Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue



Cite this: *Org. Biomol. Chem.*, 2018, **16**, 2541

# Carbon chain shape selectivity by the mouse olfactory receptor OR-I7†

Min Ting Liu,‡[a,b] Jianghai Ho,‡[c] Jason Karl Liu,[d] Radhanath Purakait,[a] Uriel N. Morzan,[d] Lucky Ahmed,[d] Victor S. Batista,[d] Hiroaki Matsunami*[c] and Kevin Ryan*[a,b,e]

Received 24th January 2018,
Accepted 14th March 2018

DOI: 10.1039/c8ob00205c

rsc.li/obc

The rodent OR-I7 is an olfactory receptor exemplar activated by aliphatic aldehydes such as octanal. Normal alkanals shorter than heptanal bind OR-I7 without activating it and hence function as antagonists *in vitro*. We report a series of aldehydes designed to probe the structural requirements for aliphatic ligand chains too short to meet the minimum approximate 6.9 Å length requirement for receptor activation. Experiments using recombinant mouse OR-I7 expressed in heterologous cells show that in the context of short aldehyde antagonists, OR-I7 prefers binding aliphatic chains without branches, though a single methyl on carbon-3 is permitted. The receptor can accommodate a surprisingly large number of carbons (*e.g.* ten in adamantyl) as long as the carbons are part of a conformationally constrained ring system. A rhodopsin-based homology model of mouse OR-I7 docked with the new antagonists suggests that small alkyl branches on the alkyl chain sterically interfere with the hydrophobic residues lining the binding site, but branch carbons can be accommodated when tied back into a compact ring system like the adamantyl and bicyclo[2.2.2]octyl systems.



## Introduction

The mammalian olfactory (a.k.a. odorant) receptors (ORs) form the largest family of G-protein coupled receptors (GPCRs) in the human and rodent genomes.[1–3] Mice, for example, are predicted to have over 1000 different OR genes, while humans have approximately 390 ORs out of a total of about 825 predicted GPCR genes. Within the context of the GPCR structure,[4,5] each OR is expected to form within its 7-transmembrane alpha helical (TM) bundle a unique binding site with distinct ligand-binding properties resulting from the convergence of receptor-specific residues, mainly from TM3, TM5, TM6 and TM7.[6–10] Some receptors, like the rodent I7 receptor (OR-I7; a.k.a. MOR103-15 and Olfr2),[11] which we study here, appear to be highly specific for odorant ligand traits such as functional group and carbon chain length, but others appear to lack ligand specificity, at least *in vitro*.[12,13] The requirement for volatility puts a limit on an odorant's molecular weight and number of polar functional groups, but terrestrial ORs have nevertheless evolved to bind innumerable small, usually hydrophobic ligands which offer a receptor limited opportunity for hydrogen-bonding and other polar interactions. Aliphatic odorants such as monoterpenoids (*e.g.* geraniol), sesquiterpenoids (*e.g.* santalols), and those derived from fatty acid biogenic precursors (*e.g.* octanal) typically have only one polar functional group, and many hydrocarbons are found among fragrant natural products. The specific manner in which the ORs interact with the hydrocarbon portion of an odorant remains entirely unknown.

The terpenoid, fatty acyl-derived and hydrocarbon odorants therefore present interesting molecular recognition puzzles because their odor character appears to depend heavily on attributes of their carbon skeletons, including features such as length, size and shape. Using calcium imaging of dissociated olfactory receptor neurons (ORNs, a.k.a. odorant sensory neurons, OSNs), evidence for a correlation was previously found between ligand conformational flexibility and the number of different ORs a ligand activated when the sole polar functional group (an aldehyde) and the number of carbons

[a]*Department of Chemistry and Biochemistry, The City College of New York, New York, NY 10031, USA. E-mail: kryan@ccny.cuny.edu*
[b]*Ph.D. Program in Chemistry, The Graduate Center of the City University of New York, New York, NY 10016, USA*
[c]*Department of Molecular Genetics and Microbiology, and Neurobiology, Duke Institute for Brain Sciences, Duke University Medical Center, Durham, NC 27710, USA. E-mail: matsu004@mc.duke.edu*
[d]*Department of Chemistry, Yale University, New Haven, CT 06520, USA*
[e]*Ph.D. Program in Biochemistry, The Graduate Center of the City University of New York, New York, NY 10016, USA*

†Electronic supplementary information (ESI) available: Synthetic procedures and characterization of analogues **2**, **5**, **6**, **7**, **8**, **9**, **10**, **11** and **12**; time-course dose response data for mOR-I7 in Hana3A cells; and mOR-I7 homology model comparison with rhodopsin and rat OR-I7 structures. This material is available free of charge *via* the journal website. See DOI: 10.1039/c8ob00205c

‡Contributed equally to this work.

(eight) were kept constant.[14] This evidence suggested that for some odorants, carbon chains with more well-defined shapes activate fewer receptors than more flexible analogs, and may thereby achieve unique olfactory codes[15] by activating smaller subsets of sensory neurons which, when mature, express only one type of OR. For example, the cyclic muscone family of odorants appears to activate only one major human OR,[13,16–19] while octanal, an odorant with an acyclic, conformationally unrestricted chain, activates 33–55 rat ORs.[20] By having better-defined shapes and also lower entropy loss upon binding, conformationally restricted ligands may be limited to binding fewer ORs than their more flexible relatives. To date, there are no atomic-level structural data on how odorants bind olfactory receptors. In this report we attempt to study the variable of carbon chain shapes in the context of a known OR-antagonist pair.

