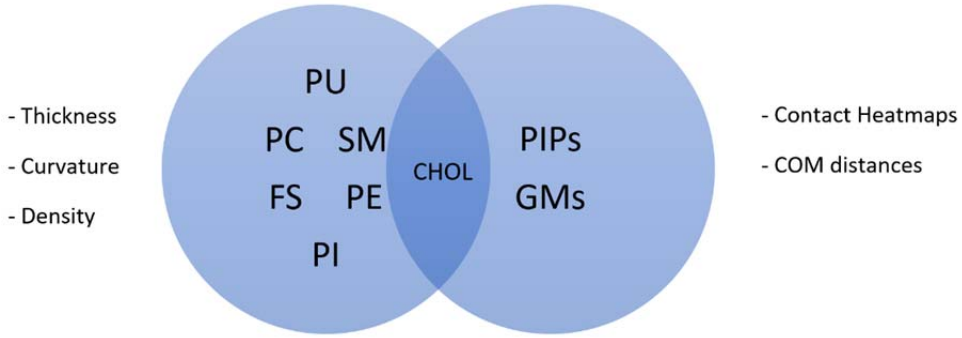
Biophysical Journal, Volume *118*

Supplemental Information

Lipid-Protein Interactions Are a Unique Property and Defining Feature of G Protein-Coupled Receptors

Besian I. Sejdiu and D. Peter Tieleman

Analysis. Considering the complexity of the setup, to comprehensively evaluate our data we analysed lipid-protein interactions by both grouping lipids based on some property as well as analysing lipid types individually. For the former, we use a combination of 2D map profiles and for the latter we calculate average contact heatmaps and distances between lipids and proteins.



We group lipids based on their headgroup type and tail saturation level. With the former grouping we define PC, PE, PS, PA, DAG, LPC, SM, CER, PI, PIPs, GM lipids. The latter allows us to croup lipids into fully-saturated (FS), poly-unsaturated (PU), cholesterol, and Other (containing lipids that are excluded from the previous three categories). PU lipids are lipids that contain above two type "D" beads and consists of: DAPC, DUPE, DAPE, DAPS, DUPS, APC, and UPC lipids. FS lipids includes SM lipids, glycolipids, ceramides, and LPC lipids.

Thickness, Curvature, and Lipid Composition. We use the same in-house tools as in our previous work(1) to calculate the thickness, curvature and lipid density profiles. The only difference is that since we are dealing with only GPCRs, to allow for direct comparison of results, we orient the structures so that H8 is facing downwards, and helices TM1-TM7 moving counter-clockwise from right to left.

Depletion-enrichment index and Equilibration Tests. For a multicomponent bilayer system with an assumed homogeneous distribution, the following relation is true:

$$[A]_{local} = [A]_{bulk}$$

where $[A]$ is the molar concentration of lipid A. Its *local* value is calculated within a cutoff distance from embedded proteins (GPCRs). There is a limit to how small this cutoff value can be, and we limit ourselves at 5Å from the protein. Rearranging the above relation, we obtain the following equation:

$$DE_{(A|GPCR)} = \frac{[A]_{local}}{[A]_{bulk}} = 1$$

That is, for a membrane system with a homogeneous lipid distribution, the expected *depletion-enrichment index* (or simply the enrichment value),(2) $DE_{(A|GPCR)}$, is unity.

With the DE index we aim to reduce what is a 3 – dimensional problem into a single number for the whole protein. As such, while it does not retain any information about the spatial distribution of the interactions, it does give a measure of the tendency of lipids at a specific point (in our case embedded proteins) to deviate from a reference distribution (a homogeneous bulk membrane distribution). From experience we have seen that it is more accurate for lipids that are either in small number in the system (e.g. PIP lipids) or change significantly during the simulation (e.g. GM lipids).

We use the average number of lipids around GPCRs to measure the direction of change in lipid distribution and estimate when the distributions have converged. We focus mainly on the 7Å radius, but the results hold for other similar values. In Figure S1 we highlight these results for $5HT_{1B}$, even though, again, the data tell a similar story for other GPCRs.

Number and Duration of Contacts. To analyse specific interactions of lipids with proteins, we calculate both the number of contacts between the corresponding lipids and each protein residue, as well as the duration of each contact. In our analysis, we use the total number of contacts (referred to simply as number of contacts) and the average duration of the longest contact for each residue. Our results, however, are not dependent on any particular analysis method or averaging statistic used. They are also independent from our cutoff choice (7Å). We use the *gmx select* utility from the GROMACS simulation package and the MDTraj package to process the trajectories for contacts and post-process the results using in-house scripts.

When calculating the contribution of TM helices in GPCR-PIP lipid interactions (as shown in Figure 4C), we have to only consider the part of the helix that is facing the intracellular membrane (since that is where PIP lipids are found, exclusively). Since it is not clear how to separate TM helices into extra- and intracellular facing residues, to ensure that we use a consistent definition between different GPCRs, we only consider the four residues that interact with PIP lipids the most. Considering all residues that form the helix, or even residues that interact at least once with PIP lipids leads to unnaturally large error bars.

For brevity and clarity, we describe only a selection of GPCRs in detail in the main text. However, we append the full analysis results for all GPCRs in the supplementary material. We provide sequence heatmaps for both the number of contacts and their duration, for GPCR-cholesterol and PIP lipid interactions in two formats: aligned and full unaligned sequence. We used GPCRdb to obtain the alignment.(3) We use the following definition for CRAC motifs: (L/V)-X(1,5)-(Y)-X(1,5)-(R/K) and its reverse for CARC motifs,(4) and use it to query our GPCR sequences using ScanProsite.(5) In these

motifs Y represents an aromatic residue (tyrosine or phenylalanine). Statistics on GPCR crystal structures solved were obtained from the GPCR-EXP database.(6)

Statistical Analysis. The expected value of the DE index assuming no depletion/enrichment is 1 – which corresponds to our null hypothesis, and we use a two-sided one-sample T-test to calculate the two-tailed p-values reported for all GPCR DE index values (reported in Table S3). The p-value of each DE index is calculated in comparison to the expected value from the null hypothesis.

The confidence intervals reported in Table S4 are calculated using one-sample t confidence intervals. If $\bar{x}$ and s are the sample mean and sample standard deviation, then the confidence interval from a random sample derived from a normal population with mean μ is:

$$\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

where $t_{\alpha/2}$ is the t critical value with $n-1$ degrees of freedom. For instance, if *pip_data* is a python list containing the DE indices of PIP lipids for each class A GPCR then the confidence interval can be calculated using the following code block:

```
import numpy as np
import scipy.stats as st

x_hat = np.mean(pip_data)
s = np.std(pip_data, ddof=1)
n = np.sqrt(len(pip_data))

t_critical_value = st.t.ppf(1-0.025, len(pip_data)-1)
t_conf_int = t_critical_value * s / n
print (x_hat-t_conf_int, x_hat+t_conf_int)
```

SciPy also provides a simpler way of doing this:

```
st.t.interval(.95, len(data)-1, loc=np.mean(data), scale=st.sem(data))
```

Lastly, we use a two-sided T-test when comparing the DE index means of different categories of GPCRs (class A vs non class A, and aminergic vs non aminergic class A GPCRs), and deriving the corresponding p-values (as reported in Table S5).

An important assumption when using t distributions is that the sample data are normally distributed. We use percent-percent (P-P) probability plots of the theoretical vs practical percentiles for each lipid type to

show that, indeed, DE indices conform to a normal distribution (Figure S3). In fact, cholesterol is the only lipid that shows a small deviation from the assumed normal distribution.

Notation. We use the TM label to denote transmembrane helices, with the exception of helix 8 which is referred to simply as H8. The exception is when we refer to H8 as part of a binding interface (TM1/8 or TM7/8). Transmembrane proteins are in contact with both leaflets of the plasma membrane and lipid interactions may occur at each side. To differentiate between binding that occurs on the extracellular and intracellular side of GPCRs we use the ec and ic notation, respectively, to label interaction sites.

Software packages. For the analysis of lipid – protein contacts we simultaneously used two different software tools: *gmx select* (part of the GROMACS simulation package) and the *compute_neighbors* method (part of the MDTraj package).(7) We use the *stats* module of *SciPy*(8) and the *statsmodels*(9) python package for the statistical analysis of our data. Lastly, plotting of the data is done using the *matplotlib*(10) and *seaborn*(11) python libraries.

GPCR-lipid interactions website. Along with the paper, we also release a dedicated webpage hosted by GitHub to interactively visualize our dataset as it pertains to cholesterol and PIP lipid interactions. It uses the NGL Viewer(12) to display the density of cholesterol and PIP lipids in a 6 nm (60Å) radius around proteins. It allows direct interactions with our dataset using a user-friendly interface and visualize the number and duration of contact maps highlighted here, as well as view 3D density profiles and 2D slices of it.

Additionally, we provide a separate application to view and interact with 3D objects representing the thickness and curvature profiles of each GPCRs. The 2D maps of these calculations are shown in Figures S7-S9, however, the online application allows the user to interactively view these profiles, customize their appearance and easily switch between mean and Gaussian curvature. These objects are generated by *g_surf* and visualized using the Three.js JavaScript library and the coloring of the curvature is done using MeshLab.(13)

We hope that this will allow for a much cleaner presentation of GPCR-lipid interactions and enable users to explore any detail of our dataset that, for reasons of text brevity and clarity, we may not have been able to do here. The webpage can be accessed through: **https://bisejdiu.github.io/GPCR-lipid-interactions** and all the code is available on the following GitHub repository: **https://github.com/bisejdiu/GPCR-lipid-interactions**

Visualization. All molecular visualizations presented in the paper are done using VMD.(14)

Protein name	Abbrev.	PDB ID	Time (μs)	Reference
5-hydroxytryptamine receptor 1B	$5HT_{1B}$	4IAQ	30	(15)
Adenosine A2A receptor	A_{2A}Ra	2YDV	30	(16)
Adenosine A2A receptor	A_{2A}Ri	3EML	30	(17)
Apelin receptor	ApelinR	5VBL	30	(18)
Angiotensin II type 2 receptor	AT_2R	5UNG	30	(19)
beta2 adrenergic receptor	b2ARa (β_2ARa)	3SN6	30	(20)
beta2 adrenergic receptor	b2ARi (β_2ARa)	2RH1	30	(21)
Cannabinoid Receptor CB1	CB_1R	5TGZ	30	(22)
Chemokine receptor CXCR1	CXCR1	2LNL	30	(23)
Dopamine D3 receptor	D_3R	3PBL	30	(24)
Endothelin ETB receptor	ETbR (ET_BR)	5X93	30	(25)
Human histamine H1	H_1R	3RZE	30	(26)
Lysophosphatidic Acid Receptor 1	LPAR1	4Z36	30	(27)
M2 muscarinic acetylcholine receptor	M_2Ra	4MQS	30	(28)
M2 muscarinic acetylcholine receptor	M_2Ri	3UON	30	(29)
mu-opioid receptor (active)	mORa (μORa)	5C1M	30	(30)
mu-opioid receptor	mORi (μORi)	4DKL	30	(31)
Human Orexin 2 Receptor	OX_2R	5WQC	30	(32)
Protease-activated receptor 2	PAR2	5NDD	30	(33)
Metarhodopsin II	RhodRa	3PQR	30	(34)
Rhodopsin	RhodRi	1GZM	30	(35)
Lysophospholipid sphingosine 1-phosphate	S1PR1	3V2Y	30	(36)
US28	US28	4XT1	30	(37)
Calcitonin receptor	CalcitoninR	5UZ7	30	(38)
Glucagon-like peptide-1 receptor	GLP1	5VEW	30	(39)
Glucagon receptor	GlucagonR	4L6R	30	(40)
Metabotropic glutamate receptor 5	mGlu5	4OO9	30	(41)
Smoothened receptor	SMO	4N4W	30	(42)

Table S1. Overview of all GPCR structures simulated. *The full name of each GPCR along with their abbreviation used in this work, as well as the respective PDB ID, simulation time and reference to the relevant paper. We believe most of our abbreviations should be obvious as to what receptor they are referring to. We also note that we differentiate between GPCRs that have been simulated in both active and inactive states by the last letter of the abbreviation (e.g.* β_2AR*i vs* β_2AR*a denoting the inactive and active state* β_2AR*, respectively).*

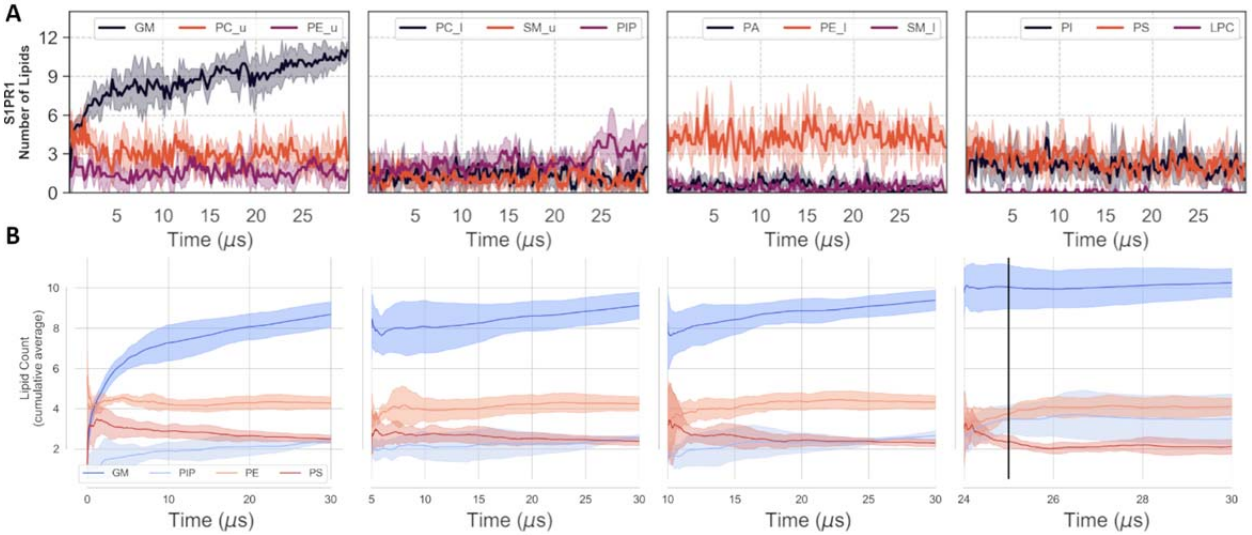


Figure S1. Convergence of the number of lipids (Lipid Count) during the course of the simulation. A. *Running average of the average number of lipids within 7Å of S1PR1 as a function of trajectory time. We chose S1PR1 to highlight here as its GM lipid count is the worst behaving among all GPCRs simulated and showcase how our focus on the last 5 μs ensures that we analyze converged lipid distributions around proteins. Lipid groups that appear in both leaflets are denoted as such (e.g. PC_u, PC_l denotes PC lipids only in the upper and lower leaflet of the bilayer, respectively)* ***B.*** *The cumulative average of several lipid groups compared to the usage of different cutoffs (that is, different points in time at which the trajectory is discarded and only the remaining part is used for analysis): no cutoff, 5 μs, 10 μs, and 24 μs. These results give us an idea of the "error" that is introduced into the calculations depending on what part, if any, of the trajectory is discarded before analysis. They also reveal the amount of simulation required to ensure converged lipid distributions. As is clear from the figure, discarding the first 5 μs of the simulation is enough for PIP, PE and PS lipid to converge, but for GM lipids even 10 μs is not enough. Simulation lengths of 20 μs and more are required to achieve converged GM lipid distributions. Please also note that, since our results are averages of 4 proteins (i.e. n is small), the initial values of the*

cumulative average are more chaotic and take some time to stabilize (hence why we're showing the values for a cutoff of 24 μs, even though in our analysis we discarded the first 25 μs, to allow for 1 μs time for the cumulative average values to stabilize). This figure is for S1PR1, and the complete dataset for all GPCRs is available on Figure S22. Highlighted areas represent ± SD (n = 4).

