



OPEN

Unusual mode of dimerization of retinitis pigmentosa-associated F220C rhodopsin

George Khelashvili^{1,2}, Anoop Narayana Pillai^{3,11}, Joon Lee^{3,11}, Kalpana Pandey^{3,11}, Alexander M. Payne⁴, Zarek Siegel⁵, Michel A. Cuendet^{1,6,7,8}, Tylor R. Lewis⁹, Vadim Y. Arshavsky⁹, Johannes Broichhagen¹⁰, Joshua Levitz³ & Anant K. Menon³

Mutations in the G protein-coupled receptor (GPCR) rhodopsin are a common cause of autosomal dominant retinitis pigmentosa, a blinding disease. Rhodopsin self-associates in the membrane, and the purified monomeric apo-protein opsin dimerizes in vitro as it transitions from detergent micelles to reconstitute into a lipid bilayer. We previously reported that the retinitis pigmentosa-linked F220C opsin mutant fails to dimerize in vitro, reconstituting as a monomer. Using fluorescence-based assays and molecular dynamics simulations we now report that whereas wild-type and F220C opsin display distinct dimerization propensities in vitro as previously shown, they both dimerize in the plasma membrane of HEK293 cells. Unexpectedly, molecular dynamics simulations show that F220C opsin forms an energetically favored dimer in the membrane when compared with the wild-type protein. The conformation of the F220C dimer is unique, with transmembrane helices 5 and 6 splayed apart, promoting widening of the intracellular vestibule of each protomer and influx of water into the protein interior. FRET experiments with SNAP-tagged wild-type and F220C opsin expressed in HEK293 cells are consistent with this conformational difference. We speculate that the unusual mode of dimerization of F220C opsin in the membrane may have physiological consequences.

The visual pigment rhodopsin is both a G protein-coupled receptor (GPCR) and a critical structural component of the outer segment of photoreceptor cells¹. Its unexpected phospholipid scramblase activity, evinced on reconstitution of purified protein into lipid vesicles, accounts for the lipid scrambling observed in isolated photoreceptor disc membranes^{2–4}. Rhodopsin can activate the visual signaling cascade as a monomer, but, like some Class A GPCRs^{5,6}, it self-associates to form dimers and higher order multimers. The quaternary structure of rhodopsin has been proposed to be important for its trafficking to the outer segment, structural contributions to disc formation, and phototransduction properties⁶. Rhodopsin dimers have been visualized in situ using atomic force microscopy of photoreceptor disc membranes adsorbed on mica⁷, and cryo-electron tomography using retinas cryofixed by high-pressure freezing⁸. Consistent with these results, chemical crosslinking studies of native disc membranes revealed a symmetric dimer interface mediated by transmembrane helix 1 (TM1) contacts⁹, and site-directed cysteine mutagenesis studies implicated transmembrane helices 4 and 5 (TM4 and TM5) in mediating dimerization of the apo-protein opsin expressed in COS1 cells¹⁰. The involvement of these TM helices in mediating dimer formation is supported by a cryo-electron microscopy study of rhodopsin reconstituted into nanodiscs¹¹, as well as a report showing that synthetic TM peptides can disrupt rhodopsin dimers in both discs and HEK cells¹².

¹Department of Physiology and Biophysics, Weill Cornell Medical College, New York, NY 10065, USA. ²Institute of Computational Biomedicine, Weill Cornell Medical College, New York, NY 10065, USA. ³Department of Biochemistry, Weill Cornell Medical College, New York, NY 10065, USA. ⁴Tri-Institutional PhD Program in Chemical Biology, Weill Cornell Medical College, New York, NY 10065, USA. ⁵Neurosciences Graduate Program, University of California San Diego, La Jolla, CA 92093, USA. ⁶Ludwig Institute for Cancer Research, University of Lausanne, Lausanne, Switzerland. ⁷Department of Oncology, University Hospital of Lausanne, 1009 Lausanne, Switzerland. ⁸Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. ⁹Department of Ophthalmology, Duke University Medical Center, Durham, NC 27710, USA. ¹⁰Leibniz-Forschungsinstitut Für Molekulare Pharmakologie, Robert-Rössle-Str. 10, 13125 Berlin, Germany. ¹¹These authors contributed equally: Anoop Narayana Pillai, Joon Lee and Kalpana Pandey. ✉email: gek2009@med.cornell.edu; jtl2003@med.cornell.edu; akm2003@med.cornell.edu

Mutations in rhodopsin are the most common cause of autosomal dominant retinitis pigmentosa (RP), a blinding disease that afflicts more than 1.5 million individuals world-wide^{13–15}. We discovered that F45L, V209M and F220C rhodopsin mutants that were identified in patients with RP appear to function normally as visual pigments and phospholipid scramblases but, unlike wild type rhodopsin, they fail to dimerize *in vitro* during detergent removal *en route* to reconstitution into lipid vesicles¹⁶. We proposed that their inability to dimerize might explain their association with retinal disease. Surprisingly, biophysical studies of the V209M and F220C mutants expressed in COS cells indicate that they self-associate similarly to wild-type rhodopsin while exhibiting a minor trafficking defect¹⁷. Furthermore, knockin mice bearing the F45L mutation are aphenotypic, while those with the F220C mutation display a small abnormality in outer segment morphology¹⁸. Given these disparate data, we decided to analyze the F220C mutant in more detail using a multi-pronged approach to characterize the protein. We used *in vitro* and cell-based fluorescence assays to verify and extend previous data, and computational approaches based on molecular dynamics (MD) simulations of the protein in detergent micelles and in the membrane. In all our experiments we compared the behavior of F220C opsin with that of the wild-type (WT) protein. As the F220C mutation is located on TM5, we also considered more generally the role of this helix in rhodopsin dimerization.

We now report that both WT and F220C opsin dimerize in the plasma membrane of HEK293 cells, but are monomeric when purified after solubilization of the cells in the detergent dodecyl- β -maltoside (DDM). The latter observation is consistent with our MD simulations showing that opsin dimerization is energetically unfavored in a DDM-micellar environment. We developed a protocol to mimic early steps of *in vitro* membrane reconstitution by treating DDM-solubilized opsins with detergent-adsorbing resin in the absence of phospholipids. With this protocol we find that DDM-solubilized WT opsin monomers self-associate on partial removal of detergent, whereas F220C opsin remains monomeric. Thus, consistent with previous results^{16,17}, the two proteins exhibit different behaviors in a micellar environment, whereas both form dimers when they are located in a membrane bilayer. These findings support the idea that WT and F220C opsins have distinct dimerization properties, motivating a deeper analysis in the membrane context.

MD simulations showed that F220C opsin forms an energetically favored dimer in the membrane when compared with the WT protein. The conformation of the F220C dimer is unique, with the TM5/TM6 helices tilted to form a splayed, criss-crossed structure. The difference in conformation of WT and F220C dimers in the membrane is supported by experiments showing that the two homo-dimers exhibit different extents of Förster resonance energy transfer (FRET) in the membrane. The splayed conformation of F220C opsin dimers promotes widening of the intracellular vestibule of each monomer in the dimer enabling water to enter the protein interior. Further analyses reveal that this unique dimerization pose is a feature of the F220C mutation and is not observed in simulations in which other positions along the TM5 interface are mutated. We speculate that the unusual mode of dimerization of F220C opsin in the membrane may underscore its physiological consequences.

Results and discussion

Dimerization of WT opsin, but not F220C opsin, in a micellar system. We previously proposed that whereas both WT and F220C opsins are monomers when purified from DDM-solubilized cells, gradual withdrawal of detergent causes WT opsin but not F220C opsin to dimerize^{3,16}. This proposal was supported by experiments showing that immobilized WT opsin could pull-down soluble WT opsin as DDM was reduced from 1.96 to 0.2 mM, just above its critical micellar concentration (cmc) (0.15 mM), whereas the same did not hold for the F220C construct¹⁶. Here we deploy fluorescence measurements to test key aspects of this proposal.

To establish that WT and F220C opsins are monomers when purified, we used Single Molecule Pulldown (SiMPull)^{19–21}. We expressed HA-SNAP-tagged WT and F220C opsins in HEK293T cells, and labeled the cell surface population with the membrane-impermeant SNAP dye LD555 before solubilizing the cells in 1% (w/v) DDM to extract the labeled proteins. The samples were subsequently transitioned to a lower detergent concentration (0.1% (w/v) DDM) and immobilized on a polyethylene glycol (PEG)-passivated coverslip via anti-HA antibodies. Individual fluorescent spots were visualized using total internal reflection (TIRF) microscopy (Fig. 1A,C). For both opsin constructs, most spots were bleached within 30 s of the onset of laser illumination (representative traces shown in Fig. 1A,C), and bleaching analysis revealed that ~90% of spots bleached in a single step (Fig. 1B,D). These results indicate that both WT and F220C opsins are monomers in 0.1% (w/v) DDM.

We next used Förster resonance energy transfer (FRET) to test the ability of the proteins to self-associate as a function of detergent removal in the absence of lipid vesicles (Fig. 1E). For this, we fluorescently labeled purified SNAP-tagged opsins with either SNAP-Surface 549 (donor (D)) or SNAP-Surface 649 (acceptor (A)) fluorophores. Equimolar mixtures of D and A-labeled WT opsin (in 0.1% (w/v) DDM (1.96 mM)) were mock-treated, or treated with different amounts of Bio-Beads (BB) to reduce detergent levels. F220C opsin samples were analyzed in parallel. Fluorescence emission spectra (Fig. 1F) indicate FRET in the WT opsin mixture treated with 5.7 mg Bio-Beads (WT + BB), but not in the F220C opsin sample (F220C + BB). No FRET was seen in mock-treated samples (dotted lines, no BB (Fig. 1F)). Based on our previous calibration of BB-mediated detergent removal³, the amount of DDM remaining after treatment of the sample with 4–6 mg BB remains above cmc. Quantification of FRET as a function of BioBead treatment, i.e. amount of detergent removed, shows that WT opsin, but not F220C opsin, dimerizes/multimerizes in samples treated with > 4 mg Bio-Beads (Fig. 1G). These results support the mechanism presented above and shown schematically in Fig. 1E, i.e. that the F220C mutation prevents opsin from dimerizing as detergent is withdrawn.

Computational analysis of the behavior of WT and F220C opsin in DDM micelles. Hypothesizing that the different behavior of WT and F220C opsin in detergent would be reflected in differences in the organization of the detergent micelle surrounding these proteins, we carried out MD simulations of opsin mono-

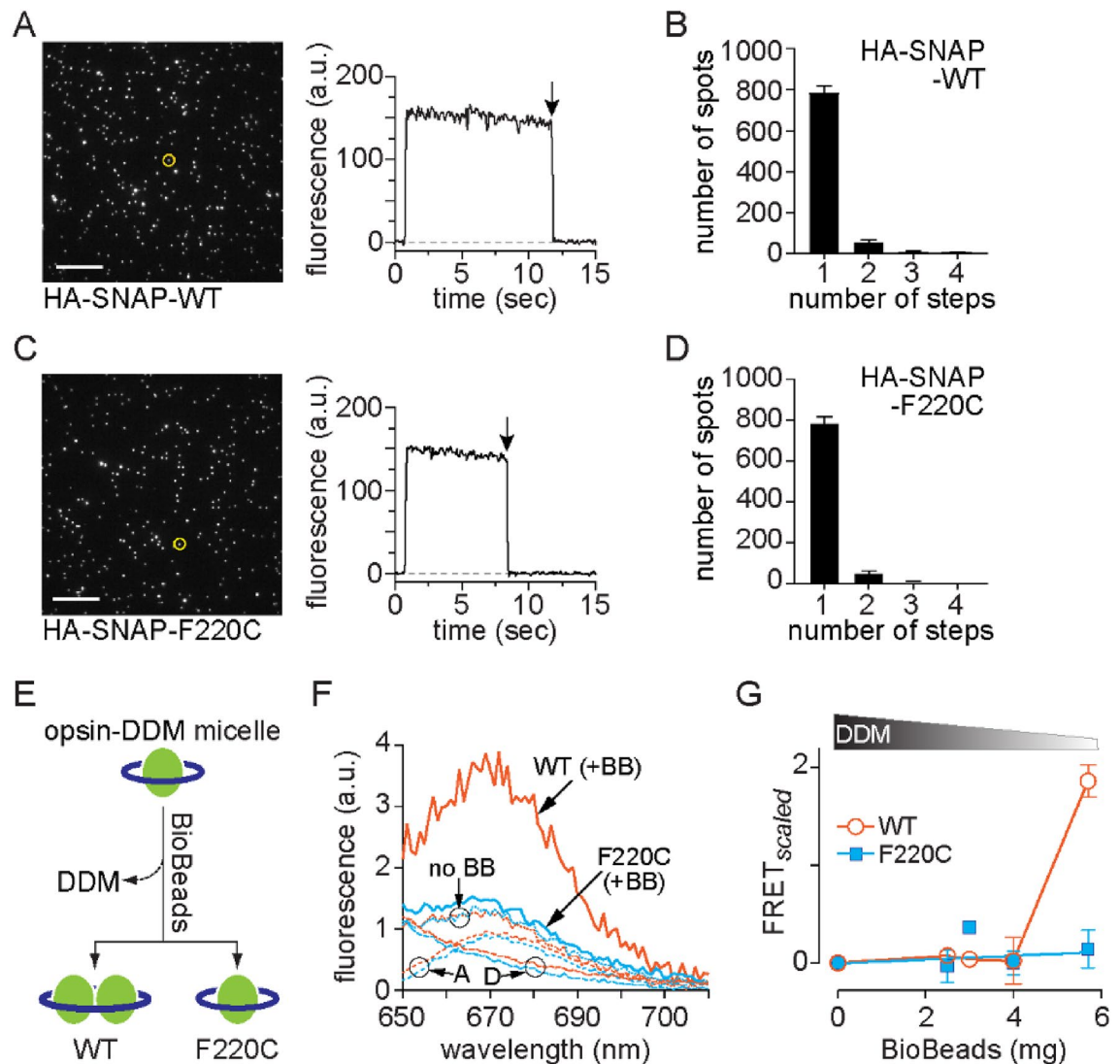
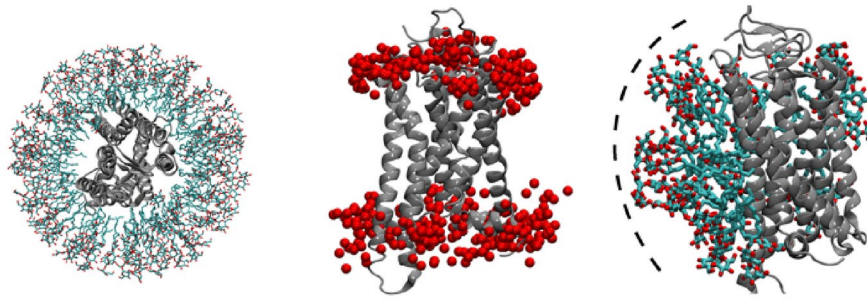


