THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 2, pp. 1216–1224, January 12, 2007

The Molecular Basis for Ligand Specificity in a Mouse Olfactory Receptor

*A NETWORK OF FUNCTIONALLY IMPORTANT RESIDUES**[S]

Received for publication, October 3, 2006, and in revised form, November 8, 2006 Published, JBC Papers in Press, November 17, 2006, DOI 10.1074/jbc.M609355200

Tatjana Abaffy[‡1], **Arun Malhotra**[§], **and Charles W. Luetje**[‡]

From the Departments of [‡]*Molecular and Cellular Pharmacology and* [§]*Biochemistry and Molecular Biology, University of Miami, Miami, Florida 33101*

Sequence differences between members of the mouse olfactory receptor MOR42 subfamily (MOR42-3 and MOR42-1) are likely to be the basis for variation in ligand binding preference among these receptors. We investigated the specificity of MOR42-3 for a variety of dicarboxylic acids. We used site-directed mutagenesis, guided by homology modeling and ligand docking studies, to locate functionally important residues. Receptors were expressed in *Xenopus* oocytes and assayed using high throughput electrophysiology. The importance of the Val-113 residue, located deep within the receptor, was analyzed in the context of interhelical interactions. We also screened additional residues predicted to be involved in ligand binding site, based on comparison of ortholog/paralog pairs from the mouse and human olfactory receptor genomes (Man, O., Gilad, Y., and Lancet, D. (2004) *Protein Sci.* 13, 240–254). A network of 8 residues in transmembrane domains III, V, and VI was identified. These residues form part of the ligand binding pocket of MOR42-3. C12 dicarboxylic acid did not activate the receptor in our functional assay, yet our docking simulations predicted its binding site in MOR42-3. Binding without activation implied that C12 dicarboxylic acid might act as an antagonist. In our functional assay, C12 dicarboxylic acid did indeed act as an antagonist of MOR42-3, in agreement with molecular docking studies. Our results demonstrate a powerful approach based on the synergy between computational predictions and physiological assays.

The interaction between a receptor and its molecular partner (the ligand) is a fundamental biological mechanism. In the olfactory system, this interaction is the interface between the central nervous system and environmental cues regarding food and danger and represents the first step in the complex perception of odor. Olfactory sensory neurons, the only true neurons in direct contact with the environment (2), reside in the main olfactory epithelium and express olfactory receptors (ORs)[2] (3). Binding of odorant ligands to these receptors triggers a signal transduction cascade that leads to olfactory sensory neuron activation and subsequent activation of neurons in the olfactory bulb and the olfactory cortex (4). Most odors are composed of multiple odorant molecules, and each odorant molecule is thought to activate several odorant receptors. This forms a combinatorial code and represents a basis for odorant recognition (5). Each olfactory sensory neuron expresses only one of the ~1000 different ORs (6–8) and typically responds to more than one odorant molecule (5) with temporally patterned bursts of action potentials (9). The axons of olfactory sensory neurons expressing the same OR converge onto distinct locations in the olfactory bulb (10, 11). Thus, the ligand specificities of individual ORs are key determinants in olfactory coding and the perception of odor. What determines the differences in ligand specificity among ORs? A hypothesis is that a group of specific residues that differ among receptors forms the basis for ligand specificity. The differences in ligand specificity among ORs have been attributed to differences in the highly variable transmembrane domains (3). Computational analysis of amino acid variability among 197 ORs from 5 different species showed that a majority of the highly variable residues are present in 3 of the 7 transmembrane domains (TM III, IV, and V) (12), indicating a potential recognition site for a large variety of odorants.

Mammalian ORs belong to the rhodopsin-like family of G protein-coupled receptors (GPCRs) with a 7-transmembrane domain helix bundle arrangement and sharing several sequence motifs and a number of highly conserved amino acids in the transmembrane domains (TMs) (13). Mouse ORs (MORs) share some of these common motifs, as well as several OR-specific motifs (14). The low overall sequence identity (~25%) among rhodopsin-like GPCRs suggests that significant deviations can occur in ligand binding pockets and in inter-helical contacts and makes the study of GPCR ligand binding sites particularly challenging. A variety of approaches have been used to identify functionally important residues within receptor proteins, such as correlated mutation analysis (15), sequence-based entropy (16), modeling of receptor ligand binding (17), and the evolutionary trace method (18). Several studies have used homology modeling and ligand docking approaches to predict the binding site in mammalian ORs (19–

* This work was supported by National Institutes of Health Grants MH66038 (to C. W. L.) and GM69972 (to A. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S4 and Table S1.

[1] To whom correspondence should be addressed: Dept. of Molecular and Cellular Pharmacology, University of Miami, P. O. Box 016189, Miami, FL 33101. Tel.: 305-243-4468; Fax: 305-243-4555; E-mail: tabaffy@med.miami.edu.

[2] The abbreviations used are: OR, odorant receptor; MOR, mouse OR; GPCR, G protein-coupled receptor; TM, transmembrane.

23). Functional specificity of proteins is assumed to be conserved among orthologs and different among paralogs. This assumption was used to identify residues that determine specificity of protein-ligand recognition sites (1). To date, only one study has combined computational and functional approaches to elucidate the structure of the ligand binding site of a mammalian OR, the mouse eugenol receptor (24).

Here we report the structural basis for ligand specificity of an OR (MOR42-3) by combining functional analysis and site-directed mutagenesis with homology modeling and ligand docking studies. MOR42-3 is closely related to MOR42-1, and both of these receptors respond to aliphatic dicarboxylic acids. However, these two receptors are differentially sensitive to odorant ligands on the basis of carbon chain length, with MOR42-3 preferring the 9-carbon nonanedioic acid and MOR42-1 preferring the 10-carbon decanedioic acid and 11-carbon undecanedioic acid (25). This discrete difference in ligand specificity and the high level of sequence identity between these receptors (97% within the transmembrane domains) allowed us to introduce mutations in MOR42-3 that changed ligand preferences. By developing a homology model of MOR42-3, based on the structure of rhodopsin (26), we can localize the critical residues identified in our mutagenesis studies and use this model as a target for molecular docking studies. In addition to providing insight into the interaction of agonists with MOR42-3, these ligand-docking studies also predict an antagonist of MOR42-3, a result confirmed by functional analysis.

EXPERIMENTAL PROCEDURES

Materials—*Xenopus laevis* frogs were purchased from Nasco (Fort Atkinson, WI). The care and use of *X. laevis* frogs in this study were approved by the University of Miami Animal Research Committee and meet the guidelines of the National Institutes of Health. The human $G\alpha_{olf}$ construct in the pcDNA3.1 vector was purchased from the University of Missouri-Rolla cDNA Resource Center. The human CFTR clone was kindly provided by Dr. Ian Dickerson (University of Rochester). Chemicals (including odorants) were from Sigma-Aldrich.