While studying conformational restriction of the octanal carbon chain as a determinant of OR-I7 activation, a relationship was previously found between restriction of the chain and whether an eight carbon aldehyde activated the receptor or bound without activation.[14] The present work is a follow-up to that study, whose findings we summarize here. Compound **1**, octanal, is a natural product activating ligand, or agonist, of OR-I7. Conceptually joining octanal's third and eighth carbons, to make cyclohexylethanal (compound **2**, Fig. 1), conformationally restricts the chain and prevents it from unfurling to an extended conformation. We found that this change converted octanal into a non-activating ligand, or antagonist (Fig. 1A).[14] Pentanal, compound **3**, which when extended has about the same chain length as **2**, was also found to be an antagonist, but was less potent than **2**. When two carbons were added back to **2**, to make (4-ethylcyclohexyl)ethanal, compound **4**, or to pentanal to make heptanal, the aldehydes regained their ability to function as agonists. These and prior findings led us to conclude that the OR-I7 receptor requires a minimum of two molecular features in a ligand, and a third feature if binding is to trigger activation. For binding, either as an agonist or antagonist, the aldehyde group is required (Fig. 1A, CHO recognition).[11,14] Next, for binding, but not necessarily activation, an aliphatic chain of at least five carbons total (as in pentanal, *e.g.*) is required.[14] We refer to octanal carbons-2 through -5 or -6, as the ligand "mid-region." For aldehydes longer than hexanal, for example, the mid-region connects the aldehyde to the third feature, a small hydrophobic group (*e.g.* carbons 7 and 8 in octanal) that must reach a putative small hydrophobic group binding pocket for OR-I7 activation. Aliphatic aldehydes shorter than the threshold of about 6.9 Å bind, but do not activate, OR-I7 and can thus function as antagonists.[14] We note that a similar dependence of activation on alkyl chain length has been found for another Class A GPCR, the cannabinoid CB1 receptor. The CB1 agonist $\Delta^9$-tetrahydrocannabinol (THC) contains a simple *n*-pentyl chain whereas the CB1 antagonist $\Delta^9$-tetrahydrocannabivarin (THCV) has an *n*-propyl chain but is otherwise identical.[21] Based on our initial short aldehyde data using rat OR-I7 there appeared to be a possible correlation between the number of carbons in the ligand's mid-region and the strength of the estimated $IC_{50}$ for antagonists that can inhibit the activation of OR-I7 by co-applied octanal.[14]

In this study, to extend the previous OR-I7 findings we have made a series of new OR-I7 antagonists to probe the mid-region requirements of aldehyde antagonists. Each compound was designed to have an extended length less than 6.9 Å to avoid receptor activation. We used methyl and ethyl groups and a variety of rings to increase the number of carbons within the mid-region. We find that the OR-I7 binding site in contact with the ligand's mid-region has a surprising capacity for aliphatic carbons, but prefers dense, compact rings with no branches, such as the bicyclo[2.2.2]octyl and adamantyl ring systems.

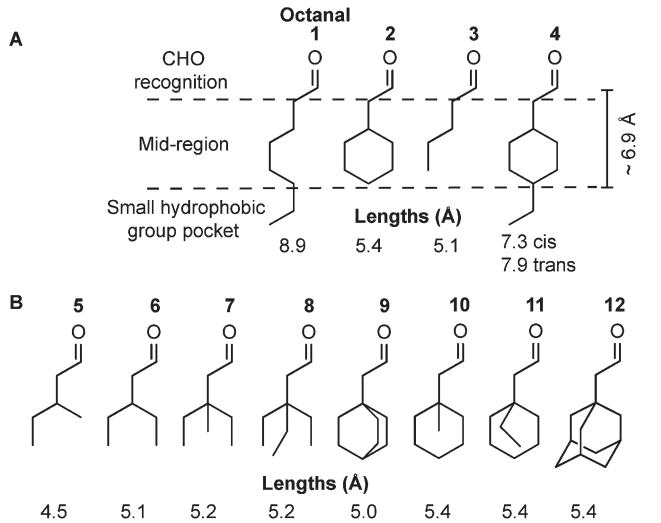



**Fig. 1** OR-I7 ligands and structural design of new antagonists. (A) The rodent OR-I7 olfactory receptor requires an aldehyde group and an aliphatic carbon chain (mid-region) of at least five carbons for antagonist ligand binding.[14] A small alkyl group such as ethyl, extending beyond ≈6.9 Å from the aldehyde, is required for receptor activation. Compounds **1** and **4** are agonists; **2** and **3** are antagonists. (B) Compounds designed in this study for antagonist structure–activity relationship analysis. Estimated lengths are for the most extended conformations after energy minimization.

# Experimental

### Method to estimate the maximum extended length of aldehydes

Chem3D Ultra 12.0 software (CambridgeSoft) was used. The structure of the aldehyde was drawn in its most extended conformation. The energy was minimized using the MM2 force field. The length was then measured from the carbonyl carbon to the most remote carbon.

### Chemical synthesis and characterization

The synthesis and characterization of the tested compounds is described in detail in the ESI.†

### Hana3A GloSensor cAMP assay

The GloSensor cAMP Assay System (Promega) was used according to the manufacturer's instructions with slight modifications. A plasmid encoding Rho-tagged mouse OR-I7 (80 ng per well) was transfected into the Hana3A cell line in 96-well plate (Biocoat; Becton Dickinson Biosciences) format along with plasmids encoding the human receptor trafficking protein, RTP1S (10 ng per well), type 3 muscarinic acetylcholine receptor (M3-R) (10 ng per well), and pGloSensor TM-22F (10 ng per well). Then, 18 to 24 h following transfection, cells were loaded with 2% GloSensor reagent for 2 h and treated with compounds in a total volume of 74 μL. Luminescence was measured using a Polarstar Optima plate reader (BMG) with a time interval of 90 seconds per well. Data were analyzed and $IC_{50}$s were estimated using Prism 5.0 and Microsoft Excel. Responses over $t$ = 3–7.5 minutes were summed, base-lined, normalized, and plotted *versus* odorant concentration.