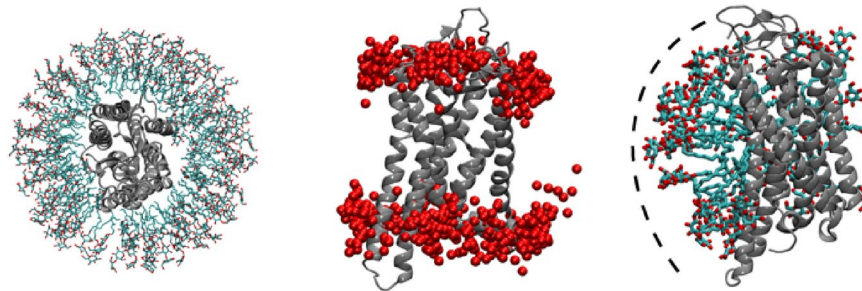
Figure 1. Dimerization of WT opsin, but not F220C opsin, in a micellar system. (A–D) SiMPull experiments with HA-SNAP-WT opsin (A,B) and HA-SNAP-F220C opsin (C,D). Opsin constructs were expressed in HEK293T cells and labeled using a membrane-impermeant SNAP fluorophore (LD555). The cells were solubilized in DDM and the proteins were captured on glass slides coated with anti-HA antibodies for visualization of fluorescence by TIRF microscopy (A,C). Bleaching of individual spots was recorded; traces for representative spots (circled in the fluorescence images) are shown, with arrows to indicate the bleaching step, and the number of spots that bleached in 1, 2, 3 or 4 steps was quantified (B,D). We note that as fluorophore bleaching is a stochastic process, the time taken for an individual spot to bleach is arbitrary. (E) Schematic of the effect of detergent removal on the quaternary structure of WT and F220C opsins. (F) Emission spectra ($\lambda_{\text{ex}} = 555 \text{ nm}$) for donor (D)-acceptor (A) pairs of WT (orange) or F220C opsin (blue) treated with 5.7 mg Bio-Beads (+BB, solid lines) or mock-treated (no BB, dotted lines). Spectra of mock treated samples of acceptor alone and donor alone ('A' and 'D', dashed lines) are also shown. Sensitized emission of the acceptor ($\lambda_{\text{ex}} = 555 \text{ nm}$, $\lambda_{\text{em}} = 670 \text{ nm}$) indicates FRET as seen in the BioBead-treated WT sample (+BB). (G) FRET was quantified after extracting increasing amounts of DDM (shown schematically above the graph) by treating samples with different amounts of Bio-Beads. A FRET measure ($\text{FRET}_{\text{scaled}}$) was obtained by calculating the ratio of the emission intensity at 670 nm for D-A mixtures relative to the acceptor-only sample, followed by an offset correction corresponding to mock-treated samples (no BB, A). Data are mean \pm SEM ($n \geq 3$) and connecting lines are drawn to guide the eye.

mers and dimers in DDM micelles. We used a detergent : protein ratio of 150:1, based on a previous estimate of the size of DDM micelles^{22,23}. As shown in Fig. 2 (top two rows), the organization of DDM molecules around monomeric WT and F220C opsins is as expected, with hydrophobic tails directed towards the protein, shielding the hydrophobic core of the protein from unfavorable exposure to the solvent. However, the arrangement of detergent around the dimeric constructs is strikingly different. Thus, as illustrated in Fig. 2 (bottom two rows), the dimer interface of both WT and F220C systems is exposed to DDM headgroups, as the detergent molecules

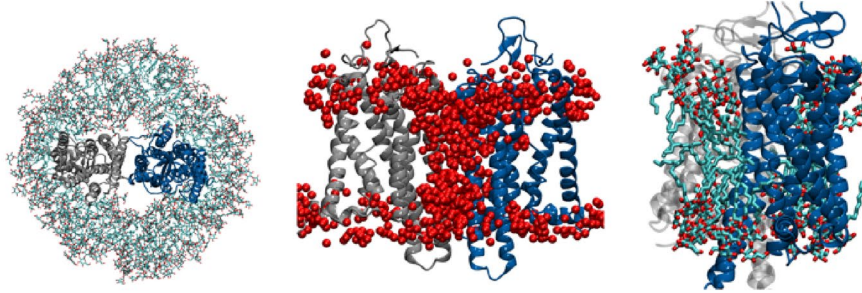
WT monomer



F220C monomer



WT dimer



F220C dimer

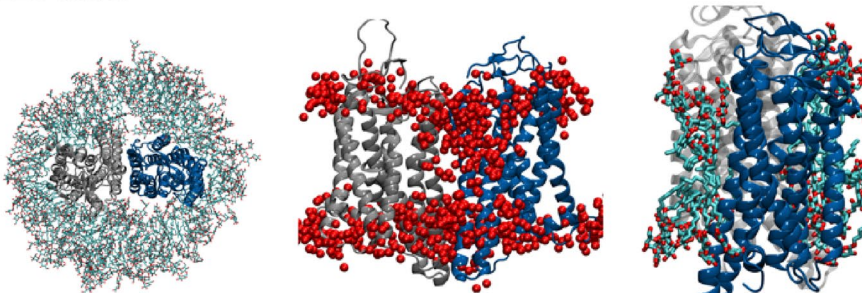


Figure 2. Structure of WT and F220C opsin monomers and dimers revealed by MD simulation. WT and F220C opsin monomers (top two rows) and dimers (bottom two rows) in DDM micelles. The snapshots on the left side of all rows show initial conformations of the respective systems: detergent molecules are shown as sticks, the protein is shown with its extracellular side on top (the monomer protein (top two rows) is in grey, and the two protomers of the dimer constructs (bottom two rows) are in grey and blue). The middle panels show a side-view overlaid with locations of the O5 oxygen atom of the DDM headgroup (red spheres) that were within 4 Å of the protein at least in one frame of the respective trajectory. The right panels show close up views of the protein constructs highlighting arrangement of the detergent molecules around TM5 and TM6 helices. Note, that DDM molecules organize around monomeric WT and F220C opsins with their hydrophobic tails directed towards the protein (the dashed arcs trace the curved surface defined by the detergent headgroups). However, in the dimeric constructs they align vertically, akin to lipids.

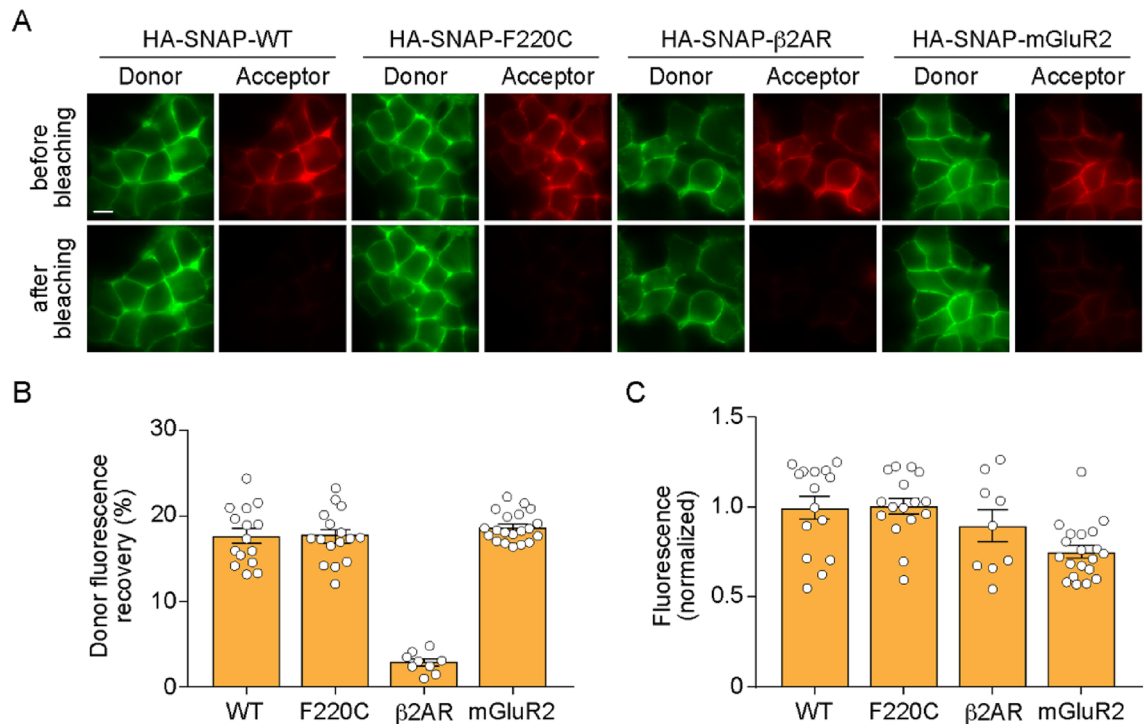


Figure 3. Dimerization of WT and F220C opsins in the plasma membrane of HEK293T cells. **(A)** Fluorescence images of HEK293T cells expressing N-terminally HA-SNAP-tagged WT opsin, F220C opsin, β2-adrenergic receptor (β2AR) or metabotropic glutamate receptor 2 (mGluR2), labeled with SNAP donor (SBG-JF₅₄₉) and acceptor (SBG-JF₆₄₆) dyes. The top and bottom rows show images of cells taken before and after bleaching of the acceptor by high intensity 640 nm illumination. Scale bar, 10 μm. **(B)** Summary of donor fluorescence recovery upon acceptor bleaching. **(C)** Summary of acceptor fluorescence as a reporter of comparable surface expression for all constructs.

align vertically, akin to lipids. Such an arrangement results in the exposure of polar DDM headgroups to hydrophobic residues at the dimer interface leading to an energetically costly hydrophobic mismatch between the protein dimer and the micelle. We calculated the residual hydrophobic mismatch (RHM) energy for each protein sidechain. The concept of RHM energy and its relation to protein–protein interactions in lipid membranes has been well established (see Refs.^{24,25}). It has been shown that the cost of RHM can be mitigated by modes of protein association in the bilayer that would cover the sites of RHM and eliminate the costly interaction of the protein with the membrane^{26,27}. Here, following the same concept, we analyzed RHM between the protein residues and DDM molecules in order to quantify the contribution of hydrophobic mismatch to the dimerization energy.

To calculate RHM, we used a computational approach based on quantifying the surface area involved in unfavorable hydrophobic-hydrophilic interactions^{24–26}. Figure S2 shows residues on TM5 and TM6 helices with high RHM values ($> 1 k_B T$) in simulations of the WT and F220C opsin monomers and dimers in micelles (the RHM values for the other parts of the protein were similar for all the systems). Overall RHM energy (F) stemming from these residues in the WT and F220C monomer systems (F_{MONO}) is $6.4 \pm 0.6 k_B T$ and $6.2 \pm 0.4 k_B T$, respectively. For the dimer systems, the RHM values (F_{DIMER}) are significantly higher, $32.8 \pm 3.5 k_B T$ and $29.7 \pm 3.4 k_B T$, for the WT and F220C dimers, respectively. Defining the RHM contribution to the dimerization energy as $F_{\text{DIMER}} - 2 * F_{\text{MONO}}$, we find that the difference in RHM for the residues in TM5 and TM6 between the WT systems is $20 k_B T$ and between the F220C systems is $17 k_B T$, suggesting that dimerization in a ‘high DDM’ environment is strongly unfavorable for both the WT and F220C systems. While our simulations consider only one of several possible modes of opsin dimerization (Table S1), these results are consistent with the experimental findings that both WT and F220C opsin proteins are monomeric at high detergent concentrations. Importantly, our results suggest that RHM may be a key factor in determining energy cost of dimerization during the in vitro reconstitution process as the detergent is gradually withdrawn from the system. Perhaps during detergent withdrawal, the energetic cost is significantly diminished for WT opsin, enabling it to self-associate. A detailed analysis of this process will be the subject of future work.