Mutagenesis—Mutations were introduced using the QuikChange II site-directed mutagenesis kit (Stratagene, La Jolla, CA). Each mutant construct was verified by sequencing.

Expression of Wild-type and Mutant Mouse Odorant Receptors in Xenopus Oocytes—Our method for functional expression of MORs in *Xenopus* oocytes has been described previously (25). Briefly, cDNA encoding each receptor (MOR42-1, MOR42-3, and various mutants of MOR42-3) are in the pCI vector (Promega) containing the 20-amino acid N-terminal sequence of human rhodopsin. This "rhodopsin tag" is required for functional expression of MOR42-3 and MOR42-1 in *Xenopus* oocytes (24). Thus, all wild-type and mutant MORs in our study contain this N-terminal sequence extension. Capped cRNA encoding each protein was generated using mMessage mMachine kits (Ambion, Austin, TX). 25 ng of cRNA encoding an MOR, together with 10 ng of $G\alpha_{olf}$ and 1 ng of cystic fibrosis transmembrane regulator cRNAs, was injected into each oocyte. Oocytes were incubated at 18 °C in Barth's saline (in mM: 88 NaCl, 1 KCl, 2.4 $NaHCO_3$, 0.3 $CaNO_3$, 0.41 $CaCl_2$, 0.82 $MgSO_4$, 15 HEPES, pH 7.6, and 12 μg/ml tetracycline) for 2–5 days prior to electrophysiological recording. Receptor expression as measured by current amplitude is highly variable in *Xenopus* oocytes. For this reason, we present our results as a ratio of responses (odorant/odorant, odorant/ isobutylmethylxanthine, etc.). Current responses of mutant receptors (in relation to isobutylmethylxanthine responses) were not different from wild-type MOR42-3 and MOR42-1.

Electrophysiology and Data Analysis—Odorant-induced Cl^- currents, resulting from cAMP-induced activation of co-expressed cystic fibrosis transmembrane regulator, were measured 2–5 days after cRNA injection using two-electrode voltage clamp in an automated parallel electrophysiology system (OpusExpress 6000A; Molecular Devices). Micropipettes were filled with 3 M KCl and had resistances of 0.2–2.0 MΩ. The holding potential was −70 mV. Current responses, filtered (4-pole, Bessel, low pass) at 20 Hz (−3 db) and sampled at 100 Hz, were captured and stored using OpusXpress 1.1 software (Molecular Devices). Initial analysis was done using Clampfit 9.1 software (Molecular Devices). Oocytes were perfused with ND96 (in mM: 96 NaCl, 2 KCl, 1 $CaCl_2$, 1 $MgCl_2$, 5 HEPES, pH 7.5). Odorants were stored under argon, and high concentration stock solutions (usually 1 M) of each odorant were freshly prepared in ethanol or Me_2SO on the day of each experiment. Each odorant, diluted in ND96, was applied for 15 s (24). Statistical analyses were done using Prism 4 (Graphpad, San Diego, CA). Statistical significance was assessed using a two-tailed unpaired *t* test, a one-tailed paired ratio *t* test, or a one-way analysis of variance followed by the Dunnett's post-test, as appropriate.

Homology Modeling—Structural models of the MOR42-3 receptor were generated using MODELLER 8v1 (27). This program, which uses spatial constraints derived from an existing atomic structure to model homologous proteins, is widely used for comparative protein modeling. Sequence alignment of MOR42-3 with rhodopsin, MORs 42–1 and 42–2, and 12 other representative GPCRs from the rhodopsin-like family A was done with ClustalW and manually improved. The 2.2 Å resolution crystal structure of bovine rhodopsin (Protein Data Bank code 1U19) (26) was used as a template to get initial coordinates for MOR42-3. This structure was then subjected to energy refinement in MODELLER with three cycles of conjugate gradient and simulated annealing to get a stereochemically plausible model. Fifteen models were generated using different random starting coordinates, and the model with the lowest modeler objective function was used.

An important consideration in our homology modeling was the location of the disulfide bond between extracellular loop 2 and the N-terminal end of TM3, which is conserved among receptors that belong to the rhodopsin family. This constraint was explicitly imposed and monitored during model refinement. Disulfide bonds are possible between several cysteine pairs, and we used modeling to sort through these combinations. The DISULFIND program (disulfind.dsi.unifi.it/) (28) initially predicted disulfide bonds between Cys-76-Cys-193, Cys-101-Cys-183, Cys-107-Cys-131, and Cys-116-Cys-173. From these, only the Cys-101-Cys-183 pair gave a model with an acceptable geometry and a low objective function. BLAST analysis showed that the Cys-101 could bind either Cys-173, Cys-183, or Cys-193; these three cysteines are conserved not

only in the MOR42 subfamily but also in most mouse (14) and human ORs (29). We generated homology models with an explicit disulfide bond between residues 101–173, 101–183, or 101–193. Because all three models had acceptable distances between the relevant Cα atoms (8.1, 6.5, and 5.5 A; an ideal disulfide bond has an α-carbon-α-carbon distance under 8 Å.), they were evaluated by DOPE (MODELLER)) and PROCHECK (nihserver.mbi.ucla.edu/SAVS/) (30). The model with a disulfide bond between residues 101–171 had the best geometry and the lowest modeler objective function and was chosen for further refinement. The relative orientation of the transmembrane helices is not affected by our choice of the 101–171 disulfide bond.

The structure was inspected with PROCHECK for inappropriate stereochemistry (clashing side chains, disallowed torsion angles, etc.). Based on the results of the PROCHECK analysis, several residues (Ile-178, Leu-198, Leu-213, Lys-322, Ser-328, Ser-97, Ile-96, and Thr-320) were manually adjusted using the O software package (31). Further minimization was then carried out using the CNS software package (32), with three cycles of conjugate gradient. This final model was also evaluated using MODELLER 8v1 to calculate the DOPE score (the energy profile graph of each residue). Both the DOPE score and PROCHECK confirm that our model is of high quality, with scores comparable with that of template structure (−48052 for rhodopsin *versus* −42544 for MOR42-3).

