Title

Human class B1 GPCR modulation by plasma membrane lipids.

Authors

Kin W Chao,[1,2]† Linda Wong[3], Affiong I Oqua[3], Jas Kalayan[2], Yusman Manchanda[3], James Gebbie-Rayet[2], George Hedger[1,*], Alejandra Tomas[3,*], Sarah L Rouse[1,*]

Affiliations

[1] Department of Life Sciences,
Sir Ernst Chain Building,
Imperial College London,
London,
SW7 2AZ,
UK.

[2] Scientific Computing Department,
Science and Technology Facilities Council,
Daresbury Laboratory,
Warrington
WA4 4AD,
UK.

[3] Department of Metabolism, Digestion and Reproduction,
Section of Cell Biology and Functional Genomics,
Imperial College London,
London,
W12 0NN,
UK.

† Present address:
School of Chemistry and Chemical Engineering,
University of Southampton,
Southampton,
SO17 1BJ,
UK

*to whom correspondence should be addressed.
Email: s.rouse@imperial.ac.uk, a.tomas-catala@imperial.ac.uk, g.hedger@imperial.ac.uk.

Abstract

The class B1 G protein-coupled receptor (GPCR) subfamily is a class of receptors known for their regulatory roles in metabolism and neuronal activity and as important drug targets. Lipids play key functional roles in modulation of GPCR signalling, yet our understanding of the molecular level detail of specific lipid interactions with class B1 GPCRs remains limited. Here we present coarse-grained molecular dynamics (MD) simulations of the active and inactive states of 15 human class B1 family members and use aiida-gromacs to capture full provenance for the set-up of simulations in complex plasma membranes. Receptors exhibit state-dependent lipid interactions with the regulatory sterol cholesterol and phospholipid phosphatidylinositiol-3,4-bisphosphate (PIP_2) at defined locations on the receptor surface. Global analysis of trends across the subfamily reveals conserved patterns of lipid interaction dynamics. The glycosphingolipid GM3 exerts a modulatory influence on the dynamics of class B1 extracellular domains in both simulations and *in vitro* time-resolved FRET assays.

Introduction

Class B1 G protein-coupled receptors (GPCRs) are a subfamily of 15 membrane proteins, which are grouped into 5 further subfamilies: Glucagon-like, PTH-like, CT-like, CRF-like and PACAP/VIP-like [1] **(**Fig. 1A**)**. These receptors are involved in regulating critical biological processes such as metabolism and neuronal activity and are important drug targets for multiple diseases [2]. The substrates of class B1 GPCRs are polypeptide hormones, and these receptors are typically characterised by a having a large extracellular domain (ECD) believed to be important in the binding affinity of peptide agonist and signal transduction [3]. Within the membrane region their structure follows the canonical GPCR seven transmembrane helix bundle. Upon activation, the

68 principal conformational change is the outward movement of the cytoplasmic portion
69 of transmembrane (TM) helix TM6. This exposes a shallow groove with which select
70 G proteins may associate, leading to nucleotide exchange and propagation of signalling
71 to downstream cascades [4,5].

72 The plasma membrane environment in which class B1 receptors reside is a key
73 modulator of membrane protein function [6,7]. Several lipid species are reported to have
74 modulatory effects on GPCRs, including cholesterol, phosphatidylinositol-4,5-
75 bisphosphate [PI4,5)P_2], and the glycosphingolipid GM3 [8–13]. These modulatory lipids
76 can influence multiple aspects of GPCR function, including the conformational
77 landscape, G protein and arrestin coupling, and oligomerisation state [14–16]. Such
78 modulation can be brought about by altering the biophysical properties of the local
79 membrane environment, as well as by specific binding of modulatory lipids to allosteric
80 sites on the GPCR [16,17]. Within the class B1 subfamily, cholesterol has been shown to
81 modulate GLP1R signalling responses and biased agonism [18]. The presence of
82 cholesterol is also required for GLP1R internalisation, optimal receptor clustering and
83 cAMP responses [19]. However, the molecular level interactions and dynamics which
84 underlie these functional effects and their conservation across the wider subfamily
85 remain obscure.

86 Recent technological development of cryo-electron microscopy (cryo-EM) approaches
87 [20] for GPCR structure determination has led to the elucidation of the active state
88 conformations of all 15 class B1 GPCRs [21]. However, interpretation of lipid interactions
89 is currently limited due to the use of detergent micelles and non-native lipid
90 compositions within nanodiscs. Likewise, the inactive states of many class B1 GPCRs
91 remain unresolved. In the present study, we leverage advances in simulation

92 methodology [22], experimental lipidomics [23], and GPU computing [24], to reunite
93 experimentally determined structures with their complex plasma membrane
94 environment using large-scale physics-based MD simulations across the entire
95 subfamily. We further leverage recent advances in protein structure prediction models
96 [25,26] to interrogate the lipid interactions of the predicted inactive states. During the
97 writing of this manuscript, a new AI foundation model, Chai-1, emerged, which enables
98 the prediction of small molecules in complex with a given protein with AF3-like
99 accuracy [27]. We applied Chai-1 to predict GPCR-cholesterol, and GPCR-PI(4,5)P_2
100 complexes, and integrated our findings with the results of physics-based MD
101 simulations. Overall, we identify a conserved pattern of PIP_2 binding at the functionally
102 important TM6/7-H8 interface, a number of deep cholesterol binding sites, and GM3
103 modulation of ECD dynamics, supported by *in vitro* time resolved FRET assays.

104 **Results**

105 We simulated each receptor in a model plasma membrane environment, using the
106 Martini 3 coarse grained force field. We generated active and inactive state models of
107 each class B1 receptor. Available experimental structures of each receptor are shown
108 in Fig. 1A and summarised in Table S1. Each simulation was performed in triplicate
109 for 10 μs, totalling approximately 1.0 ms simulation time.

110 The vast array of methods and tools used to setup initial atomic configurations and
111 compositions has resulted in complications in capturing protocols used in biomolecular
112 simulations [28]. With scientific research methodologies focused on *why* and *what* was
113 done to produce simulations, the *how* is often missing, which causes difficulties in
114 reproducing and extending research. Our study uses over 30 different simulation setups
115 with over 60 independent trajectories to analyse, necessitating novel approaches to

capture full provenance. Here we present the first application of aiida-gromacs to capture all the metadata associated with this study to allow complete MD provenance including the simulation workflow and analysis (Fig. 1B). The AiiDA python framework [29] extensively used in first-principle calculations in computational chemistry, provides a powerful framework for describing the flow of data and enabling the reproduction of complex scientific workflows. The aiida-gromacs plugin (https://aiida-gromacs.readthedocs.io) extends AiiDA to enable the capture and description of biomolecular MD simulations using the GROMACS package. By providing all the metadata as well as the production trajectories and PyLipID analyses, we aim to facilitate greater accessibility to our simulations, towards enabling reproducibility and machine learning approaches.

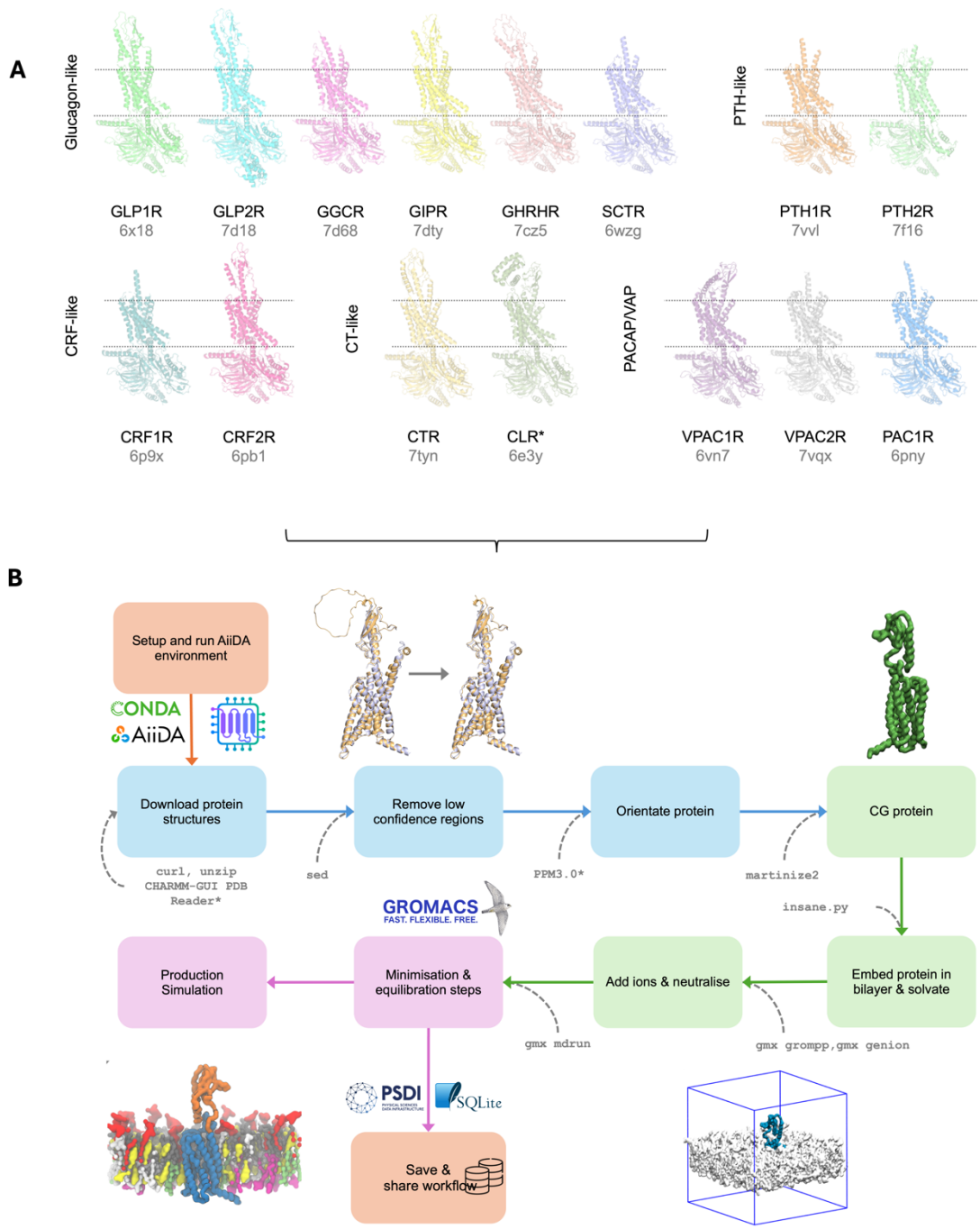


Fig. 1. Modelling and simulation workflow for class B1 GPCRs. (**A**) Structures of the 15 class B1 GPCRs in the active state used here with associated PDB accession codes. For CLR, the structure of CGRP used as the template in GPCRdb is shown. The approximate position in the membrane is shown by dotted lines. (**B**) Schematic of the B1 GPCR simulation workflow steps captured using aiida-gromacs. Orange boxes represent steps involved in enabling provenance capture and sharing, blue boxes represent steps for processing the protein structure, green boxes represent parametrisation and building of the molecular system, and purple boxes represent MD

simulation steps. The CLI tools used in each workflow step are shown in grey text. CHARMM-GUI PDB Reader[30] was used to add missing atoms, and PPM3.0[31] was used to predict membrane orientations; these steps are not yet captured by AiiDA-gromacs.

