Two mosquito odorant receptors with reciprocal specificity mediated by a single amino acid residue

Flavia P. Franco[a,1], Pingxi Xu[a,1], Brandon J. Harris[b], Vladimir Yarov-Yarovoy[b,c], and Walter S. Leal[a,*]

[a]Department of Molecular and Cellular Biology, [b]Department of Physiology and Membrane Biology, and [c]Department of Anesthesiology and Pain Medicine, University of California, Davis, Davis, CA 95616 USA

*Walter S. Leal

Email: wsleal@ucdavis.edu

Author Contributions: W.S.L., P.X., and V.Y-Y. designed research; F.P.F., P.X. B.J.H, V.Y-Y., and W.S.L. performed research and analyzed the data; W.S.L. and B.J.H wrote the paper, and F.P.F., P.X., and V.Y-Y. edited the manuscript.

Competing Interest Statement: The authors declare no competing interest.

Dedicated to the late Professor James H. Tumlinson III

This PDF file includes:

Main Text
Figures 1 to 5

Abstract

Insects perceive the environment with a repertoire of promiscuous and narrowly tuned odorant receptors (ORs). The southern house mosquito, *Culex quinquefasciatus,* utilizes two ORs, CquiOR10 and CquiOR2, which are narrowly tuned and well conserved among mosquito species. They detect the oviposition attractants skatole and indole, respectively, with reciprocal specificity. We envisioned that by reverting specificity with chimeric receptors, we could gain a better insight into the specificity determinants of insect ORs. We swapped the seven transmembrane (TM) domains of CquiOR10 and CquiOR2 and identified TM2 as a specificity determinant. With additional mutations, we showed that CquiOR10A73L behaved like CquiOR2. Conversely, CquiOR2L74A recapitulated CquiOR10 specificity. Next, we generated structural models of CquiOR10 and CquiOR10A73L using RoseTTAFold and AlphaFold and docked skatole and indole using RosettaLigand. These modeling studies suggested space-filling constraints around A73, although this residue in CquiOR10 did not directly form contacts with skatole or indole. Consistent with this hypothesis, CquiOR10 mutants with a bulkier residue (Ile, Val) were insensitive to skatole and indole, whereas CquiOR10A73G retained the specificity to skatole and showed a higher sensitivity than the wild-type receptor CquiOR10. On the other hand, Leu to Gly mutation of the indole receptor CquiOR2 reverted the specificity to skatole. Lastly, CquiOR10 was insensitive to 3-ethylindole, whereas CquiOR10A73G responded to the three ligands with 3-ethylindole eliciting a slightly higher response than indole. We then concluded that the specificity of these receptors is mediated by a single amino acid substitution leading to dilation or contraction of the binding pocket.

Keywords: *Culex quinquefasciatus, CquiOR2, CquiOR10, Rosetta, RosettaLigand*

Introduction

Mosquitoes are vectors of pathogens that cause tremendous harm to public health. Male and female mosquitoes visit plants to obtain nutrients for flight. For reproduction and survival of the species, females must acquire a blood meal to fertilize their eggs and, subsequently, oviposit in an aquatic environment suitable for the offspring to flourish. While feeding on hosts, females transmit viruses and other pathogens. The southern house mosquito, *Culex quinquefasciatus*, transmits pathogens causing filariasis and various encephalitis (1). In the United States, mosquitoes belonging to the *Culex pipiens* complex transmit the West Nile virus (2). Due to its opportunistic feeding on avian and mammalian hosts, *Cx. quinquefasciatus* is a significant bridge vector in urbanized centers in the Western United States, particularly southern California (2, 3). Female mosquitoes rely on multiple sensory modalities, including olfaction, to find plants, vertebrate hosts, and suitable environments for oviposition.

The *Cx. quinquefasciatus* genome has the largest repertoire of odorant receptor (OR) genes, thus, reflecting the importance of olfaction in the chemical ecology of the southern house mosquito. The genomes of *Anopheles darlingi,* the malaria mosquito, *Anopheles gambiae*, the yellow fever mosquito, *Aedes aegypti*, the Asian tiger mosquito, *Aedes albopictus,* and the southern house mosquito contain 18 (4), 79 (5), 117-131 (6, 7), 158 (8) and 180 (9) *OR* genes, respectively. Of those, 61 transcripts were found in *An. gambiae* (10), 107 in *Ae. aegypti (11)*, and 177 in *Cx. quinquefasciatus* (12). Given that the number of odorants in the environment is larger than the number of OR genes even in *Cx. quinquefasciatus*, it is not surprising that many odorant receptors from insects are promiscuous (13, 14). However, ORs detecting behaviorally critical compounds (semiochemicals) may be narrowly tuned. Odorant receptors narrowly tuned to 3-methylindole (=skatole) and indole have been found in mosquito species in the subfamilies Culicinae and Anophelinae. Specifically, OR10 and OR2 have been de-orphanized in the southern house mosquito (15, 16), the yellow fever mosquito (17), and the malaria mosquito (18, 19). Recently, these so-called indolergic receptors (20) have also been found in the housefly, *Musca domestica* (21). Remarkably, in all these species, OR10s are narrowly tuned to skatole and respond with lower sensitivity to indole, whereas indole is the best ligand for OR2s, which give lower responses to skatole. In *Cx. quinquefasciatus,* these receptor genes were initially named *CquiOR10* and *CquiOR2*, respectively, renamed *CquiOR21* and *CquiOR121*, but a proposition to restore the original names is under consideration (Carolyn McBride personal communication). Here, we refer to these receptors from *Cx. quinquefasciatus* as CquiOR10 (VectorBase and GenBank IDs, CPIJ002479 and GU945397, respectively) and CquiOR2 (CPIJ014392, GU945396.1).

Skatole and indole are fecal products that have been identified as oviposition attractants for the southern house mosquito (22-24). Skatole is a potent oviposition attractant in minute doses, whereas indole is active only at high doses (22). To the human nose, skatole has a pungent fecal odor. In contrast, indole has an almost floral odor when highly purified and presented at low doses (25), but a fecal odor at high doses. These odorants-detecting receptors in *Cx. quinquefasciatus*, CquiOR10 and CquiOR2 provide a suitable model to identify specific determinants as they respond to skatole and indole with reverse specificity. CquiOR10 and CquiOR2 proteins have 377 and 375 amino acid residues, respectively, and share only 50% amino acid identity. To identify the specificity determinants of these receptors, we swapped transmembrane (TM) domains and tested the chimeric receptors using the *Xenopus* oocyte recording system. Given that CquiOR10 is the second most sensitive *Cx. quinquefasciatus* OR (second only to CquiOR36 (26)), we replaced the predicted (27) transmembrane domains from CquiOR2 into CquiOR10 and measured the specificity of the chimeric receptors. With this approach, we identified TM2 as a specificity determinant. Next, we tested mutations of chimeric CquiOR10 and identified a single amino acid residue in TM2 (Ala-73) that determines the receptor's specificity. We then directly mutated the wild-type receptor and observed that CquiOR10A73L is specific to indole as CquiOR2. Additionally, CquiOR2L74A emulated the

response profile of CquiOR10. To gain a better understanding of the structural basis of this one-point specificity determinant, we generated structural models of CquiOR10 and CquiOR10A73L using RoseTTAFold (28) and AlphaFold (29), and identified the binding poses for skatole and indole using RosettaLigand molecular docking.

Results and Discussion

Chimeric OR with reversed specificity. We envisioned that studying a pair of odorant receptors with reverse specificities, like CquiOR10 (Fig. 1*A*) and CquiOR2 (Fig. 1*B*), could lead us to specificity determinants. Our approach was designed to swap TM domains using the more sensitive receptor, CquiOR10, as the acceptor. Specifically, we generated chimeric receptors by replacing CquiOR10 transmembrane domains with related domains from CquiOR2 (*SI Appendix*, Fig. S1). We referred to these chimeric receptors as CquiOR10^{Mx}, where Mx refers to TMx from CquiOR2. Then, we performed functional assays of these CquiOR10^{Mx} receptors using the *Xenopus* oocyte recording system. This long-term project could require as many as 127 possible chimeric receptors. Of note, during the life of this project, the structure of an odorant receptor coreceptor AbakOrco from the parasitic fig wasp, *Aprocrypta bakeri*, was reported (30). We then compared the Cryo-EM structure (30) with the predicted topology for AbakOrco using the same OCTOPUS method (31) we used to identify CquiOR10 and CquiOR2 TMs (*SI Appendix*, Fig. S1) (31). The almost perfect overlap between OCTOPUS prediction and the Cryo-EM structure (*SI Appendix*, Fig. S2) validated not only our TM predictions (*SI Appendix*, Fig. S2), but also the 21 chimeric ORs already tested when the structure of the coreceptor AbakOrco was reported (32, 33).

