

# Free Energy Landscape of Lipid Interactions with Regulatory Binding Sites on the Transmembrane Domain of the EGF Receptor

George Hedger,<sup>†</sup> David Shorthouse,<sup>†,‡</sup> Heidi Koldsø,<sup>\*,†,§</sup> and Mark S. P. Sansom<sup>\*,†</sup>

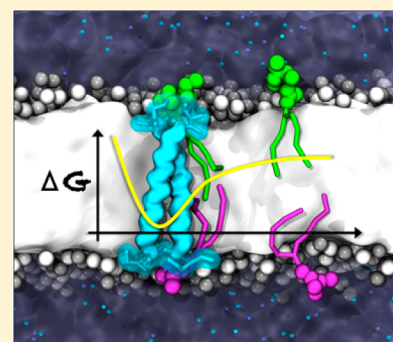
<sup>†</sup>Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom

<sup>‡</sup>MRC Cancer Unit, University of Cambridge, MRC Research Centre, Box 197, Cambridge CB2 0X1, United Kingdom

<sup>§</sup>D. E. Shaw Research, 120 West 45th Street, 39th floor, New York, New York 10036, United States

## S Supporting Information

**ABSTRACT:** Lipid molecules can bind to specific sites on integral membrane proteins, modulating their structure and function. We have undertaken coarse-grained simulations to calculate free energy profiles for glycolipids and phospholipids interacting with modulatory sites on the transmembrane helix dimer of the EGF receptor within a lipid bilayer environment. We identify lipid interaction sites at each end of the transmembrane domain and compute interaction free energy profiles for lipids with these sites. Interaction free energies ranged from ca.  $-40$  to  $-4$  kJ/mol for different lipid species. Those lipids (glycolipid GM3 and phosphoinositide PIP<sub>2</sub>) known to modulate EGFR function exhibit the strongest binding to interaction sites on the EGFR, and we are able to reproduce the preference for interaction with GM3 over other glycolipids suggested by experiment. Mutation of amino acid residues essential for EGFR function reduce the binding free energy of these key lipid species. The residues interacting with the lipids in the simulations are in agreement with those suggested by experimental (mutational) studies. This approach provides a generalizable tool for characterizing the interactions of lipids that bind to specific sites on integral membrane proteins.



## INTRODUCTION

The complexity and diversity of cell membrane compositions suggest that lipids may bind to membrane proteins at specific sites and modulate their function. Only within the past decade have structure determination techniques begun to achieve the resolution required to discern specifically bound lipids at atomic resolution, and we now possess over 100 structures of membrane proteins suggested to contain bound lipid molecules.<sup>1</sup> In a number of cases, structural data have been coupled to functional analysis, and it is now clear that several classes of membrane protein, including, e.g., K<sup>+</sup> channels, receptor tyrosine kinases (RTKs), and G-protein coupled receptors, are regulated via selective interactions with specific lipids.<sup>2–4</sup> Therefore, there is a need to better characterize these interactions to fully understand the influence of the surrounding membrane environment on membrane protein structure and function.<sup>5</sup>

The epidermal growth factor receptor (EGFR/ErbB1) is perhaps the best characterized member of the human RTK family. It resides in the plasma membrane, where it serves an essential role in transmitting information on the cellular environment to intracellular signaling networks, which subsequently elicit a response.<sup>6</sup> Mutations of the receptor have been implicated in a variety of cancers and thus the EGFR is a target for therapeutic intervention.<sup>7</sup> Of particular interest, EGFR is known to be modulated by its surrounding lipid environment. The emerging picture is that this modulation may occur via specific lipid interactions,<sup>4,8</sup> in addition to the more

general influences of lipids such as cholesterol on the biophysical properties of the membrane.<sup>9</sup>

Molecular simulation techniques provide a powerful tool for exploring lipid interactions with membrane proteins<sup>10</sup> and have previously been used to successfully predict lipid interaction sites on cytochrome C oxidase,<sup>11</sup> aquaporins,<sup>12</sup> and Kir channels.<sup>13</sup> However, there is a need to quantify these interactions to enable predictions of the effects of protein mutation and to provide a fuller understanding of the lipid selectivity of proteins such as EGFR. Molecular simulations and potential of mean force<sup>11</sup> (PMF) calculations represent one such route for the quantitative exploration of lipid–protein interactions. PMF methods are increasingly being applied to better understand the molecular basis of a range of biomolecular interactions (e.g., refs 11 and 14–17) and allow one to describe the free energy profile along a particular reaction coordinate. Sufficient sampling of configurational space along this reaction coordinate is necessary for the accurate calculation of a free energy profile. Attaining sufficient sampling and convergence of the calculation is not an insignificant challenge,<sup>18</sup> which can inhibit its application to complex systems such as membranes. These challenges may be

**Special Issue:** J. Andrew McCammon Festschrift

**Received:** February 9, 2016

**Revised:** April 21, 2016

**Published:** April 24, 2016

addressed through application of specialized sampling techniques such as umbrella sampling<sup>19</sup> and adaptive force biasing.<sup>20</sup> An additional approach is to reduce the granularity of the system (and hence the number of degrees of freedom that must be sampled) through use of coarse-grained (CG) models, which allow an increase in accessible simulation times of several orders of magnitude.<sup>21</sup> A number of comparative studies have indicated that estimates of the free energy for removal of lipid molecules from a bilayer obtained from CG and all-atom simulations are in agreement with one another and with experiment.<sup>22,23</sup> On the basis of these considerations, we chose to employ umbrella sampling together with the MARTINI CG force field (see ref 21 and references therein), which was originally parametrized based on comparison with experimental and all-atom simulation partitioning free energies.

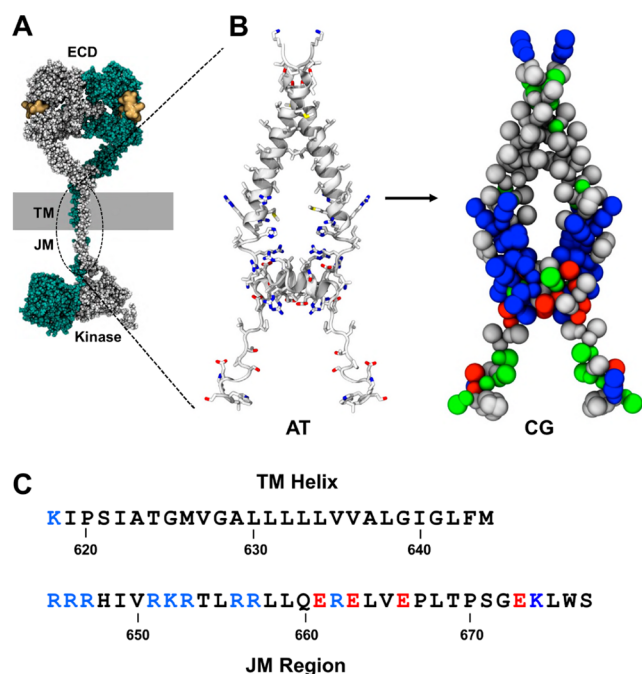
In addition to its intrinsic biological importance, the EGFR provides a well-characterized model system for the study of protein–lipid interactions. The full-length protein consists of an extracellular ligand binding ectodomain, a single transmembrane (TM) helix, a basic juxtamembrane (JM) region, and an intracellular protein kinase domain (Figure 1).<sup>6</sup> The functional activity of the receptor is known to be modulated by a glycolipid (monosialodihexosylganglioside or GM3) and by phosphatidylinositol-4,5-bisphosphate (PIP<sub>2</sub>).<sup>4,8</sup> However, despite a number of biochemical and biophysical studies, the structural and energetic bases of the interactions of these lipids

with the EGFR TM and JM regions remain to be fully characterized. This is in part due to the experimental challenges involved in studying the behavior and interactions of lipids with proteins. In this study, we have employed CG simulations and umbrella sampling to explore the free energy landscape of interactions of a model of the wild-type (WT) and mutant EGFR TM–JM dimer with a range of different lipid species present within biological membranes.