Docking Analysis—Molecular docking of ligands into the OR homology model was carried out using the AUTODOCK v3.0 program with the Lamarckian genetic algorithm search method (33). The receptor was kept rigid, while the ligand was allowed to be flexible and translate/rotate. Polar hydrogens were added to the receptor, and Kollman-united atom partial charges along with atomic solvation parameters were assigned to individual protein atoms. The three-dimensional structure of each ligand was generated using the PRODRG program (33) (davapc1.bioch.dundee.ac.uk/cgi-bin/prodrg). Ligands were drawn using the JME molecular editor and then energy minimized with GROMOS (34). For each ligand, a rigid root and rotatable bonds were assigned automatically, except for C8–C12 dioic acids where the C–C bond next to the carboxylic group was set as non-rotatable. For each ligand the non-polar hydrogens were removed and the partial charges from these were added to the carbon (Gasteiger charges). The atom type for aromatic carbons (in 1,4 phenylene dipropionic acid) was reassigned in order to use the AUTODOCK aromatic carbon grid map. Docking was carried out using a $80 \times 80 \times 80$ grid points with a default spacing of 0.375 Å. The grid was positioned to include the full ligand binding pocket in the central part of the OR and to allow extensive sampling around residue Val-113. Within this grid, the Lamarckian genetic search algorithm was used with a population size of 150 individuals, calculated using 100 different runs (*i.e.* 100 dockings). Each run had two stop criteria, a maximum of 2.7×10^6 energy evaluations or a maximum of 27,000 generations, starting from a random position and conformation; default parameters were used for the Lamarckian genetic algorithm search.

RESULTS

Initial Mutation Screen of MOR42-3—The overall sequence identity between MOR42-3 and MOR42-1 is 89%. However, within the transmembrane region (encompassing TMs I–VII, as well as the intracellular and extracellular loops) the degree of identity is 97%, with only 9 residues differing between these two receptors (see alignment in supplemental Fig. S1). These residues served as the starting point in our effort to identify the ligand-binding domain of MOR42-3 and to understand the molecular basis for the ligand preference of this receptor. We mutated each of these residues to the corresponding residues in MOR42-1. These mutations were: V82L, I112V, V113S, F114M, V202A, T205I, V206I, I208T, and A220T. MOR42-3, MOR42-1, and each of the 9 mutant MORs were expressed in *Xenopus* oocytes (see "Experimental Procedures") and screened with a series of aliphatic dicarboxylic acids, 8–12 carbons in length, each at 30 μM (Fig. 1; for structures, see supplemental Fig. S2). The V113S mutant displayed a ligand specificity that was intermediate between MOR42-3 and MOR42-1 (Fig. 1*A*). This mutant showed significantly increased responses to C10, C11, and C12 dicarboxylic acids, as compared with the responsiveness of the wild-type MOR42-3 (Fig. 1*B*, supplemental Table S1). The other mutants showed either no effect (*e.g.* V82L and I208T) or had subtle effects (*e.g.* the decrease in C8 responsiveness for I112V and V202A). We pursued a further examination of some of these positions with a series of double mutations (see below). At position 113, the change from valine to serine is a change in both the volume and hydrophobicity of the side chain. The decreased side chain volume of the serine at this position in MOR42-1 is permissive for ligands with longer carbon lengths. To further probe the role of this position, we tested a V113A mutant. The alanine residue has a volume similar to that of serine (89 *versus* 94 $Å^3$, respectively) (35) but preserves the hydrophobic character of valine. We found that the ligand specificity of the V113A mutant was also intermediate between MOR42-3 and MOR42-1 (supplemental Fig. S3). Our results suggest that Val-113 in MOR42-3 imposes a spatial restraint that prevents the receptor from responding to long chain dicarboxylic acids.

Development of a Homology Model of MOR42-3—To understand the molecular features of MOR42-3 that may be involved in ligand binding and specificity, we generated a homology model as described under "Experimental Procedures" (Fig. 1*C*). A series of highly conserved residues in the TM domains of MOR42-3 allowed unambiguous alignment with rhodopsin (supplemental Fig. S1). Our model retained the ellipsoidal central cavity lined by TM helices I–III and V–VII. Helix IV is not part of the cavity wall and only makes contacts with helix III, as seen in rhodopsin (36). Several highly conserved motifs in GPCRs are also present in the MOR42 subfamily, including the (D/E)R(Y/W) motif located near the cytoplasmic end of TMIII and the NP*XX*Y motif in TMVII. These motifs are thought to form a hydrogen bond network important for inter-helical interactions (37) and may play a role in receptor transformation from an inactive form to an active, G protein-coupled conformation (38). Conserved prolines that produce kinks in the helix structure of MOR42-3 are Pro-162 (in TMIV, equivalent to

