Supplemental information

Structural basis of ligand specificity

and channel activation

in an insect gustatory receptor

Heather M. Frank, Sanket Walujkar, Richard M. Walsh Jr., Willem J. Laursen, Douglas L. Theobald, Paul A. Garrity, and Rachelle Gaudet
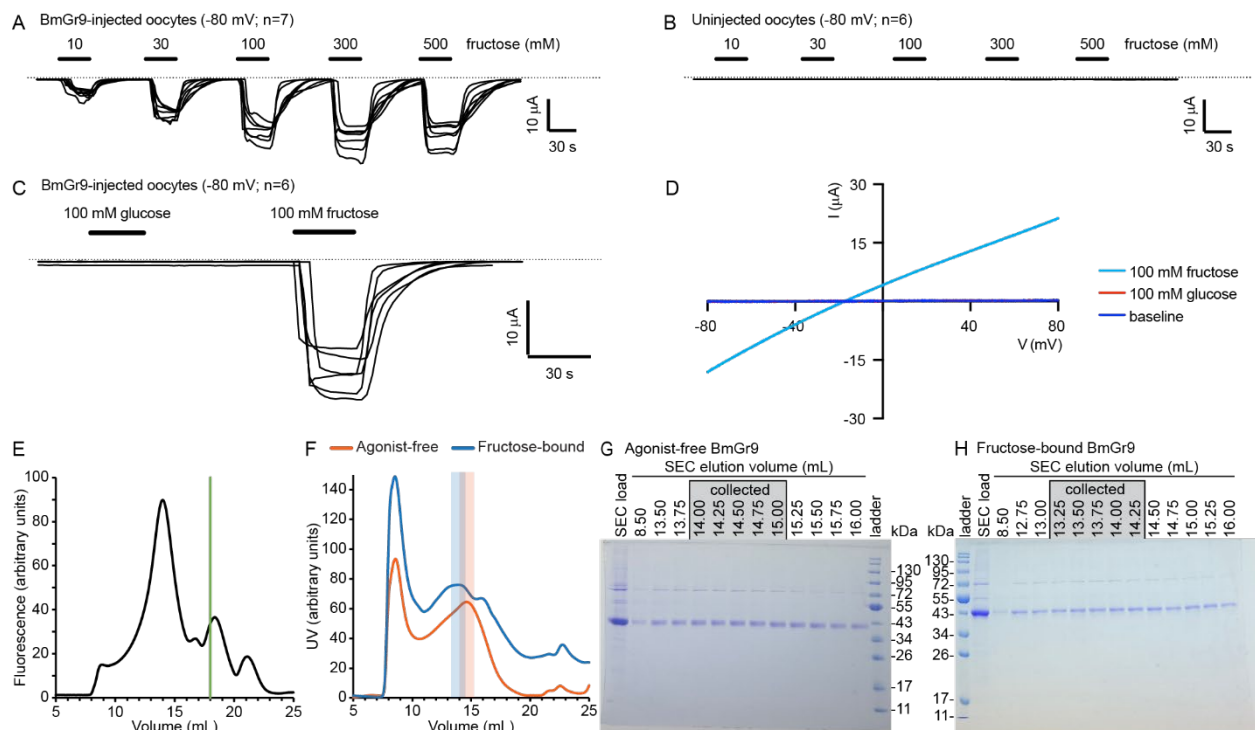
Figure S1. Functional analysis and purification of Twin-Strep-tagged BmGr9, Related to Figure 1

(A-C) BmGr9-expressing oocytes respond specifically to fructose. Individual traces show current at -80 mV in different oocytes. (A) BmGr9-expressing oocytes (n=7 oocytes). (B) Uninjected control oocytes (n=6). (C) BmGr9-expressing oocytes (n=6).

(D) Current-voltage relationships of BmGr9-expressing oocytes.

(E) FSEC trace of a clarified whole-cell lysate of HEK293 cells expressing GFP-tagged BmGr9 run on a Superose 6 10/30. The vertical green line shows where free GFP elutes.

(F) Size exclusion chromatography (SEC; Superose 6 10/30) of ligand-free (orange) and fructose-bound (blue) BmGr9 purified in GDN. The fractions highlighted by orange (ligand-free) and blue (fructose-bound) rectangles were collected and used for cryo-EM.