## METHODS

**NMR Structure-Derived Molecular Model.** The molecular model of the EGFR helix dimer used was derived from the experimentally determined NMR structure (PDB ID: 2M20).<sup>24</sup> Two native methionines (M626 and M644) had been mutated to prevent chemical cleavage of the protein during purification. These residues were restored to the WT residues using the mutagenesis tool implemented in PyMOL (<https://www.pymol.org/>). The sequence and structure of the model used in simulations are indicated in Figure 1. Models of the K618G and R645–7N mutants were also obtained using the PyMOL mutagenesis tool. Energy minimized atomistic (AT) models were converted to CG representation using the MARTINI2.2 force field.<sup>25</sup> Assignment of dihedral restraints to model the TM and short JM-A helices was achieved in an automated fashion using the martinize.py workflow, based on DSSP assignment. The N- and C-terminus of each monomer were modeled with neutral charge.

**CG Simulation Details.** Simulations were performed using the GROMACS ([www.gromacs.org](http://www.gromacs.org)) 5.0 and 4.6 simulation packages.<sup>26</sup> During the initial set of CG simulations, 50 ns self-assembly simulations<sup>27</sup> were performed to allow formation of a PC bilayer consisting of 700 PC lipid molecules around the TM region of the dimer. A locally developed script<sup>28</sup> was used to exchange PC molecules for other lipids to form an asymmetric bilayer containing 17 GM3 molecules (in the outer leaflet, corresponding to ca. 2.5% of all lipids within the system) and 17 PIP<sub>2</sub> molecules (in the inner leaflet, corresponding to ca. 2.5% of all lipids within the system). This provides a simplified model of the distribution and glycolipid/PI content of a plasma membrane.<sup>28,29</sup> Five-thousand steps of steepest descent energy minimization were applied to relax the system, followed by a 10 ns equilibration simulation. Two microsecond production runs were performed with different random initial velocity seeds. Temperature was maintained at 310 K utilizing a V-rescale thermostat<sup>30</sup> with a coupling constant of  $\tau_t = 1$  ps. Pressure was controlled at 1 bar using a Parrinello–Rahman barostat<sup>31</sup> with a coupling constant of  $\tau_p = 5$  ps and a compressibility of  $3 \times 10^{-4}$  bar<sup>-1</sup>. Protein, lipids, and solvent (water + ions) were coupled independently. Electrostatics and van der Waals interactions were shifted between 0 and 1.2 nm and 0.9 and 1.2 nm, respectively, and an integration time step of 20 was applied. Within umbrella sampling simulations, particle mesh Ewald was applied to improve modeling of long-range electrostatics.<sup>32</sup> Covalent bonds were constrained to their equilibrium values using the P-LINCS algorithm.<sup>26,33</sup> All simulations were run in the presence of conventional MARTINI water<sup>22</sup> and neutralized using 0.15 M NaCl. The MARTINI force field was used to describe all system components, other than PIP<sub>2</sub> and GM3, for which we employed locally parametrized versions previously described by us.<sup>28,34</sup> Simulations were visualized using VMD.<sup>35</sup> Contact analysis was performed using VMD. Contacts were calculated between each residue of the protein and the PO3 bead of the PIP<sub>2</sub> lipid headgroup, and the B1A



**Figure 1.** Overview of EGF receptor structure and sequence. (A) Composite structural diagram of dimeric EGFR based on the extracellular domain (ECD) crystal structure (PDB ID: 3NJP), the transmembrane (TM) and juxtamembrane (JM) domains NMR structure (PDB: 2M20), and the asymmetric tyrosine kinase crystal structure (PDB IDs: 2GS6, 3GOP, and 2JIU). (B) NMR-derived structure of the TM–JM dimer used as the basis of the simulations. The atomistic (AT) structure was converted to a coarse-grained (CG) model (basic residues in blue, acidic residues in red, polar residues in green, and hydrophobic residues in gray). (C) Sequence of the TM helix and JM region, with the residue numbering as in the processed mature protein sequence (UniProt entry P00533). An addition of +24 should be made to convert to the preprocessed numbering.

bead of the GM3 headgroup (see Figure S1 for lipid bead names). A 6 Å cutoff was applied to define a contact. Radial distribution functions for each lipid headgroup relative to the protein were calculated using the PO3 particle of PIP<sub>2</sub>, PO4 particle of PC, and B1A particle of GM3.

**PMF Calculations.** PMF is a key concept in statistical mechanics,<sup>36</sup> which allows one to describe the free energy profile along a particular reaction coordinate ( $d$ ), and is derived from the average distribution function  $p(d)$  along this reaction coordinate. A key challenge in calculating a PMF via MD simulation is attaining sufficient sampling of configurational space.<sup>37</sup> This may be addressed via a number of sampling techniques including umbrella sampling<sup>19</sup> and adaptive force biasing.<sup>20</sup> We have applied umbrella sampling in conjunction with a CG model to calculate PMF profiles for lipid interactions with a TM helix dimer.<sup>36</sup>

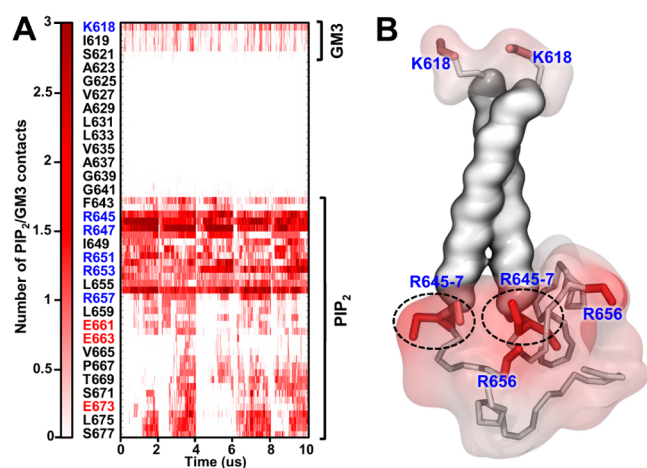
The last CG snapshot with a single bound lipid of interest was extracted from a 2  $\mu$ s production run. A 50 ns self-assembly simulation was performed to permit formation of a PC bilayer around the dimer and the bound lipid. Position restraints were applied to the protein and the lipid of interest during the self-assembly, using a force constant of 1000 kJ/mol/nm<sup>2</sup>. Protein side chains were subsequently relaxed during a 10 ns equilibration with position restraints (400 kJ/mol/nm<sup>2</sup> in the  $xy$  plane) only applied to the G625 and G628 backbone beads to prevent rotation and translation of the protein. All lipids of interest remained bound over this time course, other than PC, to which a weak positional restraint of 100 kJ/mol/nm<sup>2</sup> was applied to the PO4 bead (Figure S1). Steered MD (SMD) simulations were performed to generate a series of configurations along a reaction coordinate ranging from the lipid bound to lipid unbound states (free energy profile reaches a plateau). The lipid was pulled at a rate of 0.1 nm/ns over a distance of 2–3 nm using a force constant of 1000 kJ/mol/nm<sup>2</sup>, with the distance from the center-of-mass of G625 and G628 of the N-terminal dimerization interface of the EGFR model and a single bead of the lipid of interest defined as the 1D reaction coordinate (B1A for GM3; PO4 for PC, PS, and PG; and PO3 for PIP<sub>2</sub>). The translational and rotational motions of the protein were reduced through application of position restraints to the backbone beads of G625 and G628 using a force constant of 400 kJ/mol/nm<sup>2</sup>, with the  $z$  coordinate remaining unrestrained. For all inner leaflet SMD simulations, these restraints were also applied to M644 at the C-terminus of each helix to prevent the C-terminal portion of the helix from “following” the lipid. Additionally, a weak positional restraint of 100 kJ/mol/nm<sup>2</sup> was applied to a single bead of the lipid headgroup to limit its motional freedom along the  $y$  coordinate, in a manner similar to Arnarez et al.<sup>11,38</sup> The  $y$  coordinate within our system is defined as is the direction perpendicular to the reaction coordinate ( $x$ ) and to the bilayer normal ( $z$ ), and application of such restraints thus ensures sampling of the 1D reaction coordinate and prevents transition to 2D sampling and the associated convergence challenges. Application of these position restraints to the dimer, together with the standard dihedral restraints inherent to the MARTINI force field,<sup>39</sup> prevented any major helical transitions or conformational rearrangements within the TM region of the dimer, whereas the flexible +4 N-terminus and JM were modeled in an unrestrained fashion and could freely interact with the membrane. The subject lipid was treated separately to bulk lipids for temperature and pressure coupling. Snapshots were extracted from the SMD simulation, and these