We optimized the workflow to reduce the number of tested chimeric receptors before finding one receptor with reversed specificity. Our first attempt was to replace all seven transmembrane domains. CquiOR10^{M1,2,3,4,5,6,7}/CquiOrco-expressing oocytes did not respond to skatole or indole. Likewise, the chimeric receptor with six transmembrane domains replaced, CquiOR10^{M1,2,3,4,5,6} was also silent. Next, we replaced one TM at a time. CquiOR10^{M1} and CquiOR10^{M2} were silent, whereas CquiOR10^{M3} (Fig. 1*C*), CquiOR10^{M4} (Fig. 1*D*), CquiOR10^{M5} (Fig. 1*E*), CquiOR10^{M6} (Fig. 1*F*), and CquiOR10^{M7} (Fig. 1*G*) responded to both skatole and indole with specificity similar to that of the wild-type receptor. Although these chimeric receptors differed from CquiOR10 in terms of sensitivity, they showed the same specificity to skatole as CquiOR10 (Fig. 1*A*). While CquiOR10^{M3} and CquiOR10^{M4} were less sensitive than CquiOR10, CquiOR10^{M7} showed comparable sensitivity to CquiOR10, and CquiOR10^{M5} and CquiOR10^{M6} were more sensitive to skatole than the wild-type receptor. We remained focused on specificity and did not attempt to quantify the different responses.

Because CquiOR10^{M1} and CquiOR10^{M2} were not functional, we initially assumed that amino acid residues in these domains were not specificity determinants. We tested a chimeric receptor with the other five TM domains swapped. CquiOR10^{M3,4,5,6,7}/CquiOrco-expressing oocytes mirrored the profiles obtained with CquiOR10 in terms of sensitivity and specificity (Fig. 1*H*). We then kept the wild-type TM7, in addition to TM1 and TM2, in the next chimeric ORs tested. Albeit less sensitive than CquiOR10, CquiOR10^{M3,4,5,6} showed the same specificity of the wild-type receptor (Fig. 1*I*). Following, we tested chimeric receptors with three TM domains swapped. CquiOR10^{M4,5,6} (Fig. 1*J*), CquiOR10^{M3,5,6} (Fig. 1*K*), CquiOR10^{M3,4,6} (Fig. 1*L*), and CquiOR10^{M3,4,5} (Fig. 1*M*) did not differ from CquiOR10 in terms of specificity or sensitivity. We then surmised that multiple domains swapping might not lead us to reverse specificity because of other possible structural compensations. Consequently, we focused on chimeric receptors with two TMs swapped. Thus, we tested CquiOR10^{M3,4} (Fig. 1*N*), CquiOR10^{M3,5} (Fig. 1*O*), CquiOR10^{M3,6} (Fig. 1*P*), CquiOR10^{M3,7} (Fig. 1*Q*), CquiOR10^{M4,5} (Fig. 1*R*), CquiOR10^{M4,6} (Fig. 1*S*), CquiOR10^{M4,7} (Fig. 1*T*), CquiOR10^{M5,6} (Fig. 1*U*), CquiOR10^{M5,7} (Fig. 1*V*), and CquiOR10^{M6,7} (Fig. 1*W*). They all showed skatole-specific profiles, most of them with comparable sensitivity, whereas

CquiOR10^{M4,5}, CquiOR10^{M5,7}, and CquiOR10^{M6,7} were more sensitive and CquiOR10^{M3,6} less sensitive than CquiOR10.

Interestingly, none of the chimeric receptors having TM1 as one of the two domains swapped was functional. Specifically, oocytes co-expressing CquiOrco with either CquiOR10^{M1,2}, CquiOR10^{M1,3}, CquiOR10^{M1,4}, CquiOR10^{M1,5}, CquiOR10^{M1,6}, CquiOR10^{M1,7} did not respond to skatole or indole. Likewise, CquiOR10^{M2,3}, CquiOR10^{M2,4}, CquiOR10^{M2,5}, and CquiOR10^{M2,6} did not respond to these oviposition attractants. However, CquiOR10^{M2,7}/CquiOrco-expressing oocytes responded to both skatole and indole but with a reverse profile (Fig. 1X). This dataset shows that CquiOR10^{M2,7} (Fig. 1*X*) emulated the profile of the indole receptor CquiOR2 (Fig. 1*B*). These findings were surprising given that swapping TM2 impaired receptor function, and TM7 per se did not affect the specificity of the chimeric receptor. It has been suggested (30) and recently corroborated (34) that TM7 is not part of the binding pocket but rather essential for forming the receptor's ion channel.

A one-point mutation that reverses specificity of the skatole and indole receptors. We proceeded to identify the amino acid residues in the swapped domains of CquiOR10^{M2,7}, directly affecting the specificity of the chimeric and wild-type receptors. Given the observation that, by and large, chimeric receptors with TM5 and TM6 from CquiOR2 were more sensitive, we asked whether CquiOR10^{M2,7} with these two additional TM domains swapped would be more sensitive while keeping the same specificity to indole. Our rationale was that a more sensitive chimeric receptor might prevent us from missing essential mutations should sensitivity be reduced by substitutions/mutations. CquiOR10^{M2,5,6,7}/CquiOrco-expressing oocytes were indeed more sensitive while maintaining the selective to indole (Fig 1Y). We then used CquiOR10^{M2,5,6,7} and designed various mutants to rescue single or multiple residues in TM2 at a time (*SI Appendix*, Fig. S3). Our goal was to identify mutations in the chimeric receptor that reverts the specificity to skatole as in the wild-type receptor. We divided TM2 into outer, middle, and inner segments based on the topology predicted by OCTOPUS (31). It has been postulated that the extracellular halves of TM domains form an odorant-binding pocket (35), thus we first examined a mutant (CquiOR10^{M2,5,6,7}_Outer) having the residues in the outer segment at 59, 60, 63, and 64 restored to Glu, Val, Asn, and Ala as in the wild-type receptor (*SI Appendix*, Fig. S3). CquiOR10^{M2,5,6,7}_Outer/CquiOrco-expressing oocytes retained the specificity of CquiOR10^{M2,5,6,7} with a more robust response to indole than skatole (Fig. 2*A*). These findings indicated that the residues in the outer segment of TM2 are not specificity determinants. After that, we tested a chimeric receptor with the residues in the middle and inner part of the TM2 domain rescued to match those in the wild-type receptor (*SI Appendix*, Fig. S3). CquiOR10^{M2,5,6,7}_Mid;Inner restored the skatole-specific profile of CquiOR10 (Fig. 2*B*), thus suggesting that amino acid residues in these segments are specificity determinants. Then, we tested CquiOR10^{M2,5,6,7}_Inner (=CquiOR10^{M2,5,6,7}L73A;T78I), which had only residues at 73 and 78 rescued to the wild-type Ala and Ile, respectively. Oocytes co-expressing CquiOR10^{M2,5,6,7}_Inner and CquiOrco reverted the specificity to skatole (Fig. 2*C*). As a result, we concluded that residues in the predicted inner part of TM2 are critical for the chimeric receptor's specificity. We further probed the chimeric receptor CquiOR10^{M2,5,6,7} with one-point mutations to identify the residue(s) determining specificity. CquiOR10^{M2,5,6,7}T78I/CquiOrco-expressing oocytes showed the same specificity as the chimeric receptor CquiOR10^{M2,5,6,7} (Fig. 2*D*). Specifically, CquiOR10^{M2,5,6,7}T78I gave a more robust response to indole than skatole, suggesting that rescuing the residue at 78 did not affect CquiOR10^{M2,5,6,7} specificity. By contrast, CquiOR10^{M2,5,6,7}L73A/CquiOrco-expressing oocytes reverted the specificity to skatole, thus behaving like the wild-type receptor CquiOR10(Fig. 2*E*). To further examine the role of Ala-73 as a specificity determinant residue, we obtained a one-point mutation of CquiOR10^{M7}, which is specific to skatole (Fig. 1*G*). The responses recorded from CquiOR10^{M7}A73L/CquiOrco-expressing oocytes showed a reverse, indole-specific profile (Fig. 2*F*). Having identified a single amino acid residue in the chimeric receptor that switches the skatole/indole specificity, we proceeded to test the effect of one-point mutation on the specificity of the wild-type receptor CquiOR10. CquiOR10A73L showed a reverse specificity, with dose-dependent responses to indole (Fig. 2*G*). Taken together, these findings suggest that a single

amino acid residue in CquiOR10 determines the specificity of this receptor. Finally, we obtained an equivalent one-point mutation in the indole-specific CquiOR2 (Fig. 1*B*). Thus, CquiOR2L74A/CquiOrco-expressing oocytes gave robust and specific responses to skatole (Fig. 2*H*). As summarized in a graphical representation (*SI Appendix*, Fig. S4), these findings demonstrate that these two mosquito odorant receptors, CquiOR10 and CquiOR2, have reciprocal specificity mediated by a single amino acid residue, Ala-73 and Leu-74, respectively.

CquiOR10 computational modeling suggests space-filling constraints for indole-based odorants around A73. To gain better insights on the structural basis of the above-described unique reciprocal specificity, we generated structural models of CquiOr10, CquiOR2, CquiOR10A73L and CquiOR2L74A with RoseTTAFold (28) and AlphaFold (29) (*SI Appendix*, Fig. S5). CquiOR10 models of one homotetramer subunit were generated with AlphaFold and RoseTTAFold, each producing five models. The root mean square deviation (RMSD) of the entire subunit across all ten models was less than 1 Å. RoseTTAFold models of CquiOR10, CquiOR10A73L, CquiOR2, and CquiOR2A73L also produced an RMSD less than 1 Å across all twenty models (five models per structure). Having found no major structural deviations between AlphaFold and RoseTTAFold models, we proceeded with RoseTTAFold models for Rosetta-based small-molecule docking method RosettaLigand (36, 37).

