Supporting Information (SI) for

Conserved Residues Control Activation of Mammalian G Protein-Coupled Odorant Receptors

Claire A. de March[1]#, Yiqun Yu[2]#, Mengjue J. Ni[3], Kaylin A. Adipietro[3], Hiroaki Matsunami[3]*, Minghong Ma[2]*, Jérôme Golebiowski[1]*.

1 Institute of Chemistry - Nice, UMR 7272 CNRS - University Nice - Sophia Antipolis, 06108 Nice cedex 2, France
2 Department of Neuroscience, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA
3 Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA

jerome.golebiowski@unice.fr

hiroaki.matsunami@duke.edu

minghong@mail.med.upenn.edu

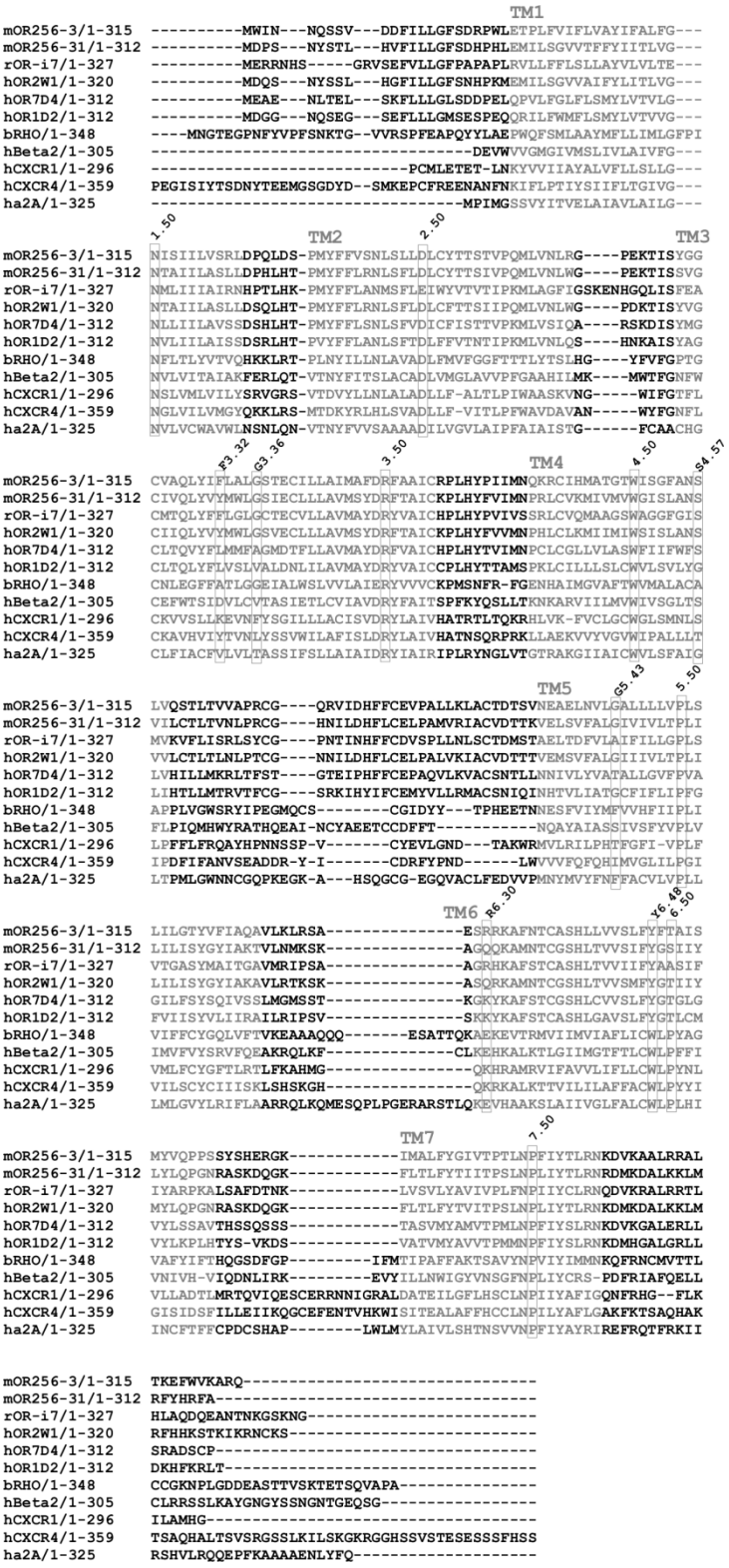

Figure S1. Alignment of MOR256-3 with several human and mouse OR sequences and the GPCR sequences used as templates for homology modeling. Transmembrane (TM) domains are highlighted in gray and the TM number is indicated. The most highly conserved residue in ORs is used as number '50' in the Ballesteros-Weinstein numbering. This residue is boxed for each TM.

Table S1. Most represented residue conservation within human and mouse OR sequences

Residue number	Human OR residue - conservation	Mouse OR residue - conservation
$108^{3.36}$	G – 47%	G – 44%
$108^{3.36}$	G/A/V/S/T – 88%	G/A/V/S/T – 85%
$121^{3.49}$	D – 98%	D – 99%
$122^{3.50}$	R – 88%	R – 97%
$234^{6.30}$	R/K – 75%	R/K – 77%
$252^{6.48}$	Y/F – 93%	Y/F – 92%

.

