Title

Structural basis of sensory receptor evolution in octopus.

Journal

Nature, 616(7956)

Authors

Bellono, Nicholas
Allard, Corey
Kang, Guipeun
et al.

Publication Date

2023-04-01

DOI

10.1038/s41586-023-05822-1

Peer reviewed

Published in final edited form as:

Nature. 2023 April ; 616(7956): 373–377. doi:10.1038/s41586-023-05822-1.

Structural basis of sensory receptor evolution in octopus

Corey A. H. Allard[1,4], **Guipeun Kang**[2,4], **Jeong Joo Kim**[2,4], **Wendy A. Valencia-Montoya**[1,3], **Ryan E. Hibbs**[2], **Nicholas W. Bellono**[1]

[1]Department of Molecular and Cellular Biology. Harvard University, Cambridge. M A, USA.

[2]Departments of Neuroscience and Biophysics, University of Texas Southwestern Medical Center, Dallas, TX, USA.

[3]Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA.

[4]These authors contributed equally: Corey A.H. Allard, Guipeun Kang, Jeong Joo Kim.

Abstract

Chemotactile receptors (CRs) are a cephalopod-specific innovation that allow octopuses to explore the seafloor via 'taste by touch'[1]. CRs diverged from nicotinic acetylcholine receptors to mediate contact-dependent chemosensation of insoluble molecules that do not readily diffuse in marine environments. Here, we exploit octopus CRs to probe the structural basis of sensory receptor evolution. We present the cryo-electron microscopy structure of an octopus CR and compare it with nicotinic receptors to determine features that enable environmental sensation versus neurotransmission. Evolutionary, structural and biophysical analyses show that channel architecture involved in cation permeation and signal transduction is conserved. In contrast, the orthosteric ligand-binding site is subject to diversifying selection, thereby mediating the detection of new molecules. Serendipitous findings in the cryo-electron microscopy structure reveal that the octopus CR ligand-binding pocket is exceptionally hydrophobic, enabling sensation of greasy compounds versus the small polar molecules detected by canonical neurotransmitter receptors. These discoveries provide a structural framework for understanding connections

Correspondence and requests for materials should be addressed to Ryan E. Hibbs or Nicholas W. Bellono. ryan.hibbs@utsouthwestern.edu; nbellono@harvard.edu.

Author contribution C.A.H.A., W.A.V-M. and N.W.B. contributed to the molecular, cellular and organismal studies. G.K., J.J.K, and R.EH. contributed to the biochemical and structural analyses. All authors were involved with writing or reviewing the manuscript.

Online content

Any methods, additional references. Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-05822-1.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-05822-1.

Peer review information *Nature* thanks Hugues Nury, Harold Zakon and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Competing interests The authors declare no competing interests.

between evolutionary adaptations at the atomic level and the emergence of new organismal behaviour.

Chemosensation is a widespread sensory modality found across the tree of life[2]. Chemosensory systems specialize to detect molecules relevant in specific ecological and behavioural contexts. For example, the main evolutionary transition of aquatic to terrestrial life has been proposed to necessitate the detection of hydrophobic airborne ligands instead of distant waterborne signalling via hydrophilic molecules[3,4]. Nonetheless, many aquatic animals use specialized chemotactile sensory systems to physically probe surfaces for insoluble chemical cues[14]. We recently uncovered a molecular basis for contact-dependent aquatic chemosensation by identifying octopus chemotactile receptors (CRs) as sensors of poorly soluble terpene molecules prevalent in marine ecosystems[1].

Octopuses are incredible sensory specialists that have evolved to thrive in their seafloor environment by using their arms to find prey in crevices inaccessible to their traditional sense organs[5,6]. Octopuses 'taste by touch' with cephalopod-specific CRs found in the sensory epithelium of arm suckers, which the octopus uses to probe surfaces (Fig. 1a). CRs assemble as homo and heteropentameric ligand-gated ion channels, homologous to 'Cys-loop' ionotropic neurotransmitter receptors[1]. Among these receptors, CRs share the highest sequence identity with nicotinic acetylcholine receptors (nAChRs). Yet, CRs have diverged from nicotinic receptors to detect, transduce and integrate distinct and diverse chemical stimuli including hydrophobic terpenes, many of which are classical volatile compounds important for terrestrial olfaction[1]. Indeed, other chemosensory receptors diverged from neurotransmitter receptors, including protostome ionotropic receptors that are related to ancestral glutamate receptors, suggesting that similar transitions could facilitate sensation in a variety of systems[7,8]. Therefore, CRs represent a striking example of biological novelty and a unique and powerful system for understanding principles of sensory receptor evolution and adaptation in the context of a lineage-specific trait. Considering the adaptive nature of octopus chemotactile sensation, we exploited CRs using structural, biophysical and behavioural app roaches to ask how a new nAChR sensory system emerges to facilitate new behaviour.

Cephalopod CRs are divergent sensory receptors

In *Octopus bimaculoides,* 26 intronless CRs are organized in a tandem array on chromosome 15, suggesting that non-homologous recombination and retrotransposition have contributed to the expansion of this family of new sensory receptors from an ancestral nicotinic receptor-like gene (Fig. 1b, Extended Data Fig. 1 and Extended Data Table 1)[9]. Although retrogenes are often non-functional pseudogenes that evolve neutrally, CRs have complete open reading frames, are expressed and translated into proteins and are under functional constraint as indicated by omega ratios lower than 1 (ref.[10]). Higher evolutionary rates observed in CR sequences compared to nicotinic receptors suggest that these receptors exhibit structural modifications that mediate sensory specialization versus neurotransmission (Fig. 1c). To explore the structural basis of sensory receptor evolution in octopus, we sought to compare structure-function relationships of recently discovered octopus CRs to the well-studied α7

nicotinic receptor. As a Foundational comparison, we chose to analyse the best characterized C R to date, chemotactile receptor for terpenes 1 (CRT1) (previously CR518). Notably, CRT1 localizes to the apical dendritic endings of native receptor cells (Fig. 1a and Extended Data Fig. 2a) and we have previously characterized it as a receptor for poorly soluble terpenes, which are chemically diverse molecules that are ecologically relevant to the octopus's chemotactile sense[1].

CRT1 architecture

To examine how the structural properties of CRT1 facilitate its adaptive function as a sensory receptor, we first determined its structure using single-particlecryo-electron microscopy (cryo-EM). We expressed and purified recombinant full-length CRT1, then made samples for cryo-EM in the mild detergent glyco-diosgenin(GDN). The cryo-EM reconstruction reached 2.6 Å overall resolution, with symmetric transmembrane domains (TMDs) and extracellular domains (ECDs) (Fig. 1d,e. Extended Data Figs. 2b and 3 and Extended Data Table 2). Like the α7 nicotinic receptor, CRT1 assembles to form a homopentameric channel complex with a central ion pore (Fig. 1d,e and Extended Data Figs. 3 and 4). Indeed, each subunit of CRT1 shares a general architecture consistent with other members of the Cys-loop receptor family. The large extracellular domain consists of a short N-terminal α-helix and ten (β-strands. The eponymous Cys-loop disulphide that connects β6-β7 is well resolved, as is 3 new disulphide near the classical Cys-loop receptor neurotransmitter site that connects β4-β7 to increase rigidity ofthe extracellular domain (Fig. 1f). Single and double alanine mutations of this second disulphide bond resulted in non-functional receptors (not shown), suggesting that this interaction is important for receptor stability or channel gating. The hydrophobic transmembrane domain exhibits clear density for M1-M3 helices (Fig. 1d,f and Extended Data Fig. 4a,e-g). Density for an M4 helix was absent in all membrane conditions tested (Extended Data Fig. 3a-d); however, it is predicted to be present by primary sequence hydrophobicity and secondary structure using AlphaFold (Extended Data Fig. 5)[11]. CRT1 has a permeation pathway similar to α7, with a comparatively open ECD vestibule lined by charged residues that could facilitate conductance tuning and a comparatively constricted TMD pore, lined by M2 helices, to gate the channel (Fig. 2a). Based on this architectural comparison, we hypothesized that the transition from neurotransmission to sensation requires broad conservation of receptor structure with divergence concentrating in functionally important areas. To explore this hypothesis, we investigated key receptor features that appeared conspicuously conserved or divergent between CRT1 and α7.

Conserved elements for signal transduction

Ion permeation is a critical feature of channels, dictating downstream signalling events such as initiation of specific neural activity. In α7 homomeric channels, but not other nicotinic receptors, a key glutamate residue E97 forms an anionic ring in the extracellular vestibule that plays a major role in the high calcium permeability of α7 (ref.[12]). The structure of CRT1 reveals that E104 occupies a similar position in the vestibule and forms its tightest construction, suggesting it could similarly influence calcium permeation (Fig. 2a-c). Indeed,

mutation of E104to alanine or lysine dramatically reduced calcium permeability but did not affect permeation of other ions (Fig. 2d,e). Beneath the outer vestibule, the transmembrane region of CRT1 forms a pore lined with hydrophobic residues in its extracellular half, which form the activation gate in α7 and in the muscle-type nicotinic receptor[12–14]. The intracellular half of the pore is lined by polar residues, culminating in a ring of anionic glutamates at the cytosolic mouth; these are determinants of cation versus anion selectivity among pentameric ligand gated ion channels. The CRT1 pore conformation is similar to the putative activated state pore of α7 (ref.[12]), with largely constant diameter from top to bottom and a minimum diameter of 6.6 Å (Fig. 2c and Extended Data Fig. 6e-i). Together, the structure of CRT1 reveals features in common with the α7 nicotinic receptor in determinants of Ca^{2+} permeation, pore diameter and chemistry.

Adaptations for sensory system function

Whereas comparisons with nicotinic receptors reveal a conserved role in signal transduction and calcium signalling, we reasoned that other structural features must underlie CRT1's adaptive function in environmental sensation. We previously determined that CRT1is activated by poorly soluble terpenes, which are adiversegroup of naturally occurring hydrocarbons produced by plants, fungi and many microbes, and often serve semiochemical roles in interspecies communication[3,4]. CRT1 agonists include the sesquiterpene polygodial and the furanoses-quiterpene atractylon, which are secreted by marine invertebrates as chemical deterrents to predation[15,16]. To determine the ligand-binding site in CRT1, we purified receptors in the presence of the terpene agonist costunolide. Because CRs were sensitive to terpenes but not ACh, and the α7 nicotinic receptor exhibited reverse ligand sensitivity (Fig. 3a), we expected to visualize costunolide bound to a structurally distinct region in CRT1. To our surprise, whereas the channel was in an open conformation suggesting an activated state (Extended Data Fig. 6), we did not visualize a density corresponding to costunolide. Instead, we observed the diosgenin moiety of the detergent GDN, which was used to stabilize CRT1 for structural analysis (Extended Data Fig. 4). Diosgenin was curiously bound to a region corresponding to the canonical neurotransmitter site of α7 (Fig. 1d and 3b,c, Extended Data Fig. 4 and Supplementary Video 1). This site is defined by three stretches of discontinuous amino acids, or loops, from the two subunits forming the interface: A, B and C from the principal side, which contributes loop C, and D, E and F on the complementary side[17]. In CRT1, loop C is shortened by six residues (Fig. 3b-d, Extended Data Fig.4b, c and Extended Data Fig. 6a,b), losing the aromatics and vicinal cysteines important in α7 and other nicotinic receptors that facilitate binding of nicotinic agonists. The result, in CRT1, is a less well-defined agonist pocket and instead a more hydrophobic, flatter surface, which cannot bind ACh (Fig. 3c and Extended Data Fig. 4b, c). Instead, diosgenin packs against this flat hydrophobic surface at the subunit interface, making most of its contacts with the complementary subunit (Fig. 4a and Extended Data Fig. 4d). This finding contrasts with agonist recognition in nicotinic receptors, in which interactions with the principal subunit are dominant.