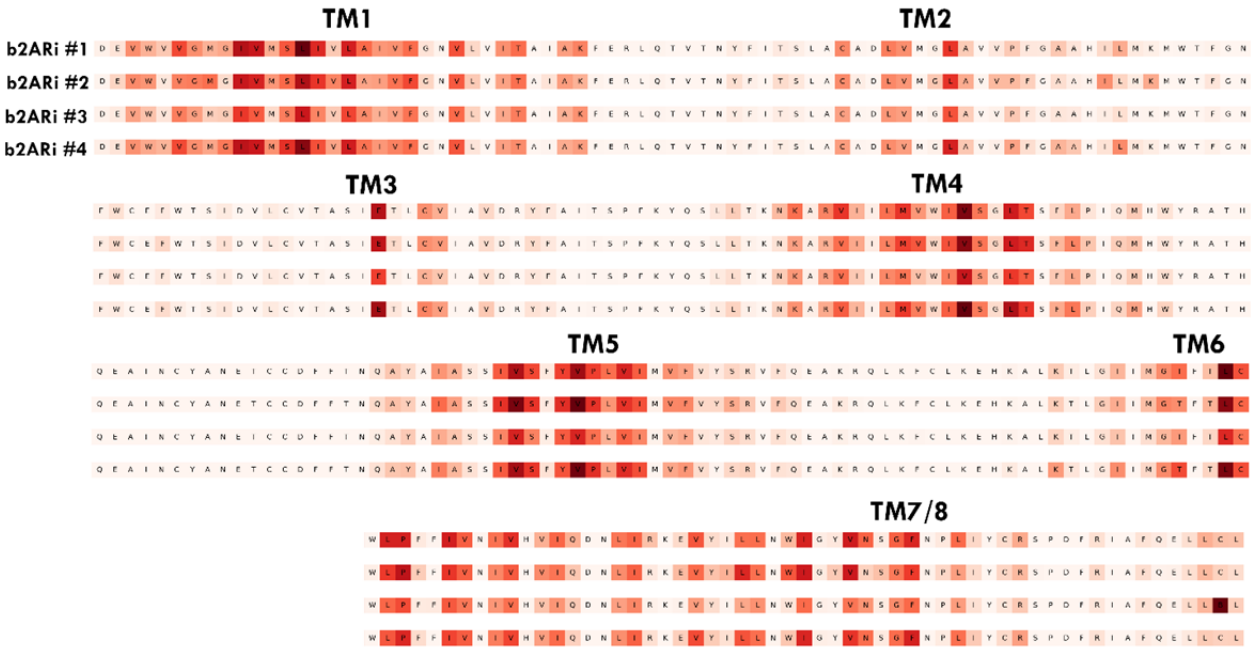


Figure S2. β_2ARi-cholesterol interactions. *Sequence heatmaps for the number of contacts with cholesterol for each β_2ARi setup. For clarity and easy comparison, we aligned the structures and are only showing residues that make up the helical core of the receptor. The systems are as follows: the setup used in our simulations (#1), including ICL2 (#2), including palmitoylated Cys-341 (#3), and pre-equilibrated system (#4). We see that our results are not affected by our simulation protocol. We also analysed the effect different strengths of the elastic network have and saw no difference. These results also show that the cholesterol distribution around proteins has converged.*

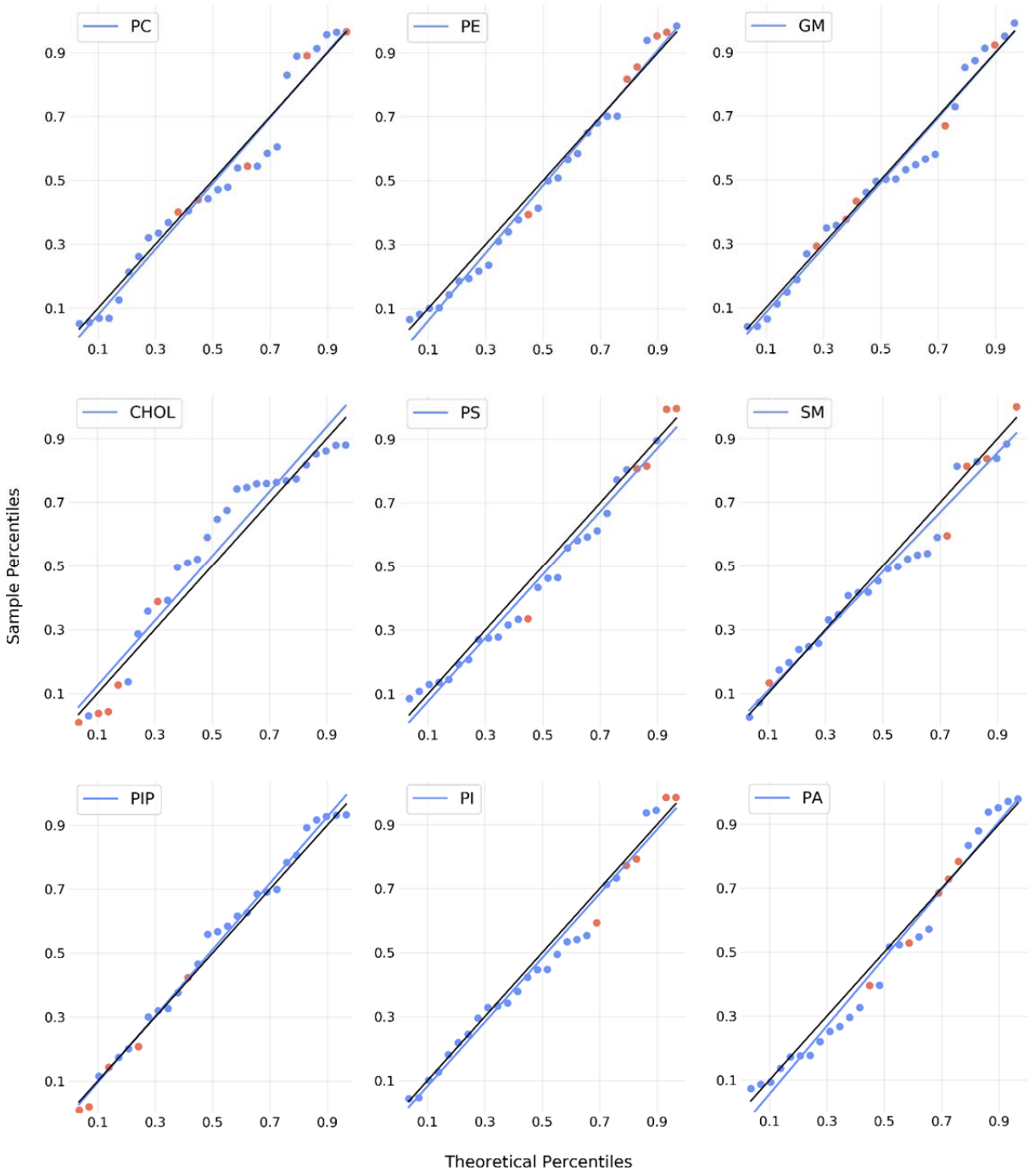


Figure S3. P-P plots of DE indices. *Each graph shows the probability plot for the DE indices calculated for all GPCRs per lipid group (PC, PE, GM, CHOL, PS, SM, PIP, PI, and PA). Each data point represents the DE index for a particular GPCR. Data points colored in red are DE indices for non class A GPCRs. In each graph, the data points are approximated with a line of best fit. The black line is drawn at 45˚ which corresponds to theoretical percentiles = sample percentiles.*

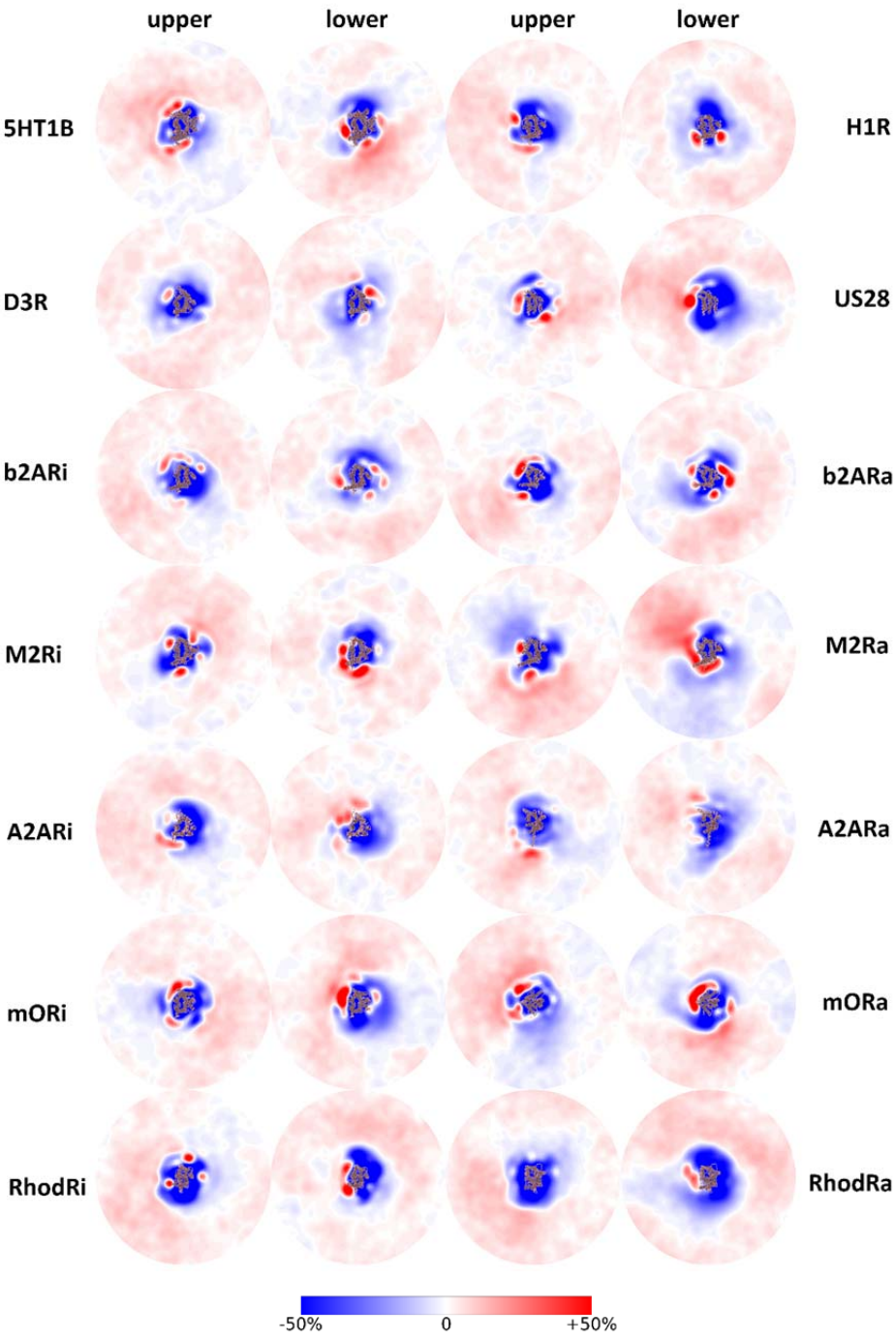


Figure S4. (continued)

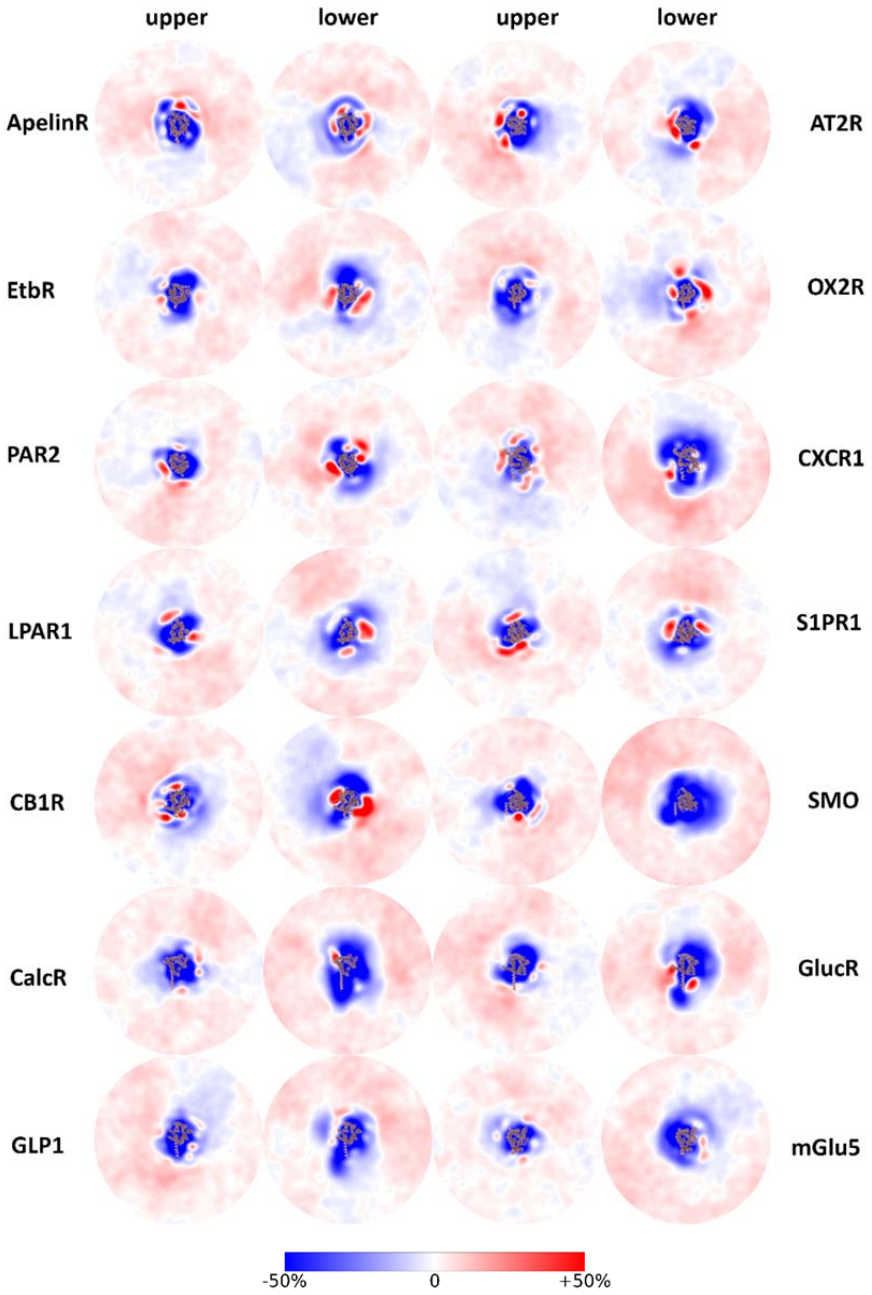


Figure S4. Cholesterol 2D density profile. *Density maps are on the x-y axis of each system and calculated for each leaflet separately. We show the density maps for cholesterol (this figure), fully-saturated lipids (Figure S3) and poly-unsaturated lipids (Figure S4). Densities are calculated for the last 5 μs and averaged over the four copies. To assist in analyzing and comparing the data, we have overlaid the atomistic structure over the approximate insertion place and orientation of proteins. The average*

density for each lipid group is shown with white, and the relative enrichment or depletion is shown if red and blue colors, respectively.