### Homology model construction and ligand docking

A mouse OR-I7 homology model was constructed beginning with a previously published rat OR-I7 ortholog model.[22] The initial rat model was based on crystallographic data taken from rhodopsin PDB entry 1U19. The mouse model included the following rat-to-mouse ortholog substitutions (single-letter amino acid abbreviations; underscore indicates a predicted helical position in TM2-TM7, superscripts are Ballesteros–Weinstein numbering[23] for TM residues): V26A$^{1.33}$, M44I$^{1.51}$, I48T$^{1.55}$, K90E, V206I$^{5.41}$, F290L$^{7.49}$, D301E and R304K. A rat/mouse OR-I7 alignment with predicted helical regions can be found in this reference.[24] The ligand binding pocket of the mOR-I7 was predicted using the SiteMap module of Maestro software package (Version 10.2, Schrödinger, LLC, New York, NY, 2015). Protonation states of amino acids were assigned by using the PROPKA module, followed by structural relaxation using the preparation wizard tool.[25] The docking configurations of the ligands were analyzed by using the Glide module of Maestro.[26] Short (20 ns) molecular dynamics simulations were performed using the CHARMM force field, as implemented in NAMD.[27] During simulations, both octanal and the antagonists were observed to sample the different aldehyde group orientations shown in Fig. 5.

## Results and discussion

### Antagonist design

To probe the structural requirements of the OR-I7 ligand mid-region, but without extending into the receptor's small hydrophobic binding pocket beyond 6.9 Å from the aldehyde group, we began with two previously identified antagonists, pentanal ($IC_{50}$ 460 μM in rat OR-I7) and cyclohexylethanal (compound **2**, $IC_{50}$ 37 μM, also in rat OR-I7).[14] As shown in Fig. 1B, carbons were added incrementally to pentanal at carbon-3, to produce compounds **5–8**. (Compound **5** is the only chiral compound in the set and was made in racemic form.) The three ethyl groups radiating from carbon-3 in **8** were also conformationally restricted by incorporating them into the bicyclo[2.2.2]octyl ring system (compound **9**). For the set based on antagonist **2**, carbons were added incrementally at carbon-3 to make **10** and **11**. Compound **11** can be viewed as an incompletely restricted version of compound **9**. Lastly, the adamantyl ring in **12**, which is larger than **2** by four carbons but, like compound **9**, is highly restricted and has no protruding small alkyl groups, was used to probe the size limit of the mid-region.

### Synthesis of OR-I7 antagonists

The synthesis and characterization of designed antagonists **5–12** is described in detail in the ESI.†

### Antagonist testing

The $IC_{50}$ values previously estimated for compounds **2** and **3** were obtained in dissociated rat ORNs infected with an adenovirus vector carrying the rat OR-I7.[14,28] Despite one seminal report that the mouse ortholog, which has 15 amino acid differences overall but only 2 predicted to be in the TM2–TM7 helical regions, prefers heptanal over octanal when expressed in HEK293 cells,[29] we and others have found that both rat and mouse orthologs are aldehyde-specific and respond similarly to octanal and heptanal.[30,31] Here, we opted to continue using the mouse ortholog but as expressed in Hana3A cells. The Hana3A system consists of specially modified HEK293T cells, including the odorant receptor specific G-protein, $G_{olf}$, and avoids the need to use live rodents.[32] We found that in these cells the mouse OR-I7 gave a more consistent octanal-induced cAMP response than rat OR-I7. To minimize the possibility of any differential evaporation among the ligands, which vary in molecular weight, we also switched from the original luciferase reporter gene system to the GloSensor™ reporter, which detects the second messenger cAMP as it is produced in response to agonist-receptor binding.[33]

The response *vs.* time plot for different concentrations of octanal applied alone is shown in Fig. 2A. To be an antagonist, a compound must not activate the receptor, and using this reporter system we confirmed that compound **2** is not a mouse OR-I7 agonist (Fig. 2B). When compound **2** was co-applied at increasing concentrations with a constant concentration of octanal (5 μM), the response to octanal was reduced in a dose-dependent manner (Fig. 2C). This result showed that **2** antagonized the activation of mouse OR-I7 by octanal, as it did in the rat homolog expressed in neurons.[14] To obtain an OR-I7 response *vs.* concentration plot for estimating $IC_{50}$s, we summed the reporter response over the period of highest cAMP production, from 3 to 7.5 minutes following addition of the odorant(s) to the stimulation medium (Fig. 2C). This measurement allowed us to estimate the $EC_{50}$ of octanal (0.7 μM, confidence interval 0.35–1.4 μM) and then to estimate the $IC_{50}$ values for the designed aldehyde antagonists. We note that day-to-day variation in the Hana3A cells, plasmid transfection efficiency and the cAMP assay response prevented the calculation of absolute $IC_{50}$ values, but relative efficacies remained consistent during preliminary testing. For this