Perhaps CRT1 is so well adapted to bind 'greasy' molecules that we unintentionally discovered an agonist site by resolving the bound steroid-like diosgenin moiety, which also

resembles known CR terpene agonists. Therefore, we tested whether GDN and structurally similar molecules including oestrogenic steroids and pentacyclic triterpenoids act as agonists of heterologously expressed CRT1. Strikingly, many of these compounds activated CRT1 with even higher potency than previously identified terpene agonists, elicited minimal desensitization with moderate pore block at high concentrations and responses could be blocked by the CR channel blocker mecamylamine (Fig. 3e,f and Extended Data Fig. 7a,b). Additionally, several of these compounds elicited axial nerve activity from amputated octopus arm tips in response to agonists perfused on arm suckers. We observed robust, slowly desensitizing responses to fish extract, the sesquiterpenoid nootkatone and the oestrogenic mycotoxin zearalenone but not steroid-like molecules or non-CR agonists, such as amino acids (Fig. 3g and Extended Data Fig. 7c). During nerve measurements, we noticed that active agonists also frequently induced autonomous behavioural responses with similar response profiles (Fig. 3h and Extended Data Fig. 7c). The octopus's distributed nervous system permits each arm to integrate environmental signals and carry out autonomous behaviours, even when removed from the rest of the organism[18,19]. Indeed, in one of our experiments, the arm responded to agonist by crawling off the recording pipette and out of the bath, emphasizing the high degree of arm autonomous behaviour elicited by the chemotactile sense. Whereas the valence of such autonomous arm responses is unclear, CRs detect both prey extracts and ligands that are aversive for other chemosensory systems, raising the possibility that these receptors mediate sensation of both appetitive and aversive cues. Collectively, these results suggest that CRT1 has adapted to bind poorly soluble molecules and transduce diverse information to the distributed nervous system. Such properties are well suited to the octopus's ability to 'taste by touch' using contact-dependent chemosensation and relatively autonomous arm behaviours.

Evolution of the agonist binding site

How do specific structural features evolve to mediate new functions? To address this question, we analysed which regions of CRT1have been subjected to diversifying selection using codon aligned CR sequences. Interestingly, residues under high positive selection are concentrated in the ECD agonist binding site (Loops A-F (Extended Data Fig. 8). Several of the same hydrophobic residues that constitute the 'greasy' CRT1 binding site and coordinate diosgenin in the structure are among those experiencing the strongest diversifying selection (Fig. 4a,b). Mutation of these residues to alanine not only altered agonist sensitivity but resulted in constitutive channel activity, emphasizing the functional connection between binding site and channel gating (Extended Data Fig. 7d). Notably, mutant receptors were inhibited by the CR channel blocker mecamylamine, indicating that mutations specifically perturb the agonist binding site and channel gating (Fig. 4c,d and Extended Data Fig. 7e). Thus, divergence at the orthosteric ligand-binding domain is critical to the adaptive transition from neurotransmission to sensation.

Collectively, our analyses demonstrate how diversifying selection at the agonist binding site, in addition to conservation of key receptor features, contributed to the evolution of a lineage-specific cephalopod sensory adaptation. The chemotactile sense appears to have evolved in parallel with an elaboration of the cephalopod nervous system and body plan relative to basal molluscs. In a companion study, we explored CR-mediated chemotactile

sensation across cephalopods to find that tuning at the orthosteric ligand-binding region is a conserved and flexible theme that facilitates diverse predatory strategies. Future work will leverage comparisons of adaptive protein structure along with natural product chemistry to understand how sensory systems evolve in the context of their specific ecological niche. Thus, our work represents a foundational basis for understanding how sensory systems adapt to facilitate the lifestyles of diverse organisms specialized to interact with their unique environments.

Methods

Animals

Adult California two-spot octopuses (*O.bimaculoides*) were wild-caught mate and female adults (1–2 years old, Aquatic Research Consultants), fed daily with fiddler crabs (*Leptuca pugilator*, Northeast Brine Shrimp) and kept on a 12-h light–dark cycle in natural seawater. Animal protocols were approved by the Harvard University Animal Care and Use Committee (protocol ID: 18-05-325).

Molecular biology and cell culture for physiology

The genes used in this study were: octopus *Crt1* (LOC106880828), and human *CHRNA7* (NM_000746) and *TMEM35* (NM_021637). Mutagenesis of pUNIV-ObCRT1 to create point mutants and epitope-tagged variants was performed by GenScript. Acustom monoclonal antibody targeting the extracellular domain of CRT1 (peptide: YGDERLLREKLLTNYSKSIRPVINLTKVVDVTALLYLQTLYDLDFVNNFIMARYYLG LIW1DEKLTWNPLDYNNITSIYLPKDKIWTPPIKMCNSMDKSEENDGVGELMLTYTG WINMWSFRLLHTYCQINAYTYPFDEHTCEIYLCVALHTINHTRIKELLYEDSKFTQNY KWDINVSGKVNGTDELFSYAFAPMYLRRKLTV) was produced in hybridomas generated from an immunized mouse by GenScript.

HEK 293 cells (authenticated and validated as negative for *Mycoplasma* by the vendor ATCC) were cultured in DMEM (Gibco) supplemented with 10% FCS (Gibco) and 50 IU ml^{-1} penicillin and 50 µg ml^{-1} streptomycin (Gibco) at 37 °C and 5% CO_2 using standard techniques. For transfection, HEK 293 cells were washed with Opti-MEM Reduced Serum Medium (Gibco) and incubated with transfection mix containing plasmid DNA and 4 µl Lipofectamine 3000 Transfection Reagent (Invitrogen) in Opti-MEM for 4–8 h at 37 °C. Cells were then replated onto glass coverslips in DMEM, incubated for 2 h at 37 °C and then moved to 30 °C incubation overnight. For CRs, 1 µg plasmid DNA was used for transfection. For α7, 1 µg receptor plasmid plus 0.5 µg of the chaperone TMEM35 were cotransfected. For the patch clamp experiments, an additional 0.3 µg green fluorescent protein plasmid was cotransfected to facilitate identification of transfected cells.

CRT1 expression and purification

CRT1was subcloned into the pEZT-BM expression vector[20] and appended with a serine linker foliowed by a single Strep tag at the C terminus for affinity chromatography. CRT1bacmam virus was produced in Sf9 cells (catalogue no.CRL-1711, ATCC) and titrated as described previously[20] for large-scale protein expression. A total of 4.81 of suspension

HEK293S GnTI^{-} (catalogue no. CRL-3022, ATCC) cells were grown at 37 °C and 8% CO_2 until they reached a cell density of approximately 4×10^6 cells per millilitre. Then, 3 mM sodium butyrate (Sigma-Aldrich) and virus were added to the culture during transduction; then, the temperature was changed to 30 °C after transduction with a multiplicity of infection of 0.5. Transduced cells were collected after 72 h by centrifugation, resuspended in TBS (20 mM Tris, 150 mM NaCl, pH 7.4) containing 1 mM phenylmethanesulfonyl fluoride (Sigma-Aldrich), then lysed using an Avestin Emulsiflex. Lysed cells were centrifuged for 20 min at 10,000*g* and the supernatant was collected and centrifuged again for h at 186,000*g*. The membrane pellet was mechanically homogenized, then solubilized for 1 h at 4 °C in TBS buffer with 40 mM *n*-dodecyl β-D-maltoside (DDM, Anatrace). Solubilized membranes were centrifuged for 40 min at 186,000*g*, then passed over high-capacity Strep-Tactin affinity resin at 0.8 ml min^{-1}. The resin was washed with TBS buffer containing 0.1 mM GDN then eluted with wash buffer supplemented with 5 mM D-desthiobiotin (Sigma-Aldrich). After affinity chromatography, the eluted fractions were combined, concentrated, then centrifuged for 20 min at 40,000 rpm to remove aggregates before size-exclusion chromatography (SEC). The resulting supernatant was separated by SEC using a Superose 6 Increase 10/300 GL column (Cytiva) equilibrated inTBS buffer containing 0.1 mM GDN (Anatrace). SEC fractions corresponding to pentameric CRT1 were pooled and concentrated and the quality of receptor protein was assessed by tryptophan fluorescence SEC before freezing the grids.

Cryo-EM sample preparation

A total of 3 μl of concentrated receptor protein (A_{280}= approximately 7) were applied to copper Rl.2/1.3200 mesh holey carbon grids (Quantifoil) that were glow-discharged at 30 mA for 80 s. The grids were immediately blotted for 3 s under 100% humidity at 4 °C and then plunge-frozen into liquid ethane cooled by liquid nitrogen using a Vitrobot Mark IV (Thermo Fisher Scientific).

Single-particle cryo-EM data acquisition and processing

Cryo-EM samples were screened on the Talos Arctica at University of Texas Southwestern Medical Center and data collection was performed at the Pacific Northwest Center for Cryo-EM. Micrographs were collected on a 300 kV Titan Krios equipped with a K3 direct-eiectron detector (Gatan) and a GIF quantum energy filter (20 e^{-}V) in super-resolution mode. Images were collected with a total electron dose of 50 e^{-}/Å^2 over 50 frames and a defocus range of −0.5 μm to −2.5 μm.

Data processing was performed by following a general workflow in Relion v.3.1 (ref.[21]). Dose-fractionated images were gain-normalized, Fourier binned twice, aligned, dose-weighted and summed with MotionCor2 (ref.[22]). The contrast transfer function correction and defocus values were estimated using GCTF[23]. Particles were picked using SPHlRE-crYOLO[24], then subjected to two-dimensional classification to remove junk particles. An ab initio model was generated and lowpass-filtered to 60 Å to use as a reference for the first round of three dimensional (3D) classification with Cl symmetry. After 3D refinement, CTF refinement and Bayesian polishing, final 3D classification with local angular searches and C5 symmetry yielded two good 3D classes. The selected 3D classes were combined

and further refined with C5 symmetry; subsequently, the resulting volume was locally sharpened in Relion[21], yielding a final reconstruction at 2.62-Å resolution. Map quality was evaluated with Phenix mtriage[25] and the resulting Fourier shell correlation curve is included in Extended Data Figure 3 Local resolution was estimated with ResMap[26]. Software packages for structural biology were compiled by SBGrid[27]. Cryo-EM data are summarized in Extended Data Table 2.

Model building, refinement and validation

An initial homology model was generated using the SwissModel[28] server based on the β2 subunit of the human α4 β2 nAChR (Protein Data Bank (PDB) ID:5KXI). This model was fitted into the density map, then rebuilt with ligands including glycans into the EM density map using Coot[29]. What we refer to as diosgenin is the ordered diosgenin component of GDN; we modelled diosgenin plus the *O*-Iinked polyether branch, based on EM density. The chemical structure of what we referto as diosgenin was adapted from the PDB (chemical ID: DU0). Subsequently, the model was iteratively refined using global real-space refinement, in Phenix[30,31] and manual adjustment in Coot. The stereochemical statistics of the final model, including Ramachandran analysis, were assessed by Molprobity[32] and the final refinement statistics, including model-map correlation and half-maps, were calculated using the comprehensive validation and mtriage analyses in Phenix[30,31].

Figures and graphical illustration

UCSF Chimera[13], UCSF ChimeraX[34] and Pymol[35] were used for graphical illustration of protein structures and density maps in figures. Protein sequences were retrieved from the UniProtKB database[36] and aligned using PROMALS3D[37]. Pore radius profiles were calculated using Hole[38] and plotted using Prism v.9.2.0 for Windows (GraphPad Software). Root mean square deviation values were calculated using the MatchMaker function in Chimera. The hydrophobicity profile along the permeation pathway was calculated with CHAP[39]. The interface area and the solvation free energy of protein-protein and protein-ligand interfaces were calculated by PDBePlSA[40] and the solvent accessibility of the agonist binding pocket was analysed using CASTp3.0 (ref.[41]). The AlphaFold model of the CRTI monomer was predicted using AlphaFold Colab[11].

Patch clamp electrophysiology

Patch clamp recordings were carried out at room temperature using a MultiClamp 700B amplifier (Axon Instruments) and digitized using a Digidata 1550B (Axon Instruments) interface and pClamp software (Axon Instruments). Whole-cell recording data were filtered at 1 kHz and sampled at 10 kHz. For whole-cell recordings in HEK 293 cells, pipettes were 3–5 MΩ. The standard extracellular solution contained): 140 mM NaCl, 5 mM KCI, 10 mM HEPES, 2 mM $CaCl_2$,2 mM $MgCl_2$, pH 7.4. The intracellular solution contained: 140 mM Cs^+ methanesul fonate, 1 mM $MgCI_2$,5 mM NaCl, 10 mM CsEGTA, 10 mM HEPES, 10 mM sucrose, pH 7.2. Agonist-evoked currents were measured at −110 mV during 500-ms ramps from −120 to 80 mV. Time courses displayed currents measured at −110 mV during successive ramps. Perfusion experiments were performed using a SmartSquirt Micro-Perfusion system (Automate Scientific) pressurized to approximately 30

kPa. In the ion substitution experiments, relative permeability was determined by measuring the shift in E_{rev} after the substitution of equimolar monovalent cations. E_{rev} was extracted from voltage ramp experiments as described for measuring agonist-elicited currents. The extracellular solution contained: 150 mM Na^+ Cs^+, *N*-methyl-D-glucamine$^+$ or 100 mM Ca^{2+} Mg^{2+}; the intracellular solution contained 150 mM CsCl and 1 mM CsEGTA. Solutions were buffered with 10 mM HEPES. Permeability ratios were estimated using the Goldman-Hodgkin-Katz equation: $P_X/P_{Na} = \left(\left(Na^+\right)\text{Luminal}/(X)_{\text{Cytoplasmic}}\right)(\exp(E_{rev}F/RT))$, $P_{Ca}/P_{Na} = \left(4\left(Ca^{2+}\right)_{\text{Luminal}}/\left(Na^+\right)_{\text{Cytoplasmic}}\right)(\exp(E_{rev}F/RT))$.