To verify if the RosettaLigand (36, 37) method could effectively sample homologous receptors in complex with odorants, we used the structure of eugenol in complex with the insect odorant receptor OR5 from *Machilis hrabei* (MhraOR5) as a control case (34). Rosetta protein-ligand docking frequently employs energy-based analyses, such as the interface energy between protein and ligand, to select the representative models. With this selection method, the RMSD of our models relative to the experimental coordinates of eugenol ranged from 0.5 – 5.0 Å (*SI Appendix,* Table S1). Of note, the largest cluster had a RMSD of 0.752 Å while the lowest interface-energy model had a RMSD of 2.4 Å. Collectively, however, these data demonstrate that our models agree with previously published data (Fig. 3, *SI Appendix,* Fig. S6).

After recapitulating the structure of the eugenol/MhraOR5 complex (Fig. 3), we docked indole and skatole to CquiOR10 and CquiOR10A73L using the same method. Thus, from our 100,000 generated docking models for each test case (CquiOR10, CquiOR10A73L, CquiOR2, and CquiOR2L73A), we clustered the 10,000 lowest interface-energy models and selected the lowest interface-energy models from the 10 most frequent clusters as representative models for our study. (*SI Appendix*, Table S2, S3). Out of the 10 lowest interface-energy models for each group, CquiOR10-skatole (cluster 3), CquiOR10A73L-skatole (cluster 5) (*SI Appendix*, Table S2**)**, CquiOR10-indole (cluster 5), and CquiOR10A73L-indole (cluster 5) (*SI Appendix*, Table S3**)** showed the lowest interface score within each group (Fig. 4). These 10 most frequent positional clusters from each structural complex suggest that both indole and skatole can readily reorient themselves in a similar position nearby Ala-73 and Leu-73 (*SI Appendix*, Fig. S7-S9). Further, skatole and indole in most positions form contacts with CquiOR10 and CquiOr10A73L in a similar plane about a center of rotation. These observations are supported by skatole and indole not containing rotatable bonds, thus relying on rigid translational movements and rotation to form favorable contacts with the rotatable and repackable receptor residues. Additionally, our models position indole and skatole within a series of nonpolar, polar-uncharged, and aromatic amino acids. Protein-ligand interaction profiler (PLIP) analysis (38) suggests that the bulk of favorable interactions are nonpolar, occasional hydrogen bonding with the odorant NH group, and occasional parallel pi stacking (*SI Appendix*, Fig. S7).

The architectures of the binding pocket in CquiOR10 (Fig. 4A,B, *SI Appendix*, Fig. S7) differ from that of a previously reported OR (34) in that no residues from TM3 forms contact with the ligands. Contrary to MhraOR5 (34), the ligand binding pocket is formed by TMs 2, 4, 5, and 6. Tyr-185 (in TM4) and Ile-69 (TM2) formed the most frequent hydrophobic contacts (*SI Appendix*, Table S4,

S5) with the ligand in the four structures analyzed (38/40 and 32/40 models, respectively, *SI Appendix*, Table S6): CquiOR10-skatole, CquiOR10-indole, CquiOR10A73L-skatole, and CquiOR10A73L-indole. Phe-261(TM5) also formed hydrophobic interactions with all four structures, but less frequently (22/40 models). Asn-72 (TM2), Tyr-287 (TM6), and Ile-291 (TM6) formed hydrophobic contacts more frequently with skatole structures compared to indole structures (Asn-72: 12/20 vs 9/20, Tyr-287: 14/20 vs 5/20, Ile-291: 8/20 vs 3/20). Notably, Leu-73 (TM2) in CquiOR10A73L formed hydrophobic contacts with indole (2/10) and skatole (5/10), but Ala-73 (TM2) did not form hydrophobic contacts with indole or skatole. Tyr-138 (TM3) formed hydrophobic contacts with skatole in CquiOR10A73L (2/10), but not in CquiOR10. Gln-294 (TM6) formed hydrophobic contacts with skatole and indole in these structures, except CquiOR10A73L-skatole; however, Gln-294 does form a hydrogen bond in one to two representative models for each structure. Of note, the range of PLIP-reported hydrophobic contacts among the ten representative models for each structure changed from an average of 6.3±0.5 contacts (Mean±SEM) in CquiOR10 to 5.2±0.3 contacts in CquiOR10A73L ($p = 0.0354$, Mann-Whitney test). Considering the number of unique residues forming hydrophobic contacts, the averages were 4.9±0.4 and 4.9±0.2 contacts, respectively ($p = 0.9704$, Mann-Whitney test). More importantly, skatole and indole did not form contacts with Ala-73 in any of the representative models. By contrast, skatole formed hydrophobic contacts with Leu-73 in in five of the ten representative models, while indole formed contacts in two representative models.

There were at most two models in each structure with hydrogen bonds (*SI Appendix*, Table S7-S9). For CquiOR10-skatole, 2 models formed hydrogen bonds with Gln-294. For CquiOR10A73L-skatole, 1 model formed hydrogen bonds with Asn-72, and one other with Gln-294 (*SI Appendix*, Table S9). For CquiOr10-indole, one model formed a hydrogen bond with the backbone of Ile69, another model with Tyr-138, another with Tyr-287, and a fourth model with Gln-294. For CquiOR10A73L-indole, only one model formed a hydrogen bond at Gln-294.

Parallel pi-stacking interactions (*SI Appendix*, Table S10-S12) were identified for CquiOR10A73L-skatole, CquiOR10-indole, and CquiOR10A73L-indole. For CquiOR10A73L-skatole, two models each formed a parallel pi-stacking interaction with Tyr-185 (*SI Appendix*, Table S12). For CquiOR10-indole, a model formed a parallel pi-stacking interaction with Tyr-138 and another model formed with Tyr-152. For CquiOR10A73L-indole, eight parallel pi-stacking interactions with Tyr-185 were identified across 7 representative models.

Ala-73 did not form direct contact with skatole or indole, but our modeling suggests that it indirectly affects specificity by modulating the volume of the binding pocket. Structurally aligning CquiOr10 and CquiOr10A73L receptors by TM7b (*SI Appendix*, Table S13), demonstrates approximately 1 Å α-Carbon outward shift of A73L (*SI Appendix*, Table S14), suggesting a tightly constrained space in CquiOR10; and expanded space in CquiOR10A73L relative to the protein backbone (Fig. 5). These contractions/dilations of the binding pocket are not ligand induced, but rather independent of skatole or indole binding. Thus, these findings are counterintuitive as the bulkier ligand (skatole) may fit better in the contracted conformation of the CquiOR10 receptor, as indicated by the receptor's response (Fig. 1A). Likewise, the smaller ligand (indole) showed an apparent higher affinity for the protein with the expanded binding pocket, as indicated by the receptor's response (Fig. 2G).

The reason for improved receptor response of skatole with Ala-73 over Leu-73, and improved receptor response of indole with Leu-73 over Ala-73 could be due to the repacking of this position to accommodate more carbon sidechain atoms that are reducing the binding pocket volume. Leu-73 contains three additional carbon sidechain atoms: γ-Carbon, δ1-Carbon, and δ2-Carbon. Adding these atoms increases the sidechain rotatable bond count from 0 (for Ala-73) to 2 (for Leu-73), causing an increase in repackable states for the binding pocket. From our representative models, the distance of Leu-73 α-Carbon to its δ1-Carbon or δ2-Carbon is approximately 2.8 to 3.9 Å. The Ala-73 α-Carbon to its β-Carbon is approximately 1.5 Å. Since the Ala-73 α-Carbon is approximately 0.9 to 1.1 Å inward relative to Leu-73 α-Carbon, Leu-73 C-δ atoms could be further

inward by approximately 0.2-1.5 Å relative to Ala-73 C-β. This could cause repacking of surrounding side chains while reducing the volume of the binding pocket, or shift key, nearby residues required for receptor response. Further, from our representative model, the Carbon-Carbon length of skatole's methyl group is approximately 1.5 Å. This distance discrepency together with the fact that indole is a rigid molecule and the methyl group of skatole is the only rotamer, could also suggest that the binding pocket around position 73 is tightly regulated. The 1.5 Å increase from skatole's methyl group could prevent skatole from occupying the appropriate configuration for receptor response with Leu-73. Conversely the additional inward 0.2-1.5 Å increase from Leu-73's sidechain could provide indole more contacts or a more favorable configuration with surrounding CquiOR10 sidechains to prompt receptor response. This is supported by seven of ten representative models of CquiOR10A73L – indole having eight total parallel pi-stacking interactions with Tyr-185; this is a considerable increase compared to two representative interactions of CquiOR10 – indole having two total parallel pi-stacking interactions with Tyr-138 and Tyr-152.

