Elucidating the structural basis for ligand binding and translocation in conserved insect odorant receptor co-receptors

Jody Pacalon, Guillaume Audic, Justine Magnat, Manon Philip, Jérôme Golebiowski, Christophe J Moreau, Jérémie Topin

Elucidating the structural basis for ligand binding and translocation in conserved insect odorant receptor co-receptors

Jody Pacalon[a,1], Guillaume Audic[b,1], Justine Magnat[b], Manon Philip[b], Jérôme Golebiowski[c], Christophe J. Moreau[b,*], Jérémie Topin[a,*]

[a] Université Côte d'Azur, Institut de Chimie de Nice UMR7272, CNRS, France

[b] Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble, France

[c] Department of Brain & Cognitive Sciences, DGIST, 333, Techno JungAng, Daero, HyeongPoong Myeon, Daegu, 711-873, Republic of Korea

* Jérémie Topin, 28 avenue Valrose, 06108 Nice France, +33 (0)4 89 15 01 32, jeremie.topin@univ-cotedazur.fr; Christophe Moreau, 71, avenue des Martyrs CS10090 F-38044 Grenoble cedex 9 France, +33 457 428 579, Christophe.moreau@ibs.fr.

Author Contributions: JT, JG and CJM. designed research; JP, GA, JM and MP performed research; JP, GA, JT and CJM analyzed data; and JT and CJM wrote the paper.

Competing Interest Statement: The authors declare no competing interest.

Classification: Biophysics and Computational Biology.

Keywords: insect olfaction, molecular modeling, electrophysiology, ion channels.

This PDF file includes:

Main Text
Figures 1 to 7
Supplementary Figures 1 to 10
Supplementary Tables 1 to 3

[1] These authors contributed equally to this work

Abstract

In numerous insects, the olfactory receptor family forms a unique class of heteromeric cation channels. Recent progress in resolving the odorant receptor structures offers unprecedented opportunities for deciphering their molecular mechanisms of ligand recognition. Unexpectedly, these structures in apo or ligand-bound states did not reveal the pathway taken by the ligands between the extracellular space and the deep internal cavities. By combining molecular modeling with electrophysiological recordings, we identified amino acids involved in the dynamic entry pathway and the binding of VUAA1 to *Drosophila melanogaster*'s odorant receptor co-receptor (Orco). Our results provide new evidence for the exact location of the agonist binding site and a detailed and original mechanism of ligand translocation controlled by a network of conserved residues. These findings would explain the particularly high selectivity of Orcos for their ligands.

Main Text

Introduction

Among all living multicellular organisms, insects represent more than half of all identified species on Earth, thus forming the most diverse group of animals (1). Insects show a remarkable capacity to adapt to a wide range of ecological niches. The rapid evolution of insect olfactory receptors is thought to contribute to this adaptation (2), endowing each insect species with the ability to selectively detect volatile chemicals associated with its specialized habitat and lifestyle. Therefore, olfaction is a vital sense necessary for them to find food, a mate, an oviposition site and a host. Moreover, the insect olfactory receptors are the main targets for the rational design of repulsive or attractive compounds for protection against vector-borne species or pest control (3, 4).

Ground breaking studies have provided a structural description of the proteins involved in odor recognition by insects (5, 6). In addition to the gustatory receptors, the repertoire of odorant receptors is mainly composed of two distinct families: i) the olfactory receptors (ORs) that form a complex with the highly conserved odorant receptor co-receptor (Orco) (7); and ii) the ionotropic receptors (IRs) that are structurally similar to the ionotropic glutamate receptor (8). The OR/Orco receptors are mainly expressed in olfactory sensory neurons (OSNs) found in insects' antennae. An individual OSN typically expresses only a single type of OR (9), which defines the neuron's response spectrum (10), even if non-canoncial co-expressions have been observed in a mosquito(11). The OR/Orco complexes are proposed to form a unique class of heteromeric cation channels composed of the two related 7-transmembrane subunits. It has been shown that Orcos could form homotetrameric channels (Fig. 1*A*), which have a different recognition spectrum than ORs (12, 13).

Orcos seems to appear late in the evolution of insects and constitute a remarkable example of an adaptive system, with a unique highly conserved signaling subunit (Orco) that can associate with a large repertoire of odorant receptor subunits that diverge to recognize specific ligands (14, 15). The evolution of ORs that led to the appearance of Orcos induced a total loss of odorant binding for this subunit, while engendering the ability to bind few synthetic ligands, like VUAA1 (16-22). On the other hand, the "ancestral" OR5 receptor from *Machilis hrabei* (MhraOR5) is activated by a large set of odorants, but not by VUAA1 (6). DmelOrco and MhraOR5 share 18.3% sequence identity and adopt the same tertiary fold (Fig. 1*B*). However, the origin of the differences in the recognition spectra of the two receptors is still not fully understood.

To decipher the molecular mechanisms governing the response of Orcos to ligands, different structure-function studies were previously employed based on site-directed mutagenesis combined with two-electrode voltage-clamp (TEVC) measurements. Fig. 1*C* summarizes the position of different residues that showed a functional impact when mutated (5, 23-26). These studies have highlighted the central role of residues from helix S7 in forming a hydrophobic gate that contributes to ion selectivity. Moreover, the structures of MhraOR5 in complex with two agonists, eugenol and DEET, revealed the ligand binding cavity of this receptor (Figs. 1*B* and *D*) (6).

Despite these highly informative structural studies, several questions remain, in particular the entry pathway and the binding site of ligands in Orcos. Their identification is essential for understanding the high specificity of action of Orco ligands and for the rational design of new molecules for

attractive or repulsive applications. To identify the binding pocket and the translocation pathway of VUAA1 from the extracellular space to the Orco binding site, we combined molecular modeling approaches with site-directed mutagenesis and functional characterization by the TEVC technique.

Results

Determination of the optimal Orco. Olfactory receptors are notorious for weakly expressing in heterologous systems, which impedes their functional characterization. Before initiating molecular dynamics (MD) simulations, we searched for the optimal Orco that generates the highest response to VUAA1 when expressed in Xenopus oocytes. Orcos from *Apocrypta bakery*, *Drosophila melanogaster*, *Aedes albopictus* and *Culex quinquefasciatus* were functionally characterized by the TEVC technique. The results (Fig. 2) clearly demonstrate that DmelOrco generated the highest current amplitude in the presence of VUAA1 and it was chosen as the model for both computations and experiments.

Orcos show a conserved cavity. A 3D model of DmelOrco was built by homology modeling using the experimental structure of AbakOrco homotetramer (pdb ID: 6C70) as a template (5). The two protein sequences are highly similar (76 % of sequence identity) prefiguring a high confidence in the accuracy of the model of DmelOrco (27). The full protocol is detailed in the Materials and Methods section. We compared it to a model obtained by AlphaFold2 (extracted from the Alpha Fold Protein database) (28). Both structures show a high similarity of transmembrane segments (RMSD = 0.7 Å). The largest deviation between the structures is observed at the intracellular loop 2 (IL2) (*SI Appendix*, Fig. S1). This loop is not resolved on the cryoEM structure of DmelOrco, which suggests a high flexibility.

The structure of AbakOrco (6) and the homology model of DmelOrco (Figs. 1*B* and *D*), revealed a cavity between helices S1 to S4 and S6 (Fig. 1*C*) that could play the role of the ligand binding site for VUAA1 and its analogues. Interestingly, this cavity has a position similar to the ligand binding site found in the structure of MhraOR5 (6) (Fig. 1*B*). The amino acids lining the two cavities are highly conserved between DmelOrco and AbakOrco with 73% identity (Fig.1*D*). Notably, the cradle of this pocket would be formed by the residue $F83^{Dmel}$, which is critical for activation by VUAA1 (26). In both structures and models, the access of VUAA1 to its putative binding site seems hindered by constrictions of the transmembrane helices, suggesting a progression of the molecule through a hidden and dynamic pathway.

MD simulations highlight a stepwise mechanism of VUAA1 entry to the embedded binding cavity.

We studied the entry of VUAA1 molecules to their putative binding site in DmelOrco by means of MD simulations. To reach this deeply embedded site, residing in the core of the transmembrane helices, the molecule must transit through a path that is assumed to be dynamic since it is closed in the structures of AbakOrco and MhraOR5. To identify this path, multiple MD simulations were performed with several ligands to enhance the sampling of rare events such as ligand migration (29, 30) and protein conformational changes (31, 32). We constructed a system containing 4 DmelOrco monomers with five VUAA1 molecules randomly placed in the extracellular part of the simulation box. Then, 22 replicas were subjected to MD simulations, leading to a total of 88 simulations on DmelOrco monomers. The total simulation time is approximately 31 µs (see Materials and Methods). A constraint was applied between each VUAA1 and the top of the channel pore to increase sampling speed without biasing the binding process. Thus, the ligands were free to sample the extracellular region of the simulation box and to diffuse into the receptor core. The migration of VUAA1 through the protein core was evaluated by the evolution of the distance between the VUAA1 center of mass and the center of mass of the binding cavity (defined as the center of mass of the eugenol molecule in MhraOR5, pdb: 7LID). We thus defined 4 distinct steps: contact (Fig. 3*A*, *area a*), entry (Fig. 3*A*, *area b*), vestibule (Fig. 3*A*, *area c*) and binding site (Fig. 3*A*, *area d*).

The results of our simulations revealed a predominant pathway of VUAA1 entry into the binding site. From the 88 trajectories, 19 showed an entry of VUAA1 within the receptor bundle (Fig. 3*A*,

area c). Out of these 19 trajectories, 7 full entries into the binding pocket (Fig. 3*A*, *area d*) were observed. The other trajectories resulted in a partial binding event, where VUAA1 remains within a vestibular site, half-way to the pocket cradle (*SI Appendix*, Figs. S2).

In all the seven observed binding events, VUAA1 consistently enters the receptor through the same gate and showed contact with residues belonging to helices S2 to S6. Interestingly, most of these residues are highly conserved among various Orcos (*SI Appendix*, Dataset S1 and Fig. S2*C*) in line with the similar action of VUAA1 observed in the majority of insect Orcos (26). The ligand does not interact with helix S7, which forms the tetrameric pore, suggesting that VUAA1 acts indirectly on the gate through conformational changes in Orco.

The migration of VUAA1 appears to be governed by stepwise hydrophobic and hydrophilic interactions throughout the ingress of the ligand towards the cradle of the binding site. The four steps have been identified based on the distance to the binding cavity and but also on the percentage of solvation of VUAA1 (Figs. 3*B* and *C*). The first step (a) is a rapid contact (few ns) VUAA1 with the extracellular side of DmelOrco and a rapid partial desolvation. The second step (b) is a stabilization of the position of VUAA1 during ~500 ns and a solvation stable at ~50%. The third step (c) is a rapid progress (less than 200 ns) of the molecule toward the cavity and a decrease of solvation up to ~20%. The fourth and last step (d) is a position of the molecule in the cavity with stable solvation around 20%. In steps (a) and (c), the desolvation of VUAA1 significantly increases, playing an essential role in the progression of the molecule toward the binding site.

The hydrophobicity and electrostatic complementarities of VUAA1 with DmelOrco in the different areas (*a* to *d*) have been evaluated in the *SI Appendix*, Table S1. Analogues of VUAA1 (VUAA0.5, 2, 3 and 4) were also incorporated and ranked by their EC_{50} (18). For all ligands considered, we noticed an increase in hydrophobic and electrostatic complementarity when the ligand was located deeper in the protein. Furthermore, the trend in hydrophobic complementarity approached that of ligand strength. Although the differences in ligand EC_{50} were minor, these observations suggest a correlation between hydrophobic complementarity in the ligand translocation pathway and ligand strength of VUAA analogues.

During its progression toward the binding site from the *area b* to *d* (Fig. 3*B*), VUAA1 is mostly orthogonal to the membrane plane (*area c*). In addition to the desolvation process, the flexibility of the molecule appears to greatly facilitate the migration of VUAA1. Thus, VUAA1 adopts several conformations to adapt to the local constraints, which allow the entrance into the protein either by its pyridine or its phenylethyl moiety. However, when considering an alternative orientation of the VUAA analogues in the binding site (*SI Appendix*, Table S1), we observe a decrease in electrostatic and hydrophobic complementarities compared to the initial orientation. This suggests that VUAA1 would have a preferred orientation in the binding site.

In the simulations, VUAA1 is stabilized by a subset of residues and must overcome an energetic barrier to reach the next metastable intermediate state. Several residues were identified as interacting with VUAA1 during its penetration into DmelOrco. A comprehensive list of these residues is provided as supplementary information (*SI Appendix*, Dataset S2). The initial binding event occurs through a contact between VUAA1 and $Y390^{S6}$ at the extracellular end of S6 (Fig. 3*B*, *area a*). Starting from this position, VUAA1 makes regular contacts with the residue side chains (Fig. 3*B*, *area b*) and undergoes a large desolvation process upon its entry into the receptor bundle (Fig. 3*B*, *area c*). The ligand then establishes additional contacts with $I79^{S2}$, $T80^{S2}$, $W150^{S3}$, $I181^{EL2}$, $V206^{S4}$, $K373^{S5}$ and $Y397^{S6}$, where it pauses for several nanoseconds (Fig. 3*B*, *area c*). The ligand finally enters the cavity (Fig. 3*B*, *area d*) that was previously identified in the structures of AbakOrco and MhraOR5, and in the model of DmelOrco (Fig. 1*D*). The final position of VUAA1 in the cavity is parallel to the membrane, and it interacts with $F83^{S2}$, $F84^{S2}$, $S146^{S3}$, $M210^{S4}$ and $Y400^{S6}$, similar to the position of the eugenol molecule in the MhraOR5 structure (*SI Appendix*, Fig. S3).

These results further guided site-directed mutagenesis experiments combined with functional assays to assess the critical role of residues identified as interacting with VUAA1 in the simulations.

Site directed mutagenesis and electrophysiological characterization support the entry mechanism of VUAA1.

To experimentally assess the functional role of residues that significantly interacted with VUAA1 in the simulations, different mutants were designed. The influence of the volume or the

physicochemical properties of their side chains was evaluated according to the response of Orco to stimulation by VUAA1. To facilitate or block the translocation process of VUAA1, the residues were mutated to smaller (alanine) or larger (tryptophan) residues, respectively. For disrupting hydrophobic interactions or hydrogen bonds between side chains and VUAA1, mutations were made in serine (small hydrophilic residue) or in phenylalanine (aromatic residue without hydroxyl group), respectively. To invert the charge at position 373^{S5}, the lysine was mutated in a negatively charged glutamate. The response to VUAA1 of each mutant was assessed by electrophysiological recordings with the TEVC method.