Broad trends in cholesterol, GM3 and PIP_2 interactions for class B1 GPCRs

Analysis of total lipid residence time, binding site occupancy (**Fig. 2**), and membrane contact properties (**Fig. S1**) in active *versus* inactive states across all Class B1 receptors revealed a number of trends for the known regulatory lipids cholesterol, GM3, and PIP_2. Calculations for the non-regulatory lipid POPC, are also shown in **Fig. S1**. We map the occupancy and binding site profile for the active and inactive states (cholesterol, GM3, and PIP_2 for each receptor onto their respective sequences using the PyLipID [32] package **(Fig. S2-S4)**. To assess if the combined 30 μs sampling time per receptor state is sufficient we used SCTR as a test case and performed longer timeframe (3 x 100 μs) simulations of SCTR in each state, using different initial membrane configurations (**Fig. S5**). These properties, as well as 3D volumetric density maps (**Fig S6**) are comparable for the 10 and 100 μs timeframes. Further analysis includes an evaluation of lipid diffusion properties across these timescales (**Fig. S7**) as well as estimated errors for calculated cholesterol occupancies per residue for 10 μs and 100 μs (**Fig. S8**). These analyses indicate that three independent replicas each of 10 μs duration is sufficient to capture the lipid binding properties of SCTR.

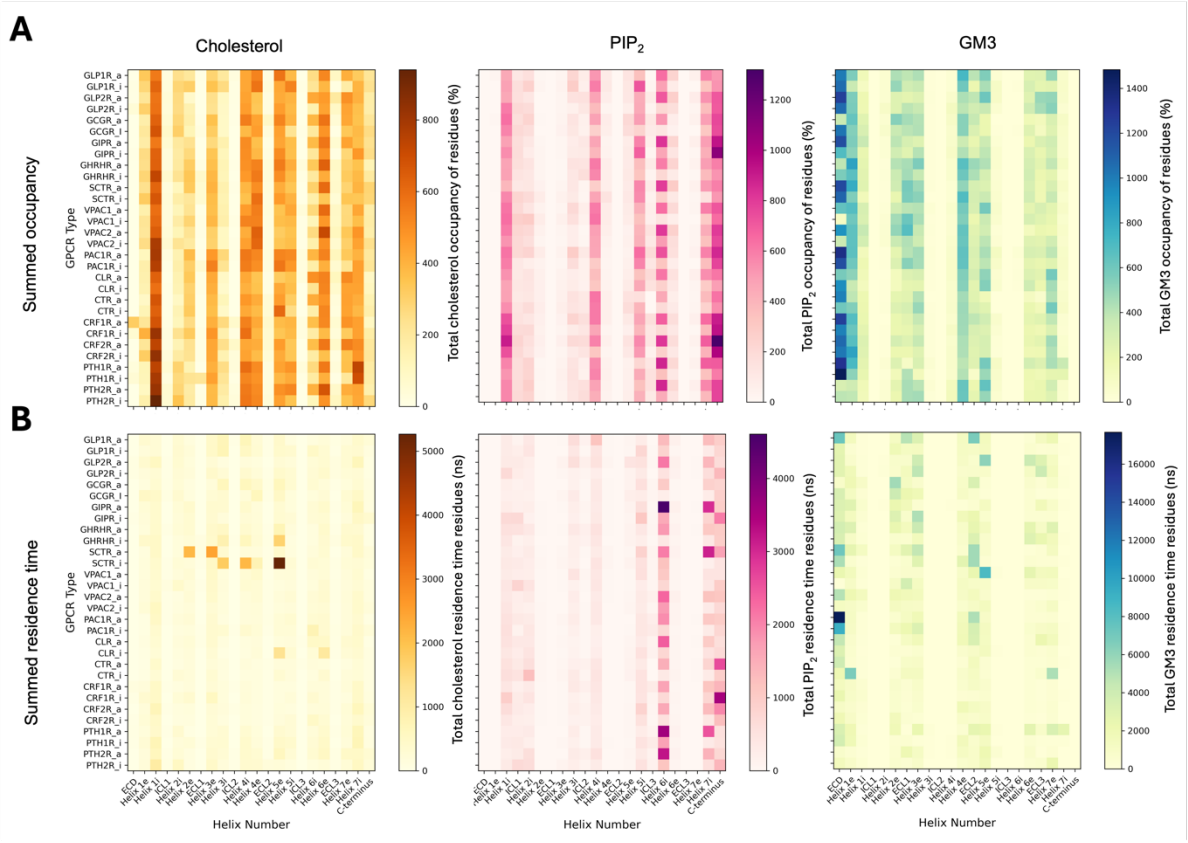


Fig. 2. Global analysis of class B1 GPCR interactions with the regulatory lipids cholesterol, PIP_2 and GM3. Lipid occupancy (**A**) and lipid residence time (**B**) is shown for each class B1 GPCR as a function of ECD, helix, and loop number. Helix 1e denotes the extracellular half of Helix 1, and Helix 1i the intracellular half. Active states are denoted by _a, and inactive by _i. Per residue lipid occupancies were computed over 3 x 10 µs of production run simulation time using PyLipID [32] and summed over the constituent residues of each region. The residue ranges defining each region correspond to those from GPCRdb [33].

All receptors show interactions with cholesterol, PIP_2 and GM3, with each lipid exhibiting distinct interaction fingerprints with defined regions of the receptors. Within the inner leaflet of the membrane, the most pronounced changes between active and inactive states for cholesterol and PIP_2 contacts occur in the functionally critical TM6 region, which undergoes an outward movement upon activation (**Fig. S1-S4**). In contrast, GM3 interactions show no apparent patterns between inactive and active states, with ECDs of both states contacting GM3 to differing degrees. Notable binding sites based on occupancy are observed for cholesterol (TM1), PIP_2 (TM6/7-H8), and

172 GM3 (ECD). The overall residence times are observed to be in the order GM3 > PIP_2
173 > cholesterol, whilst lipid occupancy is GM3 ~ PIP_2 > cholesterol. For the non-
174 regulatory lipid POPC, only PAC1R in the active state displayed high residence time
175 binding (**Fig. S1**).

176 As well as capturing the overall patterns of lipid interactions across the B1 receptors,
177 these overview plots allow us to highlight multiple outliers to characterise further. We
178 explore these properties in further detail in the following sections.

179 **Putative novel cholesterol binding sites for class B1 GPCRs**

180 We analysed the top predicted cholesterol binding sites in terms of residence time and
181 occupancy (**Fig. S9, 10**). Several class B1 receptor structures have been resolved in the
182 presence of cholesterol hemisuccinate, with cholesterol modelled into the map density
183 [34–40] (**Fig. 3A**). Across these, a single cholesterol binding site is conserved at TM4, with
184 the conserved residue $W^{4.50}$, using the Wootten numbering scheme [41], proposed to
185 mediate this interaction [39,40]. A plot of the cholesterol occupancy mapped onto the
186 aligned sequences indicates that the $W^{4.50}$ interaction is conserved during our
187 simulations (**Fig. 3B, Fig. S2**). An analysis of all cholesterol headgroup positions
188 mapped to the receptor TM bundle shows that these poses are sampled during our
189 simulations (**Fig. 3C**). However, the top cholesterol sites defined in terms of occupancy
190 and residence time corresponded to different receptor binding sites (**Fig. S9,10**).

191 Approximately 40% of the top cgMD sites across all active and inactive receptor states
192 correspond to the polar hydroxyl group of cholesterol binding in the center of the
193 membrane, with one or more such sites existing for every receptor (**Fig. S9**). This "deep
194 membrane" binding mode of cholesterol [42], with cholesterol perpendicular to the

195 membrane and its hydroxyl group buried between two TM, has been previously
196 observed in several GPCR simulation studies [42–45] but to our knowledge has not been
197 captured in X-ray crystallography or cryo-EM structures of any GPCR to date, although
198 NMR approaches have determined different affinity cholesterol binding sites [46]. We
199 also observe deep membrane binding modes where the sterol groups of the cholesterol
200 are packed against the TM bundle (**Fig. S9**). Defining contacts as transient (< 20 ns),
201 short (20-100 ns), medium (100 – 300 ns) or long-lived (> 300 ns)[47] we observe only
202 two instances of long-lived cholesterol binding, both for SCTR, which we analyse
203 further below. Across the B1 family, for deep membrane sites the
204 transient:short:medium ratio is 0:59:39, and for the canonical (bulk membrane
205 orientation) binding mode the transient:short:medium ratio is 16:68:14. Therefore the
206 deep membrane binding modes shift towards longer time-frame binding, whilst the
207 'normal' binding modes include high occupancy, transient binding of cholesterol,