Consistent with the proposed model of space constraints, mutation of Ala-73 in CquiOR10 to Ile or Val negatively affected receptor function, whereas CquiOR10A73G retained specificity and showed higher sensitivity. Specifically, CquiOR10A73I/CquiOrco- and CquiOR10A73V/CquiOrco-expressing oocytes did not respond to skatole or indole (*SI Appendix*, Fig. S10*A,B*). We concluded that mutations with these bulkier residues caused loss of binding to indole or skatole, given that these receptors were functional as indicated by the potent responses elicited by OLC12. On the other hand, CquiOR10A73G/CquiOrco-expressing oocytes showed the same skatole specificity as the wild-type receptor (*SI Appendix*, Fig. S10*C*). Interestingly, CquiOR10A73G showed a remarkably higher sensitivity to skatole than the wild-type receptor CquiOR10 (*SI Appendix*, Fig. S11), thus implying a higher affinity for skatole with a reduced constraint, or increased volume, of the binding pocket.

We indirectly examined space constraints in CquiOR2 by mutating Leu-74 to Ile, Val, or Gly. We then tested oocytes co-expressing CquiOrco with CquiOR2L74I, CquiOR2L74V, or CquiOR2L74G (*SI Appendix*, Fig. S10*D-F*). Consistent with the relaxation of the space constraints, CquiOR2L74I retained the specificity of the wild-type receptor to indole, whereas CquiOR2L74G showed a reverse specificity profile (*SI Appendix*, Fig. S10*F*), i.e. more robust response to skatole than indole. Interestingly, CquiOR2L74V/CquiOrco-expressing oocytes generated nearly equal, albeit small, currents when stimulated with the two oviposition attractants (skatole, 8.7±2.7 nA; indole, 9.0±2.5 nA at 100 µM; $n = 3$, $P>0.9999$, Wilcoxon matched-pairs signed-rank test).

Lastly, we tested the space constraints model with a bulkier ligand, 3-ethylindole. CquiOR10/CquiOrco-expressing oocytes did not respond to 3-ethylindole. By contrast, a mutant with a less space-filling residue, CquiOR10A73G, responded to indole (205±81 nA, at 100 µM), skatole (1939±142 nA), and 3-ethyl indole (526±110 nA) (*SI Appendix*, Fig. S12). We surmised that the ethyl group is too large even with the additional space provided by a smaller amino acid residue.

Conclusions

Our mutation studies demonstrated that a single amino acid substitution in two narrowly tuned odorant receptors from the southern house mosquito can revert their specificity to the oviposition attractants skatole and indole. Rosetta modeling and RosettaLigand docking studies suggest that the specificity determinant amino residue (Ala-73) did not form direct contacts with ligands but rather provides a finely-tuned, sensitive volumetric space to accommodate indole-based odorants.

Materials and Methods

Construction of Chimeric Receptors. We used a previously obtained pGEMHE-CquiOR10 (15) plasmid to amplify the sequences of the desired chimeric receptors. To obtain the full length of chimeric *OR* genes, we used specific primers, with overlapping sequences at each end to target sequences (adaptor underlined) for cloning into pGEMHE vector, which was linearized at restriction sites *XmaI* (5'-CCCGGG-3') and *XbaI* (5'-TCTAGA-3'):CqOR10_Fw 5'-GATCAATTCCCCGGGACCATGACCGCGGCACCCATTTT-3' and CqOR10_Rv 5'-CAAGCTTGCTCTAGATCAATTATAAACGCGTCTCAGCAGGGT-3'. To generate a chimeric OR by PCR amplification, we designed specific primers for the desired CquiOR2 TM domain. Simultaneously, we prepared the recipient fragments of CquiOR10 to receive the CquiOR2 TM domains. The fragments were then assembled using In-Fusion HD cloning kit (Clontech) to obtain the desired chimeric OR. Each CquiOR2 TM domain was divided into two parts, each part was synthesized with a specific primer, except for CquiOR2 TM7 domain, which used only one primer with the full domain. These primers contained an overlap with CquiOR10 sequence followed by one part of the CquiOR2 TM sequence and an overlap with the other part of the CquiOR2 TM domain sequence, which was synthesized in another PCR. The underlined sequences represent CquiOR2 TM domains and, unless otherwise specified, sequences in italic are overlaps for CquiOR2 TM domains. CqOR10^{M1}_Fw: *5'-GGTGACCGTGCTGA*ACGTGTTCCAGTTTATGAACCTGTTTCGAGCCTGGGGCAACATC-3'; CqOR10^{M1}_Rv: 5'-*TCAGCACGGTCACC*GGAATGCAGCCGAGGAGGTAACTGAGGACGTTGCCTACTGTGATCTCAAGG-3'; CqOR10^{M2}_Fw: 5'-*CCGTGCTGTACTTC*AACCTTGTGTTGAGAACCACGTTTATACTGTGCAATCGTCAGGATTATGAGG-3'; CqOR10^{M2}_Rv: 5'-*AAGTACAGCACGG*TGAAATATCCGTCGATGATGATTTTGTCGATGTTGCCCCAGGCT-3'; CqOR10^{M3}_Fw: 5'-*GTTCATCAGTGCGT*GCTTCGTGACGTATCCGCTTTTTTCACCGACACGTAGCCTCCCGTACG-3'; CqOR10^{M3}_Rv: 5'-*ACGCACTGATGAAC*GCTCCCAGCCAGAGGTTCGATTTGGACAGCAGTCGGGCACGTTTGGTGA-3'; CqOR10^{M4}_Fw: 5'-*CACGTTTCCGGCGT*GCTGCATGTACATTCCGTTTACCAGCTTCTTCGCCACGACTACTTTG-3'; CqOR10^{M4}_Rv: 5'-*ACGCCGGAAACGTG*AGGTACACTTGCAGAAAAAACACAACCTGGTACAGGGGCGTC-3'; CqOR10^{M5}_Fw: 5'-*GCTATGCGCCTTGC*TGTTTCTACTTAGCACCAGCAATAATCCCGCGCAAATTATCATCGTGG-3'; CqOR10^{M5}_Rv: 5'-*GCAAGGCGCATAGC*ATCATCCCAAACGATAGCAACTCAATCAGACAGATGTAGGTCACCAGCG-3'; CqOR10^{M6}_Fw: 5'-*TCTTTATGATTCT*GTCCCAGATGTACGCCCTGTACTGGCACGCCAACGAGCTGCG-3'; CqOR10^{M6}_Rv: 5'-*AGAATCATAAAGA*TGTACGATCCGATCATCACCATCTGCGCGGGATTTTCGATAATGTTCAGC-3'; CqOR10^{M7}_Rv: 5'-*CAAGCTTGCTCTAGA*TCAATTATAAACGCGTCTCAGCAACGTAAAGTACGAATACGAGGCATTGATCAACTTTTGAAACATTTCCAAGGTCATCGGATAGACGTTGCCTACTGTGATCTCAAGG-3' (here the italic represents the pGEMHE adaptor. The two fragments were amplified using a combination of CqOR10_Fw with CqOR10^{Mx}_Rv primers and CqOR10_Rv with CqOR10^{Mx}_Fw primers. Then, we cloned the two fragments amplified by PCR into pGEMHE vector using the In Fusion system as described below. Chimeric plasmid with a single TM swapped was used as a template to generate chimeric OR with two TM swapped and subsequently chimeric ORs with multiple CquiOR2 TM domains.

Chimeric OR Cloning and Subcloning into pGEMHE. The fragments amplification was performed using Platinum Taq DNA Polymerase High Fidelity (Thermo Fisher Scientific) with the

following conditions: 95°C for 5 min, followed by 30 cycles of 95°C for 20 s, 57°C 30 s for annealing and 68°C for 1.5 min, and extension at 68°C for 5 min. PCR products were purified by QIAquick gel extraction kit (Qiagen). The target pGEMHE plasmid was cut by *XmaI* and *XbaI* in separate reactions. Ligation was done with In Fusion (Takara Bio USA) system and the transformation was performed using Stellar competent cells (Takara Bio USA) in heat-shock. After selecting the cells, plasmids were extracted with QIAprep Spin Miniprep kit (Qiagen). The cloned gene was verified by DNA sequencing (Berkeley Sequencing Facility).

