Fully activated structure of the Smoothened GPCR-Gi protein ligand complex

Amy-Doan P. Vo[1], Soo-Kyung Kim[1], Moon Young Yang[1], Alison E. Ondrus[2,*], William A. Goddard III[1,*]

1. Materials and Process Simulation Center, Division of Chemistry & Chemical Engineering, California Institute of Technology, Pasadena, CA, USA 91125
2. Departments of Chemistry and Pharmaceutical Sciences, University of Illinois Chicago, Chicago, IL, USA 60607

*To whom correspondence should be addressed: aondrus@uic.edu; wag@caltech.edu

ORCID: ADPV (0000-0002-6697-2713); SKK (0000-0002-4498-5441); MYY (0000-0003-4436-8010); AEO (0000-0002-6023-3290); WAG (0000-0003-0097-5716)

Supplementary information

Details of MD simulations

To obtain an equilibrated system for our 370 ns simulation of SMO-CO1-cholesterol-Gi protein, we performed the following steps.

Step 1. An initial 60 steps of steepest descent energy minimization, converging to a maximum force of less than 1000 kJ mol^{-1} nm^{-1}.

Step 2. 250 ps of canonical isothermal-isochoric (NVT) simulations at 310 K, with a 1 fs time step (applied in all subsequent simulations). Three separate groups were coupled to a Berendsen thermostat with a temperature constant of 1 ps: water and ions, POPC membrane, and protein-ligand. Positional restraints for heavy atoms in the protein backbone, side chains, and POPC, as well as dihedral restraints, were applied then relaxed in 125-ps intervals. The force constants are listed below, in kJ mol^{-1} nm^{-2}:

- Protein backbone: 4000 → 2000
- Protein side chains: 2000 → 1000
- POPC z-coordinate: 1000 → 400
- Dihedrals: 1000 → 400

Step 3. 375 ps of isothermal-isobaric (NPT) simulations. Pressure was kept at 1 bar using a semi-isotropic Berendsen barostat with a pressure constant of 5 ps. Positional and dihedral restraints were relaxed in 125 ps intervals. The force constants are listed below, in kJ mol^{-1} nm^{-2}:

- Protein backbone: 1000 → 1000 → 1000
- Protein side chains: 500 → 200 → 50
- POPC z-coordinate: 400 → 200 → 40
- Dihedrals: 200 → 200 → 100

Step 4. Equilibration for 20 ns: relaxation of backbone restraints, imposition of harmonic restraints on specific interactions found via pre-MD *in vacuo* calculations (hydrogen bonding with ligands or NAG, salt bridges with G protein). At this stage, only positional restraints for the protein backbone remained. Backbone restraints were held at a force constant of 1000 kJ mol^{-1} nm^{-2} for 5 ns, then gradually relaxed in 1-ns intervals. At the 10 ns mark, no backbone restraints remained. The force constants are listed below, in kJ mol^{-1} nm^{-2}:

- 500 → 200 → 100 → 50 → 20 → 0

Simultaneously, using PLUMED (ref. S1), harmonic restraints were imposed for 20 ns upon the following interactions with a force constant of 500 kJ mol^{-1} nm^{-2}.

(a) Hydrogenated CO1 (O27) – $N521^{7.41}$ (ND2), with distance 3.0 Å
(b) Hydrogenated CO1 (O24) – $Q477^{6.58}$ (NE2), with distance 3.0 Å
(c) Hydrogenated Cholesterol (O2) – $D95^{CRD}$ (CG), with distance 3.5 Å
(d) Hydrogenated Cholesterol (O2) – $K105^{CRD}$ (NZ), with distance 3.0 Å
(e) $D95^{CRD}$ (CG) – $K105^{CRD}$ (NZ), with distance 4.0 Å
(f) $R257^{12.48}$ (CZ) – $D312^{G\beta}$ (CG), with distance 4.5 Å
(g) $D350^{G\alpha i5}$ (CG) – $R261^{2.38}$ (CZ), with distance 4.5 Å
(h) $K440^{6.21}$ (NZ) – $D341^{G\alpha i5}$ (CG), with distance 4.0 Å
(i) $K440^{6.21}$ (NZ) – $E318^{G\alpha i\text{-}\beta5/\beta6\ loop}$ (CD), with distance 4.0 Å
(j) $E160^{CRD}$ (CD) – NAG at $N493^{6.74}$ (O4), with distance 3.5 Å
(k) $R161^{CRD}$ (CZ) – NAG at $N493^{6.74}$ (O3), with distance 3.5 Å

