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

Flávia P. Franco[1], Pingxi Xu[1], Brandon J. Harris[2], Vladimir Yarov-Yarovoy[2,3], and Walter S. Leal[1]*

[1]Department of Molecular and Cellular Biology, [2]Department of Physiology and Membrane Biology, and [3]Department of Anesthesiology and Pain Medicine, University of California, Davis, Davis, CA 95616 USA

*Walter S. Leal, wsleal@ucdavis.edu

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

Dedicated to the late Professor James H. Tumlinson III

ABSTRACT

The southern house mosquito, *Culex quinquefasciatus,* utilizes two odorant receptors, CquiOR10 and CquiOR2, narrowly tuned to oviposition attractants and well conserved among mosquito species. They detect skatole and indole, respectively, with reciprocal specificity. 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. 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 more robust response than the wildtype receptor CquiOR10. On the other hand, Leu to Gly mutation of the indole receptor CquiOR2 reverted the specificity to skatole. Lastly, CquiOR10A73L, CquiOR2, and CquiOR2L74I were insensitive to 3-ethylindole, whereas CquiOR2L74A and CquiOR2L74G gained activity. Additionally, CquiOR10A73G gave more robust responses to 3-ethylindole than CquiOR10. Thus, we suggest the specificity of these receptors is mediated by a single amino acid substitution, leading to finely tuned volumetric space to accommodate specific oviposition attractants.

INTRODUCTION

Insects perceive the world with a sophisticated olfactory system essential for survival and reproduction. Their antennae are biosensors *par excellence,* allowing detection of a plethora of compounds, some with extraordinary sensitivity and selectivity (Kaissling, 2014). The insect olfactory system is comprised mainly of odorant-binding proteins, odorant-degrading enzymes, ionotropic receptors, and the ultimate gatekeepers of selectivity (Leal, 2013, 2016; Leal, 2020) – the odorant receptors (ORs) (Clyne et al., 1999; Gao & Chess, 1999; Vosshall et al., 1999). The ORs are the binding units in functional heteromeric cation channels (Neuhaus et al., 2005) formed with an odorant receptor coreceptor (Orco) (Larsson et al., 2004). Unlike mammalian olfactory receptors, insect ORs and Orco have inverse topologies compared to G-protein coupled receptors (GPCRs), with a cytosolic N-terminus and an extracellular C-terminus (Benton et al., 2006; Lundin et al., 2007). One of the major breakthroughs in the fields of insect olfaction in the last two decades since the discovery of ORs (Leal, 2020) was the determination of the cryo-electron microscopy (cryo-EM) structure of the Orco homomer from the parasitic fig wasp*, Apocrypta bakeri,* AbakOrco (Butterwick et al., 2018). Subsequently, the structure for a promiscuous OR from the evolutionarily primitive (Apterygota, wingless) jumping bristletail, *Machilis harabi*, MhOR5, was solved (Del Marmol et al., 2021). Although structures of ORs from winged insects (Pterygota) have not been solved to date, amino acid residues critical for OR specificity have been reported (Auer et al., 2020; Cao et al., 2021; Hughes et al., 2014; Leary et al., 2012; Pellegrino et al., 2011; Yang et al., 2017; Yuvaraj et al., 2021). These studies focused on one-way alteration of specificity, but did not examine how an insect detects two odorants with reverse specicity.

The genome of the southern house mosquito, *Culex quinquefasciatus* (Arensburger et al., 2010)*,* has the most extensive repertoire of *OR* genes (Leal et al., 2013) among mosquito species. However, the number of odorants in the

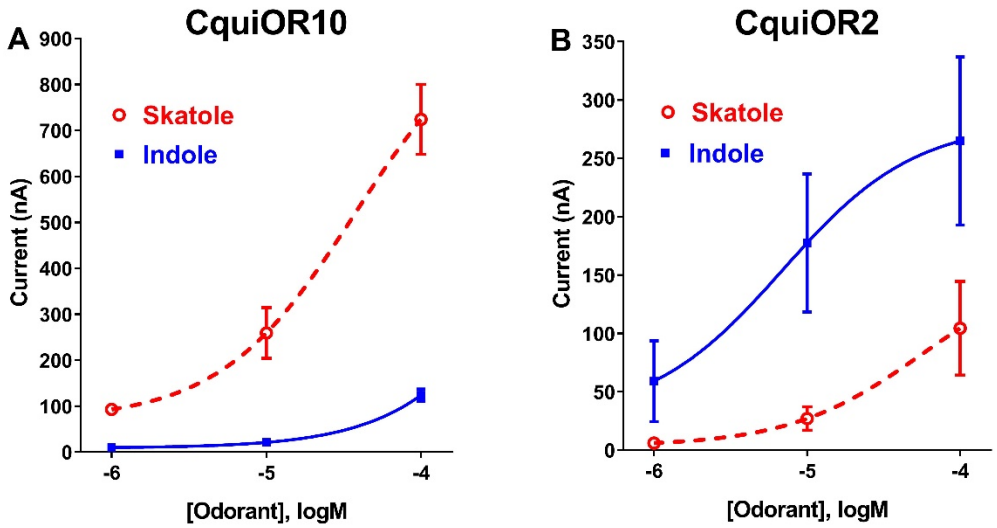
environment far exceeds the number of OR genes, even in *Cx. quinquefasciatus*. Therefore, it is unsurprising that many insect odorant receptors are promiscuous (Leal, 2013; Pask, 2020). However, ORs detecting behaviorally critical compounds may be narrowly tuned to pheromones and other semiochemicals (Hughes et al., 2010; Nakagawa et al., 2005; Stensmyr et al., 2012).

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 (Hughes et al., 2010; Pelletier et al., 2010), the yellow fever mosquito (Bohbot et al., 2011), and the malaria mosquito (Carey et al., 2010; Wang et al., 2010). Recently, these so-called indolergic receptors (Bohbot & Pitts, 2015) have also been found in the housefly, *Musca domestica* (Pitts et al., 2021). Remarkably, in all these species, OR10s are narrowly tuned to skatole and respond with lower sensitivity to indole. In contrast, indole is the most potent 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. Here, we refer to these receptors from *Cx. quinquefasciatus* as CquiOR10 (VectorBase and GenBank I.D.s, CPIJ002479 and GU945397, respectively) and CquiOR2 (CPIJ014392, GU945396.1). These oviposition-attractant-detecting receptors in *Cx. quinquefasciatus*, CquiOR10 and CquiOR2 provide a suitable model to identify specificity determinants and examine the long-standing question of reverse specificity.

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 (Choo et al., 2018)), we replaced the predicted (Reynolds et al., 2008) transmembrane

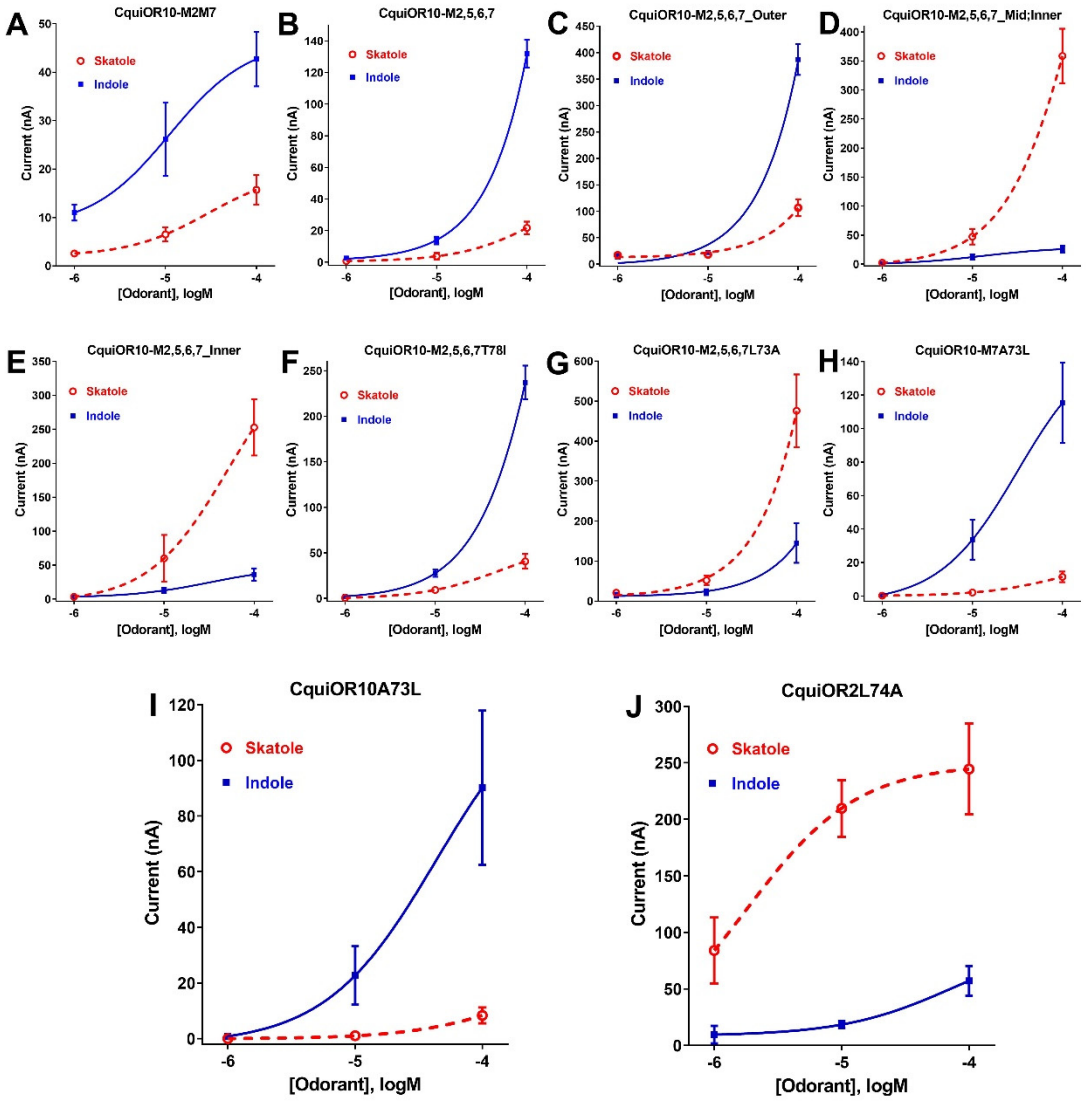
101 domains from CquiOR2 into CquiOR10 and measured the specificity of the
102 chimeric receptors. With this approach, we identified TM2 as a specificity
103 determinant. Next, we tested mutations of chimeric CquiOR10 and identified a
104 single amino acid residue in TM2 (Ala-73) that determines the receptor's
105 specificity. We then directly mutated the wildtype receptor and observed that
106 CquiOR10A73L is specific to indole as CquiOR2. Additionally, CquiOR2L74A
107 emulated the response profile of CquiOR10. To better understand the structural
108 basis of this single-point specificity determinant, we generated structural models
109 of CquiOR10 and CquiOR10A73L using RoseTTAFold (Baek et al., 2021) and
110 AlphaFold (Jumper et al., 2021). We identified the binding poses for skatole and
111 indole using RosettaLigand molecular docking and observed a finely tuned
112 volumetric space to accommodate specific oviposition attractants.
113
114
115 **RESULTS**
116
117 **Chimeric OR with reversed specificity.** We envisioned that studying a pair of
118 odorant receptors with reverse specificities, like CquiOR10 and CquiOR2, could
119 lead us to specificity determinants, which may have implications for the evolution
120 of other insect behaviors. CquiOR10 is activated by the oviposition attractant
121 skatole (Blackwell et al., 1993; Mboera et al., 2000; Millar et al., 1992) with high
122 specificity (*Figure 1A*), whereas CquiOR2 is specific to indole (*Figure 1B*). Our
123 approach was designed to swap TM domains using the more sensitive receptor,
124 CquiOR10, as the acceptor. Specifically, we generated chimeric receptors by
125 replacing CquiOR10 TM domains with related domains from CquiOR2 (*Figure*
126 *1—figure supplement 1; see Appendix 1*). We referred to these chimeric
127 receptors as CquiOR10^{Mx}, where Mx refers to TMx from CquiOR2. Then, we
128 performed functional assays of these CquiOR10^{Mx} receptors using the *Xenopus*
129 oocyte recording system. This long-term project could require as many as 127

130 possible chimeric receptors. We optimized the workflow to reduce the number of
131 tested mutants before finding a chimeric receptor with reversed specificity.
132



133
134 **Figure 1.** Concentration-response analysis for activation of wildtype ORs by
135 skatole and indole. (A) CquiOR10 and (B) CquiOR2. Lines were obtained with
136 nonlinear fit. Bars represent SEM. *n* = 4-5.
137 **Figure supplement 1.** Sequence alignment of CquiOR10 and CquiOR2.
138 **Figure supplement 2.** Concentration-response analysis for wildtype and
139 chimeric ORs.
140 **Figure supplement 3.** Overlap of CquiOR10 predicted TM domains and
141 AbakOrco cryo-EM structure.
142
143 We tested 36 chimeric receptors (see *Appendix 1* and *Figure 1—figure*
144 *supplement 2*). Fourteen chimeric receptors did not respond to skatole or indole,
145 and 21 receptors retained the specificity to skatole (*Extended—Table 1* and
146 *Figure 1—figure supplement 2).* Lastly, CquiOR10^{M2,7}/CquiOrco-expressing
147 oocytes responded to both skatole and indole with a reverse profile (*Figure 2A*).