The simulations revealed that Y390^{S6} is the first residue that has a significant interaction with VUAA1, interacting at a frequency of 0.47 averaged over all entry trajectories. Y390^{S6} was mutated into alanine (Y390A) and phenylalanine (Y390F) and both mutations did not show significant change in the response to VUAA1 (Fig. 4). Thus, the reduction of the side chain into alanine or the removal of the hydroxyl group of Y390 did not favor or abolish the action of VUAA1. Consequently, neither aromaticity nor a hydroxyl group on the aromatic ring are necessary for the interaction with VUAA1 in position 390. On the contrary, its mutation into serine led to a decrease in the activation by VUAA1 (2.44 µA *vs* 4.71 µA for the WT). A Western-blot has been performed to verify that the expression level of the Y390S mutant was similar to the WT (*SI Appendix*, Fig. S4), and the semi-quantitative analysis indicated no significant differences between both constructs. This result supported the role of Y390 in VUAA1 activation. In particular, the differences observed between the mutants emphasize the importance of the hydrophilic character of position 390^{S6}. Introduction of a serine in place of tyrosine generates a hydrophilic environment (33) that would hamper the first step of desolvation process that is crucial for the entry of VUAA1, as observed in the simulations (Fig. 3*C*).

When going deeper in the protein, VUAA1 has shown high frequencies of interaction with a planar section of seven residues: I79^{S2}, T80^{S2}, W150^{S3}, I181^{EL2}, V206^{S4}, K373^{S5} and Y397^{S6} interacting with VUAA1 (Fig. 5*A*) at frequencies of 0.44, 0.56, 0.68, 0.17, 0.11, 0.70 and 0.70, respectively (averaged over all entry trajectories). Mutations into alanine of all seven residues did not significantly affect the amplitude of activation induced by VUAA1 (Fig. 5*B*), indicating that the side chains of these residues are not critical or involved in a limiting step for the interaction with VUAA1. In contrast, mutations of the non-aromatic residues in the bulkier tryptophan significantly reduced or abolished the activation by VUAA1 (Fig. 5*B-I, red dots*). Western-blot results (*SI Appendix*, Fig. S4) showed a decrease of expression of T80W and V206W. These results suggest that these mutations not only affected the expression level of the mutants but also the response to VUAA1. In contrast, mutations I181S and I181W showed an increase of expression in Western-blot results, but still a clear loss of VUAA1 activation supporting that the ability of VUAA1 to access this region is critical for the channel response.

As these residues are pointing into the ligand pathway observed during simulations, these functional results support the implication of these residues in the entry of VUAA1. Interestingly, inserting the hydrophilic and shorter serine residue in place of the hydrophobic I181^{EL2} (I181S), significantly reduced the amplitude of activation (1.40 *vs* 4.71 µA for the WT) (Fig. 5*B*) as previously observed with Y390S mutant. This deleterious effect of the mutation into serine is site specific since the similar mutation of Y397^{S6} (Y397S) showed no significant effect on VUAA1 response (Fig. 5*B*). Mutation of the only charged residue identified in the simulations (K373^{S5}) generated unexpected responses. Thus, mutation of K373^{S5} into alanine (K373A) that profoundly modifies the physico-chemical properties by reducing the size of the side chain and by removing the positive charge, did not change the response to VUAA1 (Fig. 5*B*). Inversion of the charge by mutation of K373^{S5} into glutamate (K373E) did not abolish the response but increased it (7.47 µA), potentially by decreasing the polarity of the binding cavity (*SI Appendix*, Fig. S5). Western-blot results confirmed that the K373E was not overexpressed. All mutations made at position Y397^{S6} did not significantly change the amplitude of activation induced by VUAA1 (Fig. 5). In the simulations, VUAA1 is in

transit in this section of seven residues, and move on to a deeper cavity, which would constitute the binding site.

Site directed mutagenesis and electrophysiological characterization support the binding site of VUAA1.

In the deeper section, five residues were identified in the simulations to frequently interact with VUAA1 and formed a cavity suspected to be the binding site (*SI Appendix*, Dataset S2). The five positions F83^{S2}, F84^{S2}, S146^{S3}, M210^{S4} and Y400^{S6} (respectively interacting with VUAA1 at a frequency of 0.32, 0.02, 0.23, 0.19, 0.24, averaged on all entry trajectories) were mutated to defined more precisely the cradle of the VUAA1 binding cavity (Fig. 6). Using the site-directed mutagenesis approach, all the five residues were mutated in alanine and tryptophan (Fig. 6*B*) to reduce or increase the steric hindrance of the side chains, respectively. The Western-blot results (*SI Appendix*, Fig. S4) demonstrated that all mutants of the Fig. 6 were expressed at similar levels.

In contrast to previous results, mutation in alanine of two phenylalanine residues (F83A and F84A) decreased the response to VUAA1 (Fig. 6*B-D*) with a greater extent for F84A (medians: 2.33, 0.68 µA for F83A, F84A respectively *vs* 4.71 µA for WT). Mutation in tryptophan induced the same phenotype in position 83 (F83W) (Fig. 6*B*), while the mutation in serine had the same impact in position 84 (F84S) (Fig. 6*B*). Finally, the mutation F84W did not induce a significant change compared to the WT (Fig. 6*B*). These results indicate that these two adjacent phenylalanine residues play a critical role in the activation by VUAA1, but with different characteristics. Position 83 must be a phenylalanine and cannot be replaced by a homologous residue like tryptophan, while position 84 is more tolerant to replacement by tryptophan but much less to alanine and serine. The peripheral position of F84^{S2} in the cavity could explain this selective tolerance to large hydrophobic residues, while the central position of F83^{S2} in the cavity suggests a larger and more specific interaction with the ligand. These results are in agreement with those of Corcoran et al. (26), showing that F83^{S2} is one of the essential residues for the action of VUAA1.

On the opposite side of the cavity, S146^{S3} is also pointing toward the binding cavity. Mutation of this hydrophilic residue induced a unique phenotype of increased response to VUAA1 when mutated in alanine (medians: 7.55 µA *vs* 4.71 µA for WT). This effect is strengthen by the introduction of the bulkier and more hydrophobic residue, valine (33) (median: 11.68 µA) (Fig. 6*B* and *E*). This mutation S146V showed the highest response to VUAA1 and could be used in further studies to increase the amplitude of the response.

Mutations of M210^{S4} in shorter alanine (M210A) or leucine (M210L) residues did not change the response to VUAA1 (5.72 and 6.19 µA, respectively *vs* 4.71 µA for WT), while the mutation in the bulkier tryptophan induced a significant decrease in the amplitude of activation (2.12 µA) (Fig. 6*B* and *F*). Consequently, the methionine 210 that is in close proximity to F83^{S2} and F84^{S2} does not specifically interact with VUAA1, but this position does not tolerate steric hindrance.

Mutation of Y400^{S6} in either alanine, serine or tryptophan did not significantly change the response to VUAA1. Despite the high conservation of Y400, this result is consistent with the position of the residue, located deeply in the core of the protein, so its mutation is unlikely to change the properties of the binding cavity.

Concentration-effect curves performed on mutants with the most significant results (Fig. 6*G* and *SI Appendix*, Table S2) showed a change in Imax that was either negative (for F84A and M210W, 1.01 and 1.50 µA, respectively *vs* 3.97 µA for WT) or positive (for S146V, 6.99 µA *vs* 3.97 µA for the WT), without affecting the EC_{50}. These results suggest a dominant effect of the mutations on the efficacy of VUAA1.

Western blot results show that mutants with a significant gain or loss of function are always expressed(*SI Appendix*, Fig. S4). Gain of function means an increase of current amplitude induced by VUAA1, while loss of function is a decrease or an absence of this current amplitude. Moreover, the level of expression has no correlation with the mutant response to VUAA1. Finally, we performed control experiments on a position, which do not interact with VUAA1 during the simulations, but close to the residue S146 that is particularly sensitive to gain- and loss-of-function mutations when mutated in alanine, valine and tryptophan. Leucine 141^{S2} was mutated to these

three types of amino-acid. Contrary to what is observed on the position S146^{S2}, these mutations did not induce significant change in the channel response to VUAA1 (*SI Appendix*, Fig. S6).

Discussion

The simulations of VUAA1 binding onto DmelOrco were carried out with little knowledge about the amino acids involved in the mechanism of binding. They indicated an entry pathway and a binding site in a model of Orco that is highly conserved across species. The binding mode of VUAA1 depends on interactions mediated by different hydrophobic, aromatic, and hydrophilic residues within the pocket. Finally, the experimental results supported the predictions made by the simulations and allowed the identification of residues critical for the entry of VUAA1 and for its binding in a deeply embedded cavity that is also found in the structures of AbakOrco and MhraOR5.

The translocation of VUAA1 through the protein is highly conserved among Orcos.

The analysis of the sequence conservation reveals that the pathway followed by VUAA1 to reach the binding site of DmelOrco is highly conserved (Fig. 7). As Orcos are known to recognize a remarkably restricted number of ligands, the high conservation of the translocation pathway can be interpreted as a molecular sieve, which filters the entrance of ligands to the binding site. These residues show a high conservation in Orcos and are likely to be crucial for initiating the opening of the channel upon ligand binding. In contrast, ORs that recognize a large diversity of ligands (34-36) show a low conservation at similar positions. The chemical variation observed in residues that line the translocation pathway in ORs allows a large diversity of odorants to diffuse inside the protein and reach their binding sites.

Desolvation of VUAA1 is fundamental for its entry.

The recent advances in structural biology have led to greater insight into the role of desolvation in the thermodynamics and kinetics of binding (37-39). The importance of hydrophobic interactions as a ligand-desolvation penalty or a driving force for the induced fit of receptors is a long-term challenge in drug design (40, 41). In particular, it has been shown that water plays a crucial role in the binding kinetics (42). The binding process of VUAA1 to Orco is accompanied by a desolvation at each metastable state. The most important decrease in the number of water molecules in the first solvation shell is observed when VUAA1 enters the protein. Accordingly, the mutation of the hydrophobic Y397^{S6} to a hydrophilic serine decreased the response of DmelOrco to VUAA1, most probably by preserving water molecules around VUAA1. Our results also suggest that I181^{EL2} could be involved in the desolvation process required for entry into the transmembrane core of Orco, which would explain why no continuous translocation pathway is observed in the structures of the apo state of AbakOrco and MhraOR5.

Comparative analysis of the eugenol-bound MhraOR5 structure (pdb: 7LID) with our VUAA1-bound DmelOrco model revealed a shared binding site position with a high conservation (16 amino acid pocket: 50 % identity, 62.5 % similarity; 24 amino acid pocket: 33 % identity, 62.5 % similarity) (Fig. 7). However, the ORs show a remarkable diversity in the binding site composition. This particularity is also found in mammalian ORs, allowing for a broad detection of chemicals (43-45).

The polarity and volume of the ligand binding cavity influences the efficacy of VUAA1.

The polarity of the binding cavity appears to have a pronounced influence on the channel response to VUAA1: a decrease induces a gain of function while an increase leads to a loss of function (*SI Appendix*, Fig. S5). We further investigate this observation by evaluating the polarity of 176 Orcos from 174 species. This analysis reveals that the binding cavity of the VUAA1-insensitive MdesOrco is more polar than the responsive Orcos. When Corcoran et al. (26) replaced the hydrophilic H81^{S2} from MdesOrco by a more hydrophobic phenylalanine (H81F), it induced a response to VUAA1. In contrast, mutations that increased the polarity of the binding cavity abolished the response to VUAA1 in AsegOrco. The polarity of the cavity seems to be a good indicator to predict the response to VUAA1 of a given Orco or mutant (*SI Appendix*, Fig. S5).

The volume of the cavity also influenced the response of Orco to VUAA1 (*SI Appendix*, Fig. S7). A substantial reduction of the volume (such as the introduction of a tryptophan residue, in position

F83^{S2}, S146^{S3}, V206^{S4} or M210^{S4}) induced a significant decrease in the response to VUAA1. In contrast, mutations that increased the volume of the cavity did not rationally induce a change in the response to VUAA1. An exception was the mutation F84A, which abolished the response to VUAA1, potentially due to an indirect interaction with the ligand. These results suggest that the protein could fluctuate to accommodate bulky ligands such as VUAA1, as has already been shown for olfactory receptors (6, 43, 44).

The architecture of the ligand binding site is conserved among Orcos and ORs.

Once in the binding cavity, VUAA1 is stabilized by a combination of hydrophobic h-bond, Van der Waals and pi-stacking interactions and does not move back into the bulk within the simulation time. We compared our electrophysiological results with already published data on mutants of two ORs: MhOR1 and MhraOR5 (*SI Appendix*, Table S3). In particular, our results highlight the importance of two residues from segment 2 (F83 and F84) to form the binding site. Mutations made at similar positions in MhOR1 (Y106^{S2}A, I107^{S2}A) and MhraOR5 (Y91^{S2}A and F92 ^{S2}A) result in non-responsive channels.

The natural complex is formed by the association of Orco subunits with ORx subunits. In this configuration, it has been suggested that VUAA1 acts as an allosteric ligand but it could also act as orthosteric ligand on the ORx subunit. The absence of response to VUAA1 of the ancestral MhraOR5, that recognizes a large diversity of odorants, suggest that VUAA1 would not act on the OR subunits. A possible approach would be to co-express mutants of DmelOrco in the translocation pathway or binding site that are insensitive to VUAA1 (such as I79W, T80W or V206W) with an ORx and challenge with VUAA1.

In the final pose, VUAA1 remained in the same orientation, with the ethyl phenyl moiety located between the helices S3 and S4 and the pyridine next to S2 and S5 (Fig. 7 and *SI Appendix*, Fig. S3). This conclusion is strengthened by the increased sensitivity of S146^{S3} mutants (alanine and valine). Interestingly, decreasing the ethyl moiety to a methyl almost abolished the response of Orco (18). In contrast, the replacement of the ethyl group by an isophenyl group improved the potency of the VUAA1-derivative. All together, these observations show that increasing hydrophobicity by mutations or ligand modifications increases the response of Orco to its ligands.

