Supporting Information for

QM/MM Model of the Mouse Olfactory Receptor MOR244-3 Validated by Site-Directed Mutagenesis Experiments

Sivakumar Sekharan,[†,*] Mehmed Z. Ertem,[†,#,*] Hanyi Zhuang,[‡,£,*] Eric Block,[¶] Hiroaki Matsunami,[§,$] Ruina Zhang,[‡] Jennifer N. Wei,[†] Yi Pan,[‡] and Victor S. Batista[†,*]

[†]Department of Chemistry, Yale University, P.O. Box 208107, New Haven, CT 06520-8107, United States; [#]Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973-5000, United States; [‡]Department of Pathophysiology, Key Laboratory of Cell Differentiation and Apoptosis of the Chinese Ministry of Education, Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China 200025; [£]Institute of Health Sciences, Shanghai Jiao Tong University School of Medicine/Shanghai Institutes for Biological Sciences of Chinese Academy of Sciences, Shanghai, P. R. China 200031; [¶]Department of Chemistry, University at Albany, State University of New York, Albany, NY 12222, United States; Departments of [§]Molecular Genetics and Microbiology and [$]Neurobiology, Duke University Medical Center, Durham, NC 27710, United States.

Corresponding Authors

sivakumar.sekharan@yale.edu; mzertem@bnl.gov; hanyizhuang@sjtu.edu.cn; victor.batista@yale.edu

Total Number of Pages: S13 pages

Table of Contents

I. Experimental Details

Hana3A cells were grown in Minimum Essential Medium (Hyclone) containing 10% fetal bovine serum (Invitrogen) at 37 °C with 5% CO_2. Lipofectamine 2000 (Invitrogen) was used for transfection. Luciferase assays are performed as previously described (32). For GloSensor assay, Hana3A cells were plated onto 96-well plates (Greiner). After 18 to 24 h, OR, mRTP1S, and a GloSensor™ plasmid were transfected into cells. Twenty four hours after transfection, the cells were stimulated with odorants plus various concentrations of metal ions dissolved in HBSS. We used the GloSensor™ (Promega) kit and followed the manufacturer's protocol for measuring chemiluminescence.