configurations were used as input for umbrella sampling simulations. Windows were spaced asymmetrically with  $\sim$ 0.05 nm spacing at low-center-of-mass separations and  $\sim$ 0.1 nm in the bulk. Each window was run for between 500 and 1000 ns. Relative lipid–protein separations within each umbrella sampling window were maintained through application of a harmonic umbrella potential between the center-of-mass of G625 and G628 residues of the N-terminal dimerization interface and a single lipid headgroup bead, using a force constant of 1000 kJ/mol/nm<sup>2</sup>. Additionally, the same position restraints used in the SMD simulations were applied in the umbrella sampling simulations. Within each window, temperature and pressure were controlled using the Berendsen thermostat and Berendsen barostat.<sup>40</sup> The GROMACS implementation ( $g\_wham$ ) of the weighted histogram analysis method (WHAM) was used to combine and unbias the umbrella potentials, and the Bayesian bootstrapping method was utilized to estimate the error on each free energy profile.<sup>41</sup> Convergence was analyzed by comparing free energy profiles computed from nonintersecting periods of simulation time. On the basis of these observations, the first 100 ns of simulation time was excluded from analysis as equilibration. The possible introduction of bias caused by taking system configurations generated by pulling the lipid away from the dimer was tested by taking the end point of a 1  $\mu$ s umbrella sampling simulation at the maximum protein–lipid separation and pulling the lipid “backward” toward the dimer from the bulk bilayer along the same 1D reaction coordinate. This backward trajectory was then used to generate a new set of umbrella sampling simulations conducted in the same manner described previously.

## RESULTS

**Identification of Lipid Interaction Hotspots.** Lipid interactions were explored using a CG molecular model of the TM–JM dimer of EGFR based on an NMR structure (PDB ID: 2M20).<sup>24</sup> The monomeric unit consisted of a 4 residue N-terminal region, a 23 residue TM helix, and a 32 residue JM region (Figure 1). This model was embedded in a phosphatidylcholine (PC) bilayer via self-assembly simulations.<sup>27</sup> Selected PC lipid molecules were subsequently exchanged for other lipid species<sup>28</sup> to form an asymmetric bilayer containing 2.5% GM3 (in the outer leaflet) and 2.5% PIP<sub>2</sub> (in the inner leaflet), providing a simplified model of the distribution of these lipids in mammalian cell membranes.<sup>29,42</sup> CG-MD simulations were performed using the MARTINI force field.<sup>22,25,39</sup> Five repeat simulations of the dimer in the asymmetric membrane were run, each 2  $\mu$ s in duration (see Methods for full simulation details). The number of contacts formed over the course of the simulations between each protein residue and the headgroup of each lipid species was calculated. Within the outer leaflet, GM3 exhibited a strong preference for interaction with N-terminal membrane proximal residue K618, whereas within the inner leaflet, PIP<sub>2</sub> showed high frequencies of interaction with R645, R646, and R647, as well as with R656 within the cytoplasmic JM region (Figure 2). These lipid interaction hotspots identified by our simulations agree well with *in vivo* mutagenesis studies and *in vitro* functional assays that have shown that these residues are essential for the sensitivity of EGFR to GM3 and PIP<sub>2</sub>.<sup>4,8</sup> Visual inspection of the trajectory revealed that once associated each PIP<sub>2</sub> lipid molecule remained bound over the 2  $\mu$ s duration of the simulation. In contrast, the extracellular leaflet GM3 lipids



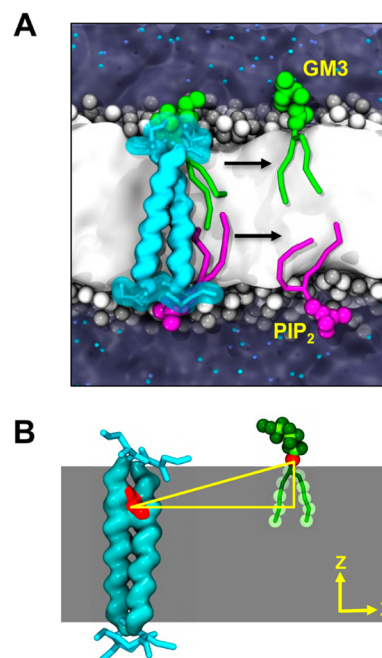


**Figure 2.** Protein–lipid interactions within a model bilayer. Lipid interaction hotspots identified by CG simulations of the TM–JM dimer in a PC bilayer containing 5% GM3 in the extracellular leaflet and 5% PIP<sub>2</sub> in the intracellular leaflet. These are shown in (A) by a contact matrix of the dynamic pattern of interaction formed between the peptide and a single bead of the headgroup moieties of PIP<sub>2</sub> and GM3 over time (see [Methods](#) for full details of contact analysis). (B) TM helix dimer (snapshot from a CG simulation) with each residue colored from white (no interaction with GM3 or PIP<sub>2</sub> headgroup) to red (high degree of interaction). Residues that formed the highest levels of contact with lipid molecules within each leaflet are labeled. Contacts were calculated over  $5 \times 2 \mu\text{s}$  simulation repeats, using a 6 Å cutoff.

exhibited a more dynamic pattern of transient receptor interactions such that multiple association and dissociation events were observed ([Figure 2A](#)).