(G-H) Coomassie-stained SDS-PAGE gels of the SEC (Superose 6 10/30) load and elution fractions of agonist-free BmGr9 (E) and fructose-bound BmGr9 (F). The corresponding elution volumes are above each lane, and the fractions pooled and concentrated for cryo-EM are indicated. The SEC load fractions were diluted 1:10.
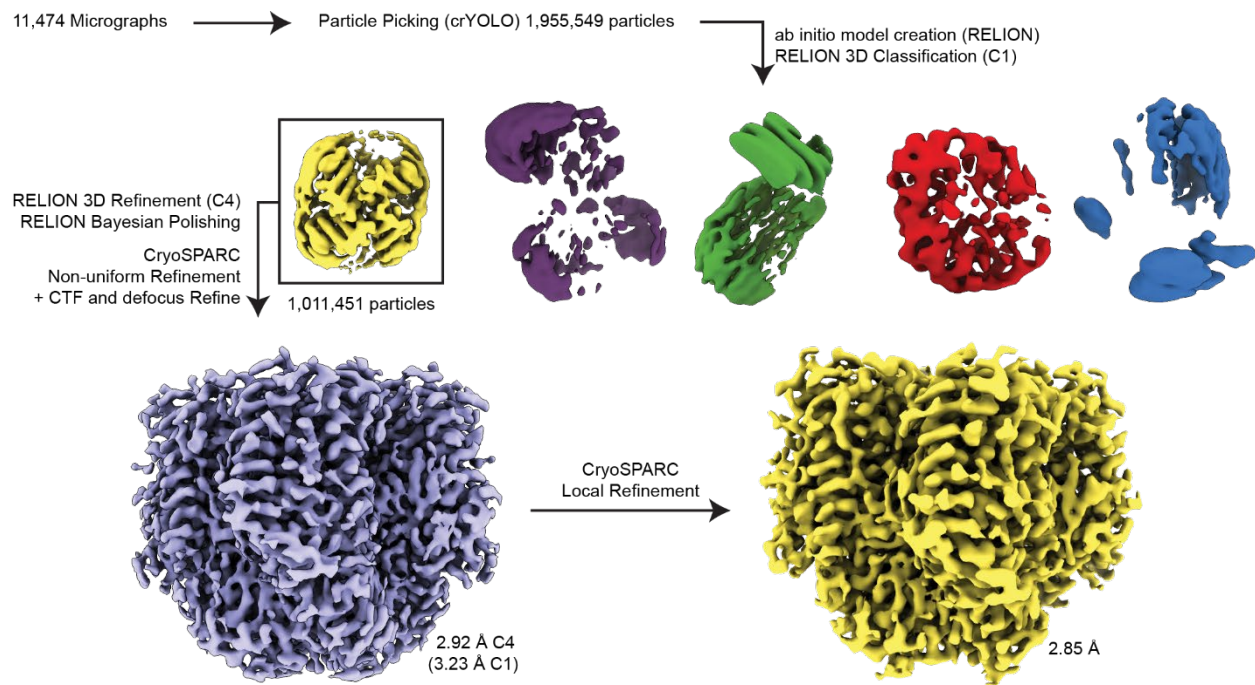
11,474 Micrographs → Particle Picking (crYOLO) 1,955,549 particles

ab initio model creation (RELION)
RELION 3D Classification (C1)

RELION 3D Refinement (C4)
RELION Bayesian Polishing

CryoSPARC
Non-uniform Refinement
+ CTF and defocus Refine

1,011,451 particles

CryoSPARC
Local Refinement

2.92 Å C4
(3.23 Å C1)