Site-Directed Point Mutagenesis and Fragment replacement.Phusion™ Site-Directed. Mutagenesis Kit (Thermo Scientific, West Palm Beach, FL) was used to generate point mutations and TM fragment replacements. Mutations were created with mismatched 5′-phosphorylated mutagenic primers and PCR amplification. The chimeric sequences (CquiOR10^{M2,5,6,7}) in pGEMHE vector were used as templates for rescues, whereas the wildtype sequences in pGEMHE vector served as templates for point-mutations. Rescue primers: CqOR10M2_M2567OUTERup: 5'-GATGATGACCTCGTCGATGTTGCCCCAGGC-3'; CqOr10M2_M2567OUTERdn: 5'-AACGCATATTTCACCGTGCTGTACTTCAACC-3'; CqOr10M2_M2567INNERup: 5'-CGCAGCACCGCGTTGAAGTACAGCACGGTG-3'; CqOr10M2_M2567INNERdn: 5'-AACAATTTTCATACTGTGCAATCGTCAGGA-3'; CqOR10M2_M2567MID-INNERup: 5'-CATCGACGGCTACTTTGCGATGATTTTCTTCAACGCG-3'; CqOR10M2_M2567MID-INNERdn: 5'-ATGATTTTGTCGATGTTGCCCCAGGCTC-3'; CqOr10M2_M2567_L73Aup: 5'-CGTTGAAGTACAGCACGGTGAAATAT-3'; CqOr10M2_M2567_L73Adn: 5'-CTGTGTTGAGAACCACGTTTATACTGT-3'; CqOr10M2_M2567_T78Iup: 5'-ATGGTTCTCAACACAAGGTTGAAGTA-3'; CqOr10M2_M2567_T78Idn: 5'-CTTTATACTGTGCAATCGTCAGGATTATG-3'; Point mutation primers: CqOR10A73up: 5'-GTTGAAGAAAATCATCGCAAAGTA-3'; CqOR10A73Ldn: 5'-CTGGTGCTGCGAACAATTTTC-3'; CqOR10A73Adn: 5'-GCGGTGCTGCGAACAATTTTC-3'; CqOR10A73Gdn: 5'-GGGGTGCTGCGAACAATTTTC-3'; CqOR10A73Vdn: 5'-GTGGTGCTGCGAACAATTTTC-3'; CqOR10A73Idn: 5'-ATCGTGCTGCGAACAATTTTC-3'; CqOR2L74up: 5'-GTTGAAGTACAGCACGGTGAAATA-3'; CqOr2L74Adn: 5'-GCTGTGTTGAGAACCACGTTTATAC-3'; CqOr2L74Idn: 5'-ATCGTGTTGAGAACCACGTTTATAC-3'; CqOr2L74Gdn: 5'-GGGGTGTTGAGAACCACGTTTATAC-3'; CqOr2L74Vdn: 5'-GTGGTGTTGAGAACCACGTTTATAC-3'. The amplified linear PCR products containing the desired modification were ligated and transformed into Stellar™ Competent Cells (Takara Bio USA). All sequences were confirmed by DNA sequencing (UC Berkeley DNA Sequencing Facility).

In Vitro Transcription, Oocyte Microinjection and Two-Electrode Voltage-Clamp assay (TEVC). Capped OR cRNA was prepared using mMESSAGE mMACHINE T7 Kit (Ambion) as previously described (39). Purified OR cRNA was resuspended in nuclease-free water at 200 ng/μL and microinjected into *Xenopus laevis* oocytes on stage V or VI (EcoCyte Bioscience) along with the same amount of CquiOrco. Injected oocytes were incubated at 18 °C for 3–7 d in a modified Barth's solution [88 mM NaCl, 1 mM KCl, 2.4 mM $NaHCO_3$, 0.82 mM $MgSO_4$, 0.33 mM $Ca(NO_3)_2$, 0.41 mM $CaCl_2$, and 10 mM HEPES at pH 7.4] supplemented with 10 μg/mL gentamycin, 10 μg/mL streptomycin and 1.8 mM sodium pyruvate. Two-electrode voltage-clamp technique (TEVC) was used to measure odorant-induced currents at a holding potential of −80 mV. Signals were amplified with an OC-725C amplifier (Warner Instruments), low-pass filtered at 50 Hz, and digitized at 1 kHz. Data acquisition and analysis were performed with Digidata 1440A and software pCLAMP 10 (Molecular Devices). Skatole (98%, CAS# 83-34-1), indole (98%, CAS# 120-72-9), and 3-ethylindole (97%, CAS# 1484-19-1) were provided by Sigma-Aldrich (Milwaukee, WI), ACROS Organics (Geel, Belgium), and AmBeed (Arlington Hts, IL), respectively.

Rosetta structural modeling and docking. We used RoseTTAFold (28) and AlphaFold (29) to generate predicted structures of CquiOR10, CquiOR2, CquiOR10A73L, and CquiOR2L74A. The OpenEye Omega toolkit (40) was used to generate conformer libraries. Molecular docking was performed using RosettaLigand (36, 37). The top models were visually analyzed using UCSF Chimera (41).

To verify RosettaLigand could effectively sample homologous receptors in complex with odorants, we used the structure of eugenol in complex with the insect olfactory receptor OR5 from *Machilis hrabei* (MhOR5) as a control test case (34). After verifying RosettaLigand could recapitulate the MhraOR5 — eugenol complex from structural studies, we docked indole and skatole to the RoseTTAFold structure models of CquiOR10 and CquiOR10A73L receptors with the same method.

Data Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Drs. David K. Wilson, Enoch Baldwin (UC Davis), and members of our research groups for enlightening discussions. The Orco ligand candidate OLC12 was provided by the Vanderbilt Institute of Chemical Biology, Chemical Synthesis Core, Vanderbilt University. Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health under award number R01AI095514. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

Figures

Figure 1. Concentration-response analysis for activation of wild-type and chimeric ORs by skatole and indole. (A) CquiOR10; (B) CquiOR2; (C)CquiOR10^{M3}; (D) CquiOR10^{M4}; (E) CquiOR10^{M5}; (F) CquiOR10^{M6}; (G) CquiOR10^{M7}; (H) CquiOR10^{M3,4,5,6,7}; (I) CquiOR10^{M3,4,5,6}; (J) CquiOR10^{M4,5,6}; (K) CquiOR10^{M3,5,6}; (L) CquiOR10^{M3,4,6}; (M) CquiOR10^{M3,4,5}; (N) CquiOR10^{M3,4}; (O) CquiOR10^{M3,5}; (P) CquiOR10^{M3,6,}; (Q) CquiOR10^{M3,7}; (R) CquiOR10^{M4,5}; (S) CquiOR10^{M4,6}; (T) CquiOR10^{M4,7}; (U) CquiOR10^{M5,6}; (V) CquiOR10^{M5,7}; (W) CquiOR10^{M6,7}; (X) CquiOR10^{M2,7}; and (Y) CquiOR10^{M2,5,6,7}. Lines were obtained with nonlinear fit. Bars represent SEM. n = 3, except for CquiOR10^{M2,5,6,7} and CquiOR10^{M2,7} (n = 4 and 7, respectively).

Figure 2. Concentration-response curves obtained with chimeric ORs stimulated with skatole and indole. (A) CquiOR10^{M2,5,6,7}_Outer; (B) CquiOR10^{M2,5,6,7}_Mid;Inner; (C) CquiOR10^{M2,5,6,7}_Inner; (D) CquiOR10^{M2,5,6,7}T78I; (E) CquiOR10^{M2,5,6,7}L73A; (F) CquiOR10^{M7}A73L; (G) CquiOR10A73L; (H) CquiOR2L74A. Lines were obtained with nonlinear fit. Bars represent SEM. The number of replicates (n) were 5, 5, 4, 3, 9, 7, 6, and 5, respectively.

Figure 3. Representative model of docked eugenol with MhraOR5 using RosettaLigand matches experimental structure. The model represented (forest green) is the lowest interface-energy pose from cluster 8. From cluster 8, the lowest RMSD pose was 0.554 Å relative to the experimental coordinates of eugenol from PDB ID: 7LID (magenta). Atoms that are not eugenol carbon atoms are color coded by atom type: carbon (grey), nitrogen (dark blue), and oxygen (red). The inset represents the clustering of the 10,000 lowest interface-energy models, plotted by Rosetta interface-energy (Rosetta Energy Units) and by RMSD (Å). The legend within the inset represents the top 10 most frequent clusters, sorted by largest to smallest (various color/shape

combinations) The grey circles represent structure models that were not clustered by the algorithm.

Figure 4. Representative models of docked skatole and indole in complex with CquiOR10 and CquiOR10A73L using RosettaLigand. Each model shown is the lowest interface-energy model from the 10 largest clusters of each docking study. CquiOR10 – skatole (forest green), CquiOR10 – indole (brown), CquiOR10A73L – skatole (light blue), and CquiOR10A73L – indole (purple). Atoms that are not indole/skatole carbon atoms are color coded by atom type: carbon (grey), nitrogen (dark blue), and oxygen (red). Ala-73 and Leu-73 indicated with space-filling representation.

Figure 5. An approximate 1 Å α-Carbon outward shift of Leu-73 (forest green) relative to Ala-73 (light blue). Models were superimposed from TM7b. Amino nitrogen (dark blue), and carboxyl oxygen (red).