To explore the potential binding modes of VUAA1, the ligand was structurally modified to VUAA2, VUAA3 and VUAA4, which display significant greater potency, and in VUAA0.5, which is less potent than VUAA1. The calculated hydrophobic and electrostatic matches between the ligands and the receptor correlated with the functional results. We then evaluated the final orientation of VUAA1 in the binding site. Because of its general shape, the cavity could only accommodate VUAA1 in two directions, one of which is the opposite of the one observed in the simulations. Therefore, we manually flipped VUAA1 into the cavity (SI Appendix, Fig. S8) and constructed the VUAA analogs. Scores of both electrostatic and hydrophobic matches were inferior to those of the initial binding mode (*SI Appendix*, Table S1), suggesting that the initial orientation is preferred.

In conclusion, this study revealed the translocation pathway and binding site of VUAA1 into DmelOrco using a combination of dynamic simulations and functional characterization. The results highlight the role of desolvation for the progression of the ligand, the role of the polarity of the binding cavity in the efficacy of VUAA1 and the lower limit of size of the cavity for VUAA1 binding. This study shows that the binding pocket location is conserved between ORs and Orcos. The striking difference between the two families is the high level of sequence conservation of the translocation pathway and binding pocket observed in Orco compared to the high diversity in ORs. The conservation and the variability are then shared in between the two subunits forming the heterodimer. This combination of the highly conserved Orcos subunit with the more versatile ORs provides the insect with extremely high chemical discrimination power.

Orcos have been shown to play a fundamental role in insect behavior such as foraging and oviposition and are thus a potential target for the development of behaviorally disruptive chemicals (46, 47). Our results provide a fine description of the particular binding process, opening the way to a rational design of orthosteric and allosteric modulators.

Materials and Methods

In silico modelling

Alignment between Orcos and ORs with MhraOR5

Alignment between MhraOR5 and Orcos was based on the alignment files for 176 Orcos and 361 ORs from Butterwick *et al.* (5). MhraOR5 was realigned with the Orcos using ClustalO (48) with default settings, then optimized by hand to conserve the existing alignment. The same process was applied for the ORs.

Orco modelling

The 176 Orcos tetramer models plus *Dmel*Orco WT and mutants were generated by SWISS-model pipeline (49) using PDB 6C70 as a template with default settings. DmelOrco alpha fold model (version 07.01.2021) was retrieved from AlphaFold Protein structure database (50). RMSD between the SWISS-model and AlphaFold model was calculated using cpptraj (51) after alignment of the structures on *i)* all the sequence, *ii)* all the sequence except IL2, and *iii)* only helices.

Cavity analysis of DmelOrco, AbakOrco and MhraOR5

Detection of the pockets of the 176 Orcos plus DmelOrco mutants (SWISS-model), AbakOrco (pdb: 6C70) and MhraOR5 (in APO form, pdb: 7LIC) cavities was carried out using fpocket3 (52) with default settings. For each receptor, visual inspection was used to identify the pocket of interest.

Molecular dynamics setup

As IL2 is not resolved in the AbakOrco (pdb: 6C70) template structure, IL2 was discarded from the structure of each DmelOrco monomer. Propka (53) was used to predict protonation states of the protein at a target pH 6.5. The DmelOrco tetramer orientation in its membrane was determined using OPM server (54). Five VUAA1 molecules were added in different orientations on the extracellular side. The system was embedded into a POPC-only model membrane using PACKMOL-memgen (55). The simulation box was completed using TIP3P water molecules and neutralized using K^+ and Cl^- ions with a final concentration of 0.15 M. The total system is made up of 286736 atoms, in a $3.4.10^6$ Å^3 periodic box. Molecular dynamics simulations were performed with the sander and pmemd.cuda modules of AMBER18, with the ff14SB force field for the proteins and the lipid14 forcefield for the membrane (56). VUAA1 parameters were generated by calculating partial atomic charges with the HF/6-31G* basis set using Gaussian 09 (57). The obtained electrostatic potential was fitted by the RESP program (58). The other parameters were taken from the General Amber Force Field 2 (gaff2). Bonds involving hydrogen atoms were constrained using the SHAKE algorithm and long-range electrostatic interactions were handled using Particle Mesh Ewald. The cut-off for non-bonded interaction was set to 10 Å. Each system was first minimized with the AMBER sander module, with 5000 steps of steepest descent algorithm then 5000 steps of conjugate gradient with a 50 kcal·mol^{-1}·Å^2 harmonic potential restraint on the protein part of the system. A second minimization of the same length without restraint was applied. The systems were then thermalized from 100 to 310 K for 10000 steps (restraining the protein and ligands with a 200 kcal·mol^{-1}·Å^2 harmonic potential). Each system underwent 50000 steps of equilibration in the NPT ensemble and 1 bar (restraining the protein and ligands with a 15 kcal·mol^{-1}·Å^2 harmonic potential) before the production phase. During equilibration and production phase, temperature was kept constant in the system at 310 K using a Langevin thermostat with a collision frequency of 5 ps^{-1}. To increase sampling, all 5 VUAA1 molecules were constrained in a sphere of 45-55 Å radius, centered on the center of mass of the Lys486 of the four Orco monomers (with a potential of 10 kcal·mol^{-1}). To avoid VUAA1 aggregation, each VUAA1's sulfur atom was constrained to be a minimum of 20 Å from each other with a soft potential penalty of 5 kcal·mol^{-1}. The VUAA1 system in water only was built solvating the molecule in a 20 Å TIP3P periodic box using the gaff2 and tip3p forcefield parameters. The system was minimized with the AMBER sander module, with 500 steps of steepest descent algorithm then 500 steps of conjugate gradient, then heated incrementally from 100 to 310K for 10000 steps. The first 10 nanoseconds of the production phase were considered as equilibration and not taken into account for analysis. The system stability was evaluated from the root mean square deviation (RMSD) evolution computed on the backbone of the full system. During the 22 replicas, the receptors underwent small fluctuations (RMSD < 3Å)

showing that they remained correctly folded during microsecond simulations (*SI Appendix*, Fig. S9). Hydration of VUAA1 was calculated using the pytraj watershell function.

Minimum distance between VUAA1 and eugenol for all trajectories

The minimum distance between VUAA1 and eugenol was calculated for all trajectories by structurally aligning MhraOR5 (pdb: 7LID) on each DmelOrco monomer using the cealign pymol command (59), then calculating the center of mass distance between eugenol and VUAA1 on each trajectory using the mindist pytraj module (51).

Selection of representative frames for b, c, and d

Representative frames of the diffusion were obtained by dividing the prototypical trajectory into 4 parts according to the curve shown in Fig. 3*C*. For each part, a frequency analysis between VUAA1 and the receptor using the get_contacts module (https://getcontacts.github.io/) identified the critical residues. These residues, plus VUAA1, were selected and used to cluster each part by kmeans clustering, using cpptraj (51) with a fixed number of 4 clusters. The representative frame of the largest cluster was then extracted as the representative frame of that part of the trajectory.

Electrostatic and hydrophobic complementarity

For each representative frame (b, c and d), the protein was extracted with VUAA1 which was then replaced with VUAA0.5, VUAA2, VUAA3 and VUAA4 (18). For the representative frame of the binding site (d), VUAA1 was also manually flipped over using the pair fitting tool in PyMol, and then replaced again with VUAA0.5, VUAA2, VUAA3 and VUAA4. Each system was then minimized using the AMBER sander module, with 5000 steps of steepest descent algorithm then 5000 steps of conjugate gradient, while restraining the backbone of the protein with a 50 $kcal \cdot mol^{-1}$ potential. Hydrophobic complementarity scores for each system were calculated using the PLATINUM web server (60) with default settings. Electrostatic complementarity scores for each system were calculated using the Flare electrostatic complementarity tool (61).

Chemicals

VUAA1 (N-(4-ethylphenyl)-2-((4-ethyl-5-(3-pyridinyl)-4H-1, 2, 4-triazol-3-yl)thio)acetamide) (CAS 525582-84-7) was purchased from Sigma-Aldrich. The stock solution was 110 mM in DMSO and subsequently diluted into appropriate buffer solution.

Molecular biology

All Orco gene sequences were optimized (62) for protein expression in Xenopus laevis oocytes with the GenSmartTM Codon optimization Tool and subcloned into a pGEMHE-derived vector. The wildtype gene of Drosophila melanogaster Orco (DmelOrco) was synthetized by Genscript and subcloned with XmaI/XhoI cloning sites. Site-directed mutagenesis of DmelOrco was done by PCR with the Q5® site directed mutagenesis kit (NEB) using primers optimized with the NEBase Changer online tool and following the supplier's protocol. After transformation of commercial competent bacteria (XL10 Gold) by standard heat-shock protocol and overnight culture in ampicillin-containing LB plates, positive clones were identified by electrophoretic restriction profile and external sequencing (Genewiz). DNAs of a positive clones were amplified with Qiagen MidiPrep Kit and the ORF fully sequenced. For in vitro transcription, DNAs were linearized with restriction enzyme NotI that cuts a unique site in the 3' region of the polyA tail. The linearized DNAs were purified by the standard phenol:chloroform extraction method and transcribed into mRNA using the T7 ultra mMessage mMachine kit (Thermo Fisher Scientific). mRNAs were purified with the NucleoSpin RNA plus XS kit (Machery-Nagel). DNA and RNA were analyzed by agarose-gel electrophoresis and quantified by spectrophotometry.

Electrophysiological recordings

Xenopus laevis oocytes were prepared as previously described (63). Briefly, oocytes were defolliculated after surgical retrieval by type 1A collagenase over 2-3h under smooth horizontal agitation. They were manually selected and incubated at 19°C in modified-Barth's solution (1 mM KCl, 0.82 mM $MgSO_4$, 88 mM NaCl, 2.4 mM NaHCO3, 0.41 mM CaCl2, Ca(NO3)2 0.3 mM, 16 mM

HEPES, pH 7.4) supplemented with 100 $U \cdot mL^{-1}$ of penicillin and streptomycin and 0.1 $mg \cdot mL^{-1}$ of gentamycin. Each oocyte was micro-injected with the Nanoject instrument (Drummond) with 50 nL of 20 ng of mRNA coding for the Orco of interest. Injected oocytes were incubated individually in 96-well plates for 4 days at 19°C in the same buffer. Different batches of oocytes have been tested per construct. The results of the mutants of interest have been confirmed by a second set of experiments comparing the amplitudes with WT in the same day and from the same batch of oocytes (SI appendix Fig. S10).

Whole cell currents were recorded with the two-electrode voltage-clamp (TEVC) technique with the HiClamp robot (MultiChannel System). The results of the mutants of interest have been confirmed by a second set of experiments comparing the amplitudes with WT in the same day and from the same batch of oocytes. Microelectrodes were filled with 3M KCl. The high K^+ buffer used for recordings was composed of 91 mM KCl, 1 mM $MgCl_2$, 1.8 mM $CaCl_2$, 5 mM HEPES, pH 7.4. Membrane voltage was clamped to -50 mV and VUAA1 was applied for 60 s. Data were extracted with M. Vivaudou's programs (64) and statistically analyzed with Prism 8 (Graphpad).

Animal handling and experiments fully conformed to European regulations and were approved by the French Ministry of Higher Education and Research (APAFIS#30915-2021040615209331 v1 to CM). The animal facility was authorized by the Prefect of Isere (Authorization #E 38 185 10 001).

Western Blots

All expression experiments were assessed on 4-20% mini-Protean TGX SDS-PAGE gels (Bio-Rad). All oocytes loaded on gel were from the same batch and injected as described above, with 4 days of incubation.

Oocytes were homogenized through several passes in a syringe with two sizes of needles (18g then 27g) into a solubilization buffer (PBS 1X, protease inhibitor cocktail tablets) and stored at -80°C. Western blots were performed by transferring proteins onto a nitrocellulose membrane using the trans-blot turbo system (BioRad). Membranes were blocked with PBS 1x-1% non-fat milk overnight at 4°C and incubated in primary antibody anti-Orco (1:500) and the secondary antibody Goat anti-rabbit IgG HRP conjugate (1:5000 Sigma-Aldrich) for 1 hour each. The immunoblot was revealed with ECL substrate kit (Abcam) and recorded on ChemiDoc (BioRad) at different times for identifying the optimal exposition time without pixel saturation. Gels were stained with standard Coomassie blue staining protocol and the pictures were taken with the Chemidoc apparatus. Relative intensities of bands in blots and volume of lanes in gels were determined with the Image Lab software (BioRad).

The polyclonal primary antibody anti-Orco was purchased from Genscript and designed against the peptide sequence SSIPVEIPRLPIKSFYPW in the second extracellular loop (ECL2). Anti-Orco was produced in rabbit and purified by antigen affinity.

Acknowledgments

This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 682286) to C. J. M. This work was supported by the Fondation Roudnitska under the aegis of Fondation de France to J.P. This work was also supported by the CNRS MITI programs: "Modélisation du vivant" and "Biomimétisme". This work was supported by the French government, through the UCAJEDI Investments in the Future project managed by the National Research Agency (ANR) under reference number ANR-15-IDEX-01. The authors are grateful to the OPAL infrastructure and the Université Côte d'Azur's Center for High-Performance Computing for providing resources and support.

Authors thank Hervé Pointu, Soumala-maya Bama Toupet and Charlène Caloud for the management and the maintenance of Xenopus and acknowledge the platform supported by GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS(ANR-17-EURE-0003). They thank Michel Vivaudou for the development of software for data analysis (64). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA).