Two-electrode voltage clamp

Defolliculated oocytes were purchased from Ecocyte or Xenopusl and stored in modified Barth's solution (88 mM NaCl, 1 mM KCI, 5 mM Tris-HCl, 1 mM $MgSO_4$, 0.4 mM $CaCl_2$, 0.33 mM $Ca(NO_3)_2$, 2.4 mM $NaHCO_3$, pH 7.4) supplemented with 0.1 mg ml^{-1} gentamycin for up to one week at 4 °C until use. For oocyte expression, plasmid pUNIV-ObCRT1 was linearized using Not1-HF (catalogue no. R3189, New England Biolabs) for 2 h at 37 °C. Linearized DNA was purified using a PCR purification kit (catalogue no. 28104, QIAGEN) and eluted in 30 µl RNase-free water. RNA synthesis was performed with 1–3 µg DNA using the mMESSAGE mMACHINE T7 Transcription Kit including 15 min of DNase treatment (catalogue no. AM1344, Ambion). RNA was treated with a Zymo Clean & Concentrator Kit and a liquoted at a concentration of approximately 10 µg ml^{-1} for injection. Oocytes were injected with 25 ng RNA using a Nanoject III (Drummond Scientific) and incubated in modified Barth's solution at 17 °C overnight. Two-electrode voltage clamp recordings were carried out at room temperature with an Oocyte Clamp OC-725C amplifier (Warner Instruments) and digitized using a Digidata 1550B interface and pClamp 11 software. Data were filtered at 1 kHz and sampled at 10 kHz. Recordings were performed using borosilicate glass pipettes with resistances of 7–10 MΩ when filled with 3 M KCI. All chemicals were diluted in ND96 extracellular solution (96 mM NaCl, 2 mM KCI, 5 mM HEPES, 1 mM $MgCI_2$, 2 mM $CaCl_2$ adjusted to pH 7.4 with NaOH). Stimuli were delivered using a micropipette positioned in the bath. Stimulus-evoked currents were obtained using 200-ms voltage ramps from −120 mV to 120 mV applied every 500 ms with an interstimulus holding potential of −40 mV. Dose-response relationships were calculated using peak currents measured at −115 mV. For the chemical screen, response amplitudes were measured at −115 mV and normalized to the response to GDN, which was the largest.

Histology

Whole-mount histology.—*O. bimaculoides* hatchlings were fixed in 4% paraformaldehyde (PFA) in PBS for 16–24 h on a rocker and subsequently stored in PBS at 4 °C until use. Samples were then washed with phosphate-buffered saline with Tween 20 (PBST) (Triton X-100 0.1%) three times and blocked for 1 h in 10% normal goat serum (NGS) in PBST and antibody solution containing custom anti-CRT1 antibody (1:100) (Genscript) and anti-HRP-Cy3 (neural-specific marker, 1:500) (catalogue no. 123–545–021, Jackson ImmunoResearch) was applied for 24 h at 4 °C in NGS. Samples were then washed with PBST (Triton X-100 0.1%) three times; goat anti-mouse IgG H&L (Alexa Fluor 488) (catalogue no. abl50113. Abeam) secondary antibodies were added (1:500). Finally, samples

were washed 3–5 times in PBST, mounted in VECTASHIELD Antifade Mounting Medium (Vector Laboratories) and imaged with an AxioZoom V16 Zoom Microscope (ZEISS) or LSM 980 Confocal Microscope with Airyscan2 (ZEISS). Images were processed using Fiji (National Institutes of Health (NIH)).

HEK cells.—HEK cells were transfected as described above. Cells were washed three times in PBS and fixed In 4% PFA in PBS for 10 min at room temperature. Coverslips were then incubated in blocking mixture containing 1% BSA and 10 mM glycine in PBS. Samples were then incubated in antibody solution containing custom anti-CRT1antibody (Genscript) (1:100) and HA Tag Polyclonal Antibody (catalogue no. ab51064–2-AP, Proteintech) (1:250) for 24 h at 4 °C in 1% BSA. Samples were then washed with PBS three times and a secondaries antibody mix containing goat anti-mouse IgG H&L (Alexa Fluor 488) (1:500) and goat anti-mouse IgG H&L(Cy3) (catalogue no. ab97035. Abeam) (1:500) was added. Finally, samples were washed three times in PBS, mounted in VECTASHIELD Antifade Mounting Medium with DAPl (Vector Laboratories) and imaged with an LSM 980 Confocal Microscope with Airyscan2 (ZEISS). Images were processed using Fiji (NIH).

Autonomous arm behaviour

Arm tips from sedated animals were transferred to a 10-cm Petri dish containing 50 ml holding solution (and held in place using suction through an appropriately sized black tygon tubing). Odorants were first dissolved in dimethyl sulfoxide (DMSO) and then diluted in holding solution to 0.1–1% final DMSO; 1 ml stimulus at the indicated concentration was perfused over the arm over 2 s using a micropipette. Arm behaviour was recorded at 500-ms intervals using a FLIR grasshopper camera (Teledyne) equipped with a 35-mm Nikon DX AF-S N1KKOR 1:1.8G lens controlled using the Spinnaker SDK software (Teledyne). Videos were segmented and motion was measured using ImageJ (NIH) by measuring Z-stack projections after performing stack difference analysis. Total arm motion in response to each chemical was measured by summing the stack difference and values were normalized such that the largest response was set to 1 for each arm.

Axial nerve recording

Arm tips from sedated animals were transferred to a 10-m Petri dish containing 50 ml holding solution (430 mM NaCl, 10 mM KCI, 10 mM HEPES, 10 mM $CaCl_2$,50 mM $MgCl_2$,10 mM D-glucose, pH 7.6). Nerve recordings were performed using a borosilicate glass suction electrode shaped and polished to fit over the entire cut end of the radial nerve. A similar reference electrode was placed in the bath. Gap-free recordings were made with 10-kHz sampling at 10,000× gain and signals were highpass-filtered at 100 Hz and lowpass-filtered at 1 kHz using a Warner DP-311A headstage and AC/DC amplifier (Warner) and digitized using a Digidata 1440A digitizer (Molecular Devices) with ClampEx software (Molecular Devices). Recordings were processed using Clampfit (Molecular Devices). For quantification of responses, the absolute value of the signal was processed using a lowpass 25-Hz Gaussian filter, the baseline signal was subtracted and the response amplitude was integrated over the response area. For controls and stimuli where no response was observed, a similar response area was measured as for the arm agonists.

Agonists

CR agonists included ACh chloride (catalogue no. A9101, Sigma-Aldrich), atractylon (catalogue no. FA74011, Carbosynth), β-oestradiol (catalogue no. 2824, Tocris), cholesterol (catalogue no. C8667, Sigma-Aldrich), cortisone (catalogue no. C2755, Sigma-Aldrich), costunolide (catalogue no. SML.0417, Sigma-Aldrich), genistein (catalogue no, G6649, Sigma-Aldrich), glyco-diosgenin (catalogue no. GDN101. Anatrace), mecamylamine hydrochloride (catalogue no. 2843, Tocris), nootkatone (catalogue no. W316620, Sigma-Aldrich), oleanic acid (catalogue no. 05504, Sigma-Aldrich), PNU-120596 (catalogue no. 2498, Tocris), polygodial (catalogue no. 14979, Cayman Chemical Company), progesterone (catalogue no. P8783, Sigma-Aldrich), resveratrol (catalogue no. RS010, Sigma-Aldrich), ursolic acid (catalogue no, U6753, Sigma-Aldrich) and zearale none (catalogue no. 3975, Tocris). Insoluble chemicals were first reconstituted in DMSO and then diluted in ND96 to 0.1–1% final DMSO. Fish extract was prepared by btending 1 g frozen zebrafish per 2 ml of holding solution, followed by centrifugation at 3,000*g* for 15 min to remove debris.

Phylogenetics and selection analysis

We mined AChR-like sequences from predicted coding sequences from the genome (BioProject: PRJNA270931)[9] and tissue transcriptomes (Gene Expression Omnibus accession no. GSE156748)[1] of *O. bimaculoides.* We used CRT1 (NCBI: NW_014712376.1) and α7 receptor (NCBI: U40583.2) as queries for iterative search of closely and distantly related AChRs homologue sequences using jack HMMER[42] with an e-value of 10×10^{-5}. We then confirmed sequence annotations using blastn[43]. Sequences were clustered at 95% identity using CD-hit[44] to remove isoforms and duplicates and aligned with MAFFT v.7 (ref.[45]). The best-fitting model for the sequence alignment was estimated using Model Finder[46], followed by maximum likelihood tree inference in IQ-TREE v.2.0 (ref.[47]). We performed 100 independent runs with different values of perturbation strength in IQ TREE v.2.0 to ensure convergence and used the tree with the highest likelihood for subsequent selection analysis. Support for clades was calculated using ultrafast bootstrap approximationUFBoot2 (ref.[48]). Tree visualization was implemented using the package phytools in R v.4.2.1 (ref.[49]).

For selection analysis, we generated codon alignments using Pal2Nal[50], Hyphy[51] and MAFFT v.7. We then manually refined the alignments in Geneious Prime v.2022,2. (https://www.geneious.com) to prevent spurious signals due to highly divergent sequences. Only genes with complete coding sequences and unambiguous mapping to the genome were used. We estimated the ratio of synonymous and non-synonymous substitutions (ω) to measure the extent of selective pressure acting across AChR-like genes using PAML v.4.9 (ref.[52]). We compared the likelihood of a single ω across branches (MO model), branch-specific ω (free ratio), distinct ω partitioning the CR clades as foreground and the remaining AChR-like genes as background (two-ratio), as well as partitioning ω between the two major CR ciades (three-ratio). Three runs with different initial ω values (0.5,1.0,1.5) were conducted to check for convergence. We used likelihood ratio tests to determine the best-fitting models and branches and sites under stronger diversifying selection. To further characterize positive selection across sites in the CRclades, we used the more sensitive and robust MEME

model[53] from HyPhy, with a default threshold of significance of LRT = 3 for positively selected sites. Finally, to explore clustering patterns and the number of introns of CRs and closely related genes, we mapped AChR-like coding sequences to the chromosome-level assembly (BioProject: PRJNA808169)[54] using GMAF[55]. Evolutionary analyses are summarized in Extended Data Table 2.

