Supporting Information for

Discrimination of Saturated Aldehydes by the Rat I7 Olfactory Receptor

Michael D. Kurland[1], Michael B. Newcomer[1], Zita Peterlin[3], Kevin Ryan[2], Stuart Firestein[3]*, and Victor S. Batista[1]*

[1]Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107; [2]Department of Chemistry, City College of New York, New York, NY 10031, [3]Department of Biological Sciences, Columbia University, New York, NY 10027

RECEIVED DATE (publisher); *E-mail: sjf24@columbia.edu; victor.batista@yale.edu

This PDF file includes:

Methods and Models

Figures

References:

Homology Structural Model of the OR-I7.

The homology structural model of OR-I7 was built from the crystal structure of bovine visual rhodopsin at 2.2 Å resolution [1]. The primary sequence of the I7 receptor [2] was aligned with the sequence of bovine rhodopsin and nonconserved amino acid residues were mutated by using Schrodinger's Maestro graphical user interface [3]. Nonconserved amino acid residues were mutated and after each mutation, the overall structure was relaxed and equilibrated.

The resulting homology structure can be compared to previously reported structural models [4, 5]. The first model of the rat OR-I7 was developed by Singer et *al* [4] from an electron density map of rhodopsin obtained at 7.5 Å resolution [6]. Hydrophobicity moments were used to identify the helical transmembrane (TM) regions from the rat I7 primary structure. Canonical helices were then created and the energy of each helix was individually minimized. Each side

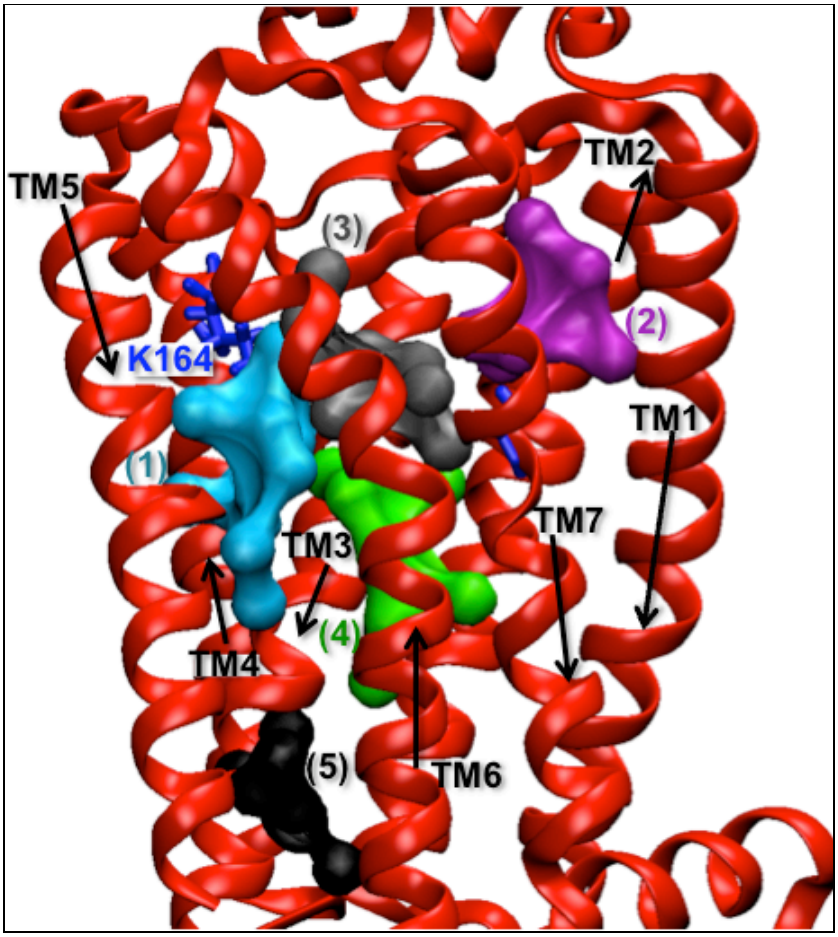

Figure S1: Homology model of the OR-I7, where the transmembrane α-helices (TM1-TM7) are indicated with black arrows. Five cavities (1)-(5) with lysine residues suitable for octanal binding were identified, including: (1) the octanal binding pocket identified in this work (light-blue), (2) the binding site of the aldehyde group of retinal in rhodopsin (purple), (3) the octanal binding pocket proposed by Singer et al[3] and Hall et al[4] (gray), and two additional pockets (4) and (5) in green and black, respectively.

chain on the helices was cast in its lowest energy conformation by accessing a rotamer library. TM domains were assembled by positioning these helices by using the electron diffraction–derived densities for rhodopsin while setting a dielectric constant (equal to 1) to mimic the surrounding membrane. The helices were then rotated to preserve the hydrophobicity construct of a GPCR assembly; that is, hydrophobicity moment was pointed directly away from the TM assembly. The entire structure thus created was then energy minimized and refined using molecular dynamics. A second model of the OR-I7 was subsequently reported by Hall et al [5]. This second model was based on the MembStruk method [7], an alternative approach to build a model of GPCRs without using the coordinates of rhodopsin but extracting information from its electron density map.