References

1. C. Mora, D. P. Tittensor, S. Adl, A. G. Simpson, B. Worm, How many species are there on Earth and in the ocean? *PLOS Biol.* **9** (2011).
2. B. S. Hansson, M. C. Stensmyr, Evolution of insect olfaction. *Neuron* **72**, 698-711 (2011).
3. J. Bohbot *et al.*, Multiple activities of insect repellents on odorant receptors in mosquitoes. *Med. Vet. Entomol.* **25**, 436-444 (2011).
4. A. Ray, Reception of odors and repellents in mosquitoes. *Curr. Opin. Neurobiol.* **34**, 158-164 (2015).
5. J. A. Butterwick *et al.*, Cryo-EM structure of the insect olfactory receptor Orco. *Nature* **560**, 447-452 (2018).
6. J. Del Marmol, M. A. Yedlin, V. Ruta, The structural basis of odorant recognition in insect olfactory receptors. *Nature* **597**, 126-131 (2021).
7. K. Sato *et al.*, Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002-1006 (2008).
8. D. Wicher, F. Miazzi, Functional properties of insect olfactory receptors: ionotropic receptors and odorant receptors. *Cell Tissue Res.*, 1-13 (2021).
9. L. B. Vosshall, A. M. Wong, R. Axel, An olfactory sensory map in the fly brain. *Cell* **102**, 147-159 (2000).
10. A. A. Dobritsa, W. v. d. G. van Naters, C. G. Warr, R. A. Steinbrecht, J. R. Carlson, Integrating the molecular and cellular basis of odor coding in the Drosophila antenna. *Neuron* **37**, 827-841 (2003).
11. M. Herre *et al.*, Non-canonical odor coding in the mosquito. *Cell* **185**, 3104-3123. e3128 (2022).
12. R. J. Pitts, C. Liu, X. Zhou, J. C. Malpartida, L. J. Zwiebel, Odorant receptor-mediated sperm activation in disease vector mosquitoes. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2566-2571 (2014).
13. P. L. Jones, G. M. Pask, D. C. Rinker, L. J. Zwiebel, Functional agonism of insect odorant receptor ion channels. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8821-8825 (2011).
14. H. M. Robertson, C. G. Warr, J. R. Carlson, Molecular evolution of the insect chemoreceptor gene superfamily in Drosophila melanogaster. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 14537-14542 (2003).
15. P. Brand *et al.*, The origin of the odorant receptor gene family in insects. *elife* **7**, e38340 (2018).
16. S. Chen, C. W. Luetje, Identification of new agonists and antagonists of the insect odorant receptor co-receptor subunit. *PLOS ONE* **7**, e36784 (2012).
17. P. L. Jones *et al.*, Allosteric antagonism of insect odorant receptor ion channels. *PLOS ONE* **7**, e30304 (2012).
18. R. W. Taylor *et al.*, Structure–activity relationship of a broad-spectrum insect odorant receptor agonist. *ACS Chem. Biol.* **7**, 1647-1652 (2012).
19. I. M. Romaine *et al.*, Narrow SAR in odorant sensing Orco receptor agonists. *Bioorg. Med. Chem. Lett.* **24**, 2613-2616 (2014).
20. S. Chen, C. W. Luetje, Phenylthiophenecarboxamide antagonists of the olfactory receptor co-receptor subunit from a mosquito. *PLOS ONE* **8**, e84575 (2013).

21. G. M. Pask, Y. V. Bobkov, E. A. Corey, B. W. Ache, L. J. Zwiebel, Blockade of insect odorant receptor currents by amiloride derivatives. *Chem. Senses* **38**, 221-229 (2013).
22. Y. Bobkov, E. Corey, B. Ache, An inhibitor of Na+/Ca2+ exchange blocks activation of insect olfactory receptors. *Biochem. Biophys. Res. Commun.* **450**, 1104-1109 (2014).
23. T. Nakagawa, M. Pellegrino, K. Sato, L. B. Vosshall, K. Touhara, Amino acid residues contributing to function of the heteromeric insect olfactory receptor complex. *PLOS ONE* **7**, e32372 (2012).
24. B. N. Kumar *et al.*, A conserved aspartic acid is important for agonist (VUAA1) and odorant/tuning receptor-dependent activation of the insect odorant co-receptor (Orco). *PLOS ONE* **8**, e70218 (2013).
25. R. M. Turner *et al.*, Mutational analysis of cysteine residues of the insect odorant co-receptor (Orco) from Drosophila melanogaster reveals differential effects on agonist- and odorant-tuning receptor-dependent activation. *J. Biol. Chem.* **289**, 31837-31845 (2014).
26. J. A. Corcoran, Y. Sonntag, M. N. Andersson, U. Johanson, C. Löfstedt, Endogenous insensitivity to the Orco agonist VUAA1 reveals novel olfactory receptor complex properties in the specialist fly Mayetiola destructor. *Sci. Rep.* **8**, 1-13 (2018).
27. M. Jaiteh, I. Rodríguez-Espigares, J. Selent, J. Carlsson, Performance of virtual screening against GPCR homology models: Impact of template selection and treatment of binding site plasticity. *PLoS Comput. Biol.* **16**, e1007680 (2020).
28. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
29. R. O. Dror *et al.*, Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13118-13123 (2011).
30. J. Topin, J. Diharce, S. b. Fiorucci, S. Antonczak, J. r. m. Golebiowski, O2 migration rates in [NiFe] hydrogenases. A joint approach combining free-energy calculations and kinetic modeling. *J. Phys. Chem. B* **118**, 676-681 (2014).
31. R. O. Dror *et al.*, Identification of two distinct inactive conformations of the β2-adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4689-4694 (2009).
32. C. A. de March *et al.*, Odorant receptor 7D4 activation dynamics. *Angew. Chem. Int. Ed.* **130**, 4644-4648 (2018).
33. W. Hoffmann *et al.*, An intrinsic hydrophobicity scale for amino acids and its application to fluorinated compounds. *Angew. Chem. Int. Ed.* **58**, 8216-8220 (2019).
34. G. Wang, A. F. Carey, J. R. Carlson, L. J. Zwiebel, Molecular basis of odor coding in the malaria vector mosquito Anopheles gambiae. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4418-4423 (2010).
35. S. A. Kreher, D. Mathew, J. Kim, J. R. Carlson, Translation of sensory input into behavioral output via an olfactory system. *Neuron* **59**, 110-124 (2008).
36. S. M. Boyle, S. McInally, A. Ray, Expanding the olfactory code by in silico decoding of odor-receptor chemical space. *elife* **2**, e01120 (2013).
37. J. Schiebel *et al.*, Intriguing role of water in protein-ligand binding studied by neutron crystallography on trypsin complexes. *Nat. Com.* **9**, 1-15 (2018).
38. L. Gajdos *et al.*, Neutron crystallography reveals mechanisms used by Pseudomonas aeruginosa for host-cell binding. *Nat. Com.* **13**, 1-9 (2022).
39. J. Mondal, R. A. Friesner, B. Berne, Role of desolvation in thermodynamics and kinetics of ligand binding to a kinase. *J. Chem. Theory Comput.* **10**, 5696-5705 (2014).

40. A. M. Davis, S. J. Teague, Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Ed.* **38**, 736-749 (1999).
41. R. F. de Freitas, M. Schapira, A systematic analysis of atomic protein–ligand interactions in the PDB. *Medchemcomm* **8**, 1970-1981 (2017).
42. P. Setny, R. Baron, P. M. Kekenes-Huskey, J. A. McCammon, J. Dzubiella, Solvent fluctuations in hydrophobic cavity–ligand binding kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1197-1202 (2013).
43. L. Charlier *et al.*, How broadly tuned olfactory receptors equally recognize their agonists. Human OR1G1 as a test case. *Cell. Mol. Life Sci.* **69**, 4205-4213 (2012).
44. J. Topin *et al.*, Discrimination between Olfactory Receptor Agonists and Non-agonists. *Chem. Eur. J.* **20**, 10227-10230 (2014).
45. X. Zhang, S. Firestein, The olfactory receptor gene superfamily of the mouse. *Nat. Neurosci.* **5**, 124-133 (2002).
46. H. Yan *et al.*, An engineered orco mutation produces aberrant social behavior and defective neural development in ants. *Cell* **170**, 736-747. e739 (2017).
47. R. A. Fandino *et al.*, Mutagenesis of odorant coreceptor Orco fully disrupts foraging but not oviposition behaviors in the hawkmoth Manduca sexta. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15677-15685 (2019).
48. F. Sievers, D. G. Higgins, Clustal omega. *Curr. Protoc. Bioinform.* **48**, 3.13. 11-13.13. 16 (2014).
49. A. Waterhouse *et al.*, SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296-W303 (2018).
50. M. Varadi *et al.*, AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439-D444 (2022).
51. D. R. Roe, T. E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084-3095 (2013).
52. V. Le Guilloux, P. Schmidtke, P. Tuffery, Fpocket: an open source platform for ligand pocket detection. *BMC Bioinform.* **10**, 1-11 (2009).
53. C. R. Søndergaard, M. H. Olsson, M. Rostkowski, J. H. Jensen, Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K a values. *J. Chem. Theory Comput.* **7**, 2284-2295 (2011).
54. M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg, A. L. Lomize, OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370-D376 (2012).
55. S. Schott-Verdugo, H. Gohlke, PACKMOL-memgen: a simple-to-use, generalized workflow for membrane-protein–lipid-bilayer system building. *J. Chem. Inf. Model.* **59**, 2522-2528 (2019).
56. I. Y. B.-S. D.A. Case, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, *et al.*, AMBER 2018, University of California, San Francisco. (2018).
57. M. Frisch *et al.*, Gaussian 09, Revision D. 01, Gaussian, Inc., Wallingford CT. *See also: URL: http://www. gaussian. com* (2009).
58. J. Wang, P. Cieplak, P. A. Kollman, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049-1074 (2000).
59. I. N. Shindyalov, P. E. Bourne, Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **11**, 739-747 (1998).

60. T. V. Pyrkov, A. O. Chugunov, N. A. Krylov, D. E. Nolde, R. G. Efremov, PLATINUM: a web tool for analysis of hydrophobic/hydrophilic organization of biomolecular complexes. *Bioinformatics* **25**, 1201-1202 (2009).
61. M. R. Bauer, M. D. Mackey, Electrostatic complementarity as a fast and effective tool to optimize binding and selectivity of protein–ligand complexes. *J. Med. Chem.* **62**, 3036-3050 (2019).
62. R. E. Roberts, J. K. Yuvaraj, M. N. Andersson, Codon optimization of insect odorant receptor genes may increase their stable expression for functional characterization in HEK293 cells. *Front. Cell. Neurosci.* **15** (2021).
63. C. J. Moreau, J. P. Dupuis, J. Revilloud, K. Arumugam, M. Vivaudou, Coupling ion channels to receptors for biomolecule sensing. *Nat. Nanotechnol.* **3**, 620-625 (2008).
64. M. Vivaudou, eeFit: a Microsoft Excel-embedded program for interactive analysis and fitting of experimental dose–response data. *Biotechniques* **66**, 186-193 (2019).

Figures and Tables

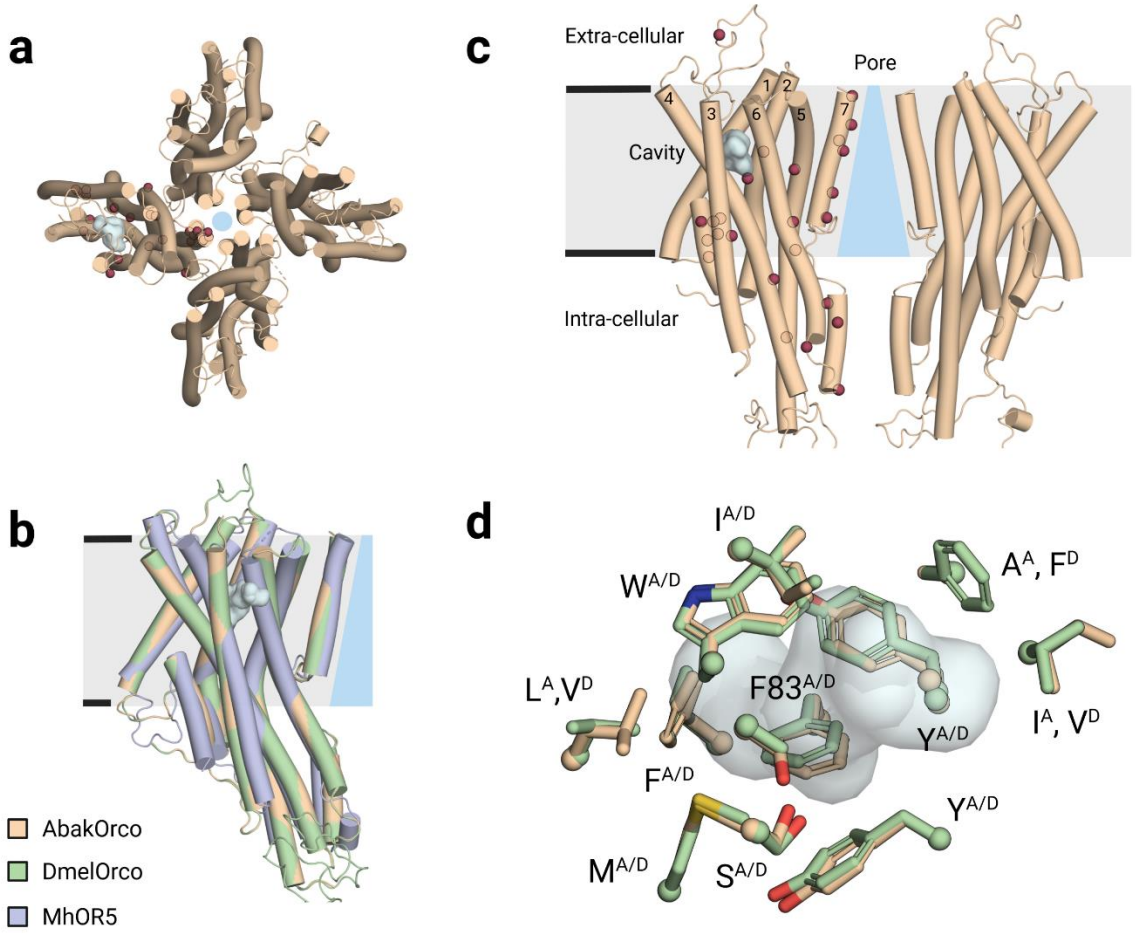


Fig. 1 | **Architecture of the DmelOrco homotetramer model and comparative representations of cavities and mutants. (a)** Extracellular view of the cryo-electronic microscopy structure of the homotetramer of *Apocrypta bakeri* Orco (AbakOrco) (pdb: 6C70). A ligand-binding pocket (in cyan) of related Orco receptors shown on the left subunit. The central pore is symbolized by a blue circle. **(b)** AbakOrco (pdb: 6C70, beige) membrane view, superposed on *Machilis hrabei* odorant receptor 5 (MhraOR5, pdb: 7LID, blue) and *Drosophila melanogaster* Orco (DmelOrco, green) homology model. Cavity analysis reveals the conserved position of a pocket (cyan) in these 3 structures. **(c)** Side view of two Orco subunits with a diagram of the channel pore (blue trapezoid). Residues shown in red spheres are equivalent to residues critical for VUAA1 response found in Orcos from point mutations that alter channel function in *Drosophila melanogaster*, *Agrotis segetum*, *Mayetiola destructor*, *Bombyx mori*, or *Apocrypta bakeri*. **(d)** Close view of the cavities (cyan) of DmelOrco and AbakOrco with their amino acids represented as sticks (respectively green and beige). DmelOrco and AbakOrco cavities share 73% of sequence identity (82% of similarity). The superscript letters A and D refer to AbakOrco and DmelOrco, respectively.