II. Computational Details

TMHMM Analysis and Homology Modeling. We performed a sequence alignment between the X-ray structures of the human M2 muscarinic receptor (1), turkey beta-adrenergic receptor (2), and mouse olfactory receptor MOR244-3 using the ClustalW2 software (3) using the default settings and determined the TM domains of MOR244-3, using the transmembrane hidden Markov Model (TMHMM) analysis (4). The extracellular region between transmembrane helix 5 and transmembrane helix 6 of the M2 receptor (residues 218-376) were removed in this alignment, and two spaces were put in its place (Fig. S1). For MOR244-3, the transmembrane domains were determined using the transmembrane hidden Markov Model (TMHMM) analysis. TMHMM is a tool that predicts the transmembrane domains of proteins based on Bayesian analysis of a pool of transmembrane proteins with resolved structures; the residues in MOR244-3 with a posterior TM probability greater than 0.1 were assigned to the transmembrane domain. For muscarinic receptor, the transmembrane domains are the helix domains designated in the corresponding PDB file with guidance from TMHMM analysis.

```
3UON    ------------MNNSTNSSNNSLALTSPYKTFEVVFIVLVAGS------LSLVTIIGNI 42
2Y00    MGDGWLPPDCGPHNRSGGGGATAAPTGSRQVSAELLSQQWEAGMSLLMALVVLLIVAGNV 60
244-3   ---------MGALNQTRVTEFIFLGLTDNWVLEILFFVPFTVTY--------MLTLLGNF 43

3UON    LVMVSIKVNRHLQTVNNYFLFSLACADLIIGVFSMNLYTLYTVIGYWPLGPVVCDLWLAL 102
2Y00    LVIAAIGRTQRLQTLTNLFITSLACADLVMGLLVVPFGATLVVRGTWLWGSFLCECWTSL 120
244-3   LIVVTIVFTPRLHNPMYFFLSNLSFIDICHSSVTVPKMLEGLLLERKTISFDNCIAQLFF 103

3UON    DYVVSNASVMNLLIISFDRYFCVTKPLTYPVKRTTKMAGMMIAAAWV---LSFILWAPAI 159
2Y00    DVLCVTASIETLCVIAIDRYLAITSPFRYQSLMTRARAKVIICTVWA---ISALVSFLPI 177
244-3   LHLFACSEIFLLTIMAYDRYVAICIPLHYSNVMNMKVCVQLVFALWLGGTIHSLVQTFLT 163
```

```
3UON   LFWQFIVG---VRTVEDGECYIQFFSNAAVTFGTAIAAFYLPVIIMTVLYWHISRAS--- 213
2Y00   MMHWWRDEDPQALKCYQDPGCCDFVTNRAYAIASSIISFYIPLLIMIFVYLRVYREA--- 234
244-3  IRLPYCGPNIIDSYFCDVPPVIKLACTDTYLTGILIVSNSGTISLVCFLALVTSYTVILF 223

3UON   -KSRI----------------------------------------PPPSREKKVTRTIL 390
2Y00   -KEQIRKIDRCEGRFYGSQEQPQPPPLPQHQPILGNGRASKRKTSRVMAMREHKALKTLG 293
244-3  SLR-------------------------------------KKSAEGRRKALSTCS 241

3UON   AILLAFIITWAPYNVMVLINTFCAPCIPNTVWTIGYWLCYINSTINPACYALCNATFKKT 450
2Y00   IIMGVFTLCWLPFFLVNIVNVFNRDLVPDWLFVFFNWLGYANSAFNPIIYCRS-PDFRKA 352
244-3  AHFMVVTLFFGPCIFLYTR----PDSSFSIDKVVSVFYTVVTPLLNPLIYTLRNEEVKTA 297

3UON   FKHLLMCHYKNIGATR---------------------------------------- 466
2Y00   FKRLLCFPRKADRRLHAGGQPAPLPGGFISTLGSPEHSPGGTWSDCNGGTRGGSESSLEE 412
244-3  MKHLRQRRICS------------------------------------------- 308
```

Figure S1. Multiple sequence alignment between the human M2 muscarinic receptor, turkey beta-adrenergic receptor and mouse olfactory receptor 244-3. The transmembrane domains of each protein are highlighted in the alignment. The DRY motif in helix 3 of each protein is also highlighted.

In TM1, there is a gap of two residues between the alignment of muscarinic receptor and MOR244-3 compared to a gap of eight residues between the alignment of beta-adrenergic receptor and MOR244-3. The predicted TM domains of the MOR244-3 are typically centered with the TM domains of muscarinic M2 receptor, and often off-centered from those of beta-adrenergic receptor (Fig. S2). Therefore, we built the homology model of MOR244-3 using the X-ray structure of the human M2 muscarinic receptor as a template. The comparative protein modeling by satisfaction of spatial restraints against the available X-ray structures in the literature yielded an expected value (E value) of $2.8*10^{-45}$ for the homology model based on the M2 receptor, which indicates a high sequence identity between the MOR244-3 and M2 receptor transmembrane regions. If the E value is greater than 1, then the structure is not suitable for homology modeling (Fig. S1, S2).

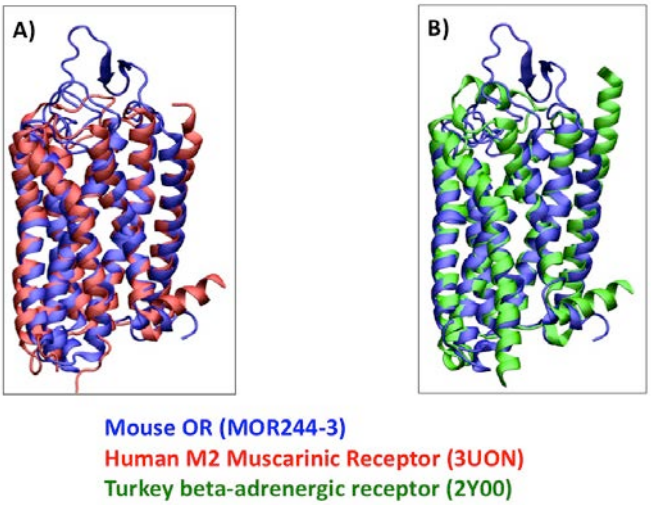


Figure S2. Multisequence Stamp alignment of the seven TM regions of the homology model of MOR244-3 with the human M2 muscarinic receptor (A) with the turkey beta-adrenergic receptor (B).