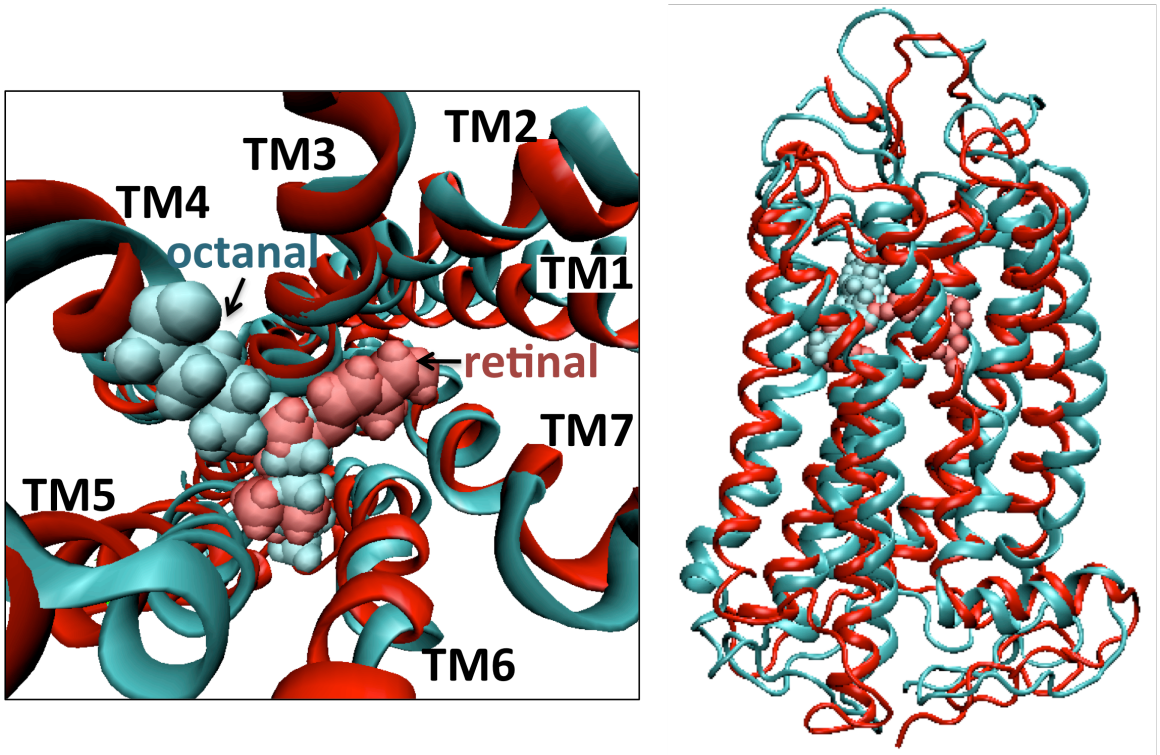

Figure S2: Top-down view (left) and side on view (right) of the homology model of the OR-I7 (light blue) with octanal hydrogen bonded to K164 (light blue spheres) in TM4, superimposed to the rhodopsin structure (red) with retinal forming a pSB with K296 in TM7 (red spheres). Note that the hydrocarbon chain of octanal is in between TM5 and TM6, partially overlapping with the β-ionone ring of retinal.

Figure S3 shows the comparison of the OR-I7 binding sites previously proposed by Singer et *al*[4] and Hall et al [5] This figure shows that the proposed octanal binding based on hydrogen bonding to K164 (see Fig. S1) agrees with these earlier models. However, the predicted binding niche localized between TM5 and TM6 (see Fig. S1) is different from earlier proposals of the binding pocket localized between TM6 and TM7. Therefore, the proposed ligand-protein interactions proposed by earlier models, with the hydrocarbon chain of the octanal ligand interacting with TM7 (see Fig. S3) differ from the model shown in Figs. S1 and S2.

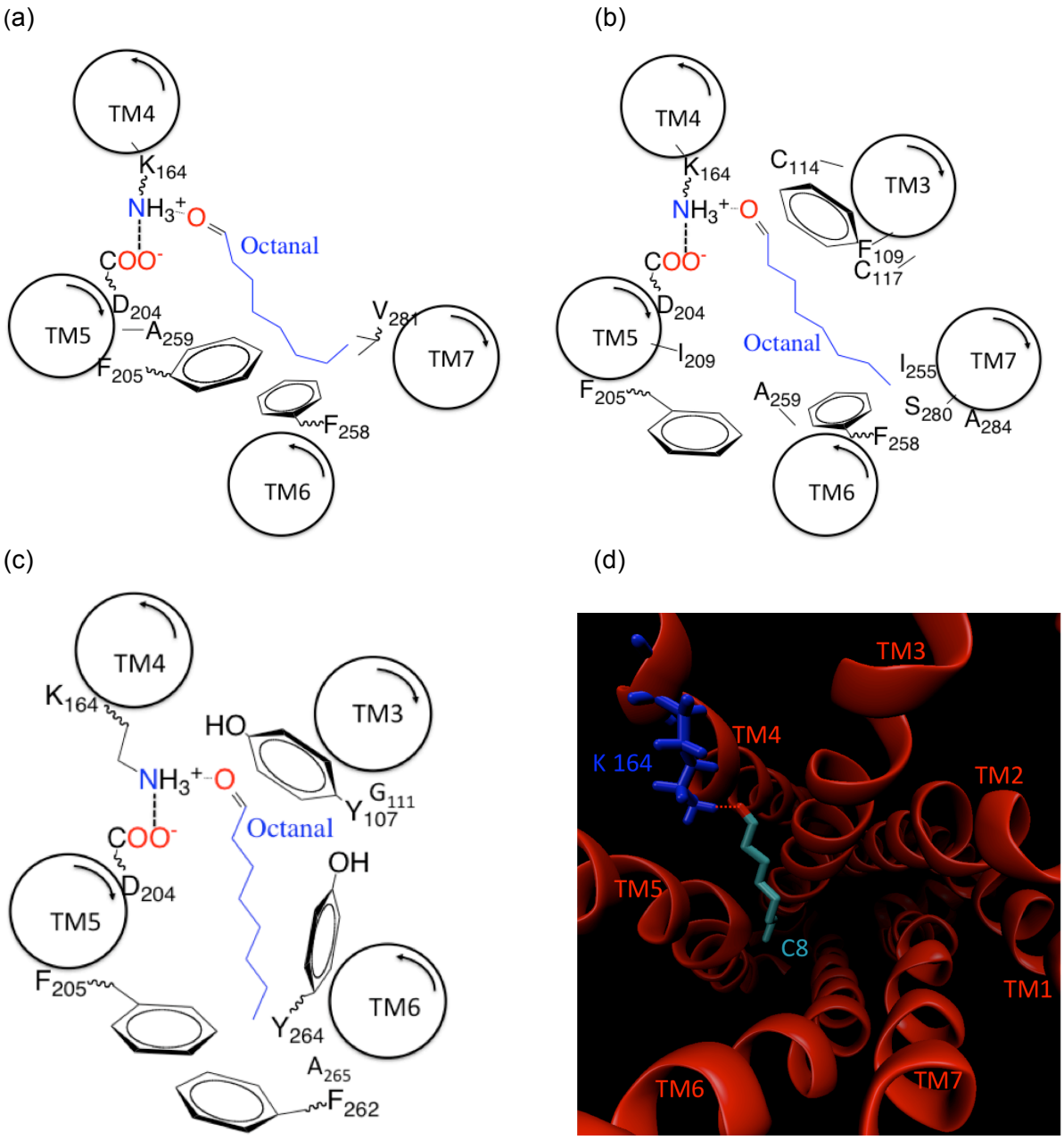

Figure S3: Schematic diagram of the octanal binding pocket (gray in Fig. S1), predicted by the structural models of the OR-I7 proposed by Singer et al[3] (a) and Hall et al[4] (b), as compared to the binding pocket proposed in this paper (c) and (d) (light-blue in Fig. S1).