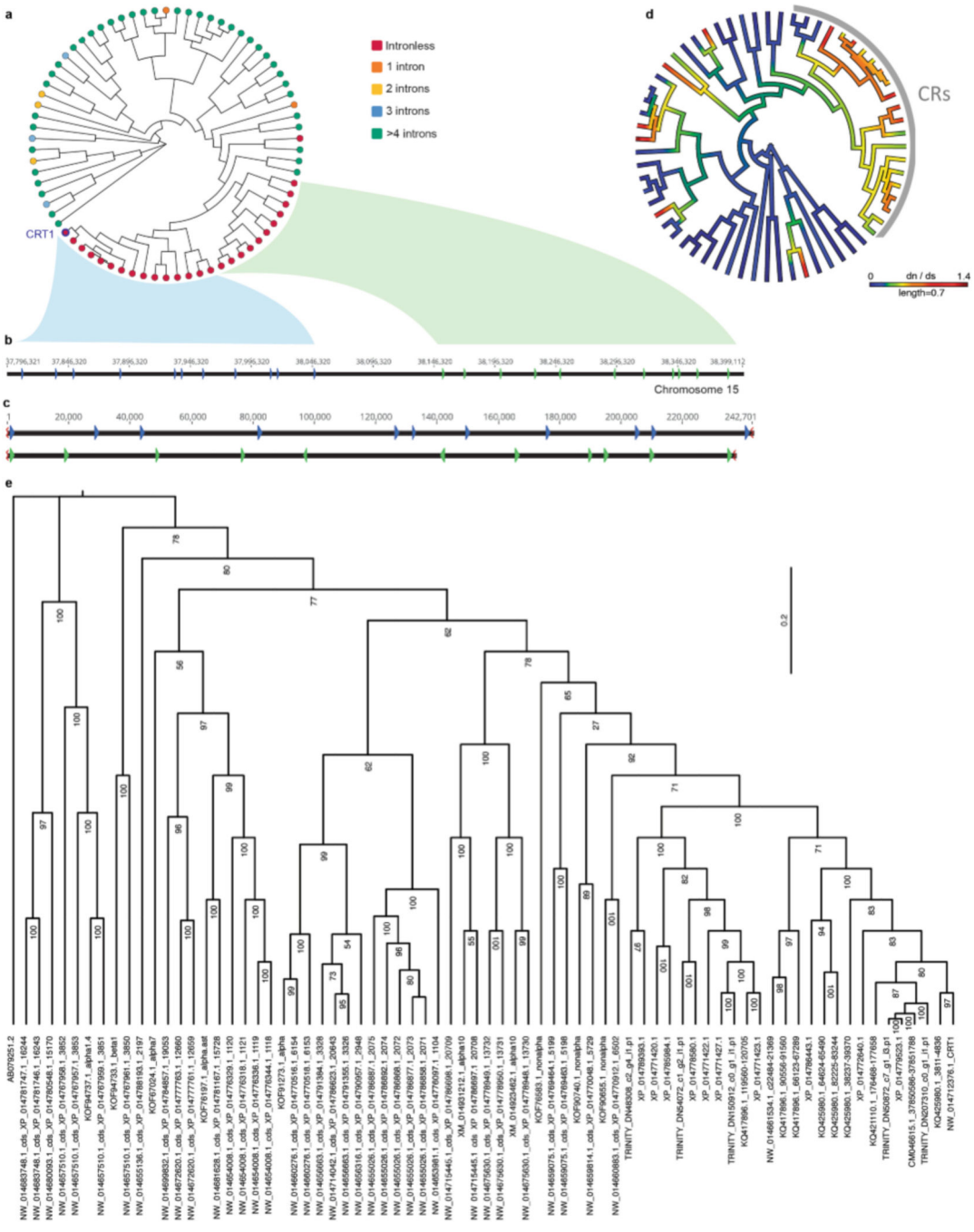
Quantification and statistical analysis

Data were analysed with Clampfit, Prism and are represented as the mean ± s.e.m. *n* represents independent experiments for the number of cells per patches or behavioural trials. Data were considered significant if $P < 0.05$ using paired or unpaired two-tailed Student's *t*-tests, Wilcoxon signed-rank test or one- or two-way ANOVAs. All significance tests were justified considering the experimental design; we assumed normal distribution and variance as is common for similar experiments. The animal behavior experiments were randomized and made blind to the chemical stimulation conditions, Sample sizes were chosen based on the number of independent experiments required for statistical significance and technical feasibility. Experiments were

Reporting summary

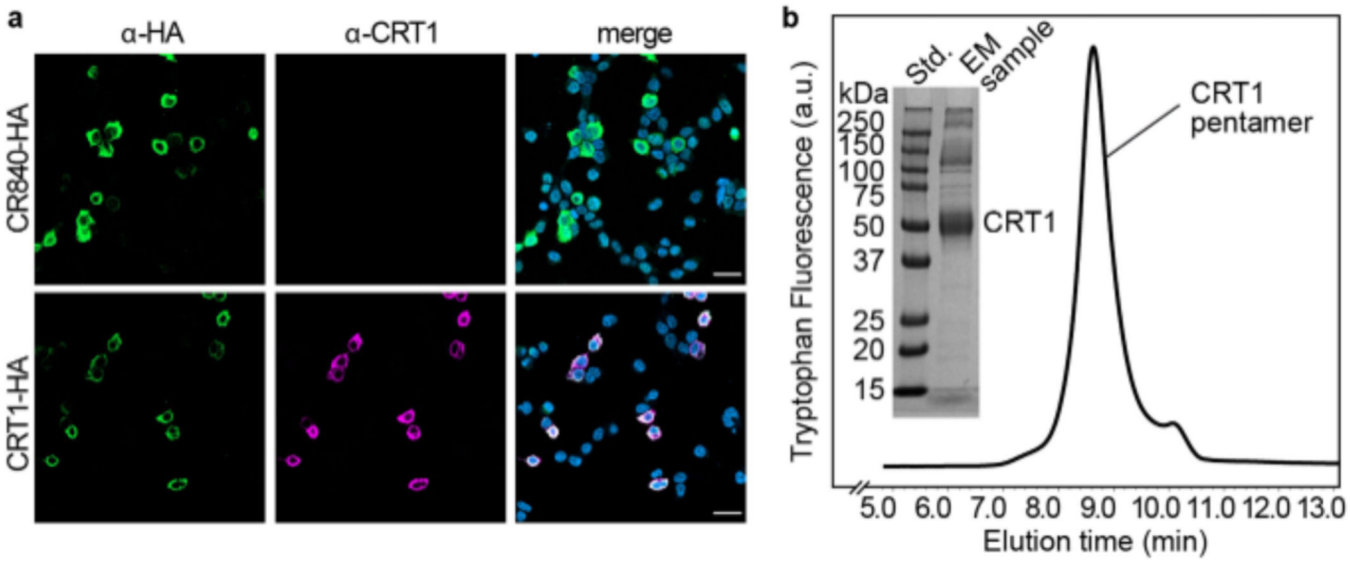
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Extended Data



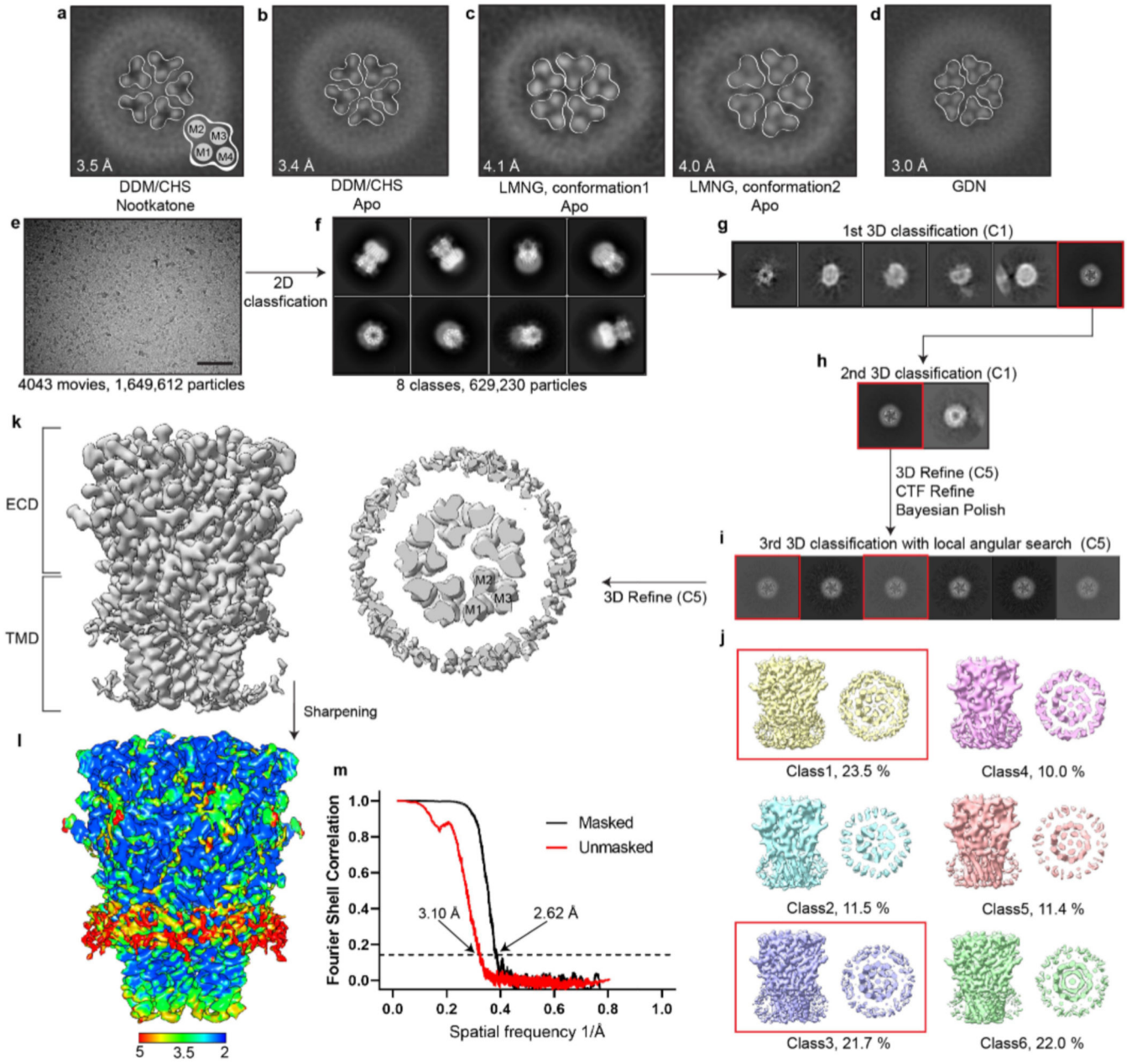
Extended Data Fig. 1 |. Octopus chemotactile receptors exhibit high rates of molecular evolution. **a**, Maximum likelihood treeof nicotinic receptor-like genes From *Octopus bimaculoides* showing the number of introns per gene. CRT1 is high lighted in the phylogeny.**b**, CRs are intronless genes arranged in tandem arrays in a single chromosome. The two main groups of CRs in octopus map to two distinct tight clusters in Chromosome 15. **c**, Alignment of the two arrays of intronless CRs in Chromosome 15 suggest diFFerences in overall architecture

between dusters. Only genes with complete coding sequences (CDS) are mapped. **d**, Phytogeny of acetylcholine receptor-like genes of *Octopus bimaculoides* depicting variation in dN/dS values for tip branches inferred underthefree-ratio model and mapped to the phylogeny. **e**, Phylogeny from Fig. 1b depicted as ultrametric and with bootstrap values.



Extended Data Figure 2 |. Antibody control and biochemistry.

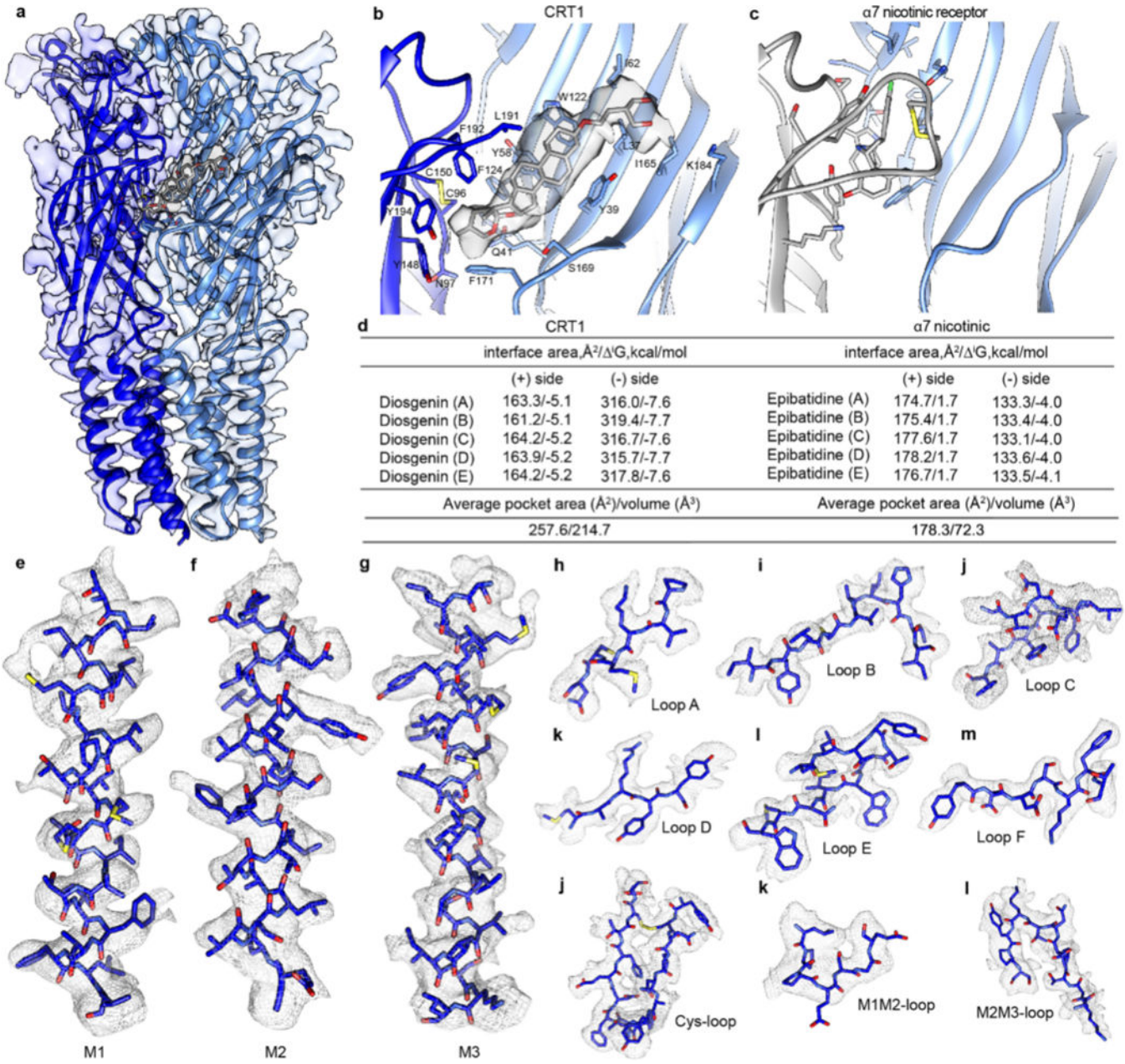
a, Anti-CRT1 is specific to expressed CRT1 as it did not label another CR (CR840) expressed in HEK293 cells. Representative of 3 independent transfections. Scale bar = 30 μm. **b**, Fluorescence-detection size-exclusion chromatography (FSEC) trace of CRT1 pentamer and SDS-page gel of final EM sample, representative of 4 independent purifications.