2.85 Å

Figure S2. Cryo-EM processing procedure for the agonist-free BmGr9 structure, Related to Figure 1
Processing scheme for classification and refinement of the agonist-free BmGr9 map. The locally filtered map used for the final reconstruction is represented with dust hidden for clarity.
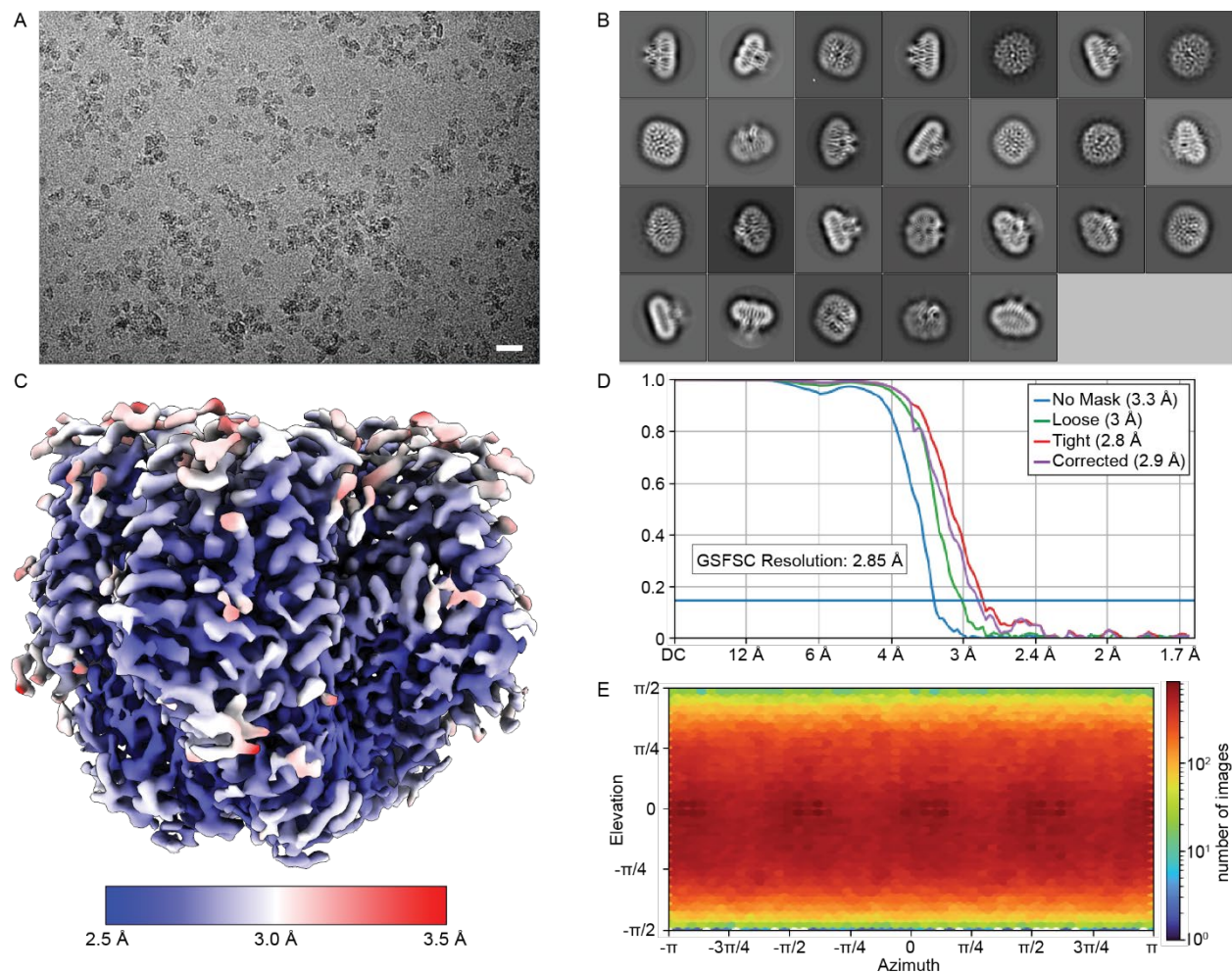
Figure S3. Cryo-EM data analysis for the agonist-free BmGr9 structure, Related to Figure 1

(A) Representative micrograph of BmGr9 embedded in vitreous ice (scale bar = 250 Å), low pass filtered for clarity.

(B) Selected 2D class averages of BmGr9.

(C) Reconstruction of BmGr9 filtered and colored by local resolution.

(D) Gold-standard Fourier shell correlation (FSC) curves from cryoSPARC.