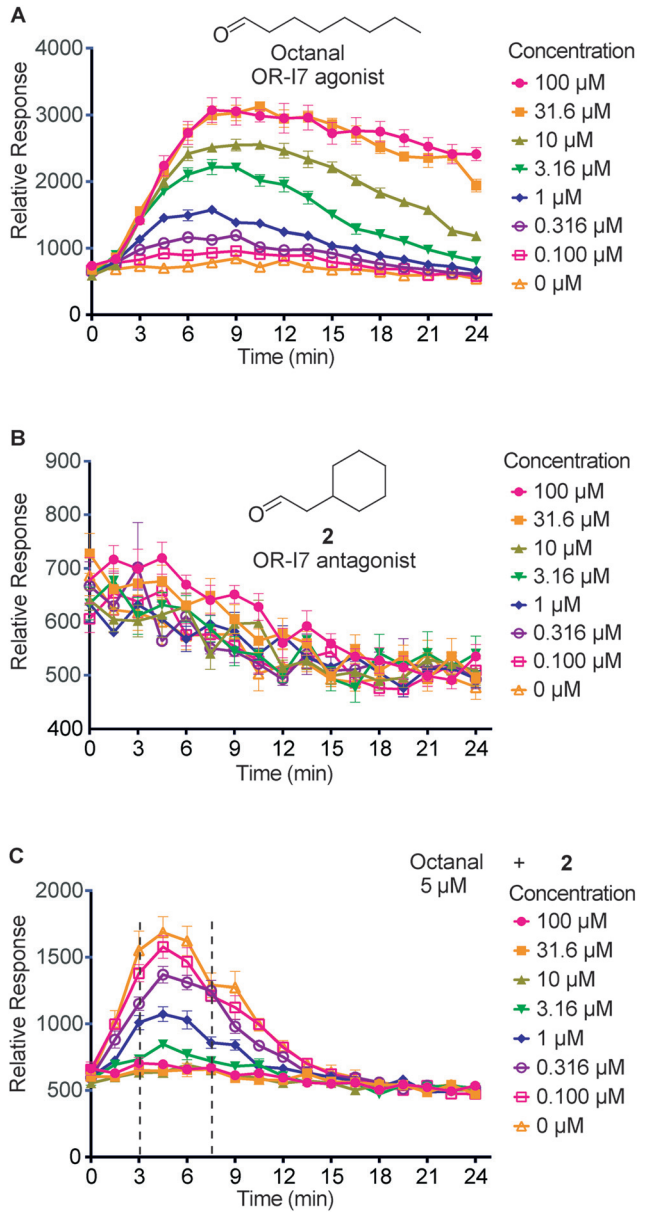
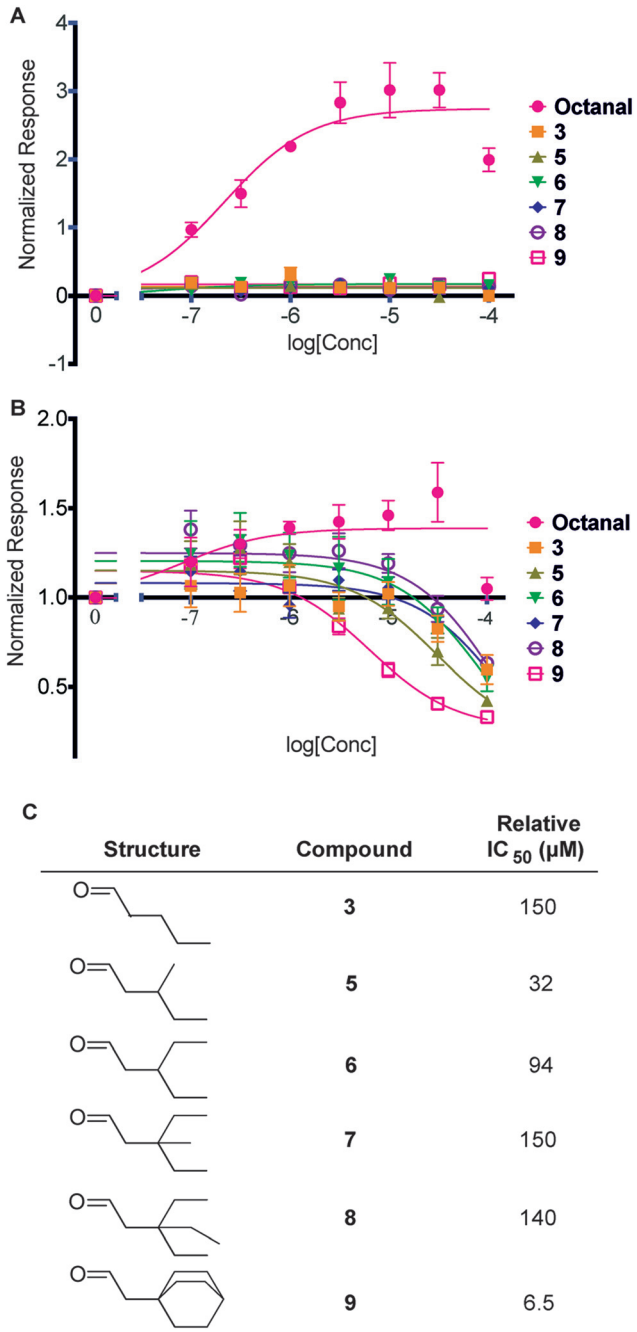



**Fig. 2** Relative mOR-I7 response elicited by octanal (an agonist) with or without antagonist **2**. Hana3A cells expressing the mouse OR-I7 receptor were exposed to increasing concentrations of octanal and/or compound **2** (cyclohexylethanal) and the rise in cAMP was monitored for 24 minutes using the GloSensor system. (A) Time course plot for increasing concentration of octanal. (B) Time course plot for increasing concentration of cyclohexylethanal, **2**. (C) Time course plot for octanal at 5 μM co-applied with increasing concentrations of **2**. The summed response between 3 min and 7.5 min (indicated by dash lines) was used to create a point for dose–response curves. Here and in Fig. 3 and 4, error bars indicate the average ± SEM of six replicates run in the same plate.