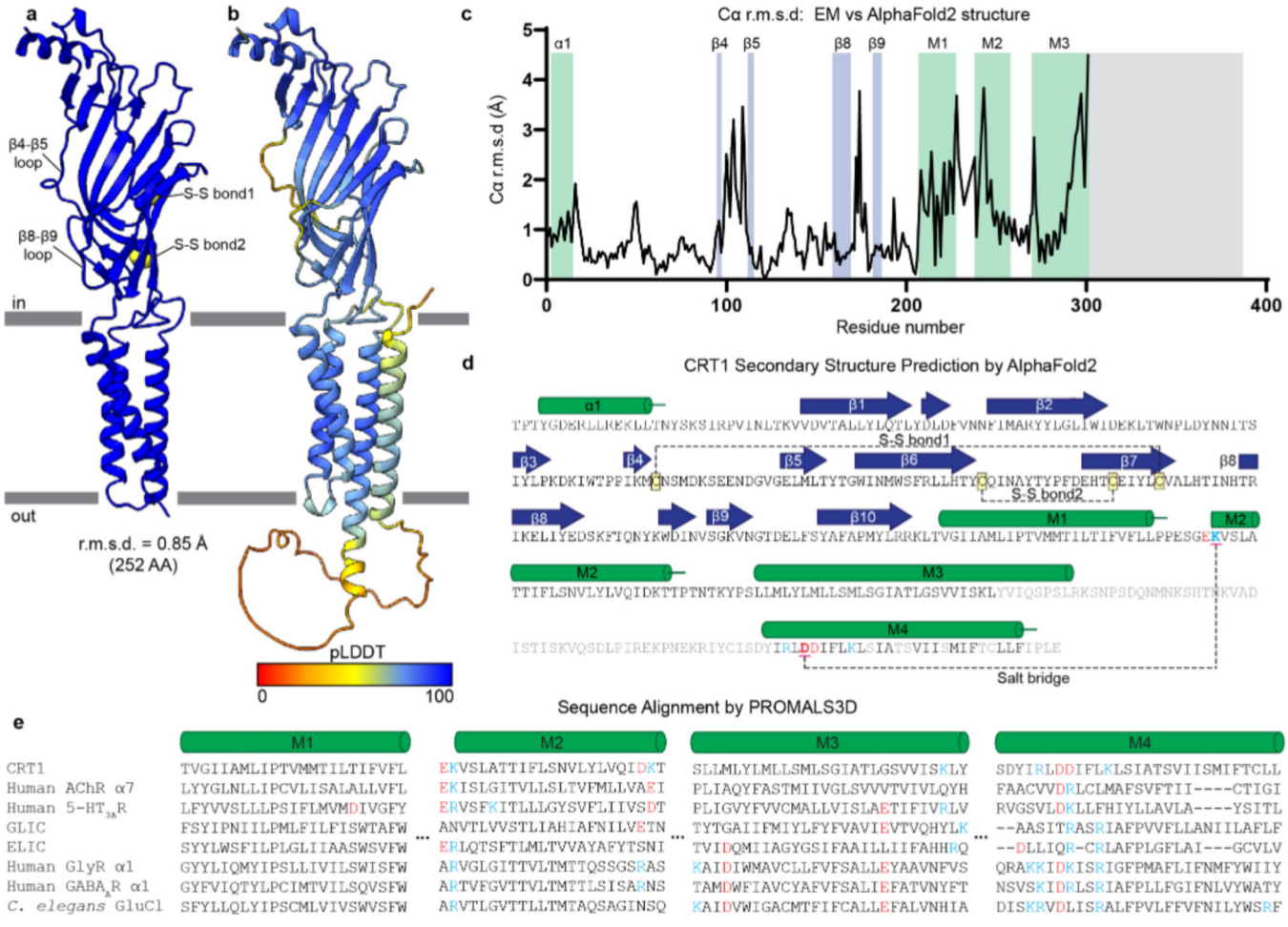
Extended Data Fig. 3 |. EM data processing.

a-d, TMD z-slices of 3D reconstructions from CRT1 in corresponding detergent conditions. Inset number indicates overall map resolution. A cartoon in panel **a** represents a schematic diagram of TMD helices, in initial purifications in DDM, both in the absence and presence of the terpenoid agonist nootkatone, the transmembrane domain (TMD) was poorly resolved, and the predicted fourth TMD helix, M4. was entirely absent. We thus collected EM datasets in two other detergents, lauryl maltose neopentyl glycol (L-MNG), and glyco-dlosgenin (GDN). L-MNG stabilized a n asymmetric TMD conformation, while GDN resulted in a well ordered and symmetric TMD, and the best overall resolution. Attempts to reconstitute the receptor in lipid nanodiscs resulted in profound aggregation. **e**,

Representative cryo-electron micrograph of CRT1 in GDN detergent micelle from dataset of 4043 dose-fractionated micrographs. Scale bar indicates 100 nm. **f**, Projection images from the final selected 2D classes. **g-i**, 3D classification results; good classes selected for further processing are boxed In red. **j**, 3D reconstructed maps from the final 3D classification, which are shown in side-view and top-view from ECD. Selected 3D classes are boxed in red to generate a final 3D map. **k**, Unsharpened 3D map where M1-M3 TMD helices from an individual subunit are labeled in black text. **l**, Sharpened map colored by local resolution. **m**, Half map FSC plot for masked and unmasked maps with resolutions indicated at FSC = 0.143.

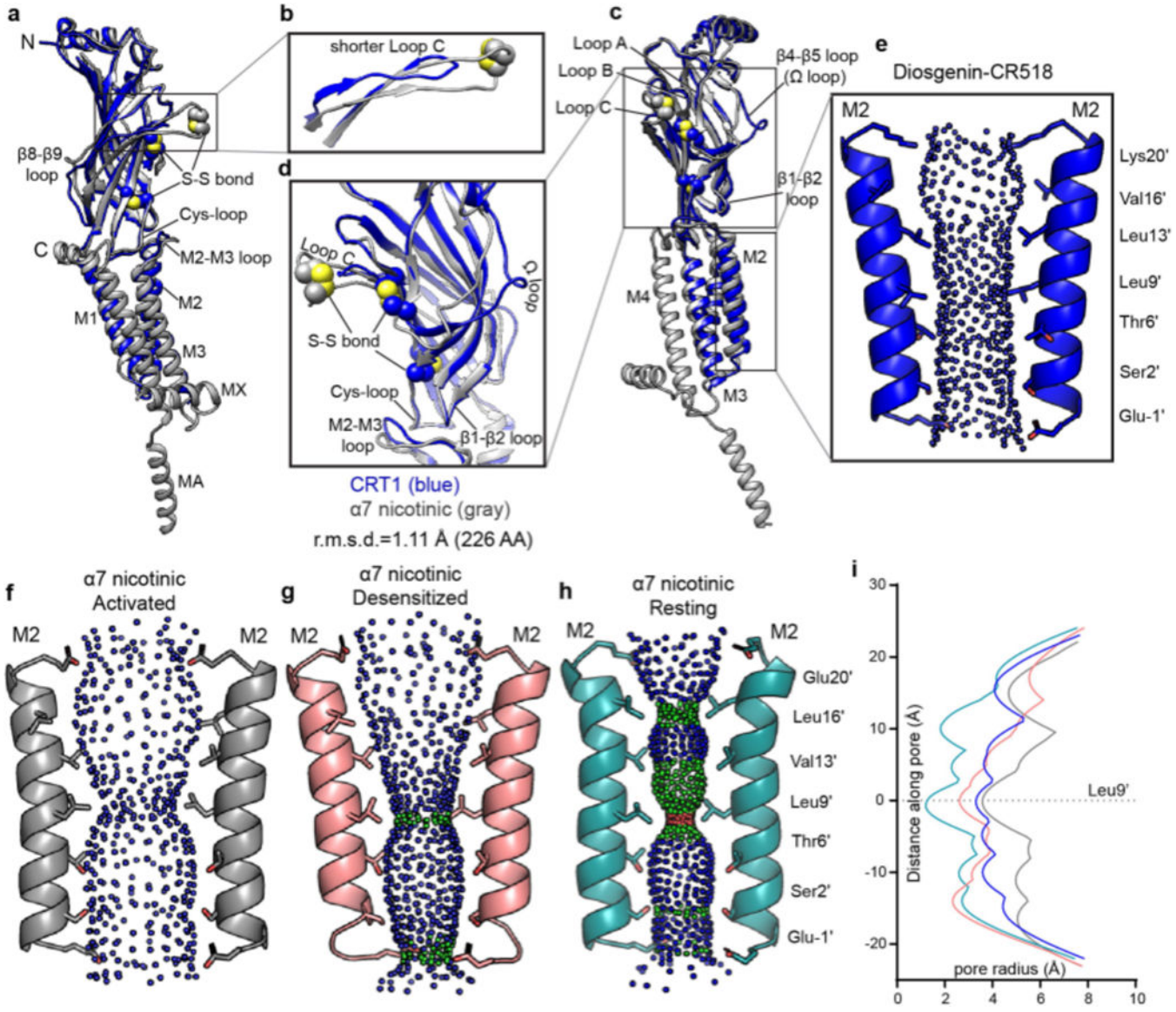


Extended Data Fig. 4 |. Cryo-EM density of the CRT1 receptor.
a, Cryo-EM density map of the CRT1 receptor tor representative adjacent subunits colored in blue and cyan. Density map of diosgenin colored in gray at a threshold level of 0.03. **b**, Orthosteric binding site of CRT1, where residues within 5 Å of diosgenin are shown as sticks. **c**, Orthosterlc binding site of α7 nicotinic receptor (PDB:7KOX), where residues within 5 Å of epibatidine are shown as sticks. **d**, Calculated interface areas and interaction energies (Δ^1G) for protein and diosgenin using PDBePISA[40]. Calculated solvent accessible area and volume of the binding pocket for CRT1 and α7 nicotinic receptor using CASTp3.0[41]. **e-g**, Cryo-EM density segments of M1-M3 helices at a threshold level of 0.02. **h-m**, Cryo-EM density segments of Loop A-F in the orthosteric binding site at a threshold level of 0.02. **j-l**, Cryo-EM density segments of Cys-loop, M1M2 loop, and M2M3 loop at a threshold level of 0.02.



Extended Data Fig. 5 |. CRT1 Secondary structure prediction from AlphaFold2.
a, A subunit of CRT1 cryo-EM structure colored in dark blue. **b**, Predicted AlphaFold2 CRT1 monomer colored by per-residue confidence score (pLDDT). Regions of high expected accuracy are colored in blue: regions of low expected accuracy are colored in red. **c**, Cα r.m.s.d between the CRT1 cryo-EM structure and the AlphaFold2 model (monomer comparisons) as a function of residue number. Helical regions are shown **in** green boxes.