Dimerization of both WT and F220C opsin in the plasma membrane revealed by ensemble FRET measurements. We next considered the self-association behavior of WT and F220C opsin in the membrane. We expressed N-terminal SNAP-tagged variants of WT and F220C opsin in HEK293T cells (Fig. 3A); as controls we used N-terminal SNAP-tagged variants of β2-adrenergic receptor (β2AR) which exhibits largely monomeric behavior^{28,29}, and metabotropic glutamate receptor 2 (mGluR2) which is a constitutive dimer in live cells³⁰. The proteins were expressed for 24 or 36 h, before labeling the cell-surface population with

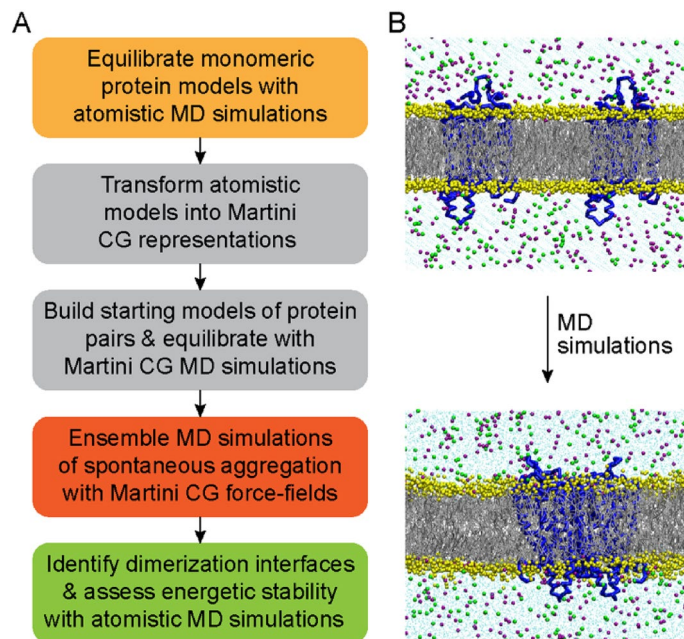


Figure 4. Set-up for computational analysis of opsin dimerization in a lipid membrane. **(A)** Multiscale computational protocol to investigate dimerization of WT and F220C opsins in a lipid membrane. **(B)** Starting (top) and final (bottom) snapshots of the WT opsin system after CG MD simulations. Initially well-separated monomers (center-of-mass distance of 50 Å) spontaneously aggregate on the 120 μ s timescale of the CG simulations. In the snapshots, opsins are shown in blue cartoon, lipid phosphate beads in yellow spheres, and lipid tails in grey lines. Also depicted are salt ions in solution (green and purple spheres) and water beads (cyan dots).

cell-impermeant red and far-red SNAP dyes (SBG-JF₅₄₉ and SBG-JF₆₄₆)³¹. A mixture of SNAP fluorophores was used to generate donor and acceptor labeled proteins in a 1:3 ratio (Fig. 3A). Quantification of acceptor fluorescence from several fields of cells revealed that all constructs were well-expressed over a several-fold range of fluorescence (Figure S1). FRET, quantified by measuring the increase in donor fluorescence upon bleaching of the acceptor with brief exposure to high intensity laser illumination, was observed at a similar level for WT and F220C opsins, and comparable to that observed for mGluR2 (Figure S1). In contrast β 2AR showed little to no FRET, with a modest linear increase in donor recovery as expression levels increased (Figure S1) indicating that the FRET readout in our experiments is not due to the random proximity of donor- and acceptor-labeled proteins. A summary of the FRET data is shown in Fig. 3B, corresponding to cells within the same range of protein expression (Fig. 3C; corresponding to \sim 2000–3500 a.u. in Figure S1). Importantly, the FRET level for WT and F220C opsin, as well as mGluR2, was saturable at high expression levels (Figure S1). These results indicate that both opsins dimerize/multimerize in the plasma membrane. However, as presented later, it is not possible to deduce the precise extent of dimerization of WT versus F220C opsin from these experiments as conformational differences between the dimers could produce different extents of FRET. Nevertheless, our results clearly indicate that the F220C mutant dimerizes/multimerizes in the plasma membrane, consistent with the result reported by Mallory et al.¹⁷.

Workflow of computational experiments to investigate the behavior of opsin in the membrane.

Thus far, we have shown that WT and F220C opsin differ dramatically in their dimerization propensity as they reconstitute into membranes from a micellar environment¹⁶ (Fig. 1E–G), yet self-associate to comparable levels in a membrane environment based on FRET measurements (Fig. 3). Because of the distinct behaviors of WT and F220C opsins in micelles, we considered the possibility that the dimers formed by these proteins in the membrane may be structurally different. To test this hypothesis, we used computational experiments in which we applied a multi-scale approach in the spirit of the DAFT (Docking Assay For Transmembrane components) methodology^{32,33}. In these experiments, 35 pairs of monomeric proteins (WT and F220C) are allowed to aggregate spontaneously in a lipid membrane from initially well-separated states. The workflow that we used for each protein system is depicted in Fig. 4A and involves the following steps: (1) Monomer protein models (WT and F220C) are equilibrated with atomistic MD simulations; (2) The atomistic structures are transformed into Martini coarse-grained (CG) representations; (3) Two copies of the CG protein are embedded into a lipid bilayer in such a way that the distance between their center-of-mass positions is \sim 50 Å (see Fig. 4B) and the system is briefly equilibrated with Martini force-fields. For this stage, we considered seven starting conformations differing in mutual orientations of the two monomers in order to randomize the initial conditions of the system; (4) Extensive ensemble CG MD simulations of spontaneous aggregation of the protein pairs are

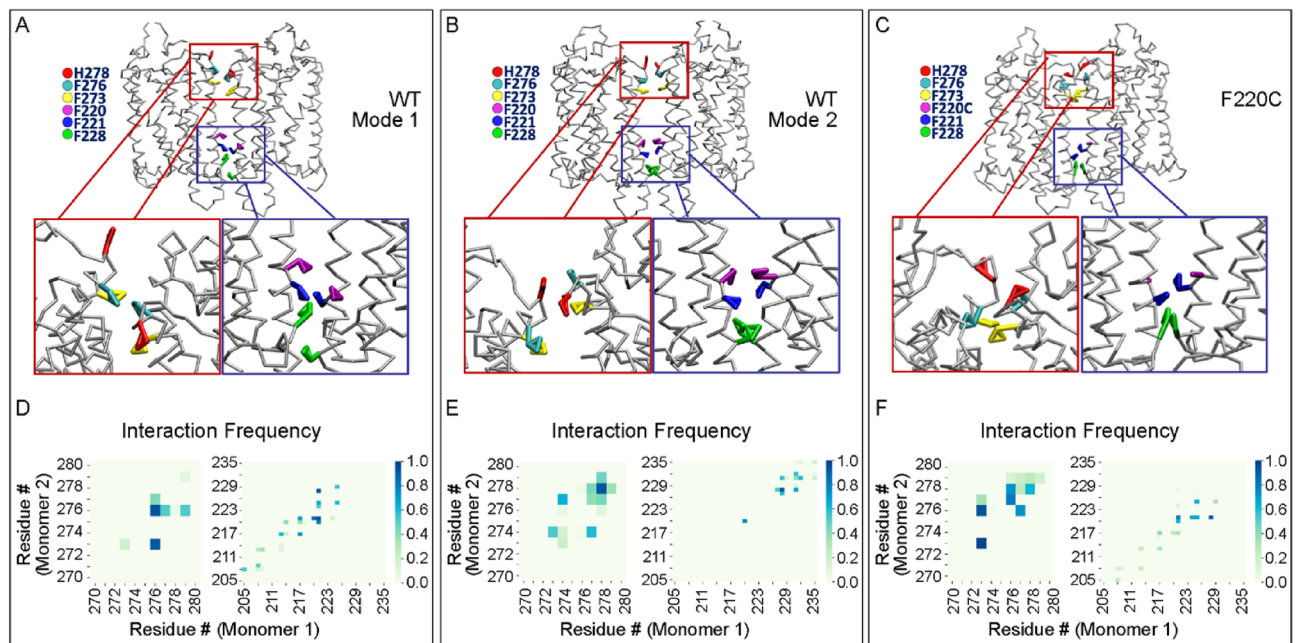


Figure 5. Modes of dimerization of the WT and F220C opsin. (A–C) Structural snapshots corresponding to the modes of dimerization of the WT (Modes 1 and 2 shown in A,B respectively) and F220C (C) opsin constructs from the CG MD simulations. Insets in red and blue rectangles show magnified views of interactions between pairs of TM6 and TM5 residues, respectively. Key residues participating in the interactions (H278^{6,61}, F276^{6,59}, F273^{6,56} in TM6 and F220^{5,55}, F221^{5,56}, F228^{5,63} in TM5) are depicted in different colors and are labeled. (D–F) Pairwise contact frequencies between residues on TM5 (right columns) and TM6 (left columns) in the dimerization modes shown in panels (A–C) respectively.

conducted in which every initial condition is simulated in 5 independent replicates, 120 μ s each (\sim 4.2 ms cumulative time in 35 replicates per construct); (5) The CG simulation trajectories are analyzed to identify preferred dimerization interfaces (as judged by monitoring pair-wise distances between TM helices); (6) The selected dimer models are transformed into all-atom representations, in order to assess their stability and to quantify the energetics of dimerization with refined atomistic MD simulations. In the following, findings from these steps are presented in detail.

WT and F220C opsin monomers spontaneously aggregate into dimers during CG MD simulations. On the extensive timescales of the CG MD simulations, we observed many instances of spontaneous aggregation of both WT and F220C opsin monomers into dimers. Indeed, as shown in Figures S3 and S4, dimerization was observed in 33 out of 35 trajectories of the WT system, and in 30 out of 35 trajectories for the F220C construct. Once formed, the dimer complexes remained stable for the remainder of the simulations. More detailed analysis of inter-protein interactions within the formed dimers revealed various modes of association involving different helices, and both symmetric and asymmetric dimer interfaces (Tables S1 and S2). Given that the F220C mutation resides in TM5, we set out to explore the impact of this mutation on modes of rhodopsin dimerization centered around the TM5 helix, focusing on those trajectories that resulted in symmetric interfaces mediated by TM5–TM5 interactions (6 trajectories for WT system and 7—for the F220C, see Tables S1 and S2). By investigating this specific interface, our studies were designed to probe the mechanistic hypothesis that the TM5–TM5 dimer interface can be modulated by the F220C mutation.

CG MD simulations identify two modes of dimerization along the TM5 helix in WT opsin. Figure S5 shows the frequency of pair-wise interactions for TM5 residues in the six CG trajectories of WT opsin in which dimerization was achieved through TM5–TM5 interactions. The data reveal two distinct modes of dimerization. One mode of dimerization (Mode 1), observed in trajectories 13 and 21 (Fig. 5A and Figure S5), features extensive interactions along the intracellular end of TM5, involving residue stretch 204^{5,39}–228^{5,63} (superscripts refer to residue labeling according to the Ballesteros-Weinstein generic residue numbering scheme for GPCRs³⁴). The most notable interactions stabilizing this mode are found between pairs of F221^{5,56} and F228^{5,63} residues from opposing monomers. This dimerization mode includes additional stabilizing interactions along the extracellular part of TM6, as seen in Fig. 5A (also Fig. 5D, left column). In the second mode of dimerization (Mode 2), observed in trajectories 0, 27, and 31 (Fig. 5B and Figure S5), dimers are mostly stabilized by pairwise interactions between a shorter stretch of residues at the intracellular end of TM5, 228^{5,63}–232^{5,67}, as well as interactions between residues 220^{5,55} from the two protomers of the dimer. This mode of association is further stabilized by contacts along the extracellular side of TM6 helix (region between 273^{6,56} and 278^{6,61}, see Fig. 5A

and also Fig. 5E, *left column*). Lastly, trajectory 7 (Figure S5) shows relatively weak interfacial interactions involving 208^{5.43}–212^{5.47} residue stretch in TM5.

F220C opsin shows a single mode of dimerization along the TM5 helix which is similar to one of the modes identified for the WT protein. Figure S6 shows the frequency of pair-wise interactions between TM5 residues for the six CG trajectories of the F220C opsin in which dimerization was achieved via TM5-TM5 interactions. The data reveal that all six trajectories evolve to a common dimerization interface in which inter-monomer contacts are stabilized through the 204^{5.39}–228^{5.63} residue stretch in TM5. This dimerization interface is similar to that seen in Mode 1 dimerization in the WT system (compare Fig. 5A,C, also Fig. 5D,F).

All-atom MD simulations structurally refine dimerization interfaces from the CG studies and reveal that F220C opsin forms energetically more stable dimers compared to the WT protein. The analysis of the CG MD simulations described above revealed the overall structural characteristics of the WT and F220C opsin dimers. In order to refine these structural features at the atomistic level and quantitatively compare the stability of the dimerization interfaces seen in the different constructs, we performed all-atom MD simulations. To this end, we selected representative structures from the CG simulations corresponding to the dimer arrangements described above (i.e. Mode 1 and 2 dimers for WT opsin, and a single dimer for F220C opsin, Fig. 5A) and converted these CG dimers into all-atom representations. The all-atom structures were then subjected to ensemble MD simulations, in which the three dimer models were each run in 6 independent replicates (900 ns cumulative time per dimer construct). The protein constructs were structurally stable in these atomistic simulations as evinced by relatively small values of root-mean-square deviation (RMSD) of backbone atoms (Figure S7). Analysis of the frequency of pair-wise interactions of the all-atom simulations showed largely the same trend as the CG simulations. Namely, as shown in Figure S8, both the Mode 1 WT dimer and the F220C dimer were stabilized by interactions in the 204^{5.39}–228^{5.63} segment of TM5, whereas the Mode 2 WT dimer was maintained by interactions in the 228^{5.63}–232^{5.67} region of the TM5 helix (although interactions between F220C residues on the two protomers were no longer observed in the all-atom simulations).

As the all-atom MD simulations preserved overall residue contacts along the dimer interface, we analyzed these trajectories to quantify dimerization energies of the different constructs. The analysis revealed that F220C opsin formed energetically more stable dimers than the WT protein. Thus, we performed Molecular Mechanics—Generalized Born Surface Area (MM-GBSA) analysis of the all-atom trajectories to calculate the strength of dimerization for each system, taking into account the energetic cost of de-solvating the binding interface from its membrane and aqueous environment. In Fig. 6, MM-GBSA binding energies are plotted separately for each all-atom trajectory. The data reveal significantly stronger binding of the F220C dimer compared to the WT dimers (compare WT1 and WT2 with F220C, see also Table S3). Analysis of sidechain and backbone atom contributions to the overall binding energetics (Figure S9) revealed that both contribute to the stability of the F220C dimers, suggesting that the energy difference between the WT and F220C dimers is the result not only of local sidechain rearrangements but also larger-scale conformational changes. To identify these changes, we next analyzed the structural differences in the WT and F220C dimers. In particular, we investigated the mutual arrangement of the two protomers in the dimers.