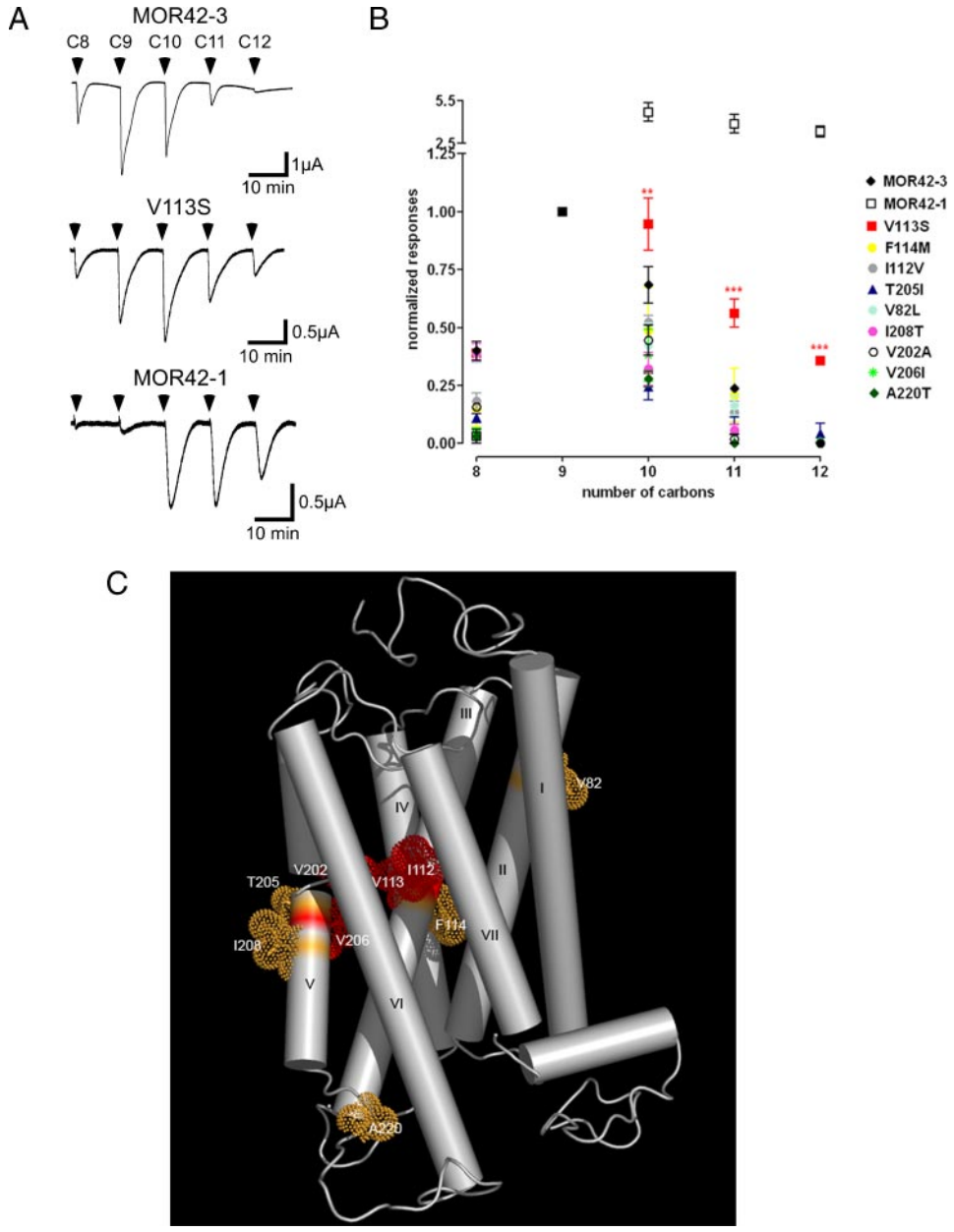

FIGURE 1. **Functional screen of mutations in MOR42-3.** *A*, oocytes expressing MOR42-3 (*top*), the V113S mutant of MOR42-3 (*middle*), or MOR42-1 (*bottom*) were challenged with 30 μM dicarboxylic acids of varying carbon length. *B*, summary of four to five recordings from oocytes expressing MOR42-3, MOR42-1, or one of nine mutants of MOR42-3. Responses were normalized to the response of the same oocyte to 30 μM C9 and were presented as mean $\pm$ S.E. Significant differences from MOR42-3: *, $p < 0.05$; **, $p < 0.01$. *C*, homology model of MOR42-3 in which each helix is represented as a *cylinder*. The *dotted* molecular surfaces show positions of the 9 residues in the transmembrane region that differ between MOR42-3 and MOR42-1: TMII (Val-82), TMIII (Ile-112, Val-113, Phe-114), TMV (V202A, T205I, V206I, and I208T) and in the long IC3 loop (A220T). Residues shown to have significant effects on ligand specificity are shown in *red*. The image was generated using PyMOL (49) with the extracellular region of the receptor at the top of the figure.

Pro-170 in rhodopsin) and Pro-303 (in TMVII, position 291 in rhodopsin). Pro-267 in the middle of the helix VI of rhodopsin also makes a significant bend of the helix structure but is not found in the members of the MOR42 subfamily. Movement of helix 6 is thought to be important for receptor activation (39), and the absence of the proline in helix VI of MOR42-3 may indicate that the activation process differs in these MORs. The short helix VIII that runs parallel to the cytoplasmic surface of the membrane is also present in our model, as are several conserved positively charged amino acids in this helix. Extracellular loop 2 (EC2) is a 24-residue loop that has 3 conserved cysteines (173, 183, 193), one of which very likely forms a disulfide bond with an equally well conserved cysteine in TMIII (at position 101 in MOR42-3). We modeled a disulfide bond between 101–173 positions, because that configuration produced the most favorable energy (see "Experimental Procedures"). This disulfide bond constrains EC2 toward the helical core and governs its topology. A disulfide bond at this position is highly conserved in rhodopsin-like GPCRs and is important in the mechanism of activation, as well as in the formation and folding of rhodopsin (40, 41). It should be noted, however, that the extracellular loops are the most imprecise features of our homology model. Much of the OR sequence is either absent from the rhodopsin crystal structure or so divergent as to preclude unambiguous alignment.

Further Mutation Analysis at the Putative Ligand Binding Site—Our initial mutation screen indicated valine 113 determines the dicarboxylic acid carbon length preference. Based on proximity to Val-113 in our model, we explored the role of additional residues that differ between MOR42-3 and MOR42-1 (Fig. 1*C*). We tested a series of double mutants consisting of the V113S mutation and each of four other mutations (I112V, F114M, V202A, and V206I). Each double mutant was screened with the C8-C12 dicarboxylic acids as described above. The results of this screen are shown in Fig. 2. The I112V/V113S, V202A/V113S, and V206A/V113S double mutants were each significantly more responsive to the C10 and C11 dicarboxylic acids than was the V113S single mutant, indicating that the ligand specificities of these double mutants were even further shifted toward the ligand specificity of MOR42-1. The ligand specificity of the F114M/V113S double mutant did not differ from V113S. Our results demonstrate that residues Ile-112, Val-113, Val-202, and Val-206 are involved in determining the carbon length preference of MOR42-3 and are likely to form part of the ligand binding pocket in this receptor.

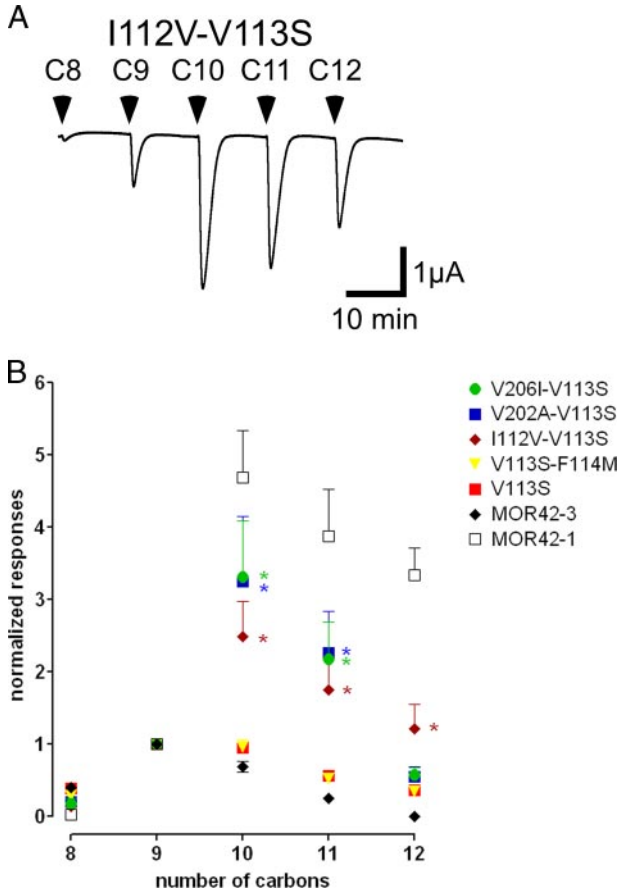

FIGURE 2. **Double mutant screen of MOR42-3.** *A*, an oocyte expressing the I112V/V113S double mutant of MOR42-3 was challenged with 30 μM dicarboxylic acids of varying carbon length. *B*, summary of four to seven recordings from oocytes expressing MOR42-3, MOR42-1, the V113S mutant of MOR42-3, or one of four double mutants of MOR42-3. Responses were normalized to the response of the same oocyte to 30 μM C9 and were presented as the mean ± S.E. Significant differences from the V113S mutant: *, $p < 0.05$.