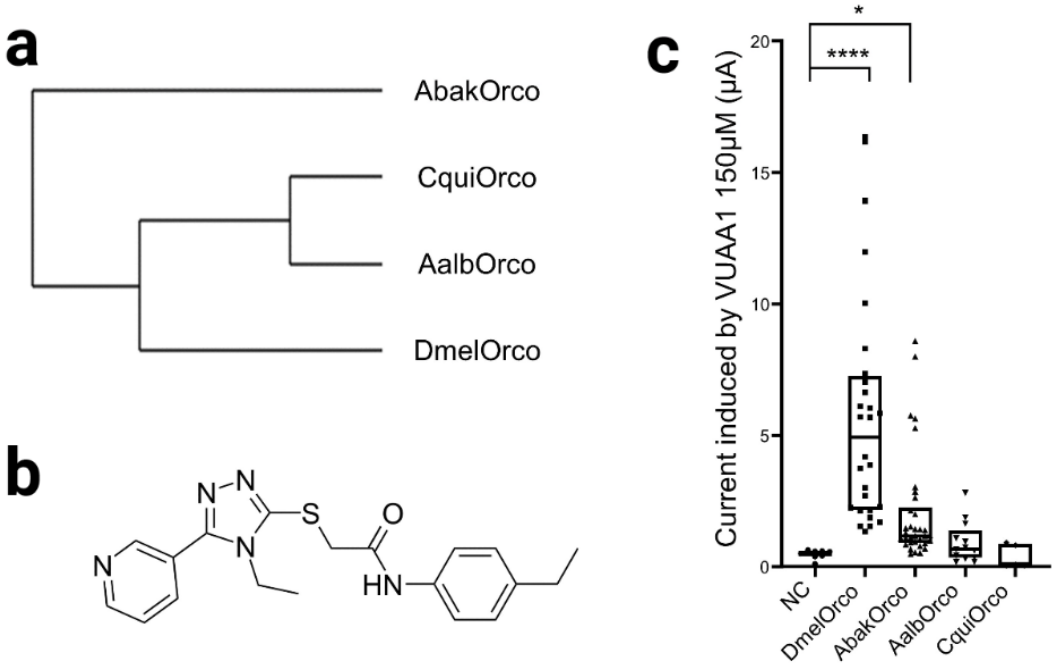


Fig. 2 | **Determination of the Orco species generating the highest response to VUAA1 in TEVC recordings. (a)** Phylogenetic tree of Orcos from different species. Abak: *Apocrypta bakeri*; Cqui: *Culex quinquefasciatus,* Aalb: *Aedes albopictus*, Dmel: *Drosophila melanogaster*. **(b)** Structure of VUAA1. **(c)** Current amplitudes induced by 500 µM VUAA1 on Orcos from different species expressed in Xenopus oocytes and recorded by TEVC. Results are median +/- SEM. P values are <0.0001 (Dmel, ****), 0.0438 (Abak, *) and 0.92 (Aalb), with NC as reference in the Kruskal-Wallis test. n= 8 Negative control, n= 28 DmelOrco, n= 34 AbakOrco, n=13 AalbOrco, n= 5 CquiOrco; NC: negative control (water-injected oocytes).

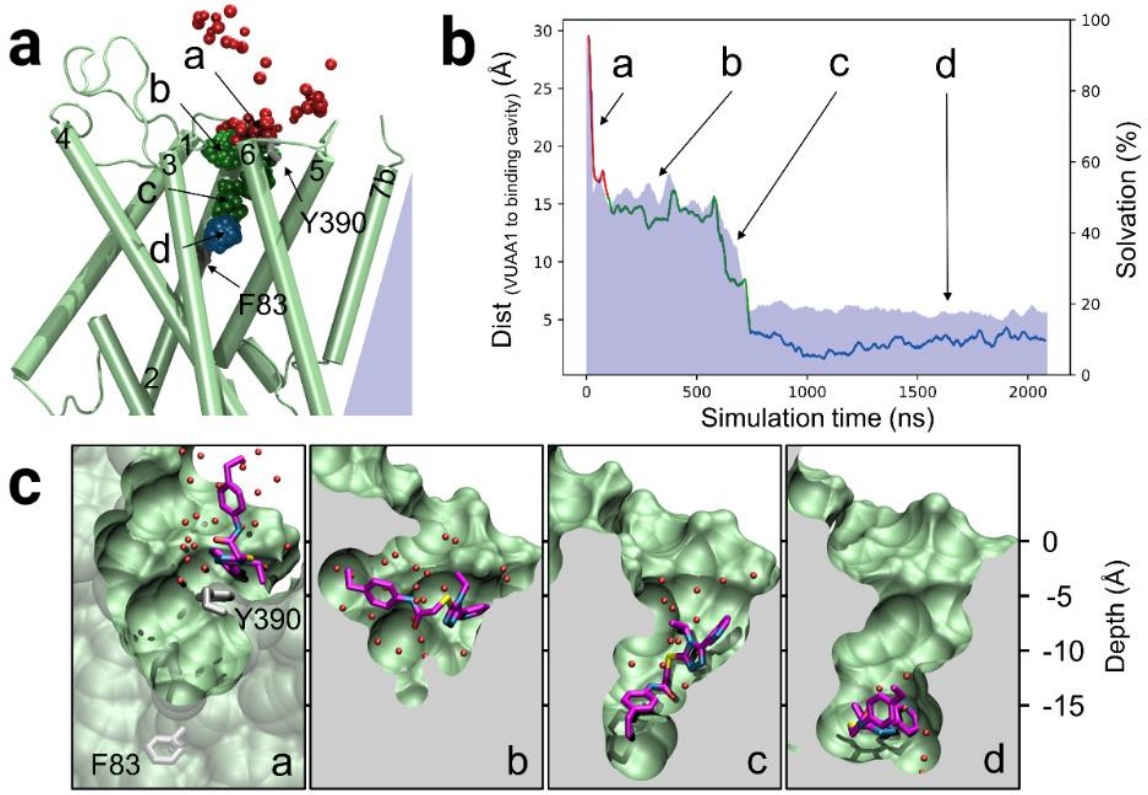


Fig. 3 | **Stepwise translocation of VUAA1 into DmelOrco model in MD simulations. (a)** Prototypical trajectory of VUAA1 binding event. The Orco monomer is shown in green. F83 and Y390 labels give their localization. The VUAA1 center of mass is represented by beads colored from red to blue according to the simulation time. a, b, c and d are sequential of VUAA1 progressing toward the binding site, from contact to Orco, entry, vestibule and binding to the internal cavity, respectively. **(b)** Close view of the progression of VUAA1 inside the Orco monomer corresponding to the positions a, b, c and d. Carbon atoms of VUAA1 are colored purple and the sulfur atom in yellow. Carbon atoms of $F83^{S2}$ and $Y390^{S6}$ are in grey and water molecules found less than 3 Å away from VUAA1 are represented by red spheres. **(c)** Evolution of the distance between the VUAA1 centers of mass and the center of mass of the binding cavity (defined as the center of mass of the eugenol molecule in MhraOR5, pdb: 7LID). The red curve represents the positions outside of the receptor to the contact (a). The green part of the curve represents the entry and vestibule events (b and c) and the blue one the sample of the binding cavity (d). The blue area shows the percentage of ligand solvation during the binding process (normalized to the solvation of the ligand outside the protein).

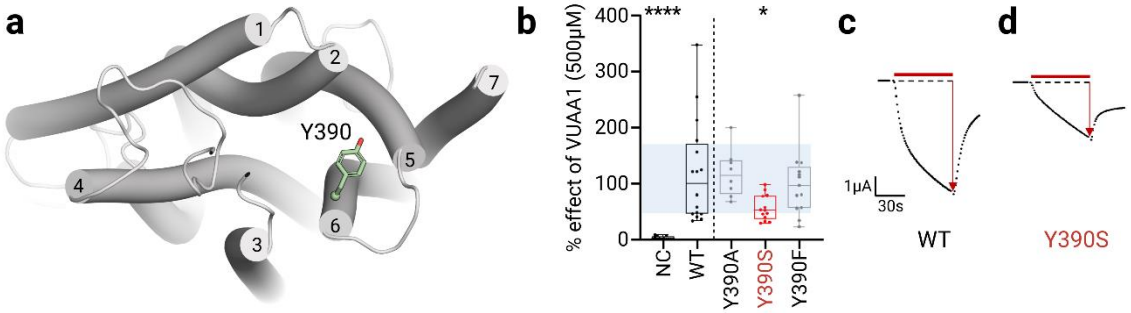


Fig. 4 | **Electrophysiological characterization of Orco mutants in the contact region**. **(a)** Top view of the entry highlighting the position of Y390. **(b)** Boxplot showing the current induced by 500µM of the ligand VUAA1 and measured by TEVC recordings on DmelOrco WT and mutants expressed by the Xenopus oocyte. n= 8 minimum recordings from different oocytes; NC: Negative Control (non-injected oocytes). Data are analysed with one-way ANOVA with α-error= 0.05 followed by Dunn's post-hoc test, with WT used as reference for the multiple comparison test. Results from mutant showing a statistical decrease from WT are coloured in red. **(c-d)** Representative current measured on WT and mutant with statistical difference from DmelOrco WT.

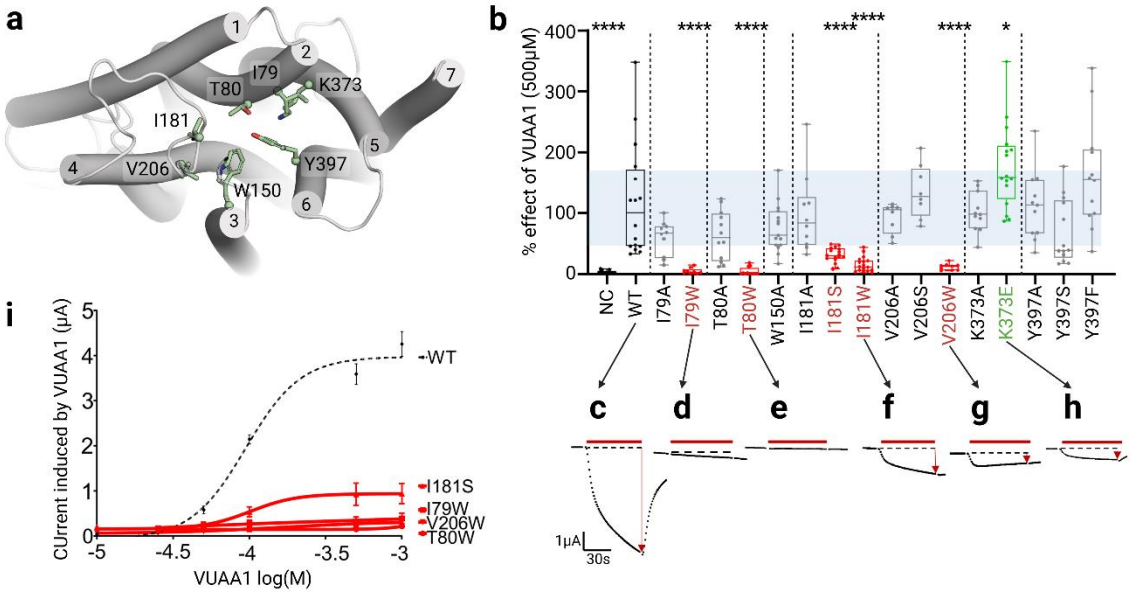


Fig. 5 | **Electrophysiological characterization of Orco mutants in the entry and vestibule region. (a)** Close view of the vestibule region highlighting the position of the different mutation sites. **(b)** Boxplot showing the current induced by 500µM of the ligand VUAA1 and measured by TEVC recordings on DmelOrco WT and mutants expressed in Xenopus oocytes. n= 8 minimum recordings from different oocytes; NC: Negative Control. Data are analysed with one-way ANOVA with α-error= 0.05 followed by Dunn's post-hoc test, with WT used as reference for multiple comparisons test. Results of mutants with statistically significant decrease or increase compared to WT are coloured in red or green, respectively. **(c-h)** Representative current measured on mutants with statistical differences from DmelOrco WT. **(i)** Dose-response curves for the mutants considered. EC_{50} are 94.5 µM for WT, 101.9 µM for I181S, not determined for the others. n≥ 8 recordings from different oocytes.