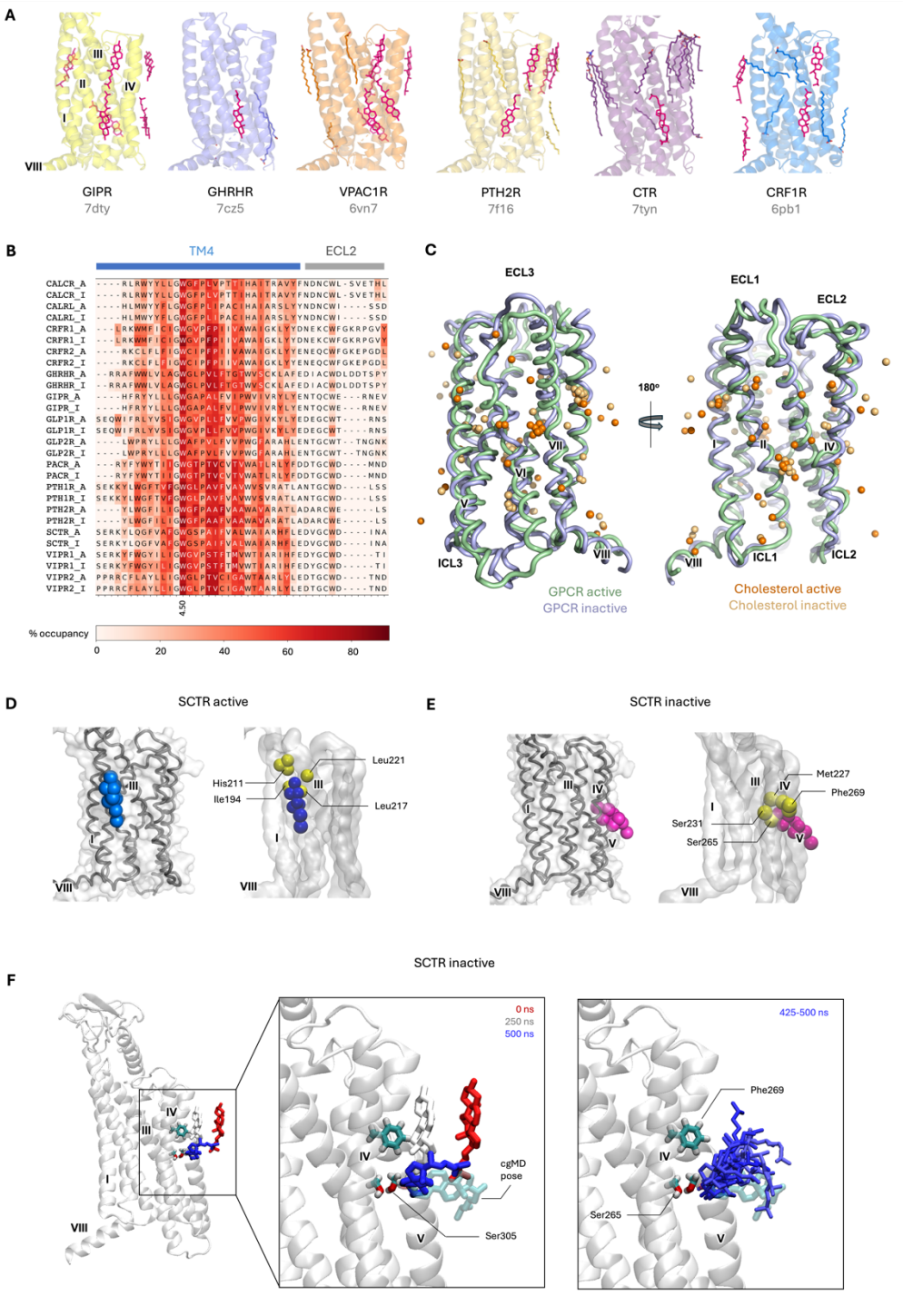


Fig. 3. Cholesterol interactions with class B1 GPCRs. (**A**) Published experimental structures of class B1 receptors containing resolved cholesterol molecules. Structures of the TM regions are aligned, with helix labels shown for GIPR. Cholesterol and cholesterol analogue molecules are shown in magenta stick representation. (**B**) Cholesterol occupancy from cgMD simulations mapped onto sequence alignment across all 15 receptors. (**C**) Positions of cholesterol molecules from simulations mapped onto the CTR receptor. Active (green) and inactive (blue) states of the CTR are shown as a backbone trace. Spheres represent hydroxyl headgroups of cholesterol for the top

ranked interaction sites extracted from simulations of active (dark orange) and inactive (light orange) states of all 15 class B1 GPCRs. (**D**) The highest residence time cholesterol site from cgMD simulations for SCTR in the active state, with residues within 6 Å of the cholesterol hydroxyl group shown as yellow spheres. (**E**) As (D) for the SCTR inactive state. (**F**) Cholesterol spontaneously entering the binding pocket during atMD simulation, with positions at given simulation frames as labelled (red = 0 ns; white = 250 ns, blue = 500 ns). Cholesterol is shown relative to the SCTR (final frame at 500 ns).

We used a recently released multi-modal foundation model AI, Chai-1 [27], to predict cholesterol-receptor complexes across all 15 class proteins (**Fig. S11, 12**). The Chai-1 webserver enables lipids to be inputted as SMILES strings, alongside the protein sequence. Chai-1 predicted the conserved TM4 site with most other predicted sites corresponding to those observed in the cryo-EM structures (**Fig. 3A, Supplementary Movie 1**). No cholesterol deep membrane binding modes were observed. This likely reflects the model training, which is in part based on PDB X-ray and cryo-EM structures, which may be considered to contain a 'biased sample' of tight binding lipids [48]. That is, those which bind tightly enough to survive the purification, detergents, and non-native temperatures [49] used during structure determination. The dynamic 'greasy patch' binding of cholesterol molecules we observe in MD simulations, and particularly deep membrane cholesterol, is likely to be challenging to resolve by these methods.

A comparison of the cholesterol site residence times across the class B1 GPCRs highlights a clear outlier in the case of the secretin receptor (SCTR, **Fig. 1A**). The top ranked replica of the active state contained a binding site with a residence time of 692 ns compared to a range of 50- 255 ns for other active state receptors). The longest residence time in the inactive the state was 1854 ns (compared to 72 – 170 ns for inactive states of other receptors). Whilst we cannot readily compute affinity from long binding events observed in single replicates, we assessed these sites further. When compared to the active state cryo-EM structure, we observe a shift in the position of

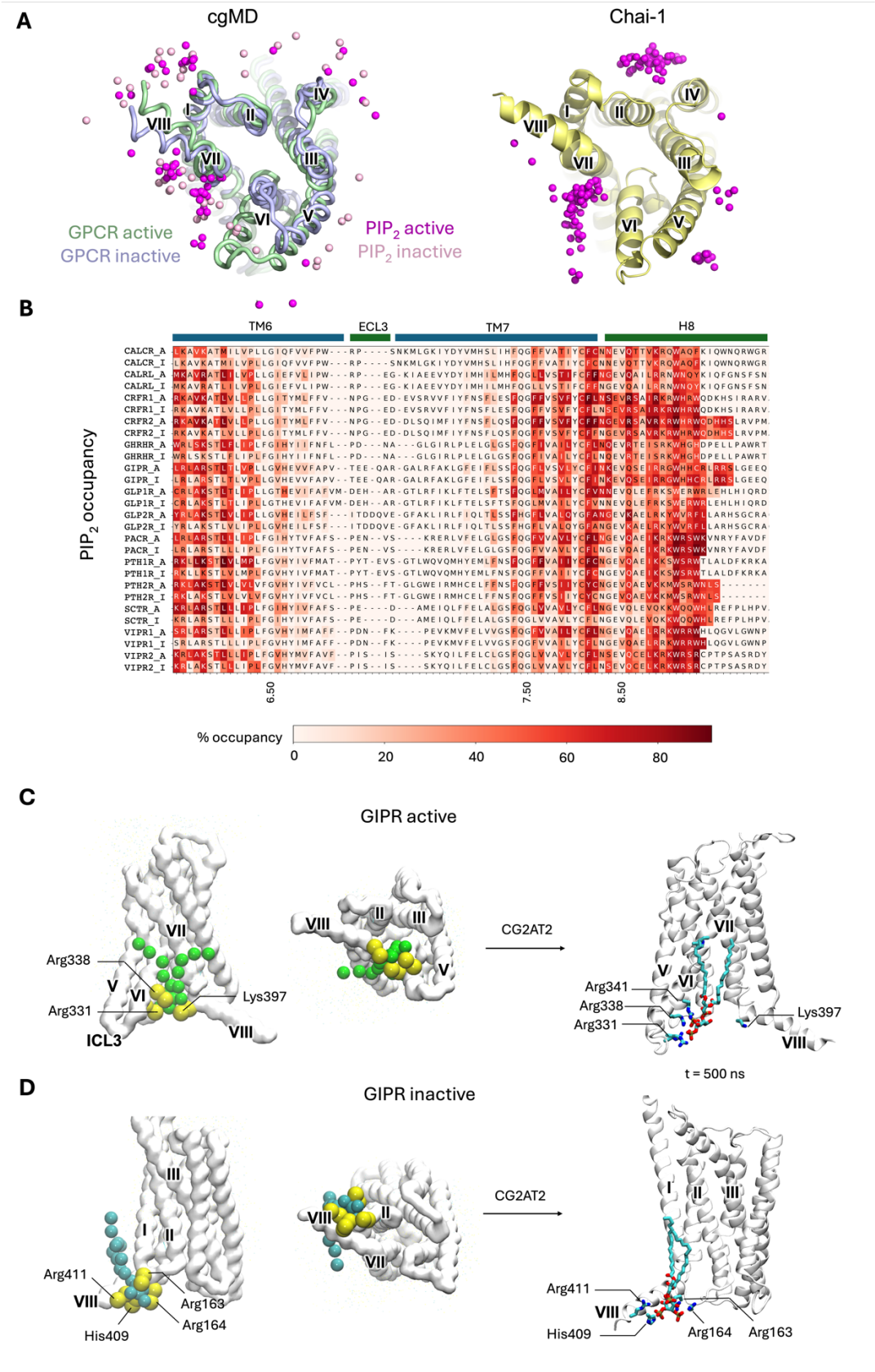
His211 in the ECL1 loop during the simulation, which forms a contact with the cholesterol hydroxyl group in the active state (**Fig. 3D**). For the inactive state, the cholesterol adopts a deep membrane binding mode and buries into the gap between TM4 and TM5 (**Fig. 3E**) The key residues responsible for mediating the contacts in both sites are shown in **Fig. 3E,D.** The Chai-1 prediction does not resolve this pose (**Fig. S13**), likely due to the model plasma membrane environment being required for the conformational change of His211.

To further evaluate the deep membrane cholesterol binding mode of SCTR we used CG2AT2[50] to convert cgMD frames corresponding to the active and inactive long residence time bound cholesterol sites (**Table S2**) as the starting point for 500 ns atomistic simulations. It is known that cholesterol accesses the membrane core more frequently in coarse-grained simulations compared to atomistic simulations[51,52]. We observe this in our simulations, where cholesterol presence in the membrane core is greatly reduced in atomistic simulations, but non-zero (**Fig. S14**). For the active state containing the long-binding time deep membrane bound cholesterol, we observed dissociation of these molecules during the equilibration and minimisation steps. For the third replicate, we observed hydrogen bonding disruption within the first 10 ns, and complete dissociation of the molecule within the first 200 ns. In this replicate, a second cholesterol molecule accessed the binding site from bulk membrane at ~420 ns and remained bound for the duration of the simulation, indicating that the site identified *via* cgMD simulations (**Fig. 3F**) may also be accessed by cholesterol at atomistic resolution. Comparative lipid occupancy heatmaps for cholesterol, PIP_2 and GM3 for SCTR at cgMD and atomistic resolution are shown in **Fig. S15**.