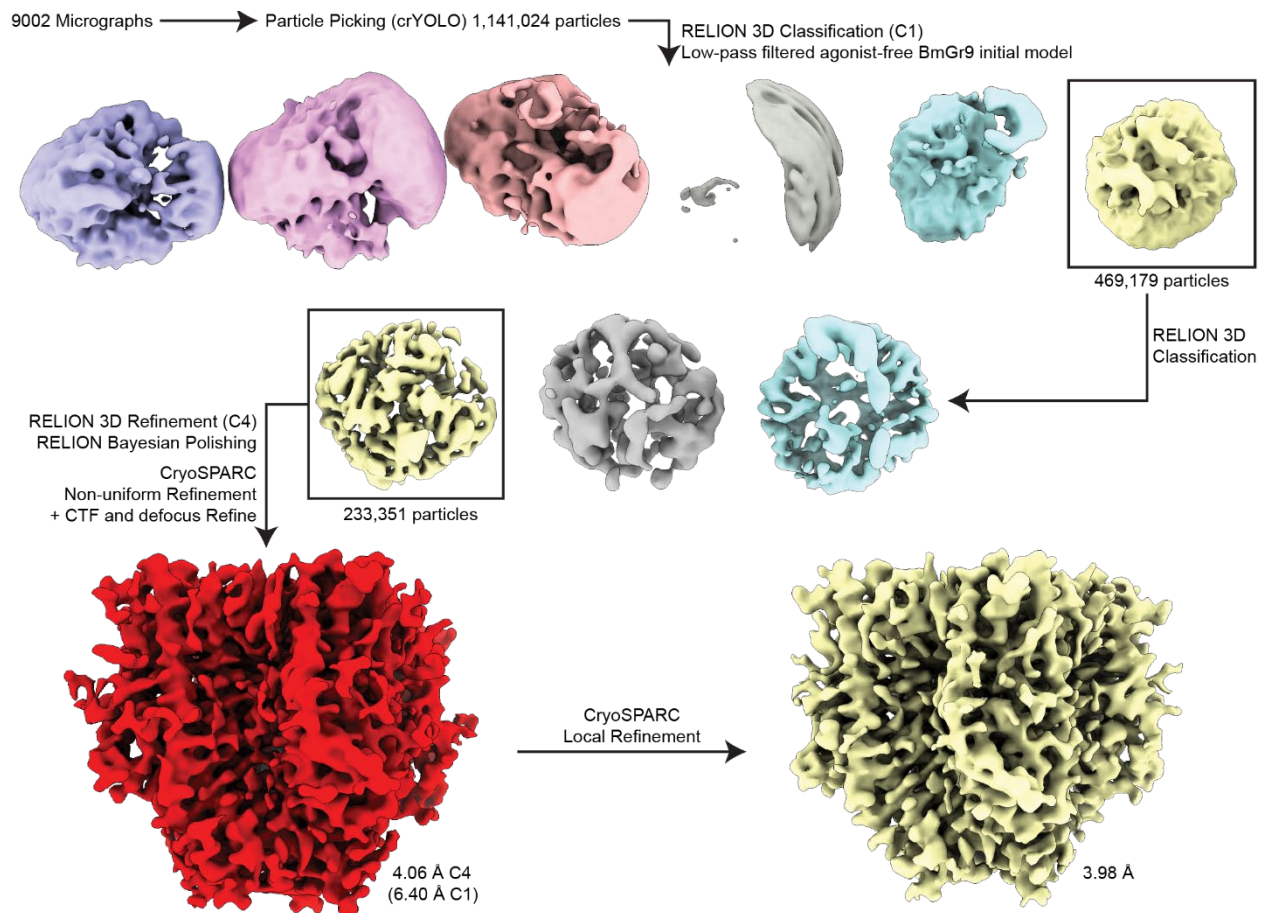(E) Viewing direction distribution plot.