QM/MM Model System Setup. The initial coordinates of the homology model of MOR244-3 structure was obtained by using the Modeler 9v7 program package (www.salilab.org/modeller) and the X-ray structure of human M2 muscarinic acetylcholine receptor bound to an antagonist (PDB: 3UON at 3.0 Å) as a template. The protonation states of all titratable residues (pH = 7) are assigned using PROPKA calculations (5) implemented in the Schrodinger's Maestro v.9.3 software package (6) and also by visual inspection. It has been shown that water molecules are as important as amino acids to the overall functionality of heptahelical transmembrane proteins (7,8) and therefore, 22 water molecules were added to fill up the cavities in the QM/MM model of MOR244-3. Based on the previous experimental evidence (9), we introduced a copper ion into the active site T203 and the overall charge of the active site (QM region) varies from neutral (*because of the presence of* Cu^{+}/S^{-}_{C109}) to cationic (*due to the presence of* Cu^{2+}/S^{-}_{C109}). To minimize the structural changes and to preserve the natural shape of the protein, the full structure was relaxed via a three-step optimization procedure. First, the structure was fully optimized at pure AMBER96 force field level (10) followed by reoptimization at the two-layer ONIOM scheme with electronic-embedding (EE) (11), as implemented in Gaussian09 (12). The QM layer included the Cu ion, H105, C109, N202 residues and MTMT ligand or its analogues treated at M06-L (13-15) level of theory using the Stuttgart [8s7p6d2f | 6s5p3d2f] ECP10MWB contracted pseudopotential basis set (16) on Cu and the 6-31G(d) (17) basis set on all other atoms. The MM

layer included the rest of the protein, as described by the AMBER96 molecular mechanics force field (10). The interface between QM and MM layers was treated using the standard hydrogen-link atom scheme (18). The ONIOM extrapolated energies for the optimized model active sites for Cu^{+} and Cu^{2+} ions are listed in Table S1. Note that this two-layer QM/MM procedure was previously employed to gain novel insights into the structural and molecular rearrangements of visual and non-visual GPCRs and of its photointermediates (19–29).

Table S1. ONIOM extrapolated energies for QM/MM optimized structures for different active site models

Cu^{+} Structures	**ONIOM Extrapolated Energy (kcal/mol)**	**Cu^{2+} Structures**	**ONIOM Extrapolated Energy (kcal/mol)**
$C109^{SH}(CH_3SCH_2SH)$ (Conformer 1)	-1647199.3	$C109^{S-}$ (CH_3SCH_2SH) (Conformer 1)	-1646811.9
$C109^{SH}(CH_3SCH_2SH)$ (Conformer 2)	-1647211.1	$C109^{S-}(CH_3SCH_2SH)$ (Conformer 2)	-1646826.8
$C109^{SH}(HSCH_2SH)$ (Conformer 1)	-1622531.5	$C109^{S-}(HSCH_2SH)$	-1622147.2
$C109^{SH}(HSCH_2SH)$ (Conformer 2)	-1622531.2	$C109^{S-}(CH_3SCH_2S^{-})$ (Conformer 1)	-1646533.6
$C109^{S-}(CH_3SCH_2SH)$ (Conformer 1)	-1646894.7	$C109^{S-}(CH_3SCH_2S^{-})$ (Conformer 2)	-1646512.1
$C109^{S-}(CH_3SCH_2SH)$ (Conformer 2)	-1646911.8	$C109^{S-}(HSCH_2S^{-})$	-1621868.1
$C109^{S-}(HSCH_2SH)$ (Conformer 1)	-1622239.4		
$C109^{S-}(HSCH_2SH)$ (Conformer 2)	-1622244.2		
$C109^{S-}(CH_3SCH_2S^{-})$ (Conformer 1)	-1646572.2		
$C109^{S-}(CH_3SCH_2S^{-})$ (Conformer 2)	-1646561.6		