Conserved PIP_2 interactions at TM6/7 and H8

269 The inner leaflet lipid PIP_2 interacted with all class B1 receptors with maximum
270 residence times in the range 955 – 6455 ns (active) and 216 to 1273 ns (inactive) (**Fig.**
271 **4A, Supplementary Movie 1**). These interactions were concentrated at the intracellular
272 tips of TM6/7, and either side of H8 (**Fig. S3**). For all receptors, PIP_2 interactions at the
273 TM6/7 site showed enhanced stability in simulations of the active state compared to the
274 inactive state (**Fig. S3, 9**). This site was also predicted by Chai-1, along with a second
275 site between TM2/4 (**Fig. 4A**). Analysis of the degree of direct contact that each residue
276 formed with PIP_2 during simulations exhibited a similar trend, with increased levels of
277 per residue contact in the active state (**Fig. S3**). We analysed the top active and inactive
278 state site for GIPR which showed one of the highest residence times for PIP_2 in the
279 active state in our simulations (**Fig. 4C, D**). Particularly high levels of interaction were
280 seen between basic residues and the anionic 3' and 4' phosphoryl moieties of the PIP_2
281 headgroup. These residues are highly conserved across class B1 GPCRs (**Fig. 4B**).
282 Binding of PIP_2 lipids was also seen on the opposing side of H8, adjacent to TM1.
283 Binding at this location is similarly driven by contact with basic residues in H8 and the
284 adjacent TM1 (**Fig. 4D**).

285 We analysed PIP_2 binding of GIPR at atomic resolution by converting the final frames
286 of each cgMD simulation replicates, and performed 500 ns atMD simulations. In these,
287 the high residence time binding sites observed in cgMD were maintained throughout
288 the simulations for both the active and inactive state, with some sidechain reorientations
289 at atomistic resolution, with the Lys397 interaction reduced on this timeframe in the
290 active state (**Fig. 4C,D**). In each case the key interactions involve charged Arg/Lys
291 residues. **Fig. S15** shows occupancy heatmaps for GIPR at cgMD and atomistic
292 resolution.



293

294 **Fig. 4. Interactions of PIP_2 with class B1 GPCRs. (A)** Alignment of top ranked
295 PyLipID PIP_2 interaction sites across all 15 receptors from cgMD simulation and Chai-

296 1. The view shown is from the cytoplasmic side onto the base of the active (green) and
297 inactive (blue) states of the CTR. Spheres represent PIP_2 headgroups for the top ranked
298 interaction sites extracted from simulations of active (dark magenta) and inactive (light
299 pink) states of all 15 class B1 GPCRs. Chai-1 predicted an active state for all 15
300 receptors. For each receptor, five Chai-1 models were generated, each containing three
301 PIP_2 poses. This yielded 15 poses per receptor, and 225 poses overall. Only poses within
302 the inner leaflet were considered.(**B**). PIP_2 occupancy mapped onto sequence alignment
303 of class B1 GPCRs for regions TMH6-8. (**C**) Highest residence time PIP_2 site in the
304 GIPR active state corresponding to TM6/7, and H8. Residues in contact with the PIP_2
305 are labelled. The final position of the PIP_2 lipid following 500 ns atomistic simulation
306 is shown. (**D**) Highest residence time PIP_2 binding site in the inactive GIPR receptor,
307 with residues in contact with the lipid headgroup shown . The final position of the PIP_2
308 lipid following 500 ns atomistic simulation is shown.

309 **The glycosphingolipid GM3 modulates the mobility of receptor ECDs**

310 To characterise the interactions of ECDs with the plasma membrane, we analysed the
311 ECD mobility of each GPCR in each state relative to the TM bundle (**Fig. S16**). These
312 visualisations demonstrate that all B1 receptors in inactive and active states exhibit
313 some degree of mobility of the ECD, but some receptors sample a much larger
314 conformational space: GIPR, CRF1R and GHRHR (active states), and PAC1R, PTH1R
315 and GLP2R (both states) showed the greatest mobility in our cgMD simulations.

316 To define the role of GM3 in ECD mobility, we chose to focus on two receptors in
317 detail: GIPR, which showed the widest conformational variability in these simulations,
318 and GLP1R which is representative of the overall B1 receptors, showing moderate
319 mobility during our simulations (**Fig. 5A,E**). For the GLP1R and GIPR, we observed
320 different modes of GM3 binding. For the active state GLP1R, GM3 bound tightly
321 (maximum residence time 1525 ns) to the base of the ECD and the orthosteric ligand
322 pocket, restricting the mobility of this domain (**Fig. 5B**). For the inactive state GLP1R,
323 GM3 bound to the base of the ECD but did not interact with the orthosteric pocket, with
324 a maximum residence time of 398 ns. For GIPR, GM3 was found to bind to the distal

325 end of the ECD, forming contacts from the bulk membrane, with maximum residence
326 times of 590 ns for the active state, and 718 ns for the inactive state (**Fig. 5F**).

327 To further assess the role of GM3 in ECD modulation, we ran comparative cgMD
328 simulations with GM3 absent from the membrane. To quantify the ECD motion, we
329 calculated the ECD bend angle [12] relative to the TM bundle (**Fig. S17**) for each state,
330 with and without GM3 present in the membrane (**Fig. 5C,G**). For active state GLP1R,
331 the absence of GM3 caused a small shift in the position of the ECD relative to the TM
332 domain. For active state GIPR, a pronounced difference was observed, with the
333 presence of GM3 restricting the motion of the ECD such that it remained membrane
334 interacting (**Fig. 5F,G**). The inactive state remained highly flexible with and without
335 GM3, able to contact the membrane by forming two different states (**Fig. 5F**).

336 To evaluate the impact of using a coarse grained model (including an elastic network)
337 on the observed ECD mobility, we converted the final frames from each cgMD replicate
338 as the starting point for 500 ns atomistic simulations, with and without GM3 present in
339 the membrane (**Table S2**). A comparison of ECD conformations sampled showed that
340 these were largely determined by their initial conformations, on the 500 ns timeframe
341 (**Fig. S17**). Accordingly, comparison of GM3 contacts showed similar trends for atMD
342 and cgMD (**Fig. S18 B,C**).

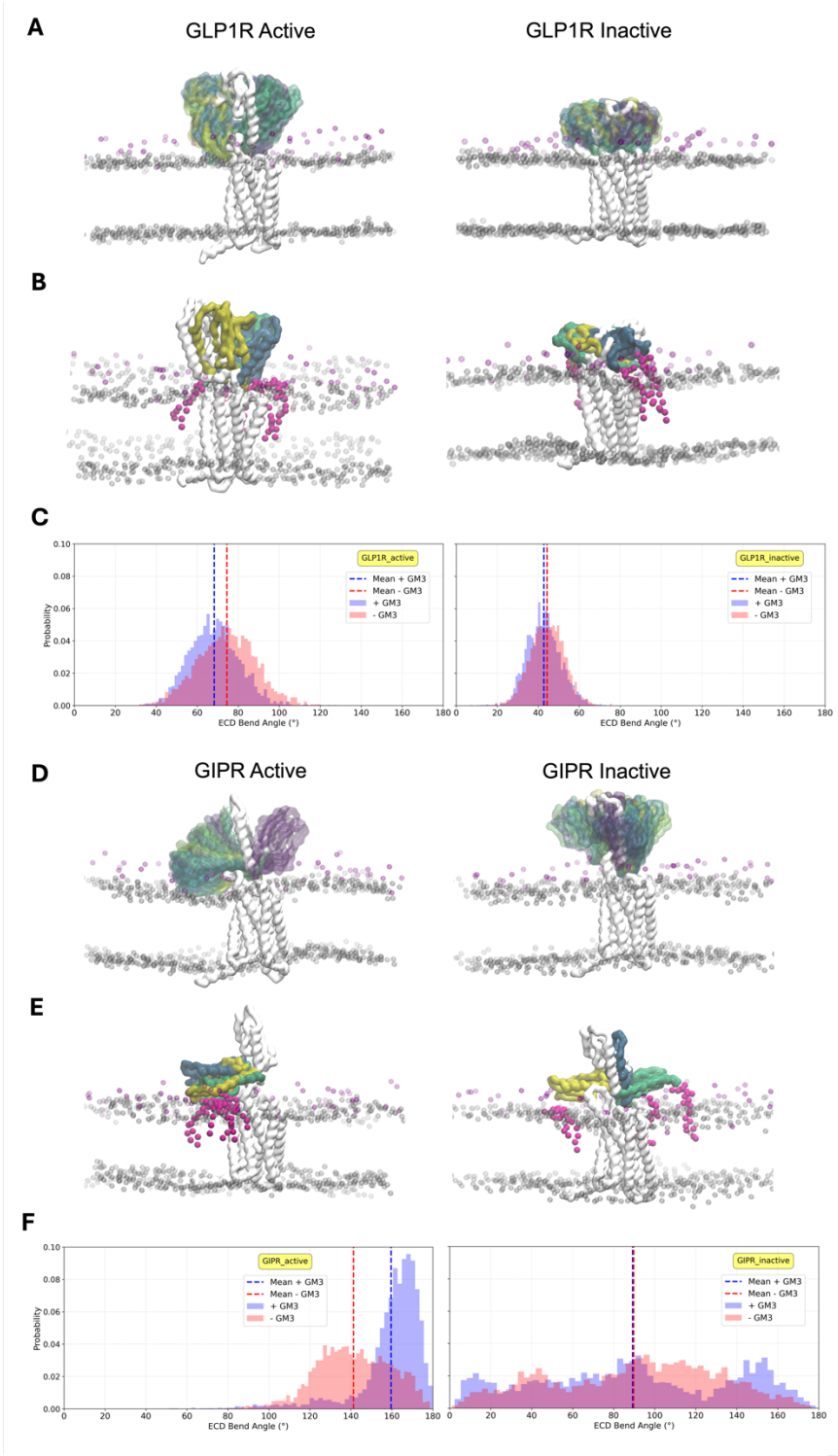


Figure 5. Modulation of GLP1R and GIPR ECDs by GM3 in cgMD simulations. (**A**) Motion of the ECD during 10 µs simulation time for the active and inactive states for GLP1R. The starting structure is shown as solid grey, whilst the ECD motion relative to the TM bundle is shown coloured by timestep (0 to 10 µs, blue-green-yellow). Phospholipid headgroup beads are shown as transparent grey spheres and GM3 GL beads are shown in magenta. (**B**) Final 10 µs conformations for each repeat. GM3

lipids within 6 Å of the ECD backbone are shown in magenta for the final frame of the first repeat. (**C**) Histogram of the angle between the ECD and the TM bundle for the simulations with GM3 present. (**D**) ECD motion of GIPR during simulation, coloured as in (A). (**E**) GIPR final conformations for each run, with all GM3 lipids contacting the ECD in the final frames shown, coloured as in (B). (**F**) Histogram of ECD angles for GIPR, as described in (C).