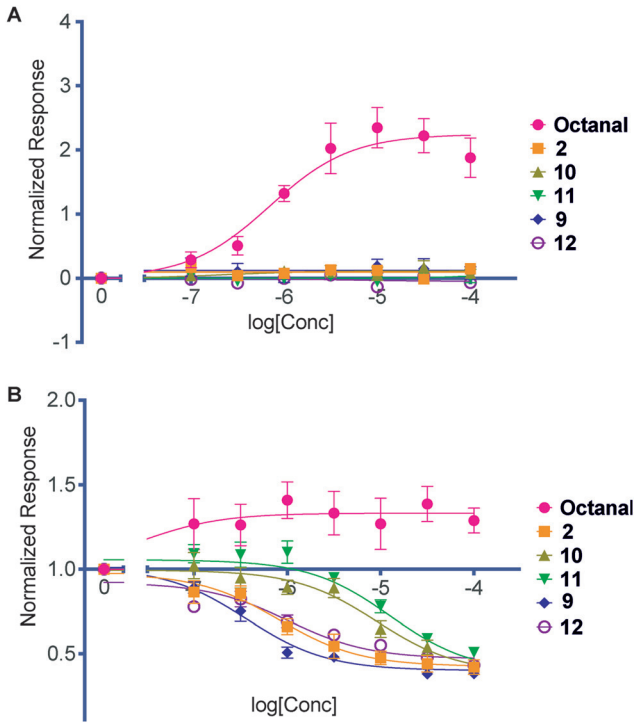
reason, we evaluated the candidate antagonists side-by-side with previously studied antagonists **2** and **3** to obtain the most accurate structure–activity comparison. Compounds grouped within Fig. 3 and 4 were tested with six replicates in the same experiment.




C

| Structure | Compound | Relative $IC_{50}$ (μM) |
|---|---|---|
|  | **3** | 150 |
|  | **5** | 32 |
|  | **6** | 94 |
|  | **7** | 150 |
|  | **8** | 140 |
|  | **9** | 6.5 |

**Fig. 3** Antagonist dose–response plots for aldehyde antagonists **3**, **5**, **6**, **7**, **8**, **9**. Activation dose–response curves for octanal, and octanal co-applied with each designed antagonist. The summed cAMP level for each concentration between 3 and 7.5 min provided one data point in the dose–response plots. (A) Compounds were applied individually to Hana3A cells expressing mouse OR-I7. The responses were compared to that of octanal, a known agonist of OR-I7. Octanal $EC_{50}$ in this experiment was estimated at 0.21 μM (confidence interval, 0.11–0.4 μM). (B) Inhibition dose–response curves. Each compound shown in panel C was tested for its ability to antagonize mouse OR-I7 in the presence of 5 μM co-applied octanal. The "octanal" curve indicates that additional octanal was added in place of antagonist (filled circles) to show whether agonist was saturating or close to saturating. (C) $IC_{50}$ values estimated from the dose–response curves.




C

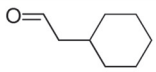
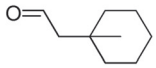
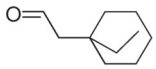
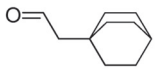
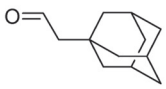
| Structure | Compound | Relative $IC_{50}$ (μM) |
|---|---|---|
|  | **2** | 0.82 |
|  | **10** | 9.1 |
|  | **11** | 13 |
|  | **9** | 0.37 |
|  | **12** | 1.0 |

**Fig. 4** Antagonist dose–response plots for aldehyde compounds **2**, **9**, **10**, **11**, and **12**. Activation dose–response curves for octanal and octanal co-applied with each designed antagonist. The summed cAMP level for each concentration between 3 and 7.5 min provided one data point in the dose–response plots. (A) Compounds were applied individually to Hana3A cells expressing mouse OR-I7. The responses were compared to that of octanal, a known agonist of OR-I7. Octanal $EC_{50}$ in this experiment was estimated at 0.7 μM (confidence interval 0.35–1.4 μM). (B) Inhibition dose–response curves. Each compound shown in panel C was tested for its ability to antagonize mouse OR-I7 in the presence of 5 μM co-applied octanal. The "octanal" curve indicates that additional octanal was added in place of antagonist (filled circles) to show whether agonist was saturating or close to saturating. (C) $IC_{50}$ values estimated from the dose–response curves.

Antagonists bind to their receptors without activating them. The first step in evaluating the new compounds was therefore to see whether any of the proposed antagonists, which are all shorter than 6.9 Å, activated mouse OR-I7. The results, summarized in Fig. 3A and 4A showed that none of these aldehydes activated Hana3A cells expressing the mouse OR-I7. These experiments provided further evidence that the aldehyde length *vs.* activation relationship previously found in the rat OR-I7 also held for the mouse OR-I7 ortholog.

Observing the response to 5 μM octanal in the presence of increasing concentrations of each inhibitor allowed us to estimate the $IC_{50}$ of each compound, subject to the limitations of the assay described above. Inhibition plots are shown for compounds **3**, and **5–9** (Fig. 3B), and for compounds **2** and **9–12** (Fig. 4B). The estimated $IC_{50}$ values are listed below the structures in the C panels of the same figures. Within each figure, all compounds were tested side-by-side on the same day on cells from the same Hana3A culture.