148 This dataset shows that CquiOR10^{M2,7} emulated the profile of the indole receptor
149 CquiOR2 (*Figure 1B*).
150



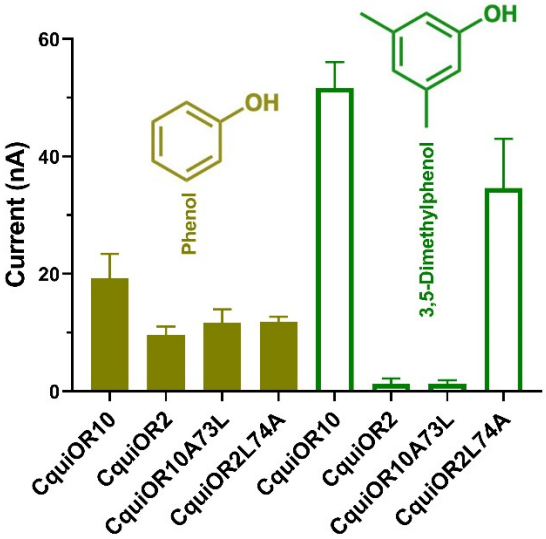
151

152 **Figure 2.** Concentration-response curves obtained with chimeric ORs stimulated
153 with skatole and indole. (A) CquiOR10^{M2,7}; (B) CquiOR10^{M2,5,6,7}; (C)
154 CquiOR10^{M2,5,6,7}_Outer; (D) CquiOR10^{M2,5,6,7}_Mid;Inner; (E)
155 CquiOR10^{M2,5,6,7}_Inner; (F) CquiOR10^{M2,5,6,7}T78I; (G) CquiOR10^{M2,5,6,7}L73A; (H)
156 CquiOR10^{M7}A73L; (I) CquiOR10A73L; (J) CquiOR2L74A. Lines were obtained

157 with nonlinear fit. Bars represent SEM. The number of replicates (*n*) were 7, 4, 5,
158 5, 4, 3, 9, 7, 6, and 5, respectively.
159 **Figure supplement 1.** Partial sequences of CquiOR10 and chimeric ORs
160 highlighting the transmembrane domain-2 (TM2).
161 **Figure supplement 2.** Schematic view of the workflow.
162
163 **A single-point mutation that reverses the specificity of the skatole and**
164 **indole receptors.** We proceeded to identify the amino acid residues in the
165 swapped domains of CquiOR10^{M2,7}, directly affecting the specificity of the
166 chimeric and wildtype receptors. Given the observation that, by and large,
167 chimeric receptors with TM5 and TM6 from CquiOR2 gave stronger responses
168 (*Figure 1—figure supplement 2*), we asked whether CquiOR10^{M2,7} responses
169 with these two additional TM domains swapped would give more robust
170 responses while keeping the same specificity to indole.
171 CquiOR10^{M2,5,6,7}/CquiOrco-expressing oocytes were indeed more sensitive while
172 maintaining the selectivity to indole (*Figure 2B*). We then used CquiOR10^{M2,5,6,7}
173 and designed various mutants to rescue single or multiple residues in TM2 at a
174 time (*Figure 2—figure supplement 1*). We focused on TM2 because swapping
175 TM7 did not affect the specificity of the receptor (*Figure 1—figure supplement*
176 *2E*). We aimed to identify mutations in the chimeric receptor that reverts the
177 specificity to skatole as in the wildtype receptor. We divided TM2 into outer,
178 middle, and inner segments based on the topology predicted by OCTOPUS
179 (Viklund & Elofsson, 2008). It has been postulated that the extracellular halves of
180 TM domains form an odorant-binding pocket (Guo & Kim, 2010), thus we first
181 examined a mutant (CquiOR10^{M2,5,6,7}_Outer) having the residues in the outer
182 segment at 59, 60, 63, and 64 restored to Glu, Val, Asn, and Ala, respectively, as
183 in the wildtype receptor (*Figure 2—figure supplement 1*).
184 CquiOR10^{M2,5,6,7}_Outer/CquiOrco-expressing oocytes retained the specificity of
185 CquiOR10^{M2,5,6,7} with a more robust response to indole than skatole (*Figure 2C*).
186 These findings indicated that the residues in the outer segment of TM2 are not

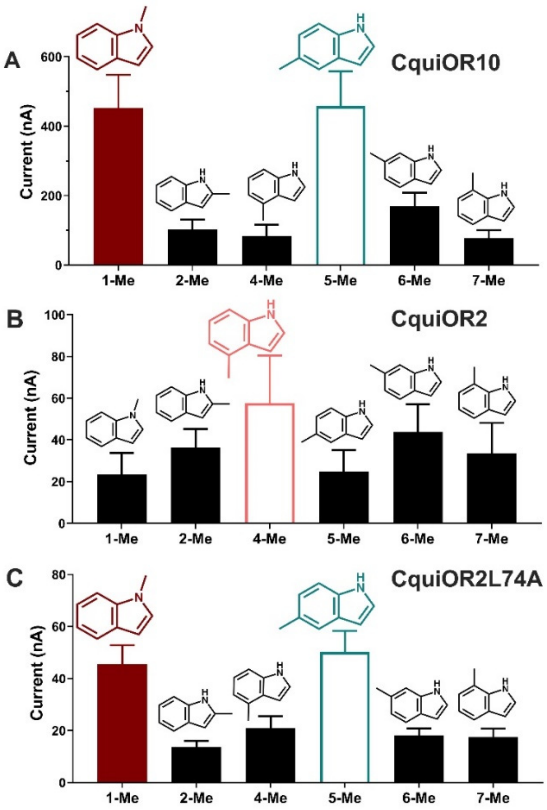
187 specificity determinants. After that, we tested a chimeric receptor with the
188 residues in the middle and inner part of the TM2 domain rescued to match those
189 in the wildtype receptor (*Figure 2—figure supplement 1*).
190 CquiOR10^{M2,5,6,7}_Mid;Inner restored the skatole-specific profile of CquiOR10
191 (*Figure 2D*), thus suggesting that amino acid residues in these segments are
192 specificity determinants. Then, we tested CquiOR10^{M2,5,6,7}_Inner
193 (=CquiOR10^{M2,5,6,7}L73A;T78I), which had only residues at 73 and 78 rescued to
194 the wildtype Ala and Ile, respectively. Oocytes co-expressing
195 CquiOR10^{M2,5,6,7}_Inner and CquiOrco reverted the specificity to skatole (*Figure*
196 *2E*). As a result, we concluded that residues in the predicted inner part of TM2
197 are critical for the chimeric receptor's specificity. We further probed the chimeric
198 receptor CquiOR10^{M2,5,6,7} with single-point mutations to identify the residue(s)
199 determining specificity. CquiOR10^{M2,5,6,7}T78I/CquiOrco-expressing oocytes
200 showed the same specificity as the chimeric receptor CquiOR10^{M2,5,6,7} (*Figure*
201 *2F*). Specifically, CquiOR10^{M2,5,6,7}T78I gave a more robust response to indole
202 than skatole, suggesting that rescuing the residue at 78 did not affect
203 CquiOR10^{M2,5,6,7} specificity. By contrast, CquiOR10^{M2,5,6,7}L73A/CquiOrco-
204 expressing oocytes reverted the specificity to skatole (*Figure 2G*), thus behaving
205 like the wildtype receptor CquiOR10 (*Figure 1A*). To further examine the role of
206 Ala-73 as a specificity determinant residue, we obtained a single-point mutation
207 of CquiOR10^{M7}, which is specific to skatole (*Figure 1—figure supplement 2E*).
208 The responses recorded from CquiOR10^{M7}A73L/CquiOrco-expressing oocytes
209 showed a reverse, indole-specific profile (*Figure 2H*), similar to the CquiOR2
210 profile (*Figure 1B*). Having identified a single amino acid residue in the chimeric
211 receptor that switches the skatole/indole specificity, we tested the effect of single
212 point mutation on the specificity of the wildtype receptor CquiOR10 (EC50:
213 skatole, 3.6 μM; indole 29.9 μM). CquiOR10A73L showed a reverse specificity,
214 with dose-dependent responses to indole (*Figure 2I*) (EC50: indole, 3.4 μM;
215 skatole 53.7 μM). Collectively, these findings suggest that a single amino acid
216 residue in CquiOR10 determines the specificity of this receptor. Additionally, we

217 obtained an equivalent single-point mutation in the indole-specific CquiOR2
218 (*Figure 1B*) (EC50: indole, 7.7 μM; skatole 16.4 μM). Thus,
219 CquiOR2L74A/CquiOrco-expressing oocytes gave robust and specific responses
220 to skatole (Figure 2J) (EC50: skatole, 8.5 μM; indole 27.6 μM).
221
222 Lastly, we recorded the response of these ORs to other phenolic ligands that
223 activate indolic receptors, albeit generating small currents. While CquiOR10 and
224 CquiOR2 responded to phenol, 3,5-dimethylphenol activated only CquiOR10
225 (*Figure 3*). A single-point mutation in CquiOR10 rendered the chimeric receptor
226 insensitive to 3,5-dimethylphenol. By contrast, an equivalent mutation in
227 CquiOR2 recapitulated the profile of CquiOR10 (*Figure 3*).