To experimentally validate the observed role of GM3 in modulation of the ECD of GLP1R and GIPR in our simulations, we used a time-resolved Förster resonance energy transfer (TR-FRET) assay to measure receptor ECD proximity to the plasma membrane under vehicle and stimulated conditions (**Fig. 6A**). Experiments were conducted in the presence and absence of the glucocerebroside synthase inhibitor Eliglustat to inhibit GM3 synthesis [53]. For GLP1R, stimulation with the agonist exendin-4 elicited a conformational rearrangement resulting in increased TR-FRET between the receptor ECD and the plasma membrane, an effect that we interpret as a proxy for ECD opening upon agonist binding. Additionally, a significant reduction in ECD opening was observed in the presence of Eliglustat, indicating that the ECD of GLP1R is more flexible when there is GM3 in the plasma membrane (**Fig. 6B**). Conversely, for GIPR, we observed that the ECD was less dynamic in the active (GIP-stimulated) compared to the inactive state (**Fig. 6C**), potentially due to the membrane-bound ECD open-state mode observed during the cgMD simulations. No significant difference in ECD opening between the untreated and the Eliglustat-treated conditions was observed for GIPR. Moreover, a higher basal state fluorescence indicated that the GIPR ECD is interacting with the membrane in the basal state more than GLP1R, consistent with our simulations (**Fig. 5F**). Our cgMD simulations partially agree with the TR-FRET results but we also find some discrepancies. For GLP-1R, simulations showed limited ECD dynamics in the inactive state under both 0% and 10% GM3 conditions (**Fig. S15**), which agrees with the experimental finding that there is no significant difference

377 between the control and Eliglustat conditions. In the active state, while experiments
378 suggest a significant reduction in ECD dynamics between the control and Eliglustat
379 conditions, our simulations did not show a strong GM3-dependent effect on ECD
380 dynamics. For GIPR, our cgMD simulations indicated that the ECD is most dynamic in
381 the inactive state, showing a wide range of bend angles under both 0% and 10% GM3
382 conditions (**Fig. S15**). This supports the experimental interpretation that the GIPR ECD
383 is contacting the membrane in the absence of ligand. However, we did not observe a
384 significant reduction in ECD dynamics upon GM3 depletion in our simulations, as was
385 seen experimentally. Taking the simulation and experimental results together, we
386 observe state-dependent modulation of GIPR and GLP1R by GM3, and our results
387 suggest that a fully closed inactive state, observed in the cryo-EM structure of GLP1R,
388 may not exist for GIPR.

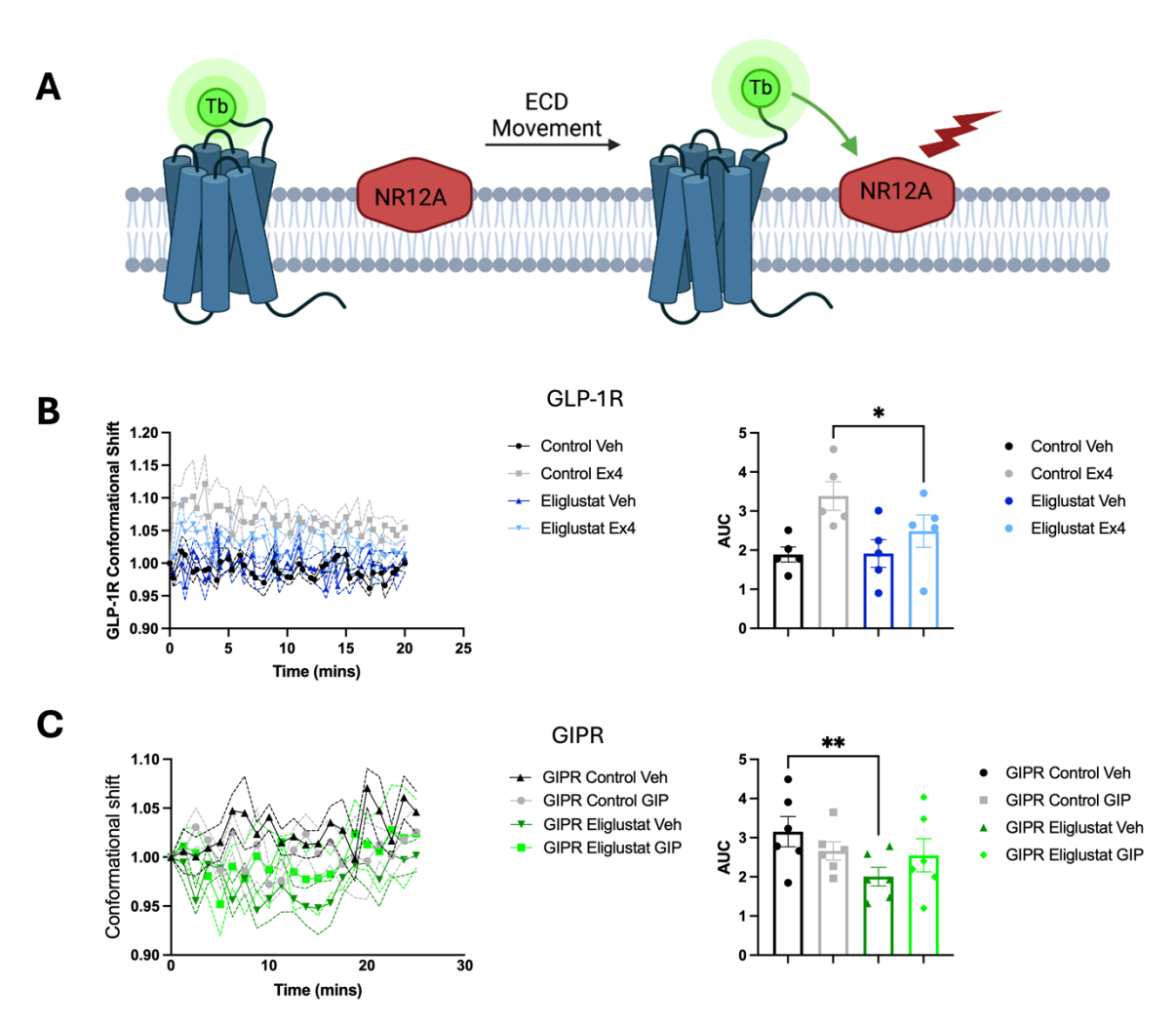


Fig. 6. TR-FRET ECD opening conformational shift assay. (**A**) Schematic diagram showing the ECD – plasma membrane conformational shift assay: SNAP-tagged receptors are expressed in pancreatic β-cells to mimic native plasma membrane lipid composition, followed by Terbium (LumiTb, Tb) labelling of the SNAP-tag and NR12A labelling of the plasma membrane prior to TR-FRET analysis. (**B**, **left**) Conformational shift response (TR-FRET ratio) over time between SNAP-Lumi4Tb-labelled GLP1R and NR12A under Control or 300 µM Eliglustat conditions following stimulation with vehicle (Veh) or 100 nM exendin-4 (Ex4); (**B**, **right**) Area under the curve (AUC) for the TR-FRET kinetic response; n = 5. (**C**, **left**) Conformational shift response (TR-FRET ratio) over time between SNAP-Lumi4Tb-labelled GIPR and NR12A under Control or 300 µM Eliglustat conditions following stimulation with vehicle (Veh) or 100 nM GIP; (**C**, **right**) Area under the curve (AUC) for the TR-FRET kinetic response; n = 5.

Rare deep membrane interactions of POPC

404 Our global analysis highlighted high-residence time binding of POPC to PAC1R in the
405 active state (**Fig. S1**). The same pattern of binding is also apparent for GHRHR in the
406 active state, albeit at lower residence times (**Fig. S1**). We analysed both binding modes,
407 which correspond to the same site at TMH3-4 (**Fig. S19A,B**), a deep membrane site
408 recently identified as occurring frequently in a large scale GPCR simulation study of
409 POPC membranes[47]. We found that in each case, in a single replicate, a POPC molecule
410 entered this pocket from the intracellular leaflet, adopting a deep membrane binding
411 mode with the POPC headgroup buried between TMH3-4 (**Fig S19A,B**). In both
412 pockets the POPC headgroup interacts with a buried Trp residue as well as interacting
413 with the surface W4.50. A sequence comparison (**Fig. S19C**) indicates that this binding
414 site may be present for all Class B1 receptors.

415

416 **Discussion**

417 This study of class B1 GPCRs has identified trends in specific lipid interactions across
418 the subfamily, highlighting avenues for further in-depth exploration. Key findings
419 include the "deep membrane" binding mode of cholesterol which has been observed in
420 multiple simulation studies at both coarse grained and atomistic resolution, but to the
421 best of our knowledge has not been captured by experimental methods. We note that a
422 recent X-ray crystallography and simulation study of aquaporin captured cholesterol in
423 a deeper position between aquaporin tetramers than bulk membrane, but the polar
424 headgroup does not access the centre of the core[54]. Our previous work on GLP1R has
425 provided evidence that mutating a cholesterol binding site impacts downstream
426 signalling and functional responses [18]. The results presented here indicate that the
427 secretin receptor may be more tightly modulated by cholesterol, with distinct high
428 residence time cholesterol binding sites in the active and inactive states potentially
429 indicating a route to modulating biased signalling of this receptor. The His211
430 identified in the long residence time active state site and the Ser231 of the inactive state
431 site are shown in the gnomAD database [55] as variants of potential interest, with
432 unknown pathogenicity.

433 Aside from the cholesterol deep binding mode, we also observe rare instances of POPC
434 lipids inserting between TMH3-4 for PAC1R and GHRHR. POPC deep membrane
435 binding has been recently described in a large scale GPCR simulation study using
436 POPC-only membranes at atomistic resolution[47]. Our simulations provide some
437 evidence that such sites may still be accessed by POPC in the presence of more complex
438 membrane compositions. The GHRHR site was identified by visual inspection and
439 comparison to the POPC occupancy plot for PAC1R, which indicates there is an

440 opportunity for machine learning pattern recognition tools to be used to extract lipid
441 binding sites from our datasets.