Step 5. Equilibration for 70 ns: simulations were briefly run for 5 ns (from the 20 to 25 ns mark) with no positional or distance restraints, then the Step 4 distance restraints were re-imposed for 5 ns. In addition, new distance restraints were imposed using PLUMED on various salt bridge interactions at the CRD-TMD interface. All distance restraints were imposed with a force constant of 500 kJ mol^{-1} nm^{-2}.

(l) $E211^{LD}$ (CD) – $R117^{CRD}$ (CZ), with distance 0.45 Å

(m) $D209^{LD}$ (CG) – $R159^{CRD}$ (CZ), with distance 0.45 Å

(n) $E158^{CRD}$ (CD) – $R151^{CRD}$ (CZ), with distance 0.45 Å

Simulations were run for 40 ns without restraints. At the 70-75 ns marks, distance restraints (h)-(i) were re-imposed. Simulations were run for 10 ns without restraints. At the 85-90 ns mark, distance restraint (e) was re-imposed.

Step 6. Production MD. Simulations were run without restraints until the 240 ns mark.

Step 7. Replacement of TMD ligand with CO1 and CRD ligand with cholesterol, energy minimization. Steps detailed in Methods.

Step 8. Production MD. Simulations were run without restraints for 130 ns, for a total simulation length of ~370 ns.

Discussion of repeat and control simulations

This work includes the results from 4 different simulations:

(a) Two independent simulations, for 370 ns and 300 ns, of the full SMO-CO1-cholesterol-Gi protein complex.

(b) Two independent simulations, each for 120 ns, of SMO-CO1-cholesterol without the Gi protein.

Based on our full-length 370 ns MD simulation of the full SMO-CO1-cholesterol-Gi protein complex (**a1**), we propose a relationship between degree of SMO activation and movement of the CRD. Specifically, our MD simulations show that the CRD makes a “bent” angle with respect to the TMD in inactive SMO but a more “upright” angle in active SMO. Compared to the starting structure, our 370 ns simulation **a1** reveals a CRD movement of 10.7 Å toward the “upright” conformation at 369 ns, as measured from the center of mass. In our repeat simulation of the SMO-CO1-cholesterol-Gi protein complex (**a2**), the CRD also moves by 19.7 Å toward the upright conformation by the 300 ns mark. By contrast, in both of our two simulations SMO-CO1-cholesterol complex without the Gi protein (**b1** and **b2**), the CRD tilts toward the bent direction, moving by 15.9 Å in one case and 9.2 Å in the other. These results strengthen our proposal that the CRD moves upright relative to the TMD during SMO-Gαi activation.

Regarding movement of the CO1 ligand, our repeat simulation of SMO-CO1-cholesterol-Gi protein complex (**a2**) is consistent with our original simulation. CO1 moves 7.1 Å downward from the starting structure, as measured from the center of mass. By contrast, our two simulations of

the SMO-CO1-cholesterol complex without the Gi protein (**b1 and b2**) give differing results for movement of the CO1 ligand. In simulation **b1**, the CO1 ligand moves slightly upward by 2.08 Å within the TMD and the cytoplasmic end of TM6 moves 2.5 Å toward the inactive position. In simulation **b2**, the CO1 moves 6.6 Å downward in the TMD while the cytoplasmic end of TM6 moves 1.5 Å toward the active position. These results show that SMO in **b1** seems to be reverting to an inactive structure, whereas SMO **b2** appears trapped in a partially active starting configuration. We propose that an energy barrier between the partially active and inactive states, which is attenuated in the presence of a G protein, may be causing this behavior.

