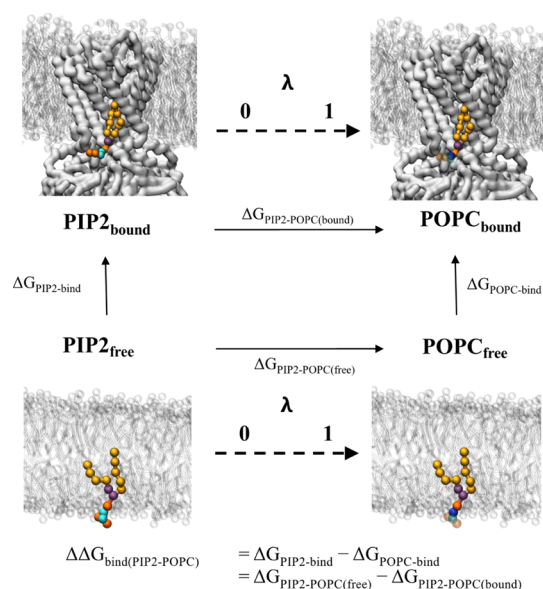

Figure 2. Thermodynamic cycle for FEP shown is a representative thermodynamic cycle used for the FEP calculations in this study. The protein receptor (here Kir2.2) is shown as white surface, with the membrane as white sticks. The target lipid is shown as coloured spheres: on the left the native PIP$_2$ molecule has lipid tails in yellow, glycerol beads in purple, phosphate beads in orange and sugar beads in cyan. On the right, the perturbed POPC lipid is coloured as PIP$_2$, but with a blue choline bead, and transparent beads for the beads which have now been decoupled from the system. The horizontal vertices represent the two FEP calculations, where the target lipid is alchemically perturbed into the bulk lipid either free in membrane ($\Delta G_{PIP2-POPC(free)}$) or bound to the receptor ($\Delta G_{PIP2-POPC(bound)}$). These reactions are represented by a λ coordinate from 0 to 1. $\Delta\Delta G_{bind}$ can be calculated as $\Delta G_{PIP2-POPC(free)} - \Delta G_{PIP2-POPC(bound)}$, which will represent the same value as calculated in the PMF calculations (Figure 1). Note that the right vertical vertex ($\Delta G_{POPC-bind}$) is the free energy required for the bulk lipid binding the receptor—in this special case of protein–lipid binding, where the ligand and the solvent are the same molecule, this value should be 0 kJ mol^{-1}.

was accounted for through two analytical terms (see Supporting eqs 1−3), and additional restraint FEP in the gas phase (see Supporting Figure 3). Simulations were run in Gromacs 2016 with restraints imposed using the Plumed v2.4 plugin.[47,48]

Metadynamics. In classical MetaD, the evolution of the system along the CV is biased by a history-dependent potential, which is the sum of the Gaussians deposited along the relevant CV.[49] After a defined period of time, the biasing potential compensates the underlying FES, allowing the real FES to be estimated. Use of a fixed Gaussian height gives classical MetaD several limitations, particularly concerning convergence, as the system can be pushed into regions of configurational space which are not physically or physiologically relevant. Here, we use WTMetaD[26,50] in which the Gaussian height is rescaled based upon an adaptive bias (ΔT), which is dependent on the system simulated.[26]

The WTMetaD simulations were made using the protein coordinates built into a simple (POPC or POPE) membrane as described above, with xyz positional restraints on select backbone particles—details provided in the Supporting Methods. The WTMetaD simulations were run using 20 walkers placed randomly within 2.5 nm of the protein. In the

$A_{2A}R$, AAC, and Kir2.2 systems, the walkers were constrained from moving greater than 3.5, 4, and 4.2 nm, respectively, from the geometric center of the protein. For LeuT, the walkers were constrained using a minimum distance of 4 nm between the lipid and the geometric center of each monomer. To prevent the cholesterol from flip-flopping between leaflets, two additional flat-bottomed restraints were applied: an angle restraint of one radian was applied between the ROH, C2 beads, and the z-axis and a wall parallel to the midpoint of the membrane offset toward the extracellular leaflet by 1 nm was applied to the ROH bead to prevent the cholesterol from moving vertically between leaflets. In all simulations, the flat bottom restraints were applied using the Plumed UPPER-WALLS of 1000 kJ mol^{-1} nm^{-2}.

The bias was added along the CV for each system, as defined in the Supporting Methods. The bias-factor was tuned to each lipid with cholesterol, CDL, and PIP$_2$ being 6, 8, and 20, respectively. The following WTMetaD parameters were applied to all systems: Gaussian width of 0.01 nm, height of 1 kJ mol^{-1}, and a 1 ps deposition rate. All WTMetaD simulations were performed at 310 K using Gromacs 2016 with the Plumed v2.4 plugin. When plotting, the bulk region of the membrane is considered to have an energy of 0 kJ mol^{-1}.

The bulk Gaussian height of 5% of the maximum value was used as a metric for the WTMetaD reaching a steady state (e.g., Supporting Figure 4), at which we see multiple association/disassociation events to the binding sites. The FES was sampled every 250 ps, whereupon the final 2D FES depicted in the text was recovered by averaging over the steady state.

■ RESULTS

We selected a panel of four experimentally well-characterized membrane proteins to represent bacterial, mitochondrial, and mammalian cell membranes. These were chosen to include examples for which experimental data for protein−lipid interactions are available, and which have been the subject of previous computational studies. They also represent three species of lipid (Figure 1C) which frequently form interactions with membrane proteins, namely: PIP$_2$, a negatively charged phospholipid which interacts with many ion channels and receptors in mammalian cell membranes; CDL, also negatively charged, which is present in mitochondrial and bacterial inner membranes, and cholesterol, which has been observed to bind to many G-protein-coupled receptors (GPCRs) and ion channels. Thus, we probe the following interactions: of PIP$_2$ with the mammalian inward rectifying potassium channel, Kir2.2; of CDL with the mitochondrial inner membrane ADP/ATP carrier protein, AAC,[21] and with the bacterial inner membrane LeuT;[35,51] and of cholesterol with a GPCR, the $A_{2A}R$.