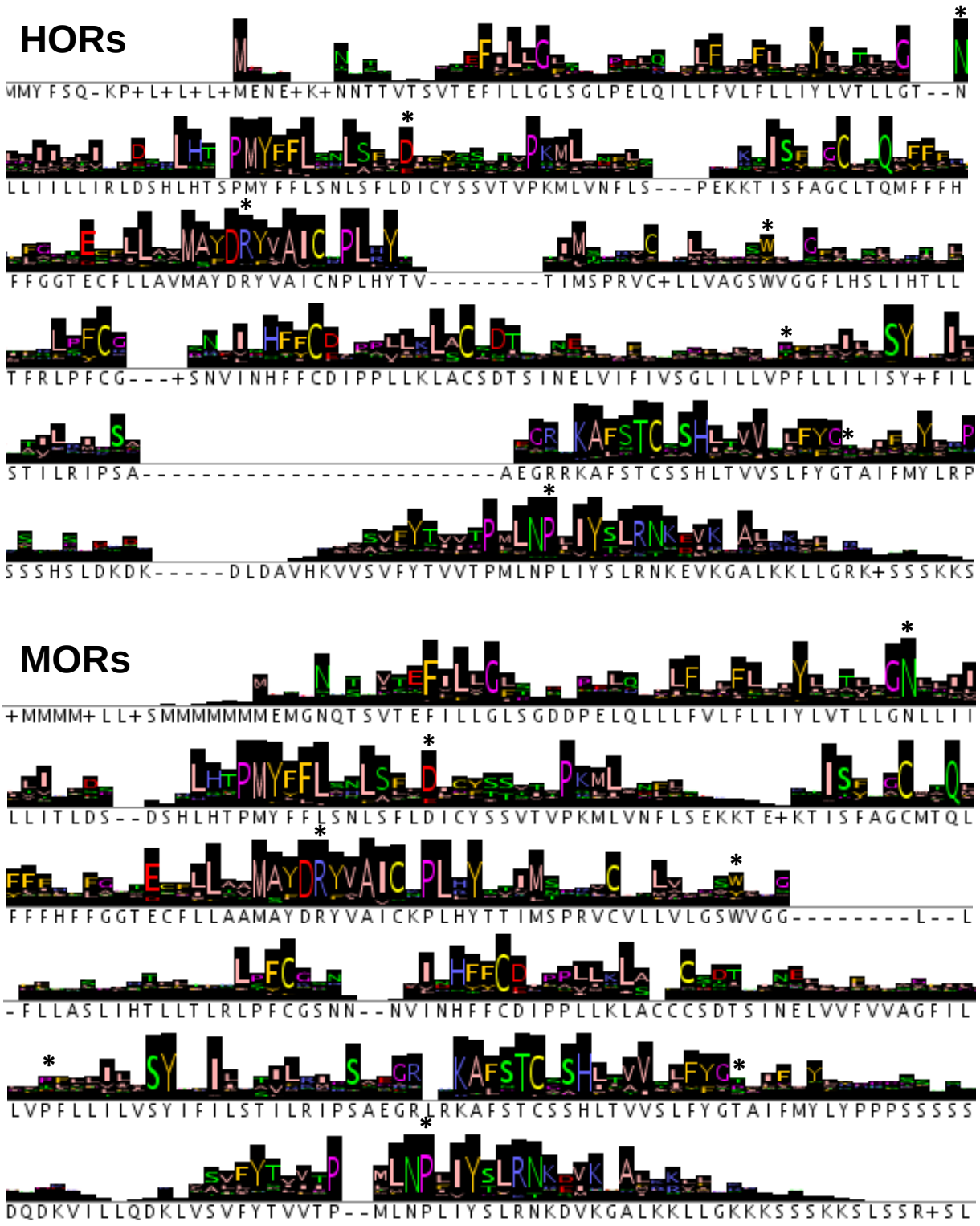

Figure S2. Logo representation of the human and mouse OR conservation. The size of the letter is proportional to the residue conservation. A star (*) highlights the residue used for the Ballesteros-Weinstein numbering within each transmembrane domain. This representation was obtained with Jalview software (http://www.jalview.org/).