Supplementary references

1. Tribello G. A., Bonomi M., Branduardi D., Camilloni C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **185**, 604 (2014).

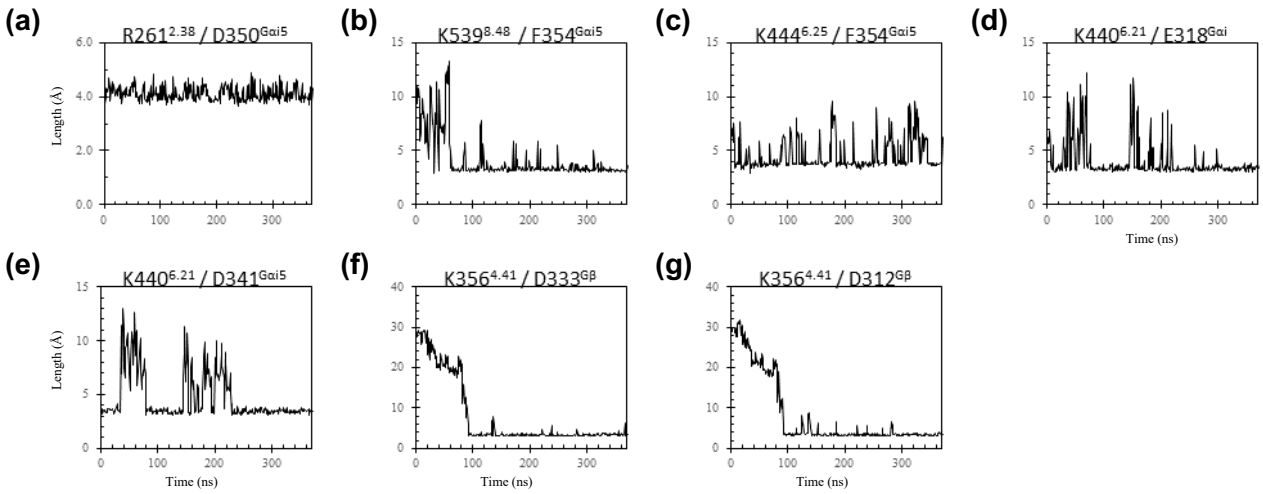

Figure S1. Distance over time between anchor point residues in the SMO-ligand-Gi complex. R-D distance is measured between the atoms CZ and CG respectively, K-F between NZ and C, K-E between NZ and CD, and K-D between NZ and CG.