Two protomers of the F220C opsin dimers orient in an unusually splayed conformation promoting widening of the intracellular vestibule and concomitant increase in hydration of the protein interior. Our analysis revealed that the F220C opsin dimer has a propensity to adopt conformations in which the TM5/TM6 helices from the two protomers are strongly tilted with respect to each other (Fig. 7A,B). To quantify this mode of interaction, we defined the splay angle (α) between vectors representing the orientation of the TM5 helix in each protomer (each vector connects C $_{\alpha}$ atoms of residues 200 and 234). We constructed histograms of α values sampled along the trajectories. As shown in Fig. 7C, the WT (Mode 1 and Mode 2) and F220C dimers sample conformations characterized by different splay angles. Whereas the dimers of WT opsin mostly sample lower splay angles ($\alpha < 40^{\circ}$), F220C dimers sample conformations with larger α angles, with the splay angle reaching values around 50° and even higher. The time-evolution of the splay angle in the all-atom trajectories (Figure S10, see WT1, WT2, and F220C) revealed that the higher value for the splay angle in the F220C dimer was already present in the initial stages of the trajectories, suggesting that the splayed conformation developed dynamically in the CG simulations and remained stable on the timescales of the all-atom trajectories. To verify this, we calculated the α angle in CG trajectories of the F220C system and compared its behavior before and after the dimer formation event in these simulations. As shown in Figure S11A–B, α samples a wide range of values when the two opsin proteins are diffusing as monomers. But once the dimer is formed, α becomes confined to a narrow range of large values (Figure S11C), suggesting that the splaying in F220C is due to dimerization.

The consequences of the observed criss-cross orientation of the monomers in the F220C dimer are seen in the widening of the area between helices TM5/TM6 and TM7, and concomitant increase in the hydration of the protein interior. Indeed, as shown in Fig. 7D, the distribution of the number of water molecules in the protein interior is shifted towards larger values in the F220C system compared to WT (Modes 1 and 2). At the same time, a small fraction of F220C dimers sample conformations in which the intracellular ends of TM6 and TM7 helices are separated by large distances, 17–20 Å (see Fig. 7E, also the snapshot in Fig. 7F). The WT dimer is seen to mostly sample shorter TM6–TM7 distances, < 15 Å. In this respect, it is important to note that the same distance measures ~ 15 Å in the opsin model from PDBID 4J4U (GaCT peptide-stabilized structure) and ~ 11 Å in the 1U19 model (retinal-bound, inactive form of the protein). Thus, our results suggest that the F220C mutation

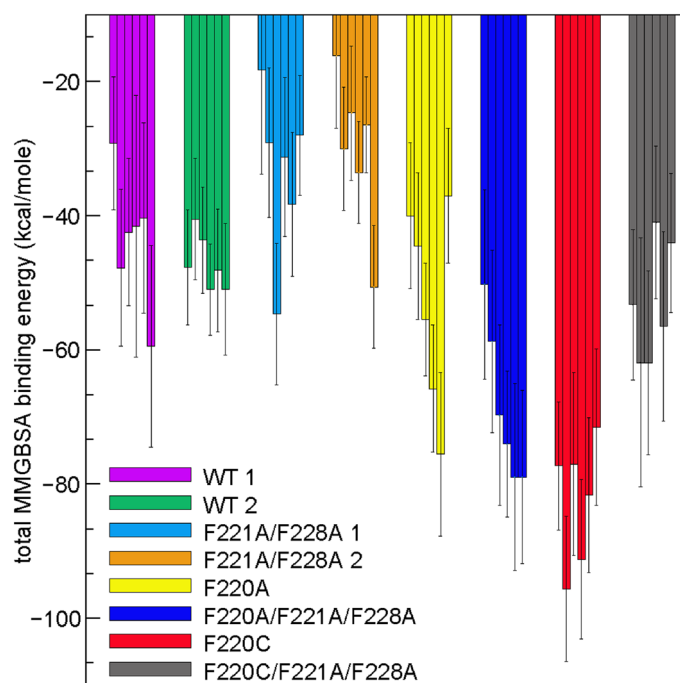


Figure 6. Binding energies of dimerization, quantified with the MMGBSA approach, for different opsin constructs. The MMGBSA binding energies are plotted for the following systems: two WT constructs (Mode 1 and Mode 2 WT), F220A and F220C single mutants, the two F221A/F228A double mutants (the mutations introduced into Mode 1 and Mode 2 WT models), as well as for F220A/F221A/F228A and F220C/F221A/F228A triple mutants (F221A and F228A mutations introduced into the F220A and F220C models, see Methods for more details). The MMGBSA calculations were performed as described in Methods. Data for six independent replicates per construct are shown separately. For each system, 25–150 ns time-intervals of the trajectories were considered for this analysis. The error bars were estimated using the approach based on determination of the effective sample size for autocorrelated data described by Chodera and co-workers⁶⁴. See also Table S3.

promotes widening of the intracellular vestibule in the corresponding opsin dimers, potentially enhancing the ability of the receptor to couple to effector proteins, such as transducin.

The splayed conformation in the F220C dimer is a direct consequence of the mutation which alters residue interactions along the dimer interface.

To identify the molecular mechanism underlying the observed splaying of the F220C dimer, we analyzed interfacial interactions in the WT and F220C dimer systems. The analysis revealed a subtle difference in the pattern of interfacial pair-wise residue interactions in the two dimer systems. Thus, in the WT dimer (Mode 1), residue F221^{5,56} in one protomer is engaged in aromatic interactions with either F220^{5,55} or F228^{5,63} in the opposite protomer (Fig. 8). In the F220C dimer, on the other hand, F221^{5,56} interacts mostly with just F228^{5,63} (Fig. 8C) since position 220 is no longer occupied by a hydrophobic aromatic amino acid. The major role of the F221^{5,56} residue in stabilizing F220C dimers is exemplified by the larger contribution of this residue to the overall binding energy in the F220C system compared to the WT system as inferred from the per-residue decomposition of the MM-GBSA binding energies. Indeed, as shown in Figure S12, the side-chains of F221^{5,56}, F228^{5,63}, and I217^{5,52} are among those that contribute the most to the stability of the dimer interfaces of the WT (Mode 1) and F220C systems. Importantly, Figure S12 reveals that per-residue binding energy for F221^{5,56} is notably larger in the F220C dimer construct. Thus, in 5/6 trajectories of the F220C system, the top contribution to dimer stabilization stems from the sidechain of residue F221^{5,56}, which contributes between -3.8 and -4.6 kcal/mole to the overall binding energy (Figure S12).

Because F221^{5,56} is about two helical turns above F228^{5,63} on the opposite protomer, stable interactions between these residues in the F220C dimer requires the phenyl ring of F228^{5,63} to position “upward” (Fig. 8G). To sustain such an orientation, the intracellular part of TM5 moves outward resulting in the overall tilt of this helix with respect to the other protomer. Thus, the splayed configuration of the F220C dimer appears to be a direct consequence of the mutation at the 220 position, which alters pair-wise residue interactions at the dimer interface.

To test whether this peculiar, splayed arrangement is specific to the F220C mutation, we computationally probed several other mutations in TM5. We considered an Ala substitution at the 220 position (F220A), and, given their importance in interfacial interactions (Fig. 8A–D), we also introduced Phe-to-Ala mutations at positions 221 and 228 using Mode 1 and Mode 2 structures of the WT dimers to create F221A/F228A 1 and F221A/F228A 2 constructs. Lastly, we created the triple mutants F220C/F221A/F228A and F220A/F221A/F228A. The new systems were subjected to all-atom MD simulations in the same manner as the F220C and the WT constructs (i.e. 6 replicates, 150 ns duration each).

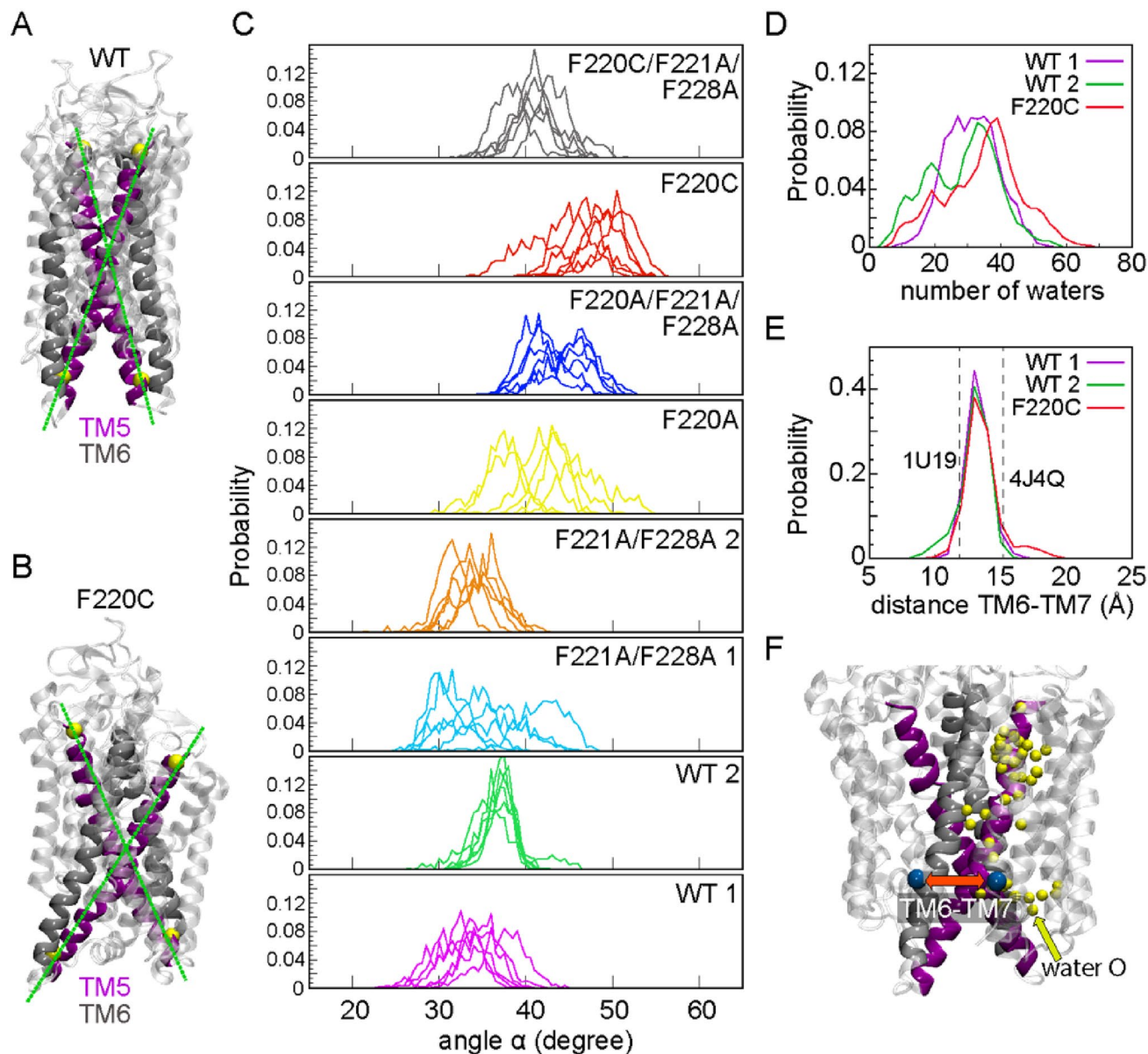


Figure 7. (A,B) Snapshots from all-atom MD simulations comparing mutual orientations of the two monomers in dimers of the WT (A) and F220C (B) constructs. TM5 and TM6 helices are shown in purple and silver, respectively. The rest of the protein is in transparent white color. Solid lines connect C_{α} atoms of residues 200 and 234 (yellow spheres) used to define the orientation of TM5 helix in each monomer (see text). (C) Probability distributions of the splay angle α between the two TM5 helices in the dimers of various constructs described in the text. The color coding of the systems follows the same pattern as in Fig. 6. The results for each replicate are shown separately (i.e. 6 plots per constructs). (D) Histogram of water counts in the interior of each monomer of the WT (Modes 1 and 2) and F220C dimers. For this analysis, in each trajectory, for a given dimer, water oxygen atoms within 7 Å of the sidechains of residues 264, 268, 272, 127, and 131 were counted in each monomer. The data was then combined for two monomers and for all the trajectories per construct (i.e. WT 1, WT 2, and F220C) and binned using a bin size of 2. (E) Histogram of C_{α} - C_{α} distance between residues 252 and 308 on each monomer of the WT (Modes 1 and 2) and F220C dimers. To construct the histogram, the data for each monomer and for all trajectories of each construct was combined and binned using a bin size of 1 Å. (F) A snapshot of the F220C dimer as in panel B, highlighting water penetration (yellow spheres show water oxygens in the protein interior) and the distance between TM6-TM7 distance reported in panel E (blue spheres are C_{α} atoms of residues 252 and 308).