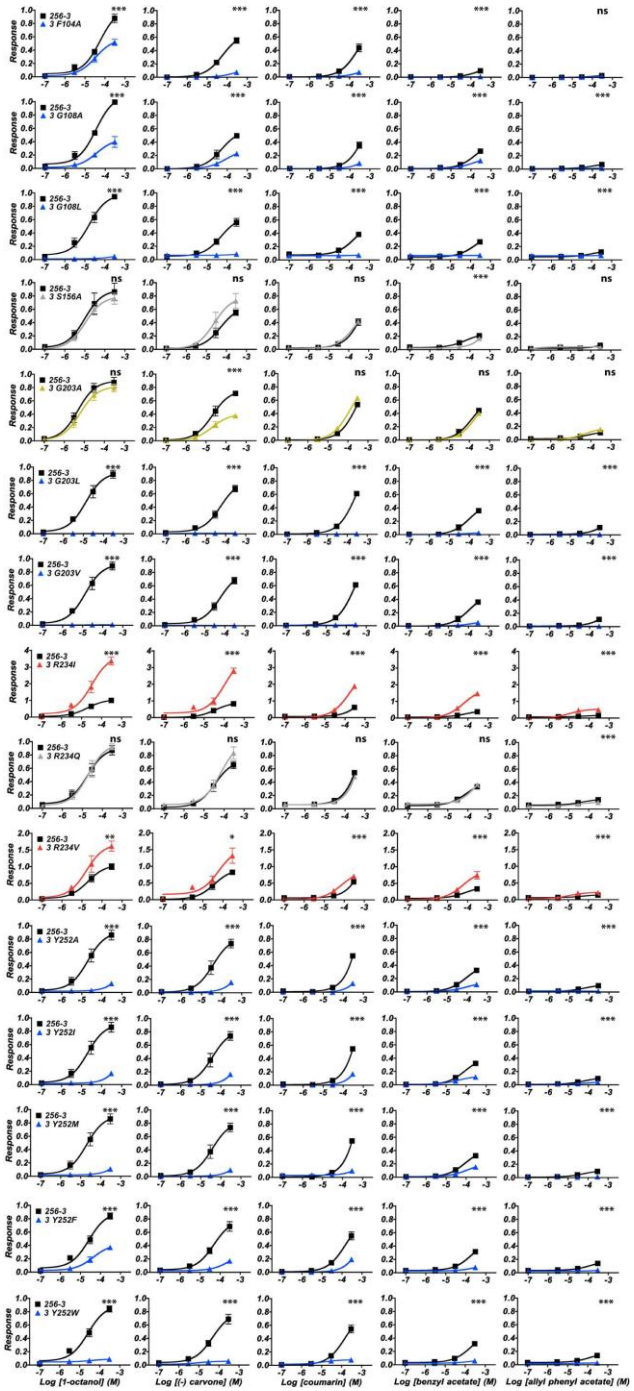

Figure S3. The dose-response curves for *wt* and mutant MOR256-3. Mutant ORs showed decreased (blue), increased (red), or unchanged (gray) responses to odorant. G203A mutation (in yellow) altered ligand selectivity. Each mutant OR was tested on the same plate as the *wt* (three repeats for each OR) and all activities were normalized to the maximum *wt* response to 1-octanol at 300μM. Two-way ANOVA tests (concentration and OR type) were performed for each mutant and *wt* pair (ns = not significantly different, * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$ for OR type).

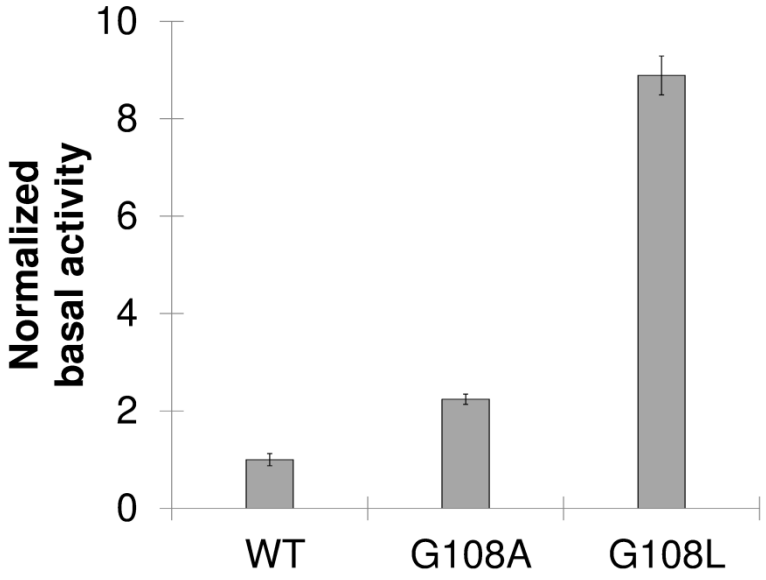

Figure S4. Basal activity of mOR256-31 mutants, normalized to the *wt*. The data are represented as mean ± s.e.m. (n = 15, 5, and 4 for wt, G108A, and G108L, respectively).