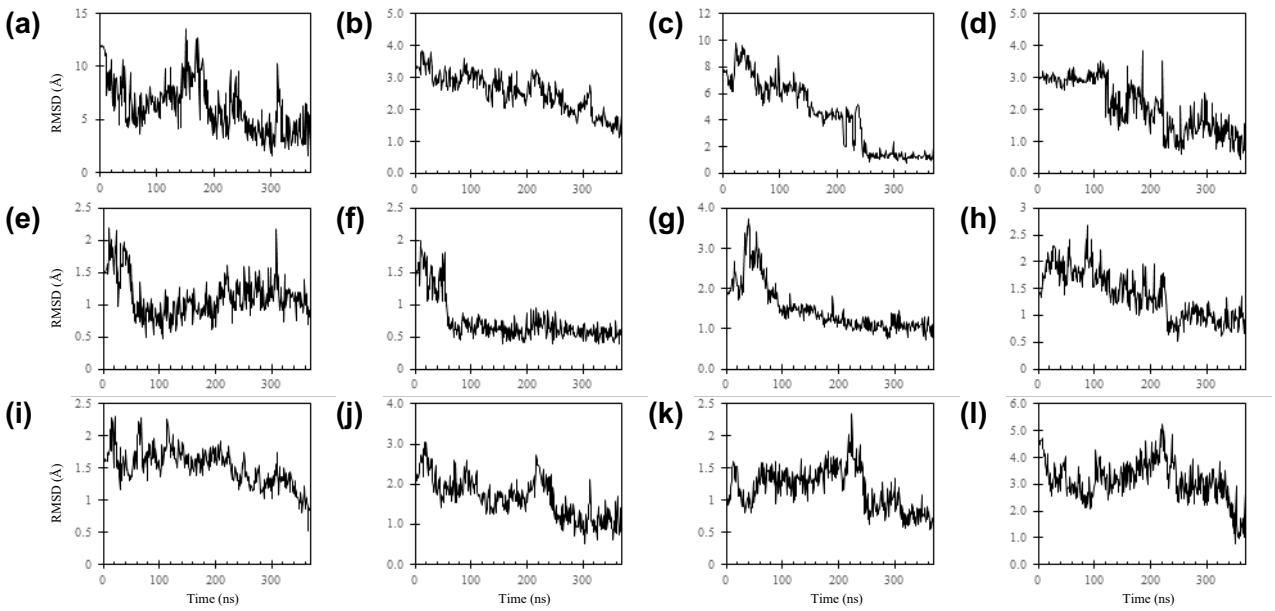

Figure S2. RMSD variation over the course of MD of the SMO-ligand-Gi complex. Unless indicated, RMSDs are measured with structures aligned to the TMD backbone. **(a)** CRD **(b)** TMD **(c)** Cholesterol ligand at the CRD (RMSD is measured with the structure aligned to the CRD backbone). **(d)** CO1 ligand at the TMD. **(e–l)** TMD helices 1-8.