228
229 **Figure 3.** Quantification of wildtype and chimeric receptors to phenol and 2,3-
230 dimethylphenol. Each receptor was co-expressed with CquiORco in *Xenopus*
231 oocytes and stimulated with the phenolic compounds at 1 mM. Bars represent
232 SEM. *n* = 3-5.
233
234 Additionally, we recorded responses elicited by methylindoles. Specifically, we
235 challenged oocytes with 1-methylindole, 2-methylindole, 4-methylindole, 5-
236 methylindole, 6-methylindole, and 7-methylindole. In these analyses, we did not
237 stimulate the oocyte preparations with 3-methylindole to avoid possible

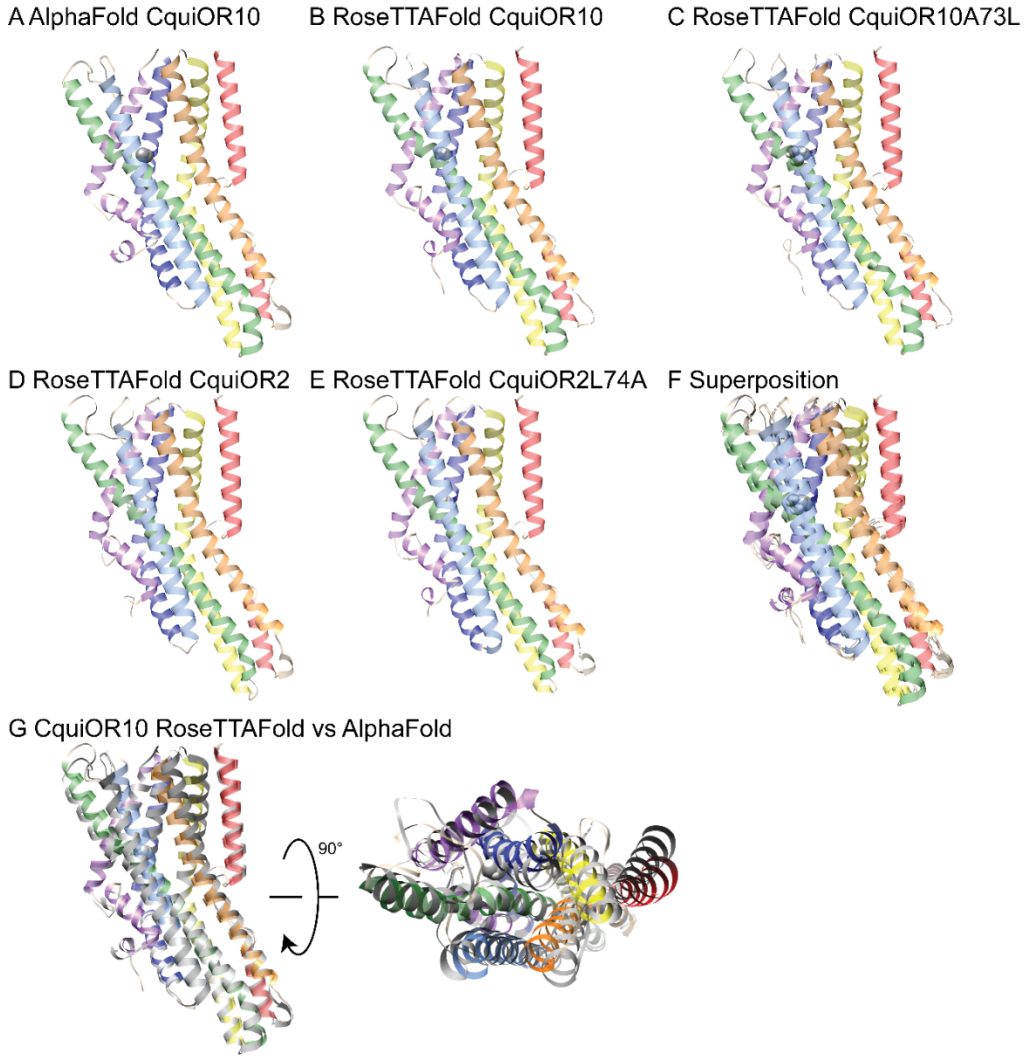
238 desensitization. CquiOR10/CquiOrco-expressing oocytes elicited stronger
239 responses when challenged with 1-methylindole and 5-methylindole than when
240 stimulated with the other methylindoles (*Figure 4A*). By contrast,
241 CquiOR2/CquiOrco-expressing oocytes elicited similarly lower responses when
242 stimulated with methylindoles (*Figure 4B*). CquiOR2L74A with a single-point
243 mutation to mimic OR10 receptor recapitulated CquiOR10 response profile
244 (*Figure 4C*). These data suggest that a single amino acid residue determines a
245 receptor's specificity toward ligands eliciting robust or small responses.
246



247

248 **Figure 4.** Quantification of responses to methylindoles presented at 100 μM. (A)
249 CquiOR10/CquiOrco-, (B) CquiOR2/CquiOrco-, and (C)-CquiOR2L74A-
250 expressing oocytes. Bars represent SEM. *n* = 9-11.
251
252 As summarized in a graphical representation (*Figure 2—figure supplement 2*),
253 these findings demonstrate that these two mosquito odorant receptors,

254 CquiOR10 and CquiOR2, have reciprocal specificity mediated by a single amino
255 acid residue, Ala-73 and Leu-74, respectively.
256
257 **CquiOR10 computational modeling suggests space-filling constraints for**
258 **indole-based odorants around A73.** To structurally hypothesize the above-
259 described reciprocal specificity, we generated structural models of CquiOr10,
260 CquiOR2, CquiOR10A73L and CquiOR2L74A using RoseTTAFold (Baek et al.,
261 2021) and a structural model of CquiOR10 using AlphaFold (Jumper et al., 2021)
262 (*Figure 5*).

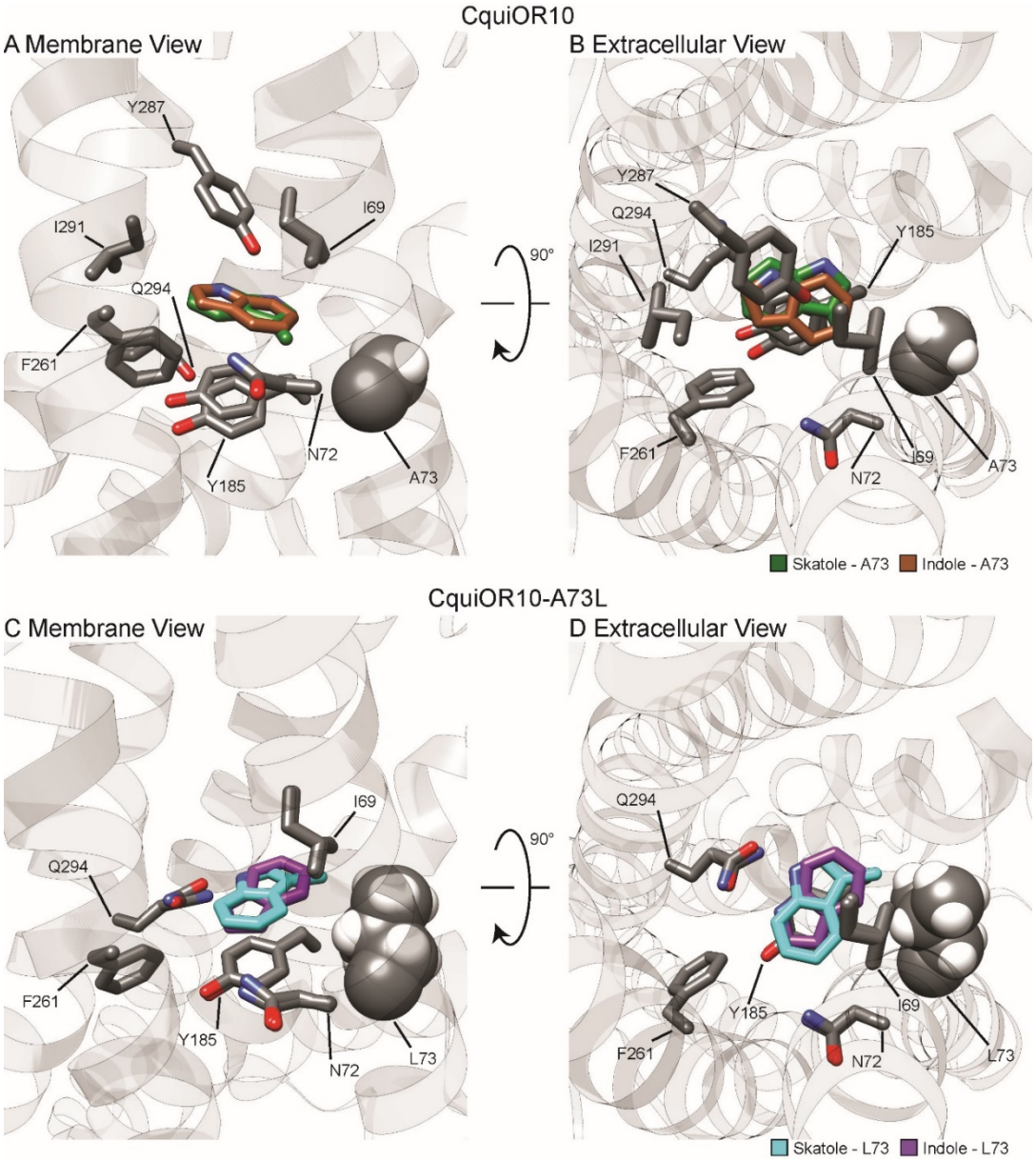


263
264 **Figure 5.** Structural models of CquiOR10 (A,B), CquiOR10A73L
265 (C),CquiOR2(D), and CquiOR2L74A(E) with AlphaFold (A) and RoseTTAFold (B-

266 E) structure prediction methods. Superposition of all RoseTTAFold models (F)
267 resulted in transmembrane helix root mean square deviation (RMSD) of 0.8 Å
268 when aligned with RoseTTAFold CquiOR10. (G) The transmembrane helix
269 RMSD of CquiOR10 RoseTTAFold (rainbow) vs AlphaFold (grey) was 1.7 Å.
270 Loops were not included in RMSD calculation due to inherent flexibility during
271 structure prediction.
272 **Figure supplement 1.** Representative model of docked eugenol with MhraOR5.
273 **Figure supplement 2.** Additional clusters of eugenol docked to MhraOR5.
274
275 CquiOR10 models of one homotetramer subunit were generated with AlphaFold
276 and RoseTTAFold, each producing five models. RoseTTAFold models of
277 CquiOR10, CquiOR10A73L, CquiOR2, and CquiOR2A73L produced a
278 transmembrane helix RMSD between α-Carbon atoms less than 1 Å across all
279 twenty models (five models per odorant receptor); this suggests that the
280 homologous CquiOR10 and CquiOR2 are structurally similar and that single-
281 point mutations should not cause great structural deviation from wild type.
282 Comparing CquiOR10 models, RoseTTAFold and AlphaFold produced a
283 transmembrane helix RMSD of 1.7 Å, with the transmembrane helices at the
284 extracellular membrane face having the largest structural deviation. Considering
285 there is little structural knowledge of insect odorant receptors, their binding
286 mecanisms, and their conformational changes, we suspected that homologous
287 odorant receptors would have similar binding modes. Using transmembrane 7b
288 (TM7b), we superimposed the top ranking CquiOR10 RoseTTAFold and
289 AlphaFold models with an experimentally resolved structure, *M. hrabei*
290 (MhraOR5) in complex with eugenol (PDB ID: 7LID, (Del Marmol et al., 2021)) to
291 identify which of our models resembled an odorant-bound conformation. With this
292 selection criteria, we proceeded with RoseTTAFold models of the odorant
293 receptors for Rosetta-based small-molecule docking method RosettaLigand
294 (Davis & Baker, 2009; DeLuca et al., 2015), as the structural similarity around the
295 hypothesized binding pocket was greater than the AlphaFold models of the

296 odorant receptors when compared to the MhraOR5 structure. We chose to select
297 conformationally similar models over modeling and docking an apo structure into
298 a bound conformation because it is a more cautious approach when there is little
299 structural information. We perceived modeling an apo structure into a bound
300 conformation to potentially yield more biologically implausible conformations than
301 docking of a structurally comparative model.
302 RosettaLigand (37, 38) is a competitive protein docking method (Smith & Meiler,
303 2020) that has been used previously to predict the binding position of ligands
304 within protein pores (Nguyen et al., 2017; Pressly et al., 2022). To verify that
305 RosettaLigand could effectively sample odorants in receptors homologous to
306 CquiOR10 and CquiOR2, we used the structure of eugenol in complex with the
307 insect odorant receptor OR5 from MhraOR5 (PDB ID:7LID) as a control (Del
308 Marmol et al., 2021). Rosetta protein-ligand docking employs energy-based
309 analyses, such as the interface energy between protein and ligand, to select the
310 representative models (Bidula et al., 2022). With this selection method, the
311 RMSD of our MhraOR5-eugenol models relative to the experimental structure
312 ranged from 0.5 – 5.0 Å (*Extended—Table 2*). RMSD values equal or below 2.0
313 Å are considered an appropriate range for validation (Park et al., 2021). After
314 using hdbscan cluster analysis (McInnes et al., 2017) to group structurally similar
315 models, the largest cluster had a RMSD of 0.75 Å while the lowest interface-
316 energy model had a RMSD of 2.4 Å. Collectively, these data demonstrate that
317 RosettaLigand paired with the hdbscan clustering method can recapitulate the
318 MhraOR5 structure and blindly select a near-native model, thus is suitable for
319 structural predictions of odorants with CquiOR10A73L and CquiOR2L74A (*Figure*
320 *5—figure supplements 1 and 2*).
321
322 For each receptor-ligand complex (CquiOR10-skatole, CquiOR10A73L-skatole,
323 CquiOR10-indole, and CquiOR10A73L-indole), we generated 100,000 docking
324 models using RosettaLigand, clustered the 10,000 lowest interface-energy
325 models, and selected the lowest interface-energy model from the 10 largest

clusters, resulting in ten models per receptor-ligand complex from which to draw structural hypotheses (*Extended—Tables 3 and 4*). Our modeling suggests that both indole and skatole can readily reorient themselves in a similar pore depth near residue 73, regardless of mutant or wildtype receptor. Comparing the lowest interface-scoring model from each receptor-ligand complex, indole and skatole are positioned in the membrane-embedded pore, flanked by transmembrane helices S2, S4, S5, and S6, and show positional overlap in both and CquiOR10-A73L CquiOR10 (Figure 6; *Figure 6—figure supplemens 1 and 2)*.