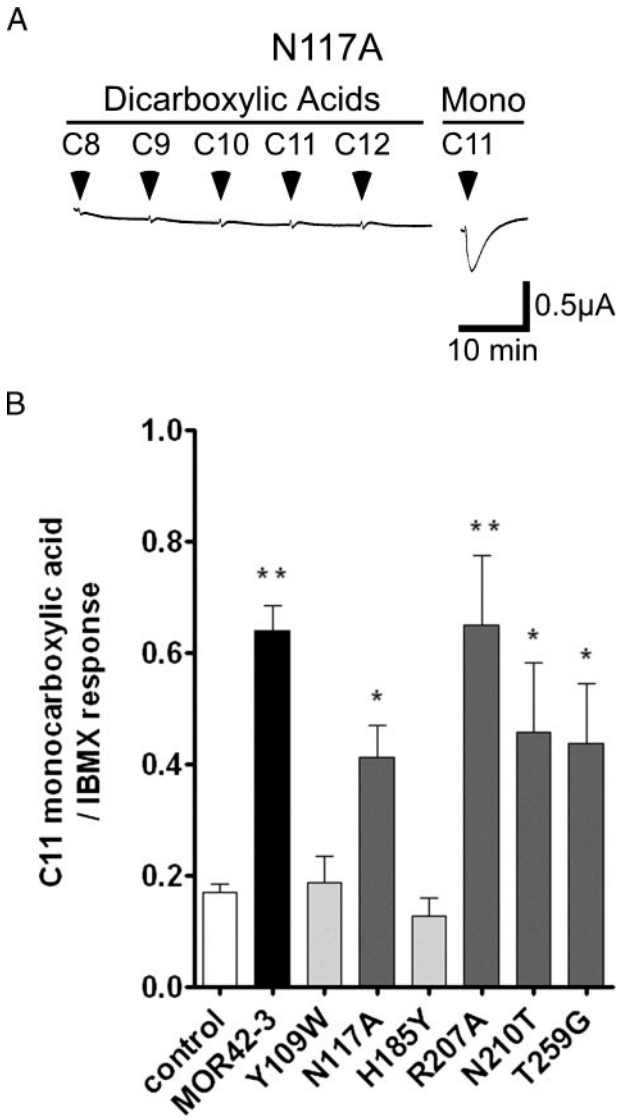

FIGURE 3. **Mutations that eliminate dicarboxylic acid sensitivity.** *A*, an oocyte expressing the N117A mutant of MOR42-3 failed to respond to 30 μM dicarboxylic acids of varying carbon length but responded to 30 μM C11 monocarboxylic acid. *B*, summary of the responses to C11 monocarboxylic acid of oocytes expressing MOR42-3 or one of six mutants of MOR42-3. Responses were normalized to the response of the same oocyte to 1 mM isobutylmethylxanthine and were presented as the mean ± S.E. ($n = 4$–12). Control oocytes were injected with no receptor cRNA, cRNA encoding the rat $5HT_3$ receptor, or cRNA encoding the human M2 muscarinic receptor (no differences were observed between these three sets, and so they were combined). 30 μM C11 monocarboxylic acid caused small currents in control samples, due to direct activation of cystic fibrosis transmembrane regulator. For this reason, all responses in MOR-expressing oocytes were compared with the control oocytes. Significant differences from the control: *, $p < 0.05$; **, $p < 0.01$.

An additional set of mutations was chosen based on proximity to Val-113 in our receptor model (R207A and N210T in TMV, T259G in TMVI) and based on inclusion in the set of 22 predicted ligand binding residues (1) (Y109W, N117A, H185Y). Residues at these six positions are identical in both MOR42-3 and MOR42-1. Thus, we mutated each residue to the corresponding residue in MOR42-2 (a more distantly related member of the MOR42 subfamily). If all three receptors had the same residue at a certain position, we mutated it to alanine. One exception was Y109W in TMIII. By introducing the bulkier hydrophobic tryptophan at this position, we sought to impede access of dicarboxylic acids to the ligand binding pocket. None of these six mutant receptors responded to dicarboxylic acids. However, we found that the N117A, R207A, N210T, and T259G mutant receptors each responded to the 11-carbon monocarbocylic acid (30 μM), indicating that the these mutant receptors are properly folded and expressed on the cell surface (Fig. 3). These mutations produced receptors able to accommodate only one carboxylic acid moiety. The other two mutants (Y109W and H185Y) did not yield specific responses to any tested ligand (including undecanoic acid), suggesting that these receptors are either non-functional or not expressed. Our mutagenesis studies yielded a network of 8 residues that are important in defining the ligand specificity of MOR42-3 (Fig. 4*A*). Residues Ile-112, Val-113, Val-202, and Val-206 are involved in conferring carbon length preference, whereas residues Asn-117, Arg-207, Asn-210, and Thr-259 appear to be involved in the accommodation of a second carboxylic acid moiety.

Docking of Ligands into MOR42-3—We performed ligand-docking simulations with the C8–C12 dicarboxylic acids (Fig. 5*A*; supplemental Fig. S4). Docking computations were performed with AUTODOCK 3.0 (see "Experimental Proce-