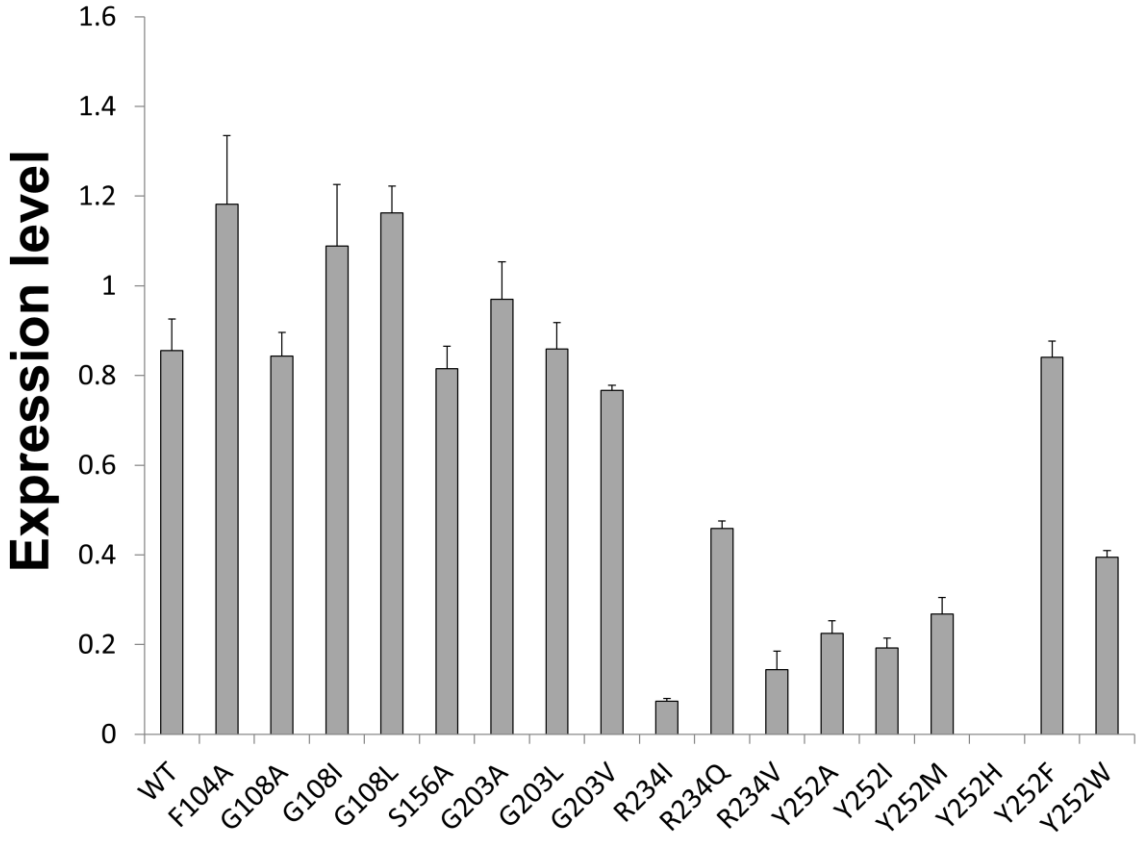

Figure S5. Expression level of mOR256-3 *wt* and mutants, expressed as ratio of Rho+/GFP+ cells. Each data point is averaged from 3-5 plates (mean ± s.e.m.).