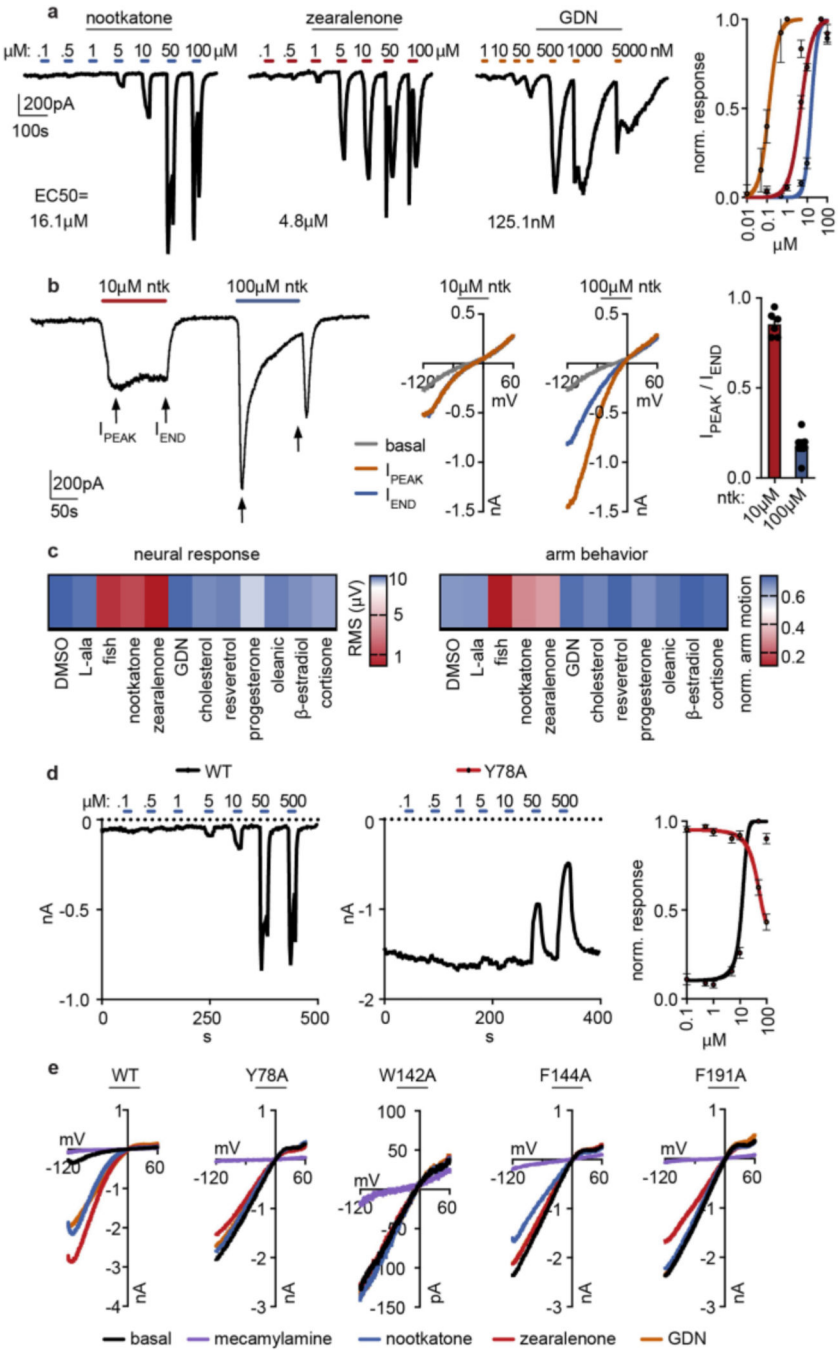
Blue boxes indicate β-strands next to the most divergent β4-β5 loop and β8-β9 loop regions. Unmodeled post-M3 helix is shaded in gray. **d**, Seque nee of CRT1 and secondary structure prediction by AlphaFold2. Helices are shown in green cylinders and strands are indicated by blue arrows. M4 helix amino acids are colored in red, blue, and black for negatively charged, positively charged, and hydrophobic residues, respectively. S-Sbond2 is in the Cys-loop. **e**, Sequence alignment of TMD helices (M1-M4) for CRT1 and other Cys-foop receptors. Same color scheme is followed as in Panel **d**.



Extended Data Fig. 6 |. Subunit and transmembrane pore conformation of CRT1 compared to the human α7 nicotinic acetylcholine receptor.

a, Comparison of single subunit structure of CRT1 and the human α7 nicotinic receptor (PDB:7KOX); MA and MX helices generally conserved in nicotinic receptors are absent in the CRT1 structure. **b**, Comparison of Loop C boxed in **a**; Disulfide bond on Loop C of

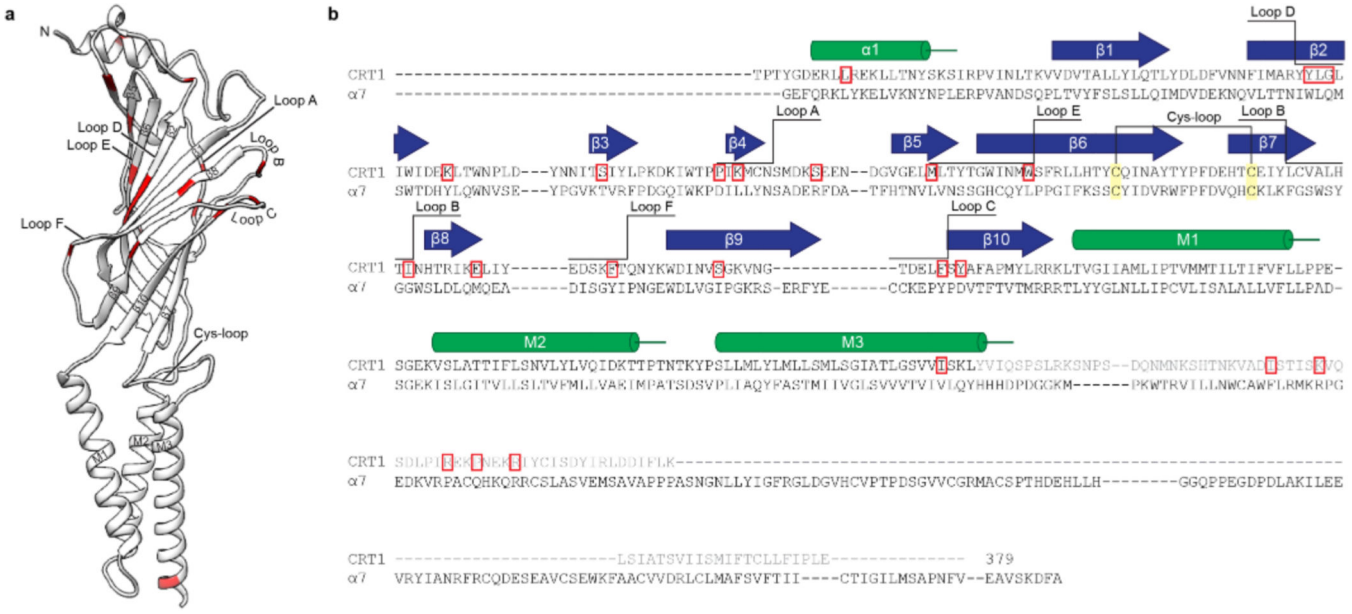
α7 is shown as spheres. **c**, View at 90° rotation about pore axis from **a. d**, Comparison of ECD boxed in **c**; disulfide bonds are shown as spheres. **e**, M2 helices of the CRT1 receptor from opposing subunits (chains A and C) with pore-lining residues shown as sticks. Colored spheres indicate the pore diameter by displaying blue spheres (pore diameter > 5.6 Å), green spheres (2.8 Å < pore diameter < 5.6 Å), and red spheres (pore diameter < 2.8 Å), **f**, M2 helices of the α7 nicotinic receptor in an activated state (PDB :7KOX). **g**, M2 helices of α7 in a desensitized state (PDB:7KOQ). **h**, M2 helices of α7 in a resting state (PDB:7KOO). **i**, Pore diameters of CRT1 and α7 in panels **e-h** as a function of a distance along the pore axis. Structures were aligned using the M2 heiix Leu9' at the midpoint of the pore, which we defined as y = 0.



Extended Data Fig. 7 |. CRT1 pharmacological properties.

a, CRT1 exhibited dose-dependent responses to nootkatone, zearalenone and GDN in patch-clamp experiments. Nootkatone EC50 = 16.1 µM, 95% CI = 13.1 – 18.7 µM, zearalenone EC50 = 4.8µM, 95% CI = 4.3 – 5.4 µM, GDN EC50 = 125.1 NM, 95% CI = 91.2 −184.6 nM. n = 8 cells per 1igand. **b**, Minimal desensitization was measured in response to low concentrations of agonist while higher concentrations produced inhibition with large wash off currents, consistent with moderate pore block. $p < 0{,}0001$ for concentration, two-tailed student's t-test (n = 6 cells), **c**, Heat map of normalized axial nerve and arm

responses to 50 μM of the indicated compound (except GDN at 5 μM to avoid micelle formation). Nootkatone and zearalenone elicited the most robust arm activity among CRT1 agonists or other molecules. Some differences in agonist efficacy in isolated receptors and arms are expected due to solubility issues at concentrations used for arm experiments, particularly with detergent molecules, unknown features of native CR signal transduction, neural integration. And differences in experimental preparations. n = 7–8 arms, **d**, WT CRT1 exhibited dose-dependent activity in response to nootkatone. Y78A mutant channels were insensitive to nootkatone except at higher concentrations which inhibited activity, n = 8 cells per condition, **e**, Current-voltage (I-V) relationships showing ligand-gated activity in WT CRT1 versus constitutive activity in agonist-binding site mutants. Ligands did not increase currents In mutant channels and all channels were sensitive to CR blocker mecamylamine (1mM). Data are represented as mean ±SEM.



Extended Data Fig. 8 |. CRT1 vs. α7 sequence alignment and residues under strongly positive selective pressure.

a, Side view of CRTI cryo-EM single subunit structure with residues under highly positive selection (LRT > 3) colored in red. **b**, Sequence alignment of CRT1 and human α7 nicotinic receptor (UniProt accession number: P36544). Unresolved post M3 region of CRT1 model is colored in gray. Red boxes indicate residues with LRT > 3 highlighted in panel **a**, and yellow boxes highlight the conserved Cys-loop disulfide bond.

Extended Data Table 1|

Likelihood ratio tests between branch models of evolution for acetylcholine receptor-like genes of *Octopus bimaculoides*

Model comparison	Degrees of freedom	Chi-square	P-value
M0 vs. Two-ratio	1	193.24	<0.0001

Model comparison	Degrees of freedom	Chi-square	P-value
Two-ratio vs. Three-ratio	1	0.009	0.75
Free-ratio vs. Two-ratio	130	865.66	<0.0001

M0: Null model, one average dN/dS for the tree. Two-ratio: dN/dS values are different for background (nicotinlc receptors) and foreground branches (CRs). Three-ratio: dN/dS values are distinct for nicotinic acetylcholine receptors and the two major clades of *O. bimaculoides* CRs. Free ratio: different dN/dS values for each branch of the tree.

Extended Data Table 2.

Cryo-EM data collection, processing and model statistics

	CRT1-diosgenin complex (EMD-28163, PDB 8EIS)
Data collection and processing	
EM Facility	PNCC
Magnification	81000
Voltage (kV)	300
Electron exposure (e^{-}/Å^2)	50
Defocus range (μm)	−0.5 to −2.5
Pixel size (Å)	1.056
Symmetry imposed	C5
Initial particle images (no.)	1,649,612
Final particle images (no.)	157,691
Map resolution (Å)	2.62
FSC threshold	0.143
Map sharpening B factor (Å^2)	−80
Refinement	
Initial model used (PDB code)	5KXI
Model Composition	
Non-hydrogen atoms	24825
Protein residues	1460
N-glycan	20
Diosgenin	5
B factors (Å^2)	
Protein	57.97
Ligand	57.25
R.m.s. deviations	
Bond length (Å^2)	0.006
Bond angle (°)	0.676
Validation	
Molprobity score	1.47 (100th %)
Clashscore	4.23 (100th %)
Poor rotamers (%)	0.00
Ramachandran plot	
Favored (%)	96.07
Allowed (%)	3.93

	CRT1-diosgenin complex (EMD-28163, PDB 8EIS)
Outliers (%)	0.00

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank C. Winkler for providing octopuses. P. Kilian and B. Walsh for assistance with behavioural experiments. L. van Giesen and S. Burke for contributions to the early stages of the project, S. Krueger for assistance with histology, J. Teng for preliminary functional tests, R. Cabuco and L. Baxter for baculovirus production, A. Grearson for photos, C. Noviello for screening grids and J. Zhou for critical reading of the manuscript. Single-particle cryo-EM grids were screened at the University of Texas Southwestern Medical Center Cryo-EM Facility, which is supported by the Cancer Prevention and Research Institute of Texas Core Facility Support Award no. RP170644. A portion of this research was supported by NIH grant no. U24GM129547 and performed at the Pacific Northwest Center at Oregon Health and Science University and accessed through the Environmental Molecular Sciences Laboratory (grid.436923.9), a Department of Energy Office of Science User Facility sponsored by the Office of Biological and Environmental Research. This research was further supported by grants to N.W.B. from the New York Stem Cell Foundation, Searle Scholars Program and NIH (nos. R35GM142697 and R01NS129060), NIH grants to R.E.H. (nos. R01NS120496 and R01NS129060), to G.K. (no. F32DA047848) and support from the American Heart Association to J J.K. (no. 20POST35200127), and from the National Science Foundation to C.A.A. (no, NSF-PRFB 2010728).