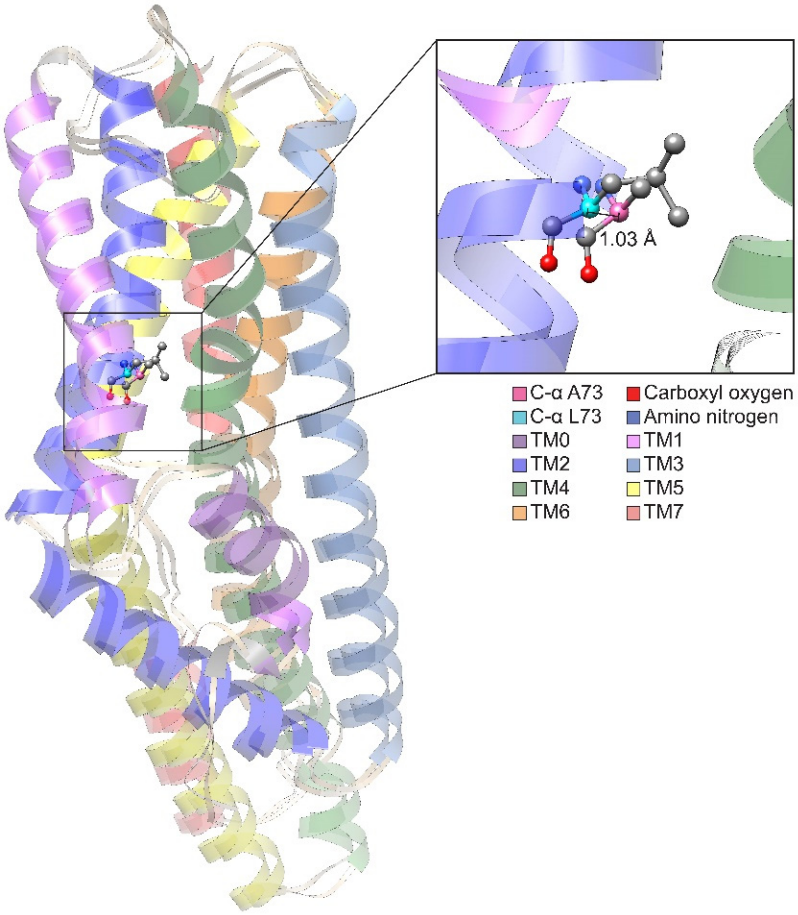
335 **Figure 6.** Representative models of docked skatole and indole in complex with
336 CquiOR10 and CquiOR10A73L using RosettaLigand. Each model shown is the
337 lowest interface-energy model from the 10 largest clusters of each docking study.
338 CquiOR10 – skatole (forest green), CquiOR10 – indole (brown), CquiOR10A73L
339 – skatole (light blue), and CquiOR10A73L – indole (purple). Atoms that are not
340 indole/skatole carbon atoms are color-coded by atom type: carbon (grey),
341 nitrogen (dark blue), and oxygen (red). Ala-73 and Leu-73 indicated with space-
342 filling representation.
343 **Figure supplement 1.** Clusters from RosettaLigand docking of skatole or indole
344 to CquiOR10 or CquiOR10A73L.
345 **Figure supplement 2.** Zoom out view of CquiOR10 or CquiOR10A73L
346 compexed to skatole or indole.
347 **Figure supplement 3.** Representative models of docked skatole or indole to
348 CquiOR10 or CquiOR10A73L with PLIC analysis.
349 **Figure supplement 4.** Superimposition of all OR-skatole and OR-indole clusters
350
351 In most models, skatole and indole form contacts with CquiOR10 and
352 CquiOr10A73L in a similar plane about a center of rotation. These observations
353 are supported by skatole and indole not containing rotatable bonds, thus relying
354 on rigid translational movements and rotation to form favorable contacts with the
355 rotatable and repackable receptor residues. Additionally, our models position
356 indole and skatole within a series of nonpolar, polar-uncharged, and aromatic
357 amino acids. Protein-ligand interaction profiler (PLIP) analysis (Adasme et al.,
358 2021) suggests that the bulk of favorable interactions are nonpolar, occasional
359 hydrogen bonding with the odorant NH group, and occasional parallel pi stacking
360 (*Figure 6—figure supplement 3*).
361
362 The ligand-binding pocket is formed by TMs 2, 4, 5, and 6. The most frequent
363 hydrophobic contacts were formed by Tyr-185 (38/40 models; TM4) and Ile-69
364 (32/40 models; TM2) (*Extended—Tables 5-7*). Phe-261 (TM5) also formed

365 hydrophobic interactions in all four receptor-ligand complex sets of models, but
366 less frequently (22/40 models). Asn-72 (TM2), Tyr-287 (TM6), and Ile-291 (TM6)
367 formed hydrophobic contacts more frequently with skatole structures compared
368 to indole structures (Asn-72: 12/20 vs 9/20, Tyr-287: 14/20 vs 5/20, Ile-291: 8/20
369 vs 3/20). Notably, Leu-73 (TM2) in CquiOR10A73L formed hydrophobic contacts
370 with indole (2/10) and skatole (5/10), but Ala-73 (TM2) did not form hydrophobic
371 contacts with indole or skatole. Tyr-138 (TM3) formed hydrophobic contacts with
372 skatole in CquiOR10A73L (2/10), but not in CquiOR10. Gln-294 (TM6) formed
373 hydrophobic contacts with skatole and indole in these structures, except
374 CquiOR10A73L-skatole; however, Gln-294 does form a hydrogen bond in one to
375 two models for each structure. Of note, the range of PLIP-reported hydrophobic
376 contacts among the ten representative models for each structure changed from
377 an average of 6.3±0.5 contacts (Mean±SEM) in CquiOR10 to 5.2±0.3 contacts in
378 CquiOR10A73L ($p = 0.0354$, Mann-Whitney test). Considering the number of
379 unique residues forming hydrophobic contacts, the averages were 4.9±0.4 and
380 4.9±0.2 contacts, respectively ($p = 0.97$, Mann-Whitney test).
381 There were at most two models in each structure with hydrogen bonds
382 (*Extended—Tables 8-10*). For CquiOR10-skatole, two models formed hydrogen
383 bonds with Gln-294. For CquiOR10A73L-skatole, one model formed hydrogen
384 bonds with Asn-72, and one other with Gln-294 (*Extended—Tables 10*). For
385 CquiOr10-indole, one model formed a hydrogen bond with the backbone of Ile-
386 69, another model with Tyr-138, another with Tyr-287, and a fourth model with
387 Gln-294. For CquiOR10A73L-indole, only one model formed a hydrogen bond at
388 Gln-294.
389 Parallel pi-stacking interactions (*Extended—Tables 11-13*) were identified for
390 CquiOR10A73L-skatole, CquiOR10-indole, and CquiOR10A73L-indole. For
391 CquiOR10A73L-skatole, two models each formed a parallel pi-stacking
392 interaction with Tyr-185 (*Extended—Table 13*). For CquiOR10-indole, a model
393 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 seven representative models.

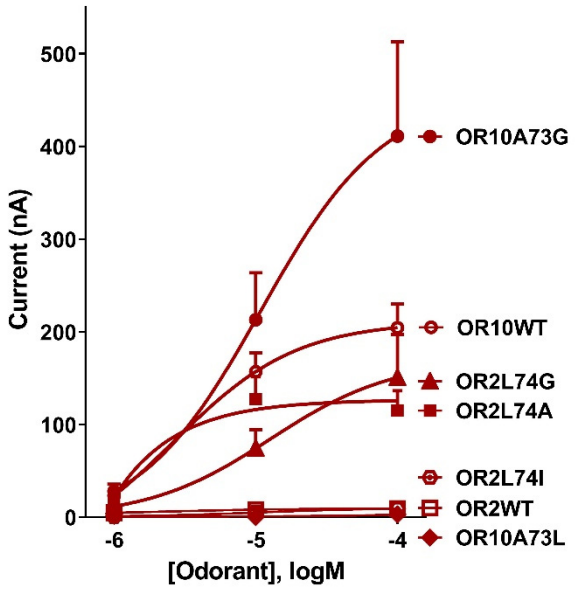
From the PLIP analysis of our docking study, we find of most importance skatole and indole not forming contacts with Ala-73 in CquiOR10 models. By contrast, skatole formed hydrophobic contacts with Leu-73 in five of the ten representative models, while indole formed contacts with Leu-73 in two representative models. This suggests that Ala-73 may indirectly affect specificity by modulating the volume of the binding pocket.

Structurally aligning CquiOR10 and CquiOR10A73L receptors by TM7b (*Extended—Table 14*), demonstrates approximately a 1 Å α-Carbon outward shift of A73L (*Extended—Table 15*), suggesting a tightly constrained space in CquiOR10 and an expanded space in CquiOR10A73L relative to the protein backbone (*Figure 7*).



408 **Figure 7.** Comparison of CquiOR10 and CquiOR10A73L models. An
409 approximate 1 Å α-Carbon outward shift of Leu-73 (forest green) in CquiOR10
410 model relative to Ala-73 (light blue) in CquiOR10A73L model. Models were
411 superimposed using the TM7b region. Residue 73 amino nitrogen is colored in
412 dark blue, and carboxyl oxygen is colored in red in each model.
413 **Figure supplement 1.** Effect of single-point mutations around A73.
414 **Figure supplement 2.** Quantification of single-point mutations around A73.
415
416 This difference in binding pocket volume is not ligand-induced, but rather
417 independent of skatole or indole binding.
418 Consistent with our structural hypothesis of space constraints, mutation of Ala-73
419 in CquiOR10 to Ile or Val negatively affected receptor function, whereas
420 CquiOR10A73G retained specificity and showed higher sensitivity. Specifically,
421 CquiOR10A73I/CquiOrco- and CquiOR10A73V/CquiOrco-expressing oocytes did
422 not respond to skatole or indole (*Figure 7—figure supplement 1A,B*). We
423 concluded that mutations with these bulkier residues caused loss of binding to
424 indole or skatole, given that these receptors were functional, as indicated by the
425 potent responses elicited by OLC12. On the other hand,
426 CquiOR10A73G/CquiOrco-expressing oocytes showed the same skatole
427 specificity as the wildtype receptor (*Figure 7—figure supplement 1C*).
428 Interestingly, skatole elicited stronger currents recorded from CquiOR10A73G
429 than the wildtype receptor CquiOR10 (*Figure 7—figure supplement 2*), further
430 implying a higher affinity for skatole when there is a reduced constraint, or
431 increased volume, of the binding pocket.
432
433 Since CquiOR2 is homologous to CquiOR10, we also propose that our space
434 constraint structural hypothesis can be applied to CquiOR2. While we did not
435 perform CquiOR2 docking, we indirectly examined space constraints in CquiOR2
436 by mutating Leu-74 to Ile, Val, or Gly. We then tested oocytes co-expressing
437 CquiOrco with CquiOR2L74I, CquiOR2L74V, or CquiOR2L74G (*Figure 7—figure*

438 *supplement 1D-F*). Consistent with the relaxation of the space constraints,
439 CquiOR2L74I retained the specificity of the wildtype receptor to indole, whereas
440 CquiOR2L74G showed a reverse specificity profile (*Figure 7—figure supplement*
441 *1F*), i.e. more robust response to skatole than indole. Interestingly,
442 CquiOR2L74V/CquiOrco-expressing oocytes generated nearly equal, albeit
443 small, currents when stimulated with the two oviposition attractants (skatole,
444 8.7±2.7 nA; indole, 9.0±2.5 nA at 100 μM; *n* = 3, P>0.9999, Wilcoxon matched-
445 pairs signed-rank test).
446



447
448 **Figure 8.** Concentration-dependent responses elicited by 3-ethylindole in
449 oocytes co-expressing CquiOrco with CquiOR10, CquiOR2, or single-point
450 mutants. Bars represent SEM (n = 4-10).
451
452 Next, we tested the space constraints hypothesis with a bulkier ligand, 3-
453 ethylindole. A CquiOR10 mutant with a less space-filling residue,
454 CquiOR10A73G, elicited responses to 3-ethylindole (526±110 nA) higher than
455 the responses to indole (205±81 nA, at 100 μM), although less robust than the
456 skatole responses (1939±142 nA; all at 100 μM) (*Figure 7—figure supplement 3*).