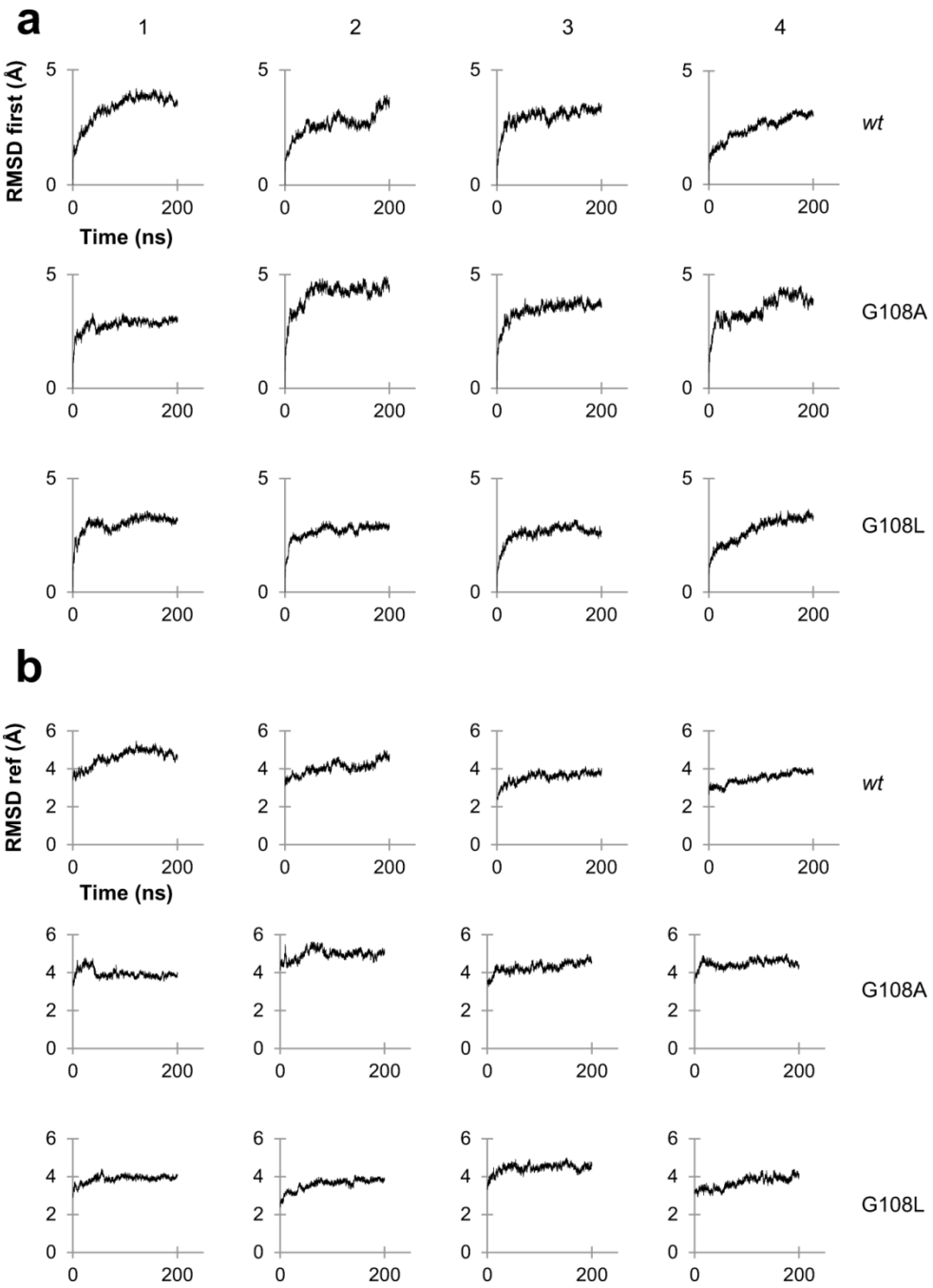

Figure S6. OR structure Root Mean Square deviation as a function of time (expressed in ns) computed for all the molecular dynamic simulations using the first frame of the production period (a) or the structure taken from Modeller (b) as a reference. The RMSD (in Å) is computed during the 200 ns of each molecular dynamics simulation on the CA atoms of the bundle (from residue 19 to 309 on each system: *wt*, G108A and G108L).

General Methods and Materials

Site-Directed Mutagenesis

The coding sequences of MOR256-3 and 256-31 were amplified from genomic DNA of C57BL/6 mice and subcloned into the pcDNA3.1/TOPO vector (Invitrogen) with an N-terminal tag of the first 20 amino acids of rhodopsin. Site-directed mutants were constructed using the Quikchange site-directed mutagenesis kit (Agilent Technologies). The sequences of all plasmid constructions were verified by both forward and reverse sequencing (DNA sequencing core facility, University of Pennsylvania).

Evaluation of OR surface expression

Live-cell immunostaining is used to evaluate OR surface expression.[1] Hana3A cells were co-transfected with the receptor and GFP plasmids 24 hours before the staining. The transfected Hana3A cells were incubated with the primary antibody solution (mouse anti-rhodopsin, Rho 4D2, Abcam) on ice for 1 h. After rinsing the cells for three times, the secondary antibody solution (Alexa Fluor 568-conjucated anti-mouse IgG) was added onto the cells, and incubated for 45 min on ice. At the end of the incubation, the cells were fixed with 2% Paraformaldehyde, and mounted with vectashield mounting medium (Vector Laboratories, Inc.). The ratio of Rho^+ cells/GFP^+ cells is used to evaluate the surface expression of each OR construct.