Kir2.2−PIP$_2$ Binding Interactions. Kir2.2 is a member of the inwardly rectifying potassium channel family, found in neuronal cell membranes, which play a key role in the regulation of plasticity and neuronal excitation.[52] As with other members of this family, opening of Kir2.2 channels can be activated through interaction with PIP$_2$.[53] Structural (X-ray) studies show PIP$_2$ binds at the interface between the transmembrane and cytoplasmic domains (Figure 3A) to bring about channel opening.[34]

Simulations were run to probe the free energy of Kir2.2−PIP$_2$ interactions in a simple phospholipid (POPC) membrane. Binding sites have previously been identified both
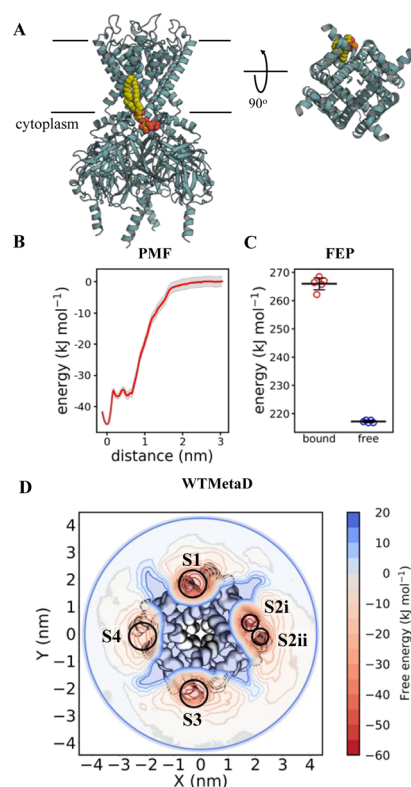

Figure 3. Calculating Kir2.2−PIP$_2$ binding energetics (A) view of Kir2.2 in cyan cartoon, with a bound PIP$_2$ molecule in yellow, orange and red spheres, as sampled with CG MD. PIP$_2$ interactions map to residues Arg 78, Arg 80, Lys 183, Arg 186, Lys 188, Lys 189 (*Gallus gallus* numbering). The approximate position of the membrane is shown with black lines. On the left is a view from the side, and on the right is a cytoplasmic view of the transmembrane region alone, with the intracellular domain removed for clarity. Note that only one PIP$_2$ binding pose is shown, but four are present around the homotetrameric Kir2.2. (B) PMF data for Kir2.2−PIP$_2$ binding. The y-axis is set to 0 for the bulk membrane, and the difference between this and the energy well (set to 0 nm on the x-axis) is $\Delta\Delta G_{bind}$, here −45 ± 2, with errors from 200 rounds of bootstrap analysis. (C) FEP data for Kir2.2−PIP2 binding, showing the energy cost for perturbing PIP$_2$ to POPC whilst bound to Kir2.2 (red) and whilst free in a POPC membrane (blue). $\Delta\Delta G_{bind}$ can be calculated from the free data minus the bound (see Figure 2), giving a value of −48 ± 2, with the error the standard deviation from five repeats. (D) 2D energy landscape for Kir2.2 and PIP$_2$ as computed using WTMetaD. The protein is shown as surface behind the data, with the large intracellular domain removed for clarity. The energetic landscape for a PIP$_2$ molecule around the protein has been computed and is shown as a red-blue contour map. Four binding regions in red can be seen around the protein, with reported $\Delta\Delta G_{bind}$ values as follow: S1 = −55 ± 7, S2i = −49 ± 4, S2ii = −45 ± 5, S3 = −45 ± 6 and S4 = −36 ± 7 kJ mol^{-1}.

structurally[34] and computationally.[5,18] We, therefore, carried out PMF and FEP simulations based on the crystallographic binding pose of PIP$_2$ in order to estimate the binding free energy. We note that these PMFs provide an interaction free energy for Kir2.2 with PIP$_2$ relative to that with POPC, that is the well depth in the PMF can be equated as $\Delta G_{PIP2\text{-}bind}$ − $\Delta G_{POPC\text{-}bind}$, modeling the vertical axes of the thermodynamic cycle in Figure 2. We shorten the product of this, $\Delta\Delta G_{bind(PIP2−POPC)}$, to $\Delta\Delta G_{bind}$. In the FEP analysis, the calculations directly compute $\Delta G_{PIP2−POPC}$ for the free and bound states of the lipid, modeling the horizontal axes of the

thermodynamic cycle in Figure 2. Therefore, FEP analysis provides us with the same $\Delta\Delta G_{bind}$ as from the PMF calculations, but calculated in a different manner.

The PMF and FEP data agree well, giving similar estimates for Kir2.2−PIP$_2$ interactions (Figure 3B,C; −45 ± 2 kJ mol^{-1} for PMF and −48 ± 2 kJ mol^{-1} for FEP). The PMF data reveal the presence of a second binding site, ∼0.5 nm from the main site, with a free energy of interaction of −36 ± 2 kJ mol^{-1}. This alternative pose of the lipid involves interactions between PIP$_2$ and residue Lys 220, and corresponds to a secondary binding mode proposed for the closely related Kir2.1 channel.[54]

Next, we carried out WTMetaD simulations to explore the multiple PIP$_2$ binding sites on the tetrameric channel structure. Through construction of a 2D CV in physical space between the geometric center of the protein to the head group (RP1, RP2, RP3, PO1, PO2, and PO3 beads) of PIP$_2$, WTMetaD allows us to calculate the same quantity as the PMF and FEP analyses, that is, the $\Delta\Delta G_{bind}$ of PIP$_2$ in relation to POPC. The data identified the primary binding modes for three of the channel subunits, with interaction free energies of −55 ± 7, −49 ± 4, and −45 ± 6 kJ mol^{-1} (Figure 3D). The latter of these values corresponds to the site probed using PMF and FEP, revealing excellent agreement between the techniques. For the fourth subunit, the secondary binding mode was recovered (−36 ± 7 kJ mol^{-1}), which again corresponds well with value from the PMF analysis.

We note that in the WTMetaD simulations of Kir2.2, as for all the systems, the protein is necessarily *xyz* restrained to keep the protein in a consistent position with relation to the CV (see Supplementary Methods for more detail). For Kir2.2, because of our input coordinates the channel is tilted (by ca. 5°) relative to the bilayer normal. We expect that this small tilt relative to the bilayer induces the shift of PIP$_2$ from the primary to the secondary binding site for the fourth channel subunit, with the primary binding mode only accessible to three of the four binding sites because of the orientation of the channel.