Data availability

Atomic model coordinates and cryo-EM density maps for the CRT1 structure have been deposited in the Protein Data Bank with accession code 8EIS and in the Electron Microscopy Data Bank with accession code EMD 28163, respectively.

References

1. van Giesen L, Kilian PB, Allard CAH & Bellono NW Molecular basis of chemotactile sensation in octopus. Cell 183, 594–604 (2020). [PubMed: 33125889]
2. Bargmann CI. Comparative chemosensation from receptors to ecology. Nature 444, 295–301 (2006). [PubMed: 17108953]
3. Mollo E. Garson MJ, Polese G, Amodeo P. & Ghiselin MT. Taste and smell in aquatic and terrestrial environments. War. Prod. Rep 34,496–513 (2017).
4. Mollo E. et al. Sensing marine biomolecules: smell, taste, and the evolutionary transition from aquatic to terrestrial life. Front. Chem 2, 92 (2014). [PubMed: 25360437]
5. Hanlon RT & Messenger JB Cephalopod Behaviour 2nd edn (Cambridge Univ. Press, 2018).
6. Wells MJ, Freeman NH & Ashburner M. Some experiments on the chemotactile sense of octopuses. J. Exp. Biol 43, 553–563 (1965). [PubMed: 5863758]
7. Croset V. et al. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. PLoS Genet. 6, e1001064 (2010).
8. Benton R, Vannice KS, Gomez-Diaz C. & Vosshall LB Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. Cell 136,149–162 (2009). [PubMed: 19135896]
9. Albertin CB et al. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524,220–224 (2015). [PubMed: 26268193]
10. Xu P. et al. Functional opsin retrogene in nocturnal moth. Mob. DNA 7,18 (2016). [PubMed: 27777631]

11. Jumper J. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). [PubMed: 34265844]
12. Noviello CM et al. Structure and gating mechanism of the α7 nicotinic acetylcholine receptor. Cell 184,2121–2134 (2021). [PubMed: 33735609]
13. Rahman MM et al. Structural mechanism of muscle nicotinic receptor desensitization and block by curare. Wat. StrucL Mot BioL 29,386–394 (2022).
14. Zarkadas E. et al. Conformational transitions and ligand-binding to a muscle-type nicotinic acetylcholine receptor. Neuron 110,1358–1370 (2022). [PubMed: 35139364]
15. Giordano G. et al. Volatile secondary metabolites as aposematic olfactory signals and defensive weapons in aquatic environments. Proc. Natl Acad. Set. USA 114, 3451–3456 (2017).
16. Long JD & Hay ME Fishes team aversions to a nudibranch's chemical defense. Mar. Ecol. Prog. Ser 307, 199 208 (2006).
17. Karlin A. Emerging structure of the nicotinic acetylcholine receptors. Nat. Rev. Neurosci 3, 102–114 (2002). [PubMed: 11836518]
18. Hochner B. An embodied view of octopus neurobiology. Curr. BioL 22, R887–892 (2012). [PubMed: 23098601]
19. Fouke KE & Rhodes HJ Electrophysiological and motor responses to chemosensory stimuli in isolated cephalopod arms. Biol. Bull, 238.1–11 (2020). [PubMed: 32163724]
20. Morales-Perez CL, Noviello CM & Hibbs RE Manipulation of subunit stoichiometry in heteromeric membrane proteins. Structure 24,797–805 (2016). [PubMed: 27041595]
21. Zivanov J. et al. New tools for automated high-resolution cryo-EM structure determination in REUON-3. eLife 7, e42166 (2018). [PubMed: 30412051]
22. Zheng SQ et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy, Nat. Methods 14, 331–332 (2017).
23. Zhang K. Gctf: real-time CTF determination and correction. J. Struct Biol 193,1–12 (2016). [PubMed: 26592709]
24. Wagner T. et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun. Biol 2, 218 (2019). [PubMed: 31240256]
25. Afonine PV et al. New tools for the analysis and validation of cryo-EM maps and atomic models. Acta Crystallogr. D Struct. Biol 74,814–840(2018). [PubMed: 30198894]
26. Kucukelbir A, Sigworth FJ & Tagare HD. Quantifying the local resolution of cryo-EM density maps. Nat. Methods 11, 63–65 (2014).
27. Morin A. et al. Collaboration gets the most out of software. eLife 2. e01456 (2013). [PubMed: 24040512]
28. Waterhouse A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303 (2018). [PubMed: 29788355]
29. Emsley P, Lohkamp B, Scott WG & Cowtan K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr 66.486–501 (2010). [PubMed: 20383002]
30. Adams PD et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr 66, 213–221 (2010). [PubMed: 20124702]
31. Liebschner D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol 75, 861–877 (2019). [PubMed: 31588918]
32. Williams CJ et al. MolProbity: more and better reference data for improved all atom structure validation. Protein Sei. 27, 293–315 (2018).
33. Pettersen EF et al. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput. Chem, 25,1605–1612(2004). [PubMed: 15264254]
34. Pettersen EF et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sei. 30, 70–82 (2021).
35. Schrödinger L. The PyMOL Molecular Graphics System, version 2.4.1 https://github.com/schrodinger/pymol-open-source (2015).
36. UniProt Consortium UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 47, D506–D515 (2019). [PubMed: 30395287]

37. Pei J, Kim B-H & Grishin NV PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 36, 2295–2300 (2008). [PubMed: 18287115]
38. Smart OS Neduvelil JG, Wang X, Wallace BA & Sansom MS HOLE: a program for the analysis of the pore dimensions of ion channel structural models. J. Mol. Graph 14,354–360(1996). [PubMed: 9195488]
39. Klesse G, Rao S, Sansom MSP & Tucker SJ CHAP: a versatile tool for the structural and functional annotation of ion channel pores. J. Mol. Bioi 431,3353–3365 (2019).
40. Krissinel E. & Henrick K. Inference of macromolecular assemblies from crystalline state. J. MoL BioL 372,774–797 (2007). [PubMed: 17681537]
41. Tian W, Chen C, Lei X, Zhao J. & Liang J. CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res. 46, W363–W367 (2018). [PubMed: 29860391]
42. Eddy SR Profile hidden Markov models. Bioinformatics 14,755–763 (1998). [PubMed: 9918945]
43. Madden T. in The NCBI Handbook 2nd edn Ch. 16 (2003).
44. Li W. & Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22,1658–1659 (2006). [PubMed: 16731699]
45. Katoh K. & Standley. D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol 30,772–780 (2013). [PubMed: 23329690]
46. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A. & Jermiin LS ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14,587–589 (2017). [PubMed: 28481363]
47. Minh BQ et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol Evol 37,1530–1534 (2020). [PubMed: 32011700]
48. Hoang DT, Chernomor O, von Haeseler A, Minh BQ & Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol. BioL Evoi 35,518–522 (2018).
49. Revell LJ phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol 3,217–223 (2012).
50. Suyama M, Torrents D, & Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34, W609–W612 (2006). [PubMed: 16845082]
51. Pond SL, Frost SD & Muse SV. HyPhy: hypothesis testing using phytogenies. Bioinformatics 21, 676–679 (2005). [PubMed: 15509596]
52. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. MoL Biol. EvoL 24, 1586–1591 (2007). [PubMed: 17483113]
53. Murrell B. et al. Delecting individual sites subject to episodic diversifying selection. PioS Genet. 8, e1002764 (2012).
54. Albertin CB et al. Genome and transcriptome mechanisms driving cephatopod evolution. Nat Commun. 13, 2427 (2022). [PubMed: 35508532]
55. Wu TD & Watanabe CK GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21,1859–1875 (2005). [PubMed: 15728110]

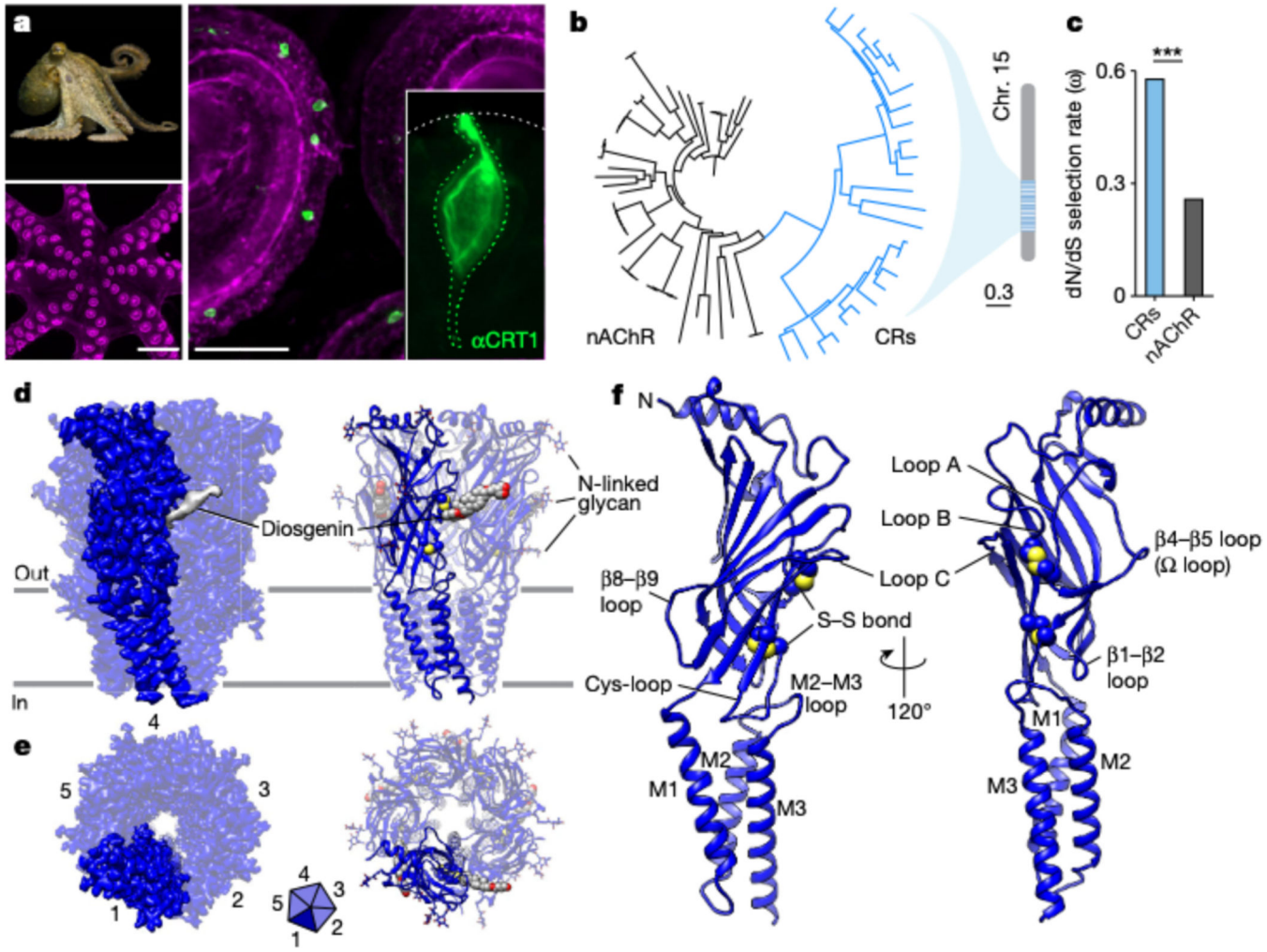


Fig. 1 |. Octopus CRs are divergent and structurally distinct from related neurotransmitter receptors.

a, Anti-CRT1 (green) localized to the extracellular dendritic ending of putative receptor cells in the sucker epithelium of *O. bimaculoides* arms. Anti horseradish peroxidase is shown in purple. Representative of three octopuses. Scale bars, 1 mm (bottom left), 100 μm (middle), **b.** Octopus CRs form a clustered expansion of 26 intronless genes that diverged from octopus nicotinic AChRs. Scale bar, branch length, **c.** The non-synonymousor synonymous substitution rate indicates accelerated evolution of CRs versus nicotinic receptors, ***$P <$ 0.001 for likelihood ratio test (LRT). **d**, Cryo-EM map and atomic model of octopus CRT1 with a single subunit highlighted from the homopentamer and diosgenin indicated in grey. Disulphide bonds in the agonist binding site and Cys-loop are shown as yellow spheres; N-linked glycans are shown as sticks. **e**. Top views per pendicular to the membrane viewed from the extraceilular space of the receptor map and model, **f.** Single subunit structure of octopus CRT1.