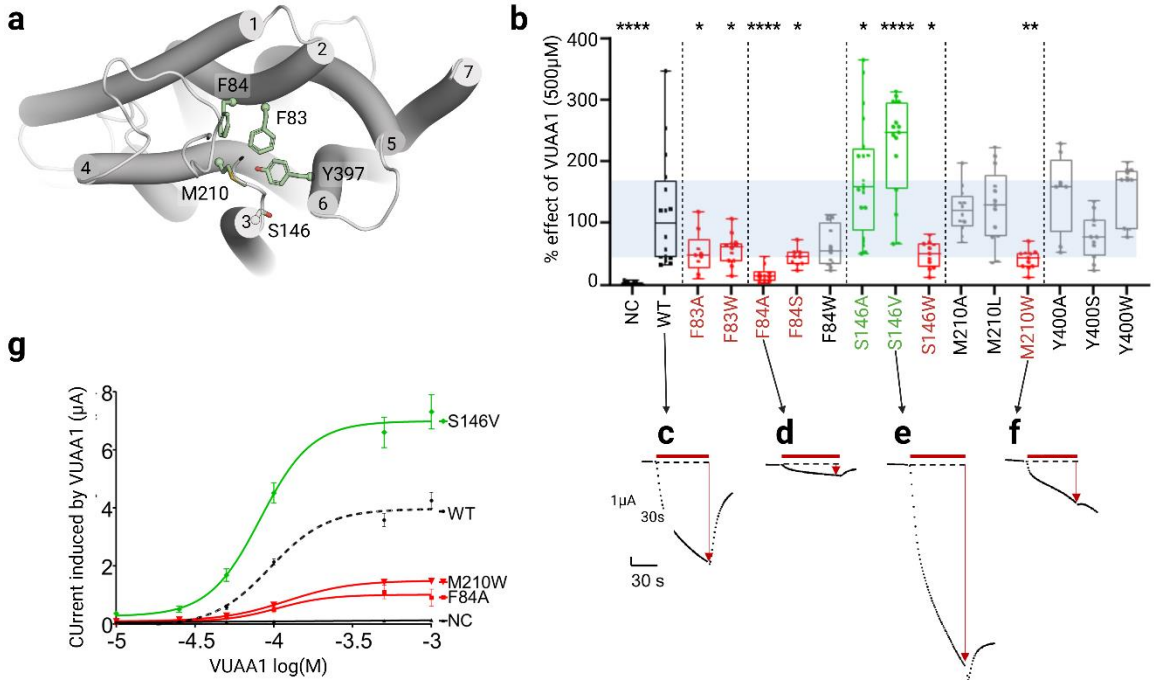


Fig. 6 | **Electrophysiological characterization of Orco mutants in the binding cavity. (a)** Close view of the binding region highlighting the position of the different mutation sites. **(b)** Boxplot showing the current induced by 500µM of the ligand VUAA1 and measured by TEVC recordings on DmelOrco WT and mutants expressed by Xenopus oocyte. n= 8 minimum recordings from different oocytes; NC: Negative Control. Data are analysed with one-way ANOVA with α-error= 0.05 followed by Dunn's post-hoc test, with WT used as reference for multiple comparisons test. Results from mutant showing a statistical decrease or increase from WT are coloured in red or green respectively. **(c-f)** Representative current measured on mutants with statistical differences from DmelOrco WT. **(g)** Dose-response curves for the mutants considered. EC_{50} are 94.5 µM for WT, 82.2 µM for S146V, 120.2 µM for M210W, 105.0 µM for F84A and not determined for NC. n≥ 7 recordings from different oocytes.

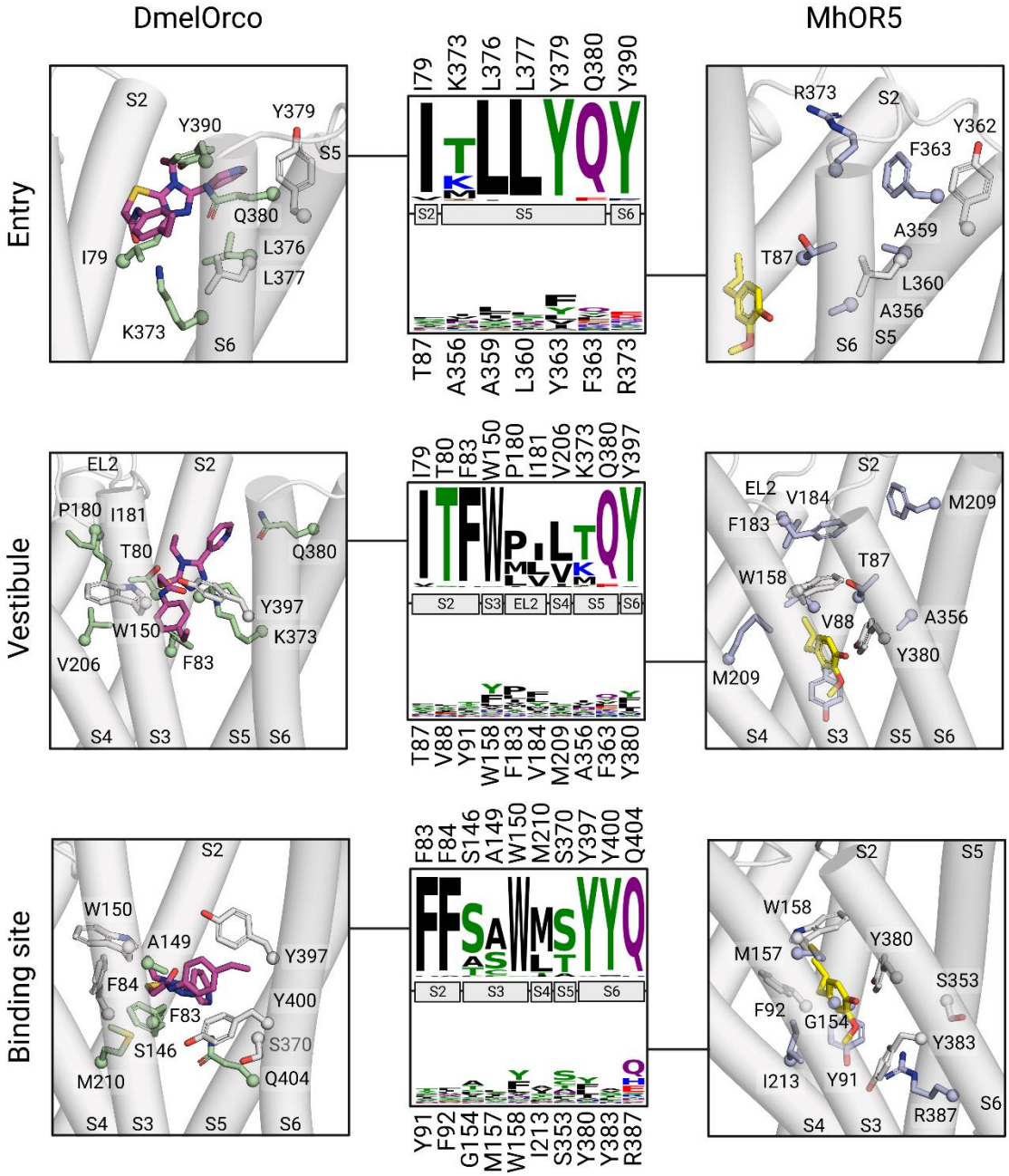


Fig. 7 | **Structural details and sequence conservation of the entry, vestibule and binding site in Orcos and ORs**. The structures for DmelOrco are extracted from MD simulations, while the structure for MhraOR5 is taken from pdb: 7LID. The amino acids were selected according to their frequencies of interaction with VUAA1 during MD simulations. Carbon atoms from amino acids conserved between Orcos and MhraOR5 are coloured in white. Carbon atoms from VUAA1 and Eugenol are shown in purple and yellow, respectively. Residue conservation among 176 Orcos from 174 species and 361 ORs from 4 species are coloured according to their side-chain chemistry. In each frame, the upper consensus logo account for Orcos and the lower one for ORs.

Supplementary Information for

Ligand of the conserved insect odorant receptor co-receptor reaches its binding site through a dynamic translocation pathway

Jody Pacalon[a,2], Guillaume Audic[b,1], Justine Magnat[b], Manon Philip[b], Jérôme Golebiowski[c], Christophe J. Moreau[b,*], Jérémie Topin[a,*]

[a] Université Côte d'Azur, Institut de Chimie de Nice UMR7272, CNRS, France

[b] Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble, France

[c] Department of Brain & Cognitive Sciences, DGIST, 333, Techno JungAng, Daero, HyeongPoong Myeon, Daegu, 711-873, Republic of Korea

* Jérémie Topin, 28 avenue Valrose, 06108 Nice France, +33 (0)4 89 15 01 32, jeremie.topin@univ-cotedazur.fr; Christophe Moreau, 71, avenue des Martyrs CS10090 F-38044 Grenoble cedex 9 France, +33 457 428 579, Christophe.moreau@ibs.fr.

This PDF file includes:

Figures S1 to S8

Tables S1 to S3

Legends for Datasets S1 to S2

SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S2

[2] These authors contributed equally to this work.

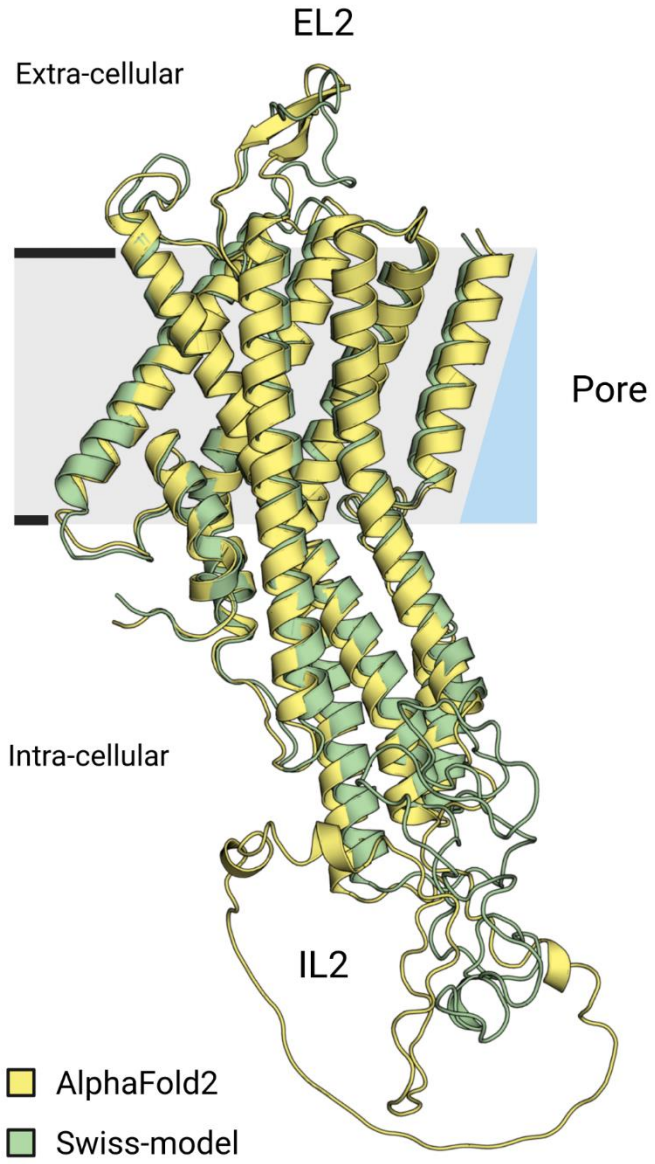


Figure S1. Superposition of the AlphaFold2 DmelOrco model with the AbakOrco structure. Superposition of the experimental structure of AbakOrco homotetramer (pdb ID: 6C70) superposed on the AlphaFold2 model retrieved from the Alpha Fold Protein database (in yellow). The two structures have an RMSD of 0.7 Å calculated on their transmembrane segments. To orient the protomer, part of the pore is shown on right as blue right-angled triangle. EL2: extracellular loop 2. IL2: intracellular loop 2.

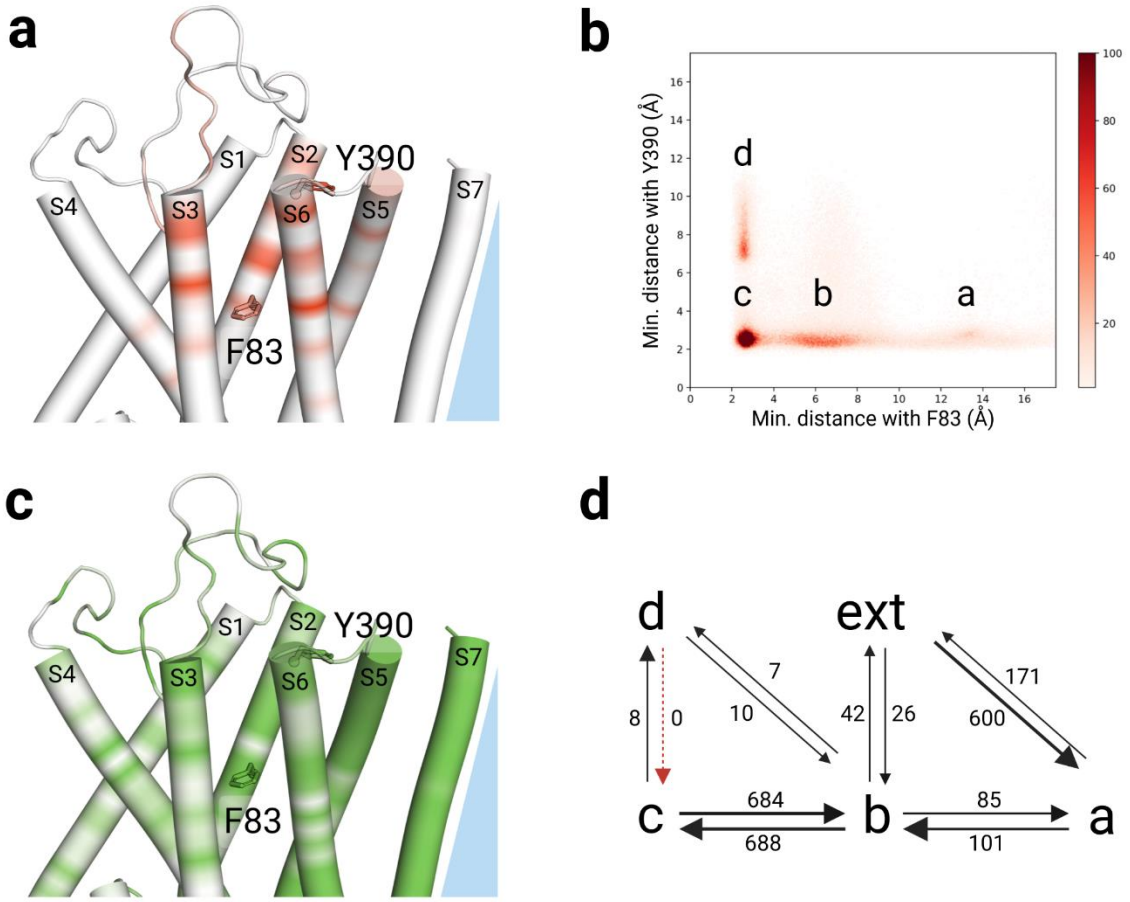


Figure S2. Summary of the interactions observed during the MD simulations between VUAA1 and DmelOrco. (a) Contact frequencies of VUAA1 mapped onto the structure of DmelOrco. The color gradient account for high frequency interaction (in red) to no interaction (in white). **(b)** Contour map of VUAA1 migration as the minimum distance from $F83^{S2}$ (distance from the entry) and minimum distance from $Y390^{S6}$ (distance from the cradle of the cavity). The four basin allow identifying four states in the VUAA1 binding: contact a, entry b, vestibule c and binding d. **(c)** Residue conservation among 176 Orcos from 174 species mapped onto the structure of DmelOrco using a colour scale from green high conservation, to white low conservation. **(d)** Proposed mechanism of VUAA1 binding. The numbers indicate the transition from one state to another.