1. R. Nasci, B. Miller, Culicine mosquitoes and the agents they transmit. *The biology of disease vectors*, 85-97 (1996).
2. T. G. Andreadis, The contribution of Culex pipiens complex mosquitoes to transmission and persistence of West Nile virus in North America. *J Am Mosq Control Assoc* **28**, 137-151 (2012).
3. Z. Syed, W. S. Leal, Acute olfactory response of Culex mosquitoes to a human- and bird-derived attractant. *Proc Natl Acad Sci U S A* **106**, 18803-18808 (2009).
4. O. Marinotti *et al.*, The genome of Anopheles darlingi, the main neotropical malaria vector. *Nucleic Acids Res* **41**, 7387-7400 (2013).
5. C. A. Hill *et al.*, G protein-coupled receptors in Anopheles gambiae. *Science* **298**, 176-178 (2002).
6. V. Nene *et al.*, Genome sequence of Aedes aegypti, a major arbovirus vector. *Science* **316**, 1718-1723 (2007).
7. J. Bohbot *et al.*, Molecular characterization of the Aedes aegypti odorant receptor gene family. *Insect Mol Biol* **16**, 525-537 (2007).
8. X. G. Chen *et al.*, Genome sequence of the Asian Tiger mosquito, Aedes albopictus, reveals insights into its biology, genetics, and evolution. *Proc Natl Acad Sci U S A* **112**, E5907-5915 (2015).
9. P. Arensburger *et al.*, Sequencing of Culex quinquefasciatus establishes a platform for mosquito comparative genomics. *Science* **330**, 86-88 (2010).
10. R. J. Pitts, D. C. Rinker, P. L. Jones, A. Rokas, L. J. Zwiebel, Transcriptome profiling of chemosensory appendages in the malaria vector Anopheles gambiae reveals tissue- and sex-specific signatures of odor coding. *BMC Genomics* **12**, 271 (2011).
11. B. J. Matthews *et al.*, Improved reference genome of Aedes aegypti informs arbovirus vector control. *Nature* **563**, 501-507 (2018).
12. W. S. Leal, Y. M. Choo, P. Xu, C. S. da Silva, C. Ueira-Vieira, Differential expression of olfactory genes in the southern house mosquito and insights into unique odorant receptor gene isoforms. *Proc Natl Acad Sci U S A* **110**, 18704-18709 (2013).
13. G. M. Pask, "Invertebrate Olfactory Receptors" in The Senses: A Comprehensive Reference (Second Edition), B. Fritzsch, Ed. (Elsevier, Oxford, 2020), doi.org/10.1016/B978-0-12-809324-5.23814-5, pp. 453-465.
14. W. S. Leal, Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* **58**, 373-391 (2013).

15. D. T. Hughes, J. Pelletier, C. W. Luetje, W. S. Leal, Odorant receptor from the southern house mosquito narrowly tuned to the oviposition attractant skatole. *J Chem Ecol* **36**, 797-800 (2010).
16. J. Pelletier, D. T. Hughes, C. W. Luetje, W. S. Leal, An odorant receptor from the southern house mosquito Culex pipiens quinquefasciatus sensitive to oviposition attractants. *PLoS One* **5**, e10090 (2010).
17. J. D. Bohbot *et al.*, Conservation of indole responsive odorant receptors in mosquitoes reveals an ancient olfactory trait. *Chem Senses* **36**, 149-160 (2011).
18. A. F. Carey, G. Wang, C. Y. Su, L. J. Zwiebel, J. R. Carlson, Odorant reception in the malaria mosquito Anopheles gambiae. *Nature* **464**, 66-71 (2010).
19. 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).
20. J. D. Bohbot, R. J. Pitts, The narrowing olfactory landscape of insect odorant receptors. *Frontiers in Ecology and Evolution* **3** (2015).
21. R. J. Pitts, R. M. Huff, S. J. Shih, J. D. Bohbot, Identification and functional characterization of olfactory indolergic receptors in Musca domestica. *Insect Biochem Mol Biol* **139**, 103653 (2021).
22. J. G. Millar, J. D. Chaney, M. S. Mulla, Identification of oviposition attractants for Culex quinquefasciatus from fermented Bermuda grass infusions. *J Am Mosquito Contr* **8**, 11-17 (1992).
23. A. Blackwell, A. J. Mordue, B. S. Hansson, L. J. Wadhams, J. A. Pickett, A behavioral and electrophysiological study of oviposition cues for Culex quinquefasciatus. *Physiol Entomol* **18**, 343-348 (1993).
24. L. E. G. Mboera, W. Takken, K. Y. Mdira, G. J. Chuwa, J. A. Pickett, Oviposition and behavioral responses of Culex quinquefasciatus to skatole and synthetic oviposition pheromone in Tanzania. *Journal of Chemical Ecology* **26**, 1193-1203 (2000).
25. G. Fenaroli, *Fenaroli's Handbook of flavor ingredients : adapted from the Italian language works of Giovanni Fenaroli* (CRC Press, Cleveland, ed. 2d, 1975).
26. Y. M. Choo *et al.*, Reverse chemical ecology approach for the identification of an oviposition attractant for Culex quinquefasciatus. *Proc Natl Acad Sci U S A* **in press** (2018).
27. S. M. Reynolds, L. Kall, M. E. Riffle, J. A. Bilmes, W. S. Noble, Transmembrane topology and signal peptide prediction using dynamic bayesian networks. *PLoS Comput Biol* **4**, e1000213 (2008).
28. M. Baek *et al.*, Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876 (2021).
29. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
30. J. A. Butterwick *et al.*, Cryo-EM structure of the insect olfactory receptor Orco. *Nature* **560**, 447-452 (2018).
31. H. Viklund, A. Elofsson, OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics* **24**, 1662-1668 (2008).
32. D. T. Hughes *et al.*, Functional and Nonfunctional Forms of CquiOR91, an Odorant Selectivity Subunit of Culex quinquefasciatus. *Chem Senses* **42**, 333-341 (2017).
33. C. W. Luetje, A. S. Nichols, A. Castro, B. L. Sherman, "Functional Assay of Mammalian and Insect Olfactory Receptors Using Xenopus Oocytes" in Olfactory Receptors: Methods and Protocols, C. J. Crasto, Ed. (Humana Press, Totowa, NJ, 2013), 10.1007/978-1-62703-377-0_14, pp. 187-202.
34. J. Del Marmol, M. A. Yedlin, V. Ruta, The structural basis of odorant recognition in insect olfactory receptors. *Nature* **597**, 126-131 (2021).

35. S. Guo, J. Kim, Dissecting the molecular mechanism of drosophila odorant receptors through activity modeling and comparative analysis. *Proteins: Structure, Function, and Bioinformatics* **78**, 381-399 (2010).
36. S. DeLuca, K. Khar, J. Meiler, Fully Flexible Docking of Medium Sized Ligand Libraries with RosettaLigand. *PLoS One* **10**, e0132508 (2015).
37. I. W. Davis, D. Baker, RosettaLigand docking with full ligand and receptor flexibility. *J Mol Biol* **385**, 381-392 (2009).
38. M. F. Adasme *et al.*, PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Research* **49**, W530-W534 (2021).
39. P. Xu *et al.*, Odorant inhibition in mosquito olfaction. *iScience* **19**, 25-38 (2019).
40. P. C. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson, M. T. Stahl, Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J Chem Inf Model* **50**, 572-584 (2010).
41. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).