457 By contrast, receptors with a bulkier residue (CquiOR2WT, CquiO10A73L, and
458 CquiOR2L74I) did not respond to 3-ethylindole (Figure 8). However, CquiOR2
459 mutant with less space-filling residues, CquiOR2L74A and CquiOR2L74G,
460 responded to 3-ethylindole in a dose-dependent manner (Figure 8). Additionally,
461 CquiOR10A73G elicited dose-dependent strong responses to 3-ethylindole than
462 CquiOR10 (Figure 8). In summary, the findings that a bulkier, non-natural ligand,
463 elicited more robust responses when CquiOR10 and CquiOR2 residues at 73
464 and 74, respectively, were mutated into Gly are consistent with the space
465 constraints hypothesis.
466
467 **DISCUSSION**
468 Our mutation studies demonstrated that a single amino acid residue substitution
469 in two narrowly tuned odorant receptors from the southern house mosquito can
470 revert their specificity to the oviposition attractants skatole and indole.
471
472 RoseTTAFold modeling and RosettaLigand docking studies suggest that the
473 specificity determinant amino acid residue (Ala-73) did not form direct contacts
474 with ligands but rather provided a finely-tuned, sensitive volumetric space to
475 accommodate odorants. Interpreting CquiOR10-A73L as an expanded space is
476 counterintuitive when compared to our oocyte recordings; the wild-type
477 CquiOR10 receptor responded better to the bulkier skatole, (*Figure 1A*), while
478 CquiOR10-A73L responded better to the smaller indole (*Figure 2I*).
479 Considering both experimental and modeling data, the reason for differing
480 receptor responses at residue 73 could be due to the repacking of this position to
481 accommodate more carbon sidechain atoms that are reducing the binding pocket
482 volume. Leu-73 contains three additional carbon sidechain atoms: γ-Carbon, δ1-
483 Carbon, and δ2-Carbon. Adding these atoms increases the sidechain rotatable
484 bond count from 0 (for Ala-73) to 2 (for Leu-73), causing an increase in
485 repackable states for the binding pocket. From our models, the distance of Leu-
486 73 α-Carbon to its δ1-Carbon or δ2-Carbon is approximately 2.8 to 3.9 Å. The

487 Ala-73 α-Carbon to its β-Carbon is approximately 1.5 Å. Since the Ala-73 α-
488 Carbon is approximately 0.9 to 1.1 Å inward relative to Leu-73 α-Carbon, Leu-73
489 C-δ atoms could be further inward by approximately 0.2-1.5 Å relative to Ala-73
490 C-β. This could cause repacking of surrounding side chains while reducing the
491 volume of the binding pocket, or shift key, nearby residues required for receptor
492 response. Further, from our models, the Carbon-Carbon length of skatole's
493 methyl group is approximately 1.5 Å. This distance discrepancy, the fact that
494 indole is a rigid molecule, and that the methyl group of skatole is the only
495 rotamer, could also suggest that the binding pocket around position 73 is tightly
496 regulated. The 1.5 Å increase from skatole's methyl group could prevent skatole
497 from occupying the appropriate configuration for receptor response with Leu-73.
498 Conversely, the additional inward 0.2-1.5 Å increase from Leu-73's sidechain
499 could provide indole more contacts or a more favorable configuration with
500 surrounding CquiOR10 sidechains to prompt increased receptor response. This
501 hypothesis is supported by seven of ten representative models of
502 CquiOR10A73L-indole having eight total parallel pi-stacking interactions with Tyr-
503 185; this is a considerable increase compared to two representative interactions
504 of CquiOR10-indole having two total parallel pi-stacking interactions with Tyr-138
505 and Tyr-152. Combined, our results suggest that the residue at 73 provide a
506 finely tuned volumetric space to accommodate specific oviposition attractants.
507
508 While modeling generally has a degree of uncertainty, RosettaFold and
509 AlphaFold are highly accurate (Baek & Baker, 2022). Given the challenge of
510 solving insect OR cryo-EM structures, the new accurate modeling methods are
511 currently the best way to advance the field and generate hypothesis for future
512 empirical tests.
513
514
515 ## Materials and methods
516 **Construction of chimeric receptors.** We used a previously obtained pGEMHE-
517 CquiOR10 (Hughes et al., 2010) plasmid to amplify the sequences of the desired

518 chimeric receptors. To obtain the full length of chimeric *OR* genes, we used
519 specific primers, with overlapping sequences at each end to target sequences
520 (adaptor underlined) for cloning into pGEMHE vector, which was linearized at
521 restriction sites *XmaI* (5'-CCCGGG-3') and *XbaI* (5'-TCTAGA-3'):CqOR10_Fw 5'-
522 GATCAATTCCCCGGGACCATGACCGCGGCACCCATTTT-3' and CqOR10_Rv
523 5'-CAAGCTTGCTCTAGATCAATTATAAACGCGTCTCAGCAGGGT-3'. To
524 generate a chimeric OR by PCR amplification, we designed specific primers for
525 the desired CquiOR2 TM domain. Simultaneously, we prepared the recipient
526 fragments of CquiOR10 to receive the CquiOR2 TM domains. The fragments
527 were then assembled using In-Fusion HD cloning kit (Clontech) to obtain the
528 desired chimeric OR. Each CquiOR2 TM domain was divided into two parts,
529 each part was synthesized with a specific primer, except for CquiOR2 TM7
530 domain, which used only one primer with the full domain. These primers
531 contained an overlap with CquiOR10 sequence followed by one part of the
532 CquiOR2 TM sequence and an overlap with the other part of the CquiOR2 TM
533 domain sequence, which was synthesized in another PCR. The underlined
534 sequences represent CquiOR2 TM domains and, unless otherwise specified,
535 sequences in italic are overlaps for CquiOR2 TM domains. CqOR10^{M1}_Fw: *5'-*
536 *GGTGACCGTGCTGA*ACGTGTTCCAGTTTATGAACCTGTTTCGAGCCTGGGG
537 CAACATC-3'; CqOR10^{M1}_Rv: 5'-
538 *TCAGCACGGTCACC*GGAATGCAGCCGAGGAGGTAACTGAGGACGTTGCCT
539 ACTGTGATCTCAAGG-3'; CqOR10^{M2}_Fw: 5'-
540 *CCGTGCTGTACTT*CAACCTTGTGTTGAGAACCACGTTTATACTGTGCAATCG
541 TCAGGATTATGAGG-3'; CqOR10^{M2}_Rv: 5'-
542 *AAGTACAGCACGG*TGAAATATCCGTCGATGATGATTTTGTCGATGTTGCCCC
543 AGGCT-3'; CqOR10^{M3}_Fw: 5'-
544 *GTTCATCAGTGCGT*GCTTCGTGACGTATCCGCTTTTTTCACCGACACGTAGC
545 CTCCCGTACG-3'; CqOR10^{M3}_Rv: 5'-
546 *ACGCACTGATGAAC*GCTCCCAGCCAGAGGTTCGATTTGGACAGCAGTCGG
547 GCACGTTTGGTGA-3'; CqOR10^{M4}_Fw: 5'-

548 *CACGTTTCCGGCGT*GCTGCATGTACATTCCGTTTACCAGCTTCTTCGCCACG
549 ACTACTTTG-3'; CqOR10^{M4}_Rv: 5'-
550 *ACGCCGGAAACGTG*AGGTACACTTGCAGAAAAAACACAACCTGGTACAGGG
551 GCGTC-3'; CqOR10^{M5}_Fw: 5'-
552 *GCTATGCGCCTTGC*TGTTTCTACTTAGCACCAGCAATAATCCCGCGCAAATT
553 ATCATCGTGG-3'; CqOR10^{M5}_Rv: 5'-
554 *GCAAGGCGCATAGC*ATCATCCCAAACGATAGCAACTCAATCAGACAGATGT
555 AGGTCACCAGCG-3'; CqOR10^{M6}_Fw: 5'-
556 *TCTTTATGATTCT*GTCCCAGATGTACGCCCTGTACTGGCACGCCAACGAGCT
557 GCG-3'; CqOR10^{M6}_Rv: 5'-
558 *AGAATCATAAAGA*TGTACGATCCGATCATCACCATCTGCGCGGGATTTTCGA
559 TAATGTTCAGC-3'; CqOR10^{M7}_Rv: 5'-
560 *CAAGCTTGCTCTAGA*TCAATTATAAACGCGTCTCAGCAACGTAAAGTACGAA
561 TACGAGGCATTGATCAACTTTTGAAACATTTCCAAGGTCATCGGATAGACGT
562 TGCCTACTGTGATCTCAAGG-3' (here the italic represents the pGEMHE
563 adaptor. The two fragments were amplified using a combination of CqOR10_Fw
564 with CqOR10^{Mx}_Rv primers and CqOR10_Rv with CqOR10^{Mx}_Fw primers. Then,
565 we cloned the two fragments amplified by PCR into pGEMHE vector using the In
566 Fusion system as described below. Chimeric plasmid with a single TM swapped
567 was used as a template to generate chimeric OR with two TM swapped and
568 subsequently chimeric ORs with multiple CquiOR2 TM domains.
569
570 **Chimeric OR cloning and subcloning into pGEMHE.** The fragments
571 amplification was performed using Platinum Taq DNA Polymerase High Fidelity
572 (Thermo Fisher Scientific) with the following conditions: 95°C for 5 min, followed
573 by 30 cycles of 95°C for 20 s, 57ºC 30 s for annealing and 68°C for 1.5 min, and
574 extension at 68°C for 5 min. PCR products were purified by QIAquick gel
575 extraction kit (Qiagen). The target pGEMHE plasmid was cut by *XmaI* and *XbaI*
576 in separate reactions. Ligation was done with In Fusion (Takara Bio USA) system
577 and the transformation was performed using Stellar competent cells (Takara Bio