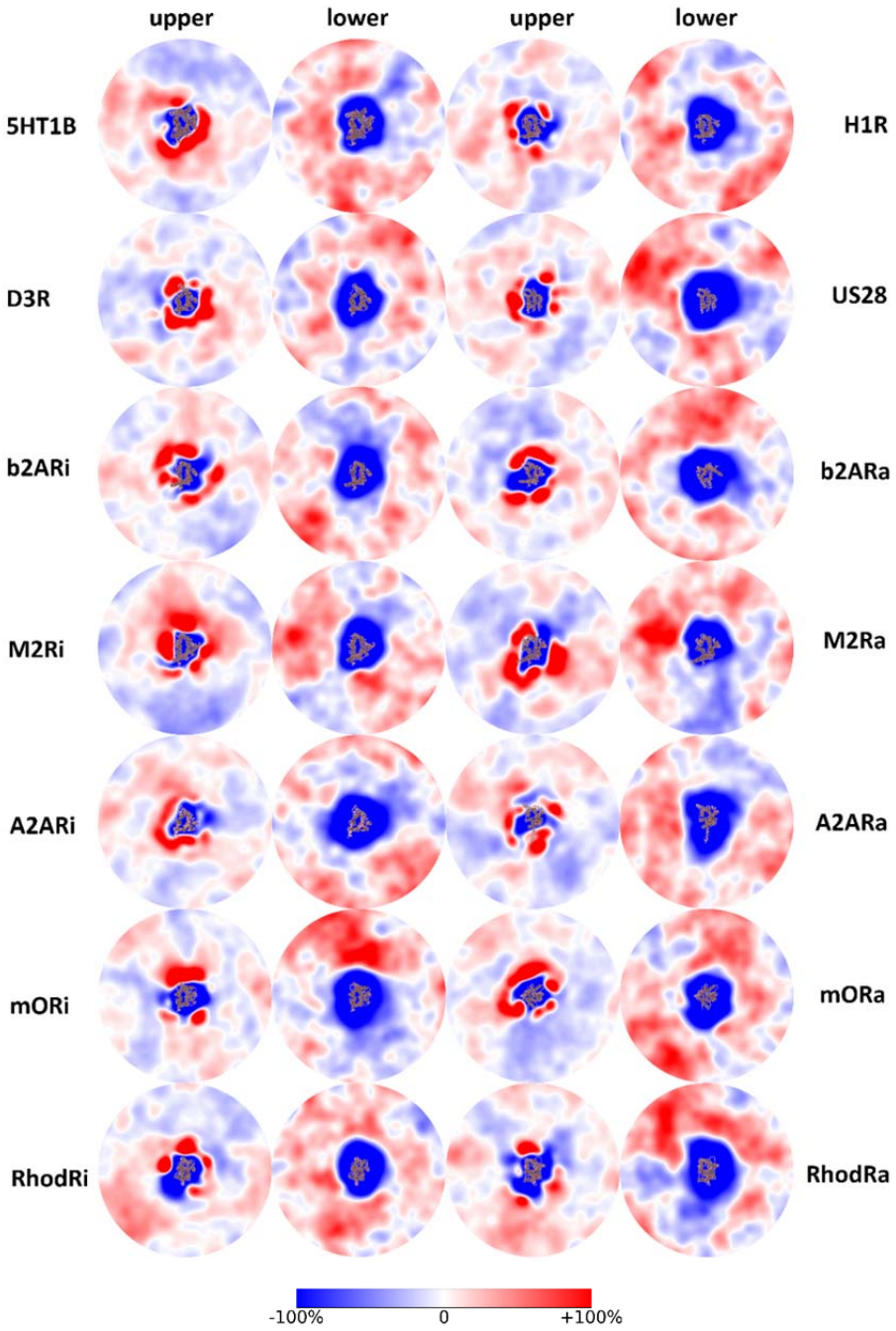


Figure S5. (continued)

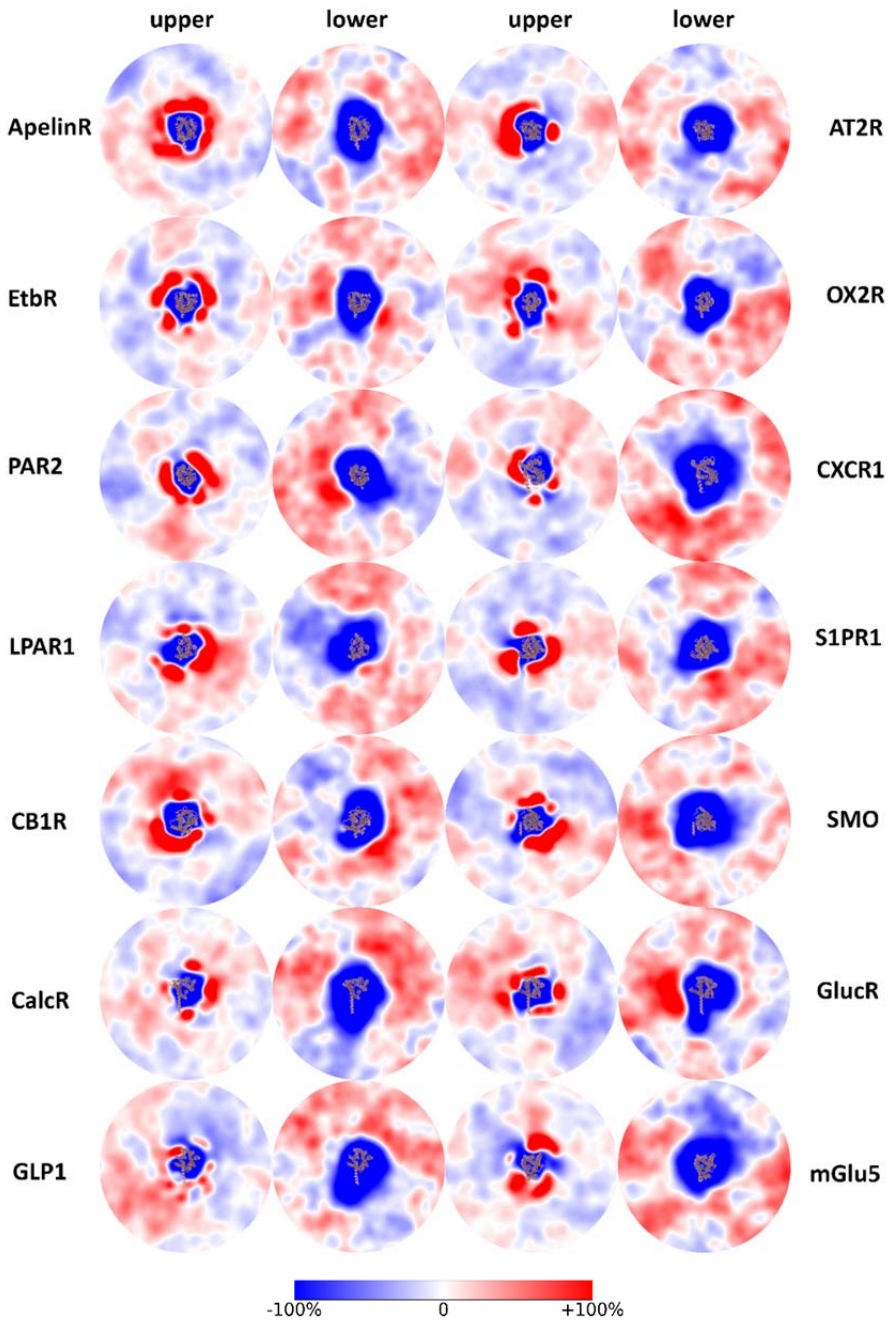


Figure S5. FS lipids density profile. *Details are similar to the cholesterol density maps (Figure S2).*

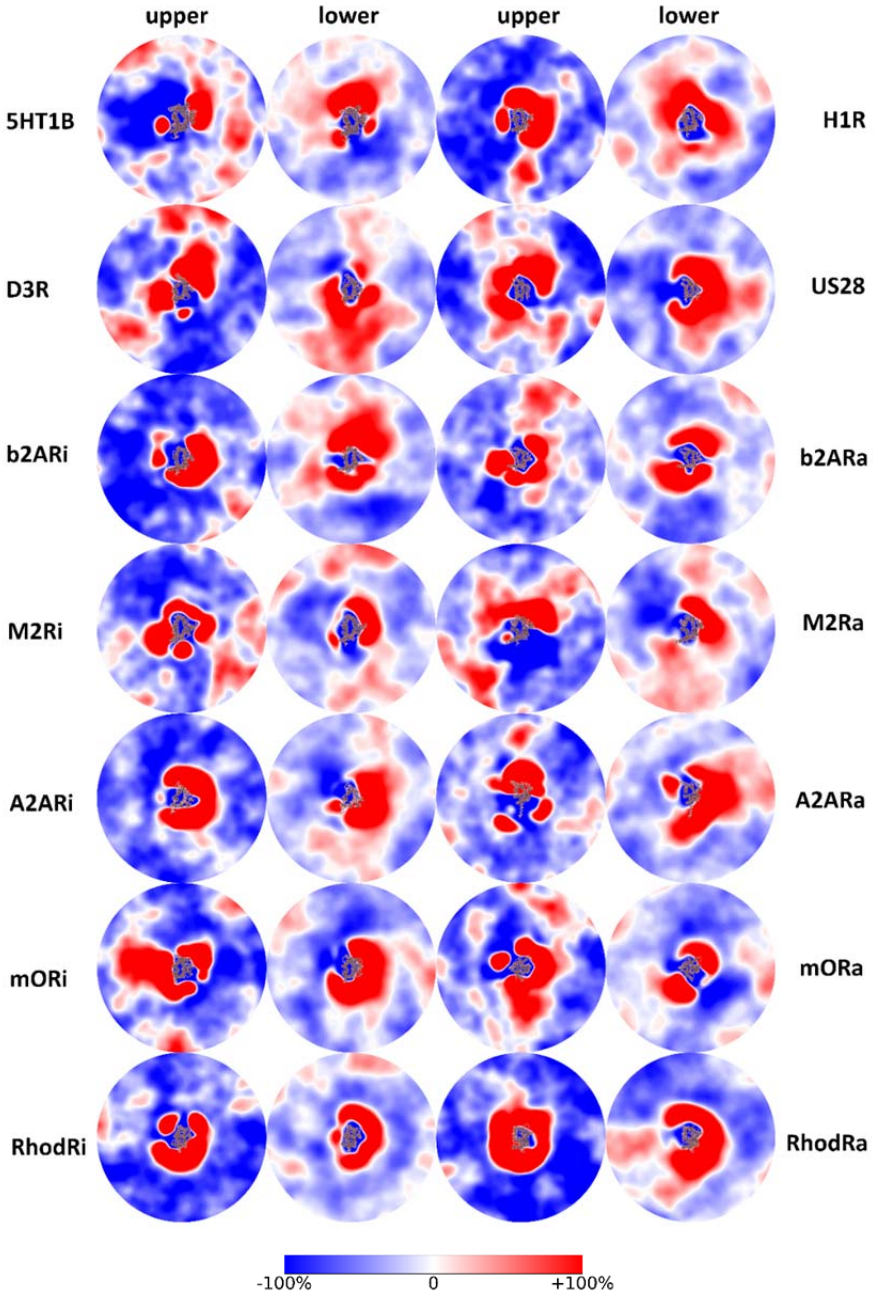


Figure S6. (continued)

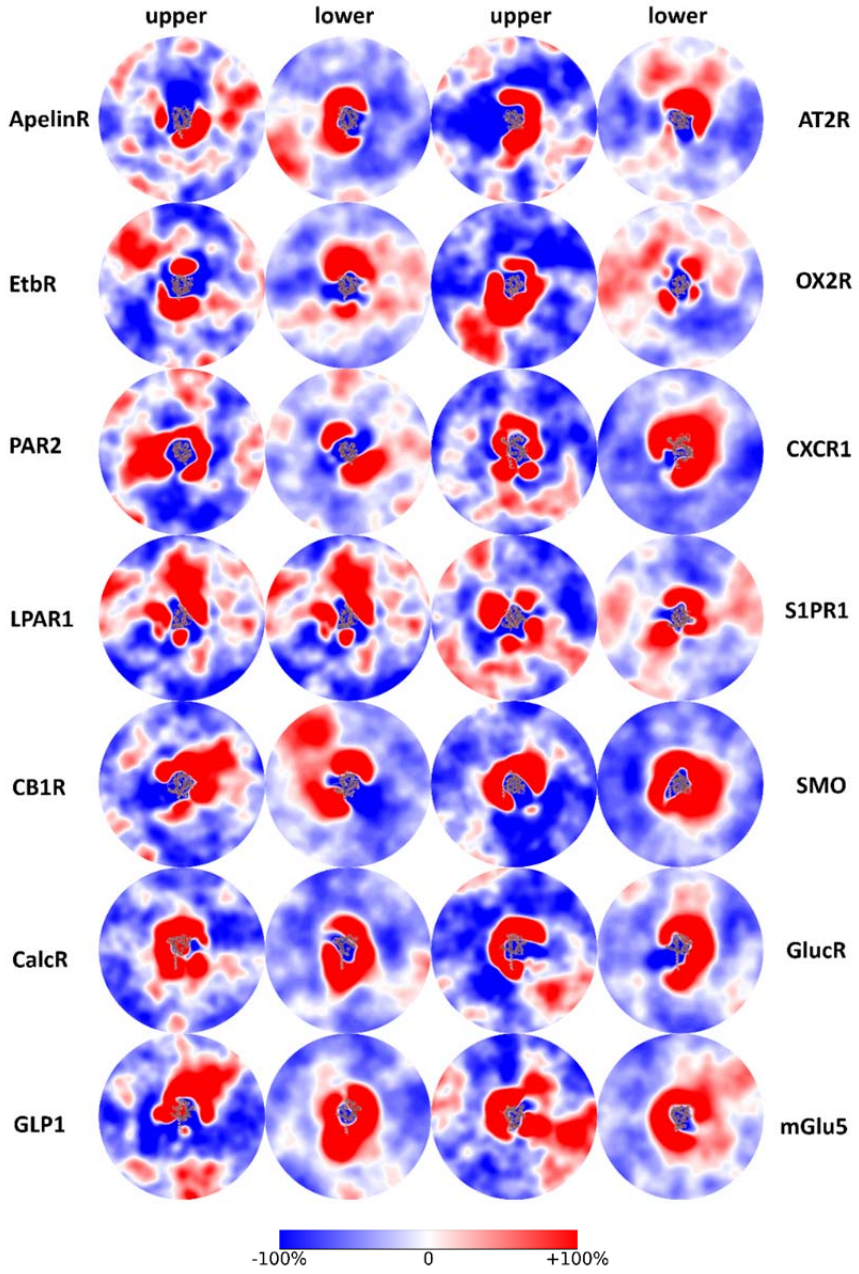


Figure S6. PU density profile. *Details are similar to the cholesterol density maps (Figure S2).*