CDL Interactions with Two Transport Proteins. The ADP/ATP carrier (AAC), also known as the adenine nucleotide translocator, is present in the inner mitochondrial membrane where it accounts for 10% of the total protein content.[55] It functions as a regulator of mitochondrial adenine nucleotide concentration, allowing flux of ATP/ADP across the mitochondrial membrane.[56] CDL is known to bind to AAC[4,57] (Figure 4A), where it results in activation of the transporter.[58,59] The protein has an approximate three-fold symmetry with three homologous, but potentially nonidentical, binding sites for CDL.

We probed the energetics of the AAC−CDL interaction at all three of these sites with the protein embedded in a POPC membrane using PMF and FEP calculations (Figure 4A). PMF analyses yielded $\Delta\Delta G_{bind}$ values of −7 ± 2, −8 ± 4 and −5 ± 2 kJ mol^{-1} for CDL binding at these sites (Figure 4B). FEP produces good agreement with the PMF data, giving $\Delta\Delta G_{bind}$ values of −9 ± 3, −11 ± 3, and −8 ± 2 kJ mol^{-1} (Figure 4C), with each technique ranking the sites the same in terms of CDL binding energies.

Next, we explored the system using WTMetaD with a CV constructed between the protein and the GL0 of CDL. This revealed clear energy wells at all three binding sites (Figure 4D). The $\Delta\Delta G_{bind}$ values for sites 1, 2, and 3 were −14 ± 3, −13 ± 3, and −11 ± 6 kJ mol^{-1}, respectively, showing reasonable agreement with the PMF and, to a greater degree,
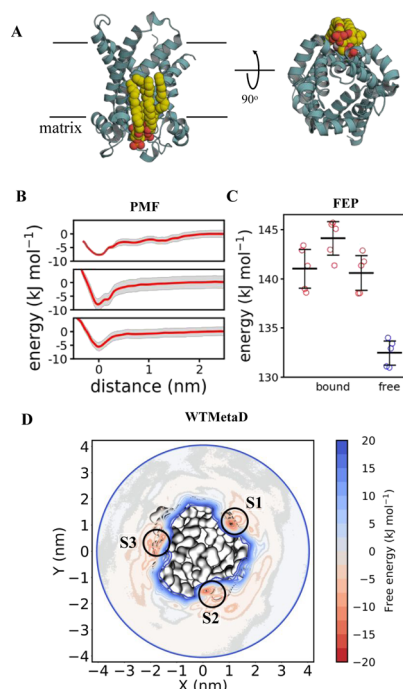

Figure 4. Calculating AAC−CDL binding energetics (A) AAC bound to CDL. Colours and views as in Figure 3A. Note that three CDL binding sites are present around AAC, which is a homotrimer. (B) PMF data for each of the AAC−CDL binding sites, as per Figure 3B. Site 1 is at the top, site 2 in the middle and site 3 at the bottom. $\Delta\Delta G_{bind}$ is −7 ± 2, −8 ± 4 and −5 ± 2 for each site respectively. (C) FEP data for AAC−CDL binding, showing the energy cost for perturbing CDL to POPC whilst bound to AAC in each of the three binding sites (red: site 1 on the left) and whilst free in a POPC membrane (blue). $\Delta\Delta G_{bind}$ is −9 ± 3, −11 ± 3 and −8 ± 2 for each site respectively. (D) 2D energy landscape for AAC and CDL, as per Figure 3D. Three binding regions in red can be seen around the protein, reporting $\Delta\Delta G_{bind}$ values as follow: S1 = −14 ± 3, S2 = −13 ± 3 and S3 = −11 ± 6 kJ mol^{-1}.

FEP estimates. Reassuringly, the sites are again similarly ranked, with the site 3 lower in energy than sites 1 or 2.

We also explored CDL interactions with LeuT, a bacterial homologue of the solute carrier family 6 (SLC6) class of proteins, which includes the human serotonin transporter.[60] LeuT catalyzes sodium-driven small hydrophobic amino acid transport across the bacterial inner membrane,[61] and has been shown to bind to CDL (Figure 5A), which stabilizes the dimeric form of the transporter.[51] We probed the energetics of the CDL−LeuT interaction in a POPE membrane using both PMF and FEP calculations. The initial CDL pose was based on a previously identified likely binding site at the dimer interface (see ref 51 for more details). PMF analysis yielded a $\Delta\Delta G_{bind}$ value of −6 ± 3 kJ mol^{-1} for CDL at this site (Figure 5B), in agreement with the FEP $\Delta\Delta G_{bind}$ value of −9 ± 2 kJ mol^{-1} (Figure 5C). WTMetaD not only identifies the dimer interface binding site (with a binding free energy of −7 ± 3 and −10 ± 6 kJ mol^{-1} for the equivalent and opposing sites, respectively) but also additional sites around the complex (Figure 5D). For example, sites 2 and 4 appear to bind CDL with a lower energy than the dimer-interface site (−3 ± 2 kJ mol^{-1} for each site). Nonspecific of CDL with a low free energy of ca. −3 kJ mol^{-1} is also seen over a 3 nm area of each monomers (Figure 5D; dotted lines).
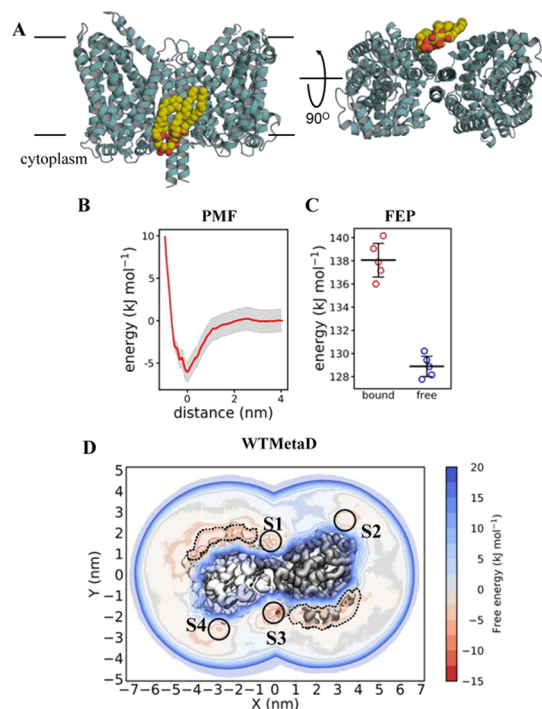
Figure 5. Calculating LeuT−CDL binding energetics (A) view of LeuT bound to CDL. Colours and views as in Figure 3A. Note that a second equivalent CDL binding site is present on the other side of the homodimeric LeuT. (B) PMF data for AAC−LeuT binding, as per Figure 3B. $\Delta\Delta G_{bind}$ is -6 ± 3. (C) FEP data for AAC−LeuT binding, perturbing CDL to POPE whilst bound to LeuT (red) and whilst free in a POPE membrane (blue). $\Delta\Delta G_{bind}$ is -9 ± 2. (D) 2D energy landscape for LeuT and CDL, as per Figure 3D. Two binding regions in red can be seen around the protein, reporting $\Delta\Delta G_{bind}$ values as follow: S1 = -7 ± 3, S2 = -3 ± 2, S3 = -10 ± 6 and S4 = -3 ± 2 kJ mol^{-1}. Two indiscriminate binding regions of ca. -3 kJ mol^{-1} are highlighted with dotted lines.