Figure S4. Cryo-EM processing procedure for the fructose-bound BmGr9 structure, Related to Figure 1
Processing scheme for classification and refinement of the fructose-bound BmGr9 map. The locally filtered map used for the final reconstruction is represented with dust hidden for clarity.
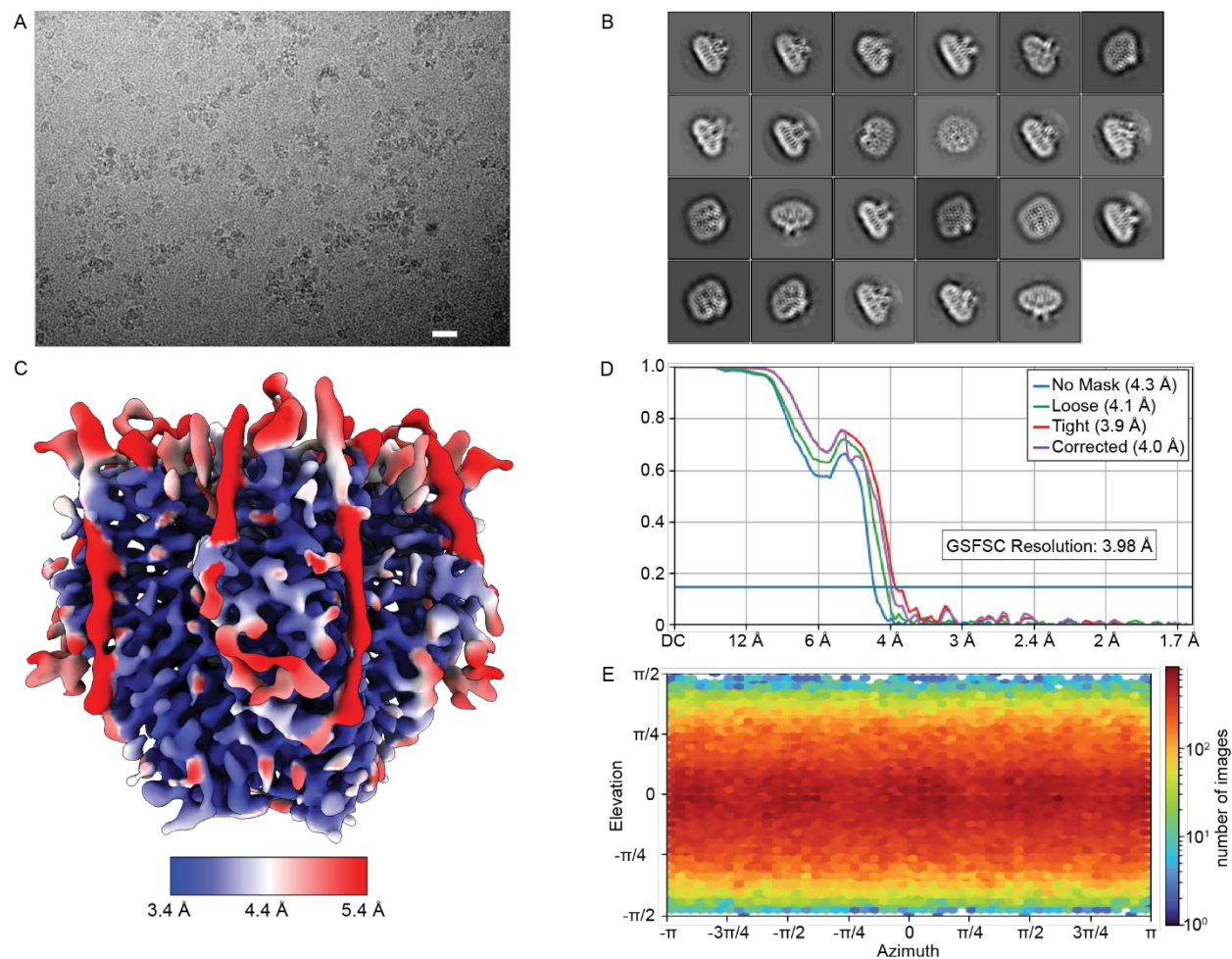
Figure S5. Cryo-EM data analysis for the fructose-bound BmGr9 structure, Related to Figure 1

(A) Representative micrograph of BmGr9 in the presence of fructose embedded in vitreous ice (scale bar = 250 Å), low pass filtered for clarity.

(B) Selected 2D class averages of fructose-bound BmGr9.

(C) Reconstruction of fructose-bound BmGr9 filtered and colored by local resolution.

(D) Gold-standard Fourier shell correlation (FSC) curves from cryoSPARC.

(E) Viewing direction distribution plot.