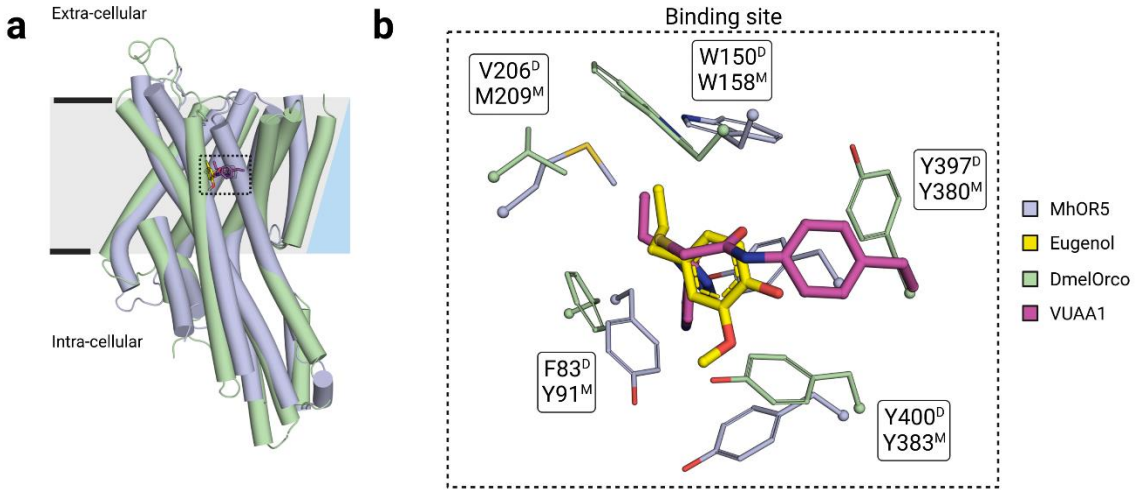
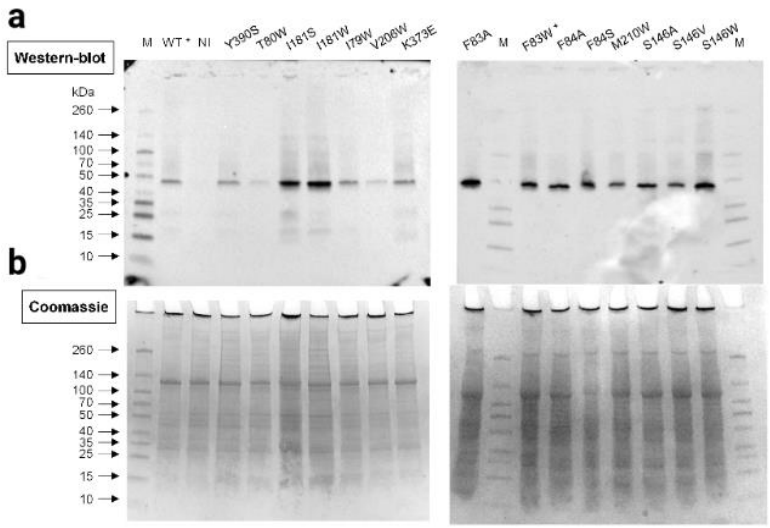


Figure S3. Comparison of the binding site location predicted in DmelOrco and observed in MhraOR5 structure. (a) Superposition of DmelOrco (green tubes) and MhraOR5 (blue tubes), with their respective ligands VUAA1 (magenta sticks) and eugenol (yellow sticks). **(b)** Close view on the binding site of DmelOrco (green licorice) and MhraOR5 (blue licorice). The superscripts D or M on the amino acids names represent DmelOrco and MhraOR5, respectively.



c

Relative intensities

	Western-blot	Coomassie	WB/Coomassie
WT *	1.0	1.0	100%
NI		1.3	
Y390S	1.0	1.0	109%
T80W	0.3	1.5	18%
I181S	5.3	2.5	211%
I181W	6.6	2.2	296%
I79W	1.2	1.5	78%
V206W	0.3	1.1	30%
K373E	1.2	1.4	85%
F83A	1.5	1.0	158%
F83W *	1.0	1.0	100%
F84A	1.0	0.8	131%
F84S	0.8	0.7	118%
M210W	0.6	0.8	72%
S146A	0.9	0.8	108%
S146V	0.9	0.8	108%
S146W	1.3	1.0	134%

Figure S4. Western-blot analysis of DmelOrco WT and mutant expression in Xenopus oocytes. **(a)** Western blots with a polyclonal primary antibody directed against DmelOrco. The constructs indicated above the lanes were expressed in Xenopus oocytes and the samples are crude membrane extracts. M: ThermoScientific Spectra Multicolor Broad Range Protein Ladder; WT*: wild-type used as reference for determining the relative intensities of bands of the first blot and gel, while the mutant F83W* was used for the second blot and gel. **(b)** Coomassie blue – stained SDS PAGE gels (4-20%). **(c)** Values of relative intensities of the bands corresponding to Orco (Western-blot column) in panel A and to the lanes (Coomassie) in panel B. Corrected intensities of bands relative to the intensities of lanes are indicated in the column WB/Coomassie.

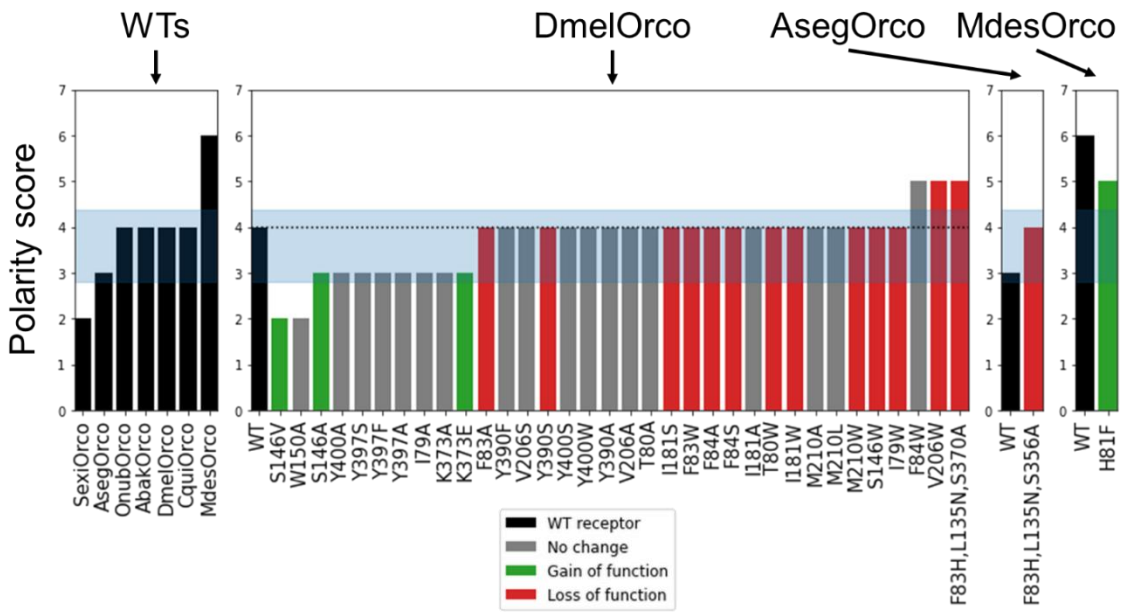


Figure S5. Evolution of the polarity score of DmelOrco's mutants and from other species computed with fpocket. On the left, the results for WT Orco; for clarity, only the VUAA1-responsive and VUAA1-non-responsive Orcos are shown. SexiOrco and MdesOrco are the receptors with the lowest and highest scores, respectively. The central panel gathers the results for DmelOrco mutants. The two right panel gathers results for AsegOrco and MdesOrco. The blue bands account for the mean and standard deviation among 176 Orcos from 174 species.

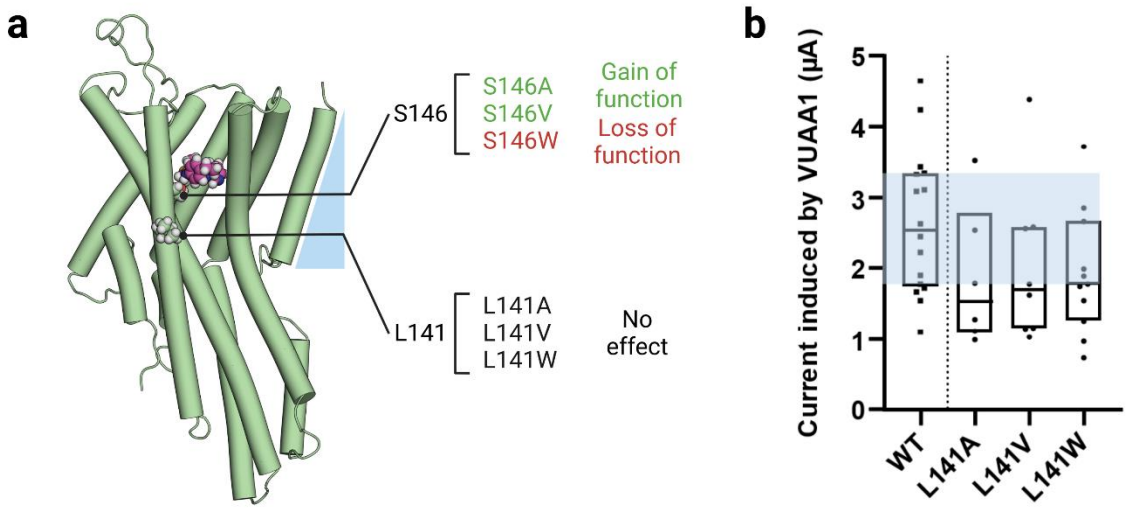


Figure S6. Characterization of mutations at the position L141 that is not involved in the binding site nor in the translocation pathway of VUAA1. (a) Location of the L141 in the model of DmelOrco. For comparison, S146 was also depicted. The carbon atoms from the residue are in pale green, while the ones of VUAA1 are in purple. **(b)** Boxplot showing the current induced by 500µM of the ligand VUAA1 and measured by TEVC recordings on DmelOrco WT and mutants expressed by the Xenopus oocyte (n≥ 5). P values are 0.290, 0.396, 0.192 for L141A, L141V and L141W, respectively. Data are analysed with one-way ANOVA with α-error= 0.05 followed by Dunn's post-hoc test, with WT used as reference for multiple comparisons test.

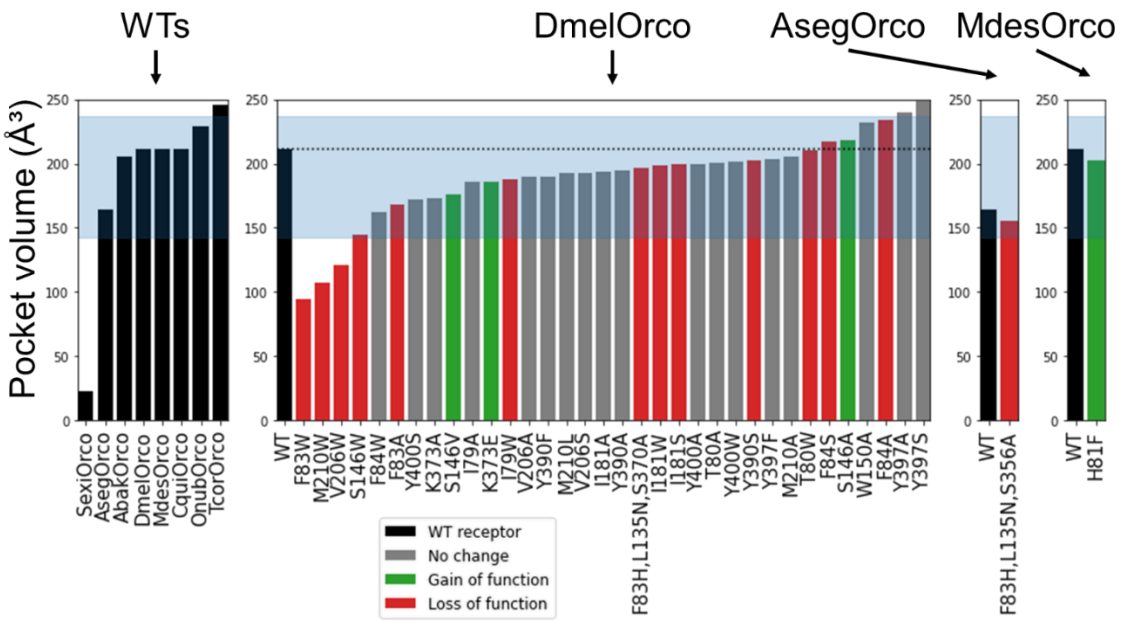


Figure S7. Evolution of the binding cavity volume of DmelOrco's mutants and from other species computed with fpocket. On the left, the results for WT Orcos. For more clarity, only the VUAA1-responsive and VUAA1-non-responsive Orco are shown. SexiOrco and TcorOrco are the receptors with the lowest and highest volume, respectively. The central panel gathers the results for DmelOrco mutants. The two right panel gathers results for AsegOrco and MdesOrco. The blue bands account for the mean and standard deviation among 176 Orcos from 174 species.