Luciferase assay in Hana3A cells

The Dual-Glo Luciferase Assay (Promega) was used to determine the activities of firefly and Renilla luciferase in Hana3A cells [1]. Firefly luciferase, driven by a cAMP response element promoter (CRE-Luc; Stratagene), was used to determine OR activation levels. Renilla luciferase, driven by a constitutively active SV40 promoter (pRLSV40; Promega), functioned as an internal control for transfection efficiency and cell viability. Hana3A cells stably expressing RTP1L, RTP2, REEP1, and $G_{\square olf}$ were plated on poly-D-lysine-coated 96-well plates (Nalge Nunc) and incubated overnight in minimum essential medium eagle (Sigma) with 10% FBS at 37°C and5%CO_2. The following day, cells were transfected using Lipofectamine 2000 (Invitrogen). For each 96-well plate, 1 µg pRL-SV40, 1 µg CRE-Luc, 1 µg mouse RTP1s, and 6 µg of receptor plasmid DNA were transfected. After transfection (24 h), medium was replaced with 25 µl of odorant solution diluted in CD293 chemically defined medium (Invitrogen), and cells were further incubated for 4 h at 37°C and 5% CO_2. The manufacturer's protocols were followed to measure firefly luciferase and Renilla luciferase activities. A Wallac Victor 1420 plate reader (Perkin-Elmer) was used to measure luminescence. Data were analyzed using Microsoft Excel and GraphPad Prism. Normalized activity was further calculated using the following formula: [Luc/RLuc(N)-Luc/RLuc(lowest)]/[Luc/RLuc(highest)-Luc/RLuc(lowest)],

where Luc/RLuc(N) = luminescence of firefly luciferase divided by luminescence of Renilla luciferase in a certain well; Luc/RLuc(lowest) = lowest firefly luminescence divided by Renilla luminescence of a plate or set of plates; Luc/RLuc(highest) = highest firefly luminescence divided by Renilla luminescence of a plate. To facilitate comparison between OR responses from multiple plates, the Rho-tag empty vector and *wt* MOR256-3 were always included as negative and positive control, respectively. The basal activity of an OR was averaged from four wells in the absence of odorants and further corrected by subtracting that of the control empty vector. An odorant-induced activity was averaged from at least three wells and further corrected by subtracting the basal activity of that receptor. All odorant-induced activities were normalized to *wt* MOR256-3 response to 300 μM 1-octanol. Both basal activity and odorant-induced responses were corrected for the surface expression ratio (Rho+/GFP+ when Hana3A cells were co-transfected with a Rho-tagged OR and GFP) normalized to that of *wt*.

Molecular modeling

Model building. The protocol follows a previously published method[2]. Sequences of MOR256-3, -8, -17, -22 and -31, I7 (olfr2), mOR-EG (olfr73), and S25 (olfr480) are aligned with 396 human ORs[3] and nine sequences of X-ray elucidated GPCRs: bovine rhodopsin (PDB: 1U19)[4], human beta 2 adrenergic (PDB: 2RH1)[5], turkey beta 1 adrenergic (PDB: 2VT4)[6], human chemokine receptors CXCR4 (PDB: 3ODU)[7] and CXCR1 (PDB: 2LNL)[8], human dopamine receptor D3 (PDB: 3PBL)[9], human adenosine a2A receptor (PDB: 2YDV)[10], human histamine H1 receptor (PDB: 3RZE)[11] and muscarinic acetylcholine receptor M2 (PDB: 3UON)[12]. Highly conserved motifs in ORs are considered as constraints for the alignment: GN in helix 1, PMYFFLXXLSXXD in helix 2, MAYDRYXAICXPLXY in helix 3, SYXXI in helix 5, KAFSTCASH in helix 6, LNPXIY in helix 7 and a pair of conserved cysteines $97^{3.25}$-$179^{4.80}$ which constitute a known disulfide bridge between the beginning of helix 3 and the extracellular loop 2. Four experimental GPCR structures (1U19, 3ODU, 2YDV and 2LNL) are selected as templates to build MOR256-3 and its G108A and G108L mutants by homology modeling with Modeller.[13] The N-terminal structure is omitted to avoid perturbing the modeling protocol. Five models are obtained and the one fulfilling several constraints (binding cavity sufficiently large, no large folded structure in extra-cellular loops, all TMs folded as α-helices, a small α-helix structure between TM3 and TM4) is kept for further molecular dynamics simulations.