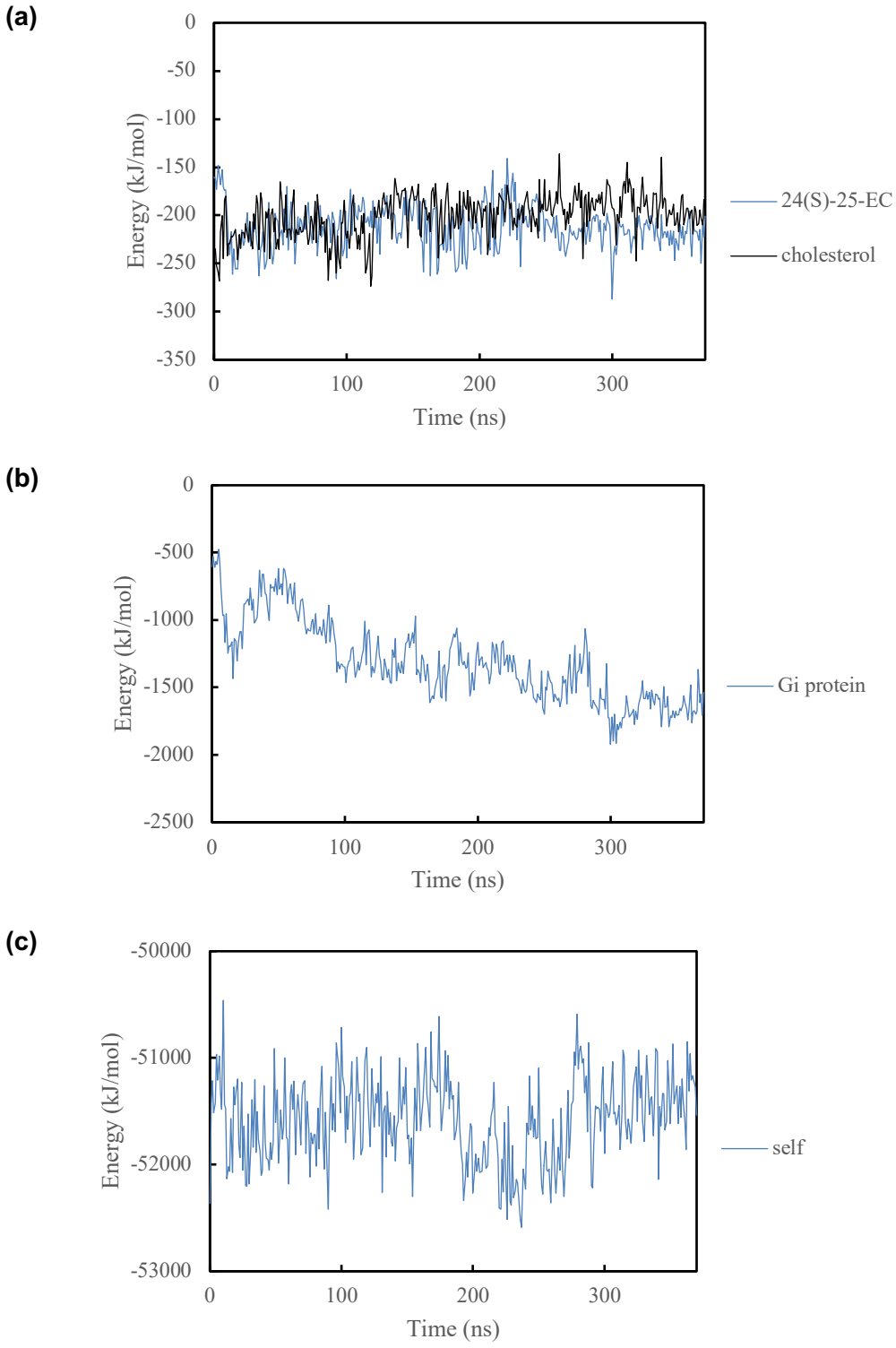

Figure S3. Interaction energy between SMO the SMO-ligand-Gi complex. Interaction energy between SMO and **(a)** sterol ligands **(b)** Gi protein **(c)** self.

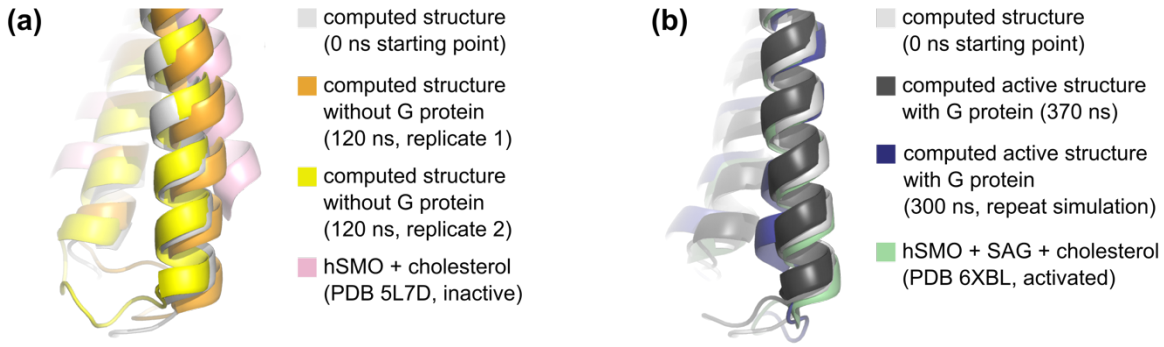

Figure S4. Comparison of TM6 position in active and inactive structures. (a) Computed hSMO structures from 2 independent simulations without the Gi protein (0 ns, light gray; 120 ns replicate 1, orange; 120 ns replicate 2, yellow) and an inactive experimental hSMO structure (PDB 5L7D, light pink). The distances between TM6 in the 0 ns structure and TM6 in replicate 1, replicate 2, and PDB 5L7D, respectively, are 2.5 Å, 1.5 Å, and 6.0 Å. **(b)** Active computed hSMO structures from two independent simulations including the Gi protein (0 ns, light gray; 370 ns simulation, dark gray; 300 ns simulation, navy) and an activated experimental hSMO structure (PDB 6XBL, green). The distances between TM6 in the 0 ns structure and TM6 in the 370 ns structure, the 300-ns structure, and PDB 6XBL, respectively, are 1.4 Å, 3.2 Å, and 1.1 Å.