### Structure activity relationship of the antagonists

Pentanal, **3**, but not butanal, was previously shown to function as a weak rat OR-I7 antagonist,[14] while compound **2**, which is about the same length as pentanal but has eight carbons, was eight-fold more potent as an antagonist. One simplistic interpretation of this difference is that increasing the number of carbons on the aldehyde buries more hydrophobic surface area upon binding OR-I7 (*i.e.* a favorable hydrophobic effect contribution) and also increases the van der Waals contact with the part of the receptor binding site in contact with the ligand's mid-region. As shown in Fig. 3, we tested this possibility by adding carbons to pentanal, beginning at carbon-3 rather than carbon-2 so as not to risk interfering with aldehyde recognition.[11] Adding one carbon to make compound **5** increased potency 5-fold, but additional carbons progressively decreased potency (compounds **6**, **7**, **8**) until the three alkyl groups of compound **8** were tied back into the conformationally restricted bicyclo[2.2.2]octyl ring system (compound **9**), which was the most potent antagonist in this series. Thus, adding one carbon to carbon-3 was favorable but additional carbons were unfavorable unless they were conformationally restricted in the bicyclo[2.2.2]octyl ring system.

We used a similar approach beginning with antagonist **2**, and a similar trend was observed: the methyl and ethyl groups of **10** and **11**, respectively, were unfavorable, but tying compound **11**'s ethyl group back, as in compound **9**, was once again favorable in comparison (Fig. 4). We expanded the size of the bicyclo[2.2.2]octyl ring system – without adding methyl or ethyl groups and without exceeding 6.9 Å in length – by attaching the adamantyl group to ethanal (compound **12**). This aldehyde, which was noticed to have a distinct camphoraceous odor, was 3-fold weaker in potency than **9**, suggesting that while the OR-I7 binding site mid-region can accommodate this large ring system, it may be approaching the site's size limit.

In combination with previous reports on the rodent OR-I7,[11,14,20,22,31,34,35] our results are consistent with the view that the part of the OR-I7 binding site in contact with the ligand's mid-region has evolved to accommodate carbon chains with unbranched alkyl chains, *i.e.* a chain of methylene

groups, such as carbons-3 through -6 of octanal. This interpretation suggests that the small alkyl protrusions of compounds **6**, **7**, **8**, **10** and **11** may be interpreted by the receptor as alkyl chain branches, which are not found on the typical fatty acyl chain aldehyde odorants that OR-I7 is known to detect.[11] Thus, the middle of the OR-I7 site accommodates a chain of methylene groups passing through it, but anchored at one end by the aldehyde recognition site and, for activation only, at the other end by the small hydrophobic group binding site. The ability to accommodate the bicyclo[2.2.2]octyl system may reflect accommodation of a "jump rope" movement of the octanal's methylene groups between the anchors. An apparent exception to the unbranched chain preference was compound **5**, whose single 3-methyl group improved potency compared to pentanal. Compound **5** is identical to the first five carbons in the carbon chain of the terpene citronellal. Interestingly, citronellal, which has an aldehyde group and a length exceeding 6.9 Å, is almost as good a rat OR-I7 agonist as octanal.[11] We suggest that the OR-I7 features that allow citronellal binding in the mid-region also allow compound **5** to bind – but as an antagonist because its chain does not extend beyond the 6.9 Å threshold required for activation.[14] The idea that shortened forms of good OR-I7 agonists make good OR-I7 antagonists (compare **5** with citronellal, and **2** with **4**) raises the possibility that the binding site's mid-region is the same for agonists and antagonists alike, and the ligand does not change its location during activation. Alternatively, longer aldehydes may be able to move into a different location where they stabilize the active conformation of the receptor. Without structural information it is impossible to discern between these scenarios. Overall, these data support an interpretation where OR-I7 detects unbranched aliphatic aldehydes and citronellal-like terpene aldehydes, but when the chain is shorter than about 6.9 Å, binding fails to stabilize the activated form of the receptor and the ligand acts as an antagonist.

### Structural model of the mouse OR-I7 binding site

We previously built a rhodopsin-based homology model of the unactivated rat OR-I7 and predicted an orthosteric binding site by looking for voids large enough to accommodate octanal.[22] To understand the *in vitro* data presented above, we adapted this model to the mouse I7 ortholog. The ligand binding sites for aldehydes **2**, **8**, **9**, **10** and octanal were predicted using the SiteMap module of the Maestro software package.

The predicted ligand binding site, modeled using compound **10**, is shown in relation to the overall predicted receptor structure in Fig. 5A. The site is in the upper half of the receptor (closer to the extracellular side) and formed by TM3-TM6. The site is approximately the same binding site predicted for octanal in the rat I7 ortholog and that found experimentally for retinal in rhodopsin (see the ESI Fig. S5† for a comparison).[22] In Fig. 5A we have highlighted four residues of mouse OR-I7 homologous to residues predicted in the mouse MOR256-3 to mark the ligand binding site in that odorant receptor: F109$^{3.32}$, G113$^{3.36}$, A208$^{5.43}$ and Y257$^{6.48}$,