We sought to quantify the strength of the observed lipid–protein interactions and test the importance of residue identity at the interaction sites identified as hotspots through calculation of the PMF (i.e., free energy profiles) for lipid association with both WT and mutant receptors. A final CG snapshot with a single bound lipid of interest at either of the previously determined sites was extracted from the  $2 \mu\text{s}$  simulation trajectory to provide starting structures for the PMF calculations. In each case, the snapshot corresponded to a structure in which the lipid of interest was bound between the two TM helices, approximately equidistant from each helix axis. In the case of PIP<sub>2</sub>, the flexible region of the JM after R647 was removed to leave just the first three JM residues (R645–7) present within a truncated JM region. This enabled us both to expedite data collection and to combat the technical challenge of achieving convergence in PMF calculations with a highly flexible JM region. A new PC bilayer was subsequently formed via self-assembly simulations around the position-restrained protein (see [Methods](#) for details of restraints used) and the bound lipid. The TM region of the dimer was thus conformationally “frozen” in terms of degrees of freedom, whereas the flexible +4 N-terminus and C-terminal JM regions remained unrestrained and free to dynamically interact with the membrane. The choice of membrane composition served to combat the additional convergence challenges that arise from calculations in more complex multicomponent bilayers. To generate a 1D reaction coordinate ranging from the lipid-bound state to the corresponding lipid (GM3 or PIP<sub>2</sub>) being free in the PC bilayer, a steered MD simulation was performed in which a force was applied to pull the lipid of interest away from the restrained TM helix dimer over a distance of 2–3 nm

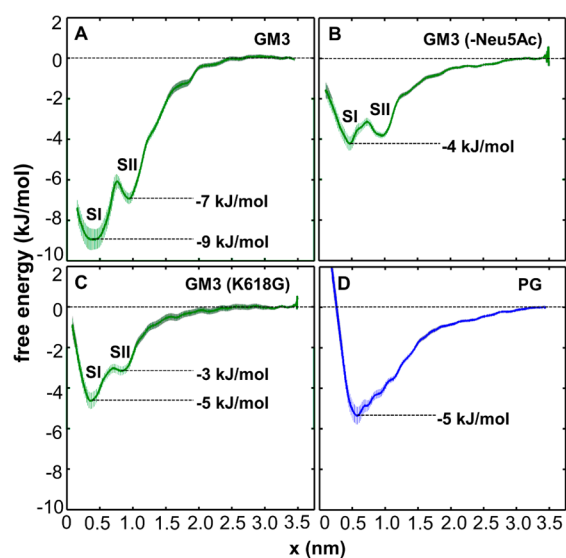
([Figure 3A,B](#)). A series of 20–40 system configurations were extracted along the reaction coordinate, each spaced by ca. 0.1



**Figure 3.** Reaction coordinate for probing the energetics of lipid interactions with the EGFR TM–JM dimer in a lipid bilayer. (A) Schematic illustration of a dimer in a membrane, with the arrows showing the approximate pathways (i.e., reaction coordinates) for evaluating the free energy profiles (i.e., potentials of mean force, PMFs) of GM3 and PIP<sub>2</sub> interactions with the protein. (B) Reaction coordinate for exploring the energetics of protein–lipid interactions illustrated for GM3. The reaction coordinate is defined as the separation along the  $x$  coordinate of the center-of-mass (red) of the two glycine residues within the N-terminal GxxGA motif of the dimer and of the B1A particle (red) of GM3.

nm. These configurations were used as starting points for umbrella sampling simulations. The subject lipid was maintained on the 1D reaction coordinate during umbrella sampling through application of lateral position restraints in a manner previously described in refs 11 and 38. Simulations were performed using Gromacs 4.6,<sup>26</sup> and free energy profiles were constructed using the *g\_wham* utility.<sup>41</sup> (See the [Methods](#) for full details of how PMF calculations were conducted.)

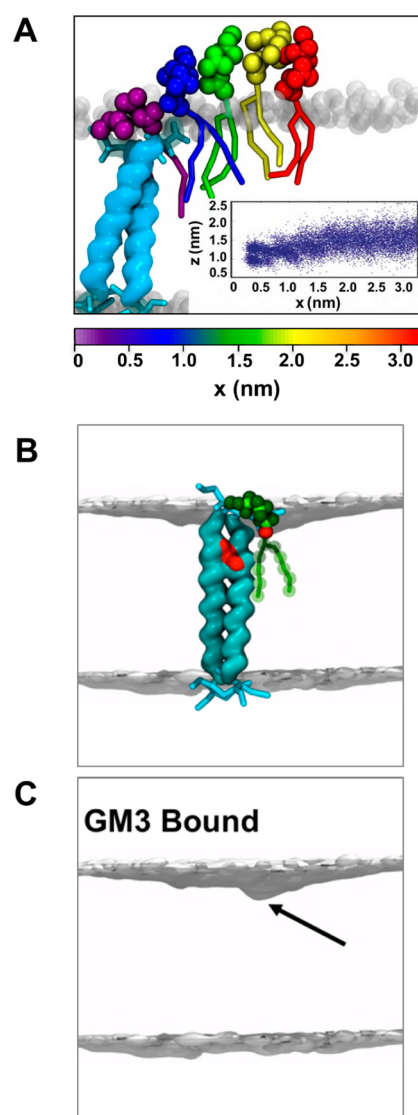
**Binding of GM3 within the Extracellular Leaflet.** Free energy profiles were calculated for GM3 and for phosphatidylglycerol (PG) interaction with the WT dimer within the outer leaflet and for GM3 interactions with a mutant dimer (K618G) ([Figure 4](#)). Additionally, the importance of the N-acetyl neuraminic acid (Neu5Ac) moiety of GM3 for protein interactions was tested by calculation of free energy profiles for GM3 lacking the Neu5Ac moiety (GM3 (–Neu5Ac)). GM3 was found to interact with the EGFR dimer with an overall minimum (state I) free energy of  $-9 \text{ kJ/mol}$ . A second local energy minimum (state II) with a free energy of  $-7 \text{ kJ/mol}$  was also observed ([Figure 4A](#)). These minima both represent bound states of GM3. In state II, the lipid headgroup is erect and parallel to the membrane normal. In contrast, in state I the GM3 lipid headgroup is tilted and penetrates further into the bilayer ([Figures 5A, S2, and S3](#)). This enables formation of optimal interactions between the N-terminal four residues of the helix dimer and the GM3 headgroup. In state I,



**Figure 4.** Outer leaflet free energy profiles. (A) GM3, (B) GM3 (–Neu5Ac), i.e., GM3 from which the terminal N-acetyl neuraminic acid has been removed, (C) GM3 and the K618G mutant dimer, and (D) phosphatidylglycerol (PG). The standard deviation estimated from bootstrapping is shown as the shaded area behind the curve.

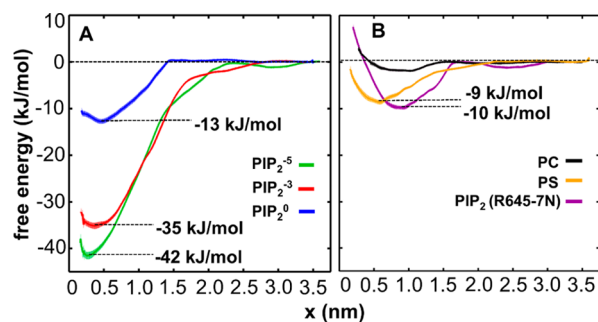
there is a degree of local bilayer deformation seen as bilayer thinning (Figure 5B,C). This local deformation is absent from state II. The second minimum thus most likely represents an energetically favorable adjustment of the tilt angle and depth of GM3 within the lipid bilayer upon moving away from the protein and the consequent release of local deformation in the surrounding PC bilayer. The ability of PMF calculations to detect these subtle differences in the behavior of a single lipid illustrates the sensitivity of the method. Mutation of residue K618 to glycine decreased the binding affinity of GM3 by 4 kJ/mol compared to the WT (Figure 4). In addition, removal of the Neu5Ac sugar from the GM3 headgroup caused a reduction in binding affinity of 5 kJ/mol (Figure 4). Both the K618G mutant and GM3 (–Neu5Ac) truncation remove one of the formal charge components of the interaction between the lysine side chain and the carboxylate group of GM3. The reduction in binding affinity by  $\sim 50\%$  demonstrates that both the cationic lysine side chain at position 618 and the anionic headgroup moiety of the glycolipid are significant contributors to the GM3–EGFR interaction. In agreement with this prediction, *in vitro* assays have shown that treatment of GM3 with neuraminidase (which cleaves Neu5Ac from the GM3 headgroup) abolishes its ability to modulate receptor activity and that K618 is likewise essential for GM3 interaction and the sensitivity of the receptor to GM3 levels within membranes.<sup>4</sup> Significantly, the simple anionic lipid PG exhibited a weaker binding affinity than GM3 (Figure 4), despite both lipids bearing a formal charge of  $-1$ . This suggests that the additional polar moieties of the GM3 headgroup play a role in providing interaction sites that allow EGFR to form selective interactions with GM3. As anticipated, the zwitterionic lipid PC exhibited no significant interaction (Figure S4).