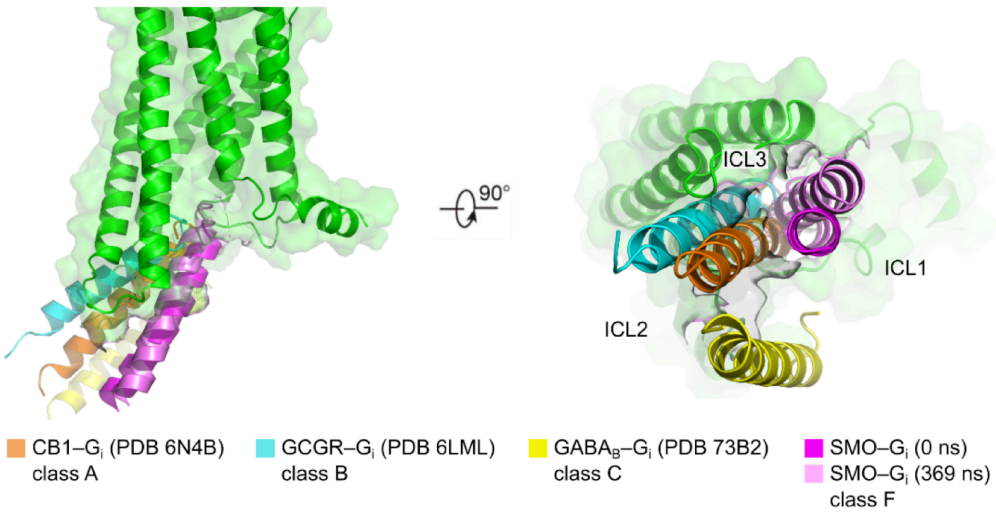

Figure S5. Comparison of the Gαi5 helix position across different GPCR classes. The 370 ns equilibrated SMO-ligand-Gi complex is shown in green. Figure adapted from Shen et al. (ref. 22).