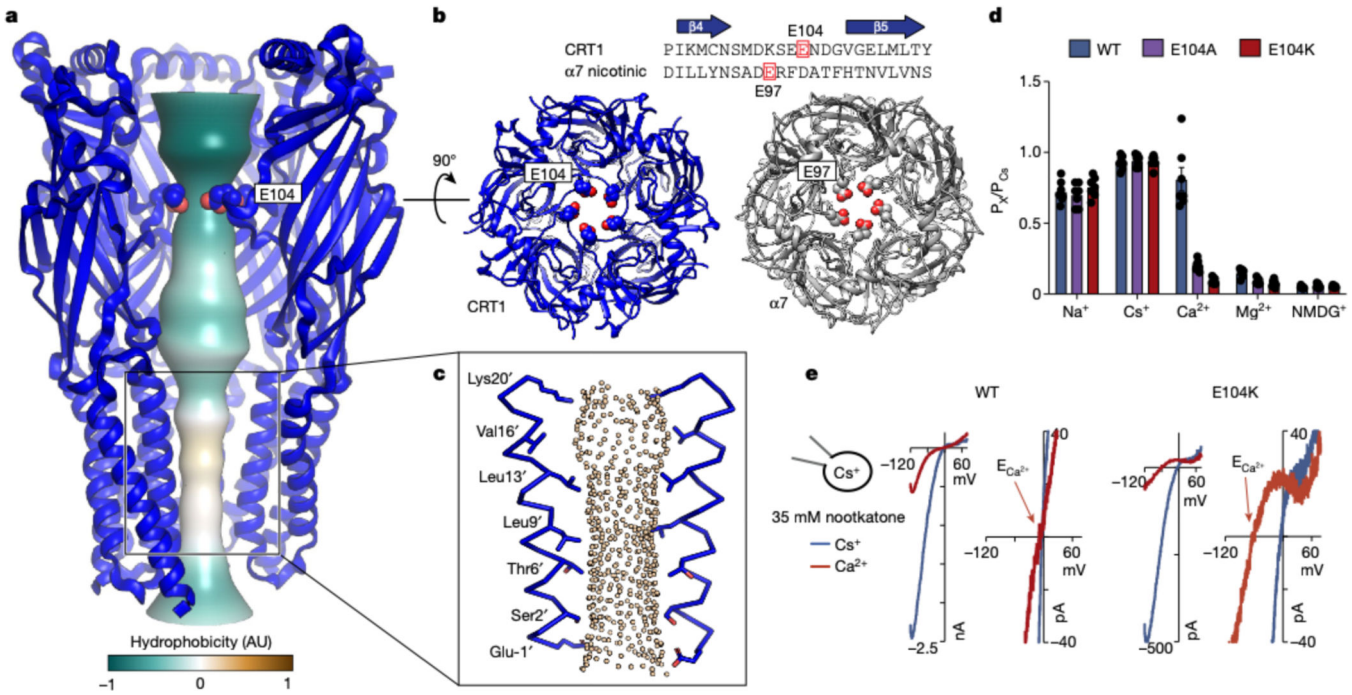


Fig. 2 |. Octopus chemotactile receptor permeation pathway.

a, CRT1 ion permeation pathway coloured by hydrophobicity with El04 indicated as spheres; the front subunit has been removed for clarity. **b**, Top view comparisons of CRT1 and human α7 nicotinic receptor show negatively charged El04 points towards the channel pore. **c**, Residues lining the permeation pathway from two M2 helices; the tan spheres illustrate pore shape. **d, e**, Mutating E104 to alanine or lysine decreased Ca^{2+} selectivity (**d**), leaving El04K with the least Ca^{2+} permeation (**e**). $P < 0.001$ for WT versus E104Aor E104K, $P < 0.01$ for E104A versus E104K. Two way analysis of variance (A NOVA) with post hoc Tukey's test ($n = 7$ cells). Data are represented as the mean± s.e.m.

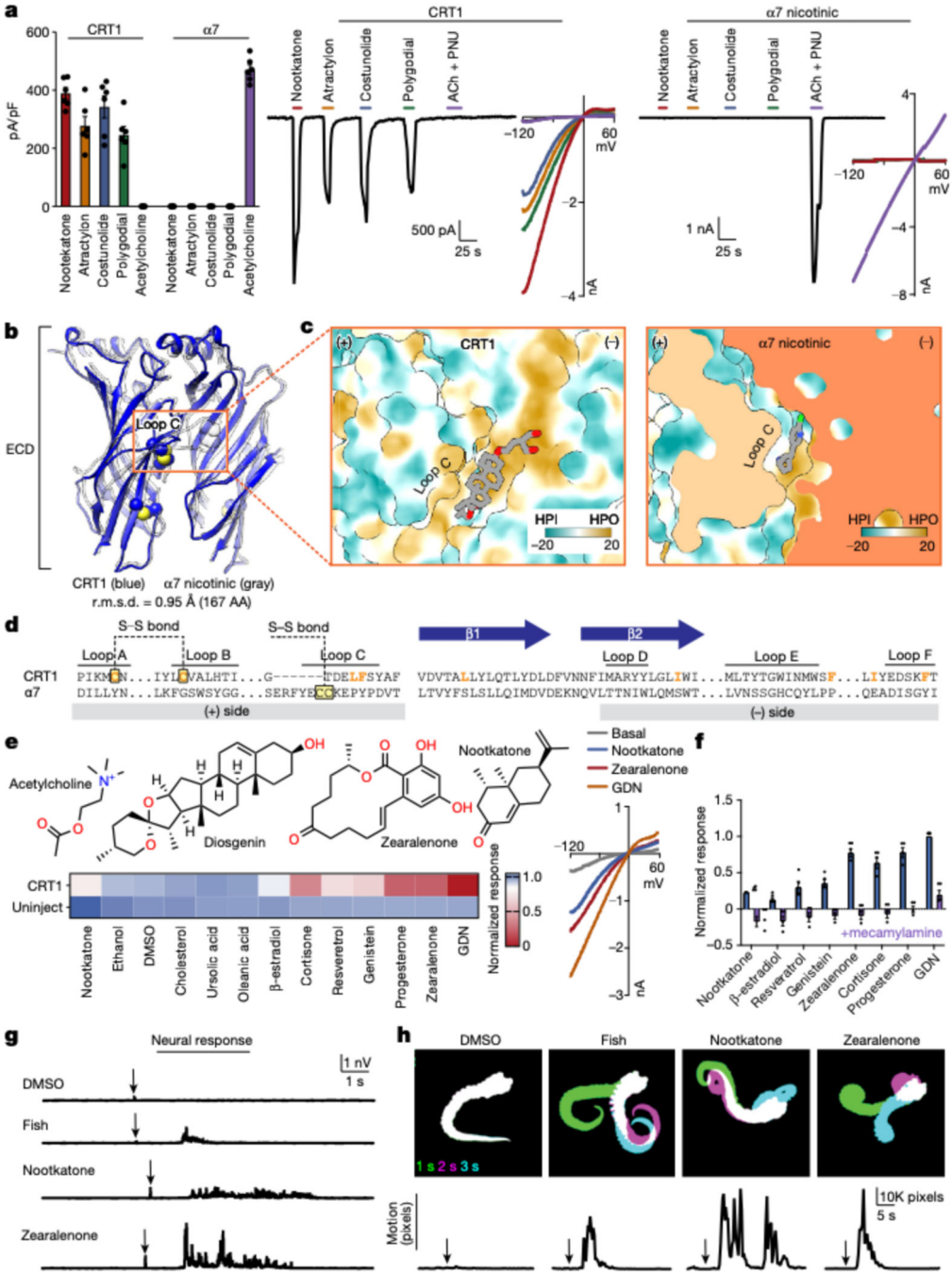


Fig. 3 |. Octopus chemotactile receptor binds poorly soluble molecules.

a, CRT1 was sensitive to terpenes (35 μM nootkatone, 35 μM atrac tylon, 25 μM costunolide, 10 μM polygodial) but not ACh (100 μM ACh + modulator 10 μM PNU) in patch clamp recordings. The α7 nicotinic receptor was activated by ACh but not terpenes. $P < 0.001$, two-way ANOVA with post-hoc Bonferroni test ($n = 6$ cells), **b**, Extracellular domains of octopus CRT1 (blue) and human α7 nicotinic receptor (grey). The orange box indicates partially overlapping agonist binding sites. Disulphide bonds are shown as blue and yellow spheres for CRT1. **c**, The surface of the agonist site is coloured

by molecular lipophilicity potential from hydrophilic (KPI, green) to hydrophobic (HPO, brown). Agonists are shown as sticks. Part of Loop C from α7 was removed to visualize bound epibatidine. **d**, Structure-based sequence alignment of agonist binding site elements. Hydrophobic residues in 5 Å of bound diosgenin in theCRT1model are coloured orange. The yellow box Indicates a disulphide bond present only in CRT1. **e**, Screening of *Xenopus* oocytes expressing CRT1 identified active ligands with structural simiiarities to GDN and terpenes ($n = 4$).Current-voltage relationship ofCRT1 activity in response to 15 μM nootkatone, 5 μM zearalenone or 125 nMGDN. **f**, The CR blocker mecamylamine(1mM) abolished ligand-evoked activity. $P < 0.001$, two-way ANOVA with post-hoc Bonferroni test ($n = 4$ cells), **g**, Arm axial nerve innervating suckers responded to 3 kDa filtered fish extract, 50 μM nootkatone or zearalenone. $P < 0.001$, one-way ANOVA with post-hoc Tukeys test ($n = 7 – 8$). **h**, Amputated octopus arms exhibit autonomous responses to fish extract or 50 μM ligands but not control. $P < 0.001$, one-way ANOVA with post-hoc Tukey s test ($n = 8$). Data represent the mean ± s.e.m.

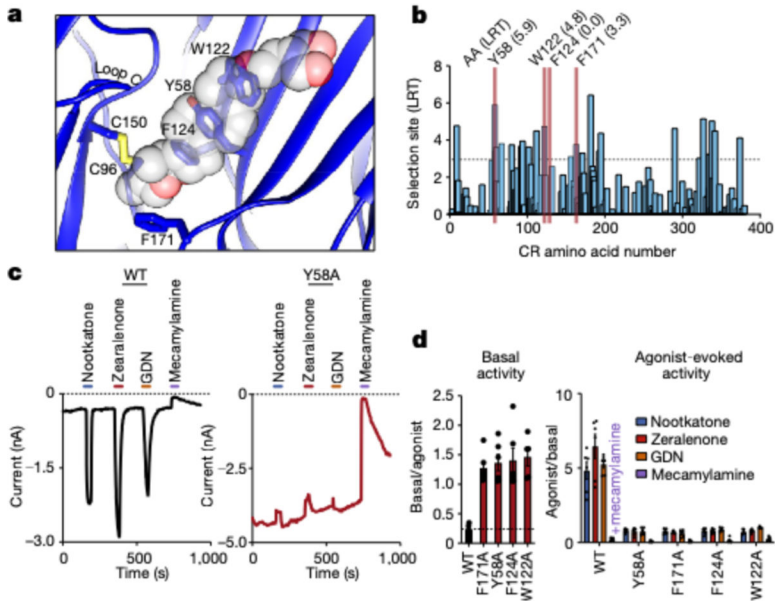


Fig. 4 |. Evolution of the orthosteric binding site facilitates sensory adaptation.

a, Side view of the CRT1 orthosteric binding site showing residues involved in hydrophobic interaction with the diosgenin moiety of GDN. **b**, LRT to diversify selection indicated that ligand interaction sites (red) are under positive selective pressure. Threshold for significance, LRT = 3 ($P < 0.05$). **c**, Mutating hydrophobic residues coordinating GDN and under selective pressure resulted in constitutively active CRs with reduced sensitivity to 15 μM nootkatone, 5 μM zearalenone 125 nMGDN. Wild-type (WT) and mutant activities were blocked by 1 mM mecamylamine. Compared with WT, mutations increased average basal activity; $P < 0.001$, one way ANOVA with post-hoc Tukey'stest ($n = 5 – 6$ cells), **d**, Mutations reduced ligand-dependent currents. $P < 0.001$, two-way ANOVA with post-hoc Tukey's test ($n = 5–6$ cells). Data represent the mean ± s.e.m.