$C109^{S-}(HSCH_2S^-)$ (Conformer 1)	-1621906.6		
$C109^{S-}(HSCH_2S^-)$ (Conformer 2)	-1621882.4		

Binding Energy Calculations. Vibrational frequency calculations were carried out on QM/MM optimized structures to obtain zero point energy (ZPE) and thermal contributions to the electronic energies and to compute the enthalpy (H) of the systems (Table S2). Single point calculations were carried out at M06-L level of theory, with SDD basis set on Cu and 6-31G(d) basis on all other atoms, on model complexes built from QM/MM optimized structures (Fig. S4) to predict the free energy of solvation (G_{sol}) values using SMD (30) aqueous continuum solvation model (Table S3). The final enthalpy values for the systems were obtained by adding the G_{sol} to the enthalpy values obtained via QM/MM calculations (Table S2) or DFT calculations on free ligands (Table S4). These final composite enthalpy values were then used to compute the binding enthalpies. Similar QM/MM protocols with thermodynamic cycles were employed in the literature to calculate different contributions to the final binding energies (31).

For instance, the binding enthalpy of MTMT to the binding site is calculated using Tables S2-S4 as follows:

$\Delta H_{sol} = \Delta H_{gas} + \Delta\Delta E_{sol}$ (model)

$\Delta H_{sol} = \Delta H_{gas} + [\Delta E_{sol} - \Delta E_{gas}]$ (model)

ΔH_{sol} = [-1638327.6 – (-1088569.5) – (-549727.5)] + [-1533770.3 – (-983988.7) – (-549782.8)] - [-1533738.7 – (-983952.8) – (-549779.9)]

ΔH_{sol} = -30.6 + 1.2 + 6.0

ΔH_{sol} = -23.4 kcal/mol

For relative binding energy of $HSCH_2SH$ versus MTMT (CH_3SCH_2SH) obtained via Table S1 and Table S4:

ΔE_{sol} = (-1622244.2 + (-549782.8) - (-1646911.8) - (-525115.8)

ΔE_{sol} = 0.6 kcal/mol disfavoring binding of $HSCH_2SH$

Table S2. ONIOM extrapolated energies for QM/MM optimized structures and enthalpies for different active site models (see Computational Methods for details)

Model	ONIOM Extrapolated Energy (kcal/mol)	Enthalpy (H) (kcal/mol)
$C109^{S-}Cu^{+}$	-1097100.7	-1088569.5
$C109^{S-}(CH_3SCH_2SH)Cu^{+}$ (Conformer 2)	-1646911.8	-1638327.6

Table S3. Electronic energies (E) and electronic energies including free energy of solvation (E_{sol}) in units of kcal/mol obtained via single point calculations on model structures at M06-L level of theory (see Computational Methods for details)

Model	Energy (kcal/mol)	E_{sol} (kcal/mol)
$C109^{S-}Cu^{+}$	-983952.8	-983988.7
$C109^{S-}(CH_3SCH_2SH)Cu^{+}$ (Conformer 2)	-1533738.7	-1533770.3

Table S4. Electronic energies (E), enthalpies (H) and electronic energies including free energy of solvation (E_{sol}) in units of kcal/mol for ligands optimized at M06-L level of theory (see Computational Methods for details)

Model	Energy (kcal/mol)	Enthalpy (kcal/mol)	E_{sol} (kcal/mol)
CH_3SCH_2SH	-549779.9	-549727.5	-549782.8
$HSCH_2SH$	-525112.2	-525079.2	-525115.8