Ligand Docking and Calculations of Binding Free Energies.

Ligand Binding: Several cavities were identified to have lysine residues suitable for aldehyde binding (see (1)-(5) in Fig. S1). First, protein cavities that could host multiple water molecules were identified as follows. The protein was soaked in water, by using the 'solvate' function of VMD [8]. The water molecules had to be at least within 3.0 Å from any atom in the protein to be placed by the script. The clusters of water molecules that fitted inside the protein were then used to identify all potential places where an octanal ligand could fit. For each of these pockets, the water clusters were substituted by an octanal moiety, if the cavity had a lysine residue where the octanal could form a hydrogen bond.

Annealing molecular dynamics simulations were performed to compare the relative stability of octanal in the various pockets. Langevin molecular dynamics simulations were started at 1000 K and run for 100,000 steps of 0.5 fs. The complete simulation was divided in 30 time windows, quenching the kinetic energy at the end of each window by reducing the temperature of the system by a factor of 1.5 until reaching a state of no motion. Each window ran for 100,000 steps (meaning 3,000,000 total for each annealing process). All α-carbons were fixed throughout these annealing simulations. Each possible pocket was probed multiple times by running simulated annealing of octanal starting with several independent initial conditions.

Free Energy Calculations: Calculations of binding free energies were based on the free energy perturbation (FEP) method [9], as implemented in NAMD [10], following the sequence of alchemical transformations outlined in Fig. S4. The FEP method computes the free energy difference, associated with the transformation of state A into state B, by using the Zwanzig equation:

$$\Delta G\left(A \to B\right) = -k_B T \ln\left\langle \exp\left(-\frac{\left(E_B - E_A\right)}{k_B T} \right) \right\rangle_A \quad , \tag{1}$$

where E_A and E_B is the energy of the system in state A and B, respectively, k_B is the Boltzmann's constant and $\langle ... \rangle_A$ denotes the thermal ensemble average at temperature T for the system equilibrated in state A.

In practice, the transformation between the two thermodynamic states A and B is replaced by a series of N_w transformations between non-physical intermediate states with energy $E_\lambda = E_A + \lambda_k (E_B - E_A)/N_w$, with $k = 1 – N_w$, along an alchemical pathway that connects A to B. Such a fragmentation of the pathway in N_w windows is essential to ensure that for each sub-transformation the initial and final states are sufficiently similar, so that the ensemble of N_c configurations sampled for the initial state includes a statistically complete set of configurations of the system in state B. Each transformation was performed using N_w=24 windows, consisting of N_c=320000 configurations sampled by molecular dynamics (MD) simulations at 300 K, after 80000 MD steps of equilibration. Each MD step propagated the system for 0.5 femtoseconds. The windows for the alchemical pathway were defined with λ=(0, 0.0000001, 0.000001, 0.00001, 0.0001, 0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 0.999, 0.9999, 0.99999, 0.999999, 0.9999999, 1) and the overall free energy change was obtained as follows:

$$\Delta G\left(A \to B\right) = -k_B T \sum_{k=1}^{N_w} \ln \frac{1}{N} \sum_{j=1}^{N_c} \exp\left(-\frac{\left(E^{(j)}_{\lambda_{k+1}} - E^{(j)}_{\lambda_k}\right)}{k_B T} \right)_{\lambda_k} \quad (2)$$

Starting with the equilibrated structure of the octanal-bound OR-I7, the free energy change associated with the octanal → heptanal transformation was computed, yielding the binding affinity of heptanal relative to octanal (ΔΔG).

Analogously, free energy changes were computed for the following transitions: heptanal → hexanal → pentanal → butanal, octanal → nonanal → decanal → undecanal → dodecanal depicted in Fig. S4. All binding energies were computed relative to the energy of the ligands in a low dielectric environment (ε=2). The underlying assumption is that ligands are first dissolved in a low-dielectric reservoir, embedded in the membrane, and then partitioned into the transmembrane OR-I7 binding site.

Conformational flexibility: Representative configurations of the ligand binding to the OR-I7 for the homologous series of n-alkyl saturated aldehydes are shown in Fig. S5. It is shown that conformational flexibility is critical to accommodate one or more kinks in the hydrocarbon chains longer than octanal, an essential feature aspect to adopt bent conformations and fit in the limited volume of the binding site.

Conformational flexibility was also found to be important, even for aldehydes as small as octanal, to remain hydrogen bonded to K164 in the presence of room-temperature thermal fluctuations in the binding pocket. Figure S6 shows representative configurations of *n*-octanal, under room-temperature conditions, responding to thermal fluctuations that affect the configuration of the side chain of F205. Note that due to the flexibility of octanal, the ring of F205 can displace the tail of the hydrocarbon chain (see 4th snapshot in Fig. S6), without breaking the hydrogen-bond with K164.

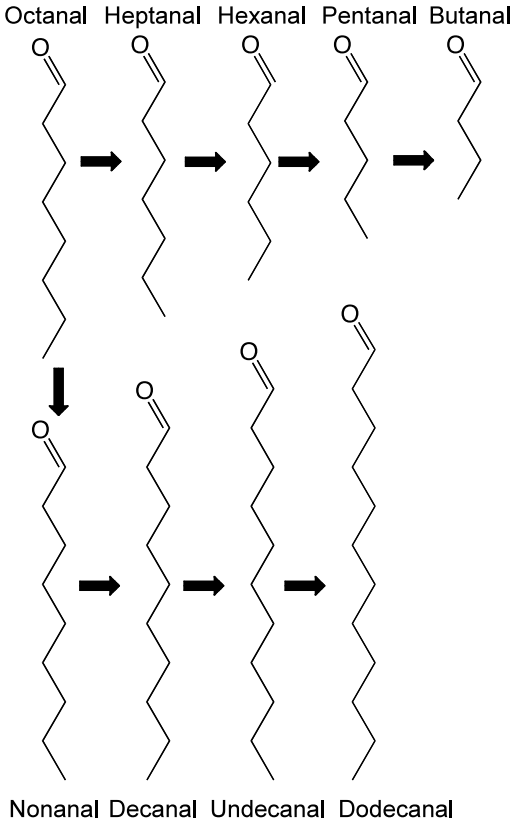

Figure S4: Alchemical transformations implemented for calculations of aldehyde binding free energies in OR-I7, using the FEP method.