Calculation of the α splay angle in the new mutant systems revealed that all the constructs that had a mutation at position 220 were, on average, characterized by larger splay angle compared to the ones in which this site was not altered (Figs. 7C and S9). Interestingly, F220C still showed the largest splay angle among all the

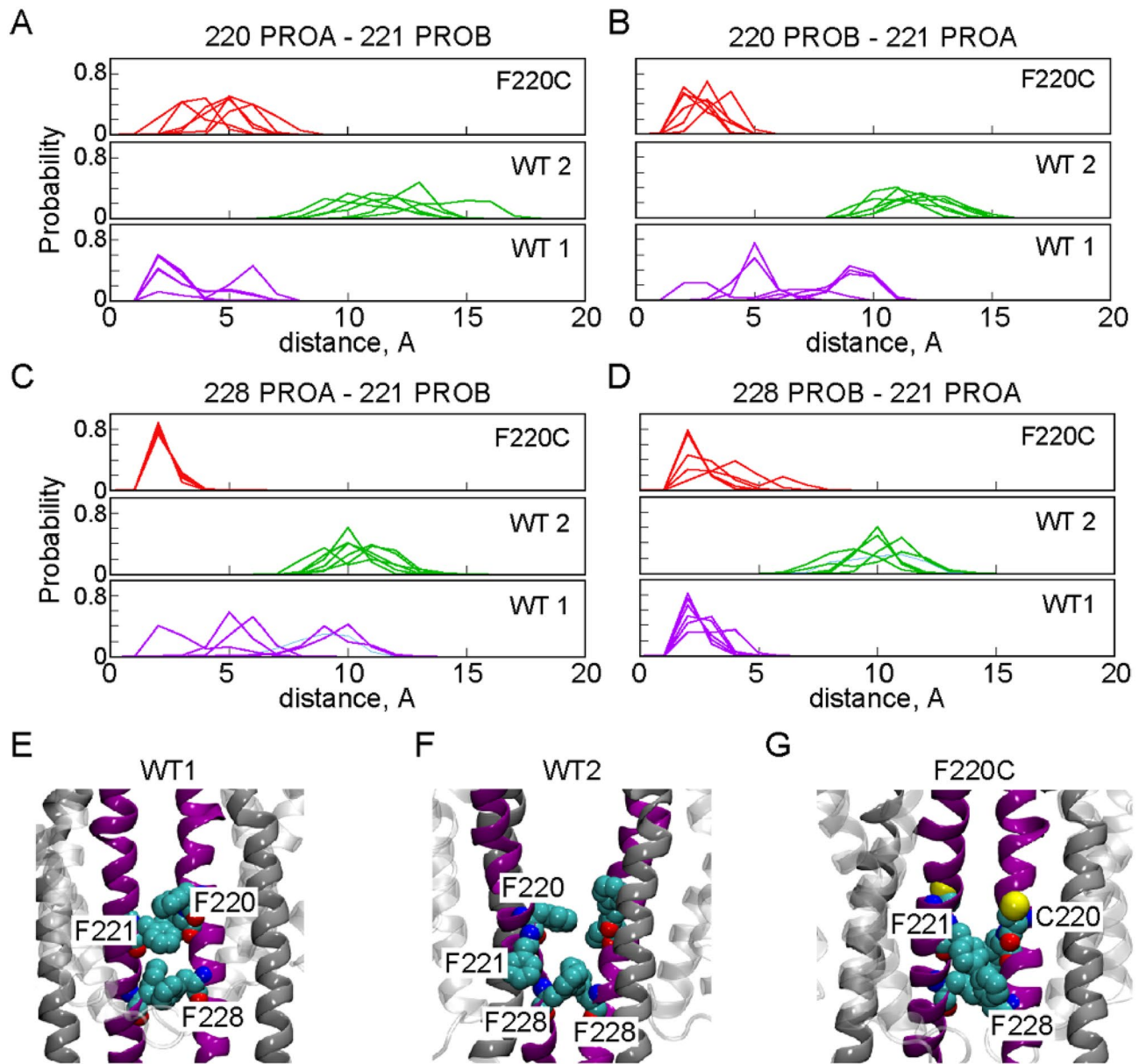


Figure 8. (A,B) Probability distribution of the minimum distance between residues F/C220^{5.55} and F221^{5.56} on the opposite monomers of dimers formed by WT opsin (Modes 1 and 2—WT1, WT2) and F220C. (C,D) Probability distribution of the minimum distance between residues F228^{5.63} and F221^{5.56} on the opposite monomers of dimers formed by WT opsin (Modes 1 and 2—WT1, WT2) and the F220C mutant. PROA (A,C) and PROB (B,D) denote two monomers of the dimer, and the results are shown separately for all six replicates for each construct. (E–G) Close-up view of molecular interactions along TM5 helix in the F220C dimer (G) and in the Mode 1 and Mode 2 WT dimers (E,F respectively). Relevant residues are shown in space-fill and labeled. TM5 and TM6 helices are shown in purple and silver, respectively.

systems. MM-GBSA calculations (Fig. 6) revealed that F221A/F228A 1 and F221A/F228A 2 constructs had WT-like binding energies which were significantly lower compared to the binding energies for the F220C system (see Table S3). On the other hand, all the constructs in which position 220 was mutated were characterized by binding energies more similar to that found for F220C compared with constructs in which the 220 position was unaltered. These results suggest that a side chain at position 220 position is crucial for the stability of the TM5-TM5 dimerization interface.

To test experimentally the effects of the mutations at 220, 221, and 228 positions, we constructed single, double and triple mutants of opsin (F1 (F228A), F2 (F221A, F228A), and F3 (F220C, F221A, F228A)), expressed them along with WT protein in HEK293 GnTI^{-/-} cells and purified them by affinity chromatography (Fig. 9A). The yield of the constructs was similar, indicating that the various mutations do not have a detrimental effect on protein structure. We reconstituted the proteins into large unilamellar vesicles and used our end-point scramblase assay to measure the fraction of vesicles able to scramble lipids as a function of the amount of protein that

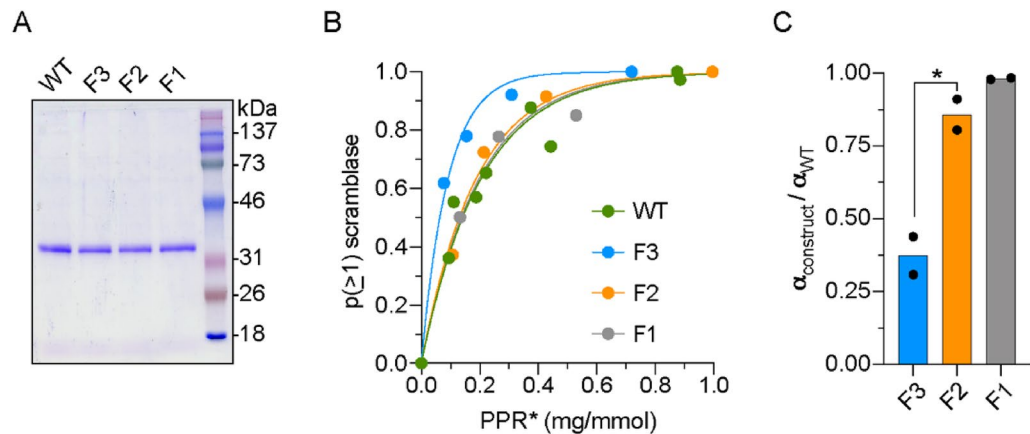


Figure 9. Reconstitution of WT opsin and TM5 mutants into liposomes. **(A)** Coomassie-stained SDS-PAGE of purified WT opsin and the F1 (F228A), F2 (F221A, F228A) and F3 (F220C, F221A, F228A) opsin constructs. Molecular weight markers are indicated. **(B)** Protein dependence of the reconstitution of vesicles with scramblase activity. NBD-PC-containing vesicles were reconstituted with different amounts of the indicated opsin construct and the extent of fluorescence reduction on dithionite treatment was determined by curve-fitting. The data were processed to calculate the probability "p(≥ 1) scramblase" that an individual vesicle in the ensemble possesses a functional scramblase. The graph shows "p(≥ 1) scramblase" versus the corrected protein/phospholipid ratio (PPR*) of the sample (the correction eliminates the contribution of vesicles that are refractory to reconstitution). **(C)** Ratio of mono-exponential fit constants (α) obtained from analyses of the reconstitution of WT and mutant opsins similar to (and including) the analysis shown in **(B)**. The ratio $\alpha_{\text{construct}}/\alpha_{\text{WT}} \sim 1$ for F1 and F2 opsins, indicating that these proteins reconstitute similarly to WT opsin as dimers or higher order multimers. However, $\alpha_{\text{F3}}/\alpha_{\text{WT}} \sim 0.4$, consistent with reconstitution of F3 as a monomer ($\alpha_{\text{WT}} = 0.25 \pm 0.07$, mean \pm S.D.). The bar chart is constructed from the results of three independent experiments, each involving 4–5 reconstitutions at different protein/phospholipid ratios (*, $p = 0.03$, unpaired t-test (two-tail)).

was reconstituted into a fixed number of liposomes. The data clearly show that F3 reconstitutes as a monomer, whereas F1 and F2 opsins that lack the F220C mutation reconstitute as dimers, similar to WT protein (Fig. 9B,C).

As an additional test of specificity of the splayed arrangement of the F220C substitution, we considered the splayed conformation of the F220C dimer ($\alpha \sim 48^\circ$), and introduced in this construct the wild type Cys residue into position 220 (C220F mutant). We then initiated 6 independent MD simulations of this dimer system. As shown in Figure S13, during ~ 300 ns runs (excluding the standard equilibration protocol described in Methods), in 3/6 trajectories α started to decrease from its initial high value, suggesting that the splayed arrangement indeed is stabilized by the F220C mutation.

Conformational differences in WT and F220C opsin dimers revealed by ensemble FRET experiments using constructs with a C-terminal SNAP-tag. The MD simulations described above indicate that the conformational pose of the F220C dimer differs from that of the WT dimer. We considered the possibility that this difference, not evident in the FRET experiments with N-terminal SNAP-tagged constructs (Fig. 3), might be revealed if the SNAP tag were located at the C-terminus. Accordingly, we expressed C-terminally SNAP-tagged WT and F220C opsins in HEK293 cells, labeled the tags using membrane-permeant donor and acceptor fluorophores, and assessed FRET (Fig. 10). The constructs were expressed at a similar level (Fig. 10A,B), and the relative presence of intracellular fluorescence (as seen in the representative images in Fig. 10A) was the same, but now the extent of donor recovery was significantly greater for F220C opsin compared with WT (Fig. 10C). Importantly, the same fluorophores were used for both N- and C-terminal SNAP-tagged constructs, eliminating the possibility that the observed difference is due to different photophysical properties of the labels. While this result does not permit a detailed structural comparison of WT and F220C dimers, it clearly shows that the SNAP tags are positioned differently in the F220C dimer compared with WT, consistent with the prediction of the MD simulation that the F220C mutation stabilizes the intracellular end of the protein juxtaposed to the C-terminus in a different conformation compared to the WT protein.

Conclusion

The F220C rhodopsin mutant was originally discovered in a study of 88 patients/families with a history of RP³⁵. Based on in vitro reconstitution studies we originally proposed¹⁶ that F220C differed from WT opsin in its ability to dimerize en route to reconstitution into lipid vesicles. We suggested that this deficiency in an otherwise normal rhodopsin could account for its disease association. However, subsequent work¹⁷ unexpectedly indicated that F220C opsin could dimerize in the membrane, and recent studies of the F220C-knockin mouse model revealed only minor phenotypic consequences of this mutation¹⁸. We now present an expansive characterization of this enigmatic opsin mutant, using a combination of experimental and in silico studies. Using new, fluorescence-based

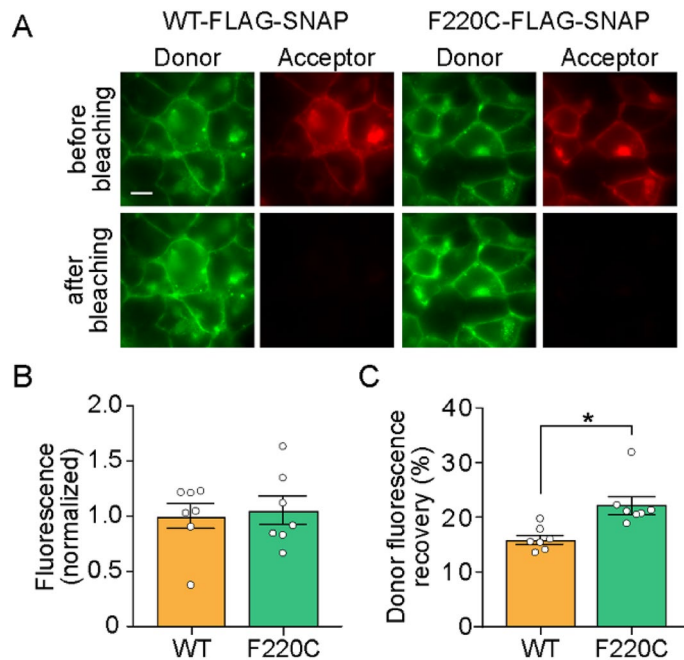


Figure 10. Distinct dimer conformations of WT and F220C opsins in the plasma membrane of HEK293 cells. **(A)** Fluorescence images of HEK293 cells expressing WT-FLAG-SNAP opsin and F220C-FLAG-SNAP opsin labeled with SNAP donor (BG-JF₅₄₉) and acceptor (BG-JF₆₄₆) dyes. The top and bottom rows show images of cells taken before and after bleaching of the acceptor by high intensity 640 nm illumination. Scale bar, 10 μ m. **(B)** Summary of acceptor fluorescence as a reporter of comparable surface expression for both constructs. **(C)** Summary of donor fluorescence recovery upon acceptor bleaching, indicative of FRET (*, $p=0.007$, unpaired t-test (two-tail)).

experimental approaches we show that F220C opsin, like WT opsin, is a monomer when expressed in HEK293 cells and isolated using 1% (w/v) DDM. However, reduction of detergent levels causes the WT protein to dimerize whereas F220C remains monomeric. Remarkably, MD simulations showed that the micellar environment is energetically unsuited to opsin dimers, raising the question of why the WT protein dimerizes at all under these conditions. Further work will be needed to understand this phenomenon. The behavior of the F220C mutant in the membrane also contrasted with that of the WT protein. Whereas we could show that F220C dimerizes in the membrane as previously reported, MD simulations revealed that it forms an energetically favored dimer when compared with the WT protein. The conformation of the F220C dimer is unique, with the TM5/TM6 helices tilted to form a splayed, criss-crossed structure promoting widening of the intracellular vestibule of each monomer in the dimer. This could potentially enhance the ability of the receptor to couple to effector proteins, such as transducin, and also influence the higher order assembly of rhodopsins in disc membranes. While it is intriguing that the difference in quaternary structure between WT and F220C opsins does not immediately translate into a more dramatic phenotypic consequence in the F220C-knockin mouse model, the unusual nature of the F220C dimer may account, at least in part, for the observed phenotype and disease association in humans.