442 We found that, in a model plasma membrane, class B1 ECDs are highly mobile. Our
443 analysis of GM3 ECD modulation of GLP1R and GIPR by both simulation and
444 experiment demonstrate that GM3 has receptor-specific effects on ECD modulation,
445 adding to the growing body of evidence that GM3 influences GPCR behaviour.
446 Previous studies have shown that both GCGR and GHRHR are modulated by GM3[12,56].
447 Taking our simulation and experimental results together, our findings suggest that GM3
448 can modulate ECD–membrane interactions in a receptor- and state-dependent manner
449 for GLP-1R and GIPR. Our simulation data indicates that this is a general feature of the
450 class B1 GPCRs and may be a valuable pathway to explore for future therapeutic design
451 efforts.

452 We additionally identify a highly conserved PIP_2 interaction site formed by TM6/7 and
453 H8 *via* physics-based simulation. PIP_2 binding at this site shows a strong preference for
454 the active over the inactive state, with the outward movement of TM6 exposing
455 additional surface area and residues, as well as a 'groove-like' architecture into which
456 a PIP_2 lipid may 'slot'. This site has previously been reported for isolated studies of the
457 glucagon receptor (GCGR) [12,57]. The markedly high degree of conservation that we
458 observe across all B1 receptors, the strength of interaction, and the location of this site
459 adjacent to TM6, which undergoes the key movement seen during the inactive/active
460 conformational change, is particularly intriguing. This raises the prospect that PIP_2 may
461 modulate the kinetics of the transition and/or stabilise the active state, as has been seen
462 for select class A GPCRs [56,58]. Similarly, binding of PIP_2 at this site in class B1 GPCRs
463 may have a role in the recruitment of soluble intracellular signalling partners. In class
464 A GPCRs, $PI(4,5)P_2$ binding at TM4 and TM1 has been shown to stabilise the active

465 state, and enhance coupling to mini-$G\alpha_s$ over mini-$G\alpha_i$ and mini-$G\alpha_{12}$ by acting as a
466 'molecular glue' to bridge complex formation [17]. Functional experiments have also
467 reported a role of $PI(4,5)P_2$ in enhancing β-arrestin recruitment for select GPCRs
468 [11,45,59,60]. The PIP_2 site identified in the present study of class B1 GPCRs is particularly
469 conserved across the subfamily compared to PIP_2 sites in class A GPCRs [17,45], and
470 merits further investigation as to its possible functional consequences. We note that two
471 antagonists of the GCGR have been resolved bound at this same interaction site (PDB
472 ids: 5EE7 and 5XEZ) [61,62], raising the prospect of direct phosphoinositide / drug
473 competition and synergistic binding effects. Similarly, a positive allosteric modulator
474 (PAM) used to stabilise structures of GIPR in complex with the antidiabetic therapeutic
475 tirzepatide (Mounjaro) (PDB id: 7RBT) binds proximal to the TM6/7-H8 site and may
476 interact with lipids in this site [63].

477 Taken together, our comparison of cgMD simulations to the artificial intelligence-based
478 predictions shows that interactions requiring model membrane environments or subtle
479 conformational changes are not yet captured by structure prediction-based approaches,
480 highlighting that the molecular simulation approaches remain a powerful tool in studies
481 of membrane protein-lipid interactions. Tools such as aiida-gromacs, alongside efforts
482 to standardise data in the simulation field, including MDVerse [64]and the Molecular
483 Dynamics Data Bank (https://mddbr.eu), may enable training of AI models and
484 improve predictive capacity within complex bilayer environments.

485 **Limitations**

486 It is important to discuss the limitations of the present study. Our study was enabled by
487 the availability of all 15 receptor cryo-EM structures, but these were mostly in the
488 monomeric active state, whilst most inactive states had no experimental structures
489 available. There is therefore a need for further class B1 GPCR structural studies to

resolve inactive states, as well as intermediate states to enable the role of specific lipids throughout the full conformational cycle to be captured[65]. We used the Martini 3 forcefield which has been extensively employed for specific protein-lipid interactions. This has the advantage of stabilising inactive states in substrate and G protein-free conformation, but with the consequence that it limits the conformational space explored. This is particularly relevant in the observation of ECD motion, where the elastic network used to maintain protein structure will impact conformational mobility. Our simulations captured some elements of our experimental findings, but we may attribute differences to a combination or subset of: the different temporal and spatial scales; a single receptor in a membrane patch in the simulations compared to the full biological complexity in experiment; restricted sampling (including protein conformations) in Martini 3 simulations.

Sampling remains a challenge for coarse grained and atomistic simulation studies[66]. Our analyses of 10 μs and 100 μs timeframes showed comparable global properties across lipids, yet we observe instances of long residence time binding of cholesterol (to SCTR), and deep membrane POPC binding (PAC1R, GHRHR) events in individual replicates, limiting our ability to compute reliable error estimates[67]. This is particularly the case for PyLipID defined lipid site residence times, where different residues may be assigned to a given site across replicates. Our observations are intended to inform future experiments (and simulation studies), such as mutagenesis of the observed "deep membrane" cholesterol and POPC sites identified in SCTR, PAC1R and GHRHR.

There are multiple routes to extend this work, including experiments to validate the novel lipid binding sites described here, as well as confirming the family-wide modulation of the ECDs by GM3. We have made our set up and simulation files available to the community to build on this study. Future work will lead to building a

515 more complete picture of class B1 GPCR signalling, by including heterodimerisation
516 states [68,69], substrates, G protein/ β-arrestin complexes and post translational
517 modifications [70].

518 **Materials and Methods**

519 **GPCR Model Building**

520 B1 GPCRs structures in active and inactive states were taken from cryo-electron
521 microscopy (cryo-EM) structures and AlphaFold multistate models deposited in
522 GPCRdb [71]. For the inactive state, the GPCRdb models were used when there were no
523 experimental structures available (13/15 inactive states). For active state structures,
524 GPCRdb models were used for consistency when the Cα RMSD with the experimental
525 structure was < 0.8 Å. Above this cut off the ECD in GPCRdb models did not align
526 with the experimental structure, and so the ECD from the experimental structure was
527 used instead (**Table S1**). The low-confidence region in the N-terminal ECD and C-
528 terminal tail were removed to match the existing experimental structures for all
529 GPCRdb models used. Modeller v10.1 [72] was used to add missing loops and the
530 CHARMM-GUI PDB reader [30] was used to model any missing atoms in experimental
531 structures. All models used in this study used the native human sequence. Further
532 information about each model, along with the associated PDB accession codes [73] can
533 be found in **Table S1**. [747576]

534 **Coarse-grained MD Simulations**

535 GPCR structures were converted to a coarse-grained representation using Martinize2 [74]
536 and the Martini 3 force field [22]. ElNeDyn elastic network restraints were applied to each
537 GPCR using an elastic bond force constant of 500 kJ/mol/nm^2 and an upper cut-off of

538 0.9 nm [75]. The transmembrane region of GPCR was predicted using PPM 3.0 Web
539 Server [31] and embedded into an asymmetric lipid membrane bilayer in a 25 x 25 x 19
540 nm^3 box using the insane.py script [76]. The mammalian plasma membrane composition
541 previously described [77] was used as follows: POPC (25%), DOPC (25%), POPE (8%),
542 DOPE (7%), GM3 (10%) and cholesterol (25%) in the upper leaflet, and POPC (5%),
543 DOPC (5%), POPE (20%), DOPE (20%), POPS (8%), DOPS (7%), POPI(3,4)P_2 (10%)
544 and cholesterol (25%). For the simulations of GLP1R and GIPR in a 0% GM3
545 membrane, POPC and DOPC were increased to 30% each in the upper leaflet with all
546 other components the same. The most recent Martini 3 parameters for PI(3,4)P_2 [78] and
547 cholesterol [79] were used. The parameters for PI(3,4)P_2 were used as they correspond to
548 the PIP_2 version used in previous Martini 2 studies[12,45]. We also ran comparative
549 simulations using PI(4,5)P_2 for GLP1R and GIPR in both inactive and active states as
550 test cases, with limited variation observed across these simulations (Fig. S20). Thus, at
551 coarse grained resolution, we are not likely to resolve differences in binding of
552 PI(3,4)P_2 and PI(4,5)P_2. Each system was solvated using Martini water [22] and 0.15 M
553 NaCl to neutralise the system, followed by minimisation and equilibration steps using
554 protocol files from CHARMM-GUI Martini Maker [80]. Briefly, the system was
555 subjected to steepest descents energy minimisation, followed by six sequential NPT
556 equilibration steps with restraints with decreasing force constant in each subsequent
557 step from 1000 kJ/(mol nm^2) to 50 kJ/(mol nm^2) for the protein backbone and 200
558 kJ/(mol nm^2) to 10 kJ/(mol nm^2) for lipid headgroups.

559 Production run simulations were performed for 10 μs of cgMD. Three repeats were run
560 for each active and inactive state, initiated from the same initial particle configurations,
561 and random initial velocity seeds. A summary of simulations performed can be found
562 in Table 1. The v-rescale thermostat (tau 1.0 ps) [81] and the Parrinello–Rahman barostat

563 (tau 12.0 ps) [82] were used to maintain temperature (303.15 K) and pressure (1 bar) on
564 all production runs. Nonbonded interactions were treated with reaction-field
565 electrostatics with a cutoff of 1.1 nm[83]. All simulations were performed using
566 GROMACS 2022.4 (https://doi.org/10.5281/zenodo.7323393). Lipid interaction
567 profiles were calculated using PyLipID [32] with a cut-off of 0.7 nm to calculate lipid
568 residence time and occupancy profiles across all three repeats. VMD [84] and PyMOL [85]
569 was used for trajectory visualisation and analysis. 3D volumetric surfaces were
570 generated with the VMD VolMap plugin[86]. For the longer timeframe simulations of
571 SCTR, the same protocol was used except that the insane.py script was used to generate
572 three independent starting configurations, and production runs were performed for 100
573 μs each.

574 **Atomistic simulations**

575 The selected frames (shown in Table S2) from the coarse-grained simulations were
576 backmapped to atomistic resolution using the CG2AT2 package[50], in which the protein
577 and the membrane were backmapped. For GIPR, the last frame of each of the three
578 simulation repeats was selected for backmapping, as these represent the two extremes
579 of ECD motion (ECD bend angles ranging from 30° to 180°), as shown in Figure S10.
580 The same final frames were used for GLP-1R to ensure consistency. For SCTR, the
581 selected frames corresponded to the bound pose with the highest residence time,
582 identified using the PyLipID package[32].

583 The system was solvated using the CHARMM TIP3P water model[83] and 0.15 M of
584 NaCl was added to give a neutral system. The system was then subjected to steepest
585 decent energy minimisation steps, followed by six sequential NPT equilibration steps
586 with decreasing restraints on the lipid headgroups and protein backbone.