Experimental Binding Free Energies: Single olfactory neurons were isolated from infected epithelia and calcium imaging recordings were performed as described in previous work [18-20]. The experimental binding free energies were estimated as $\Delta\Delta G^{\dagger} = RT \ln([C_n]/[C_8])$ from the concentrations $[C_n]$ necessary for half-maximal activation of the OR-I7 (EC_{50}), for C_n=C_6-C_{11}, or the concentration at which 50% inhibition of activation is achieved with C5, simultaneously applied octanal (IC_{50}).

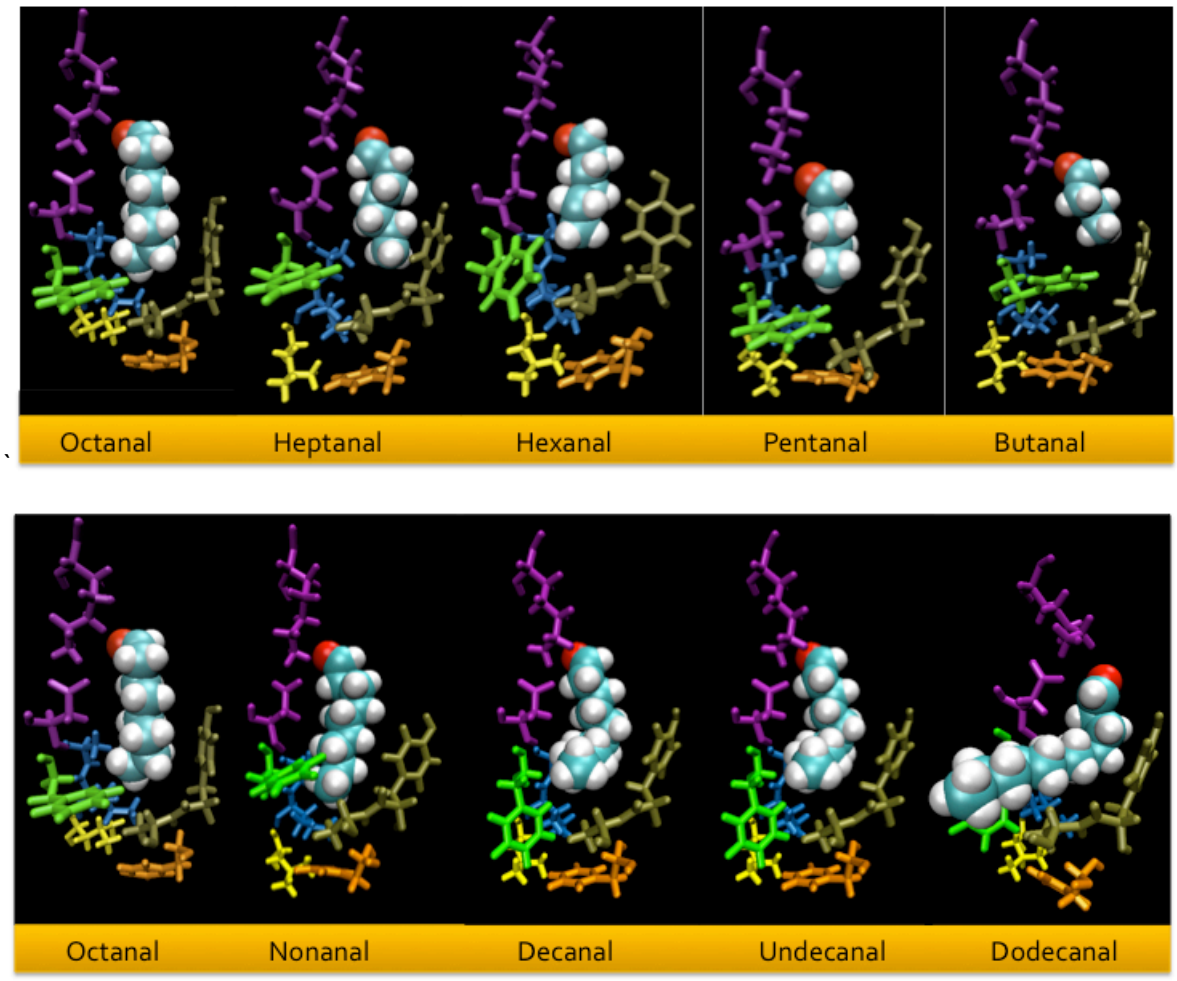

Figure S5: Comparative analysis of ligand binding to the OR-I7 for the homologous series of n-alkyl saturated aldehydes (space filled) Amino acid residue color key: Green: F205; Purple: K164-D204 pair; Yellow: A265; Orange: F262; Gray: Y264.

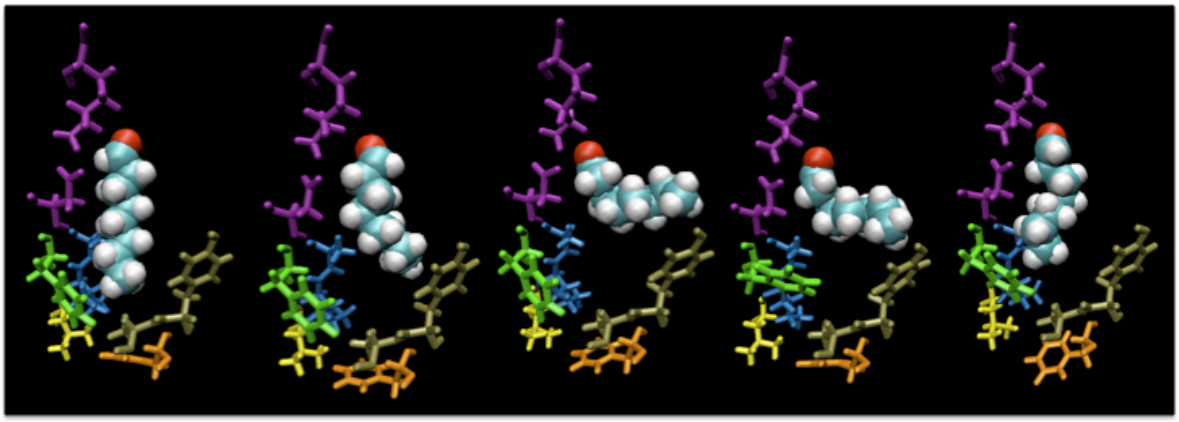

Figure S6: Conformational flexibility exhibited by octanal, bound to the OR-I7 at room temperature. Green: F205; Purple: K164-D204 pair; Yellow: A265; Orange: F262; Gray: Y264.