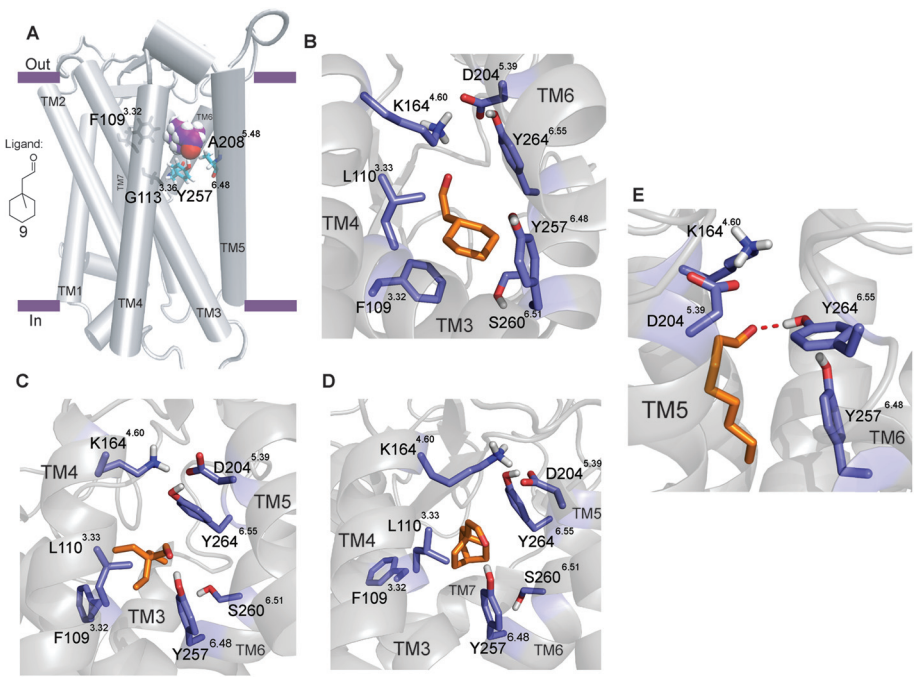



**Fig. 5** A rhodopsin-based mouse OR-I7 homology model docked with selected antagonists. Representative docking configurations for the mouse OR-I7 homology model and antagonist **10** (panel A), **2** (panel B), **8** (panel C), **9** (panel D) and octanal (panel E). In panel A, the global location of the predicted binding site is shown, with ligand presented as a space-filling model. The four numbered OR-I7 residues, 109, 113, 208 and 257, correspond to residues predicted in homology models for other odorant receptors to define the most likely orthosteric ligand-binding site, as noted in the text.

which correspond in MOR256-3 to F104$^{3.32}$, G108$^{3.36}$, G203$^{5.43}$ and Y252$^{6.48}$, respectively.[36] Though not coinciding exactly, this comparison predicts that the binding cavity is close to that of MOR256-3 and the binding sites predicted for several other odorant receptors for which models have been made.[8,37–39] In the continuing absence of any OR structural biology data, a consensus among binding site predictions is building increased confidence in their validity. Closer inspection of the mOR-I7 site reveals a binding cavity lined with hydrophobic amino acids, such as F109$^{3.32}$, L110$^{3.33}$, and the aromatic rings of Y257$^{6.48}$ and Y264$^{6.55}$ (Fig. 5B–E). Hydrogen-bonding interactions with the aldehyde were predicted for these two tyrosines and K164$^{4.60}$, a protonated amino acid residue that is also capable of forming a hydrogen bond with Y264$^{6.55}$ and a salt-bridge with the negatively charged D204$^{5.39}$. Based on this model, we speculate that the conformationally flexible ethyl groups found in relatively lower potency ligands like **8** (*e.g.* Fig. 5C) and **11** sterically interfere with some of the hydrophobic residues lining the site, *e.g.* L110$^{3.33}$, while the conformationally restricted ring systems of **2** (Fig. 5B) and **9** (Fig. 5D), being more compact and unbranched, are better accommodated by mOR-I7. For comparison, a representative view of octanal in the model's binding site is shown in Fig. 5E.

## Conclusions

The new aldehyde odorants studied here were designed to probe the carbon chain requirements for antagonizing the mouse OR-I7 receptor. The results show that the receptor prefers chains of methylene groups, disfavors branches except for a single methyl on carbon-3 and can accommodate a surprisingly large number of carbons (*e.g.* ten in adamantyl) as long as they are part of conformationally constrained ring system like cyclohexyl, bicyclo[2.2.2]octyl or adamantyl. Thus, in the context of antagonist ligands, the part of the receptor in contact with the mid-region imposes shape selectivity for compact carbon rings. In the context of an agonist, the ligand mid-region has to also serve to spatially orient the two end groups – the aldehyde and last two carbons of octanal, separated optimally by five carbons – as required for activation. A homology model predicts the location of the antagonist binding site, which is close to the ligand site predicted for several other ORs and rhodopsin.

## Abbreviations

| | |
|---|---|
| GPCR | G protein-coupled receptor |
| OR | Olfactory or odorant receptor |
| ORN | Odorant receptor neuron, aka OSN, olfactory sensory neuron |
| TM | Transmembrane |
| cAMP | Cyclic adenosine monophosphate |
| $IC_{50}$ | Half maximal inhibition constant |
| $EC_{50}$ | Half maximal binding constant |

## Conflicts of interest

The authors have no conflicts of interest to declare.

## Acknowledgements

This work was supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-13-1-0148 (to K. R.), NIH grants DC012095 and DC014423 (to H. M.) and NSF grant CHE-1465108 (to V. S. B.). Additional infrastructural support at the City College of New York was provided through grant 3G12MD007603-30S2 from the National Institute on Minority Health and Health Disparities. R. P. gratefully acknowledges support from a National Science Foundation REU grant (DBI-1560384). We thank Dr Lijia Yang for mass spectroscopy analysis, and NERSC for high-performance computing time. We thank an anonymous reviewer for bringing to our attention the THC/THCV analogy.