578 USA) in heat-shock. After selecting the cells, plasmids were extracted with
579 QIAprep Spin Miniprep kit (Qiagen). The cloned gene was verified by DNA
580 sequencing (Berkeley Sequencing Facility).
581
582 **Site-directed point mutagenesis and fragment replacement.** Phusion™ Site-
583 Directed Mutagenesis Kit (Thermo Scientific, West Palm Beach, FL) was used to
584 generate point mutations and TM fragment replacements. Mutations were
585 created with mismatched 5′-phosphorylated mutagenic primers and PCR
586 amplification. The chimeric sequences (CquiOR10^{M2,5,6,7}) in pGEMHE vector
587 were used as templates for rescues, whereas the wildtype sequences in
588 pGEMHE vector served as templates for point-mutations. Rescue primers:
589 CqOR10M2_M2567OUTERup: 5'-
590 GATGATGACCTCGTCGATGTTGCCCCAGGC-3';
591 CqOr10M2_M2567OUTERdn: 5'-
592 AACGCATATTTCACCGTGCTGTACTTCAACC-3'; CqOr10M2_M2567INNERup:
593 5'-CGCAGCACCGCGTTGAAGTACAGCACGGTG-3';
594 CqOr10M2_M2567INNERdn: 5'-AACAATTTTCATACTGTGCAATCGTCAGGA-
595 3'; CqOR10M2_M2567MID-INNERup: 5'-
596 CATCGACGGCTACTTTGCGATGATTTTCTTCAACGCG-3';
597 CqOR10M2_M2567MID-INNERdn: 5'-
598 ATGATTTTGTCGATGTTGCCCCAGGCTC-3'; CqOr10M2_M2567_L73Aup: 5'-
599 CGTTGAAGTACAGCACGGTGAAATAT-3'; CqOr10M2_M2567_L73Adn: 5'-
600 CTGTGTTGAGAACCACGTTTATACTGT-3'; CqOr10M2_M2567_T78Iup: 5'-
601 ATGGTTCTCAACACAAGGTTGAAGTA-3'; CqOr10M2_M2567_T78Idn: 5'-
602 CTTTATACTGTGCAATCGTCAGGATTATG-3'; Point mutation primers:
603 CqOR10A73up: 5'-GTTGAAGAAAATCATCGCAAAGTA-3'; CqOR10A73Ldn: 5'-
604 CTGGTGCTGCGAACAATTTTC-3'; CqOR10A73Adn: 5'-
605 GCGGTGCTGCGAACAATTTTC-3'; CqOR10A73Gdn: 5'-
606 GGGGTGCTGCGAACAATTTTC-3'; CqOR10A73Vdn: 5'-
607 GTGGTGCTGCGAACAATTTTC-3'; CqOR10A73Idn: 5'-

608 ATCGTGCTGCGAACAATTTTC-3'; CqOR2L74up: 5'-
609 GTTGAAGTACAGCACGGTGAAATA-3'; CqOr2L74Adn: 5'-
610 GCTGTGTTGAGAACCACGTTTATAC-3'; CqOr2L74Idn: 5'-
611 ATCGTGTTGAGAACCACGTTTATAC-3'; CqOr2L74Gdn: 5'-
612 GGGGTGTTGAGAACCACGTTTATAC-3'; CqOr2L74Vdn: 5'-
613 GTGGTGTTGAGAACCACGTTTATAC-3'. The amplified linear PCR products
614 containing the desired modification were ligated and transformed into Stellar™
615 Competent Cells (Takara Bio USA). All sequences were confirmed by DNA
616 sequencing (UC Berkeley DNA Sequencing Facility).
617
618 **In vitro transcription, oocyte microinjection and two-electrode voltage-**
619 **clamp assay (TEVC).** Capped OR cRNA was prepared using mMESSAGE
620 mMACHINE T7 Kit (Ambion) as previously described (Xu et al., 2019). Purified
621 OR cRNA was resuspended in nuclease-free water at 200 ng/μL and
622 microinjected into *Xenopus laevis* oocytes on stage V or VI (EcoCyte Bioscience)
623 along with the same amount of CquiOrco. Injected oocytes were incubated at 18
624 °C for 3–7 d in a modified Barth's solution [88 mM NaCl, 1 mM KCl, 2.4 mM
625 $NaHCO_3$, 0.82 mM $MgSO_4$, 0.33 mM $Ca(NO_3)_2$, 0.41 mM $CaCl_2$, and 10 mM
626 HEPES at pH 7.4] supplemented with 10 μg/mL gentamycin, 10 μg/mL
627 streptomycin and 1.8 mM sodium pyruvate. Two-electrode voltage-clamp
628 technique (TEVC) was used to measure odorant-induced currents at a holding
629 potential of −80 mV. Signals were amplified with an OC-725C amplifier (Warner
630 Instruments), low-pass filtered at 50 Hz, and digitized at 1 kHz. Data acquisition
631 and analysis were performed with Digidata 1440A and software pCLAMP 10
632 (Molecular Devices). Skatole (98%, CAS# 83-34-1), indole (98%, CAS# 120-72-
633 9), and 3-ethylindole (97%, CAS# 1484-19-1) were provided by Sigma-Aldrich
634 (Milwaukee, WI), ACROS Organics (Geel, Belgium), and AmBeed (Arlington Hts,
635 IL), respectively. Phenol (99.5%, CAS# 108-95-2), 3,5-dimethylphenol (99%,
636 CAS# 108-68-9), 1-methylindole (97%, CAS# 603-76-9), 2-methylindole (98%,
637 CAS# 95-20-5), 5-methylindole (99%, CAS# 614-96-0), 6-methylindole (97%,

638 CAS# 3420-02-8) and 7-methylindole (97%, CAS# 933-67-5) were acquired from
639 Sigma-Aldrich, and 4-methylindole (99%, CAS# 16096-32-5) was provided by
640 ACROS.
641
642 **Rosetta structural modeling and docking.** We used RoseTTAFold (Baek et
643 al., 2021) and AlphaFold (Jumper et al., 2021) to generate predicted structures of
644 CquiOR10, CquiOR2, CquiOR10A73L, and CquiOR2L74A. The OpenEye
645 Omega toolkit (Hawkins et al., 2010) was used to generate conformer libraries.
646 Molecular docking was performed using RosettaLigand (Davis & Baker, 2009;
647 DeLuca et al., 2015). The top models were visually analyzed using UCSF
648 Chimera (Pettersen et al., 2004).
649
650 To verify RosettaLigand could effectively sample homologous receptors in
651 complex with odorants, we used the structure of eugenol in complex with the
652 insect olfactory receptor OR5 from *Machilis hrabei* (MhOR5) as a control test
653 case (Del Marmol et al., 2021). After verifying RosettaLigand could recapitulate
654 the MhraOR5 — eugenol complex from structural studies, we docked indole and
655 skatole to the RoseTTAFold structure models of CquiOR10 and CquiOR10A73L
656 receptors with the same method. Briefly, the initial placement of all ligands into
657 respective structures was guided by the position of eugenol in a complex with
658 MhOR5 as a basis since it is the only homologous structure with an odorant. For
659 an unbiased receptor sampling, an initial transformation was performed on the
660 ligand relative to the receptor site using Rosetta's Transform mover. Within a 7 Å
661 sphere, the ligand underwent a Monte Carlo simulation, whereby the ligand was
662 allowed to translate up to 0.2 Å and rotate up to 20 degrees. This was performed
663 500 times on the ligand, with the lowest scoring pose used as the starting pose
664 for docking. This form of docking is termed local docking and is commonplace
665 when there is experimental evidence and homologous structures as templates.
666 physiologically relevant binding modes can be identified with local, high-

667 resolution ligand docking (Kaufmann et al., 2009). For additional methods, see
668 Appendix 2.
669

670 ## Appendix 1

671

672 **Preliminary screening of chimeric receptors.** During the life of this project, the
673 structure of an odorant receptor coreceptor AbakOrco was reported (Butterwick
674 et al., 2018). We then compared the Cryo-EM structure (Butterwick et al., 2018)
675 with the predicted topology for AbakOrco using the same OCTOPUS method
676 (Viklund & Elofsson, 2008) we used to identify CquiOR10 and CquiOR2 TMs
677 (*Figure 1—figure supplement 1*) (Viklund & Elofsson, 2008). The almost perfect
678 overlap between OCTOPUS prediction and the Cryo-EM structure (*Figure 1—*
679 *figure supplement 3*) validated not only our TM predictions (*Figure 1—figure*
680 *supplement 1*), but also the 21 chimeric ORs already tested when the structure of
681 the coreceptor AbakOrco was reported (Hughes et al., 2017; Luetje et al., 2013).
682 CquiOR10^{M1,2,3,4,5,6,7}/CquiOrco-expressing oocytes did not respond to skatole or
683 indole. Likewise, the chimeric receptor with six transmembrane domains
684 replaced, CquiOR10^{M1,2,3,4,5,6} was also silent. Next, we replaced one TM at a
685 time. CquiOR10^{M1} and CquiOR10^{M2} were silent, whereas CquiOR10^{M3} (*Figure*
686 *1—figure supplement 2A*), CquiOR10^{M4} (*Figure 1—figure supplement 2B*),
687 CquiOR10^{M5} (*Figure 1—figure supplement 2C*), CquiOR10^{M6} (*Figure 1—figure*
688 *supplement 2D*), and CquiOR10^{M7} (*Figure 1—figure supplement 2E*) responded
689 to both skatole and indole with specificity similar to that of the wildtype receptor.
690 Although these chimeric receptors differed from CquiOR10 in terms of sensitivity,
691 they showed the same specificity to skatole as CquiOR10 (*Figure 1A*). While
692 CquiOR10^{M3} and CquiOR10^{M4} were less sensitive than CquiOR10, CquiOR10^{M7}
693 showed comparable sensitivity to CquiOR10, and CquiOR10^{M5} and CquiOR10^{M6}
694 were more sensitive to skatole than the wildtype receptor. We remained focused
695 on specificity and did not attempt to quantify the different responses.

696 We tested a chimeric receptor with the other five TM domains swapped.
697 CquiOR10^{M3,4,5,6,7}/CquiOrco-expressing oocytes mirrored the profiles obtained
698 with CquiOR10 in terms of sensitivity and specificity (*Figure 1—figure*
699 *supplement 2F*). We then kept the wildtype TM7, in addition to TM1 and TM2, in
700 the next chimeric ORs tested. Albeit less sensitive than CquiOR10,
701 CquiOR10^{M3,4,5,6} showed the same specificity of the wildtype receptor (*Figure 1—*
702 *figure supplement 2G*). Following, we tested chimeric receptors with three TM
703 domains swapped. CquiOR10^{M4,5,6} (*Figure 1—figure supplement 2H*),
704 CquiOR10^{M3,5,6} (*Figure 1—figure supplement 2I*), CquiOR10^{M3,4,6} (*Figure 1—*
705 *figure supplement 2J*), and CquiOR10^{M3,4,5} (*Figure 1—figure supplement 2K*) did
706 not differ from CquiOR10 in terms of specificity or sensitivity. Next, we focused
707 on chimeric receptors with two TMs swapped. Thus, we tested CquiOR10^{M3,4}
708 (*Figure 1—figure supplement 2L*), CquiOR10^{M3,5} (*Figure 1—figure supplement*
709 *2M*), CquiOR10^{M3,6} (*Figure 1—figure supplement 2N*), CquiOR10^{M3,7} (*Figure 1—*
710 *figure supplement 2O*), CquiOR10^{M4,5} (*Figure 1—figure supplement 2P*),
711 CquiOR10^{M4,6} (*Figure 1—figure supplement 2Q*), CquiOR10^{M4,7} (*Figure 1—figure*
712 *supplement 2R*), CquiOR10^{M5,6} (*Figure 1—figure supplement 2S*), CquiOR10^{M5,7}
713 (*Figure 1—figure supplement 2T*), and CquiOR10^{M6,7} (*Figure 1—figure*
714 *supplement 2U*). They all showed skatole-specific profiles, most of them with
715 comparable sensitivity, whereas CquiOR10^{M4,5}, CquiOR10^{M5,7}, and CquiOR10^{M6,7}
716 were more sensitive and CquiOR10^{M3,6} less sensitive than CquiOR10.
717 Interestingly, none of the chimeric receptors having TM1 as one of the two
718 domains swapped was functional. Specifically, oocytes co-expressing CquiOrco
719 with either CquiOR10^{M1,2}, CquiOR10^{M1,3}, CquiOR10^{M1,4}, CquiOR10^{M1,5},
720 CquiOR10^{M1,6}, CquiOR10^{M1,7} did not respond to skatole or indole (*Extended—*
721 *Table 1*). Likewise, CquiOR10^{M2,3}, CquiOR10^{M2,4}, CquiOR10^{M2,5}, and
722 CquiOR10^{M2,6} did not respond to these oviposition attractants (*Extended—Table*
723 *1*).
724
725

726 # Appendix 2

727

728 **Ligand conformer library generation.** The Protein DataBank file (.pdb) of
729 eugenol was extracted from the MhOR5 structure (PDB ID: 7LID), while indole
730 and skatole were extracted as Structure-Data Files files (.sdf) from PubChem
731 (Kim et al., 2021). Each ligand structure was bond-corrected, protonated at pH
732 7.4, energy minimized using the Merck Molecular Force Field (Halgren, 1996),
733 and saved as Tripos Mol2 files (.mol2) with the Avogadro software (Hanwell et
734 al., 2012). Next, partial charge, atom, and bond type assignment were performed
735 with Antechamber AM1BCC correction using Amber Tools (Case et al., 2021;
736 Salomon-Ferrer et al., 2013); AM1BCC partial-charge correction is common in
737 Rosetta ligand docking protocols (Park et al., 2021) and has been shown to be
738 similarly ranked with other Rosetta-based partial-charge methods for
739 RosettaLigand (Smith & Meiler, 2020). An in-house script using the OpenEye
740 Omega toolkit and Rosetta was used to generate the conformer library and
741 associated Rosetta-readable ligand parameters file.