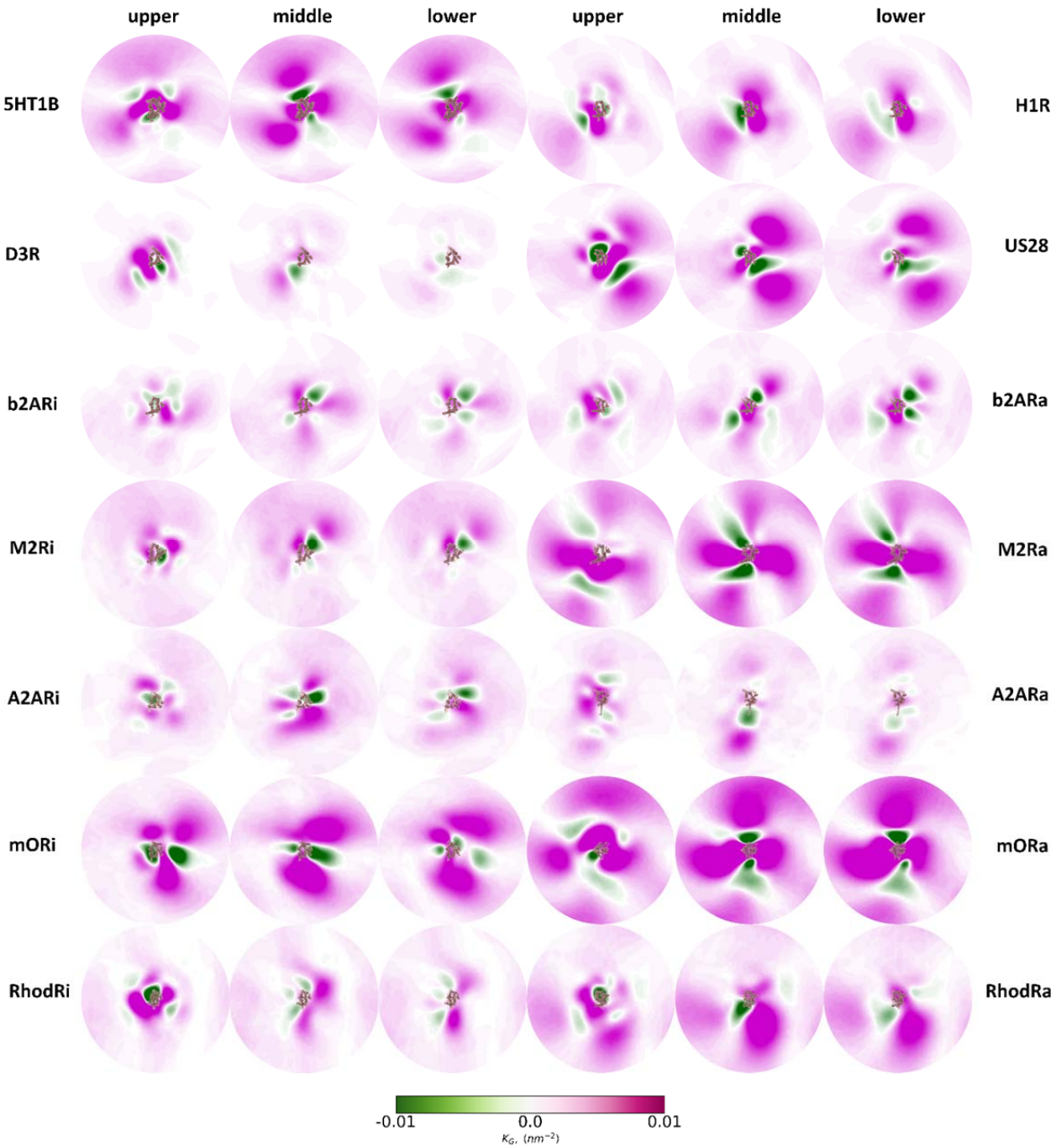


Figure S7. (continued)

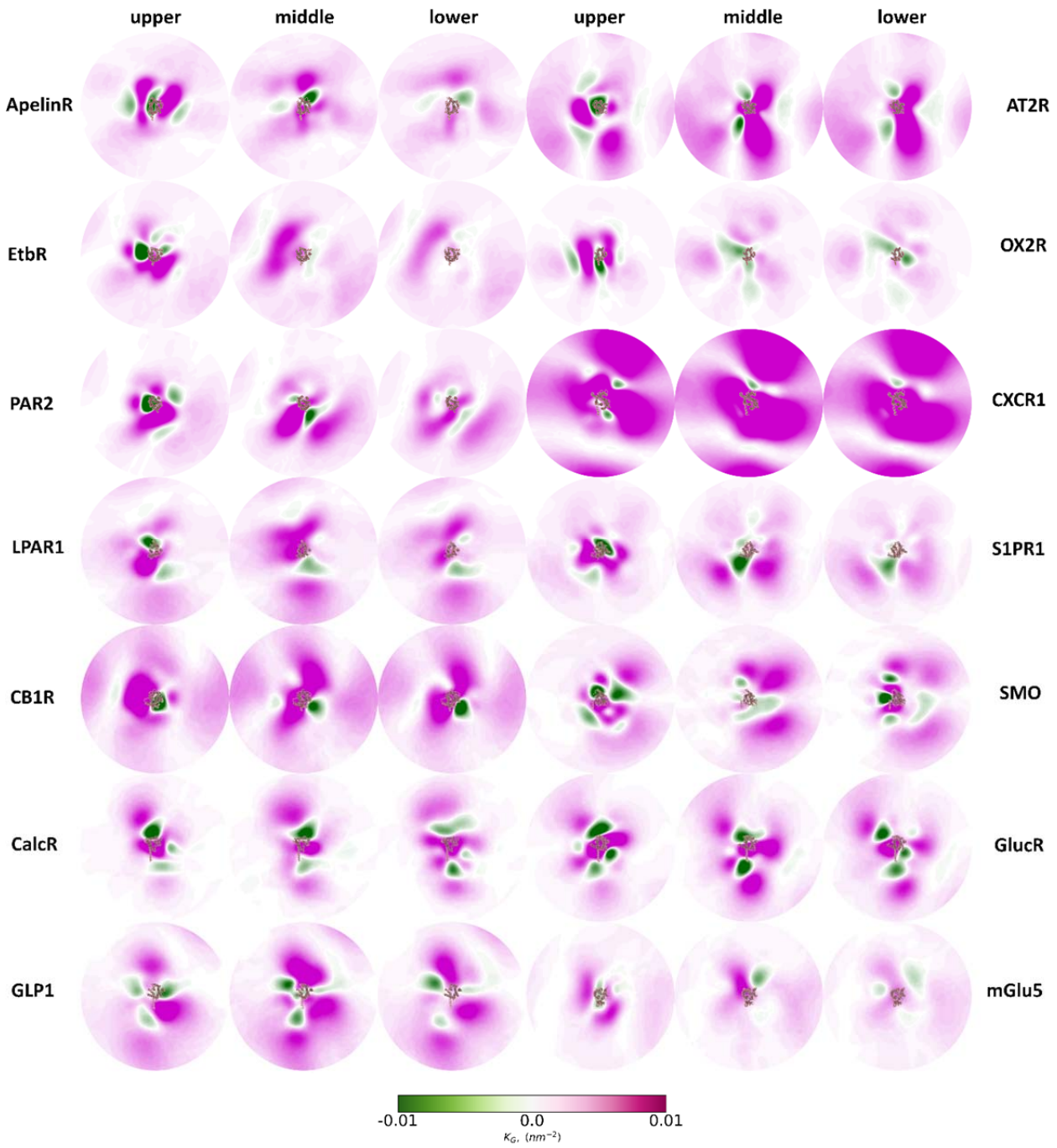


Figure S7. Gaussian curvature (K_G) maps. *2D K_G curvature maps are calculated for each system simulated over the last 5 μs. We define three surfaces for which we calculate curvature and thickness maps: upper, middle and lower. We use the PO4 and GM1 beads to define the upper surface and the PO4 and CP beads to define the lower surface. For the middle surface we use the last lipid tail beads. Saddles (negative K_G) are colored magenta and convex/concave regions (positive K_G) are colored green. Atomistic structures over overlaid at the approximate insertion place and orientation of proteins.*

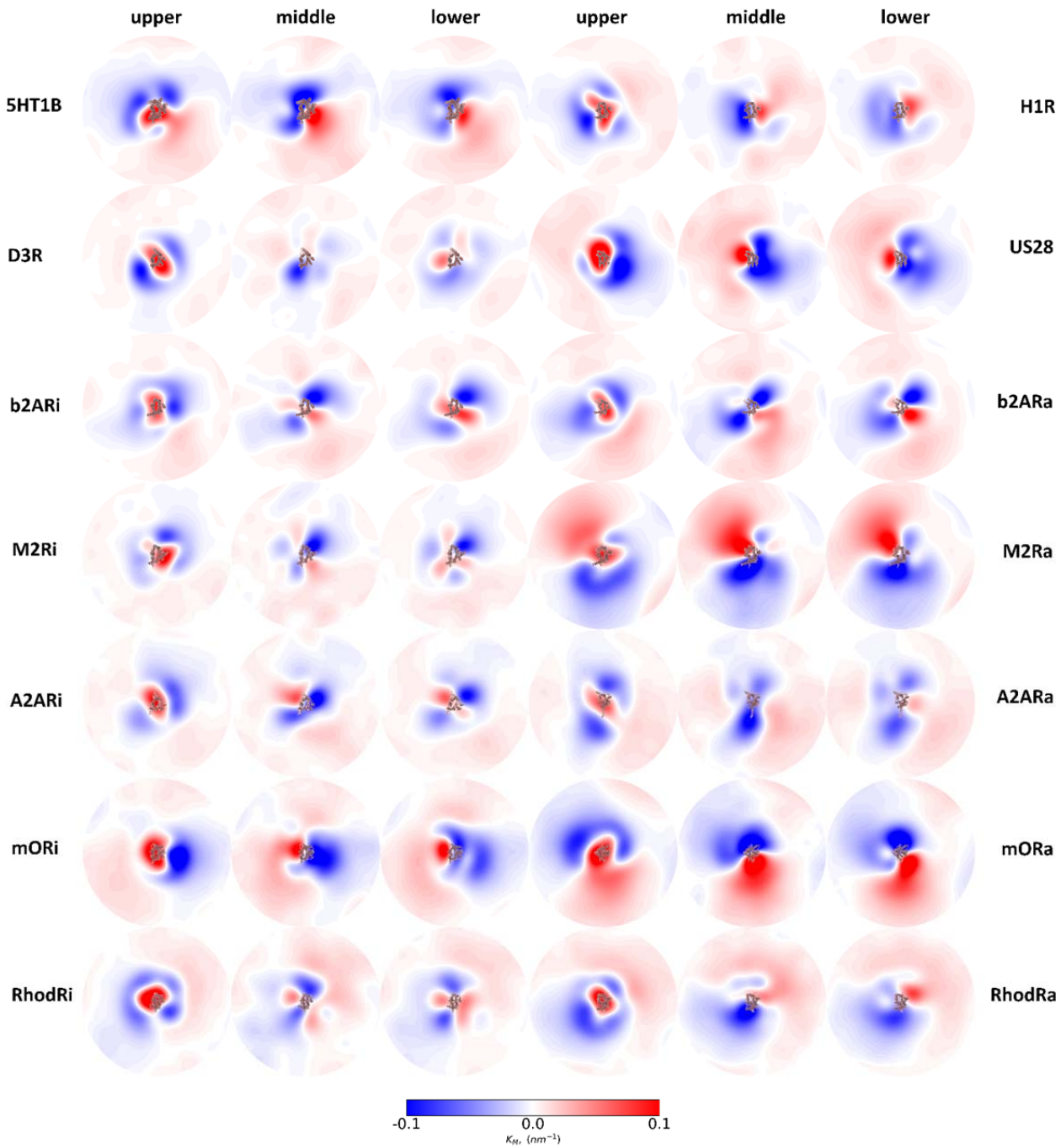


Figure S8. (continued)

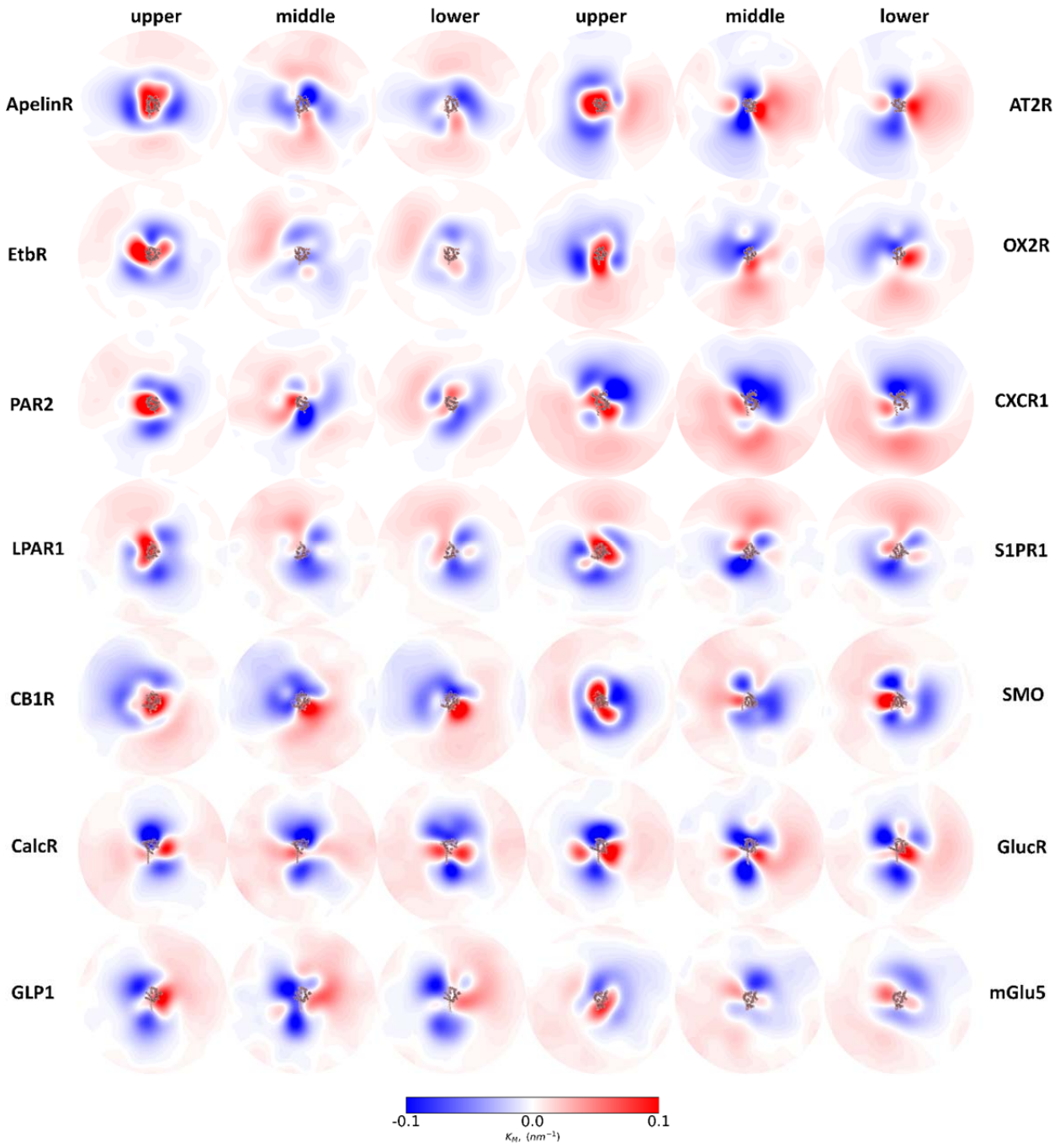


Figure S8. Mean curvature (K_M) maps. *2D K_M curvature maps are calculated for each system simulated over the last 5 μs. We color negative curvature and positive curvature with blue and red, respectively, using the RWB coloring scheme. white present zero curvature. Atomistic structures are overlaid at the approximate insertion place and orientation of proteins.*