## References

1 L. Buck and R. Axel, *Cell*, 1991, **65**, 175–187.
2 G. Glusman, I. Yanai, I. Rubin and D. Lancet, *Genome Res.*, 2001, **11**, 685–702.
3 T. Olender, D. Lancet and D. W. Nebert, *Hum. Genomics*, 2008, **3**, 87–97.
4 V. Katritch, V. Cherezov and R. C. Stevens, *Annu. Rev. Pharmacol. Toxicol.*, 2013, **53**, 531–556.
5 B. K. Kobilka, *Biochim. Biophys. Acta*, 2007, **1768**, 794–807.
6 T. Abaffy, A. Malhotra and C. W. Luetje, *J. Biol. Chem.*, 2007, **282**, 1216–1224.
7 C. A. de March, Y. Yu, M. J. Ni, K. A. Adipietro, H. Matsunami, M. Ma and J. Golebiowski, *J. Am. Chem. Soc.*, 2015, **137**, 8611–8616.
8 L. Gelis, S. Wolf, H. Hatt, E. M. Neuhaus and K. Gerwert, *Angew. Chem., Int. Ed.*, 2012, **51**, 1274–1278.
9 O. Man, Y. Gilad and D. Lancet, *Protein Sci.*, 2004, **13**, 240–254.
10 K. Schmiedeberg, E. Shirokova, H. P. Weber, B. Schilling, W. Meyerhof and D. Krautwurst, *J. Struct. Biol.*, 2007, **159**, 400–412.
11 R. C. Araneda, A. D. Kini and S. Firestein, *Nat. Neurosci.*, 2000, **3**, 1248–1255.
12 X. Grosmaitre, S. H. Fuss, A. C. Lee, K. A. Adipietro, H. Matsunami, P. Mombaerts and M. Ma, *J. Neurosci.*, 2009, **29**, 14545–14552.
13 K. Nara, L. R. Saraiva, X. Ye and L. B. Buck, *J. Neurosci.*, 2011, **31**, 9179–9191.
14 Z. Peterlin, Y. Li, G. Sun, R. Shah, S. Firestein and K. Ryan, *Chem. Biol.*, 2008, **15**, 1317–1327.
15 B. Malnic, J. Hirono, T. Sato and L. B. Buck, *Cell*, 1999, **96**, 713–723.
16 E. Block, V. S. Batista, H. Matsunami, H. Zhuang and L. Ahmed, *Nat. Prod. Rep.*, 2017, **34**, 529–557.

17 E. Block, S. Jang, H. Matsunami, S. Sekharan, B. Dethier, M. Z. Ertem, S. Gundala, Y. Pan, S. Li, Z. Li, S. N. Lodge, M. Ozbil, H. Jiang, S. F. Penalba, V. S. Batista and H. Zhuang, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E2766–E2774.
18 T. S. McClintock, K. Adipietro, W. B. Titlow, P. Breheny, A. Walz, P. Mombaerts and H. Matsunami, *J. Neurosci.*, 2014, **34**, 15669–15678.
19 M. Shirasu, K. Yoshikawa, Y. Takai, A. Nakashima, H. Takeuchi, H. Sakano and K. Touhara, *Neuron*, 2014, **81**, 165–178.
20 R. C. Araneda, Z. Peterlin, X. Zhang, A. Chesler and S. Firestein, *J. Physiol.*, 2004, **555**, 743–756.
21 A. Thomas, L. A. Stevenson, K. N. Wease, M. R. Price, G. Baillie, R. A. Ross and R. G. Pertwee, *Br. J. Pharmacol.*, 2005, **146**, 917–926.
22 M. D. Kurland, M. B. Newcomer, Z. Peterlin, K. Ryan, S. Firestein and V. S. Batista, *Biochemistry*, 2010, **49**, 6302–6304.
23 J. A. Ballesteros and H. Weinstein, *Methods Neurosci.*, 1995, **25**, 366–428.
24 Y. Li, Z. Peterlin, J. Ho, T. Yarnitzky, M. T. Liu, M. Fichman, M. Y. Niv, H. Matsunami, S. Firestein and K. Ryan, *ACS Chem. Biol.*, 2014, **9**, 2563–2571.
25 G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, *J. Comput.-Aided Mol. Des.*, 2013, **27**, 221–234.
26 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177–6196.
27 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
28 H. Zhao and S. Firestein, *Science*, 1998, **279**, 237–242.
29 D. Krautwurst, K. W. Yau and R. R. Reed, *Cell*, 1998, **95**, 917–926.
30 T. Bozza, P. Feinstein, C. Zheng and P. Mombaerts, *J. Neurosci.*, 2002, **22**, 3033–3043.
31 Y. Li, Z. Peterlin, J. Ho, T. Yarnitzky, M. T. Liu, M. Fichman, M. Y. Niv, H. Matsunami, S. Firestein and K. Ryan, *ACS Chem. Biol.*, 2014, **9**, 2563–2571.
32 H. Saito, M. Kubota, R. W. Roberts, Q. Chi and H. Matsunami, *Cell*, 2004, **119**, 679–691.
33 F. Fan, B. F. Binkowski, B. L. Butler, P. F. Stecha, M. K. Lewis and K. V. Wood, *ACS Chem. Biol.*, 2008, **3**, 346–351.
34 M. S. Singer, *Chem. Senses*, 2000, **25**, 155–165.
35 R. S. Smith, Z. Peterlin and R. C. Araneda, *Methods Mol. Biol.*, 2013, **1003**, 203–209.
36 C. A. de March, Y. Yu, M. J. Ni, K. A. Adipietro, H. Matsunami, M. Ma and J. Golebiowski, *J. Am. Chem. Soc.*, 2015, **137**, 8611–8616.
37 O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel and H. Pick, *Biochemistry*, 2011, **50**, 843–853.
38 S. Katada, T. Hirokawa, Y. Oka, M. Suwa and K. Touhara, *J. Neurosci.*, 2005, **25**, 1806–1815.
39 S. Sekharan, M. Z. Ertem, H. Zhuang, E. Block, H. Matsunami, R. Zhang, J. N. Wei, Y. Pan and V. S. Batista, *Biophys. J.*, 2014, **107**, L5–L8.