Fig. 1

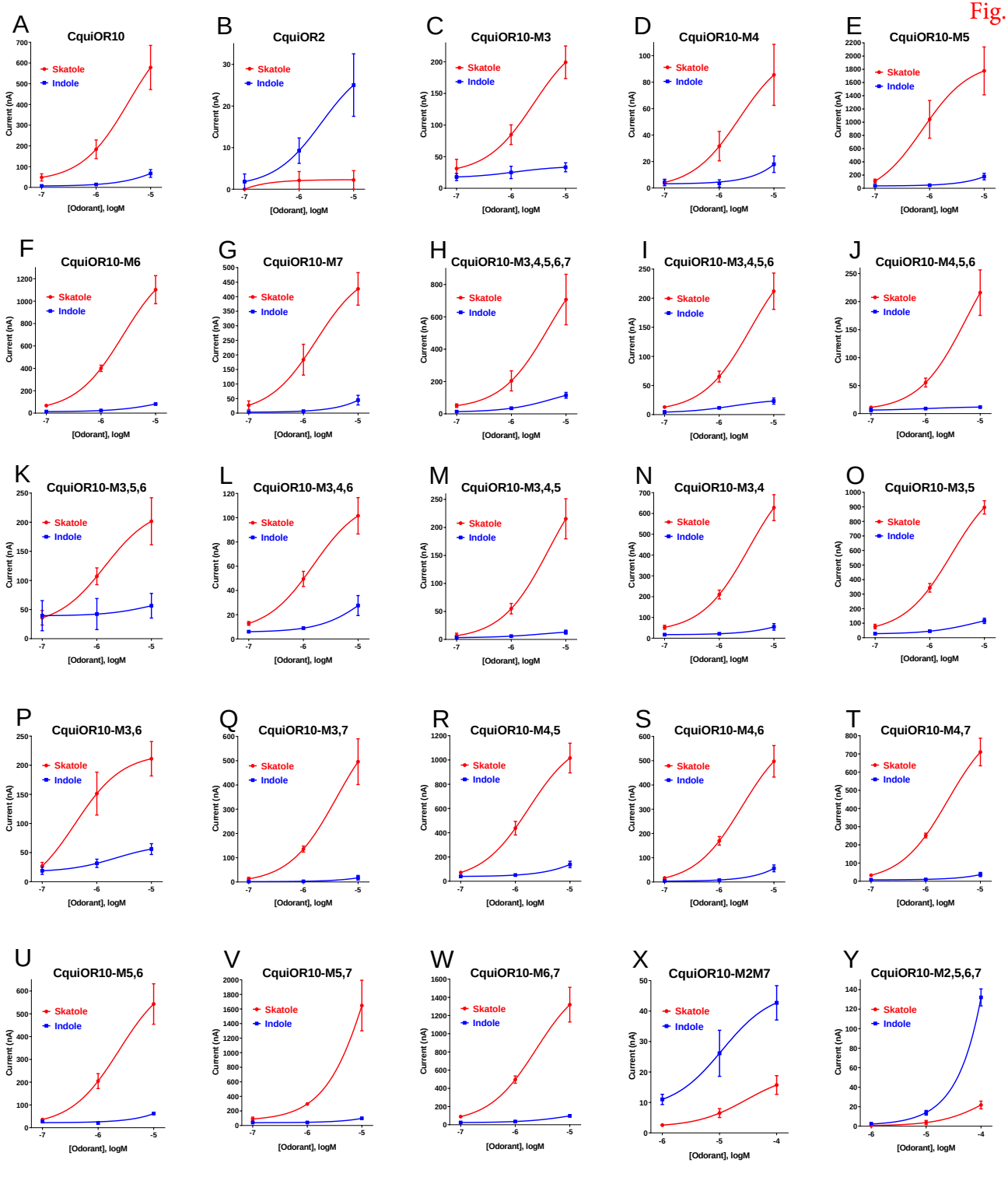

A
CquiOR10
B
CquiOR2
C
CquiOR10-M3
D
CquiOR10-M4
E
CquiOR10-M5
F
CquiOR10-M6
G
CquiOR10-M7
H
CquiOR10-M3,4,5,6,7
I
CquiOR10-M3,4,5,6
J
CquiOR10-M4,5,6
K
CquiOR10-M3,5,6
L
CquiOR10-M3,4,6
M
CquiOR10-M3,4,5
N
CquiOR10-M3,4
O
CquiOR10-M3,5
P
CquiOR10-M3,6
Q
CquiOR10-M3,7
R
CquiOR10-M4,5
S
CquiOR10-M4,6
T
CquiOR10-M4,7
U
CquiOR10-M5,6
V
CquiOR10-M5,7
W
CquiOR10-M6,7
X
CquiOR10-M2M7
Y
CquiOR10-M2,5,6,7
Skatole
Indole
Current (nA)
[Odorant], logM

Fig. 2

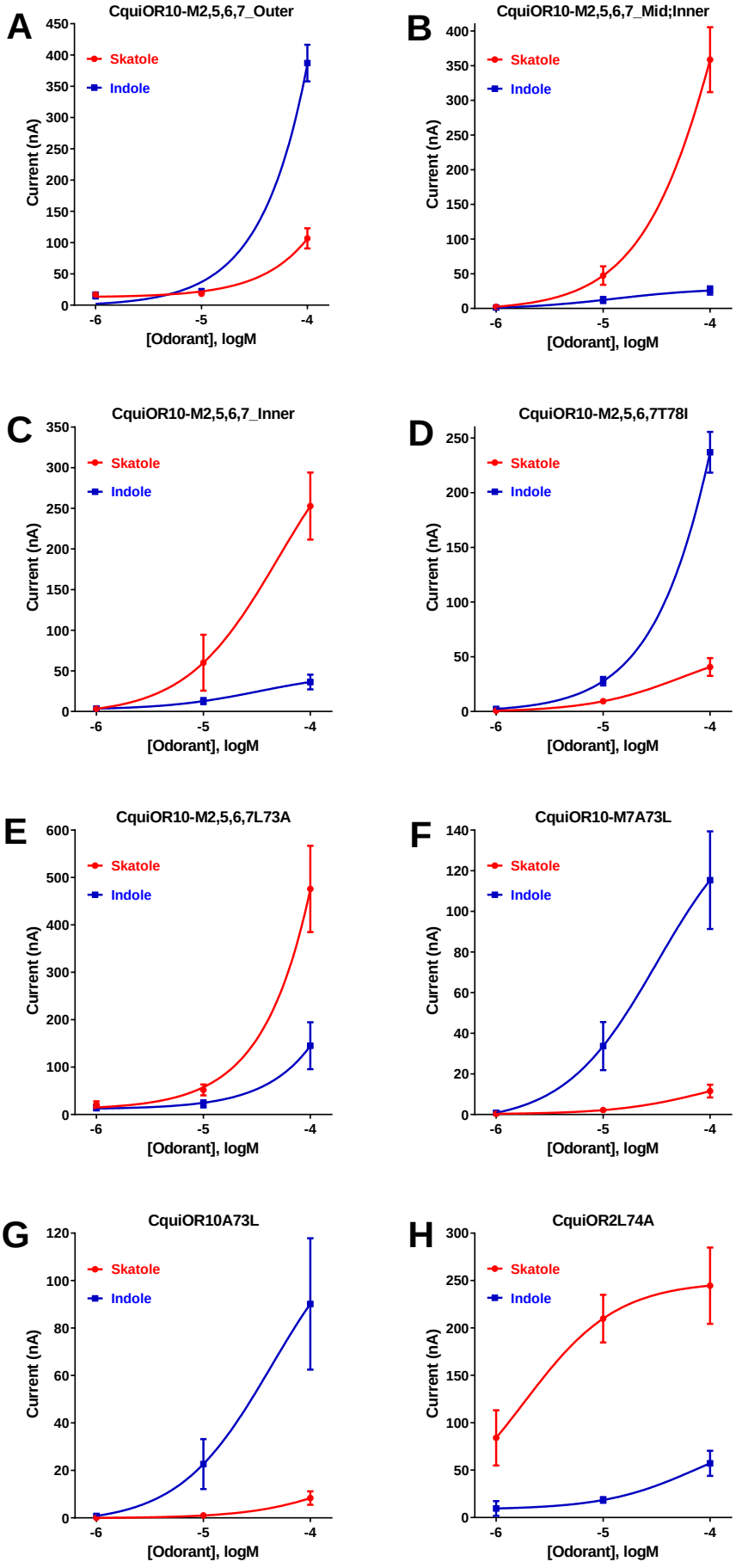

A
CquiOR10-M2,5,6,7_Outer
Skatole
Indole
Current (nA)
[Odorant], logM
B
CquiOR10-M2,5,6,7_Mid;Inner
Skatole
Indole
Current (nA)
[Odorant], logM
C
CquiOR10-M2,5,6,7_Inner
Skatole
Indole
Current (nA)
[Odorant], logM
D
CquiOR10-M2,5,6,7T78I
Skatole
Indole
Current (nA)
[Odorant], logM
E
CquiOR10-M2,5,6,7L73A
Skatole
Indole
Current (nA)
[Odorant], logM
F
CquiOR10-M7A73L
Skatole
Indole
Current (nA)
[Odorant], logM
G
CquiOR10A73L
Skatole
Indole
Current (nA)
[Odorant], logM
H
CquiOR2L74A
Skatole
Indole
Current (nA)
[Odorant], logM