The calculated binding free energies were obtained by considering the model of equilibrium processes shown in Fig. S7.

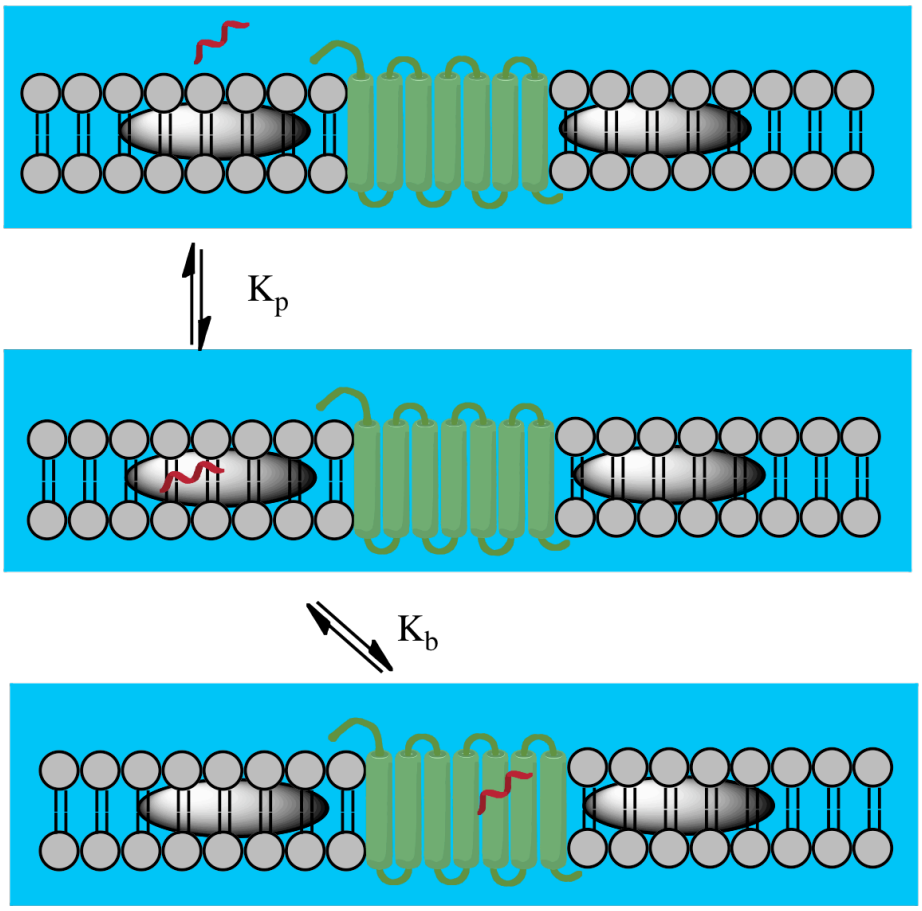

Figure S7: Equilibrium processes describing the partitioning the odorants between the aqueous solution and the low dielectric membrane of olfactory sensory neurons, in the nasal epithelium, as quantified by the equilibrium constant K_p, and the binding of odorants to the olfactory receptor OR-I7, as quantified by K_b.

At 50% activity, the concentration of activated receptors (*i.e.*, OR-I7 with bound odorants) is $[ORI7\text{-}C^{(n)}]=[ORI7]_m/2$, where $[ORI7]_m$ is the concentration giving maximum response.

The concentration of odorant in the membrane is

$$[C_m^{(n)}] = K_p^{(n)}[C_s^{(n)}], \tag{4}$$

where $K_p^{(n)}$ is the equilibrium partition constant and $\left[C_s^{(n)}\right] = EC_{50}^{(n)}$ for n = 6-11, or $\left[C_s^{(n)}\right] = IC_{50}^{(n)}$ for n=5 is the concentration of odorant in solution. The concentration of empty receptors (without bound odorants) is $[ORI7] = [ORI7]_0\text{-}[ORI7]_m/2$, where $[ORI7]_0$ is the initial

concentration of receptors in the absence of odorants. Using these concentrations, we compute the odorant binding constants, as follows:

$$\begin{aligned} K_b^{-1}(n) = e^{\Delta G_n / RT} &= \frac{\left[C_m^{(n)}\right][ORI7]}{\left[ORI7 - C^{(n)}\right]} \\ &= \left[C_s^{(n)}\right] K_p^{(n)} \left(\frac{[ORI7]_0 - [ORI7]_m / 2}{[ORI7]_m / 2} \right) \end{aligned} \tag{5}$$

where

$$\Delta G_n = RT \ln[C_s^{(n)}] + RT \ln[K_p^{(n)}] + RT \ln\left(\frac{[ORI7]_0 - [ORI7]_m / 2}{[ORI7]_m / 2} \right) \tag{6}$$

Therefore, substituting Eq. (4) into Eq. (6), it is obtained that $\Delta\Delta G_n$ can be computed as follows:

$$\begin{aligned} \Delta\Delta G_n &= -RT \left[\ln \frac{K_b^{(n)}}{K_b^{(8)}} \right] \\ &= \Delta\Delta G_n^{\dagger} + RT \left[\ln \frac{K_p^{(n)}}{K_p^{(8)}} \right] \end{aligned} \tag{7}$$

where

$$\Delta\Delta G_n^{\dagger} = RT \ln \frac{[C_s^{(n)}]}{EC_{50}^{(8)}} . \tag{8}$$

Figure 2 in the main manuscript shows that $\Delta\Delta G_n$ is linearly correlated with $\Delta\Delta G_n^{\dagger}$, with a proportionality constant close to ½ (e.g., $\Delta\Delta G^{\dagger} = \Delta\Delta G^{\dagger} / \alpha_n$, with $\alpha \sim 2$). Such a correlation indicates that

$$\Delta\Delta G_n^{\dagger} = \frac{\Delta\Delta G_n^{\dagger}}{\alpha} - RT \ln \frac{K_p^{(n)}}{K_p^{(8)}} \tag{9}$$

and substituting Eqs. (4) and (8) into Eq. (9), we obtain:

$$\left[C_m^{(n)}\right] \propto \left[C_s^{(n)}\right] \approx \sqrt{C_s^{(n)}} \tag{10}$$

for the partitioning of aldehydes between the aqueous and the membrane environments. A similar dependence has been recently reported for the partitioning of aldehydes in reactive extraction experiments, where aldehydes are extracted from the aqueous solution with an organic phase and form a Schiff-base with a primary amine [11].