Cholesterol Interactions with a GPCR. The GPCR class of proteins constitutes the largest family of membrane proteins and account for 35% of all drug targets.[62] Many members are thought to be functionally modulated by cholesterol binding,[63] including the highly studied $A_{2A}R$. This provides an example of protein−lipid interactions with an uncharged and relatively rigid lipid molecule.

Cholesterol binding to $A_{2A}R$ has been previously identified structurally[64] and explored using molecular dynamics.[7,20,65] To probe the energetics of this process, we applied WTMetaD to human $A_{2A}R$ in a POPC membrane to produce a full 2D map of cholesterol binding on the extracellular leaflet around the receptor, using the protein and ROH bead of cholesterol as a CV. We identified three binding sites between TM 1−2, 1−7, and 6−7 (Figure 6A,B), with $\Delta\Delta G_{bind}$ of -8 ± 3, -9 ± 3, and -5 ± 2 kJ mol^{-1}. Of these, only the TM 6−7 cholesterol binding site has been captured in structural data for $A_{2A}R$,[36,66,67] although the other two binding sites correspond to electron density identified as acyl tails in several $A_{2A}R$ structures.[36,66,67]

We then probed these sites using ABFE, using restraints and a thermodynamic cycle outlined in a recent study by Salari et al.[46] (see Supporting Methods and Supporting Figure 3). We obtained $\Delta\Delta G_{bind}$ values of -5 ± 2, -8 ± 1, and -2 ± 1 kJ mol^{-1} for the three sites (Figure 6C), in reasonable agreement with the WTMetaD data. Importantly, we get an identical
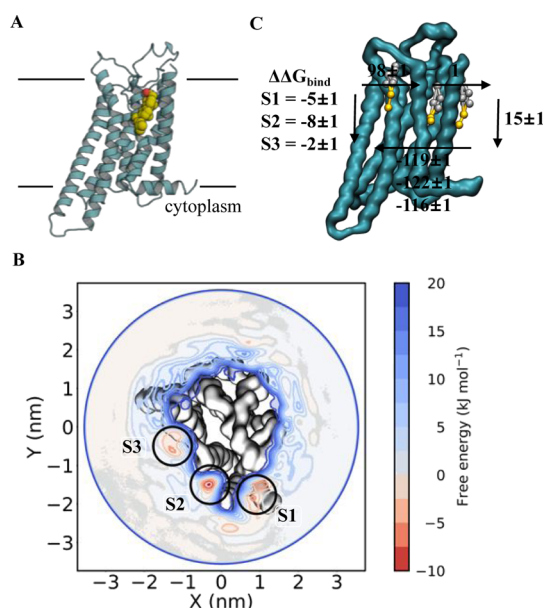

Figure 6. Calculating $A_{2A}R$−cholesterol binding energetics (A) view of $A_{2A}R$ bound to cholesterol, in one of several possible binding sites. Colours as in Figure 3A. (B) WTMetaD 2D energy landscape for $A_{2A}R$ and cholesterol, for the extracellular leaflet only. Three binding regions in red can be seen around the protein, reporting $\Delta\Delta G_{bind}$ values as follow: S1 = -8 ± 3, S2 = -9 ± 3 and S3 = -5 ± 2 kJ mol^{-1}. (C) ABFE analysis the sites from panel B. Here, the bound cholesterol is fully decoupled from a POPC membrane whilst bound to $A_{2A}R$ or whilst free in the membrane. The cycle represented here is a simplified version of the cycle in Supporting Figure 3. The final $\Delta\Delta G_{bind}$ values are -5 ± 1, -8 ± 1 and -2 ± 1 kJ mol^{-1}.

ranking of the sites between the techniques, with site 2 the highest energy and site 3 the lowest. Note that we were unable to probe these sites using PMF calculations, as the energies become swamped by background thermal fluctuations.

We can compare our estimates of the strength of $A_{2A}R$ cholesterol interactions with those from other simulation studies of GPCRs. Genheden et al.[68] estimated free energies of interaction with b2AR and $A_{2A}R$ of the order of -10 to -15 kJ mol^{-1} from CG simulations. Lee and Lyman[7] estimated free energies for the $A_{2A}R$ of -3 to -5 kJ mol^{-1} from atomistic MD simulations. Thus, our estimates are in broad agreement with those from previous studies of the $A_{2A}R$, both of which estimated free energies directly from cholesterol occupancies following extended (but possibly under sampled) equilibrium MD simulations.