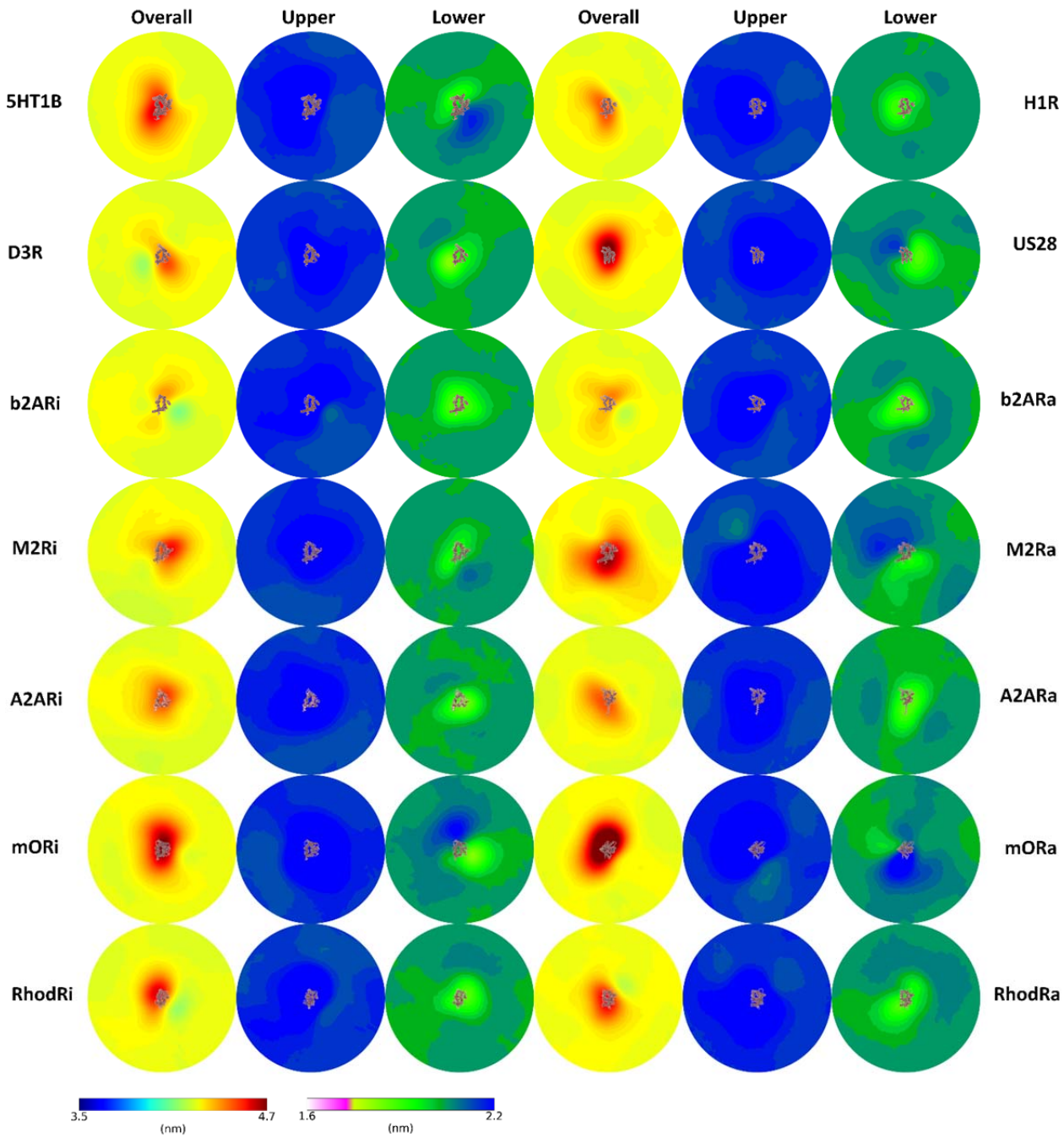


Figure S9. (continued)

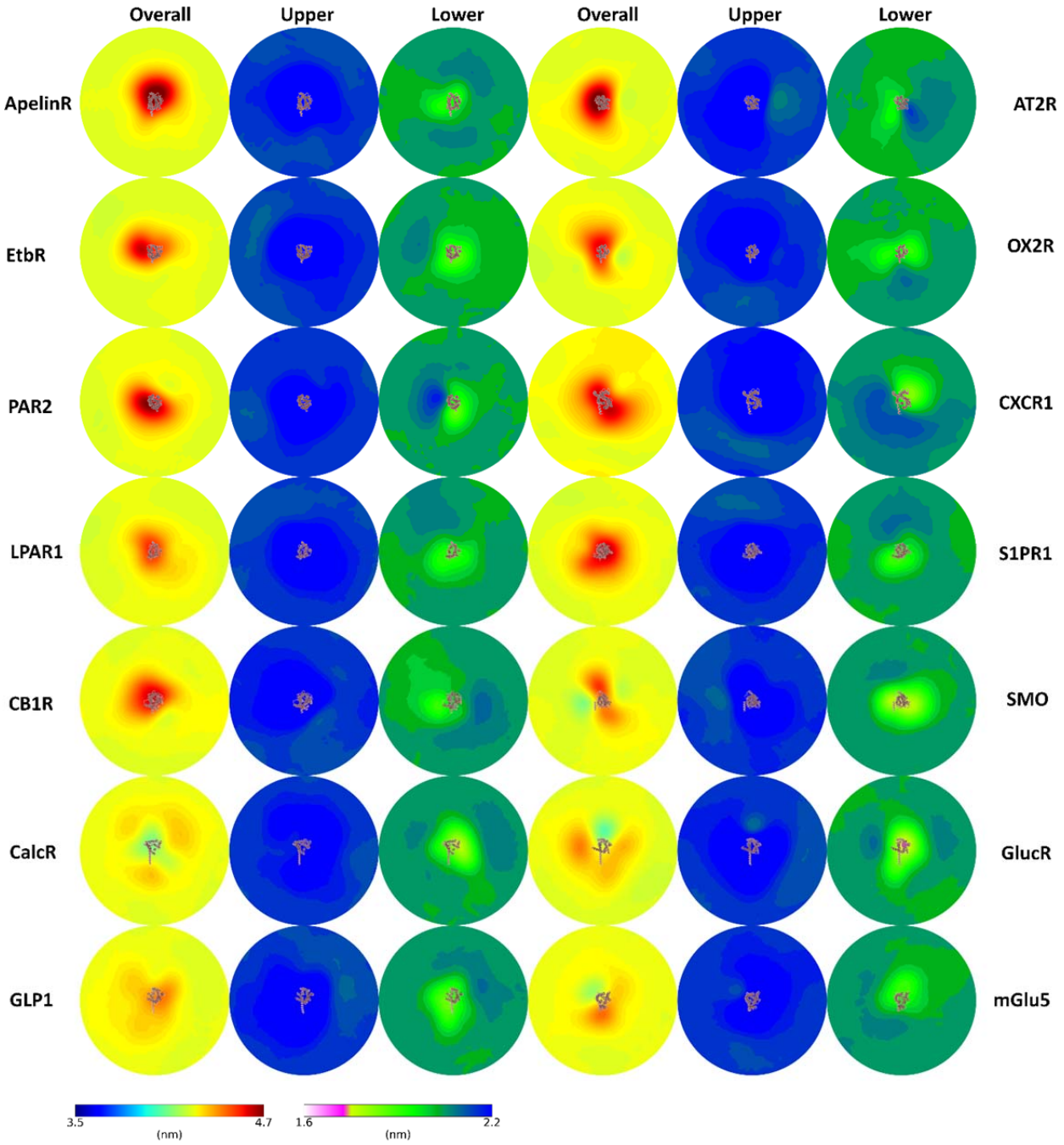


Figure S9. GPCR membrane thickness. *2D thickness maps are calculated for each system simulated over the last 5 μs. The same three surfaces defined for curvature analysis are used here as well. Overall thickness is calculated as the distance from the upper to the lower surface. Upper and lower thickness are the distance between the middle surface to the upper and lower surfaces, respectively. Atomistic structures over overlaid at the approximate insertion place and orientation of proteins.*

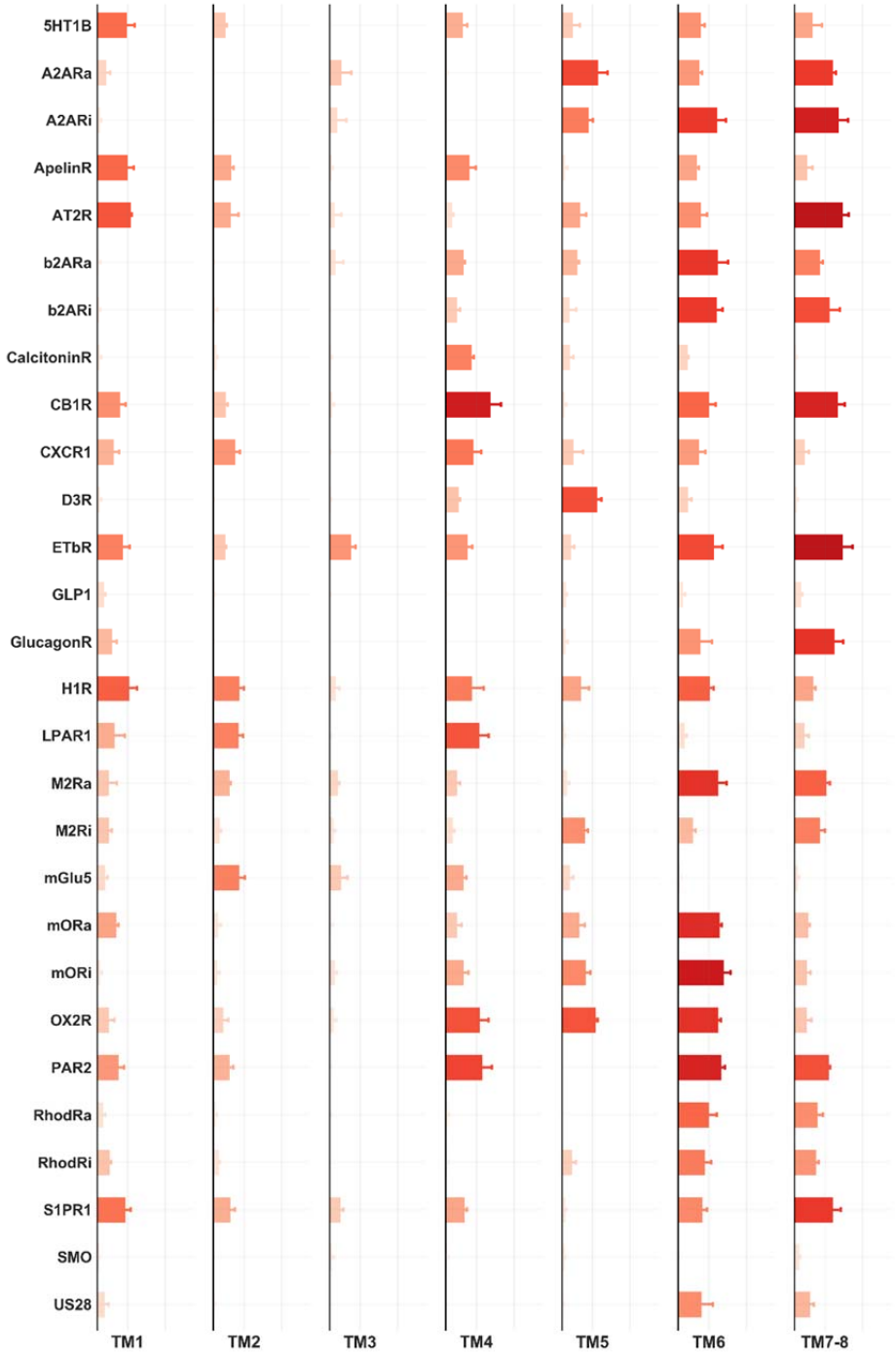


Figure 10. PIP lipid - TM helix interactions. *A bar plot showing the interactions between each GPCR TM helix with PIP lipids. For each helix, only the four highest residues in terms of interactions with PIP lipids are considered. The data show how different GPCRs are in their interactions with PIP lipids, in particular we note that non class A GPCRs (GLP1, mGlu5, SMO, CalcitoninR and GlucagonR) have strikingly fewer interactions with PIP lipids compared to their class A counterparts. A color gradient from white-to-red is used to color the increase in the number of contacts with PIP lipids.*

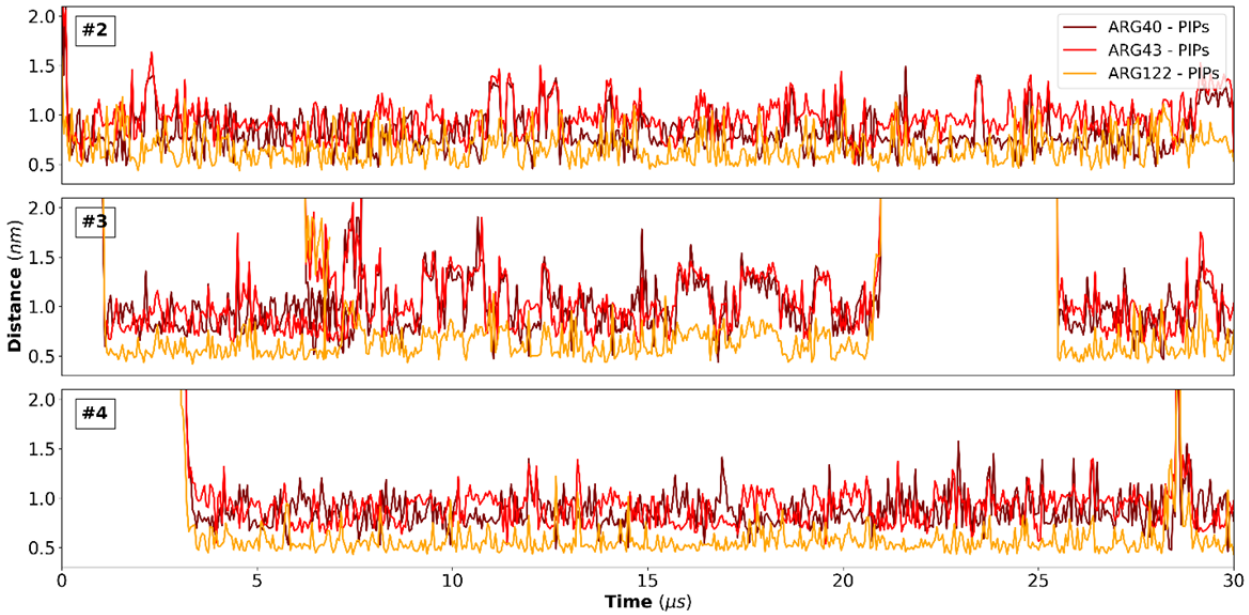


Figure S11. CXCR1-PIP interactions. *Interactions between bound PIP lipids and arginine residues lining up the CXCR1 binding site. We show the data for protein #2, #3, and #4 in our system. In each case, there is a PIP lipid that is tightly bound at the TM1/2/4 interface.*