**PIP<sub>2</sub> Binding to the EGFR TM Domain.** Within the inner leaflet of the bilayer, free energy profiles were computed for the interactions of PIP<sub>2</sub><sup>-5</sup>, PC, and PS. The importance of PIP<sub>2</sub> ionization state<sup>43</sup> was then evaluated by calculation of free energy profiles for PIP<sub>2</sub><sup>-3</sup> and PIP<sub>2</sub><sup>0</sup>, and the significance of the side chain properties at positions within the JM was



**Figure 5.** Repositioning of GM3 at the protein–lipid interface. (A) Position of the GM3 molecule (purple to red) relative to the upper leaflet of the lipid bilayer (gray) as a function of the displacement,  $x$ , along the reaction coordinate. The inset shows the position along the bilayer normal ( $z$ ) for the B1A particle (red) of GM3 as a function of the  $x$  coordinate of the lipid. (B,C) Density map of the location of the PC head groups averaged over the course of the first umbrella sampling window simulation (i.e., that with the GM3 closest to the protein). A degree of bilayer thinning is observed around the dimer and bound lipid. The approximate locations of the protein and GM3 are indicated in (B).

investigated by calculation of a profile for the interaction of PIP<sub>2</sub><sup>-5</sup> with a R645–7N mutant dimer (Figure 6). In contrast to the relatively small interaction free energies observed within the upper leaflet, PIP<sub>2</sub><sup>-5</sup> bound to the dimer with a free energy of  $-42$  kJ/mol, whereas PIP<sub>2</sub><sup>-3</sup> was bound with a free energy of  $-35$  kJ/mol. The significantly deeper wells observed in these profiles are consistent with observations from the contact analysis (Figure 2) and with visual inspection of trajectories revealing that PIP<sub>2</sub> lipids remained bound for the duration of the initial  $2 \mu\text{s}$  CG simulations. *In vivo*, PIP<sub>2</sub> lipids are suggested to exist in a mixture of the  $-3$ ,  $-4$ , and  $-5$  ionization states,<sup>44</sup> although the exact charge may depend on the lipid micro-environment.<sup>45</sup> Thus, we would expect the interaction free

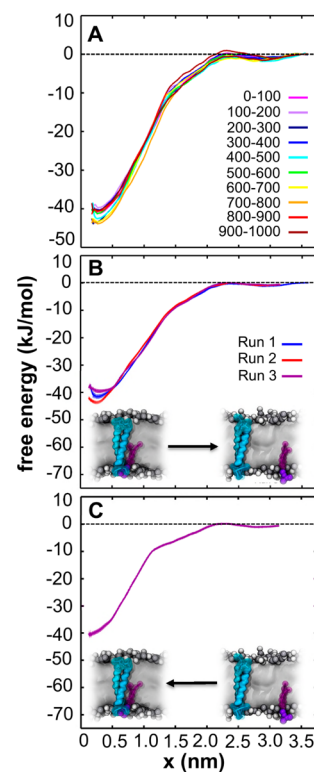


**Figure 6.** Inner leaflet free energy profiles. (A) Free energy profiles for PIP<sub>2</sub><sup>-5</sup>, PIP<sub>2</sub><sup>-3</sup>, and PIP<sub>2</sub><sup>0</sup> interactions with the WT protein. (B) Profiles for PS and PC interactions with the WT protein and PIP<sub>2</sub> interactions with the R645–7N mutant. The standard deviation estimated from bootstrapping is shown as the shaded area behind the curve.

energy of PIP<sub>2</sub> with the dimer to be about  $-40$  kJ/mol, i.e., 4 times stronger than the interaction with GM3 in the opposite leaflet of the bilayer. Mutation of residues R645–7 within the JM cationic cluster of each monomer to three asparagines resulted in a significant reduction in the strength of interaction to about  $-10$  kJ/mol. The free energy of the PIP<sub>2</sub> interaction with this asparagine (NNN) mutant is comparable of that of phosphatidylserine (PS) with the WT receptor (Figure 6), demonstrating that these residues are essential for the strong binding of PIP<sub>2</sub> lipids as well as the selectivity of EGFR for this lipid over other species present within cell membranes. Furthermore, as expected, these mutations show that the formal charge interactions between PIP<sub>2</sub> phosphate groups and the basic arginine residues contribute to a significant portion of the binding energy. In support of this, the interaction of neutral charge state PIP<sub>2</sub> lipids was ca. 30 kJ/mol weaker compared to that of the  $-5$  ionization state (Figure 6). The residues that we identify as important for protein–lipid interactions have been experimentally shown to be required for normal receptor function, with R-to-N and R-to-A mutants exhibiting significantly impaired activity. In addition, basic residues within the early JM region have been shown to be crucial in facilitating PIP<sub>2</sub> interactions and the ability of PIP<sub>2</sub> lipids to modulate receptor activity *in vivo*.<sup>8,46</sup>

When judging the utility of PMF calculations, it is important to consider their convergence.<sup>18</sup> The convergence of the PIP<sub>2</sub><sup>-5</sup> PMF calculations was judged by comparison of energy profiles calculated from multiple nonoverlapping segments of simulation time (Figure 7A). The well depth oscillated around  $-42$  kJ/mol, with no overall trend during the 1  $\mu$ s simulation window ( $30 \times 1 \mu$ s for the whole profile). The efficient convergence seen in these profiles is likely a function of the CG model that we employed, as the reduction in degrees of freedom<sup>22</sup> allows the system to attain convergence on faster time scales. The approach was found to be reproducible to within 5 kJ/mol (i.e., ca. 2 kT) by comparison of three independently calculated PMFs of the PIP<sub>2</sub><sup>-5</sup> system (Figure 7B). The direction of the force applied in the initial steered MD simulation (i.e., whether the 1D reaction coordinate was generated by pulling the lipid away from the protein or toward it) was also tested and did not show any significant influence on the profile obtained (Figure 7C), indicating the absence of a hysteresis.

In addition to more specific protein–lipid interactions, we have previously revealed lipid clustering induced by disordered

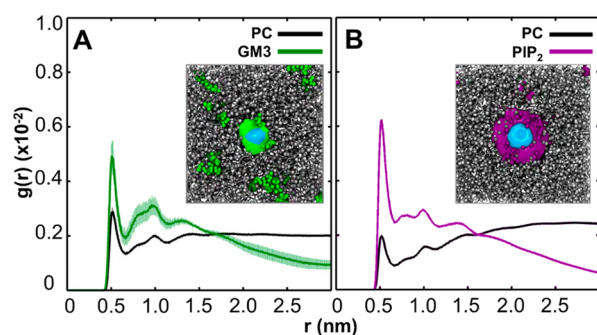


**Figure 7.** Convergence of the free energy landscape. (A) Free energy profiles for the interaction of PIP<sub>2</sub><sup>-5</sup> with the WT dimer, each calculated from nonoverlapping 100 ns segments of the simulation time within each 1  $\mu$ s window. (B) Three PIP<sub>2</sub><sup>-5</sup> profiles, each calculated from an independent repeat. In each case, a different structure was extracted from an equilibrium simulation, and a new set of SMD simulations and umbrella sampling calculations was conducted using that structure as the starting point. The well depths span a range of ca. 5 kJ/mol. (C) Profile obtained by pulling a PIP<sub>2</sub><sup>-5</sup> lipid from the bulk membrane “backward” along the 1D coordinate toward the protein during an SMD simulation and using that trajectory to define the reaction pathway for a new set of umbrella sampling windows.