DISCUSSION

Integral membrane proteins are strongly affected by the lipid environment in which they reside (for a recent review, see ref 1). Information on the structural basis of these interactions can be determined by X-ray crystallography, cryoelectron microscopy, and NMR. In contrast, determining the free energies of these interactions remains very challenging either experimentally[15,16,30] or computationally.[5,11,18,20,21]

One of the main challenges facing computational estimation of free energies of protein−lipid interactions is presented by the slow timescale of relaxation of lipid molecules in a bilayer. This means that for free energy calculations, we need to be confident that we have adequately sampled those interactions, that is, the simulations have converged. The agreement

Table 1. Summary of Binding Energies for the Systems Tested Here[a]

system	site	lipid	membrane	PMF	FEP/ABFE	WTMetaD
Kir2.2	3	PIP$_2$	POPC	−45 ± 2	−48 ± 2	−45 ± 6
AAC	1	CDL	POPC	−7 ± 2	−10 ± 4	−14 ± 3
	2			−8 ± 4	−11 ± 2	−13 ± 3
	3			−7 ± 4	−9 ± 4	−11 ± 6
LeuT	1	CDL	POPE	−6 ± 3	−9 ± 2	−7 ± 3
A2AR	1	chol	POPC		−5 ± 1	−8 ± 3
	2				−8 ± 1	−9 ± 3
	3				−2 ± 1	−5 ± 2

[a]Reported are the measured $\Delta\Delta G_{bind}$ values, with the errors calculated as defined in the text and figure legends. These data have been used to populate the chart in Figure 7.

between all three techniques for lipid binding energetics gives us confidence that this is indeed the case (Table 1 and Figure 7). To our surprise, the analyses here provide lower energies
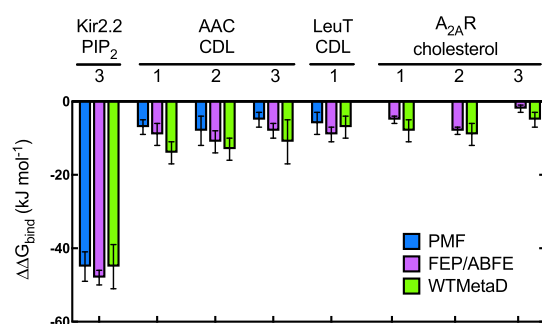

Figure 7. Comparison of different techniques. Bar chart showing $\Delta\Delta G_{bind}$ for the different protein−lipid systems described above. Shown are energies calculated with PMF (blue), FEP/ABFE (maroon) and WTMetaD (green). The error bars shown here are from the 200 rounds of bootstrap analysis for the PMF, standard deviations of 5 repeats for the FEP, or standard deviations of the energies for each site for the WTMetaD (green).

for CDL binding to AAC[21] and other proteins[11,18,69] than previously reported; this discrepancy most likely reflects a different handling of long-range electrostatic interactions. Nonetheless, when comparing like-for-like systems, as here, it is evident that all three techniques converge to a common value and provide a similar degree of accuracy.

In addition to accuracy, it is of interest to compare the computational efficiency of the three methods, especially if they are to be employed in an automated pipeline (see e.g. ref 32) to characterize and compare protein−lipid interactions across a wide range of membrane protein structures. As can be seen from Table 2, the least computationally demanding technique is FEP/ABFE. This is because each alchemical pathway can be adequately described in 21 (PIP2), 29

Table 2. Computational Cost of Each Technique as Applied Here[a]

technique	system applied to here	required simulation time (μs)
PMF	Kir2.2, AAC, LeuT	50−75
FEP	Kir2.2, AAC, LeuT	25−40
ABFE	A$_2$AR	35
WTMetaD	Kir2.2, AAC, LeuT, A$_2$AR	120−190

[a]Reporting the simulation time used in each of the analysis measures here. Note that, as described in the Discussion, these values may overestimate the time required for convergence.

(cholesterol) or 31 (CDL) windows, and 250 ns of simulation time per window is sufficient for good convergence (see Supporting Figure 2, 6 and 9). Running 5 independent repeats (resulting in ∼25−40 μs simulation time in total) permits statistical analysis of the data. In fact, the data suggest that this cost could be further reduced to 150 ns per window with only 3 repeats (∼9.5 μs in total; achievable in about 1−2 days with a mid to high range GPU), as this agrees well with the more extensively sampled data (e.g., −48 ± 2 kJ mol^{-1} vs −51 ± 1 kJ mol^{-1}; Supporting Figure 13A). Note that these estimates do not account for the initial set of simulations that are required for the perturbation of each lipid species in the bulk membrane, as this benchmark may be used for subsequent calculations.

PMFs are generally less cost-efficient than FEP, taking at least 50 μs to converge. This is largely because equilibration of a lipid as it diffuses within a bilayer is relatively slow, meaning that each window needs to be simulated for 1 μs (see Supporting Figures 1, 5 and 8). In addition, a 0.05 nm window spacing from the bound state to bulk membrane usually requires 60 or more windows to get from the bound state to bulk. It should be noted, however, that a less frequent spatial sampling of 0.1 nm intervals, while resulting in a lower statistical certainty across the PMF, produces a similar estimation of $\Delta\Delta G_{bind}$ (for instance −45 ± 2 kJ mol^{-1} vs −48 ± 2 kJ mol^{-1}; Supporting Figure 13B). In contrast to FEPs, however, PMFs offer a reaction coordinate in 1D space, which may be useful in certain cases. An example of this is for Kir2.2, where the PMF was able to detect the second, previously identified,[54] binding site (Figure 3B).

Generally, WTMetaD is the most expensive technique, taking >100 μs to converge (see Supporting Figures 4, 7, 10 and 12), which is compounded by the requirement to read/write the Gaussian depositions to file, together taking ∼3× longer to sample the same timescales as a standard MD simulation. However, in instances where a binding site is known, the simulations could be scaled back to focus upon the single binding site, saving considerable resources in the process. In addition, WTMetaD allows the mapping of the full 2D energy landscape around the target protein. This proved to be especially powerful for the A$_2$AR and LeuT systems, where multiple asymmetric binding sites were revealed. In addition to this, the WTMetaD is not restricted by needing to know the highest energy pose before analysis, which was of particular use for the study of A$_2$AR here.

For each system, the 2D landscapes were additionally compared to the PMF data through extraction of 1D coordinates from the data (see the Supporting Methods for details, and e.g. Supporting Figure 4F,G for data). This is

potentially very powerful as it allows the construction of pathways from the bulk membrane to the binding site, providing a greater degree of insight into the dynamics of binding.