742

743 **Protein receptor preparation.** The MhOR5 structure was relaxed with backbone
744 constraints using the RosettaRelax protocol (Nivon et al., 2013) prior to
745 RosettaLigand docking. This is commonplace to allow the repacking of protein
746 sidechains and minimization of the structure into the Rosetta score function for
747 comparison between structures. For RoseTTAFold-generated structures, the
748 RosettaRelax protocol was not performed since the Rosetta score function was
749 used to generate the structures.

750

751 **RosettaLigand docking.** RosettaLigand docking was performed using
752 RosettaScripts protocols described previously (Davis & Baker, 2009; DeLuca et
753 al., 2015). Briefly, initial placement of all ligands into respective structures was
754 guided by the position of eugenol in complex with MhOR5 as a basis. For an
755 unbiased sampling of the receptor, an initial transformation was performed on the

ligand relative to the receptor site using Rosetta's Transform mover. Within a 7 Å sphere, the ligand underwent a Monte Carlo simulation, whereby the ligand was allowed to translate up to 0.2 Å and rotate up to 20 degrees. This was performed 500 times on the ligand, with the lowest scoring pose used as the starting pose for docking. Six cycles of high-resolution docking were performed with repacking of sidechains every third iteration using Rosetta's HighResDocker mover. Lastly, the protein/ligand complex was minimized using Rosetta's FinalMinimizer mover and interface scores reported using the InterfaceScoreCalculator.

For every test case, 100,000 docking poses were generated. The top 10,000 models were selected by Rosetta interface energy (Interface_delta_X) metric, followed by ligand clustering using Biomol2Clust (D. et al., 2021), the hdbscan algorithm (McInnes et al., 2017), and a minimum cluster threshold of 50 poses. The 10 most frequently sampled clusters were visually analyzed using UCSF Chimera (Pettersen et al., 2004) by using the lowest interface energy pose in each cluster as the representative for that cluster. The Protein Ligand Interaction Profiler (Salentin et al., 2015) was performed on these representative poses to identify potential hydrophobic interactions, hydrogen bonds, π-Stacking, and salt bridges. Hydrogen bond and pi-stacking interactions were filtered by previously reported bond distances (Bissantz et al., 2010) to present a conservative estimate of potential contacts. Atom types from PLIP analysis are reported using IDATM nomenclature (Meng & Lewis, 1991).

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.

Competing Interest Statement: The authors declare no competing interest.

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.

Data Availability. All data generated during this study are included in the manuscript and supporting files.

References

Adasme, M. F., Linnemann, K. L., Bolz, S. N., Kaiser, F., Salentin, S., Haupt, V J., & Schroeder, M. (2021). PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Research, 49*(W1), W530-W534. https://doi.org/10.1093/nar/gkab294

Arensburger, P., Megy, K., Waterhouse, R. M., Abrudan, J., Amedeo, P., Antelo, B., Bartholomay, L., Bidwell, S., Caler, E., Camara, F., Campbell, C. L., Campbell, K. S., Casola, C., Castro, M. T., Chandramouliswaran, I., Chapman, S. B., Christley, S., Costas, J., Eisenstadt, E., Feschotte, C., Fraser-Liggett, C., Guigo, R., Haas, B., Hammond, M., Hansson, B. S., Hemingway, J., Hill, S. R., Howarth, C., Ignell, R., Kennedy, R. C., Kodira, C. D., Lobo, N. F., Mao, C., Mayhew, G., Michel, K., Mori, A., Liu, N., Naveira, H., Nene, V., Nguyen, N., Pearson, M. D., Pritham, E. J., Puiu, D., Qi, Y., Ranson, H., Ribeiro, J. M., Roberston, H. M., Severson, D. W., Shumway, M., Stanke, M., Strausberg, R. L., Sun, C., Sutton, G., Tu, Z. J., Tubio, J. M., Unger, M. F., Vanlandingham, D. L., Vilella, A. J., White, O., White, J. R., Wondji, C. S., Wortman, J., Zdobnov, E. M., Birren, B., Christensen, B. M., Collins, F. H., Cornel, A., Dimopoulos, G., Hannick, L. I., Higgs, S., Lanzaro, G. C., Lawson, D., Lee, N. H., Muskavitch, M. A., Raikhel, A. S., & Atkinson, P. W. (2010, Oct 1). Sequencing of Culex quinquefasciatus establishes a platform for mosquito comparative genomics. *Science, 330*(6000), 86-88. https://doi.org/10.1126/science.1191864

Auer, T. O., Khallaf, M. A., Silbering, A. F., Zappia, G., Ellis, K., Alvarez-Ocana, R., Arguello, J. R., Hansson, B. S., Jefferis, G., Caron, S. J. C., Knaden, M., & Benton, R. (2020, Mar). Olfactory receptor and circuit evolution promote host specialization. *Nature, 579*(7799), 402-408. https://doi.org/10.1038/s41586-020-2073-7

Baek, M., & Baker, D. (2022, Jan). Deep learning and protein structure modeling. *Nat Methods, 19*(1), 13-14. https://doi.org/10.1038/s41592-021-01360-8

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millan, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., van Dijk, A. A., Ebrecht, A. C., Opperman, D. J., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M. K., Dalwadi, U., Yip, C. K., Burke, J. E., Garcia, K. C., Grishin, N. V., Adams, P. D., Read, R. J., & Baker, D. (2021, Aug 20). Accurate prediction of protein structures and interactions using a three-track neural network. *Science, 373*(6557), 871-876. https://doi.org/10.1126/science.abj8754

Benton, R., Sachse, S., Michnick, S. W., & Vosshall, L. B. (2006, Feb). Atypical membrane topology and heteromeric function of Drosophila odorant receptors in vivo. *PLoS Biol, 4*(2), e20. https://doi.org/10.1371/journal.pbio.0040020

Bidula, S., Nadzirin, I. B., Cominetti, M., Hickey, H., Cullum, S. A., Searcey, M., Schmid, R., & Fountain, S. J. (2022, Jan). Structural basis of the negative allosteric modulation of 5-BDBD at human P2X4 receptors. *Mol Pharmacol, 101*(1), 33-44. https://doi.org/10.1124/molpharm.121.000402

Bissantz, C., Kuhn, B., & Stahl, M. (2010, Jul 22). A medicinal chemist's guide to molecular interactions. *J Med Chem, 53*(14), 5061-5084. https://doi.org/10.1021/jm100112j

Blackwell, A., Mordue, A. J., Hansson, B. S., Wadhams, L. J., & Pickett, J. A. (1993, Dec). A behavioral and electrophysiological study of oviposition cues for Culex quinquefasciatus. *Physiological Entomology, 18*(4), 343-348. <Go to ISI>://WOS:A1993MP27600003

Bohbot, J. D., Jones, P. L., Wang, G., Pitts, R. J., Pask, G. M., & Zwiebel, L. J. (2011, Jan). Conservation of indole responsive odorant receptors in mosquitoes reveals an ancient olfactory trait. *Chemical Senses, 36*(2), 149-160. https://doi.org/10.1093/chemse/bjq105

Bohbot, J. D., & Pitts, R. J. (2015, 2015-April-10). The narrowing olfactory landscape of insect odorant receptors [Hypothesis and Theory]. *Frontiers in Ecology and Evolution, 3*. https://doi.org/10.3389/fevo.2015.00039

Butterwick, J. A., Del Marmol, J., Kim, K. H., Kahlson, M. A., Rogow, J. A., Walz, T., & Ruta, V. (2018, Aug). Cryo-EM structure of the insect olfactory receptor Orco. *Nature, 560*(7719), 447-452. https://doi.org/10.1038/s41586-018-0420-8

Cao, S., Liu, Y., Wang, B., & Wang, G. R. (2021, Sep 24). A single point mutation causes one-way alteration of pheromone receptor function in two Heliothis species. *Iscience, 24*(9). https://doi.org/ARTN 102981

10.1016/j.isci.2021.102981

Carey, A. F., Wang, G., Su, C. Y., Zwiebel, L. J., & Carlson, J. R. (2010, Mar 04). Odorant reception in the malaria mosquito Anopheles gambiae. *Nature, 464*(7285), 66-71. https://doi.org/10.1038/nature08834

Case, D. A., Aktulga, H. M., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham III, T. E., Cisneros, G. A., Cruzeiro, V. W. D., Darden, T. A., & Duke, R. E. (2021). *Amber 2021*. https://ambermd.org/doc12/Amber21.pdf

Choo, Y. M., Xu, P., Hwang, J. K., Zeng, F., Tan, K., Bhagavathy, G., Chauhan, K., & Leal, W. S. (2018). Reverse chemical ecology approach for the identification of an oviposition attractant for Culex quinquefasciatus. *Proc Natl Acad Sci U S A, 115*(4), 714-719.

Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., & Carlson, J. R. (1999, Feb). A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in Drosophila. *Neuron, 22*(2), 327-338. https://doi.org/10.1016/s0896-6273(00)81093-4

D., T., Y., S., V., Š., & D., S. (2021). *Biomol2Clust.* In https://biokinet.belozersky.msu.ru/Biomol2Clust

Davis, I. W., & Baker, D. (2009, Jan 16). RosettaLigand docking with full ligand and receptor flexibility. *J Mol Biol, 385*(2), 381-392. https://doi.org/10.1016/j.jmb.2008.11.010

Del Marmol, J., Yedlin, M. A., & Ruta, V. (2021, Sep). The structural basis of odorant recognition in insect olfactory receptors. *Nature, 597*(7874), 126-131. https://doi.org/10.1038/s41586-021-03794-8

DeLuca, S., Khar, K., & Meiler, J. (2015). Fully flexible docking of medium sized ligand libraries with RosettaLigand. *PLOS ONE, 10*(7), e0132508. https://doi.org/10.1371/journal.pone.0132508

Gao, Q., & Chess, A. (1999, Aug 15). Identification of candidate Drosophila olfactory receptors from genomic DNA sequence. *Genomics, 60*(1), 31-39. https://doi.org/10.1006/geno.1999.5894

Guo, S., & Kim, J. (2010). Dissecting the molecular mechanism of drosophila odorant receptors through activity modeling and comparative analysis. *Proteins: Structure, Function, and Bioinformatics, 78*(2), 381-399. https://doi.org/https://doi.org/10.1002/prot.22556

Halgren, T. A. (1996, Apr). Merck molecular force field .1. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry, 17*(5-6), 490-519. https://doi.org/Doi 10.1002/(Sici)1096-987x(199604)17:5/6<490::Aid-Jcc1>3.0.Co;2-P

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012, Aug 13). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics, 4*. https://doi.org/Artn 17

10.1186/1758-2946-4-17

Hawkins, P. C., Skillman, A. G., Warren, G. L., Ellingson, B. A., & Stahl, M. T. (2010, Apr 26). Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and