III. Other Supporting Figures

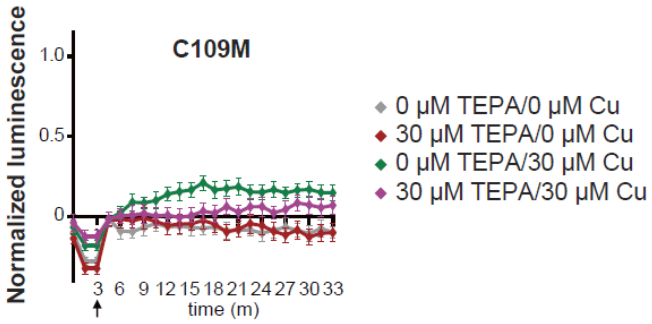
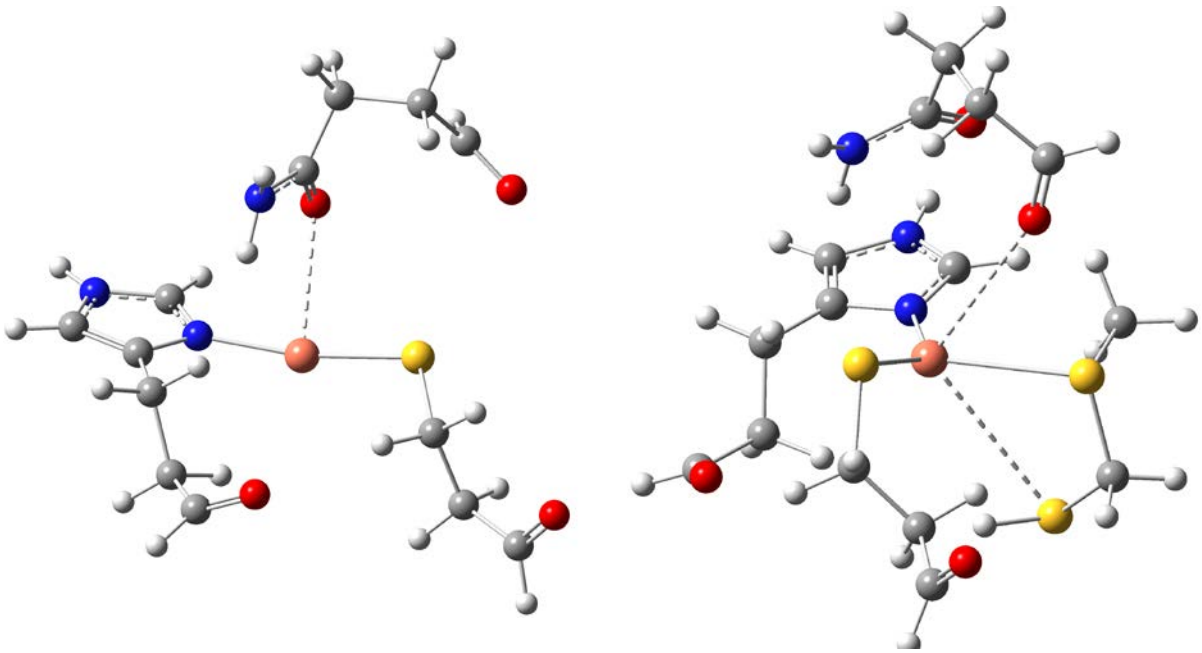


Figure S3. Real-time measurement of the activity of the C109M mutant in the presence of 30 μM Cu^{2+} and/or 30 μM tetraethylenepentamine (TEPA) using the GloSensor™ assay. The arrow indicates the time point of odorant addition. The *y*-axis represents mean luminescence ±SEM, normalized to the response of wild type MOR244-3 to 30 μM Cu^{2+} and 30 μM MTMT (N = 6).



(a) (b)

Figure S4. Active site models of MOR244-3 without ligand (a) and with MTMT ligand (b) employed to compute the free energy of solvation contributions.