At present, we have compared these methods for estimation of lipid interaction free energies using the Martini[22,23] force field. We are aware that the free energies estimated are therefore an approximation because of smoothing of the free energy landscape, an inevitable consequence of coarse-graining. In the future, comparison of these data with corresponding atomistic simulations will be necessary, although achieving convergence of equivalent atomistic simulations is currently extremely challenging for systems of the size and complexity described here. These analyses would likely benefit from the recently described use of a funnel potential for WTMetaD simulations,[70] allowing a much more focused CV.

Finally, comparison of these methods to experimental analyses will be of particular importance. Currently, there have been only a few examples of experimentally determined protein–lipid binding affinities: these include using nMS,[15] SPR,[16] and FRET.[30] The majority of these approaches, however, involve both the protein and lipid being solubilized beforehand in detergent micelles. This will likely provide an inaccurate picture of the true energetics of lipid binding/unbinding in a lipid membrane, as the acyl tails will likely contribute substantially to the binding energies. Our work therefore highlights a pressing need for accurate measurement of lipid-binding affinities to be applied to a number of well characterized membrane protein systems.

In summary, in this study we have shown that three distinct methods for CG free energy calculations are able to provide robust estimates of the strength and specificity of lipid-binding sites on membrane proteins. We envisage that these simulations can be readily performed, taking protein–lipid binding sites identified from structures, long equilibrium simulations ("fingerprinting")[9,71] or WTMetaD, and enumerating energies for specific lipid association in a semiautomated manner. The characterization of lipid-binding sites on membrane proteins offers the prospect of discovery of potential druggable allosteric sites on a wide range of membrane proteins, including almost all potential target species within the current membrane protein structural proteome of ca. 4000 structures.[33]

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00548.

Detailed descriptions of how the different free energies were implemented in Gromacs and Plumed (PDF)

Kir2.2 PMF convergence analysis; Kir2.2 FEP convergence analysis; thermodynamic cycle for ABFE; Kir2.2 WTMetaD supporting data; AAC site 1 convergence analysis; AAC FEP convergence analysis; AAC WTMetaD supporting data; LeuT PMF convergence analysis; LeuT FEP convergence analysis; LeuT WTMetaD supporting data; A2AR ABFE supporting data; A2AR WTMetaD supporting data; sampling frequency for FEP and PMF; calculation of binding energies from the thermodynamic cycle; modified form of calculation of binding energies;

calculation of second analytical term in the modified form of calculation of binding energies (PDF)

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: phillip.stansfeld@bioch.ox.ac.uk. Phone: +44-1865-613362.

ORCID Ⓘ
Mark S. P. Sansom: 0000-0001-6360-7959
Phillip J. Stansfeld: 0000-0001-8800-7669

Author Contributions
[†]R.A.C. and O.N.V. contributed equally to this work.

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Corradi, V.; Sejdiu, B. I.; Mesa-Galloso, H.; Abdizadeh, H.; Noskov, S. Y.; Marrink, S. J.; Tieleman, D. P. Emerging Diversity in Lipid-Protein Interactions. *Chem. Rev.* **2019**, *119*, 5775.

(2) Norimatsu, Y.; Hasegawa, K.; Shimizu, N.; Toyoshima, C. Protein-phospholipid interplay revealed with crystals of a calcium pump. *Nature* **2017**, *545*, 193.

(3) Whorton, M. R.; MacKinnon, R. Crystal Structure of the Mammalian GIRK2 K+ Channel and Gating Regulation by G Proteins, PIP2, and Sodium. *Cell* **2011**, *147*, 199−208.

(4) Pebay-Peyroula, E.; Dahout-Gonzalez, C.; Kahn, R.; Trézéguet, V.; Lauquin, G. J.-M.; Brandolin, G. Structure of Mitochondrial ADP/ATP Carrier in Complex with Carboxyatractyloside. *Nature* **2003**, *426*, 39.

(5) Stansfeld, P. J.; Hopkinson, R.; Ashcroft, F. M.; Sansom, M. S. P. PIP2-Binding Site in Kir Channels: Definition by Multiscale Biomolecular Simulations. *Biochemistry* **2009**, *48*, 10926−10933.

(6) Schmidt, M. R.; Stansfeld, P. J.; Tucker, S. J.; Sansom, M. S. P. Simulation-Based Prediction of Phosphatidylinositol 4,5-Bisphosphate Binding to an Ion Channel. *Biochemistry* **2013**, *52*, 279−281.

(7) Lee, J. Y.; Lyman, E. Predictions for Cholesterol Interaction Sites on the A2Aadenosine Receptor. *J. Am. Chem. Soc.* **2012**, *134*, 16512−16515.

(8) Stansfeld, P. J.; Jefferys, E. E.; Sansom, M. S. P. Multiscale Simulations Reveal Conserved Patterns of Lipid Interactions with Aquaporins. *Structure* **2013**, *21*, 810−819.

(9) Corradi, V.; Mendez-Villuendas, E.; Ingólfsson, H. I.; Gu, R.-X.; Siuda, I.; Melo, M. N.; Moussatova, A.; DeGagné, L. J.; Sejdiu, B. I.; Singh, G.; Wassenaar, T. A.; Delgado Magnero, K.; Marrink, S. J.; Tieleman, D. P. Lipid-Protein Interactions Are Unique Fingerprints for Membrane Proteins. *ACS Cent. Sci.* **2018**, *4*, 709−717.

(10) Arnarez, C.; Mazat, J.-P.; Elezgaray, J.; Marrink, S.-J.; Periole, X. Evidence for Cardiolipin Binding Sites on the Membrane-Exposed Surface of the Cytochrome Bc1. *J. Am. Chem. Soc.* **2013**, *135*, 3112−3120.

(11) Arnarez, C.; Marrink, S. J.; Periole, X. Identification of Cardiolipin Binding Sites on Cytochrome c Oxidase at the Entrance of Proton Channels. *Sci. Rep.* **2013**, *3*, 1263.

(12) Hedger, G.; Shorthouse, D.; Koldsø, H.; Sansom, M. S. P. Free Energy Landscape of Lipid Interactions with Regulatory Binding Sites on the Transmembrane Domain of the EGF Receptor. *J. Phys. Chem. B* **2016**, *120*, 8154−8163.