Experimental procedures

Computational methods. *Molecular constructs for all-atom MD simulations of opsin monomer systems.* All computations were based on the X-ray structure of retinal-free opsin (PDBID 4J4Q)³⁶. This structure also contains bound synthetic GaCT peptide which was not considered here. Using the CHARMM-GUI web interface³⁷, the wild type (WT) and F220C opsin models were prepared and embedded in an all-atom phospholipid bilayer containing a 9:1 mixture of POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) and POPG (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)) lipids, mimicking the composition of the reconstituted vesicles used for scrambling activity assays³. The protein to lipid ratio was 1:362. After adding a solvation box containing 150 mM KCl the total system size was ~ 130,000 atoms.

All-atom MD simulations of opsin monomer systems. All-atom systems containing the WT and F220C mutant opsin monomers were subjected to ~ 130 ns MD simulations using NAMD 2.12 software and the latest all-atom CHARMM 36m force fields for protein, lipids and ions³⁸. The simulations followed the multi-step equilibration protocol prescribed by CHARMM-GUI before switching to unbiased runs. These simulations implemented the *all* option for rigid bonds, 2 fs integration time-step, PME for electrostatics interactions³⁹, and were carried out in NPT ensemble under semi-isotropic pressure coupling conditions, at a temperature of 303 K. The Nose-Hoover Langevin piston algorithm⁴⁰ was used to control the target $P=1$ atm pressure with the Langevin piston period set to 100 fs and decay time constant set to 50 fs. The van der Waals interactions were

calculated applying a cutoff distance of 12 Å and switching the potential from 10 Å. In addition, the `vdwforces-witching` option was set to *on*.

Molecular constructs for coarse-grained (CG) Martini simulations. Using CHARMM-GUI, the opsin monomer models (WT and F220C) from the final frame of the respective all-atom MD trajectories were transformed into Martini-based coarse-grained (CG) structural representation^{41,42}. For each construct, a second copy of the CG protein was then created by translating the center-of-mass of the first one in *x/y* (membrane) plane by 50 Å. The two CG opsins were randomly rotated with respect to each other and around *z*-axis to create 5 different mutual orientations of the pair of proteins. Using again CHARMM-GUI, these 5 opsin pairs were embedded into a 1099 lipid-size membrane containing 9:1 mixture of POPC and POPG lipids, the membrane composition used for the proteoliposomes in the experimental scramblase activity assay (see below). The total system size, after solvating and ionizing steps was 38,280 CG beads.

CG MD simulations of spontaneous aggregation of opsin monomers into dimers. For each protein construct (i.e. WT or F220C) and mutual orientation 7 independent CG MD simulations of 120 μs in length were performed using Gromacs version 5.1.4, resulting in 35 total simulations per construct with a cumulative time of 4.2 ms. The simulations implemented 20 fs timestep, and were carried out in NPT ensemble under semi-isotropic pressure coupling conditions, at a temperature of 303 K. The Berendsen algorithm was used to control the target $P = 1$ atm pressure with the τ_p and *compressibility* parameters set to 5.0 ps and $3e-4$ bar⁻¹, respectively. The temperature was controlled by the *v-rescale* algorithm (separated for protein, membrane and solute).

Molecular constructs for all-atom MD simulations of opsin dimer systems. Using CHARMM-GUI, selected opsin dimer CG models for the WT and F220C system (see Results) were transformed back into all-atom representation. For the WT system, two models were considered, termed, Mode 1 and Mode 2. For F220C, a single CG model was transformed. To create an all-atom model for the F221A/F228A double mutant system, Phe-to-Ala substitutions were introduced to Mode 1 and Mode 2 WT models, resulting in two starting conformations for the double mutant system, F221A/F228A 1 and F221A/F228A 2. The all-atom model for the F220A mutant was created by replacing Cys220 by Ala in the all-atom F220C dimer. Lastly, two triple mutant systems were created F220A/F221A/F228A and F220C/F221A/F228A by introducing F221A and F228A mutations into the F220A and F220C models. Together, 6 different dimer constructs were created, totaling 8 systems: WT (two starting models), F220C, F220A, F221A/F228A double mutant (two starting models), F220A/F221A/F228A, and F220C/F221A/F228A.

Using CHARMM-GUI, each of these 8 models were embedded in a 900 lipid-size membrane consisting of 9:1 mixture of POPC/POPG lipids. After adding a solvation box containing 150 mM KCl the total system size was ~ 320,000 atoms.

All-atom MD simulations of opsin dimer systems in lipid membranes. The eight all-atom dimer systems were subjected to ensemble MD simulations using NAMD 2.12 software⁴⁰. Each system was run in six 150 ns long independent replicates, resulting in cumulative time of 900 ns per construct. For these runs as well, the latest all-atom CHARMM 36m force fields for protein, lipids and ions, and the simulations used the same protocols and run parameters as described above for the all-atom monomer systems.

All-atom MD simulations of opsin dimer systems in detergent micelles. WT and F220C opsin monomers and dimers were simulated in micelles composed of DDM detergent. The initial structures of the dimer systems were taken from the all-atom MD simulations in lipid membranes in which the F220C dimer was in splayed conformation whereas in the WT dimer the two subunits were more aligned with respect to each other (see text). The same structures served as starting points for the monomer simulations. The protein constructs were inserted into DDM micelles, solvated and ionized using CHARMM-GUI. Protein-to-detergent ratio was 1:150 for the monomers and 1:300 for the dimers. Each system was first equilibrated using standard CHARMM-GUI equilibration protocol and then run with OpenMM 7.4⁴³ in six independent replicates (WT systems – 150 ns each, F220C monomer – 260 ns each, F220C dimer – 300 ns each). For these simulations as well, the latest all-atom CHARMM 36m force fields for protein, lipids and ions were used. The simulations used the same protocols and run parameters as described above for the all-atom systems except 4 fs integration timestep (with hydrogen mass repartitioning), and isotropic pressure coupling.

Calculation of dimerization energies from all-atom MD simulations using the MM-GBSA method. To quantify strength of opsin dimerization, we calculated binding energies using the MM-GBSA (Molecular Mechanics, Generalized Born Surface Area)⁴⁴. The approach combines molecular mechanics interaction energies from the CHARMM36 force field⁴⁵ with a generalized Born and surface area continuum solvation model to estimate solvation effects upon binding of macromolecules (in this case two monomers of the protein dimer). Protein conformations are sampled from all-atom MD trajectories, from which all water molecules, lipids and ions are discarded. Here we used the single-trajectory variant of the method, where unbound conformations of one protomer are obtained by discarding the atoms of the other protomer. Thus, the results do not take into account any reorganization energies associated with internal conformational changes upon binding. Despite this strong assumption, the approach has been widely used to compare strength of binding of different macromolecular complexes, with the useful possibility of performing approximate per-residue binding energy decompositions^{46,47} that compare well to experimental alanine scanning^{48–50}.

MM-GBSA analysis was carried out on 25–150 ns time-intervals of the all-atom trajectories using the implementation in the CHARMM software⁵¹ complemented with in-house scripts, as described previously⁵². Briefly, the dielectric constant inside the protein was set to 2 and we used the accurate Generalized Born Molecular Volume II (GBMVII) method⁵³ to calculate the polar solvation free energies. The non-polar solvation free energy was proportional to the solvent accessible surface area (SASA) with a coefficient of $-0.0072 \text{ kcal/mol/Å}^2$ ⁵⁴. We used the Debye-Hückel correction to account for ionic screening⁵⁵, with screening constant $\kappa = 0.316\sqrt{[\text{salt}]}$ (expressed in Å^{-1})⁵⁶, where $[\text{salt}]$ is the monovalent ion concentration taken to be 0.154 mol/l. The membrane was treated as a continuum slab with a dielectric constant varying from low in the membrane core to medium in the head-group region using the Heterogeneous Dielectric Generalized Born (HDGB) method⁵⁷. We used the improved dielectric profile and non-polar correction profile from Ref.⁵⁸.

Using this framework, to calculate binding energy in a specific dimer construct, MM-GBSA calculations were performed to obtain the solvation free energies separately for three molecular systems: for the dimer (G_{DIMER}), and for the two monomers of the dimer separately (G_{PROA} and G_{PROB}). Entropic contributions in the proteins were neglected, so that the dimerization free energy was then estimated from: $\Delta G_{\text{bind}} = E_{\text{int}} + G_{\text{DIMER}} - (G_{\text{PROA}} + G_{\text{PROB}})$, where E_{int} is the electrostatic and Van der Waals interaction energy from CHARMM36 with no distance cutoff. All terms in ΔG_{bind} being either local or pair-wise interactions, a decomposition is obtained by summing all local contribution and half of the pair-wise interactions for all atoms in a group (typically a residue or a side chain)^{46,47}.

Calculation of the RHM energies from all-atom MD simulations in micelle systems. RHM values for each protein residue was calculated as described previously^{25–27,59}. Briefly, the residual exposure of each residue ($SA_{\text{res},i}$) was quantified from Solvent Accessible Surface Areas (SASA) calculated from corresponding MD trajectories, as follows:

For hydrophobic residues, $SA_{\text{res},i}$ is the surface area of the hydrophobic residues exposed to the micelle head-group or water and is calculated as the SASA with the solute taken to be the protein and hydrophobic core of the detergent micelle (i.e. hydrocarbon tails of DDM molecules, atoms C1-C12 in CHARMM36m nomenclature).

For hydrophilic residues, $SA_{\text{res},i}$ is the surface area of the residue embedded in the hydrophobic core of the micelle and is calculated as the difference of the SASA with the solute taken to be the protein only and the SASA with the solute taken to be the protein and the hydrophobic part of the micelle.

The corresponding energy penalty ($\Delta G_{\text{res},i}$) can be approximated as linearly proportional to $SA_{\text{res},i}$: $\Delta G_{\text{res},i} = \sigma_{\text{res}} SA_{\text{res},i}$ with the constant of proportionality σ_{res} taken to be $0.028 \text{ kcal}/(\text{mol} \cdot \text{Å}^2)$ ^{60,61}. For more details of the method please see Ref.^{25–27,59}.

Experimental methods. *Materials.* Anti-FLAG-M2 affinity agarose gel and 3X FLAG peptide were obtained from Sigma Aldrich, n-dodecyl β -D-maltoside (DDM) was from Anatrace, protease inhibitor cocktail Set V was from Calbiochem/EMD Millipore Corp, Transporter 5™ transfection reagent (PEI MAX) was from Polysciences Inc., Antibiotic–Antimycotic (100X, containing 10,000 units penicillin, 10,000 $\mu\text{g/ml}$ streptomycin and 25 $\mu\text{g/ml}$ amphotericin) was from Gibco, and spin columns (product number 69705) were from Pierce Biotechnology. HEK293S GnTI⁺ cells were a kind gift of the Meyerson laboratory (Weill Cornell Medical College). NEBuilder HiFi DNA Assembly Cloning Kit, Phusion polymerase, Dpn1 and T4 DNA ligase were from New England Biolabs, and primers were obtained from IDT. Phospholipids (1-palmitoyl-2-oleoyl-glycerol-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-*sn*-glycerol-3-phospho-(1'-*rac*-glycerol)(POPG) and 1-palmitoyl-2- $\{6-[(7\text{-nitro-2-1,3-benzoxadiazol-4-yl)amino]hexanoyl\}$ -*sn*-glycerol-3-phosphocholine (NBD-PC)) were obtained as stock solutions in chloroform from Avanti Polar Lipids. Sodium dithionite was from Sigma. SNAP-Surface 549 and SNAP-Surface 649 were from New England Biolabs. Bio-Beads SM2 resin was from Bio-Rad Laboratories; the resin was washed twice with methanol (25 ml per g of resin), twice with water and once with buffer (as indicated) prior to use.

Opsin constructs. We previously described a bovine opsin construct containing N2C and D282C mutations for enhanced thermostability, and a C-terminal tandem 3X-FLAG tag¹⁶. The construct, here termed wild-type (WT), was prepared in the pMT3 expression vector. The F220C construct, derived from WT, was also previously described¹⁶. Additional mutations were introduced into the WT construct to generate F1 (F228A) and F2 (F221A, F228A), and into F220C to generate F3 (F220C, F221A, F228A). Site-directed mutagenesis was achieved using the following primers: for F2, forward primer—**GCC** TGC TAT GGC CAG CTG GTG **GCC** ACC GTC AAG GAG GCT GCA GCC C, reverse primer—GAA GAT GAC AAT CAG CGG GAT GAT G, and for F3, forward primer—**GCC** TGC TAT GGC CAG CTG GTG **GCC** ACC GTC AAG GAG GCT GCA GCC C, reverse primer—GCA GAT GAC AAT CAG CGG GAT GAT G, and for F228A, forward primer—**GCC** ACC GTC AAG GAG GCT GCA GCC C, and reverse primer—CAC CAG CTG GCC ATA GCA GAA GA. Polymerase chain reactions (PCR) were performed using High Fidelity Phusion polymerase. PCR-amplified products (verified by electrophoretic separation using a 1% agarose gel) were treated with the Dpn1 restriction enzyme to remove the parent plasmid before being subjected to ligation (using T4 DNA ligase) and transformed into *Escherichia coli* DH5 α cells. All constructs were sequenced to confirm the desired mutations.

N-terminally tagged HA-SNAP-Opsin (WT and F220C) clones were made by replacing the mGluR2 gene from a previously described N-terminal HA-SNAP-mGluR construct⁶² with thermostable opsin-FLAG, using the NEBuilder HiFi DNA Assembly Cloning Kit. The HA-SNAP containing vector backbone was amplified using the forward and reverse primers TGAAGATCCCACACTCCTGCCC and ACGCGTGCCAGCCCAGG, respectively, and opsin genes (WT and F220C) were amplified from the corresponding pMT3-opsin constructs (described above) using forward (**AGCCTGGGCTGGGCACGCGTTGCGGTACCGAAGGCCAAAC**) and reverse (**GCAGGAGTGTGGGATCTTCACTTGTCATCGTCATCCTTGATGTCG**) primers containing the

overlapping region (underlined) from the HA-SNAP-mGluR construct. The constructs were confirmed by sequencing.