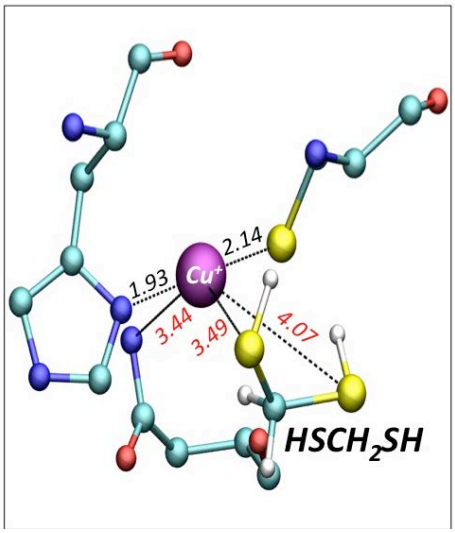


Figure S5. QM/MM optimized model of methanedithiol ($HSCH_2SH$) and Cu^+ in the active site. The values in black represent the bonding distances, while the values in red correspond to the distances between two atoms.

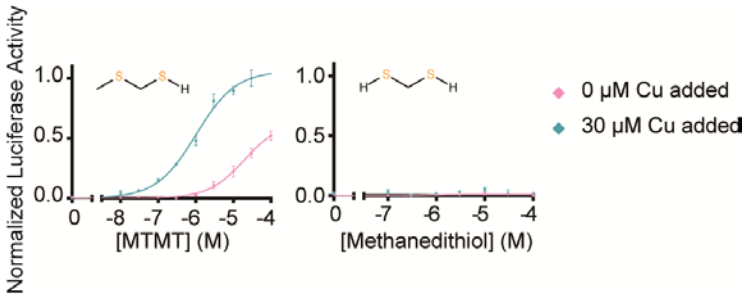


Figure S6. Dose-response curves of MOR244-3 to (methylthio)methanethiol (CH_3SCH_2SH) and methanedithiol ($HSCH_2SH$) with and without 30 µM exogenous Cu^{2+}addition as measured in luciferase reporter gene assay. The *y*-axis represents mean luciferase activity ±SEM, normalized to the response to (methylthio)methanethiol ($N = 3$). The figure is modified with permission from ref. (9).