Bound life-times: The stability of odorants bound to the OR-I7 can be correlated to the exchange rates $k=\tau^{-1}$ with odorants in the membrane (*i.e.*, the inverse of the bound life-times), according to the Arrhenius equation:

$$\tau = A^{-1} \exp(E_a/RT), \tag{11}$$

where E_a is the activation energy for displacing the odorant from the pocket into the membrane, and $RT \approx 0.6$ kcal mol^{-1} at T = 29 ºC. According to Eq. (11),

$$\tau_1 = \tau_0 10^{\Delta Ea / 1.38}. \tag{12}$$

Therefore, increasing the activation free energy by only ΔE_a=1.38 kcal mol^{-1} extends the life-time of the bound state by an order of magnitude (10 times).

Imine Formation in OR-I7.

In order to create the protonated Schiff base (pSB), the octanal was moved by hand in 0.1Å steps to reduce the distance between the carbon of the aldehyde and the nitrogen of Lys164 from 3Å down to the carbon-nitrogen bond distance of 1.5Å. After each movement, the protein

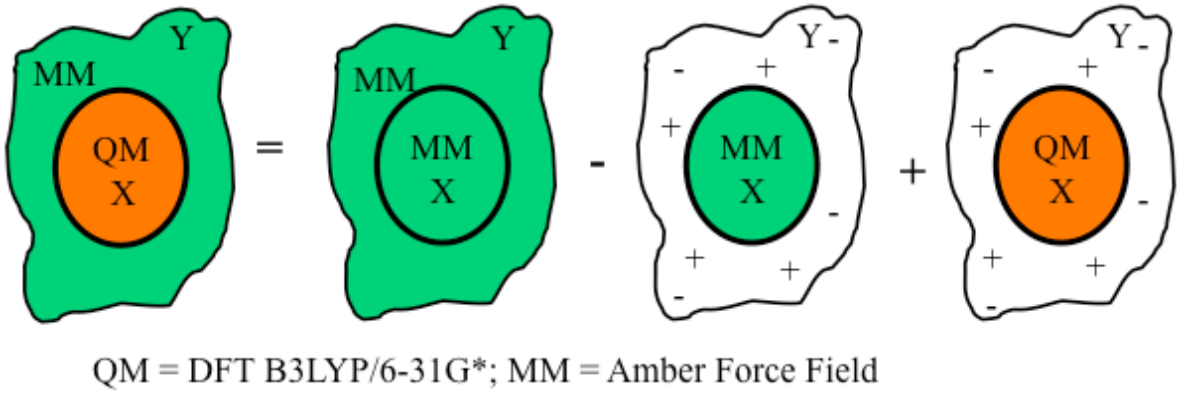

Figure S8: Schematic representation of the ONIOM-EE method for QM/MM calculations.

was relaxed at the molecular mechanics level of theory using the OPLS-AA force field as found in the Schrodinger suite of packages[3]. Once this distance was achieved, the carbon nitrogen bond was formed, and the condensation of the water was created. The newly formed water was made to hydrogen bond with the pSB. The quantum mechanics/molecular mechanics (QM/MM) model of OR-I7 is then built at the ONIOM Electronic-Embedding (DFT B3LYP/6-31G*:Amber) level, as previously reported for bovine rhodopsin.[12-14] The complete system is formally partitioned (see Fig. S8) according to the two-layer ONIOM Electronic-Embedding (EE) method.[15] The reduced system *X* separates the potentially reactive part of the binding site from the rest of the system (region *Y*). When octanal is hydrogen bonded to K164, the reduced system *X* includes the octanal ligand, the carboxylate group of D204 and the side chain of K164 with the QM/MM boundary at the Cδ-Cε bond. (see Fig. S9). For the protonated Schiff base, the reduced system includes the imine form of octanal condensed with K164, up to the Cδ-Cε bond defining the QM/MM boundary, the bound water molecule formed by condensation, and the carboxylate group of D204. The total energy E is obtained for that partitioning scheme according to the link-hydrogen atom method,[15] as implemented in Gaussian03:[16]

$$E = E^{MM,X+Y} + E^{QM,X} - E^{MM,X}$$

where $E^{MM,X+Y}$ is the energy of the complete system computed at the molecular mechanics level of theory, while $E^{QM,X}$ and $E^{MM,X}$ correspond to the energy of the reduced-system *X* computed at the QM and MM levels of theory, respectively. Electrostatic interactions between regions X and Y are included in the calculation of both $E^{QM,X}$ and $E^{MM,X}$ at the quantum mechanical and molecular mechanics levels, respectively. Thus, the electrostatic interactions computed at the MM level in $E^{MM,X}$ and $E^{MM,full}$ cancel and the resulting DFT QM/MM evaluation of the total energy involves a quantum mechanical description of the polarization of the reduced system due to the electrostatic influence of the surrounding protein environment. Polarization of the amino acid residues in the binding pocket due to the redistribution of charge upon imine formation is introduced by correcting the atomic charges according to the MoD-QM/MM self-consistent polarization protocol.[17] The energy difference between the pSB model and the hydrogen-bonded model of OR-I7 was then calculated by extracting the QM layer shown in Figure S9. A constrained minimization where only the hydrocarbon tail of octanal was allowed to relax was then conduced at the B3LYP/6-31g* level of theory using Gaussian[16]. The energy of the pSB was found to be 4.8 kcal/mol more stable than the hydrogen bonded model of OR-I7.