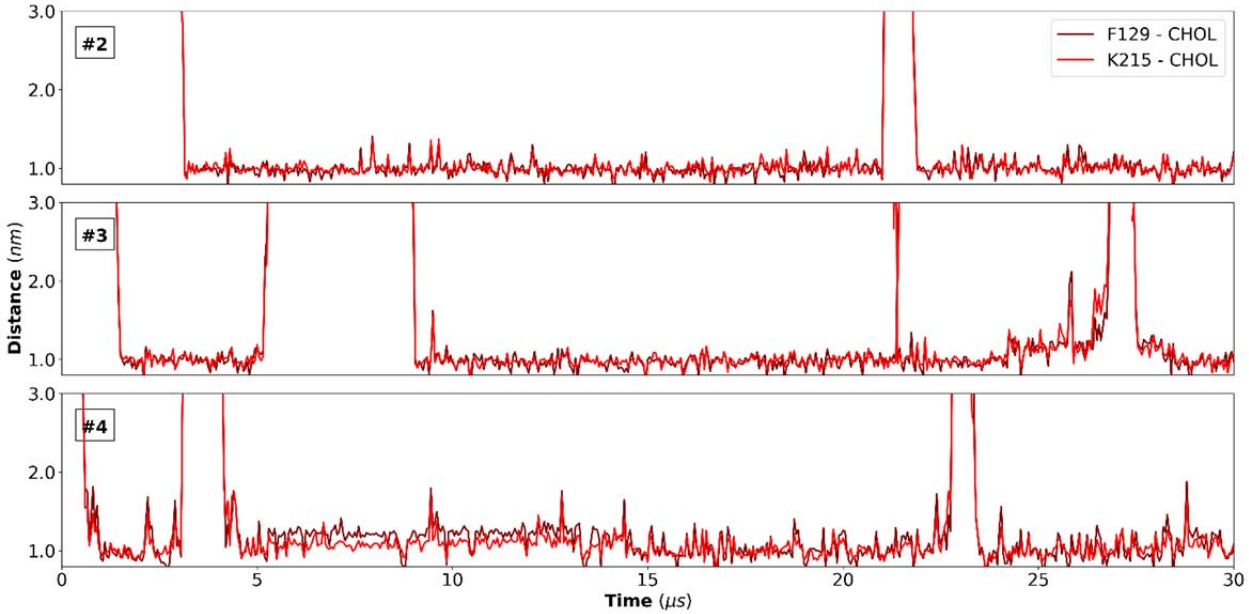


Figure S12. AT2R-cholesterol interactions. *Centre-of-mass distances between F129 and K215 of AT_2R and bound cholesterol molecules. We show the data for protein #2, #3, and #4 in our system. We see that*

in each case there is a cholesterol molecule tightly bound at the TM4/5 interface. We also observe multiple binding/unbinding events.

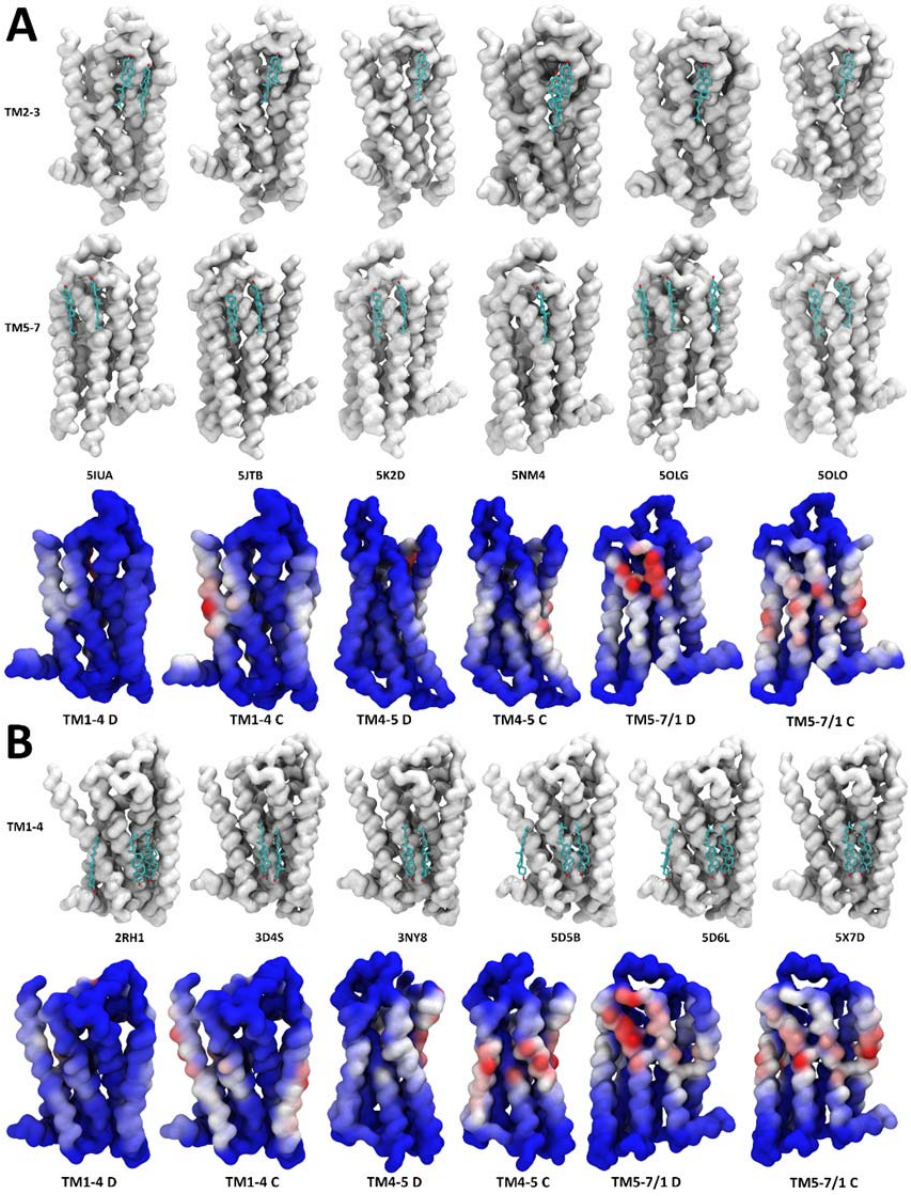


Figure S13. Cholesterol interactions with $A_{2A}R$ and β_2AR. *Crystal structures are shown as a surface representation to allow for easy comparison with simulation results. The latter are mapped on the surface of proteins in a Red-White-Blue scale denoting an increase in either cholesterol number of contacts (denoted with the letter C) or duration of contacts (letter D). We only show one crystal structure per study published.* ***A. $A_{2A}R$****. 27 of the 45 $A_{2A}R$ crystal structures solved contain bound cholesterol. They are found at the TM2/3 and TM5-7 interface. Song et al.(43) show cholesterol interacting with A_2AR at 7 different*

interaction sites: TM1/2(ec), TM1/2/4 (ic), TM2/3 (ec), TM3/4 (ec), TM3/4 (ic), TM5/6 (ec), and TM6/7. The number of these interaction sites is dependent on the conformational state of $A_{2A}R$. Rouviere et al.(44) as well performed coarse-grained MD simulations and showed several cholesterol interaction sites. Using long-scale all-atom simulations they noted two sites with a particularly high cholesterol-binding affinity: TM5/6(ic) and TM6(ec). While we do observe cholesterol interactions at the TM2-4 interface, the majority of these interactions are seen at the TM5-7/1 interface. Accounting for the longevity of cholesterol-contacts, the $A_{2A}R$ TM6-7 interface has a distinctly higher affinity for cholesterol. We also find the following interaction sites: TM5-6, TM7-8/1, TM1-2, and TM2-4. The TM5-6 interface noted here and by Rouviere et al,(44) has also been found to be the entry site for cholesterol insertion inside the receptor by Guixa et al.(45) Such lipid entry events are unlikely to occur in CGMD simulations employing an elastic network. ***B. β_2AR.*** *They account for 11 of the 64 cholesterol – co-crystalized GPCR structures. There are two cholesterol binding sites in the β_2AR crystal structures: TM1-4 and TM1/8. In microsecond long all-atom simulations of β_2AR, Cang et al.(46) observe 8 cholesterol interaction sites: 5 in the lower leaflet and 3 in the upper leaflet side of the receptor. In particular, they note interaction sites between helices TM1/7, TM6/7 and TM1/8. These simulations also show that the TM5-7 interface of β_2AR is quite involved in cholesterol interactions, despite this not being reflected in solved crystal structures. Manna et al.(47) also find several high-density cholesterol interaction sites in their simulations. In the lower half of the receptor they find interaction sites at TM1-TM4 and TM5/6, and in the upper half of the receptor they find one cholesterol binding at the TM5/6 and another one at the TM6/7 interface. Agreeing with previous simulations, we find several cholesterol interaction sites. The TM5-7 interface of β_2AR, in particular, displays a cholesterol-interaction profile that is quite similar to that of $A_{2A}R$. The β_2AR TM1-4 interface, however, is distinctly different. It interacts with cholesterol molecules in the lower half of the receptor where helices TM1/2/4 converge and corresponds to the cholesterol binding site from crystal structures. We also note the existence of an interaction site between helices TM4/3/5, which is not observed for $A_{2A}R$. Lastly, we note the TM1/8 interaction site, which matches the second cholesterol binding site found in crystal structures.*

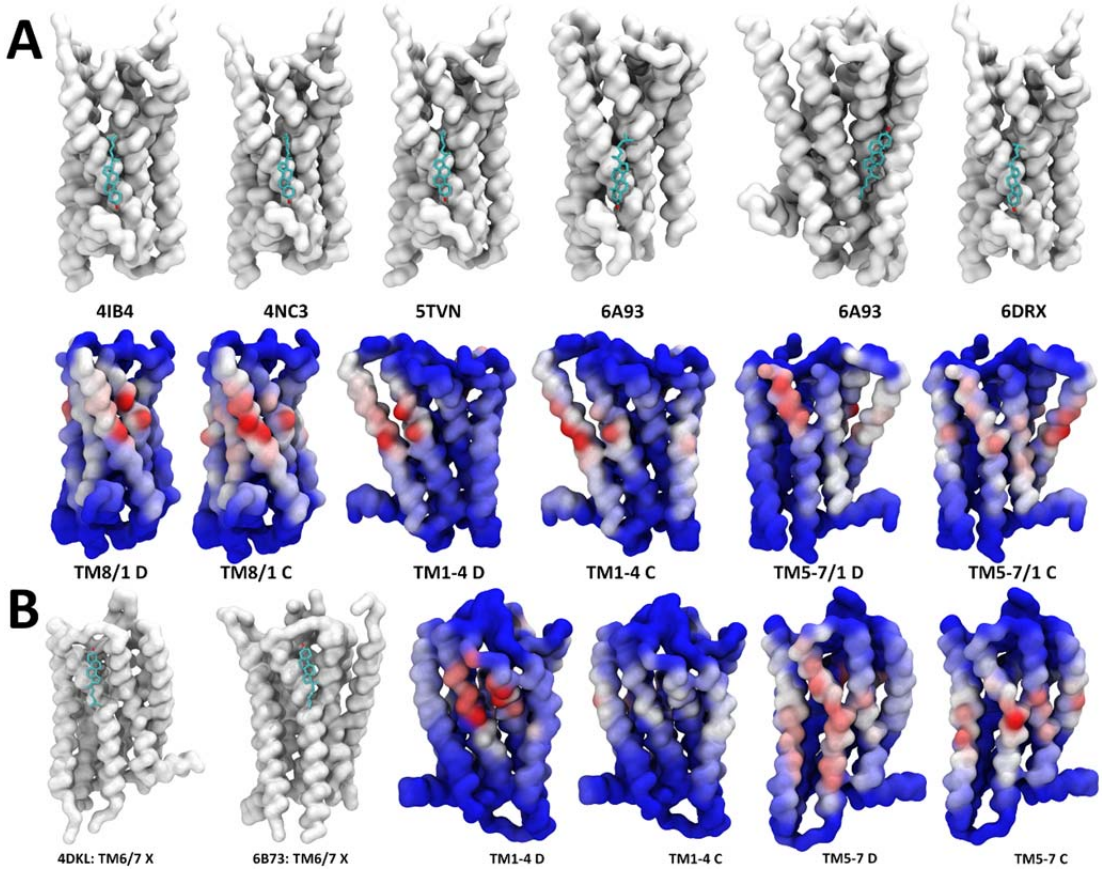


Figure S14. Cholesterol interactions with Serotonin and μOR. A. Serotonin. *There are 9 cholesterol – co-crystalized structures of serotonin receptors. The majority belong to $5HT_{2B}$ receptor, and while we did not simulate this particular subtype, we simulated $5HT_{1B}$ which shares a high sequence similarity to $5HT_{2B}$. $5HT_{2B}$ crystal structures reveal a cholesterol molecule bound in the lower half of the receptor at TM8/1. One of the structures solved has a cholesterol molecule in a crevice between helices TM2/3/4 not unlike the binding site observed for β_2AR, however and surprisingly, here cholesterol is found in the opposite direction with its OH group facing upward and, positioned at the midplane of the bilayer. Shan et al.(48) carried out simulations of 5-HT2AR and found several cholesterol interaction sites: TM1/2/4 (ic) TM2/3 (ec) and TM6/7 (ec and ic side). In our simulations, one of the most pronounced interaction sites observed is between TM1/8 which matches crystallographic experimental evidence. We also note the existence of other interaction sites located at TM1/2, TM2/4, TM6/7 and TM7/TM1. The interface between helices TM6/7 and TM8/1 are observed for β_2AR, $A_{2A}R$ and $5HT_{1B}$. Cholesterol interaction sites are also observed at the TM1-4 interface, on the ic side for β_2AR and $5HT_{1B}$, and on the ec side for $A_{2A}R$.* ***B. μOR.*** *3 opioid receptors have been crystalized with bound cholesterol: kappa-Opioid (1) receptor and mu-Opioid receptor (2). Marino et al.(49) using CGMD and a setup similar to ours to simulate μOR show the existence of an interaction site in a hydrophobic region close to TM6 and TM7 of the receptor which*

corresponds to the position of cholesterol resolved in the crystal structure of μOR (4DKL). In our simulations, we reproduce this cholesterol-binding site at TM6/7 quite well. Here, again, we find the TM1-4 interface to be involved in cholesterol interactions. The interaction site we identify is formed from helices TM2/3 with partial involvement of TM4 in the ec side of the receptor. We also note an interaction site at TM5/6 ic side.