Fig. 3

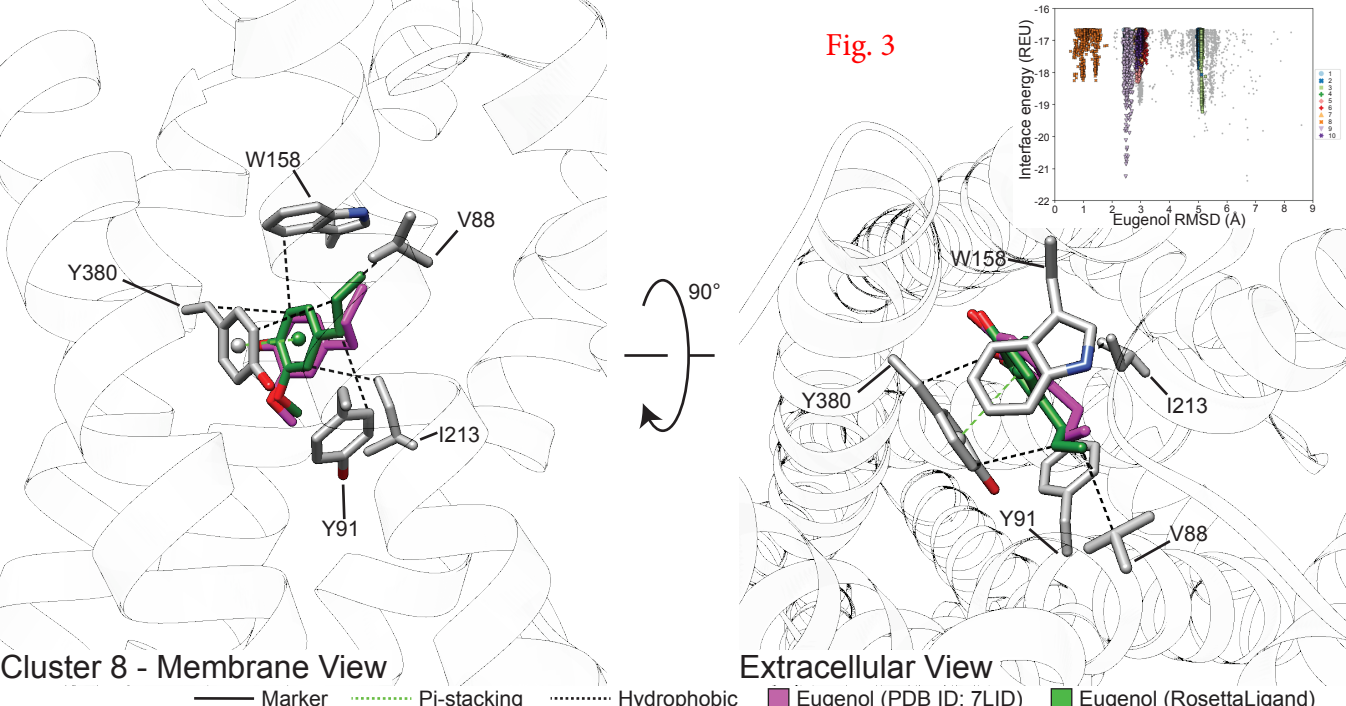
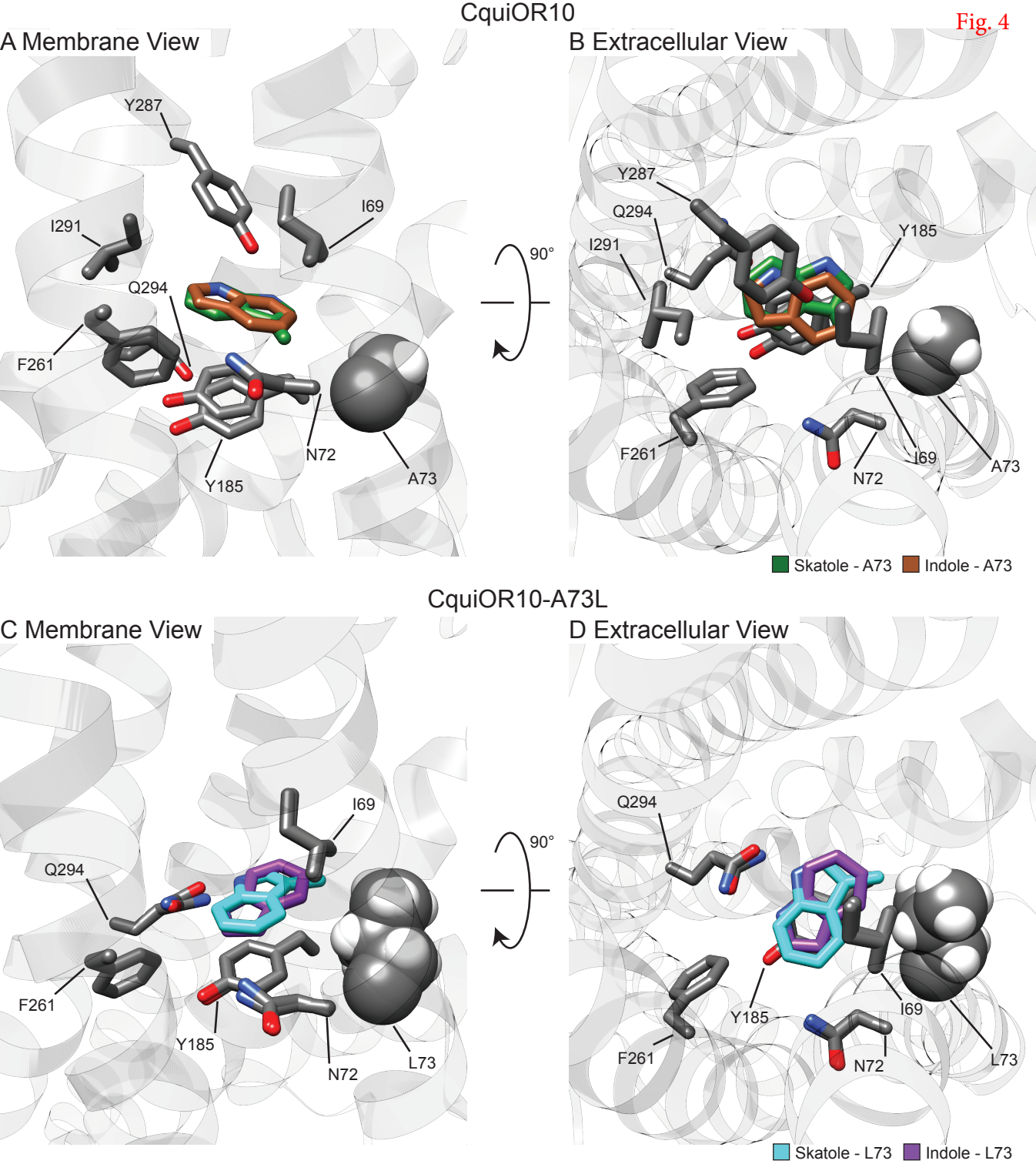

W158
V88
Y380
I213
Y91
90°
W158
Y380
I213
Y91
V88
Interface energy (REU)
Eugenol RMSD (Å)
-16
-17
-18
-19
-20
-21
-22
0
1
2
3
4
5
6
7
8
9
1
2
3
4
5
6
7
8
9
10
Cluster 8 - Membrane View
Extracellular View
Marker
Pi-stacking
Hydrophobic
Eugenol (PDB ID: 7LID)
Eugenol (RosettaLigand)

Fig. 4


CquiOR10
A Membrane View
Y287
I69
I291
Q294
F261
N72
Y185
A73
90°
B Extracellular View
Y287
Q294
Y185
I291
F261
N72
I69
A73
Skatole - A73
Indole - A73
CquiOR10-A73L
C Membrane View
I69
Q294
F261
Y185
N72
L73
90°
D Extracellular View
Q294
Y185
I69
F261
N72
L73
Skatole - L73
Indole - L73

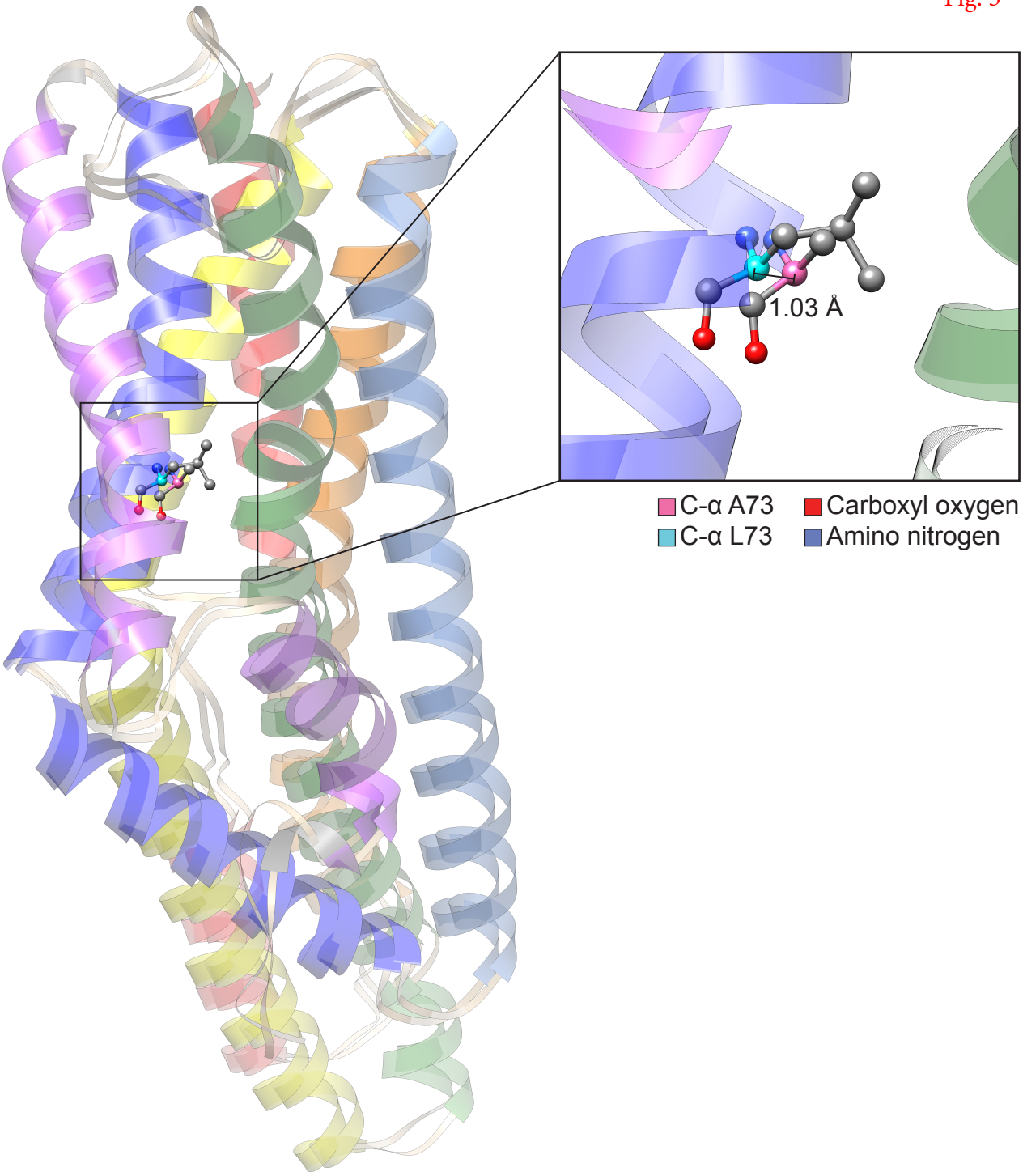



Fig. 5