Cell culture and transfection. Homogeneously *N*-glycosylated FLAG-tagged opsins were produced by transfection of HEK293S GnTI⁻ cells using the Transporter 5[™] transfection reagent (PEI MAX). Briefly, cells were harvested from a confluent (80–90%) 10-cm plate cultured at 37 °C in a 5% CO₂ atmosphere in Dulbecco's Modified Eagle Medium (DMEM) containing 10% fetal bovine serum (FBS) and antibiotic–antimycotic, and split into four 10-cm plates using the same medium. After 24 h when the cells were 60–70% confluent, the culture medium was replaced with 10-mL (per plate) fresh DMEM containing 4% FBS. The cells were incubated for a further 2 h before transfection. For each plate, a transfection mix was prepared by combining 1 ml 150 mM NaCl, plasmid DNA (20 µg of a 4:1 (w/w) mixture of the pMT3 plasmid carrying the opsin gene and pRSVTag carrying a gene encoding SV40 large tumour (T) antigen), and 75 µl PEI MAX. The mix was incubated at room temperature for 20 min before being added drop-wise to the culture plate while gently swirling it. The plates were incubated for 72–96 h. At the end of the incubation, the medium was removed and the cell monolayer was washed once in phosphate buffered saline, before scraping off the cells using a rubber policeman, and recovering them by centrifugation. Cell pellets were either snap frozen and stored at -80 °C, or directly taken for purification of opsin.

For SiMPull and ensemble FRET experiments, opsin constructs with N-terminal HA-SNAP or C-terminal FLAG-SNAP were expressed in HEK293T cells or HEK293S GnTI⁻ cells grown on poly-L-lysine coated coverslips (18 mm in diameter, Calorina) and processed 24–36 h after transfection.

SiMPull experiments. SNAP-tagged opsins were labeled by incubating expressing cells with 1.5 µM LD555 SNAP fluorophore in EX solution containing (in mM): 10 HEPES, 135 NaCl, 5.4 KCl, 2 CaCl₂, 1 MgCl₂, pH 7.4 for 45 min at 37 °C. After the labeling, cells were washed with EX solution and harvested using Ca²⁺ free-DPBS for 20 min at 37 °C. After pelleting at 10,000 × *g*, 4 °C for 1 min, the cells were lysed with 1% (w/v) DDM for 1 h at 4 °C. The lysate was clarified by centrifugation at 16,000 × *g* (4 °C, 20 min) and diluted to 0.1% DDM for SiMPull. A glass flow chamber was prepared as previously described^{20,21} using mPEG/biotin-PEG-passivated quartz slides and coverslips coated with neutravidin and subsequently decorated with biotinylated anti-HA antibodies. The flow chamber was placed on an inverted TIRF microscope (Olympus IX83 with cellTIRF system) and LD555-labeled HA-SNAP-opsin constructs were immobilized at a density that allowed clear resolution of individual spots (The fluorophore was excited using a DPSS 561 nm laser, imaged through a 100× objective (NA = 1.49) and the data was collected at room temperature. Images were acquired with a sCMOS detector camera (Hamamatsu ORCA-flash4.0 V3) at 20 Hz using Olympus cellSens software. Spots were analyzed using a previously described program⁶³ and manually characterized as bleaching in 1, 2, 3, or 4 steps (or deemed uncountable). Data in bar graphs represent averages of the step distributions for different movies.

Ensemble FRET experiments. Cells expressing SNAP-tagged proteins were labeled with donor and acceptor fluorophores at a 1:3 molar ratio with SBG- (sulfonated, membrane-impermeant) or BG- (cell-permeable) derivatives of JF549 and JF646 dyes³¹ in EX solution for 45 min at 37 °C. After the labeling cells were washed with EX solution. For the cells labeled with cell-permeable dyes, cells were incubated in EX solution > 30 min at 37 °C so that unbound dyes can diffuse out of the cells. Images were taken using a 60× objective on an inverted microscope using an sCMOS detector camera (see SiMPull methods) in both donor and acceptor channels using donor excitation with low intensity 561 nm laser illumination before and after bleaching the acceptor with high-intensity 640 nm illumination. Regions of interest were manually drawn around the same clusters of cells and the percentage increase in donor fluorescence intensity following acceptor bleaching was calculated for at least 7 images per condition. Surface expression was calculated by measuring fluorescence intensity in the acceptor channel prior to bleaching using low intensity 640 nm illumination. A correction (~10%) was added to the donor recovery values for each data point to account for incidental donor recovery, which was calculated from donor-only controls. For all comparisons the same laser intensities and exposure times (100 ms) were used.

Purification and quantification of opsin. Anti-FLAG-M2 agarose (~10 µl bed volume per 10-cm plate of cells) was pre-equilibrated by washing 3 times with 10 bed volumes of wash buffer (50 mM HEPES pH 7.4 100 mM NaCl, 1.0% (w/v) DDM). Cell pellets were incubated in 400 µl (per 10-cm plate of cells) of lysis buffer (10 mM HEPES pH 7.4, 100 mM NaCl, 10% (w/v) DDM, protease inhibitor cocktail) at 4 °C for 40 min with end-over-end mixing. The lysate was clarified by ultracentrifugation at 150,000 × *g* for 20 min at 4 °C, and the supernatant was transferred to the washed anti-FLAG-M2 agarose and incubated at 4 °C for 2 h with end-over-end mixing. The resin was collected by centrifugation using a Pierce spin column and washed 3 times with 10 bed volumes of wash buffer, before incubating at 4 °C for 1 h with wash buffer containing 0.15 mg/mL 3X-FLAG peptide (20 µl per 10-cm plate) to elute bound opsin. The eluate was collected by centrifugation, and the elution step was repeated. Purified opsin (typically a 4 µl aliquot) was analyzed by Coomassie-stained SDS-PAGE, in parallel with in-gel bovine serum albumin (BSA) standards (200–1000 ng). Band intensities were quantified by ImageJ and opsin was estimated in comparison with the linear calibration plot obtained with BSA standards. The yield of the purified opsin (similar for all constructs) was ~6 µg per 10-cm plate of cells, at a concentration of ~150 ng/µl. The purified protein was kept at 4 °C for up to a week or snap-frozen in aliquots and stored at -80 °C.

FRET measurements in vitro. WT and F220C opsins with a C-terminal FLAG-SNAP tag¹⁶ were captured onto anti-FLAG resin and labeled with either SNAP-Surface 549 (donor (D)) or SNAP-Surface 649 (acceptor (A)). The resin was washed to remove unreacted dye, and the labeled proteins were eluted with FLAG peptide as

described above. The labeled proteins (in 0.1% (w/v) DDM) were mixed in 1:1 ratio in a total volume of 400 μ l. Samples are treated with 0 to 8 mg amounts of washed Bio-Beads SM2 for 1 h at room temperature with end-over-end mixing. After the incubation period, the sample was removed and taken for fluorescence spectroscopy. Emission spectra ($\lambda_{\text{ex}} = 555$ nm) were acquired for donor (D) alone, acceptor (A) alone, and D-A pairs of WT or F220C opsin. Buffer subtraction was done after first using the Rayleigh peak to scale the buffer spectrum to that of the sample (scale factor > 1); the scaled buffer spectrum was used for subtraction. Buffer-subtracted spectra were further adjusted for protein loss due to Bio-Beads treatment by directly exciting the acceptor ($\lambda_{\text{ex}} = 650$ nm), determining the area under the spectrum over the range 660–700 nm, calculating the ratio of the area under the spectrum for the D-A sample versus that of the corresponding A-only sample, and using this ratio as a correction factor. A FRET measure ($\text{FRET}_{\text{scaled}}$) was obtained by calculating the ratio of the emission intensity at 670 nm ($\lambda_{\text{ex}} = 555$ nm) for D-A mixtures relative to the A-only sample, followed by an offset correction corresponding to mock-treated samples.

Liposome preparation. A mixture of POPC, POPG (9:1) and NBD-PC (0.3 mol percentage) (total lipids 52.5 μ mol) was dried in a round-bottomed flask using a rotary evaporator. The lipids were dissolved in diethyl ether and dried again in the rotary evaporator, followed by incubation in a vacuum desiccator overnight (or for at least 3 h). The lipid film was rehydrated in 10 ml Buffer A (50 mM HEPES–NaOH, pH 7.4, 100 mM NaCl) by rotating (250 rpm) at 37 °C for 20 min using the rotary evaporator (no vacuum) and then subjected to five freeze–thaw cycles (freezing in liquid nitrogen, thawing in a 30 °C bath). The resulting suspension was extruded using a LIPEX extruder (Northern Lipids, Inc., Burnaby, BC, Canada) by 10 passes through 400-nm track-etch polycarbonate membranes followed by 4 passes through 200-nm track-etch polycarbonate membranes to generate unilamellar liposomes. The liposomes were kept at 4 °C and used within 2 weeks. Phospholipids were quantified using a colorimetric assay as described previously^{2,3}, and the size distribution of the liposomes was determined by dynamic light scattering using a Litesizer 500 instrument (Anton Paar, USA).

Reconstitution of opsin—preparation of proteoliposomes. Unilamellar liposomes (800 μ l, ~ 3.3 mM phospholipid) were destabilized in an 840- μ l reaction containing 8.2 mM DDM in a 2-ml Eppendorf tube. After incubation for 3 h at room temperature (RT) with end-over-end mixing, DDM-solubilized opsin and Buffer A were added and the sample was mixed for an additional hour at room temperature. When titrating the amount of opsin to be reconstituted, the DDM levels of all samples was adjusted to that of the sample containing the maximum amount of opsin. To form sealed vesicles, DDM was removed by adding washed Bio-Beads SM2 as follows. First, 80 mg Bio-Beads were added, followed by end-over-end mixing for 1 h at RT. This was followed by the addition of 160 mg of beads and incubation for an additional 2 h at RT. The sample was then transferred to a glass screw-cap tube containing 160 mg of Bio-Beads SM2 and mixed end-over-end overnight at 4 °C. We previously showed that the resulting proteoliposomes have undetectable levels of DDM, i.e. a DDM-to-phospholipid ratio of < 1:100. The recovery of both opsin and phospholipids was ~ 65% as determined by quantitative western blotting and a colorimetric assay, respectively³. Protein-free liposomes were reconstituted in parallel.

Scramblase activity assay. Liposomes or proteoliposomes (50 μ l) were added to 1950 μ l Buffer A in a plastic fluorimeter cuvette equipped with a stir bar. Time-based fluorescence was recorded ($\lambda_{\text{ex}} = 470$ nm; $\lambda_{\text{em}} = 530$ nm, excitation and emission slits 2.0 nm, data acquisition frequency 1 Hz, 25 °C, stirring 900 rpm) in a temperature-controlled Horiba Fluoromax + spectrofluorometer. After measurement for at least 50 s to ensure a stable signal, 40 mM dithionite (40 μ l of a freshly prepared 1 M solution of sodium dithionite in unbuffered 0.5 M Tris) was added through the injection port in the lid of the fluorimeter chamber, and fluorescence was monitored for at least 500 s. Fluorescence traces were normalized to the starting value of fluorescence (F_0) at the time of dithionite addition and analyzed by fitting to a double-exponential function. As described previously³, the fast component of the fit with $t_{1/2} = 16.8 \pm 3.4$ s (mean \pm S.D. (n = 27)), accounted for the majority of the fluorescence change in both liposomes and proteoliposomes. The $t_{1/2}$ of the slow component was an order of magnitude greater; its origin is unknown³ but may be due to a combination of dithionite entry and slow spontaneous scrambling that is present in all traces irrespective of the presence of reconstituted protein. The plateau value of fluorescence (F_{min}) obtained from the fit was taken as the end-point of the assay and used to calculate the fraction of vesicles equipped with a functional scramblase ($F_{\text{min}} = 0.55$ for protein-free liposomes as previously reported³). Briefly, values of $(0.55 - F_{\text{min}})$ were calculated for all proteoliposome samples and normalized to the value obtained for samples reconstituted at a high protein-phospholipid ratio (PPR) (typically > 0.48 mg opsin/mmol phospholipid) to obtain $p(\geq 1)$ scramblase as described³. PPR values were scaled to PPR^* to include only those vesicles that could be reconstituted with protein. Thus, $\text{PPR}^* = \text{PPR}/x$ where x is twice the F_{min} value obtained for samples reconstituted at a high protein-phospholipid ratio¹⁶. Plots of $p(\geq 1)$ scramblase versus PPR^* were fitted to a mono-exponential function^{3,16}, yielding a fit constant (α) that was used to compare the dimerization propensity of various opsin constructs.

Data availability

All data generated or analyzed during this study are included in this published article and available from GK (gek2009@med.cornell.edu) or AKM (akm2003@med.cornell.edu) on reasonable request.