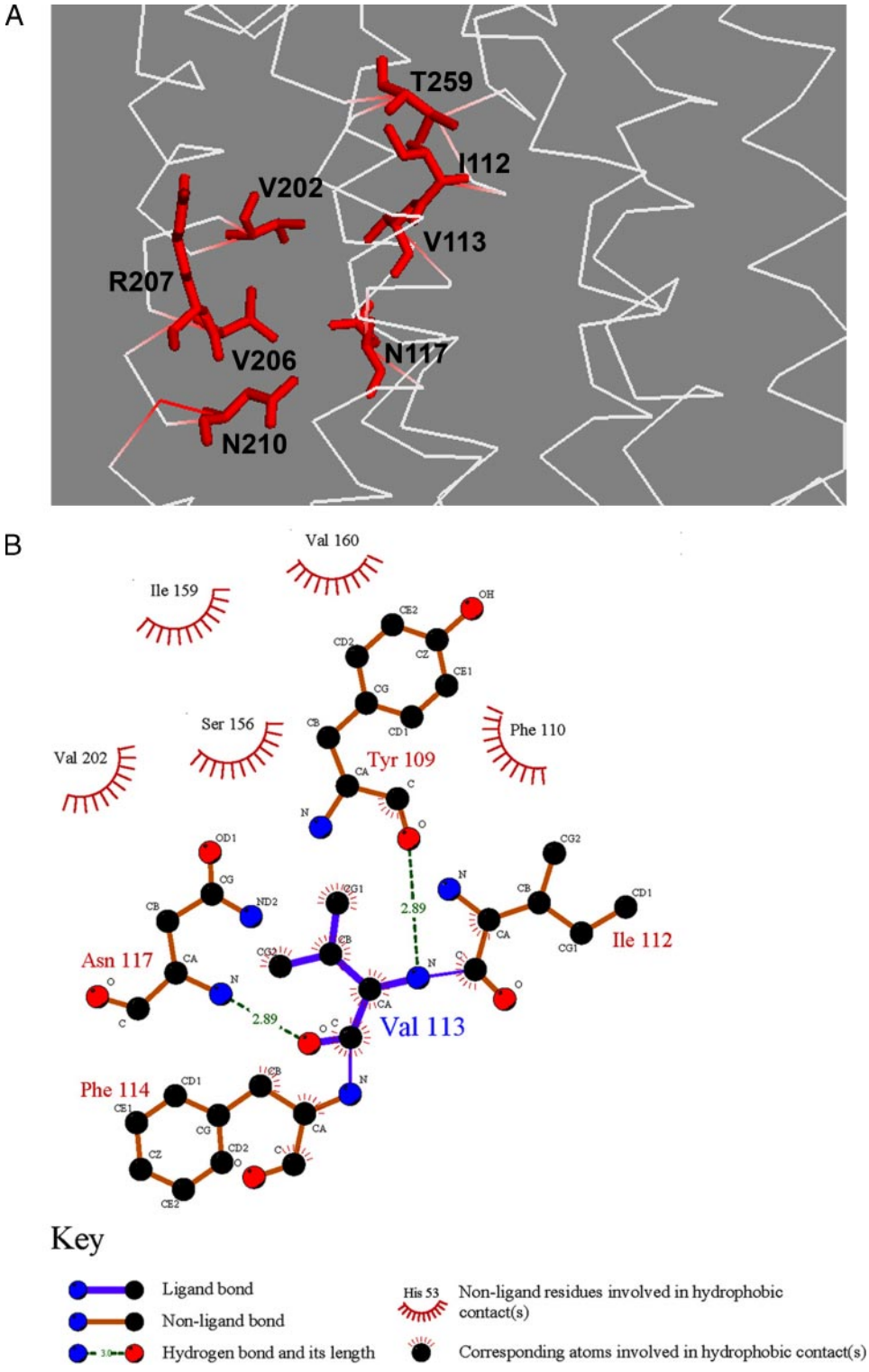

FIGURE 4. **A network of 8 functionally important residues that constitute part of the binding pocket for dicarboxylic acids in MOR42-3.** *A*, the 8 residues shown as *ball and stick* on a Cα trace of the TM region of the MOR42-3 homology model. The image was generated using PyMOL (49) and was oriented with the extracellular region on the top. *B*, intramolecular interactions of valine 113 in the MOR42-3 model were shown using LIGPLOT (50). Valine 113 forms hydrophobic interactions with Phe-110 (TMIII), Ser-156, Ile-159, Val-160 (TMIV), and Val-202 (TMV) and hydrogen bonds with Tyr-109 and Asn-117.

dures"). AUTODOCK calculates two kinds of free energies, free binding energy that includes the intermolecular energy and torsional free energy and docking energy that includes the intermolecular and intramolecular energy. The former is presented at the end of docking simulations, whereas the latter is used for selecting better individuals from a population during a docking analysis. For each ligand, the results from docking analysis yielded 100 conformations. Each conformation is compared with one another. Conformations with a root mean square deviation <0.5 Å were clustered together. Clusters were ranked in order of increasing docking energy. The optimal result is for the lowest energy cluster to be the most populated cluster; however, this was not always the case. For nonanedioic acid, the preferred ligand for MOR42-3, 100 docking conformations were grouped into 50 clusters. The lowest docking energy cluster (-12.16 kcal/mol) was also the most populated cluster with 18 conformations, and this vertical conformation represents the predominant docking orientation for this ligand (Fig. 5*A*, supplemental Fig. S4). Docking analysis of the C8 and C10 ligands also yielded vertical conformations as the lowest docking energy clusters. However, for both ligands a horizontal orientation was observed in a highly populated but higher docking energy cluster (supplemental Fig. S4). Docking analysis of the C11 and C12 ligands produced 84 and 97 different clusters, respectively, reflecting the higher number of rotatable bonds in these ligands that made it difficult to find the best binding mode. The large number of rotatable bonds in all tested ligands also explains small differences, both in docking and binding energies. In agreement with our mutagenesis data, the results of docking analyses indicate a ligand binding pocket located in the vicinity of residue 113 of MOR42-3.

Identification of a Partial Agonist and an Antagonist for MOR42-3—Our docking simulations indicate that the C11 and C12 dicarboxylic acids bind to MOR42-3 and that the binding domains of these ligands partially overlap with the C9 binding site. However, in the functional assay, MOR42-3 responded poorly to the C11 and not at all to the C12 (Fig. 1; see also Ref. 25), suggesting that C11 and C12 might be a partial agonist and an antagonist, respectively.

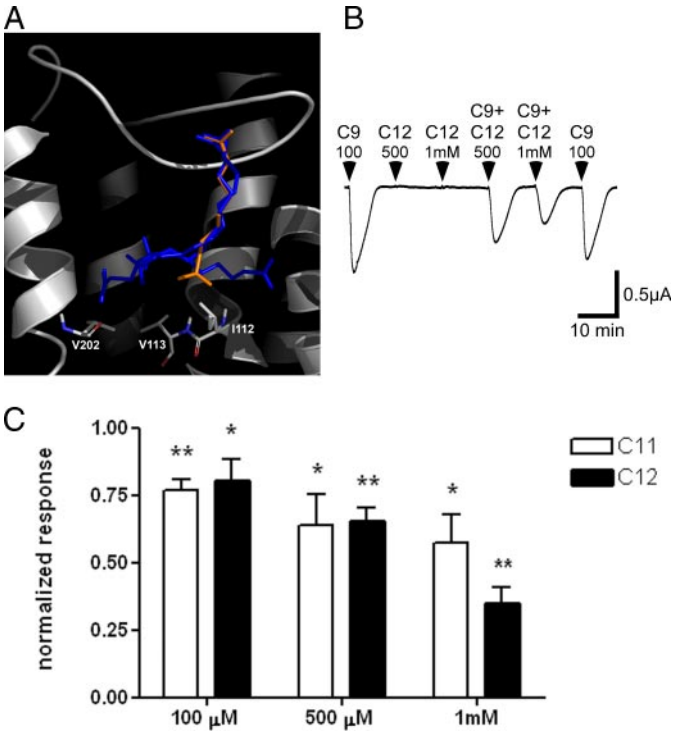

FIGURE 5. **Identification of an antagonist of MOR42-3.** *A*, overlapping conformations of C9 (*orange*) and C12 (*blue*, three conformations) in MOR42-3 model derived from docking simulations. Residues Ile-112, Val-113, and Val-202 were indicated. *B*, an oocyte expressing MOR42-3 responds to 100 μM C9 dicarboxylic acid, but not 500 μM or 1 mM C12 dicarboxylic acid. Co-application of C12 antagonizes the response to C9. The C9 response in the presence of C12 was compared with the mean of the first and last response to C9 alone. *C*, antagonism of C9 responses by co-application of increasing concentrations of C11 or C12 dicarboxylic acids ($n = 4$). Responses were normalized to 100 μM C9 response in each oocyte and presented as the mean ± S.E. Significant differences from the C9 response: *, $p < 0.05$; **, $p < 0.01$.