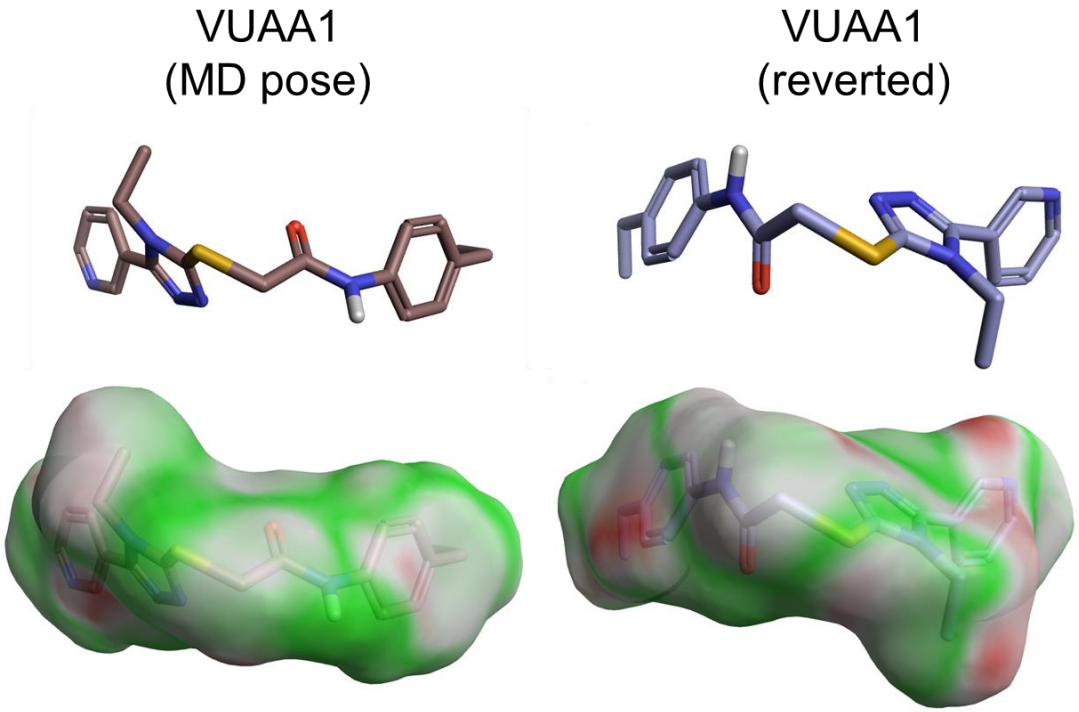


Figure S8. Comparison of the orientation of VUAA1 in the binding site obtained during the MD simulations (MD pose) and the manually reversed orientation. The electrostatic complementarity of the ligand to the protein appears as a surface. The areas where the protein-ligand electrostatics are favorable or unfavorable are colored from green to red.

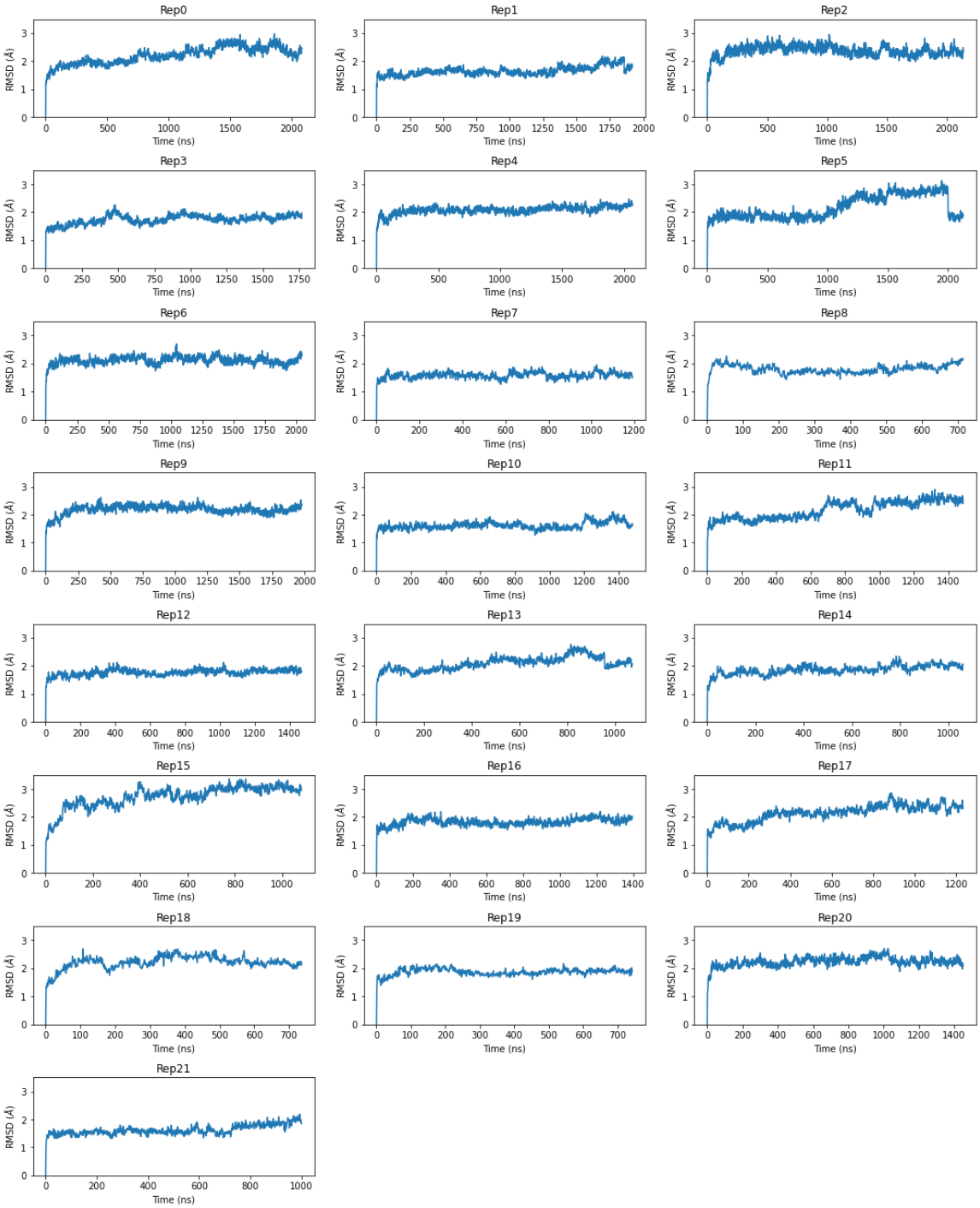
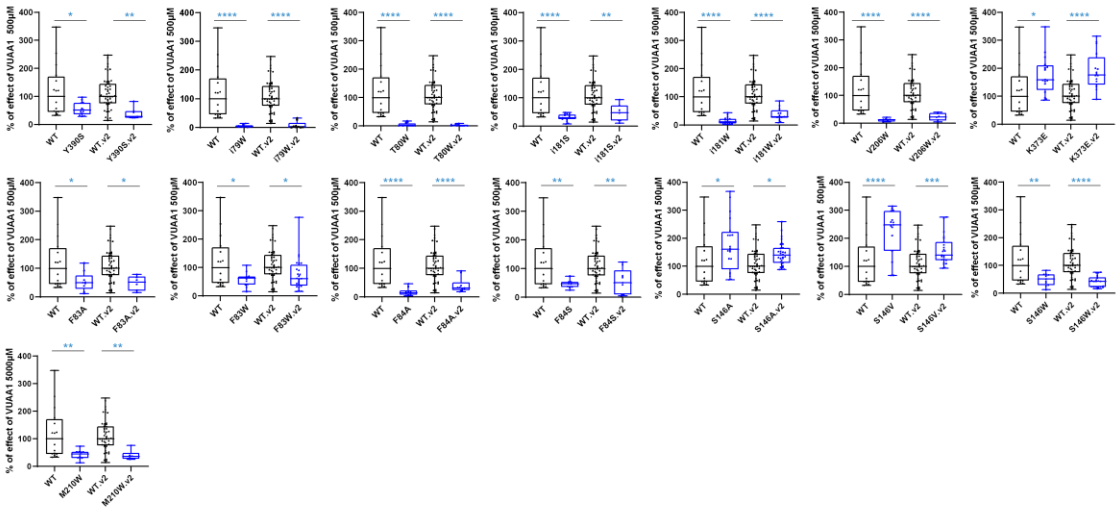


Figure S9. Stability of the DmelOrco model in 22 replicas. The RMSD is computed for the receptor backbone (CA, C, N atoms) with respect to the initial model structure.



WT or mutant Dme/Orco	WT	Y390S	I79W	T80W	I181S	I181W	V206W	K373E	F83A	F83W	F84A	F84S	S146A	S146V	S146W	M210W	WT.v2	Y390S.v2	I79W.v2	T80W.v2	I181S.v2	I181W.v2	V206W.v2	K373E.v2	F83A.v2	F83W.v2	F84A.v2	F84S.v2	S146A.v2	S146V.v2	S146W.v2	M210W.v2
median current (µA)	4,71	2,44	0,14	0,03	1,40	0,54	0,59	7,47	2,33	2,94	0,68	2,19	7,55	11,68	2,44	2,12	3,37	0,97	0,07	0,03	1,59	0,97	0,77	5,95	1,74	2,06	1,07	1,70	4,70	4,74	1,45	1,22
% of activation	100,00	51,86	2,98	0,64	29,76	11,48	12,54	158,66	49,52	62,38	14,35	46,55	160,47	248,25	51,86	45,04	100,00	28,78	2,08	0,89	47,18	28,78	22,85	176,56	51,63	61,13	31,60	50,45	139,47	140,65	42,88	36,05

Figure S10. Relative responses to 500 µM VUAA1 of mutants (blue bars) compared to WT (black bars) in 2 different sets of experiments. Percentages are calculated based on the median of WT current amplitudes as 100%. The experiments in the 2nd set (noted –v2) have been performed in the same day and with the same batch of oocytes as WT-v2. The number of recordings are between 6 to 31 for mutants-v2 and 49 for WT-v2. P values are symbolized over bars as blue stars, each star being one decimal. Data are analysed with one-way ANOVA with α-error= 0.05 followed by Dunn's post-hoc test, with WT used as reference for multiple comparisons test.

936 **Table S1 | Electrostatic and hydrophobic complementary between the series of VUAA1**
937 **analogues and different stable locations in the protein.** The location are defined according to
938 Figure 3: entry (b), vestibule (c) and binding site (d). The analogues have been ranked according
939 to the EC_{50} measured by Tailor *et. al*. on AgamOrco/AgamOR65 (1). The reversed ligand account
940 for the reversed orientation described in Figure S7.

Location	Ligand (EC_{50})	Hydrophobic complementarity	Electrostatic complementarity
Entry (b)	VUAA0.5	0.37	0.26
	VUAA1	0.41	0.29
	VUAA2	0.44	0.31
	VUAA3	0.46	0.33
	VUAA4	0.45	0.32
Vestibule (c)	VUAA0.5	0.61	0.28
	VUAA1	0.62	0.30
	VUAA2	0.63	0.31
	VUAA3	0.64	0.29
	VUAA4	0.66	0.29
Binding site (d)	VUAA0.5 (110 µM)	0.65	0.33
	VUAA1 (37 µM)	0.64	0.34
	VUAA2 (9.2 µM)	0.64	0.36
	VUAA3 (8.4 µM)	0.71	0.36
	VUAA4 (2.1 µM)	0.72	0.36
Binding site (d) with reversed ligand	VUAA0.5	0.58	0.28
	VUAA1	0.58	0.31
	VUAA2	0.58	0.31
	VUAA3	0.65	0.30
	VUAA4	0.64	0.32

941

942 The electrostatic complementarity of the VUAA series in the binding site show little differences in
943 EC values, although it shows a correct trend (pearson correlation of -0.93). These small differences
944 are in the same order of magnitude with a previous study from Bauer et al.(2) In their study, the 6
945 ligands are discriminated based on an EC score with subtle EC variations (lig18: IC50 = 2nM,
946 EC = 0.425; lig23: IC50 = 447nM, EC = 0.385).

947

Table S2 | EC_{50} and Imax of VUAA1 from WT and mutated DmelOrco. ND means Not Determined due to the absence of plateau. Separation lines differentiate WT from mutants in the vestibule (Figure 5) and the binding site (Figure 6), in this respective order. Within separation lines, mutants are ranked in descending order of Imax, which is measured in concentration-effect curves and slightly lower than the Imax measured in single-concentration experiments.

	EC_{50} (µM)	I_{max} (µA)
WT	94.5	3.97
I181S	101.9	0.94
I79W	ND	0.65
V206W	ND	0.33
T80W	ND	0.23
S146V	82.2	6.99
M210W	120.2	1.50
F84A	105.0	1.01

Table S3 | Comparison of the mutation effects on MhraOR5, MhraOR1 eugenol-induced channel response and DmelOrco VUAA1-induced channel response. A, D, I, na and nt mean Abolished, Decreased, Increased, non-affected and not tested, respectively. Results for MhraOR5 and MhraOR1 are taken from ref(3). Each line corresponds to the same position in the three considered proteins MhraOR5, MhraOR1 and DmelOrco. Alignment between the sequences was taken from ref(2).

	MhraOR5	Effect on Eugenol	Effect on DEET	MhraOR1	Effect on Eugenol	DmelOrco	Effect on VUAA1
Ligand diffusion	T87A	D		L102A	n.a	I79A	na
	L379A	na		L398A	n.a	G399	nt
Binding site	V88A	D		T103	A	T80A	na
	Y91A	D		Y106A	A	F83A	D
	F92A	A		I107A	A	F84A	D
	S151A	A		S166A	A	S146A	I
	G154A	A		G169A	I	A149	nt
	W158A	A		W173A	A	W150A	na
	M209A	A	D	L227A	D	V206A	na
	M209V	D	I			V206S	n.a
	M209L	D	D			V206W	A
	I213A	A	A	M231A	A	M210A	na
	I231M	D	A			M210W	D
	L379A	na		L398A	na	G399	nt
	Y380A	A		Y339A	A	Y400A	na
	Y383A	A		C402A	na	A403	nt

Dataset S1 (Dataset S1.xlx). Sequence alignment of different Orco. The known mutations are indicated in color on the sequence.

Dataset S2 (Dataset S2.xlx). Contact frequencies between VUAA1 and amino-acids from Orco during the MD simulations. The first sheet gathers the contact frequencies for the simulations in which the ligand sampled the binding site. The second sheet gather the frequency when the ligand visits cavity b, c or d but do not reach the binding site.

SI References

1. R. W. Taylor *et al.*, Structure–activity relationship of a broad-spectrum insect odorant receptor agonist. *ACS Chem. Biol.* **7**, 1647-1652 (2012).

2. M. R. Bauer, M. D. Mackey, Electrostatic complementarity as a fast and effective tool to optimize binding and selectivity of protein–ligand complexes. *J. Med. Chem.* **62**, 3036-3050 (2019).

3. J. Del Marmol, M. A. Yedlin, V. Ruta, The structural basis of odorant recognition in insect olfactory receptors. *Nature* **597**, 126-131 (2021).