587 The CHARMM36m forcefield was used[87]. All the simulations were performed using
588 GROMACS 2022.4 (https://doi.org/10.5281/zenodo.7323393). 500 ns of production
589 simulations were performed for each system, and the temperature was maintained at
590 303.15 K using the Nose-Hoover thermostat[88,89] and the pressure at 1.0 bar using the
591 semiisotropic Parrinello-Rahman barostat[82] (τT = 1.0 ps and compressibility = 4.5 x 10^{-}
592 ^{5} bar^{-1}). The periodic boundary condition was applied and a time step of 2 fs was used.
593 The Particle Mesh Ewald (PME) method[90] was used for the long-range electrostatic
594 treatment. The LINCS algorithm was used to constrain H-bonds[91].

595 **Reproducible simulation protocols with aiida-gromacs**

596 These simulations were produced using an automated workflow with each step for
597 setting up the system captured using aiida-gromacs, allowing others to reuse or extend
598 our dataset to more GPCRs. An exemplar trajectory and starting files are available on
599 zenodo (https://doi.org/10.5281/zenodo.14359056) along with the single AiiDA
600 archive file, which contains all the steps involved to reproduce the simulation
601 methodology described above (https://doi.org/10.5281/zenodo.14728414) . Along with
602 the raw files and simulation metadata, a Jupyter Notebook tutorial
603 (https://github.com/PSDI-UK/aiida-gromacs/tree/master/notebooks, https://aiida-
604 gromacs.readthedocs.io/en/latest/tutorials/PTH2R_CGMD.html) is available,
605 showcasing the simplicity in provenance capture using aiida-gromacs in this study.

606 **AI models**

607 The AI models generated for this study utilized the Chai-1 multi-modal foundation
608 model described in [27](https://github.com/chaidiscovery/chai-lab/). Models were
609 generated using the web interface (https://lab.chaidiscovery.com/), with the multiple

610 sequence alignment (MSA) option active for enhanced accuracy. The model we utilized
611 had a training cutoff date of 2021-01-12 and will not have seen experimental GPCR-
612 cholesterol structures after this date. Primary sequences of the same class B1 constructs
613 used in MD simulations were extracted and used as input for the Chai-1 model, along
614 with the SMILES strings for cholesterol, and PI(4,5)P_2. SMILES strings were obtained
615 from PubChem [92]. The PI(4,5)P_2 variant used was the 1-stearoyl-2-arachidonoyl, which
616 is the most common tail saturation pattern for PI(4,5)P_2 in human cells [93]. Five AI
617 models were generated for each cholesterol run, with three cholesterol molecules in
618 each run. This yielded a total of 15 cholesterol poses for each receptor, and 225
619 cholesterol poses globally across all receptors. PI(4,5)P_2 predictions followed the same
620 protocol.

621 **Conformational Assay by TR-FRET**

622 INS-1 832/3 pancreatic β-cells stably expressing SNAP/FLAG-hGLP1R or
623 SNAP/FLAG-hGIPR were pre-incubated in suspension with 40 nM of SNAP-Lumi4-
624 Tb (SSNPTBC, Cisbio), a lanthanide fluorophore probe for TR-FRET, or 40 nM
625 SNAP-Lumi4Tb plus 300 μM Eliglustat, an inhibitor of GM3 gangliosides[94] for 1 hour
626 at 37°C in full culture media. Cells were pelleted and washed with HBSS and then
627 resuspended in 100 nM NR12A, a Nile red-based plasma membrane fluorescent probe
628 [95], diluted in HBSS, incubated for 5 minutes, and then seeded unto white opaque 96-
629 well half-area plates. TR-FRET was carried out as previously described [96]. Briefly, a
630 baseline reading was taken for 5 minutes using a Flexstation 3 plate reader with an
631 excitation of 335 nm, emission of 490 nm and 590 nm at 37°C. The cells were
632 stimulated with vehicle or either 100 nM Exendin-4 (GLP1R) or 100 nM GIP (GIPR)
633 and the plate was read for a further 20 minutes at 37°C. Receptor conformational shift

was calculated as the ratio of the fluorescent signal of both wavelengths (590/490 nm) after normalising to basal signal. Area under the curve (AUC) for the TR-FRET kinetic response was calculated from n = 5 individual experiments with 4 technical repeats per experiment. Data presented as mean +/- Standard Error of the Mean (SEM) and analysed using one way ANOVA with Šídák's multiple comparisons test $p^* < 0.05$. Data was collected using SoftMax Pro software on a Flexstation Plate reader (Molecular Devices) and copied into Excel (Microsoft). Data was analysed using Prism 10 (GraphPad).

Statistics and reproducibility

Each molecular dynamics simulation condition had at least three independent replicates. Per residue lipid occupancy mean values and standard deviations were calculated from three replicates. The simulations are reproducible using the same settings and input files provided. TR-FRET assays used 5 individual experiments with 4 technical replicates per experiment.

Data and materials availability:

All data are available in the main text or the supplementary materials. Additional data including raw data for plots presented has been deposited on Zenodo: https://doi.org/10.5281/zenodo.14359056; https://doi.org/10.5281/zenodo.14728414.

References

1. Cong, Z. *et al.* Structural perspective of class B1 GPCR signaling. *Trends in Pharmacological Sciences* vol. 43 Preprint at https://doi.org/10.1016/j.tips.2022.01.002 (2022).

2. Karageorgos, V. *et al.* Current understanding of the structure and function of family B GPCRs to design novel drugs. *Hormones* vol. 17 Preprint at https://doi.org/10.1007/s42000-018-0009-5 (2018).

3. Cary, B. P. *et al.* New Insights into the Structure and Function of Class B1 GPCRs. *Endocrine Reviews* vol. 44 Preprint at https://doi.org/10.1210/endrev/bnac033 (2023).

4. Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR dynamics: Structures in motion. *Chemical Reviews* vol. 117 Preprint at https://doi.org/10.1021/acs.chemrev.6b00177 (2017).

5. Hauser, A. S. *et al.* GPCR activation mechanisms across classes and macro/microscales. *Nat Struct Mol Biol* **28**, (2021).

6. Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**, 172–5 (2014).

7. Yeagle, P. L. Non-covalent binding of membrane lipids to membrane proteins. *Biochimica et Biophysica Acta - Biomembranes* vol. 1838 Preprint at https://doi.org/10.1016/j.bbamem.2013.11.009 (2014).

8. Byrne, E. F. X. *et al.* Structural basis of Smoothened regulation by its extracellular domains. *Nature* **535**, (2016).

9. Oates, J. & Watts, A. Uncovering the intimate relationship between lipids, cholesterol and GPCR activation. *Current Opinion in Structural Biology* vol. 21 Preprint at https://doi.org/10.1016/j.sbi.2011.09.007 (2011).

10. Kiriakidi, S. *et al.* Effects of cholesterol on GPCR function: Insights from computational and experimental studies. in *Advances in Experimental Medicine and Biology* vol. 1135 (2019).

11. Janetzko, J. *et al.* Membrane phosphoinositides regulate GPCR-β-arrestin complex assembly and dynamics. *Cell* **185**, (2022).

12. Ansell, T. B., Song, W. & Sansom, M. S. P. The Glycosphingolipid GM3 Modulates Conformational Dynamics of the Glucagon Receptor. *Biophys J* **119**, (2020).

13. Hedger, G. & Yen, H.-Y. The influence of phosphoinositide lipids in the molecular biology of membrane proteins: recent insights from simulations. *J Mol Biol* 168937 (2025) doi:https://doi.org/10.1016/j.jmb.2025.168937.

14. 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* **12**, (2016).

15. Guixà-González, R. *et al.* Membrane omega-3 fatty acids modulate the oligomerisation kinetics of adenosine A2A and dopamine D2 receptors. *Sci Rep* **6**, (2016).

16. Dawaliby, R. *et al.* Allosteric regulation of G protein-coupled receptor activity by phospholipids. *Nat Chem Biol* **12**, (2016).

17. Yen, H. Y. *et al.* PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **559**, (2018).

18. Oqua, A. I. *et al.* Molecular mapping and functional validation of GLP-1R cholesterol binding sites in pancreatic beta cells. (2024) doi:10.7554/elife.101011.1.

19. Buenaventura, T. *et al.* Agonist-induced membrane nanodomain clustering drives GLP-1 receptor responses in pancreatic beta cells. *PLoS Biol* **17**, (2019).

20. Cheng, Y. Membrane protein structural biology in the era of single particle cryo-EM. *Current Opinion in Structural Biology* vol. 52 Preprint at https://doi.org/10.1016/j.sbi.2018.08.008 (2018).

21. Zhang, M. *et al.* G protein-coupled receptors (GPCRs): advances in structures, mechanisms and drug discovery. *Signal Transduct Target Ther* **9**, 88 (2024).

22. Souza, P. C. T. *et al.* Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat Methods* **18**, (2021).

23. Han, X. & Gross, R. W. The foundations and development of lipidomics. *Journal of Lipid Research* vol. 63 Preprint at https://doi.org/10.1016/j.jlr.2021.100164 (2022).

24. Páll, S. *et al.* Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. *Journal of Chemical Physics* **153**, (2020).

25. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).

26. Heo, L. & Feig, M. Multi-state modeling of G-protein coupled receptors at experimental accuracy. *Proteins: Structure, Function and Bioinformatics* **90**, (2022).

27. Discovery, C. *et al.* Chai-1: Decoding the molecular interactions of life. *bioRxiv* (2024) doi:10.1101/2024.10.10.615955.

28. Bereau, T., Walter, L. J. & Rudzinski, J. F. Martignac: Computational workflows for reproducible, traceable, and composable coarse-grained Martini simulations. Preprint at https://arxiv.org/abs/2409.15478 (2024).

29. Huber, S. P. *et al.* AiiDA 1.0, a scalable computational infrastructure for automated reproducible workflows and data provenance. *Sci Data* **7**, (2020).

30. Park, S. J., Kern, N., Brown, T., Lee, J. & Im, W. CHARMM-GUI PDB Manipulator: Various PDB Structural Modifications for Biomolecular Modeling and Simulation. *J Mol Biol* **435**, (2023).

31. Lomize, A. L., Todd, S. C. & Pogozheva, I. D. Spatial arrangement of proteins in planar and curved membranes by PPM 3.0. *Protein Science* **31**, (2022).

32. Song, W. *et al.* PyLipID: A Python Package for Analysis of Protein-Lipid Interactions from Molecular Dynamics Simulations. *J Chem Theory Comput* **18**, (2022).