JM regions in simulations of the monomeric TM–JM domains of RTKs<sup>47</sup> and cytokine receptors.<sup>48</sup> Comparable clustering of anionic lipids was seen in the current simulations of the dimeric EGFR TM–JM (Figure 8A,B). Spherical radial distribution functions (RDFs) for each lipid headgroup relative to the protein exhibited peaks at 0.5 nm, a second peak extending out to 1 nm, and a third weaker peak at 1.5 nm, indicating a degree of local lipid reorganization and enrichment of certain lipid species around the protein dimer. In the outer leaflet, the RDF revealed clustering of GM3; in the inner leaflet, clustering of PIP<sub>2</sub> is seen. The degree of local clustering (i.e., the first peak in the RDF) is greater for PIP<sub>2</sub> than for GM3, reflecting the stronger electrostatic interactions for the underlying interaction, as seen in the free energy profiles.

**Limitations.** It is important to consider the possible limitations of the model that we employed. In particular, we might expect that inclusion of the extracellular and intracellular domains of the EGFR (along with the glycans that decorate the extracellular domain<sup>49</sup>) to exert some effect on the free energy landscape of lateral lipid interaction with the TM domain. However, the use of RTK TM domain constructs and mutational studies within these systems have previously been shown to be relevant to the full-length receptor.<sup>50</sup> Furthermore,





**Figure 8.** Clustering of GM3 and PIP<sub>2</sub> lipids. Clustering of (A) GM3 in the outer leaflet and (B) PIP<sub>2</sub> in the inner leaflet around the TM + JM dimer. In each case, the radial distribution function (RDF;  $g(r)$ ) for the headgroup of PC (black) and of GM3 (green) or PIP<sub>2</sub> (purple) is shown. Calculations were performed over  $5 \times 2 \mu\text{s}$  of CG simulation, with the standard deviation indicated as the shaded area behind each curve. The area under each curve is normalized to unity to aid comparison between the different lipid species. The two insets are density maps showing the average spatial occupancy of each lipid headgroup over the outer or inner leaflet surface. GM3 is indicated in green, PIP<sub>2</sub> is in purple, and lysine or arginine density is in cyan. Calculations were performed over  $5 \times 2 \mu\text{s}$  simulation repeats. In these images, the protein has been omitted to show the lysine and arginine density maps more clearly.

by its nature, the CG MARTINI model involves a degree of approximation arising from loss of atomic detail and is likely to cause deviation from binding free energies estimated from AT simulations and/or measured experimentally. In particular, some local changes in secondary structure might be induced upon binding of regulatory lipids. The standard protein backbone dihedral restraints employed within the MARTINI model<sup>39</sup> and the lack of hydrogen-bond directionality arising from coarse-graining the system mean that possible lipid-induced conformational changes are not captured within our calculations. Along the same line, it is possible that protein and lipid protonation states may change as a function of the reaction coordinate. To capture such effects, constant-pH simulations would be required (see, e.g., ref 51). In lieu of modeling dynamic protonation, some indication of the possible effects of protonation state may be provided by our choice to calculate PIP<sub>2</sub> free energy profiles for three different ionization states, along with neutralizing protein and lipid mutagenesis. The standard MARTINI water model<sup>22</sup> is likely to oversimplify the role of water in these interactions, particularly for lipids with large solvent-exposed headgroups such as GM3. Polarizable water models<sup>52</sup> have been suggested to improve the accuracy of free energy calculations for partitioning amino acid side chains into hydrophobic environments. However, there remains an incomplete understanding of its influence on lipid–water and protein–water interactions, together with a substantial increase in the convergence time associated with its use (see, e.g., ref 53).

Despite these limitations, we note that the model has been shown to correctly predict lipid interaction sites on a number of membrane proteins,<sup>10–13</sup> and also that the MARTINI force field may be considered to be lipid-oriented as it was parametrized based on the reproduction of experimental partitioning free energies and atomic simulation parameters.<sup>39</sup> For example, free energy profiles for phospholipid removal from a DPPC bilayer show reasonable agreement in all-atom and CG simulations, which have respectively been used to

obtain estimates of ca.  $-80$  and  $-90$  kJ/mol.<sup>22,54</sup> These values may in turn be compared with an experimental estimate of ca.  $-70$  kJ/mol derived from the critical micelle concentration.<sup>14</sup> Turning to proteins, free energy profiles for the dimerization of the TM domain of glycoprotein A have yielded well-depth estimates of  $-30$  to  $-40$  kJ/mol<sup>55,56</sup> in MARTINI CG and of  $-45$  to  $-60$  kJ/mol<sup>57,58</sup> in all-atom simulations. Considering the differences in bilayer composition, helix length, and the errors associated with both methods, this suggests that broad agreement is seen between the two levels of granularity for both lipid–lipid and protein–protein interactions. Taking this together with a significant reduction in the sampling challenges that are routinely faced in AT simulations,<sup>18</sup> we expect the CG free energy profiles that we calculate to provide a reasonable first approximation of EGFR–lipid interaction free energies and of the relative effects of protein mutation.

## DISCUSSION

We have demonstrated that simulations can reveal both the strength and specificity of the interactions of regulatory lipids with the TM domain of a RTK. We show that K618, R645, R646, and R647 play an important role in EGFR interactions with modulatory lipids GM3 and PIP<sub>2</sub>. Comparison to experimental mutations confirms that CG simulations are indeed capable of identifying lipid interaction hotspots on integral membrane proteins. The significant (ca. 4-fold) difference in receptor interaction energies of modulatory lipids GM3 and PIP<sub>2</sub> and the differences in the dynamics of these interactions are suggestive of possible differential modes of receptor modulation. To date, these lipids have been proposed to modulate EGFR activity via influences on dimerization propensity,<sup>4</sup> direct conformational stabilization and orientation of intracellular JM and kinase regions,<sup>59,60</sup> and larger scale effects on receptor clustering at the cell surface.<sup>47,61</sup> RTK helix dimerization energies have previously been estimated to be on the order of ca.  $-20$  kJ/mol for EGFR<sup>62</sup> to ca.  $-60$  kJ/mol for EphA1.<sup>63</sup> The strength of interactions of PIP<sub>2</sub> lipids revealed in the current simulations (ca.  $-40$  kJ/mol) is supportive of the notion that lipid modulation of EGFR activity could occur via direct conformational stabilization and effects on EGFR helix dimerization rather than solely via effects on receptor clustering.<sup>61</sup> In contrast, the weaker and more transient interactions of GM3 that we observe may be consistent with a more raft-like model of receptor modulation, which requires only weak lateral interactions between receptors and lipids.<sup>64</sup> This is consistent with observations that GM3 inhibits EGFR activity only within lipid-disordered/lipid-ordered (i.e., raft-like) membranes containing sphingomyelin and cholesterol<sup>4</sup> and not in lipid-disordered phase bilayers. Characterization of the underlying energetics of EGFR–lipid interactions is thus an essential first step toward a mechanistic understanding of these interactions. In particular, a key prediction arising from our results is that the comparable magnitudes of PIP<sub>2</sub> binding energies and those of helix dimerization indicate that PIP<sub>2</sub> lipids may be able to modulate receptor dimerization by direct competition with TM helices, as suggested by, e.g., Stangl and Schneider.<sup>65</sup> This could occur both by influencing the relative monomer/dimer populations of the EGFR and by driving possible conformational transitions within existing dimers, as hinted at by recent structural studies of a presumed inactive state of the TM domain.<sup>66</sup> This prediction could be tested either via reconstitution of the EGFR in membranes of defined lipid composition<sup>4</sup> or via modulation of the lipid composition

of the membranes of living cells.<sup>67</sup> Depending on the conformation the extracellular domain adopts and the type of glycosylation, a number of glycans can be positioned spatially close to the N-termini of the helices with which GM3 interacts.<sup>49</sup> It therefore seems likely that in certain states these glycans may influence lateral lipid interaction with the TM domain. Further combined simulation and experimental studies would be welcome in elucidating the influence of glycans and their relationship with regulatory lipid species and the TM domain. Ultimately, structural studies may reveal lipid-binding sites at high resolution and provide further insight into the mechanism of lipid modulation.