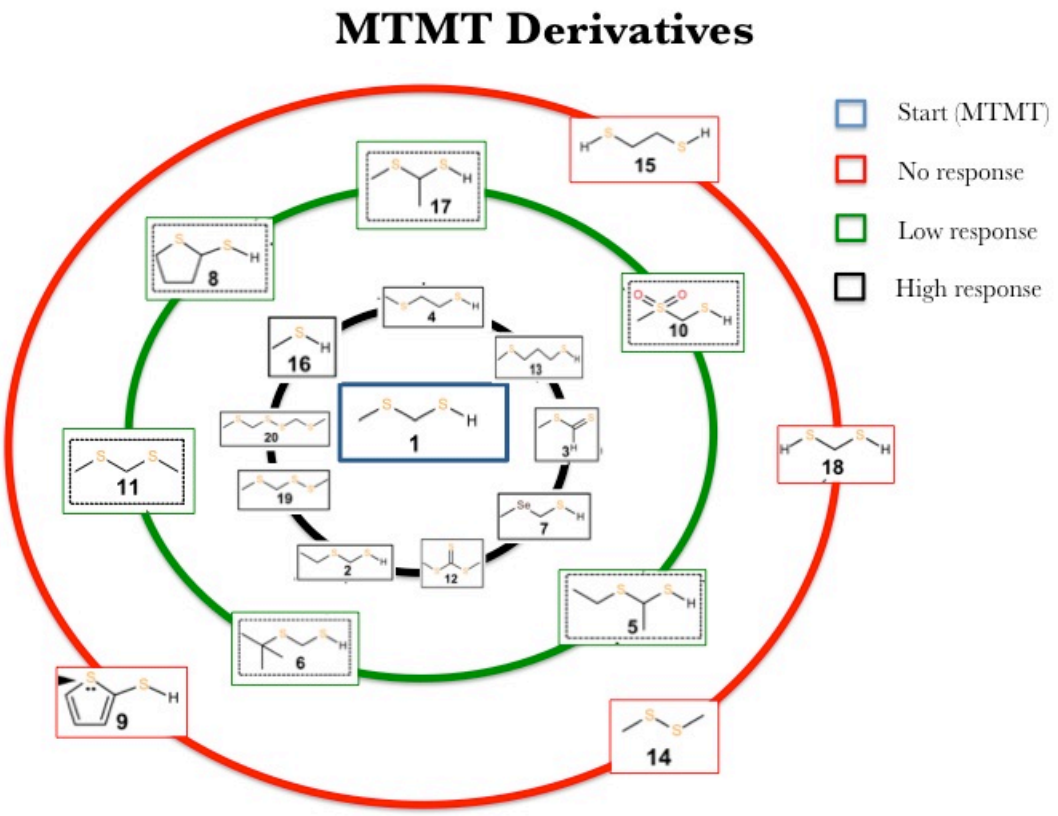


Figure S7. Schematic representation of response profiles of the MTMT derivatives from ref. (9).

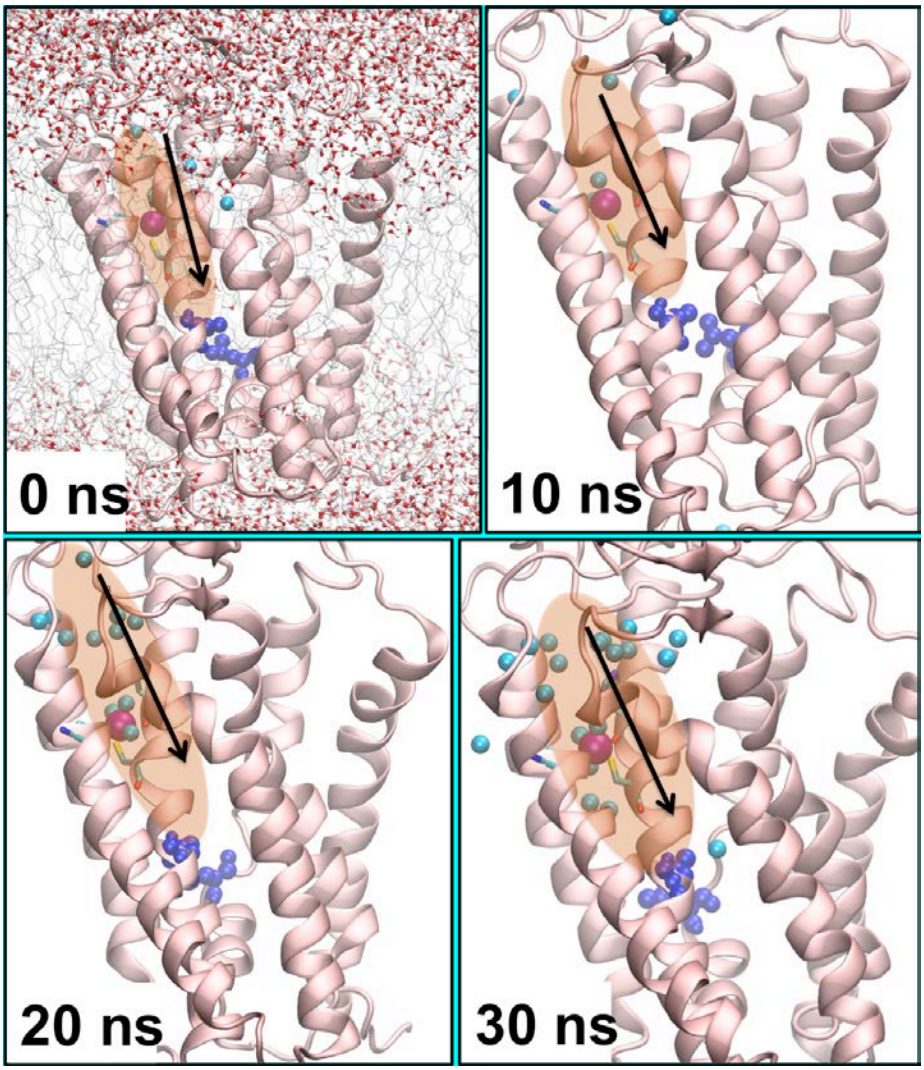


Figure S8. Molecular dynamics (MD) simulations of the MOR244-3 system inside the lipid bilayer. Representative MD snapshots of water molecules (cyan) passing through the aqueous channel (orange) interrupted by a hydrophobic cap (blue) formed by L66 and L114 residues are presented.