33. Pándy-Szekeres, G. *et al.* GPCRdb in 2023: state-specific structure models using AlphaFold2 and new ligand resources. *Nucleic Acids Res* **51**, (2023).

34. Wang, X. *et al.* Molecular insights into differentiated ligand recognition of the human parathyroid hormone receptor 2. *Proc Natl Acad Sci U S A* **118**, (2021).

35. Duan, J. *et al.* Cryo-EM structure of an activated VIP1 receptor-G protein complex revealed by a NanoBiT tethering strategy. *Nat Commun* **11**, (2020).

36. Cao, J. *et al.* A structural basis for amylin receptor phenotype. *Science (1979)* **375**, (2022).

37. Zhou, F. *et al.* Structural basis for activation of the growth hormone-releasing hormone receptor. *Nat Commun* **11**, (2020).

38. Zhao, F. *et al.* Structural insights into hormone recognition by the human glucosedependent insulinotropic polypeptide receptor. *Elife* **10**, (2021).

39. Zhao, L. H. *et al.* Structure and dynamics of the active human parathyroid hormone receptor-1. *Science (1979)* **364**, (2019).

40. Ma, S. *et al.* Molecular Basis for Hormone Recognition and Activation of Corticotropin-Releasing Factor Receptors. *Mol Cell* **77**, (2020).

41. Wootten, D., Simms, J., Miller, L. J., Christopoulos, A. & Sexton, P. M. Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations. *Proc Natl Acad Sci U S A* **110**, (2013).

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

43. Guixà-González, R. *et al.* Membrane cholesterol access into a G-protein-coupled receptor. *Nat Commun* **8**, (2017).

44. Morra, G., Razavi, A. M., Menon, A. K. & Khelashvili, G. Cholesterol occupies the lipid translocation pathway to block phospholipid scrambling by a G protein-coupled receptor. *Structure* **30**, (2022).

45. Sejdiu, B. I. & Tieleman, D. P. Lipid-Protein Interactions Are a Unique Property and Defining Feature of G Protein-Coupled Receptors. *Biophys J* **118**, (2020).

46. Gater, D. L. *et al.* Two classes of cholesterol binding sites for the β2AR revealed by thermostability and NMR. *Biophys J* **107**, (2014).

47. Aranda-García, D. *et al.* Large scale investigation of GPCR molecular dynamics data uncovers allosteric sites and lateral gateways. *Nat Commun* **16**, 2020 (2025).

48. Hedger, G. & Sansom, M. S. P. Lipid interaction sites on channels, transporters and receptors: Recent insights from molecular dynamics simulations. *Biochim Biophys Acta Biomembr* **1858**, (2016).

49. Hu, J. *et al.* Physiological temperature drives TRPM4 ligand recognition and gating. *Nature* **630**, 509–515 (2024).

50. Vickery, O. N. & Stansfeld, P. J. CG2AT2: an Enhanced Fragment-Based Approach for Serial Multi-scale Molecular Dynamics Simulations. *J Chem Theory Comput* **17**, 6472–6482 (2021).

51. Jo, S., Rui, H., Lim, J. B., Klauda, J. B. & Im, W. Cholesterol Flip-Flop: Insights from Free Energy Simulation Studies. *J Phys Chem B* **114**, 13342–13348 (2010).

52. Pedersen, K. B. *et al.* The Martini 3 Lipidome: Expanded and Refined Parameters Improve Lipid Phase Behavior. *ACS Cent Sci* (2025) doi:10.1021/acscentsci.5c00755.

53. Cox, T. M. Eliglustat tartrate, an orally active glucocerebroside synthase inhibitor for the potential treatment of Gaucher disease and other lysosomal storage diseases. *Current Opinion in Investigational Drugs* vol. 11 Preprint at (2010).

54. Chiu, P.-L., Orjuela, J. D., de Groot, B. L., Aponte Santamaría, C. & Walz, T. Structure and dynamics of cholesterol-mediated aquaporin-0 arrays and implications for lipid rafts. *Elife* **12**, RP90851 (2024).

55. Chen, S. *et al.* A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).

56. Damian, M. *et al.* Allosteric modulation of ghrelin receptor signaling by lipids. *Nat Commun* **12**, (2021).

57. Kjølbye, L. R. *et al.* Lipid Modulation of a Class B GPCR: Elucidating the Modulatory Role of PI(4,5)P2Lipids. *J Chem Inf Model* **62**, (2022).

58. Thakur, N. *et al.* Anionic phospholipids control mechanisms of GPCR-G protein recognition. *Nat Commun* **14**, (2023).

59. Huang, W. *et al.* Structure of the neurotensin receptor 1 in complex with β-arrestin 1. *Nature* **579**, (2020).

60. Gomes, A. *et al.* LIPIDS MODULATE THE DYNAMICS OF GPCR:β-ARRESTIN INTERACTION. *bioRxiv* (2024) doi:10.1101/2024.03.16.585329.

61. Zhang, H. *et al.* Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature* **546**, (2017).

62. Jazayeri, A. *et al.* Extra-helical binding site of a glucagon receptor antagonist. *Nature* **533**, (2016).

63. Sun, B. *et al.* Structural determinants of dual incretin receptor agonism by tirzepatide. *Proc Natl Acad Sci U S A* **119**, (2022).

64. Tiemann, J. K. *et al.* MDverse, shedding light on the dark matter of molecular dynamics simulations. *Elife* **12**, (2024).

65. Cong, Z. *et al.* Molecular features of the ligand-free GLP-1R, GCGR and GIPR in complex with Gs proteins. *Cell Discov* **10**, (2024).

66. Knapp, B., Ospina, L. & Deane, C. M. Avoiding False Positive Conclusions in Molecular Simulation: The Importance of Replicas. *J Chem Theory Comput* **14**, (2018).

67. Sexton, R., Fazel, M., Schweiger, M., Pressé, S. & Beckstein, O. Bayesian Nonparametric Analysis of Residence Times for Protein-Lipid Interactions in Molecular Dynamics Simulations. *J Chem Theory Comput* **21**, 4203–4220 (2025).

68. Hay, D. L. & Pioszak, A. A. Receptor Activity-Modifying Proteins (RAMPs): New Insights and Roles. *Annu Rev Pharmacol Toxicol* **56**, (2016).

69. Chetverikov, N., Janoušková-Randáková, A., Nelic, D. & Jakubík, J. Cholesterol differentially modulates the activity of opioid and muscarinic receptors via common non-canonical binding site. *bioRxiv* 2024.10.09.617372 (2025) doi:10.1101/2024.10.09.617372.

70. Naglekar, A., Chattopadhyay, A. & Sengupta, D. Palmitoylation of the Glucagon-like Peptide-1 Receptor Modulates Cholesterol Interactions at the Receptor–Lipid Microenvironment. *J Phys Chem B* **127**, 11000–11010 (2023).

71. Pándy-Szekeres, G. *et al.* GPCRdb in 2023: state-specific structure models using AlphaFold2 and new ligand resources. *Nucleic Acids Res* **51**, (2023).

72. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779–815 (1993).

73. M Berman, H. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242 (2000).

74. Kroon, P. C. *et al.* Martinize2 and Vermouth: Unified Framework for Topology Generation. *Elife* **12**, (2023).

75. Periole, X., Cavalli, M., Marrink, S. J. & Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: Structure, dynamics, and intermolecular recognition. *J Chem Theory Comput* **5**, (2009).

76. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations. *J Chem Theory Comput* **11**, (2015).

77. Song, W., Yen, H. Y., Robinson, C. V. & Sansom, M. S. P. State-dependent Lipid Interactions with the A2a Receptor Revealed by MD Simulations Using In Vivo-Mimetic Membranes. *Structure* **27**, (2019).

78. Borges-Araújo, L., Souza, P. C. T., Fernandes, F. & Melo, M. N. Improved Parameterization of Phosphatidylinositide Lipid Headgroups for the Martini 3 Coarse-Grain Force Field. *J Chem Theory Comput* **18**, (2022).

79. Borges-Araújo, L. *et al.* Martini 3 Coarse-Grained Force Field for Cholesterol. *J Chem Theory Comput* **19**, (2023).

80. Qi, Y. *et al.* CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J Chem Theory Comput* **11**, (2015).

81. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J Chem Phys* **126**, 014101 (2007).

82. Parrinello, M. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* **52**, 7182 (1981).

83. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, (1983).

84. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33–8, 27–8 (1996).

85. Schrödinger LLC. *The PyMOL Molecular Graphics System, Version~1.8*. (2015).

86. Cohen, J., Arkhipov, A., Braun, R. & Schulten, K. Imaging the migration pathways for O2, CO, NO, and Xe inside myoglobin. *Biophys J* **91**, (2006).

87. Huang, J. & Mackerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J Comput Chem* **34**, (2013).

88. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys* **52**, (1984).

89. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A (Coll Park)* **31**, (1985).

90. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089 (1993).

91. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* **18**, 1463–1472 (1997).

92. Kim, S. *et al.* PubChem 2023 update. *Nucleic Acids Res* **51**, (2023).

93. Barneda, D., Cosulich, S., Stephens, L. & Hawkins, P. How is the acyl chain composition of phosphoinositides created and does it matter? *Biochem Soc Trans* **47**, (2019).

94. Shayman, J. A., Hinkovska-Galcheva, V. & Shu, L. Inhibitors of Glucosylceramide Synthase. in *Methods in Molecular Biology* vol. 2613 (2023).

95. Klymchenko, A. S. Fluorescent Probes for Lipid Membranes: From the Cell Surface to Organelles. *Acc Chem Res* **56**, (2023).

96. Lucey, M. *et al.* Acylation of the incretin peptide exendin-4 directly impacts glucagon-like peptide-1 receptor signaling and traffickings. *Mol Pharmacol* **100**, (2021).

Acknowledgments

We thank Joseph Barritt for helpful discussion.

Funding:

UKRI Future Leaders Fellowship (MR/Y01975X/1): SLR
MRC Project Grant (MR/X021467/1): AT (YM)
Wellcome Trust Discovery Award (301619/Z/23/Z): AT (AIO), SLR (GH)

Author contributions:

Conceptualization: KWC, GH, SLR, AT, JGR, JK
Methodology: KWC, LW, JK, AT
Investigation: KWC, LW, GH, SLR, YM, AIO
Visualization: KWC, LW, GH, SLR, AT, AIO, YM
Supervision: GH, AT, JGR, SLR
Writing—original draft: KWC, GH, JK, JGR, AT, SLR
Writing—review & editing: LW, AIO, JK, JGR, GH, AT, SLR

Competing interests:

Authors declare that they have no competing interests.

Supplementary Materials

Supplementary Materials Fig. S1-S20, Table S1,2, Supplementary Movie 1.