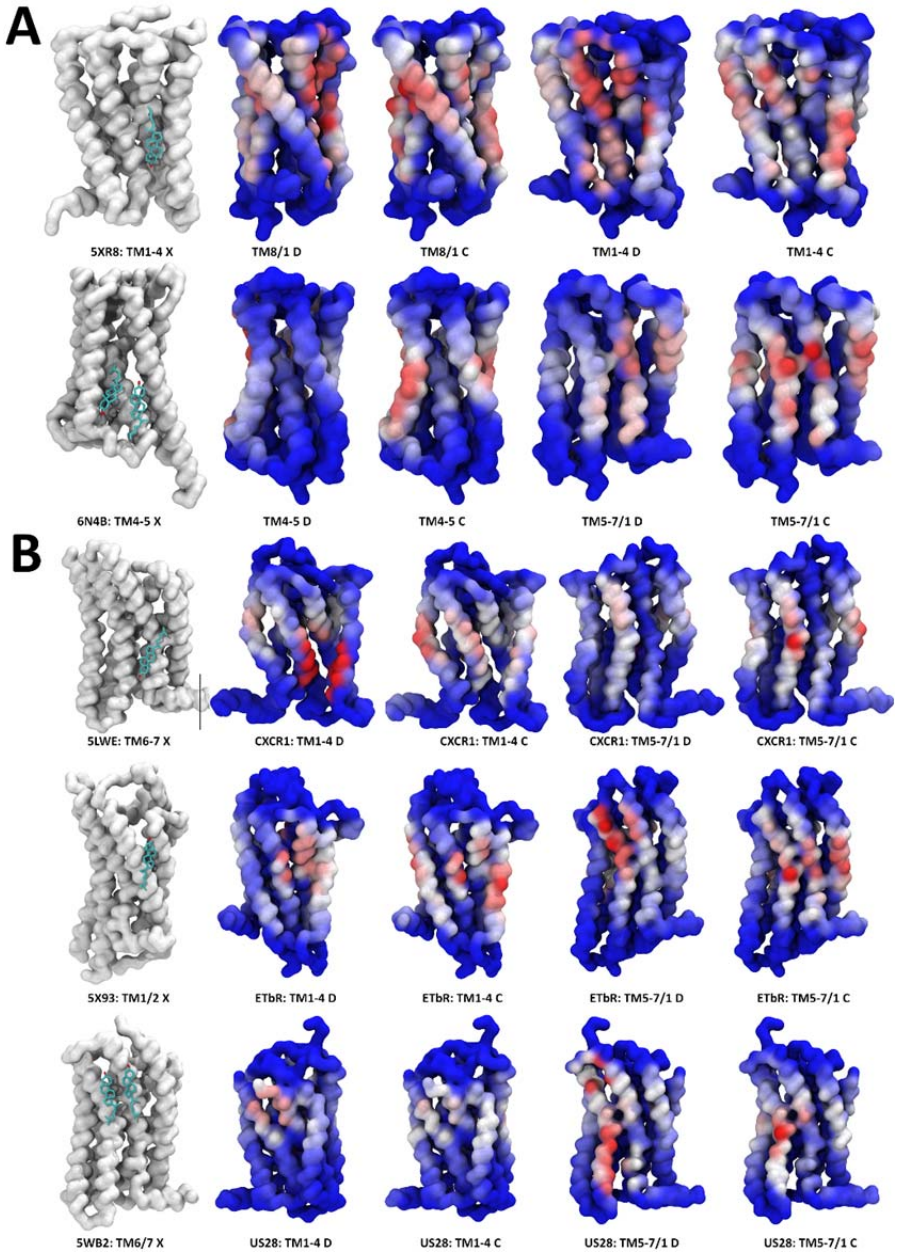


Figure S15. Cholesterol interactions with CB1R, chemokine receptor, ET_BR, and US28. A. CBR. *Currently, there are three structures of the cannabinoid receptor (CB1) that have been solved with bound*

cholesterol. Two of them (5XR8 and 5XRA) have cholesterol bound at the TM6/7 interface and the third (6N4B) has two cholesterol molecules bound at the TM4/5 interface. CB_1R differs from other GPCRs so far in that the TM1-4 interface is very heavily involved in interacting with cholesterol. We observe cholesterol interaction sites at: TM7/1, TM1/2 (ec and ic side), TM2/3, TM4/5 and TM6/7 (ic side). We also observe the interaction site at TM4/5 seen in the crystal structure, although in terms of the duration of cholesterol contacts it is of lesser "strength" than the other sites. ***B. CCR9, ET_BR and US28.*** *We summarize the data for three additional GPCRs with co-crystalized cholesterol molecules (from top row to down row): CCR9 (a chemokine), ET_BR and US28. They feature cholesterol bound at different sites, namely: TM6/TM (ic), TM1/2 (ec) and TM6/7 (ec). We observe the cholesterol interaction sites from the crystal structures in all three cases, and in particular for ET_BR and US28. We observe a stronger interaction between cholesterol and helices TM2/4 – perhaps because we simulated CXCR1 and not CCR9.*

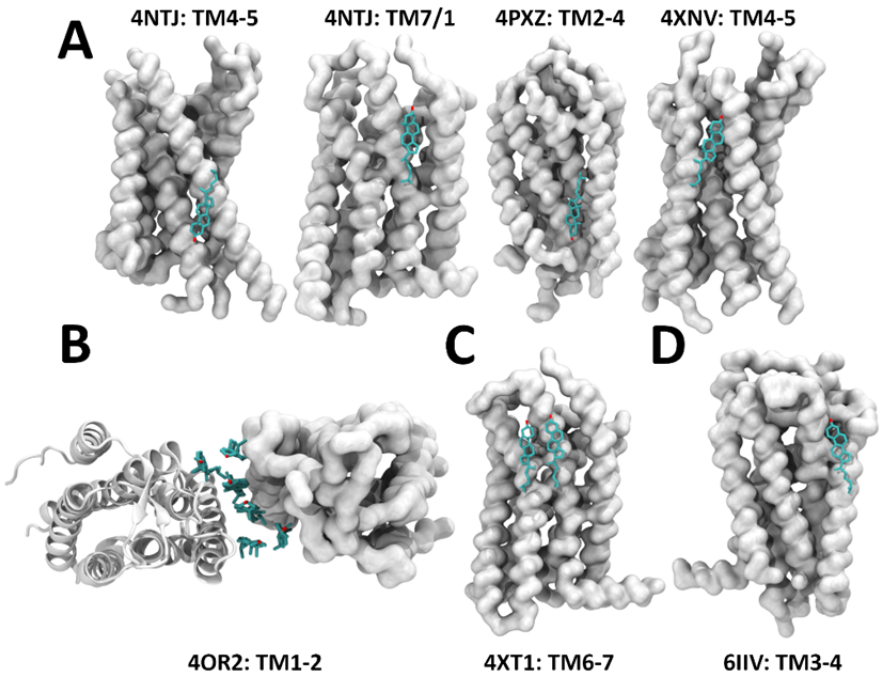


Figure S16. Cholesterol interactions with other GPCR. *We show the rest of GPCR crystal structures that have been solved with cholesterol molecules bound to them.* ***A. P2Y12 and P2Y1.*** *Cholesterol is bound at three P2Y12 sites: TM3/4 interface, which hasn't been observed in other crystal structures, TM7/1 which is observed in our simulations (notably, RhodRi), and TM1-4 (ic). P2Y1, on the other hand, features a cholesterol molecule bound at the TM4-5 interface, at the exact location where we observe cholesterol binding inside AT2R.* ***B. mGlu1.*** *mGlu1 has been solved as a dimer and the interface between*

monomers is lined up with 6 cholesterol molecules. These are found at the TM1/2/3 interface. ***C. CX3CL1.*** *Two cholesterol molecules are found at the TM6/7 interface.* ***D. Thromboxane A2.*** *One cholesterol molecule is found at the TM2/3 interface, which is not observed elsewhere, and is also rare in our simulations.*

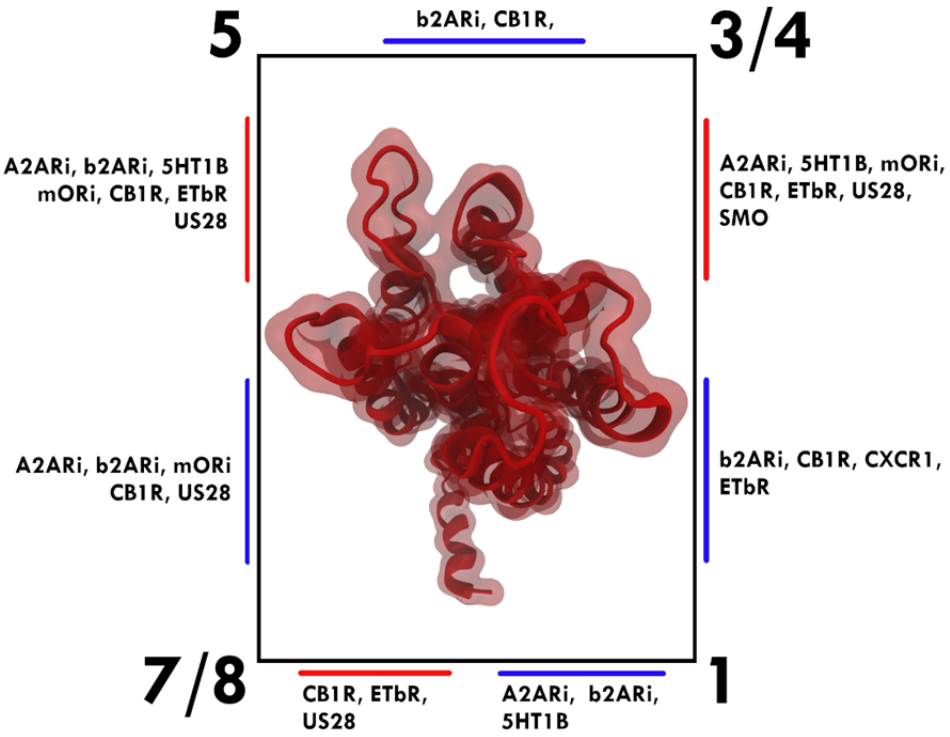


Figure S17. GPCR-lipid interactions. *We approximate GPCRs with a rectangular parallelogram (or cuboid) and define two large interfaces (TM1-4 and TM5-7/8), and two smaller interfaces (TM7/8-1 and TM3/4 – 5). TM helices as part of a cuboid face (shown as a rectangle in the picture) do not form any interface with TM helices part of other faces. We also use colors red and blue to differentiate between ec and ic located binding sites, respectively. For the TM7/8 – 1 interface, red and blue, in fact, represent the TM7/1 and TM1/8 interfaces respectively. The figure shows the major interaction sites for a selection of GPCRs in a schematic way.*

SUPPORTING REFERENCES

1. Corradi V., E. Mendez-Villuendas, H. I. Ingolfsson, R. X. Gu, I. Siuda, M. N. Melo, A. Moussatova, L. J. DeGagne, B. I. Sejdiu, G. Singh, T. A. Wassenaar, K. D. Magnero, S. J. Marrink, D. P. Tieleman. Lipid-Protein Interactions Are Unique Fingerprints for Membrane Proteins. Acs Central Sci. 2018;4(6):709-717.
2. Gu R.-X., S. Baoukina, D. P. Tieleman. Cholesterol Flip-Flop in Heterogeneous Membranes. Journal of chemical theory and computation. 2019;15(3):2064-2070.
3. Isberg V., S. Mordalski, C. Munk, K. Rataj, K. Harpsoe, A. S. Hauser, B. Vroling, A. J. Bojarski, G. Vriend, D. E. Gloriam. GPCRdb: an information system for G protein-coupled receptors. Nucleic Acids Res. 2016;44(D1):D356-D364.
4. Fantini J., F. J. Barrantes. How cholesterol interacts with membrane proteins: an exploration of cholesterol-binding sites including CRAC, CARC, and tilted domains. Front Physiol. 2013;4:9.
5. De Castro E., C. J. Sigrist, A. Gattiker, V. Bulliard, P. S. Langendijk-Genevaux, E. Gasteiger, A. Bairoch, N. Hulo. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res. 2006;34(suppl_2):W362-W365.
6. Jianyi Y., Z. Yang. GPCR-EXP: a database for experimentally solved GPCR structures.
7. McGibbon R. T., K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L.-P. Wang, T. J. Lane, V. S. Pande. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. Biophys J. 2015;109(8):1528-1532.
8. Jones E., T. Oliphant, P. Peterson. {SciPy}: Open source scientific tools for {Python}. 2014.
9. Seabold S., J. Perktold, editors. Statsmodels: Econometric and statistical modeling with python. Proceedings of the 9th Python in Science Conference; 2010: Scipy.
10. Hunter J. D. Matplotlib: A 2D graphics environment. Computing in science & engineering. 2007;9(3):90.
11. Waskom M., O. Botvinnik, P. Hobson, J. Warmenhoven, J. Cole, Y. Halchenko, J. Vanderplas, S. Hoyer, S. Villalba, E. Quintero. Seaborn: statistical data visualization. Seaborn: Statistical Data Visualization Seaborn 0. 2014;5.
12. Rose A. S., P. W. Hildebrand. NGL Viewer: a web application for molecular visualization. Nucleic Acids Res. 2015;43(W1):W576-W579.
13. Cignoni P., M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, G. Ranzuglia, editors. Meshlab: an open-source mesh processing tool. Eurographics Italian chapter conference; 2008.

14. Humphrey W., A. Dalke, K. Schulten. VMD: visual molecular dynamics. Journal of molecular graphics. 1996;14(1):33-38.

15. Wang C., Y. Jiang, J. Ma, H. Wu, D. Wacker, V. Katritch, G. W. Han, W. Liu, X.-P. Huang, E. Vardy. Structural basis for molecular recognition at serotonin receptors. Science. 2013;340(6132):610-614.

16. Lebon G., T. Warne, P. C. Edwards, K. Bennett, C. J. Langmead, A. G. Leslie, C. G. Tate. Agonist-bound adenosine A 2A receptor structures reveal common features of GPCR activation. Nature. 2011;474(7352):521.

17. Jaakola V. P., M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. T. Chien, J. R. Lane, A. P. Ijzerman, R. C. Stevens. The 2.6 Angstrom Crystal Structure of a Human A(2A) Adenosine Receptor Bound to an Antagonist. Science. 2008;322(5905):1211-1217.

18. Ma Y., Y. Yue, Y. Ma, Q. Zhang, Q. Zhou, Y. Song, Y. Shen, X. Li, X. Ma, C. Li. Structural basis for apelin control of the human apelin receptor. Structure. 2017;25(6):858-866. e854.

19. Zhang H., G. W. Han, A. Batyuk, A. Ishchenko, K. L. White, N. Patel, A. Sadybekov, B. Zamlynny, M. T. Rudd, K. Hollenstein. Structural basis for selectivity and diversity in angiotensin II receptors. Nature. 2017;544(7650):327.

20. Rasmussen S. G., B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara, B. K. Kobilka. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. Nature. 2011;477(7366):549-555.