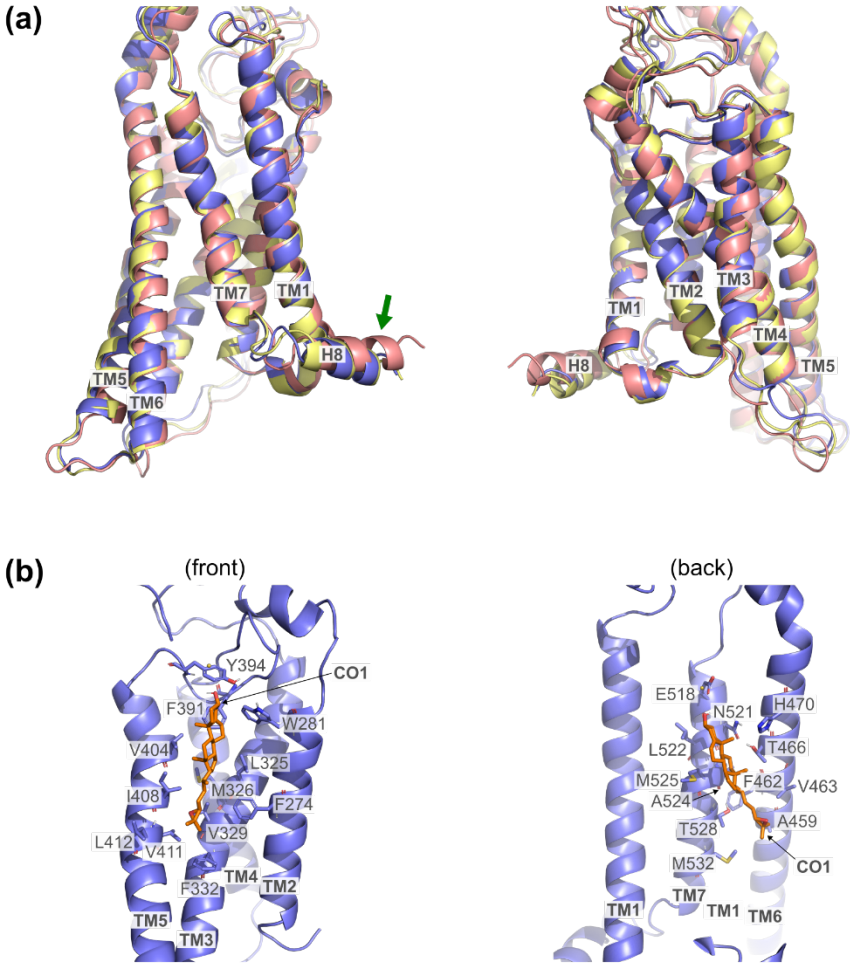

Figure S6. TMD structure and binding interactions with 24(*S*),25-epoxycholesterol (CO1) in the equilibrated MD structure of the SMO-ligand-Gi complex. (a) Comparison between MD time frames after equilibration (344 ns, red; 354 ns, yellow; 364 ns, blue). **(b)** Interactions between CO1 and residues within 4 Å in the TMD binding site at 364 ns.

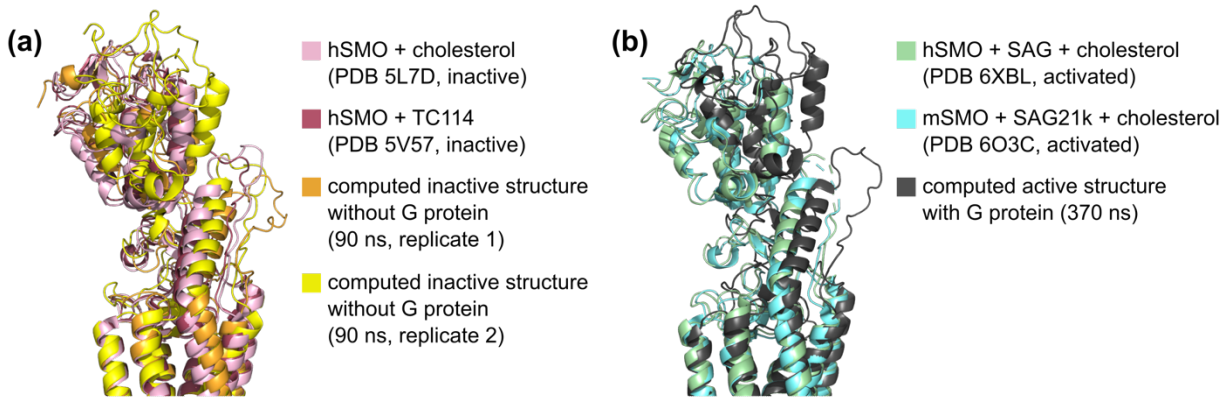

Figure S7. Movement of the CRD relative to the TMD. (a) Inactive experimental hSMO structures (PDB 5L7D, light pink; PDB 5V57, dark pink) and computed hSMO structures without the Gi protein (120 ns; orange, yellow). **(b)** Activated experimental structures: hSMO (PDB 6XBL, green), mSMO (PDB 6O3C, blue). Active computed hSMO structures from 2 independent simulations (370 ns simulation, dark gray; 300 ns simulation, navy).