We tested this possibility in the functional assay (Fig. 5). Whereas C12 was unable to activate MOR42-3, co-application of C12 reduced the response to C9, indicating that C12 is an antagonist (Fig. 5, *B* and *C*). Similarly, whereas C11 yielded only modest receptor activation, co-application of C11 with C9 reduced the response to C9, indicating that C11 is a partial agonist of MOR42-3 (Fig. 5*C*).

We also tested a series of cyclic dicarboxylic acids for potential antagonist activity. These compounds were: 1,4 phenylene diacetic, dipropionic and diacrylic acids, 2,2′ bipyridine 5,5′ dicarboxylic, and 2,2′ biquinoline 4,4′ dicarboxylic acid (supplemental Fig. S2). Each compound (500 μM) was tested for the ability to activate or antagonize MOR42-3. As we have previously reported (25), 1,4 phenylene dipropionic acid displayed modest agonist activity (~30% of the response to 100 μM C9) for MOR42-3. However, this compound failed to reduce the response to C9, indicating that it is a low potency agonist. Docking simulations of 1,4 phenylene dipropionic acid showed docking in the vicinity of Val-113 (data not shown). MOR42-1 is more responsive to this compound (25) and has the smaller serine at this position, suggesting that 1,4 phenylene dipropionic acid would be better accommodated. None of the other cyclic compounds could activate or antagonize MOR42-3.

DISCUSSION

We have identified 8 residues that participate in determining the ligand specificity of MOR42-3, Ile-112, Val-113, Asn-117, Val-202, Val-206, Arg-207, Asn-210, and Thr-259 (Fig. 4*A*). These residues are located in TMs III, V, and VI. The most structurally and functionally important of these residues is valine 113. What is the role of Val-113 residue? We postulate that the methyl groups of Val-113 form hydrophobic interactions with nearby residues from other helixes, thereby decreasing the size of the lower portion of the binding pocket. In our model, Val-113 shows hydrophobic interactions with Phe-110 (TMIII), Ser-156, Ile-159, Val-160 (TMIV), and Val-202 (TMV), bringing helixes III, IV, and V closer together (Fig. 4*B*). Upon introduction of the polar serine at this position, these hydrophobic interactions are lost, thereby opening the pocket and permitting accommodation of larger molecules. In addition, the smaller side chain volumes resulting from the I112V, V113S, and V202A mutations would also contribute to enlargement of the binding pocket. Our observation that the effects of the I112V and V202A mutations are only apparent in the context of the V113S mutation (compare the double mutant results in Fig. 2 with the single mutant results in Fig. 1) suggests that the effect of the side chain volume reduction is less important than the loss of the interhelical hydrophobic interactions effected by the V113S mutation. This postulated difference in binding site size would also explain our previous experimental results (25) showing much larger responses to 1,4-phenylene dipropionic acid with MOR42-1 than with MOR42-3. Docking analysis shows that this cyclic compound binds deep inside the receptor where it can be better accommodated if the TM hydrophobic interactions are weakened and the binding site enlarged (data not shown).

Based on our homology model and proximity to the Val-13 residue, we identified four additional residues that help determine the ligand specificity of MOR42-3, Asn-117 in TMIII, Arg-207 and Asn-210 in TMV, and Thr-259 in TMVI. In contrast to rhodopsin, TMV of MOR42-3 does not form a single continuous helix but has a short stretch of structurally flexible residues. This offers the potential for movement during receptor activation and might allow residues located further down the helix (such as Arg-207 and Asn-210) to become exposed to the binding pocket located above the centrally located Val-113. Similar observations have been documented for Ser-193, Ser-194, and Ser-197 in TMV of D2 dopamine receptor (42).

The concepts of orthology and paralogy have become important as whole genome comparison allows their identification in complete genomes. The functional specificity of proteins is assumed to be conserved among orthologs but not among paralogs. Lancet and co-workers (1) have used this assumption, comparing human-mouse ortholog/paralog pairs, to identify a series of 22 residue positions as candidate determinants of ligand recognition by ORs. Several of these predicted binding site residues have been shown to participate in ligand recognition in mouse euginol receptor (24). Only 2 of the 22 predicted binding site residues (Ile-112 and Val-113) differ between MOR42-3 and MOR42-1. Our results provide experimental confirmation for the involvement of these residues in forming a

ligand-binding domain for dicarboxylic acids in MOR42-3. In addition, we identified the Asn-117 position, also a predicted binding site residue, as important in the ability of the receptors to accommodate a second carboxylic acid moiety.

Our docking simulations show dicarboxylic acids docking into a binding pocket located toward the extracellular end of the receptor. Although a homology model based on the inactive structure of rhodopsin can be a valuable tool in receptor research and ligand design (43), the modeled structure may not be representative of the additional conformations that the receptor assumes upon activation. This is important, because agonists bind to these active receptor conformations (44). Thus, understanding the structural basis for agonist binding and efficacy will involve understanding the conformational changes associated with receptor activation (45). Our studies cannot address this issue because our model represents an inactive form of the receptor and is kept static during docking simulations, whereas the binding sites of ORs may display structural plasticity during receptor activation. We also do not exclude the possibility that the quality of our modeled structure may be constrained by the limited amount of available structural data (in particular, the low sequence similarity between the rhodopsin template and MOR42-3). For these reasons, information about the binding of ligands to ORs based solely on docking simulations in models should be considered with caution. It is critical that docking predictions be tested in a functional context. The close agreement between our mutagenesis and functional studies, and our simulated docking results, increases our confidence that we have identified the ligand binding pocket of MOR42-3. Recently, docking simulations were used to predict citric acid as a ligand for MOR42-3 (46). When tested in our system, 500 μM citric acid did not activate or antagonize MOR42-3 (data not shown). Our results with C12 carboxylic acid suggest that binding energies alone are not sufficient to predict agonist or antagonist activity of a ligand.

The various issues that make interpretation of simulated agonist docking difficult suggest that it may be easier to predict the binding domain of an antagonist, which is likely to bind best to the inactive conformation of the receptor. The ability of the C12 ligand to bind in our simulated docking studies, despite the fact that C12 does not activate the receptor, suggested that this compound might be an antagonist. This prediction was confirmed in our functional assay. The existence of antagonists for ORs has been shown in a few studies (47, 48). Our results with the C12 dicarboxylic acid provide further insight into the role of OR antagonism. We find that one ligand (C12) can act on two closely related receptors in a concerted fashion, activating one receptor (MOR42-1) while at the same time inhibiting the response of the other receptor (MOR42-3). Thus, the relative proportions of compounds in an odorant mixture can contribute to the fine-tuning of the receptor response. This is an example of functional tuning of olfactory receptors that may represent a previously unrecognized mechanism critical to the overall perception of odor.