(13) Gupta, K.; Li, J.; Liko, I.; Gault, J.; Bechara, C.; Wu, D.; Hopper, J. T. S.; Giles, K.; Benesch, J. L. P.; Robinson, C. V. Identifying Key Membrane Protein Lipid Interactions Using Mass Spectrometry. *Nat. Protoc.* **2018**, *13*, 1106.

(14) Yen, H.-Y.; Hoi, K. K.; Liko, I.; Hedger, G.; Horrell, M. R.; Song, W.; Wu, D.; Heine, P.; Warne, T.; Lee, Y.; Carpenter, B.; Plückthun, A.; Tate, C. G.; Sansom, M. S. P.; Robinson, C. V. PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **2018**, *559*, 423.

(15) Patrick, J. W.; Boone, C. D.; Liu, W.; Conover, G. M.; Liu, Y.; Cong, X.; Laganowsky, A. Allostery Revealed within Lipid Binding Events to Membrane Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 2976−2981.

(16) Inada, M.; Kinoshita, M.; Sumino, A.; Oiki, S.; Matsumori, N. A Concise Method for Quantitative Analysis of Interactions between Lipids and Membrane Proteins. *Anal. Chim. Acta* **2019**, *1059*, 103−112.

(17) Chipot, C. Frontiers in Free-Energy Calculations of Biological Systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 71−89.

(18) Domański, J.; Hedger, G.; Best, R. B.; Stansfeld, P. J.; Sansom, M. S. P. Convergence and Sampling in Determining Free Energy Landscapes for Membrane Protein Association. *J. Phys. Chem. B* **2017**, *121*, 3364−3375.

(19) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational Modeling of Realistic Cell Membranes. *Chem. Rev.* **2019**, *119*, 6184.

(20) Song, W.; Sansom, M. S. P. State Dependent Interactions of Lipids with GPCR Revealed by MD Simulations Using In Vivo-Mimetic Membranes. *Biophys. J.* **2018**, *114*, 274a.

(21) Hedger, G.; Rouse, S. L.; Domański, J.; Chavent, M.; Koldsø, H.; Sansom, M. S. P. Lipid-Loving ANTs: Molecular Simulations of Cardiolipin Interactions and the Organization of the Adenine Nucleotide Translocase in Model Mitochondrial Membranes. *Biochemistry* **2016**, *55*, 6238−6249.

(22) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.

(23) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.

(24) Shirts, M. R.; Mobley, D. L. An Introduction to Best Practices in Free Energy Calculations. *Methods Mol. Biol.* **2013**, *924*, 271−311.

(25) Aldeghi, M.; Bluck, J. P.; Biggin, P. C. Absolute Alchemical Free Energy Calculations for Ligand Binding: A Beginner's Guide. *Methods Mol. Biol.*, **2018**, *1762*, 199−232.

(26) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 20603.

(27) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. *Chem. Soc. Rev.* **2013**, *42*, 6801.

(28) Duncan, A. L.; Robinson, A. J.; Walker, J. E. Cardiolipin Binds Selectively but Transiently to Conserved Lysine Residues in the Rotor of Metazoan ATP Synthases. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 8687−8692.

(29) Dahlberg, M.; Maliniak, A. Mechanical properties of coarse-grained bilayers formed by cardiolipin and zwitterionic lipids. *JCTC* **2010**, *6*, 1638−1649.

(30) Corey, R. A.; Pyle, E.; Allen, W. J.; Watkins, D. W.; Casiraghi, M.; Miroux, B.; Arechaga, I.; Politis, A.; Collinson, I. Specific cardiolipin-SecY interactions are required for proton-motive force stimulation of protein secretion. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 7967−7972.

(31) Pluhackova, K.; Gahbauer, S.; Kranz, F.; Wassenaar, T. A.; Böckmann, R. A. Dynamic Cholesterol-Conditioned Dimerization of the G Protein Coupled Chemokine Receptor Type 4. *PLoS Comput. Biol.* **2016**, *12*, No. e1005169.

(32) Stansfeld, P. J.; Goose, J. E.; Caffrey, M.; Carpenter, E. P.; Parker, J. L.; Newstead, S.; Sansom, M. S. P. MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* **2015**, *23*, 1350−1361.

(33) Newport, T. D.; Sansom, M. S. P.; Stansfeld, P. J. The MemProtMD Database: A Resource for Membrane-Embedded Protein Structures and Their Lipid Interactions. *Nucleic Acids Res.* **2018**, *47*, D390−D397.

(34) Hansen, S. B.; Tao, X.; MacKinnon, R. Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2. *Nature* **2011**, *477*, 495.

(35) Yamashita, A.; Singh, S. K.; Kawate, T.; Jin, Y.; Gouaux, E. Crystal Structure of a Bacterial Homologue of Na+/Cl−Dependent Neurotransmitter Transporters. *Nature* **2005**, *437*, 215.

(36) Segala, E.; Guo, D.; Cheng, R. K. Y.; Bortolato, A.; Deflorian, F.; Doré, A. S.; Errey, J. C.; Heitman, L. H.; Ijzerman, A. P.; Marshall, F. H.; et al. Controlling the Dissociation of Ligands from the Adenosine A2A Receptor through Modulation of Salt Bridge Strength. *J. Med. Chem.* **2016**, *59*, 6470−6479.

(37) Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.; Pieper, U.; Sali, A. Comparative Protein Structure Modeling Using Modeller. *Curr. Protoc. Bioinf.* **2006**, *15*, 5−6.

(38) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46*, W296.

(39) Melo, M. N.; Ingólfsson, H. I.; Marrink, S. J. Parameters for Martini Sterols and Hopanoids Based on a Virtual-Site Description. *J. Chem. Phys.* **2015**, *143*, 243152.

(40) Souaille, M.; Roux, B. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.* **2001**, *135*, 40−57.

(41) Hub, J. S.; De Groot, B. L.; Van Der Spoel, D. g_wham-A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theory Comput.* **2010**, *6*, 3713−3720.

(42) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; Van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem. Phys. Lett.* **1994**, *222*, 529−539.

(43) Steinbrecher, T.; Mobley, D. L.; Case, D. A. Nonlinear Scaling Schemes for Lennard-Jones Interactions in Free Energy Calculations. *J. Chem. Phys.* **2007**, *127*, 214108.

(44) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the Analysis of Free Energy Calculations. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 397−411.

(45) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.