Molecular dynamics simulations. The *wt*, G108A and G108L mutants are embedded in a model membrane made-up of POPC lipids solvated by TIP3P water molecules using Maestro.[14] The total system is made up of ~48,650 atoms in a periodic box of 91*89*98 $Å^3$.

Molecular dynamics simulations are performed with sander and pmemd.cuda modules of AMBER12 with the ff03 force-field for the protein and the gaff.lipid for the membrane. Hydrogen atoms bond are constrained by SHAKE algorithm and long-range electrostatics interactions are handled with Particle

Mesh Ewald (PME). The cutoff for non-bonded interactions is set at 8 Å. Temperature and pressure are maintained constant with a Langevin thermostat with a collision frequency of 2 ps^{-1}. In addition, a weak coupling anisotropic algorithm with a relaxation time of 1 ps^{-1} is applied. Snapshots are saved every 20 ps.

Two energy minimizations are performed during 10,000 steps with the 5,000 first steps using a conjugate gradient algorithm. The first one is run with a restraint of 200 $kcal.mol^{-1}$ applied on all atoms of the membrane and water and the second one with the same restraint on all atoms of the receptor. This last constraint is kept for the heating phase of 20 ps (NTP, 100K to 310K, Langevin thermostat with collision frequency of 5 ps^{-1}) and equilibration of 15 ns (NTP, 310K). Restraints are then reduced by 5 $kcal.mol^{-1}Å^{-2}$ and another cycle of minimization-equilibration is performed. The systems (*wt*, G108A and G108L mutants) are replicated four times and 200 ns-long production molecular dynamics are performed after an equilibration period of 50 ns.

(1) Zhuang, H.; Matsunami, H. Nat Protoc 2008, 3, 1402.
(2) Charlier, L.; Topin, J.; de March, C. A.; Lai, P. C.; Crasto, C. J.; Golebiowski, J., In Olfactory Receptors, Crasto, C. J., Ed. Humana Press: New York, 2013.
(3) Zozulya, S.; Echeverri, F.; Nguyen, T. Genome Biol 2001, 2, 1.
(4) Okada, T.; Sugihara, M.; Bondar, A. N.; Elstner, M.; Entel, P.; Buss, V. J Mol Biol 2004, 342, 571.
(5) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. Science 2007, 318, 1258.
(6) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G.; Tate, C. G.; Schertler, G. F. Nature 2008, 454, 486.
(7) Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Science 2010, 330, 1066.
(8) Park, S. H.; Das, B. B.; Casagrande, F.; Tian, Y.; Nothnagel, H. J.; Chu, M.; Kiefer, H.; Maier, K.; De Angelis, A. A.; Marassi, F. M.; Opella, S. J. Nature 2012, 491, 779.
(9) Chien, E. Y.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G. W.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V.; Stevens, R. C. Science 2010, 330, 1091.
(10) Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G.; Tate, C. G. Nature 2011, 474, 521.
(11) Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; Kobayashi, T.; Stevens, R. C.; Iwata, S. Nature 2011, 475, 65.
(12) Haga, K.; Kruse, A. C.; Asada, H.; Yurugi-Kobayashi, T.; Shiroishi, M.; Zhang, C.; Weis, W. I.; Okada, T.; Kobilka, B. K.; Haga, T.; Kobayashi, T. Nature 2012, 482, 547.
(13) Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.-y.; Pieper, U.; Sali, A., In Current Protocols in Bioinformatics, Bateman, A.; Draghici, S.; Pearson, W. R.; Stein, L. D.; Yates, J. R. I.; , Eds. John Wiley & Sons, Inc.: New York, 2006.
(14) Schrödinger 1: Maestro, version 9.4, 2013.