Acknowledgments—We thank Ana Mederos and Floyd Maddox for expert technical assistance.

REFERENCES

1. Man, O., Gilad, Y., and Lancet, D. (2004) *Protein Sci.* **13,** 240–254
2. Zelano, C., and Sobel, N. (2005) *Neuron* **48,** 431–454
3. Buck, L., and Axel, R. (1991) *Cell* **65,** 175–187
4. Schild, D., and Restrepo, D. (1998) *Physiol. Rev.* **78,** 429–466
5. Malnic, B., Hirono, J., Sato, T., and Buck, L. B. (1999) *Cell* **96,** 713–723
6. Lewcock, J. W., and Reed, R. R. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 1069–1074
7. Serizawa, S., Ishii, T., Nakatani, H., Tsuboi, A., Nagawa, F., Asano, M., Sudo, K., Sakagami, J., Sakano, H., Ijiri, T., Matsuda, Y., Suzuki, M., Yamamori, T., Iwakura, Y., and Sakano, H. (2000) *Nat. Neurosci.* **3,** 687–693
8. Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y., and Sakano, H. (2003) *Science* **302,** 2088–2094
9. Duchamp-Viret, P., and Duchamp, A. (1997) *Prog. Neurobiol.* (*Oxf.*) **53,** 561–602
10. Firestein, S. (2001) *Nature* **413,** 211–218
11. Mombaerts, P. (1999) *Annu. Rev. Neurosci.* **22,** 487–509
12. Pilpel, Y., and Lancet, D. (1999) *Protein Sci.* **8,** 969–977
13. Baldwin, J. M., Schertler, G. F. X., and Unger, V. M. (1997) *J. Mol. Biol.* **272,** 144–164
14. Zhang, X., and Firestein, S. (2002) *Nat. Neurosci.* **5,** 124–133
15. Oliveira, L., Paiva, A. C., and Vriend, G. (2002) *Chembiochem.* **3,** 1010–1017
16. Suel, G. M., Lockless, S. W., Wall, M. A., and Ranganathan, R. (2003) *Nat. Struct. Biol.* **10,** 59–69
17. Hobrath, J. V., and Wang, S. (2006) *J. Med. Chem.* **49,** 4470–4476
18. Madabushi, S., Yao, H., Marsh, M., Kristensen, D. M., Philippi, A., Sowa, M. E., and Lichtarge, O. (2002) *J. Mol. Biol.* **316,** 139–154
19. Singer, M. S., Oliveira, L., Vriend, G., and Shepherd, G. M. (1995) *Receptors Channels* **3,** 89–95
20. Singer, M. S. (2000) *Chem. Senses* **25,** 155–165
21. Vaidehi, N., Floriano, W. B., Trabanino, R., Hall, S. E., Freddolino, P., Choi, E. J., Zamanakos, G., and Goddard, W. A., III (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 12622–12627
22. Floriano, W. B., Vaidehi, N., Goddard, W. A., III, Singer, M. S., and Shepherd, G. M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 10712–10716
23. Hall, S. E., Floriano, W. B., Vaidehi, N., and Goddard, W. A., III (2004) *Chem. Senses* **29,** 595–616
24. Katada, S., Hirokawa, T., Oka, Y., Suwa, M., and Touhara, K. (2005) *J. Neurosci.* **25,** 1806–1815
25. Abaffy, T., Matsunami, H., and Luetje, C. W. (2006) *J. Neurochem.* **97,** 1506–1518
26. Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P., and Buss, V. (2004) *J. Mol. Biol.* **342,** 571–583
27. Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol.* **234,** 779–815
28. Vullo, A., and Frasconi, P. (2004) *Bioinformatics* **20,** 653–659
29. Zozulya, S., Echeverri, F., and Nguyen, T. (2001) *Genome Biol.* **2,** RESEARCH0018
30. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
31. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** Pt. 2, 110–119
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., and Pannu, N. S. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
33. Morris, G. M., Goodsell, D. S., Halliday, R. S., Ruth Huey, W. W., Hart, R. K., and Olson, B. A. A. J. (1998) *J. Comput. Chem.* **19,** 1639–1662
34. van Gunsteren, W. F., and Mark, A. E. (1992) *Eur. J. Biochem.* **204,** 947–961
35. Tsai, J., Taylor, R., Chothia, C., and Gerstein, M. (1999) *J. Mol. Biol.* **290,** 253–266
36. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) *Science* **289,** 739–745
37. Menon, S. T., Han, M., and Sakmar, T. P. (2001) *Physiol. Rev.* **81,** 1659–1688

38. Robinson, P. R., Cohen, G. B., Zhukovsky, E. A., and Oprian, D. D. (1992) *Neuron* **9,** 719–725
39. Shi, L., Liapakis, G., Xu, R., Guarnieri, F., Ballesteros, J. A., and Javitch, J. A. (2002) *J. Biol. Chem.* **277,** 40989–40996
40. Davidson, F. F., Loewen, P. C., and Khorana, H. G. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 4029–4033
41. Hwa, J., Klein-Seetharaman, J., and Khorana, H. G. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 4872–4876
42. Javitch, J. A., Fu, D., and Chen, J. (1995) *Biochemistry* **34,** 16433–16439
43. Varady, J., Wu, X., Fang, X., Min, J., Hu, Z., Levant, B., and Wang, S. (2003) *J. Med. Chem.* **46,** 4377–4392
44. Gether, U., and Kobilka, B. K. (1998) *J. Biol. Chem.* **273,** 17979–17982
45. Dima, R. I., and Thirumalai, D. (2006) *Protein Sci.* **15,** 258–268
46. Floriano, W. B., Vaidehi, N., and Goddard, W. A., III (2004) *Chem. Senses* **29,** 269–290
47. Oka, Y., Omura, M., Kataoka, H., and Touhara, K. (2004) *EMBO J.* **23,** 120–126
48. Shirokova, E., Schmiedeberg, K., Bedner, P., Niessen, H., Willecke, K., Raguse, J. D., Meyerhof, W., and Krautwurst, D. (2005) *J. Biol. Chem.* **280,** 11807–11815
49. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, San Carlos, CA
50. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Eng.* **8,** 127–134

The Molecular Basis for Ligand Specificity in a Mouse Olfactory Receptor: A NETWORK OF FUNCTIONALLY IMPORTANT RESIDUES

Tatjana Abaffy, Arun Malhotra and Charles W. Luetje

J. Biol. Chem. 2007, 282:1216-1224.
doi: 10.1074/jbc.M609355200 originally published online November 17, 2006