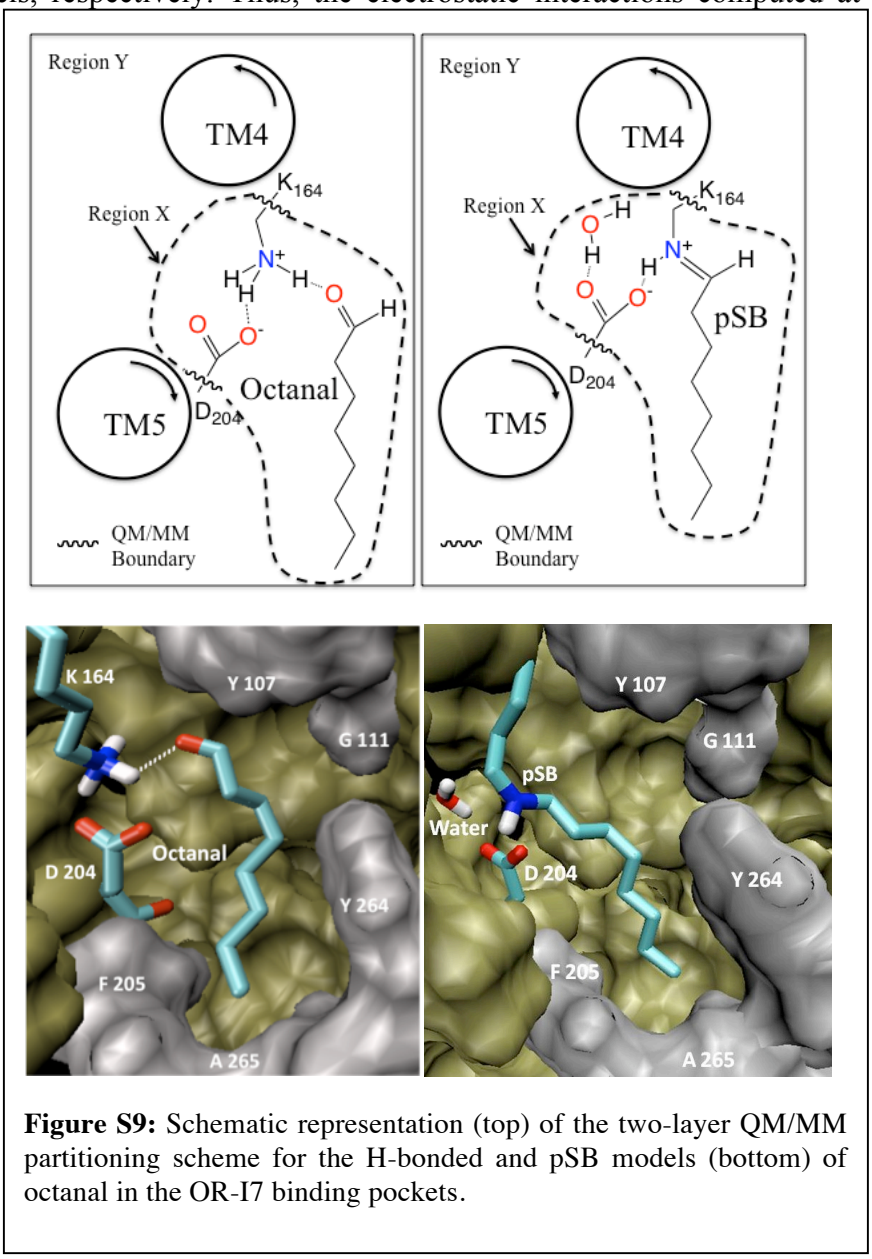

Figure S9: Schematic representation (top) of the two-layer QM/MM partitioning scheme for the H-bonded and pSB models (bottom) of octanal in the OR-I7 binding pockets.

Primary Sequence Alignment.

The primary sequences of bovine visual rhodopsin [1] and the rat I7 receptor [2] were aligned by using the standard alignment algorithm by Larkin et al. [21]. Two alignments, shown below, were generated by using the Gonnet [22] and BLOcks of Amino Acid SUBstitution (BLOSUM) [23] alignment matrices:

Gonnet Matrix Alignment:

```
BOVINRhodopsin      MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLG-FPINFLTL 59
RATI7               MERRNHSGRVSEFVLLGFPAPAPLRVLLFFLSLLAYVLVLTENMLIIIAIRNHPTLHKPM 60
                    *:  : ..    *       . :*:..  ::*:    . :*:  *:::* :  .*  . .:

BOVINRhodopsin      YVTVQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGFFATL 119
RATI7               YFFLAN------MSFLEIWYVTVTIPKMLAGFIGSKEN--HGQLISFEACMTQLYFFLGL 112
                    *. : :      :.:: :  ... :  ::.**  :  .  ** ::   :  .   **  *

BOVINRhodopsin      GGEIALWSLVVLAIERYVVVCKPMSN-FRFGENHAIMGVAFTWVMALACAAPPLVGWSRY 178
RATI7               G-CTECVLLAVMAYDRYVAICHPLHYPVIVSSRLCVQMAAGSWAGGFGISMVKVFLISRL 171
                    *       *.*:* :***.:*:*:   . .... .:  .* :*. .:. :   :.  **

BOVINRhodopsin      IPEG----MQCSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAA 234
RATI7               SYCGPNTINHFFCDVSPLLNLSCTDMSTAELTDFVLAIFI-LLGPLSVTGASYMAITGAV 230
                       *     :  *.:.     . *: .:  :  **: ::* *:  :   *   :::. *.


BOVINRhodopsin      AQQQESATTQKAEKEVTRMVIIMVIAFLICWLPYAG-VAFYIFTHQGSDFGPIFMTIPAF 293
RATI7               MRIPSAAGRHKAFSTCASHLTVVIIFYAASIFIYARPKALSAFDTN--------KLVSVL 282
                     :  .:*  :** .  :  : :::* :  . : **   *:  *  :          :..:


BOVINRhodopsin      FAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQVAPA      348
RATI7               YAVIVPLFNPIIYCLRNQDVKRALRRTLHLAQ----DQEANTNKGSKNG------      327
                    :*   .::**:** : *::.:..:  **  .:    *:**.*. ...:
```