IV. References for the Supporting Information

1. 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.
2. Warne, A.; Moukhametzianov, R.; Baker, J. G.; Nehme, R.; Edwards, P. C.; Leslie, A. G. W.; Schertler, G. F. X.; Tate, C. G. *Nature* **2011**, *469*, 241.
3. Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins. D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948.
4. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. L. *J. Mol. Biol.* **2001**, *305*, 567.
5. Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct. Funct. Bioinf.* **2005**, *61*, 704.
6. Maestro, version 9.0, Schrödinger, LLC, New York, NY, 2009.
7. Gerwert, K.; Freier, E.; Wolf, S. *Biochem. Biophys. Acta* **2014**, *1837*, 606.
8. Angel, T. E.; Gupta, S.; Jastrzebska, B.; Palczewski, K.; Chance, M. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**,*106*, 14367.
9. Duan, X.; Block, E.; Li, Z.; Connelly, T.; Zhang, J.; Huang, Z.; Su, X.; Pan, Y.; Wu, L.; Chi, Q.; Thomas, S.; Zhang, S.; Ma, M.; Matsunami, H.; Chen, G. Q.; Zhuang, H. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 3492.
10. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
11. Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760.
12. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; B. Mennucci; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; J. A. Montgomery, J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; J. J. Heyd, E. B.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.
13. Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.
14. Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2008**, *41*, 157.
15. Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2008**, *120*, 215.
16. Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123.
17. Hehre, W. J.; Radom, L.; Schleyer, PvR.; Pople, J. (**1986**) *Ab Initio Molecular Orbital Theory*; Wiley: New York.
18. Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580.
19. Sekharan, S.; Altun, A.; Morokuma, K. *Chem.–Eur. J.* **2010**, *16*, 1744.
20. Sekharan, S.; Morokuma, K. *J. Phys. Chem. Lett.* **2010**, *1*, 668.
21. Sekharan, S.; Altun, A.; Morokuma, K. *J. Am. Chem. Soc.* **2010**, *132*, 15856.
22. Sekharan, S.; Morokuma, K. *J. Am. Chem. Soc.* **2011**, *133*, 4734.
23. Sekharan, S.; Morokuma, K. *J. Am. Chem. Soc.* **2011**, *133*, 19052.

24. Sekharan, S.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2011**, *115*, 15380.
25. Sekharan, S.; Altun, A.; Morokuma, K. *Annu. Rep. Comput. Chem.* **2011**, *7*, 215.
26. Sekharan, S.; Katayama, K.; Kandori, H.; Morokuma, K. *J. Am. Chem. Soc.* **2012**, *134*, 10706.
27. Sekharan, S.; Wei, J. N.; Batista, V. S. *J. Am. Chem. Soc.* **2012**, *134*, 19536.
28. Pal, R.; Sekharan, S.; Batista, V. S. *J. Am. Chem. Soc.* **2013**, *135*, 9624.
29. Sekharan, S.; Mooney, V. L.; Rivalta, I.; Kazmi, M.; Neitz, M.; Neitz, J.; Sakmar, T. P.; Yan. E. C. Y.; Batista, V. S. *J. Am. Chem. Soc.* **2013**, *135*, 19064.
30. Marenich A. V.; Cramer C. J.; Truhlar, D. G. *J. Phys.Chem. B* **2009**, *113*, 6378.
31. Rao, L.; Cui, Q.; Xu, X. *J. Am. Chem. Soc.* **2010**, *132*, 18092.
32. Saito, H.; Kubota, M.; Roberts, R. W.; Chi, Q.; Matsunami, H. *Cell* **2004**, *119*, 679.
33. Pirrung, M. C.; Bleecker, A. B.; Inoue, Y.; Rodriguez, F. I.; Sugarawa, N.; Wada, T.; Zou, Y.; Binder, B. M. *Chem. Biol.* **2008**, *15*, 313.