Cambridge Structural Database. *J Chem Inf Model, 50*(4), 572-584. https://doi.org/10.1021/ci100031x

Hughes, D. T., Pelletier, J., Luetje, C. W., & Leal, W. S. (2010, Aug). Odorant receptor from the southern house mosquito narrowly tuned to the oviposition attractant skatole. *J Chem Ecol, 36*(8), 797-800. https://doi.org/10.1007/s10886-010-9828-9

Hughes, D. T., Pelletier, J., Rahman, S., Chen, S., Leal, W. S., & Luetje, C. W. (2017, May 01). Functional and nonfunctional forms of CquiOR91, an odorant selectivity subunit of Culex quinquefasciatus. *Chemical Senses, 42*(4), 333-341. https://doi.org/10.1093/chemse/bjx011

Hughes, D. T., Wang, G., Zwiebel, L. J., & Luetje, C. W. (2014, Nov). A determinant of odorant specificity is located at the extracellular loop 2-transmembrane domain 4 interface of an Anopheles gambiae odorant receptor subunit. *Chemical Senses, 39*(9), 761-769. https://doi.org/10.1093/chemse/bju048

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., & Hassabis, D. (2021, Aug). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589. https://doi.org/10.1038/s41586-021-03819-2

Kaissling, K. E. (2014). Pheromone reception in Insects: The example of silk moths. In C. Mucignat-Caretta (Ed.), *Neurobiology of Chemical Communication*. https://www.ncbi.nlm.nih.gov/pubmed/24830037

Kaufmann, K. W., Dawson, E. S., Henry, L. K., Field, J. R., Blakely, R. D., & Meiler, J. (2009, Feb 15). Structural determinants of species-selective substrate recognition in human and Drosophila serotonin transporters revealed through computational docking studies. *Proteins, 74*(3), 630-642. https://doi.org/10.1002/prot.22178

975 Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.
976 A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2021,
977 Jan 8). PubChem in 2021: new data content and improved web interfaces.
978 *Nucleic Acids Res, 49*(D1), D1388-D1395.
979 https://doi.org/10.1093/nar/gkaa971

980

981 Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M. E., Amrein, H., &
982 Vosshall, L. B. (2004, Sep 2). Or83b encodes a broadly expressed
983 odorant receptor essential for Drosophila olfaction. *Neuron, 43*(5), 703-
984 714. https://doi.org/10.1016/j.neuron.2004.08.019

985

986 Leal, W. S. (2013). Odorant reception in insects: roles of receptors, binding
987 proteins, and degrading enzymes. *Annu Rev Entomol, 58*, 373-391.
988 https://doi.org/10.1146/annurev-ento-120811-153635

989

990 Leal, W. S. (2016). Molecular biology of reception. *Pheromone Communication in*
991 *Moths: Evolution, Behavior, and Application*, 127-138. <Go to
992 ISI>://WOS:000388302800009

993

994 Leal, W. S. (2020). Mechanism of action of insect pheromones and other
995 semiochemicals. In H.-W. Liu & T. P. Begley (Eds.), *Comprehensive*
996 *Natural Products III* (pp. 222-236). Elsevier.
997 https://www.sciencedirect.com/science/article/pii/B9780124095472147353

998

999 Leal, W. S., Choo, Y. M., Xu, P., da Silva, C. S., & Ueira-Vieira, C. (2013, Nov
1000 12). Differential expression of olfactory genes in the southern house
1001 mosquito and insights into unique odorant receptor gene isoforms. *Proc*
1002 *Natl Acad Sci U S A, 110*(46), 18704-18709.
1003 https://doi.org/10.1073/pnas.1316059110

1004

1005 Leary, G. P., Allen, J. E., Bunger, P. L., Luginbill, J. B., Linn, C. E., Jr.,
1006 Macallister, I. E., Kavanaugh, M. P., & Wanner, K. W. (2012, Aug 28).
1007 Single mutation to a sex pheromone receptor provides adaptive specificity
1008 between closely related moth species. *Proc Natl Acad Sci U S A, 109*(35),
1009 14081-14086. https://doi.org/10.1073/pnas.1204661109

1010

1011 Luetje, C. W., Nichols, A. S., Castro, A., & Sherman, B. L. (2013). Functional
1012 assay of mammalian and Insect olfactory receptors using Xenopus
1013 oocytes. In C. J. Crasto (Ed.), *Olfactory Receptors: Methods and*

Protocols (pp. 187-202). Humana Press. https://doi.org/10.1007/978-1-62703-377-0_14

Lundin, C., Kall, L., Kreher, S. A., Kapp, K., Sonnhammer, E. L., Carlson, J. R., Heijne, G., & Nilsson, I. (2007, Dec 11). Membrane topology of the Drosophila OR83b odorant receptor. *FEBS Lett, 581*(29), 5601-5604. https://doi.org/10.1016/j.febslet.2007.11.007

Mboera, L. E. G., Takken, W., Mdira, K. Y., Chuwa, G. J., & Pickett, J. A. (2000, May). Oviposition and behavioral responses of Culex quinquefasciatus to skatole and synthetic oviposition pheromone in Tanzania. *Journal of Chemical Ecology, 26*(5), 1193-1203. https://doi.org/Doi 10.1023/A:1005432010721

McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *J. Open Source Softw., 2*, 205.

Meng, E. C., & Lewis, R. A. (1991, Sep). Determination of Molecular Topology and Atomic Hybridization States from Heavy-Atom Coordinates. *Journal of Computational Chemistry, 12*(7), 891-898. https://doi.org/DOI 10.1002/jcc.540120716

Millar, J. G., Chaney, J. D., & Mulla, M. S. (1992, Mar). Identification of oviposition attractants for Culex quinquefasciatus from fermented Bermuda grass infusions. *Journal of the American Mosquito Control Association, 8*(1), 11-17. <Go to ISI>://WOS:A1992HM11400002

Nakagawa, T., Sakurai, T., Nishioka, T., & Touhara, K. (2005, Mar 11). Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science, 307*(5715), 1638-1642. https://doi.org/10.1126/science.1106267

Neuhaus, E. M., Gisselmann, G., Zhang, W., Dooley, R., Stortkuhl, K., & Hatt, H. (2005, Jan). Odorant receptor heterodimerization in the olfactory system of Drosophila melanogaster. *Nat Neurosci, 8*(1), 15-17. https://doi.org/10.1038/nn1371

Nguyen, H. M., Singh, V., Pressly, B., Jenkins, D. P., Wulff, H., & Yarov-Yarovoy, V. (2017, Apr). Structural insights into the atomistic mechanisms of action

of small molecule inhibitors targeting the KCa3.1 channel pore. *Mol Pharmacol, 91*(4), 392-402. https://doi.org/10.1124/mol.116.108068

Nivon, L. G., Moretti, R., & Baker, D. (2013). A Pareto-optimal refinement method for protein design scaffolds. *PLOS ONE, 8*(4), e59004. https://doi.org/10.1371/journal.pone.0059004

Park, H., Zhou, G., Baek, M., Baker, D., & DiMaio, F. (2021, Mar 9). Force field optimization guided by small molecule crystal lattice data enables consistent sub-Angstrom protein-ligand docking. *J Chem Theory Comput, 17*(3), 2000-2010. https://doi.org/10.1021/acs.jctc.0c01184

Pask, G. M. (2020). Invertebrate olfactory receptors. In B. Fritzsch (Ed.), *The Senses: A Comprehensive Reference (Second Edition)* (pp. 453-465). Elsevier. https://doi.org/doi.org/10.1016/B978-0-12-809324-5.23814-5

Pellegrino, M., Steinbach, N., Stensmyr, M. C., Hansson, B. S., & Vosshall, L. B. (2011, Oct 27). A natural polymorphism alters odour and DEET sensitivity in an insect odorant receptor. *Nature, 478*(7370), 511-514. https://doi.org/10.1038/nature10438

Pelletier, J., Hughes, D. T., Luetje, C. W., & Leal, W. S. (2010). An odorant receptor from the southern house mosquito Culex pipiens quinquefasciatus sensitive to oviposition attractants. *PLOS ONE, 5*(4), e10090. https://doi.org/10.1371/journal.pone.0010090

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004, Oct). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry, 25*(13), 1605-1612. https://doi.org/10.1002/jcc.20084

Pitts, R. J., Huff, R. M., Shih, S. J., & Bohbot, J. D. (2021, Dec). Identification and functional characterization of olfactory indolergic receptors in Musca domestica. *Insect Biochem Mol Biol, 139*, 103653. https://doi.org/10.1016/j.ibmb.2021.103653

Pressly, B., Lee, R. D., Singh, V., Pessah, I. N., & Wulff, H. (2022, May). The seizure-inducing plastic explosive RDX inhibits the alpha1beta2gamma2

GABAA receptor. *Ann Clin Transl Neurol, 9*(5), 600-609. https://doi.org/10.1002/acn3.51536

Reynolds, S. M., Kall, L., Riffle, M. E., Bilmes, J. A., & Noble, W. S. (2008, Nov). Transmembrane topology and signal peptide prediction using dynamic bayesian networks. *PLoS Comput Biol, 4*(11), e1000213. https://doi.org/10.1371/journal.pcbi.1000213

Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F., & Schroeder, M. (2015, Jul 1). PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res, 43*(W1), W443-447. https://doi.org/10.1093/nar/gkv315

Salomon-Ferrer, R., Case, D. A., & Walker, R. C. (2013, Mar-Apr). An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews-Computational Molecular Science, 3*(2), 198-210. https://doi.org/10.1002/wcms.1121

Smith, S. T., & Meiler, J. (2020). Assessing multiple score functions in Rosetta for drug discovery. *PLOS ONE, 15*(10), e0240450. https://doi.org/10.1371/journal.pone.0240450

Stensmyr, M. C., Dweck, H. K., Farhan, A., Ibba, I., Strutz, A., Mukunda, L., Linz, J., Grabe, V., Steck, K., Lavista-Llanos, S., Wicher, D., Sachse, S., Knaden, M., Becher, P. G., Seki, Y., & Hansson, B. S. (2012, Dec 7). A conserved dedicated olfactory circuit for detecting harmful microbes in Drosophila. *Cell, 151*(6), 1345-1357. https://doi.org/10.1016/j.cell.2012.09.046

Viklund, H., & Elofsson, A. (2008). OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics, 24*(15), 1662-1668. https://doi.org/10.1093/bioinformatics/btn221

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., & Axel, R. (1999, Mar 5). A spatial map of olfactory receptor expression in the Drosophila antenna. *Cell, 96*(5), 725-736. https://doi.org/10.1016/s0092-8674(00)80582-6

1129 Wang, G., Carey, A. F., Carlson, J. R., & Zwiebel, L. J. (2010, Mar 02). Molecular
1130 basis of odor coding in the malaria vector mosquito Anopheles gambiae.
1131 *Proc Natl Acad Sci U S A, 107*(9), 4418-4423.
1132 https://doi.org/10.1073/pnas.0913392107

1133
1134 Xu, P., Choo, Y. M., Chen, Z., Zeng, F., Tan, K., Chen, T. Y., Cornel, A. J., Liu,
1135 N., & Leal, W. S. (2019, Jul 12). Odorant inhibition in mosquito olfaction.
1136 *Iscience, 19*, 25-38. https://doi.org/10.1016/j.isci.2019.07.008

1137
1138 Yang, K., Huang, L. Q., Ning, C., & Wang, C. Z. (2017, Oct 24). Two single-point
1139 mutations shift the ligand selectivity of a pheromone receptor between two
1140 closely related moth species. *Elife, 6*. https://doi.org/10.7554/eLife.29100

1141
1142 Yuvaraj, J. K., Roberts, R. E., Sonntag, Y., Hou, X. Q., Grosse-Wilde, E.,
1143 Machara, A., Zhang, D. D., Hansson, B. S., Johanson, U., Lofstedt, C., &
1144 Andersson, M. N. (2021, Jan 26). Putative ligand binding sites of two
1145 functionally characterized bark beetle odorant receptors. *BMC Biol, 19*(1),
1146 16. https://doi.org/10.1186/s12915-020-00946-6

1147

1148