## CONCLUSIONS

In summary, this well-studied model system allows us to demonstrate that calculation of free energy profiles for lipid interactions via CG simulation represents a powerful generalizable tool for the dissection and prediction of protein–lipid interactions and selectivity within cell membranes. A comparable approach has been previously applied to probe the specific binding of cardiolipin to cytochrome C oxidase<sup>11</sup> and, more recently, for the binding of other lipid species.<sup>38</sup> Taken together, these studies suggest that the CG lipid PMF methodology may be generally applicable to a wide range of membrane proteins and lipid types. This approach provides important biophysical quantification of the relative strengths of lipid interaction at specific binding sites and thus complements structural insights from crystallographic,<sup>68</sup> computational,<sup>69</sup> and biophysical<sup>70</sup> approaches.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.6b01387.

Additional information on CG lipids used in the simulations and particle nomenclature (Figure S1); positioning of the GM3 lipid in the bilayer (Figures S2 and S3); and a control showing a free energy profile for PC interaction with the N-terminus of the dimer (Figure S4) (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

\*(H.K.) E-mail: heidi.koldso@deshawresearch.com.

\*(M.S.P.S.) E-mail: mark.sansom@bioch.ox.ac.uk. Phone: +44 (0)1865-613212.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

Research in M.S.P.S.'s group is funded by grants from the BBSRC, EPSRC, and the Wellcome Trust. H.K. is an Alfred Benzon research fellow. G.H. is funded by the MRC. This work used the ARCHER UK National Supercomputing Service (<http://www.archer.ac.uk>). The authors thank Dr. Ünal Coskun for useful discussions.

## REFERENCES

(1) Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochim. Biophys. Acta, Biomembr.* **2014**, *1838*, 1548–1559.

(2) Zocher, M.; Zhang, C.; Rasmussen, S. G. F.; Kobilka, B. K.; Müller, D. J. Cholesterol increases kinetic, energetic, and mechanical stability of the human beta(2)-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E3463–E3472.

(3) Hansen, S. B.; Tao, X.; Mackinnon, R. Structural basis of PIP<sub>2</sub> activation of the classical inward rectifier K<sup>+</sup> channel Kir2.2. *Nature* **2011**, *477*, 495–498.

(4) Coskun, Ü.; Grzybek, M.; Drechsel, D.; Simons, K. Regulation of human EGF receptor by lipids. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 9044–9048.

(5) Laganowsky, A.; Reading, E.; Hopper, J. T. S.; Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* **2013**, *8*, 639–651.

(6) Bessman, N. J.; Freed, D. M.; Lemmon, M. A. Putting together structures of epidermal growth factor receptors. *Curr. Opin. Struct. Biol.* **2014**, *29*, 95–101.

(7) Gschwind, A.; Fischer, O. M.; Ullrich, A. Timeline - The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat. Rev. Cancer* **2004**, *4*, 361–370.

(8) Michailidis, I. E.; Rusinova, R.; Georgakopoulos, A.; Chen, Y.; Iyengar, R.; Robakis, N. K.; Logothetis, D. E.; Baki, L. Phosphatidylinositol-4,5-bisphosphate regulates epidermal growth factor receptor activation. *Pfluegers Arch.* **2011**, *461*, 387–397.

(9) Pike, L. J.; Casey, L. Cholesterol levels modulate EGF receptor-mediated signaling by altering receptor function and trafficking. *Biochemistry* **2002**, *41*, 10315–10322.

(10) Hedger, G.; Sansom, M. S. P. Lipid interaction sites on channels, transporters and receptors: recent insights from molecular dynamics simulations. *Biochim. Biophys. Acta, Biomembr.* **2016**, DOI: 10.1016/j.bbmem.2016.02.037.

(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) 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.

(13) 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.

(14) Tieleman, D. P.; Marrink, S. J. Lipids out of equilibrium: Energetics of desorption and pore mediated flip-flop. *J. Am. Chem. Soc.* **2006**, *128*, 12462–12467.

(15) Lemkul, J. A.; Bevan, D. R. Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics. *J. Phys. Chem. B* **2010**, *114*, 1652–1660.

(16) Corry, B.; Thomas, M. Mechanism of ion permeation and selectivity in a voltage gated sodium channel. *J. Am. Chem. Soc.* **2012**, *134*, 1840–1846.

(17) Huang, X.; Zhao, X.; Zheng, F.; Zhan, C.-G. Cocaine esterase-cocaine binding process and the free energy profiles by molecular dynamics and potential of mean force simulations. *J. Phys. Chem. B* **2012**, *116*, 3361–3368.

(18) Gumbart, J. C.; Roux, B.; Chipot, C. Standard binding free energies from computer simulations: what is the best strategy? *J. Chem. Theory Comput.* **2013**, *9*, 794–802.

(19) Torrie, G. M.; Valleau, J. P. Monte-Carlo free-energy estimates using non-Boltzmann sampling - application to subcritical Lennard-Jones fluid. *Chem. Phys. Lett.* **1974**, *28*, 578–581.

(20) Hénin, J.; Chipot, C. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J. Chem. Phys.* **2004**, *121*, 2904–3004.

(21) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini model. *Chem. Soc. Rev.* **2013**, *42*, 6801–6822.

(22) Marrink, S. J.; Risselada, 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) Bennett, W. F. D.; MacCallum, J. L.; Hinner, M. J.; Marrink, S. J.; Tieleman, D. P. Molecular view of cholesterol flip-flop and chemical