BLOSUM Matrix Alignment:

```
BOVINRhodopsin      MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLY 60
RATI7               ------MERRNHSGRVSEFVLLGFPAPAPLRVLLFFLSLLAYVLVLTENMLIIIAIRNHP 54
                           :     .. : .*   * **    .  : :*:**  :.*   : : * : .

BOVINRhodopsin      VTVQHKKLRTP-LNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGFFATL 119
RATI7               TLHKPMYFFLANMSFLEIWYVTVTIPKMLAGFIGSKEN--HGQLIS-FEACMTQLYFFLG 111
                    . :    :  . :.:: :  ... :  ::.**  :  .  ** ::    .*   : :*

BOVINRhodopsin      GGEIALWSLVVLAIERYVVVCKPMSN-FRFGENHAIMGVAFTWVMALACAAPPLVGWSRY 178
RATI7               LGCTECVLLAVMAYDRYVAICHPLHYPVIVSSRLCVQMAAGSWAGGFGISMVKVFLISRL 171
                     *      *.*:* :***.:*:*:   . .... .:  .* :*. .:. :   :. **

BOVINRhodopsin      IPEGMQCSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAAAQQQ 238
RATI7               SYCGPNT---INHFFCDVSPLLNLSCTDMSTAELTDFVLAIFILLGPLSVTGASYMAITG 228
                       * :    *:::  . ..  :     * ...:   ::.**:  * * .*  .  *

BOVINRhodopsin      ESATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTIPAFFAKTS 298
RATI7               -AVMRIPSAAGRHKAFSTCASHLTVVIIFYAASIFIYARPKALSAFDTNKLVSVLYAVIV 267
                     :.    :     : .   .  : :  : **.  *     :. .       :..::*

BOVINRhodopsin      AVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGDDEASTTVSKTETSQVAPA           348
RATI7               PLFNPIIYCLRNQDVKRALRRTLHLAQ-----DQEANTNKGSKNG-----           327
                    .::**:** : *::.:..:  **  .:     *: :.* . ::..
```

In the above alignments, yellow highlight indicates a rhodopsin TM helix while green highlight indicates an OR-I7 helix[24]. Both alignments result in overlap of the Rhodopsin TM helices with the OR-I7 helices. Full models of OR-I7 were constructed using both alignments to test the robustness of the methodology. After MM relaxation, both models yielded the same binding pocket, suggesting the methodology is resilient to changes in the alignment used for the homology model construction process. Overall, the BLOSUM alignment shows better alignment for TM helices 1 and 2 while the Gonnet matrix results in better alignment of TM helices 5 and 6. TM helices 3,4, and 7 are aligned almost identically. Since TM helices 5 and 6 directly define the binding pocket, while TM helices 1 and 2 do not, the alignment of TM helices 5 and 6 is preferred. Therefore, the Gonnet alignment model was used for the free energy perturbation calculations.

Complete Reference (9):

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., 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., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C. and Pople, J. A. (2004) Gaussian, Inc. ed.)^eds.), Wallingford, CT

References for Supporting Information:

1 Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P. and Buss, V. (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 angstrom crystal structure. J. Mol. Biol. **342**, 571-583

2 Buck, L. and Axel, R. (1991) A novel multigene family may encode odorant receptors - A molecular basis for odor recognition. Cell. **65**, 175-187

3 Schrodinger, I. (2006) ed.)^eds.), Maestro, New York

4 Singer, M. S. (2000) Analysis of the molecular basis for octanal interactions in the expressed rat I7 olfactory receptor. Chem. Senses. **25**, 155-165

5 Hall, S. E., Floriano, W. B., Vaidehi, N. and Goddard, W. A. (2004) Predicted 3-D structures for mouse 17 and rat 17 olfactory receptors and comparison of predicted odor recognition profiles with experiment. Chem. Senses. **29**, 595-616

6 Schertler, G. F. X. (1998) Structure of rhodopsin. Eye. **12**, 504-510

7 Floriano, W. B., Vaidehi, N., Goddard, W. A., Singer, M. S. and Shepherd, G. M. (2000) Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor. Proc. Natl. Acad. Sci. U.S.A. **97**, 10712-10716

8 Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD - Visual Molecular Dynamics. J. Mol. Graphics Modell. **14**, 33-38

9 Zwanzig, R. W. (1954) High-temperature equation of state by a perturbation method. 1. Nonpolar gases. J. Chem. Phys. **22**, 1420-1426

10 Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L. and Schulten, K. (2005) Scalable molecular dynamics with NAMD. J. Comput. Chem. **26**, 1781-1802

11 Babic, K., van der Ham, A. G. J. and de Haan, A. B. (2009) Reactive extraction of aldehydes from aqueous solutions with Primene (R) JM-T. Sep. Purif. Technol. **66**, 525-531
12 Gascon, J. A. and Batista, V. S. (2004) QM/MM study of energy storage and molecular rearrangements due to the primary event in vision. Biophys. J. **87**, 2931-2941
13 Gascon, J. A., Sproviero, E. M. and Batista, V. S. (2005) QM/MM study of the NMR spectroscopy of the retinyl chromophore in visual rhodopsin. J. Chem. Theory Comput. **1**, 674-685
14 Gascon, J. A., Sproviero, E. M. and Batista, V. S. (2006) Computational studies of the primary phototransduction event in visual rhodopsin. Acc. Chem. Res. **39**, 184-193
15 Vreven, T. and Morokuma, K. (2000) The ONIOM (our own N-layered integrated molecular orbital plus molecular mechanics) method for the first singlet excited (S-1) state photoisomerization path of a retinal protonated Schiff base. J. Chem. Phys. **113**, 2969-2975
16 Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., 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., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C. and Pople, J. A. (2004) Gaussian, Inc. ed.)^eds.), Wallingford, CT
17 Gascon, J. A., Leung, S. S. F., Batista, E. R. and Batista, V. S. (2006) A self-consistent space-domain decomposition method for QM/MM computations of protein electrostatic potentials. J. Chem. Theory Comput. **2**, 175-186
18 Yuste, R., Lanni, F. and Konnerth, A. (2000) Imaging Neurons: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York
19 Araneda, R. C., Kini, A. D. and Firestein, S. (2000) The molecular receptive range of an odorant receptor. Nat. Neurosci. **3**, 1248-1255
20 Araneda, R. C., Peterlin, Z., Zhang, X., Chesler, A. and Firestein, S. (2004) A pharmacological profile of the aldehyde receptor repertoire in rat olfactory epithelium. J. Physiol. (Lond.). **555**, 743-756
21 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. and Higgins, D. G. (2007) Clustal W and clustal X version 2.0. Bioinformatics. **23**, 2947-2948
22 Gonnet, G. H., Cohen, M. A. and Benner, S. A. (1992) Exhaustive Matching of the Entire Protein-Sequence Database. Science. **256**, 1443-1445
23 Henikoff, S. and Henikoff, J. G. (1992) Amino-Acid Substitution Matrices from Protein Blocks. Proc. Natl. Acad. Sci. U.S.A. **89**, 10915-10919
24 Lai, P. C., Bahl, G., Gremigni, M., Matarazzo, V., Clot-Faybesse, O., Ronin, C. and Crasto, C. J. (2008) An olfactory receptor pseudogene whose function emerged in humans: a case study in the evolution of structure-function in GPCRs. Journal of Structural and Functional Genomics. **9**, 29-40