(46) Salari, R.; Joseph, T.; Lohia, R.; Hénin, J.; Brannigan, G. A Streamlined, General Approach for Computing Ligand Binding Free Energies and Its Application to GPCR-Bound Cholesterol. *J. Chem. Theory Comput.* **2018**, *14*, 6560−6573.

(47) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.

(48) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.

(49) Bussi, G.; Laio, A.; Parrinello, M. Equilibrium Free Energies from Nonequilibrium Metadynamics. *Phys. Rev. Lett.* **2006**, *96*, 090601.

(50) Dama, J. F.; Parrinello, M.; Voth, G. A. Well-Tempered Metadynamics Converges Asymptotically. *Phys. Rev. Lett.* **2014**, *112*, 240602.

(51) Gupta, K.; Donlan, J. A. C.; Hopper, J. T. S.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The Role of Interfacial Lipids in Stabilizing Membrane Protein Oligomers. *Nature* **2017**, *541*, 421.

(52) Tempel, B. L.; Jan, Y. N.; Jan, L. Y. Cloning of a Probable Potassium Channel Gene from Mouse Brain. *Nature* **1988**, *332*, 837.

(53) Hilgemann, D. W.; Feng, S.; Nasuhoglu, C. The Complex and Intriguing Lives of PIP2 with Ion Channels and Transporters. *Sci. STKE* **2001**, *2001*, re19.

(54) Lee, S.-J.; Wang, S.; Borschel, W.; Heyman, S.; Gyore, J.; Nichols, C. G. Secondary Anionic Phospholipid Binding Site and Gating Mechanism in Kir2. 1 Inward Rectifier Channels. *Nat. Commun.* **2013**, *4*, 2786.

(55) Brand, M. D.; Pakay, J. L.; Ocloo, A.; Kokoszka, J.; Wallace, D. C.; Brookes, P. S.; Cornwall, E. J. The Basal Proton Conductance of Mitochondria Depends on Adenine Nucleotide Translocase Content. *Biochem. J.* **2005**, *392*, 353−362.

(56) Palmieri, F. The Mitochondrial Transporter Family (SLC25): Physiological and Pathological Implications. *Pflüg. Arch.* **2004**, *447*, 689−709.

(57) Beyer, K.; Klingenberg, M. ADP/ATP Carrier Protein from Beef Heart Mitochondria Has High Amounts of Tightly Bound Cardiolipin, as Revealed by Phosphorus-31 Nuclear Magnetic Resonance. *Biochemistry* **1985**, *24*, 3821−3826.

(58) Brandolin, G.; Doussiere, J.; Gulik, A.; Gulik-Krzywicki, T.; Lauquin, G. J. M.; Vignais, P. V. Kinetic, Binding and Ultrastructural Properties of the Beef Heart Adenine Nucleotide Carrier Protein after Incorporation into Phospholipid Vesicles. *Biochim. Biophys. Acta, Bioenerg.* **1980**, *592*, 592−614.

(59) Ruprecht, J. J.; King, M. S.; Zögg, T.; Aleksandrova, A. A.; Pardon, E.; Crichton, P. G.; Steyaert, J.; Kunji, E. R. S. The Molecular Mechanism of Transport by the Mitochondrial ADP/ATP Carrier. *Cell* **2019**, *176*, 435−447.

(60) Periole, X.; Zeppelin, T.; Schiøtt, B. Dimer Interface of the Human Serotonin Transporter and Effect of the Membrane Composition. *Sci. Rep.* **2018**, *8*, 1−15.

(61) Singh, S. K.; Piscitelli, C. L.; Yamashita, A.; Gouaux, E. A Competitive Inhibitor Traps LeuT in an Open-to-out Conformation. *Science* **2008**, *322*, 1655−1661.

(62) Sriram, K.; Insel, P. A. GPCRs as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **2018**, *96*, mol-117 111062.

(63) Oates, J.; Watts, A. Uncovering the Intimate Relationship between Lipids, Cholesterol and GPCR Activation. *Curr. Opin. Struct. Biol.* **2011**, *21*, 802−807.

(64) Gimpl, G. Interaction of G Protein Coupled Receptors and Cholesterol. *Chem. Phys. Lipids* **2016**, *199*, 61−73.

(65) Rouviere, E.; Arnarez, C.; Yang, L.; Lyman, E. Identification of Two New Cholesterol Interaction Sites on the A2A Adenosine Receptor. *Biophys. J.* **2017**, *113*, 2415−2424.

(66) Liu, W.; Chun, E.; Thompson, A. A.; Chubukov, P.; Xu, F.; Katritch, V.; Han, G. W.; Roth, C. B.; Heitman, L. H.; IJzerman, A. P.; et al. Structural Basis for Allosteric Regulation of GPCRS by Sodium Ions. *Science* **2012**, *337*, 232−236.

(67) Melnikov, I.; Polovinkin, V.; Kovalev, K.; Shevchenko, V.; Shevtsov, M.; Borshchevskiy, V.; Cherezov, V.; Leonard, G.; Gordeliy, V.; Popov, A. Fast Iodide-SAD Phasing for Membrane Protein Structure Determination. *Acta Crystallogr., Sect. A: Found. Adv.* **2016**, *72*, s199.

(68) Genheden, S.; Essex, J. W.; Lee, A. G. G protein coupled receptor interactions with cholesterol deep in the membrane. *Biochim. Biophys. Acta, Biomembr.* **2017**, *1859*, 268−281.

(69) Arnarez, C.; Marrink, S. J.; Periole, X. Molecular Mechanism of Cardiolipin-Mediated Assembly of Respiratory Chain Supercomplexes. *Chem. Sci.* **2016**, *7*, 4435−4443.

(70) Yuan, X.; Raniolo, S.; Limongelli, V.; Xu, Y. The Molecular Mechanism Underlying Ligand Binding to the Membrane-Embedded Site of a G-Protein-Coupled Receptor. *J. Chem. Theory Comput.* **2018**, *14*, 2761−2770.

(71) Shearer, J.; Jefferies, D.; Khalid, S. Outer Membrane Proteins OmpA, FhuA, OmpF, EstA, BtuB, and OmpX Have Unique Lipopolysaccharide Fingerprints. *J. Chem. Theory Comput.* **2019**, *15*, 2608−2619.