- potential in different membrane environments. *J. Am. Chem. Soc.* **2009**, *131*, 12714–12720.
- (24) Endres, N. F.; Das, R.; Smith, A. W.; Arkhipov, A.; Kovacs, E.; Huang, Y. J.; Pelton, J. G.; Shan, Y. B.; Shaw, D. E.; Wemmer, D. E.; et al. Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **2013**, *152*, 543–556.
- (25) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periolo, X.; Tieleman, D. P.; Marrink, S. J. Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theory Comput.* **2013**, *9*, 687–697.
- (26) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (27) Scott, K. A.; Bond, P. J.; Ivetac, A.; Chetwynd, A. P.; Khalid, S.; Sansom, M. S. P. Coarse-grained MD simulations of membrane protein-bilayer self-assembly. *Structure* **2008**, *16*, 621–630.
- (28) Koldsø, H.; Shorthouse, D.; Hélie, J.; Sansom, M. S. P. Lipid clustering correlates with membrane curvature as revealed by molecular simulations of complex lipid bilayers. *PLoS Comput. Biol.* **2014**, *10*, e1003911.
- (29) van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 112–124.
- (30) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (31) Parrinello, M.; Rahman, A. Polymorphic transitions in single-crystals - a new molecular-dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (32) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (33) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINC: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (34) Stansfeld, P. J.; Hopkinson, R. J.; Ashcroft, F. M.; Sansom, M. S. P. The PIP<sub>2</sub> binding site in Kir channels: definition by multi-scale biomolecular simulations. *Biochemistry* **2009**, *48*, 10926–10933.
- (35) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (36) Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **1995**, *91*, 275–282.
- (37) Gambin, Y.; Lopez-Esparza, R.; Reffay, M.; Sierecki, E.; Gov, N. S.; Genest, M.; Hodges, R. S.; Urbach, W. Lateral mobility of proteins in liquid membranes revisited. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 2098–2102.
- (38) Arnarez, C.; Marrink, S. J.; Periolo, X. Molecular mechanism of cardiolipin-mediated assembly of respiratory chain supercomplexes. *Chem. Sci.* **2016**, DOI: 10.1039/C5SC04664E.
- (39) Monticelli, L.; Kandasamy, S. K.; Periolo, 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.
- (40) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (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) van Meer, G.; de Kroon, A. I. P. M. Lipid map of the mammalian cell. *J. Cell Sci.* **2011**, *124*, 5–8.
- (43) Kooijman, E. E.; King, K. E.; Gangoda, M.; Gericke, A. Ionization properties of phosphatidylinositol polyphosphates in mixed model membranes. *Biochemistry* **2009**, *48*, 9360–9371.
- (44) McLaughlin, S.; Wang, J.; Gambhir, A.; Murray, D. PIP<sub>2</sub> and proteins: interactions, organization, and information flow. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 151–175.
- (45) Suh, B. C.; Hille, B. PIP<sub>2</sub> is a necessary cofactor for ion channel function: How and why? *Annu. Rev. Biophys.* **2008**, *37*, 175–195.
- (46) Red Brewer, M.; Choi, S. H.; Alvarado, D.; Moravcevic, K.; Pozzi, A.; Lemmon, M. A.; Carpenter, G. The juxtamembrane region of the EGF receptor functions as an activation domain. *Mol. Cell* **2009**, *34*, 641–651.
- (47) Hedger, G.; Sansom, M. S. P.; Koldsø, H. The juxtamembrane regions of human receptor tyrosine kinases exhibit conserved interaction sites with anionic lipids. *Sci. Rep.* **2015**, *5*, 9198.
- (48) Koldsø, H.; Sansom, M. S. P. Local lipid reorganization by a transmembrane protein domain. *J. Phys. Chem. Lett.* **2012**, *3*, 3498–3502.
- (49) Kaszuba, K.; Grzybek, M.; Orłowski, A.; Danne, R.; Rog, T.; Simons, K.; Coskun, U.; Vattulainen, I. N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 4334–4339.
- (50) Sarabipour, S.; Hristova, K. FGFR3 transmembrane domain interactions persist in the presence of its extracellular domain. *Biophys. J.* **2013**, *105*, 165–171.
- (51) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. Constant-pH molecular dynamics using continuous titration coordinates. *Proteins: Struct., Funct., Genet.* **2004**, *56*, 738–752.
- (52) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, *6*, e1000810.
- (53) Naughton, F. B.; Kalli, A. C.; Sansom, M. S. P. Association of peripheral membrane proteins with membranes: Free energy of binding of GRP1 PH domain with PIP-containing model bilayers. *J. Phys. Chem. Lett.* **2016**, *7*, 1219–1224.
- (54) Sapay, N.; Bennett, W. F. D.; Tieleman, D. P. Thermodynamics of flip-flop and desorption for a systematic series of phosphatidylcholine lipids. *Soft Matter* **2009**, *5*, 3295–3302.
- (55) Janosi, L.; Prakash, A.; Doxastakis, M. Lipid-modulated sequence-specific association of glycoprotein A in membranes. *Biophys. J.* **2010**, *99*, 284–292.
- (56) Sengupta, D.; Marrink, S. J. Lipid-mediated interactions tune the association of glycoprotein A helix and its disruptive mutants in membranes. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12987–12996.
- (57) Hénin, J.; Pohorille, A.; Chipot, C. Insights into the recognition and association of transmembrane  $\alpha$ -helices. The free energy of  $\alpha$ -helix dimerization in glycoprotein A. *J. Am. Chem. Soc.* **2005**, *127*, 8478–8484.
- (58) Kuznetsov, A. S.; Polyansky, A. A.; Fleck, M.; Volynsky, P. E.; Efremov, R. G. Adaptable lipid matrix promotes protein-protein association in membranes. *J. Chem. Theory Comput.* **2015**, *11*, 4415–4426.
- (59) Arkhipov, A.; Shan, Y.; Das, R.; Endres, N. F.; Eastwood, M. P.; Wemmer, D. E.; Kuriyan, J.; Shaw, D. E. Architecture and membrane interactions of the EGF receptor. *Cell* **2013**, *152*, 557–569.
- (60) Abd Halim, K. B.; Koldsø, H.; Sansom, M. S. P. Interactions of the EGFR juxtamembrane domain with PIP-containing lipid bilayers: Insights from multiscale molecular dynamics simulations. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 1017–1025.
- (61) Wang, Y.; Gao, J.; Guo, X. D.; Tong, T.; Shi, X. S.; Li, L. Y.; Qi, M.; Wang, Y. J.; Cai, M. J.; Jiang, J. G.; et al. Regulation of EGFR nanocluster formation by ionic protein-lipid interaction. *Cell Res.* **2014**, *24*, 959–976.
- (62) Prakash, A.; Janosi, L.; Doxastakis, M. Self-association of models of transmembrane domains of ErbB receptors in a lipid bilayer. *Biophys. J.* **2010**, *99*, 3657–3665.
- (63) Chavent, M.; Chetwynd, A. P.; Stansfeld, P. J.; Sansom, M. S. Dimerization of the EphA1 receptor tyrosine kinase transmembrane domain: Insights into the mechanism of receptor activation. *Biochemistry* **2014**, *53*, 6641–6652.
- (64) Lingwood, D.; Simons, K. Lipid rafts as a membrane-organizing principle. *Science* **2010**, *327*, 46–50.
- (65) Stangl, M.; Schneider, D. Functional competition within a membrane: Lipid recognition vs. transmembrane helix oligomerization. *Biochim. Biophys. Acta, Biomembr.* **2015**, *1848*, 1886–1896.
- (66) Bocharov, E. V.; Lesovoy, D. M.; Pavlov, K. V.; Pustovalova, Y. E.; Bocharova, O. V.; Arseniev, A. S. Alternative packing of EGFR

transmembrane domain suggests that protein-lipid interactions underlie signal conduction across membrane. *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858*, 1254–1261.

(67) Coban, O.; Zanetti-Dominguez, L. C.; Matthews, D. R.; Rolfe, D. J.; Weitsman, G.; Barber, P. R.; Barbeau, J.; Devaughes, V.; Kampmeier, F.; Winn, M.; et al. Effect of phosphorylation on EGFR dimer stability probed by single-molecule dynamics and FRET/FLIM. *Biophys. J.* **2015**, *108*, 1013–1026.

(68) Palsdottir, H.; Hunte, C. Lipids in membrane protein structures. *Biochim. Biophys. Acta, Biomembr.* **2004**, *1666*, 2–18.

(69) Stansfeld, P. J.; Goose, J. E.; Caffrey, M.; Carpenter, E. P.; Parker, J. L.; Newstead, N.; Sansom, M. S. P. MemProtMD: automated insertion of membrane protein structures into explicit lipid membranes. *Structure* **2015**, *23*, 1350–1361.

(70) Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **2014**, *510*, 172–175.