21. Cherezov V., D. M. Rosenbaum, M. A. Hanson, S. G. F. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka, R. C. Stevens. High-resolution crystal structure of an engineered human beta(2)-adrenergic G protein-coupled receptor. Science. 2007;318(5854):1258-1265.

22. Hua T., K. Vemuri, M. Pu, L. Qu, G. W. Han, Y. Wu, S. Zhao, W. Shui, S. Li, A. Korde. Crystal structure of the human cannabinoid receptor CB1. Cell. 2016;167(3):750-762. e714.

23. Park S. H., B. B. Das, F. Casagrande, Y. Tian, H. J. Nothnagel, M. Chu, H. Kiefer, K. Maier, A. A. De Angelis, F. M. Marassi. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. Nature. 2012;491(7426):779.

24. Chien E. Y., W. Liu, Q. Zhao, V. Katritch, G. W. Han, M. A. Hanson, L. Shi, A. H. Newman, J. A. Javitch, V. Cherezov. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. Science. 2010;330(6007):1091-1095.

25. Shihoya W., T. Nishizawa, K. Yamashita, A. Inoue, K. Hirata, F. M. N. Kadji, A. Okuta, K. Tani, J. Aoki, Y. Fujiyoshi, T. Doi, O. Nureki. X-ray structures of endothelin ETB receptor bound to clinical antagonist bosentan and its analog. Nat Struct Mol Biol. 2017;24(9):758-+.

26. Shimamura T., M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G. W. Han. Structure of the human histamine H 1 receptor complex with doxepin. Nature. 2011;475(7354):65.

27. Chrencik J. E., C. B. Roth, M. Terakado, H. Kurata, R. Omi, Y. Kihara, D. Warshaviak, S. Nakade, G. Asmar-Rovira, M. Mileni. Crystal structure of antagonist bound human lysophosphatidic acid receptor 1. Cell. 2015;161(7):1633-1643.

28. Kruse A. C., A. M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hübner, E. Pardon, C. Valant, P. M. Sexton. Activation and allosteric modulation of a muscarinic acetylcholine receptor. Nature. 2013;504(7478):101.

29. Haga K., A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, T. Haga. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. Nature. 2012;482(7386):547.

30. Huang W. J., A. Manglik, A. J. Venkatakrishnan, T. Laeremans, E. N. Feinberg, A. L. Sanborn, H. E. Kato, K. E. Livingston, T. S. Thorsen, R. C. Kling, S. Granier, P. Gmeiner, S. M. Husbands, J. R. Traynor, W. I. Weis, J. Steyaert, R. O. Dror, B. K. Kobilka. Structural insights into mu-opioid receptor activation. Nature. 2015;524(7565):315-+.

31. Manglik A., A. C. Kruse, T. S. Kobilka, F. S. Thian, J. M. Mathiesen, R. K. Sunahara, L. Pardo, W. I. Weis, B. K. Kobilka, S. Granier. Crystal structure of the mu-opioid receptor bound to a morphinan antagonist. Nature. 2012;485(7398):321-U170.

32. Suno R., K. T. Kimura, T. Nakane, K. Yamashita, J. Wang, T. Fujiwara, Y. Yamanaka, D. Im, S. Horita, H. Tsujimoto. Crystal structures of human orexin 2 receptor bound to the subtype-selective antagonist EMPA. Structure. 2018;26(1):7-19. e15.

33. Cheng R. K., C. Fiez-Vandal, O. Schlenker, K. Edman, B. Aggeler, D. G. Brown, G. A. Brown, R. M. Cooke, C. E. Dumelin, A. S. Doré. Structural insight into allosteric modulation of protease-activated receptor 2. Nature. 2017;545(7652):112.

34. Choe H.-W., Y. J. Kim, J. H. Park, T. Morizumi, E. F. Pai, N. Krauss, K. P. Hofmann, P. Scheerer, O. P. Ernst. Crystal structure of metarhodopsin II. Nature. 2011;471(7340):651.

35. Li J., P. C. Edwards, M. Burghammer, C. Villa, G. F. Schertler. Structure of bovine rhodopsin in a trigonal crystal form. J Mol Biol. 2004;343(5):1409-1438.

36. Hanson M. A., C. B. Roth, E. Jo, M. T. Griffith, F. L. Scott, G. Reinhart, H. Desale, B. Clemons, S. M. Cahalan, S. C. Schuerer. Crystal structure of a lipid G protein–coupled receptor. Science. 2012;335(6070):851-855.

37. Burg J. S., J. R. Ingram, A. J. Venkatakrishnan, K. M. Jude, A. Dukkipati, E. N. Feinberg, A. Angelini, D. Waghray, R. O. Dror, H. L. Ploegh, K. C. Garcia. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. Science. 2015;347(6226):1113-1117.

38. Liang Y. L., M. Khoshouei, M. Radjainia, Y. Zhang, A. Glukhova, J. Tarrasch, D. M. Thal, S. G. B. Furness, G. Christopoulos, T. Coudrat, R. Danev, W. Baumeister, L. J. Miller, A. Christopoulos, B. K. Kobilka, D. Wootten, G. Skiniotis, P. M. Sexton. Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. Nature. 2017;546(7656):118-+.

39. Song G., D. Yang, Y. Wang, C. de Graaf, Q. Zhou, S. Jiang, K. Liu, X. Cai, A. Dai, G. Lin. Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators. Nature. 2017;546(7657):312.

40. Siu F. Y., M. He, C. De Graaf, G. W. Han, D. Yang, Z. Zhang, C. Zhou, Q. Xu, D. Wacker, J. S. Joseph. Structure of the human glucagon class B G-protein-coupled receptor. Nature. 2013;499(7459):444.

41. Doré A. S., K. Okrasa, J. C. Patel, M. Serrano-Vega, K. Bennett, R. M. Cooke, J. C. Errey, A. Jazayeri, S. Khan, B. Tehan. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. Nature. 2014;511(7511):557.

42. Wang C., H. Wu, T. Evron, E. Vardy, G. W. Han, X.-P. Huang, S. J. Hufeisen, T. J. Mangano, D. J. Urban, V. Katritch. Structural basis for Smoothened receptor modulation and chemoresistance to anticancer drugs. Nat Commun. 2014;5:4355.

43. Song W., H. Y. Yen, C. V. Robinson, M. S. P. Sansom. State-dependent Lipid Interactions with the A2a Receptor Revealed by MD Simulations Using In Vivo-Mimetic Membranes. Structure. 2019;27(2):392-403 e393.

44. Rouviere E., C. Arnarez, L. W. Yang, E. Lyman. Identification of Two New Cholesterol Interaction Sites on the A(2A) Adenosine Receptor. Biophys J. 2017;113(11):2415-2424.

45. Guixa-Gonzalez R., J. L. Albasanz, I. Rodriguez-Espigares, M. Pastor, F. Sanz, M. Marti-Solano, M. Manna, H. Martinez-Seara, P. W. Hildebrand, M. Martin, J. Selent. Membrane cholesterol access into a G-protein-coupled receptor. Nat Commun. 2017;8:14505.

46. Cang X. H., Y. Du, Y. Y. Mao, Y. Y. Wang, H. Y. Yang, H. L. Jiang. Mapping the Functional Binding Sites of Cholesterol in beta(2)-Adrenergic Receptor by Long-Time Molecular Dynamics Simulations. J Phys Chem B. 2013;117(4):1085-1094.

47. Manna M., M. Niemela, J. Tynkkynen, M. Javanainen, W. Kulig, D. J. Muller, T. Rog, I. Vattulainen. Mechanism of allosteric regulation of beta2-adrenergic receptor by cholesterol. eLife. 2016;5.

48. Shan J., G. Khelashvili, S. Mondal, E. L. Mehler, H. Weinstein. Ligand-dependent conformations and dynamics of the serotonin 5-HT(2A) receptor determine its activation and membrane-driven oligomerization properties. PLoS Comput Biol. 2012;8(4):e1002473.

49. Marino K. A., D. Prada-Gracia, D. Provasi, M. Filizola. Impact of Lipid Composition and Receptor Conformation on the Spatio-temporal Organization of mu-Opioid Receptors in a Multi-component Plasma Membrane Model. PLoS Comput Biol. 2016;12(12):e1005240.

Figure S18. Contact heatmap of GPCR - PIP lipids interactions. We show full-sequence heatmaps and corresponding bar graphs of the number of contacts with PIP lipids for each GPCR simulated.

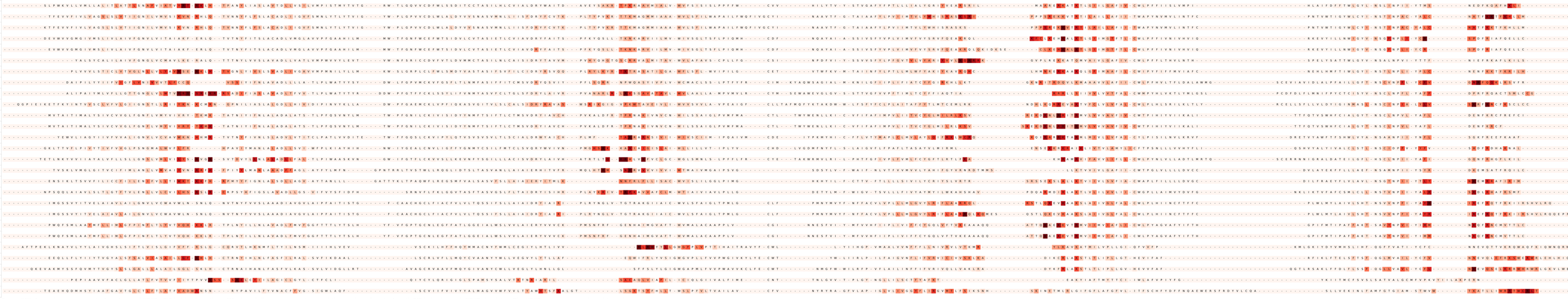


Figure S19. Aligned sequence contact heatmap of GPCR - PIP lipids interactions. We show sequence heatmaps of the number of contacts with PIP lipids for each GPCR simulated. The GPCR sequences are aligned to make comparison easy.

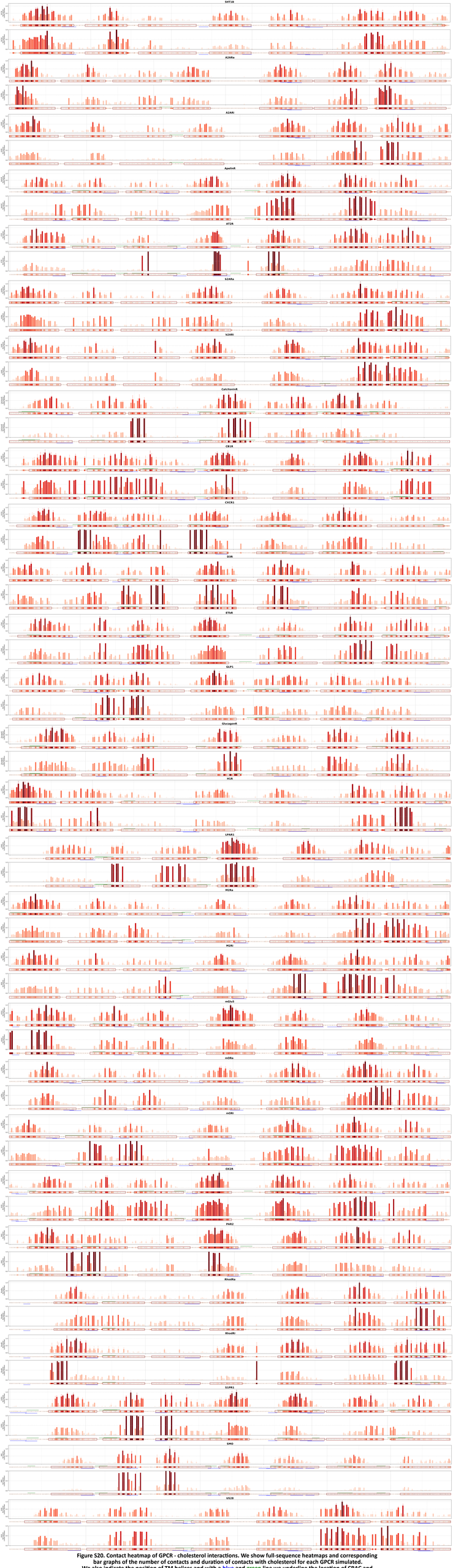


Figure S20. Contact heatmap of GPCR - cholesterol interactions. We show full-sequence heatmaps and corresponding bar graphs of the number of contacts and duration of contacts with cholesterol for each GPCR simulated. We also indicate the position of TM helices and with a blue and green line we underline the location of CRAC and CARC motifs, respectively.

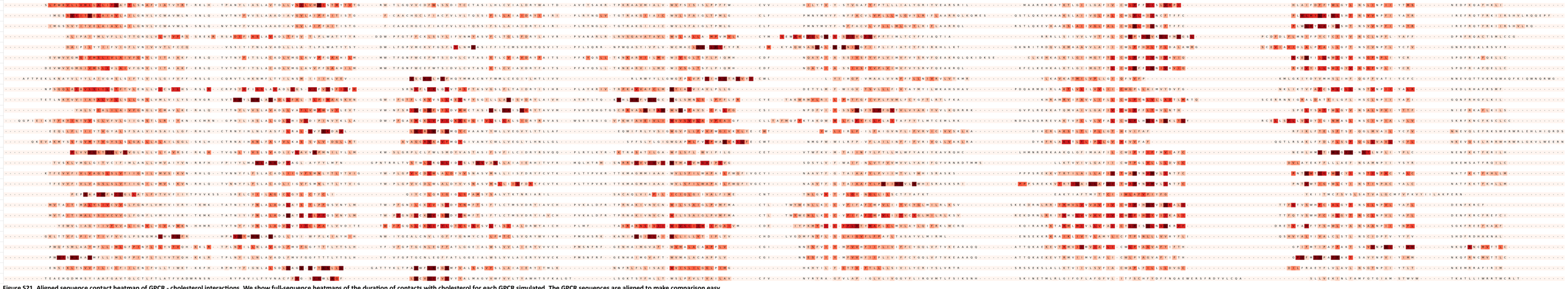


Figure S21. Aligned sequence contact heatmap of GPCR - cholesterol interactions. We show full-sequence heatmaps of the duration of contacts with cholesterol for each GPCR simulated. The GPCR sequences are aligned to make comparison easy.

5HT1B
A2AR1
A2AR1
ApelinR
AT2R
b2AR1
b2ARI
CalcitoninR
CB1R
CXCR1
D3R
ETbR
GLP1
GlucagonR
H1R
LPAR1
M2Ra
M2RI
mGlu5
mOra
mORI
OX2R
PAR2
RhodRa
RhodRi
S1PR1
SMO
US28
Number of Lipids
GM
PC_u
PE_u
PC_l
SM_u
PIP
PA
PE_l
SM_l
PI
PS
LPC
Time (μ)


5HT1B
A2ARa
A2ARi
ApelinR
AT2R
b2ARa
b2ARi
CalcitoninR
CB1R
CXCR1
D3R
ETbR
GLP1
GlucagonR
H1R
LPAR1
M2Ra
M2Ri
mGlu5
mORa
mORi
OX2R
PAR2
RhodRa
RhodRi
S1PR1
SMO
US28
Time (microseconds)