Received: 2 March 2021; Accepted: 29 April 2021

Published online: 18 May 2021

References

- Palczewski, K. G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.* **75**, 743–767. <https://doi.org/10.1146/annurev.biochem.75.103004.142743> (2006).
- Menon, I. *et al.* Opsin is a phospholipid flippase. *Curr. Biol.* **21**, 149–153. <https://doi.org/10.1016/j.cub.2010.12.031> (2011).
- Goren, M. A. *et al.* Constitutive phospholipid scramblase activity of a G protein-coupled receptor. *Nat. Commun.* **5**, 5115. <https://doi.org/10.1038/ncomms6115> (2014).
- Ernst, O. P. & Menon, A. K. Phospholipid scrambling by rhodopsin. *Photochem. Photobiol. Sci.* **14**, 1922–1931. <https://doi.org/10.1039/c5pp00195a> (2015).
- Shah, U. H., Toneatti, R., Gaitonde, S. A., Shin, J. M. & Gonzalez-Maeso, J. Site-specific incorporation of genetically encoded photo-crosslinkers locates the heteromeric interface of a GPCR complex in living cells. *Cell Chem. Biol.* **27**, 1308–1317 e1304. <https://doi.org/10.1016/j.chembiol.2020.07.006> (2020).
- Milligan, G., Ward, R. J. & Marsango, S. GPCR homo-oligomerization. *Curr. Opin. Cell Biol.* **57**, 40–47. <https://doi.org/10.1016/j.ceb.2018.10.007> (2019).
- Fotiadi, D. *et al.* Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **421**, 127–128. <https://doi.org/10.1038/421127a> (2003).
- Gunkel, M. *et al.* Higher-order architecture of rhodopsin in intact photoreceptors and its implication for phototransduction kinetics. *Structure* **23**, 628–638. <https://doi.org/10.1016/j.str.2015.01.015> (2015).
- Knepp, A. M., Periolo, X., Marrink, S. J., Sakmar, T. P. & Huber, T. Rhodopsin forms a dimer with cytoplasmic helix 8 contacts in native membranes. *Biochemistry* **51**, 1819–1821. <https://doi.org/10.1021/bi3001598> (2012).
- Kota, P., Reeves, P. J., Rajbhandary, U. L. & Khorana, H. G. Opsin is present as dimers in COS1 cells: Identification of amino acids at the dimeric interface. *Proc. Natl. Acad. Sci. U S A* **103**, 3054–3059. <https://doi.org/10.1073/pnas.0510982103> (2006).
- Zhao, D. Y. *et al.* Cryo-EM structure of the native rhodopsin dimer in nanodiscs. *J. Biol. Chem.* **294**, 14215–14230. <https://doi.org/10.1074/jbc.RA119.010089> (2019).
- Jastrzebska, B. *et al.* Disruption of rhodopsin dimerization with synthetic peptides targeting an interaction interface. *J. Biol. Chem.* **290**, 25728–25744. <https://doi.org/10.1074/jbc.M115.662684> (2015).
- Farrar, G. J., Kenna, P. F. & Humphries, P. On the genetics of retinitis pigmentosa and on mutation-independent approaches to therapeutic intervention. *EMBO J.* **21**, 857–864. <https://doi.org/10.1093/emboj/21.5.857> (2002).
- Hartong, D. T., Berson, E. L. & Dryja, T. P. Retinitis pigmentosa. *Lancet* **368**, 1795–1809. [https://doi.org/10.1016/S0140-6736\(06\)69740-7](https://doi.org/10.1016/S0140-6736(06)69740-7) (2006).
- Mendes, H. F., van der Spuy, J., Chapple, J. P. & Cheetham, M. E. Mechanisms of cell death in rhodopsin retinitis pigmentosa: Implications for therapy. *Trends Mol. Med.* **11**, 177–185. <https://doi.org/10.1016/j.molmed.2005.02.007> (2005).
- Ploier, B. *et al.* Dimerization deficiency of enigmatic retinitis pigmentosa-linked rhodopsin mutants. *Nat. Commun.* **7**, 12832. <https://doi.org/10.1038/ncomms12832> (2016).
- Mallory, D. P. *et al.* The retinitis pigmentosa-linked mutations in transmembrane helix 5 of rhodopsin disrupt cellular trafficking regardless of oligomerization state. *Biochemistry* **57**, 5188–5201. <https://doi.org/10.1021/acs.biochem.8b00403> (2018).
- Lewis, T. R. *et al.* The F220C and F45L rhodopsin mutations identified in retinitis pigmentosa patients do not cause pathology in mice. *Sci. Rep.* **10**, 7538. <https://doi.org/10.1038/s41598-020-64437-y> (2020).
- Jain, A. *et al.* Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484–488. <https://doi.org/10.1038/nature10016> (2011).
- Jain, A., Liu, R., Xiang, Y. K. & Ha, T. Single-molecule pull-down for studying protein interactions. *Nat. Protoc.* **7**, 445–452. <https://doi.org/10.1038/nprot.2011.452> (2012).
- Pandey, K. *et al.* An engineered opsin monomer scrambles phospholipids. *Sci. Rep.* **7**, 16741. <https://doi.org/10.1038/s41598-017-16842-z> (2017).
- Strop, P. & Brunger, A. T. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci.* **14**, 2207–2211. <https://doi.org/10.1110/ps.051543805> (2005).
- Ernst, O. P., Gramse, V., Kolbe, M., Hofmann, K. P. & Heck, M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. *Proc. Natl. Acad. Sci. U S A* **104**, 10859–10864. <https://doi.org/10.1073/pnas.0701967104> (2007).
- Marsh, D. Energetics of hydrophobic matching in lipid-protein interactions. *Biophys. J.* **94**, 3996–4013 (2008).
- Mondal, S., Khelashvili, G. & Weinstein, H. Not just an oil slick: How the energetics of protein-membrane interactions impacts the function and organization of transmembrane proteins. *Biophys. J.* **106**, 2305–2316. <https://doi.org/10.1016/j.bpj.2014.04.032> (2014).
- Mondal, S. *et al.* Membrane driven spatial organization of GPCRs. *Sci. Rep.* **3**, 2909. <https://doi.org/10.1038/srep02909> (2013).
- Mondal, S., Khelashvili, G., Shan, J., Andersen, O. S. & Weinstein, H. Quantitative modeling of membrane deformations by multi-helical membrane proteins: Application to G-protein coupled receptors. *Biophys. J.* **101**, 2092–2101 (2011).
- Gutzzeit, V. A. *et al.* Conformational dynamics between transmembrane domains and allosteric modulation of a metabotropic glutamate receptor. *Elife* **8**, <https://doi.org/10.7554/eLife.45116> (2019).
- Lan, T. H. *et al.* BRET evidence that beta2 adrenergic receptors do not oligomerize in cells. *Sci. Rep.* **5**, 10166. <https://doi.org/10.1038/srep10166> (2015).
- Levitz, J. *et al.* Mechanism of assembly and cooperativity of homomeric and heteromeric metabotropic glutamate receptors. *Neuron* **92**, 143–159. <https://doi.org/10.1016/j.neuron.2016.08.036> (2016).
- Poc, P. *et al.* Interrogating surface versus intracellular transmembrane receptor populations using cell-impermeable SNAP-tag substrates. *Chem. Sci.* **11**, 7871–7883. <https://doi.org/10.1039/D0SC02794D> (2020).
- Wassenaar, T. A. *et al.* High-throughput simulations of dimer and trimer assembly of membrane proteins. The DAFT approach. *J. Chem. Theory Comput.* **11**, 2278–2291. <https://doi.org/10.1021/ct5010092> (2015).
- Pluhackova, K., Gahbauer, S., Kranz, F., Wassenaar, T. A. & Bockmann, R. A. Dynamic cholesterol-conditioned dimerization of the G protein coupled chemokine receptor type 4. *PLoS Comput Biol* **12**, e1005169. <https://doi.org/10.1371/journal.pcbi.1005169> (2016).
- Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).
- Bunge, S. *et al.* Molecular analysis and genetic mapping of the rhodopsin gene in families with autosomal dominant retinitis pigmentosa. *Genomics* **17**, 230–233. <https://doi.org/10.1006/geno.1993.1309> (1993).
- Park, J. H. *et al.* Opsin, a structural model for olfactory receptors?. *Angew. Chem. Int. Ed. Engl.* **52**, 11021–11024. <https://doi.org/10.1002/anie.201302374> (2013).
- Lee, J. *et al.* CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413. <https://doi.org/10.1021/acs.jctc.5b00935> (2016).
- Huang, J. *et al.* CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73. <https://doi.org/10.1038/nmeth.4067> (2017).
- Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
- Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

41. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
42. Monticelli, L. *et al.* The MARTINI coarse grained forcefield: extension to proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).
43. Eastman, P. *et al.* OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **13**, e1005659. <https://doi.org/10.1371/journal.pcbi.1005659> (2017).
44. Qiu, D., Shenkin, P. S., Hollinger, F. P. & Still, W. C. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate born radii. *J. Phys. Chem. A* **101**, 3005–3014 (1997).
45. Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273. <https://doi.org/10.1021/ct300400x> (2013).
46. Gohlke, H., Kiel, C. & Case, D. A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RalGDS complexes. *J. Mol. Biol.* **330**, 891–913. [https://doi.org/10.1016/s0022-2836\(03\)00610-7](https://doi.org/10.1016/s0022-2836(03)00610-7) (2003).
47. Zoete, V., Meuwly, M. & Karplus, M. Study of the insulin dimerization: Binding free energy calculations and per-residue free energy decomposition. *Proteins* **61**, 79–93. <https://doi.org/10.1002/prot.20528> (2005).
48. Zoete, V. & Meuwly, M. Importance of individual side chains for the stability of a protein fold: Computational alanine scanning of the insulin monomer. *J. Comput. Chem.* **27**, 1843–1857. <https://doi.org/10.1002/jcc.20512> (2006).
49. Zoete, V. & Michielin, O. Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: Application to the TCR-p-MHC complex. *Proteins* **67**, 1026–1047. <https://doi.org/10.1002/prot.21395> (2007).
50. Zoete, V., Irving, M. B. & Michielin, O. MM-GBSA binding free energy decomposition and T cell receptor engineering. *J. Mol. Recognit.* **23**, 142–152. <https://doi.org/10.1002/jmr.1005> (2010).
51. Brooks, B. R. *et al.* CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614. <https://doi.org/10.1002/jcc.21287> (2009).
52. Doktorova, M. *et al.* Cholesterol promotes protein binding by affecting membrane electrostatics and solvation properties. *Biophys. J.* **113**, 2004–2015. <https://doi.org/10.1016/j.bpj.2017.08.055> (2017).
53. Lee, M. S., Feig, M., Salisbury, F. R. Jr. & Brooks, C. L. 3rd. New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. *J. Comput. Chem.* **24**, 1348–1356. <https://doi.org/10.1002/jcc.10272> (2003).
54. Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127–6129 (1990).
55. Srinivasan, J., Trevathan, M. W., Beroza, P. & Case, D. A. Application of a pairwise generalized Born model to proteins and nucleic acids: Inclusion of salt effects. *Theoret. Chem. Acc.* **101**, 426–434 (1999).
56. Onufriev, A., Bashford, D. & Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **55**, 383–394. <https://doi.org/10.1002/prot.20033> (2004).
57. Tanizaki, S. & Feig, M. A generalized Born formalism for heterogeneous dielectric environments: Application to the implicit modeling of biological membranes. *J. Chem Phys* **122**, 124706. <https://doi.org/10.1063/1.1865992> (2005).
58. Mirjalili, V. & Feig, M. Interactions of amino acid side-chain analogs within membrane environments. *J. Phys. Chem. B* **119**, 2877–2885. <https://doi.org/10.1021/jp511712u> (2015).
59. Mondal, S., Khelashvili, G., Shi, L. & Weinstein, H. The cost of living in the membrane: A case study of hydrophobic mismatch for the multi-segment protein LeuT. *Chem. Phys. Lipids* **169**, 27–38 (2013).
60. Choe, S., Hecht, K. A. & Grabe, M. A continuum method for determining membrane protein insertion energies and the problem of charged residues. *J. Gen. Physiol.* **131**, 563–573 (2008).
61. Ben-Tal, N., Ben-Shaul, A., Nicholls, A. & Honig, B. Free-energy determinants of alpha-helix insertion into lipid bilayers. *Biophys. J.* **70**, 1803–1812 (1996).
62. Doumazane, E. *et al.* A new approach to analyze cell surface protein complexes reveals specific heterodimeric metabotropic glutamate receptors. *FASEB J.* **25**, 66–77. <https://doi.org/10.1096/fj.10-163147> (2011).
63. Ulbrich, M. H. & Isacoff, E. Y. Subunit counting in membrane-bound proteins. *Nat. Methods* **4**, 319–321. <https://doi.org/10.1038/nmeth1024> (2007).
64. Chodera, J. D., Swope, W. C., Pitera, J. W., Seok, C. & Dill, K. A. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.* **3**, 26–41. <https://doi.org/10.1021/ct0502864> (2007).

Acknowledgements

We thank the Meyerson and Andersen laboratories (Weill Cornell Medical College) for a gift of HEK293S GnT1^{-/-} cells and use of a dynamic light scattering instrument, respectively, and Scott Blanchard (St. Jude Children's Research Hospital, Memphis, TN) for the LD555 dye.

Author contributions

G.K., T.R.L., V.Y.A., J.Lev. and A.K.M. conceptualized the project. A.N.P. generated the opsin constructs, expressed, purified and reconstituted the proteins, and performed and analyzed scramblase assays. J. Lee performed ensemble FRET and SiMPull experiments. K.P. performed in vitro FRET experiments. G.K., A.M.P., Z.S. and M.A.C. performed and analyzed MD simulations. A.K.M. oversaw the design and analysis of scramblase reconstitution and in vitro FRET experiments. G.K. oversaw the design and analysis of MD simulation experiments. J.Lev. oversaw the design and analysis of ensemble FRET and SiMPull experiments. G.K. and A.K.M. wrote the first draft of the manuscript. All authors edited the manuscript.

Funding

This work was supported by National Institutes of Health grants R01 EY027969 (GK, VYA, AKM), P30 EY005722 (VYA), R35 GM124731 (J Levitz) and F32 EY029929 (TRL), and a Rohr Family Research Scholar Award (J Levitz). GK is also supported by the HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute of Computational Biomedicine at Weill Cornell Medical College through a gratefully acknowledged support from the 1923 Fund. The following computational resources are gratefully acknowledged: resources of the Oak Ridge Leadership Computing Facility (Summit allocation BIP109) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725; and the computational resources of the David A. Cofrin Center for Biomedical Information in the HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine at Weill Cornell Medical

College. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90039-3>.

Correspondence and requests for materials should be addressed to G.K., J.L. or A.K.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021