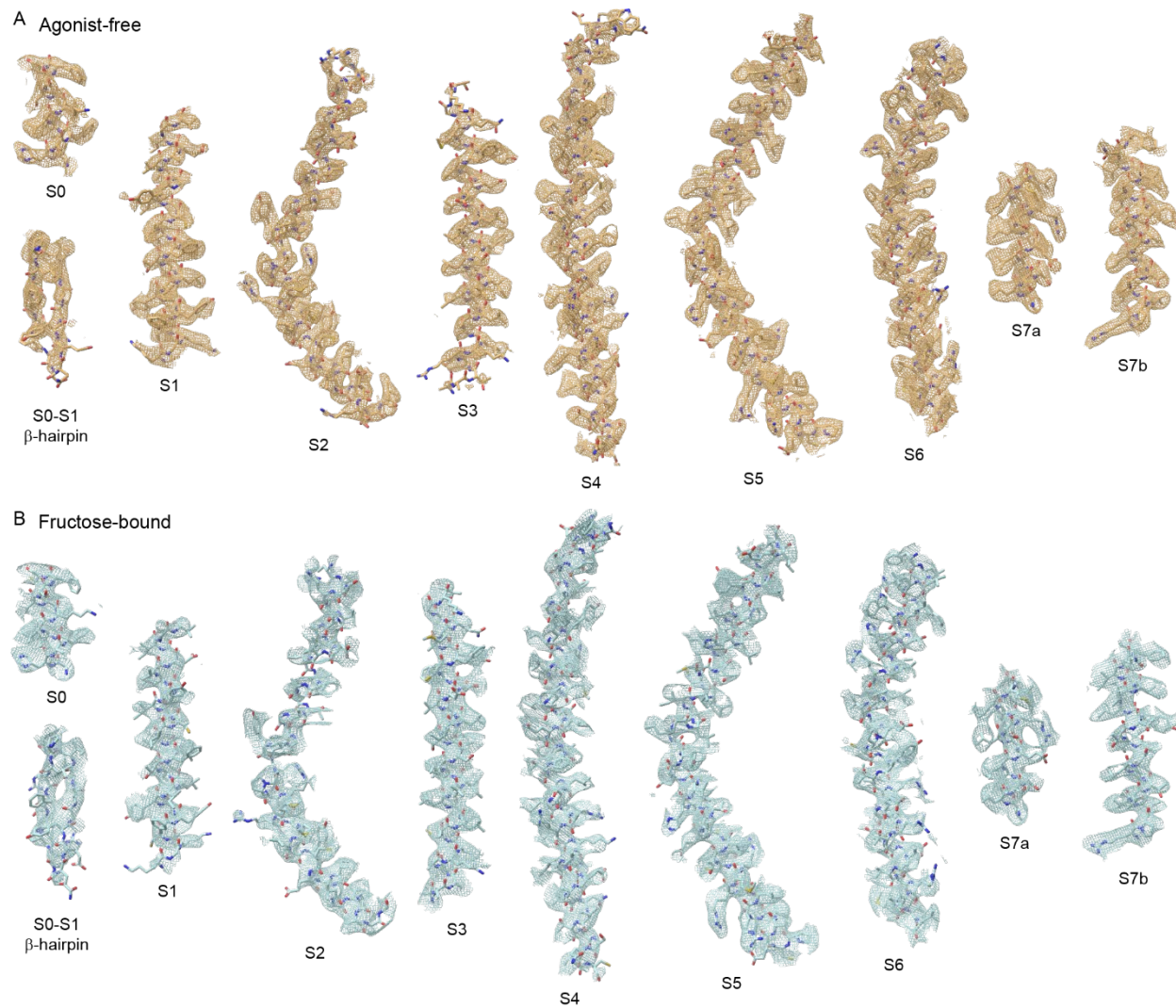**Figure S6. Structural data supporting secondary structure and topology analyses, Related to Figure 2**
Cryo-EM densities of individual helices for the agonist-free (A) and fructose-bound (B) BmGr9 structures. Protein is shown in stick representation with density contoured to 3.5σ and 4.5σ for the agonist-free and fructose-bound maps, respectively. The helices (and S0-S1 β-hairpin) are labeled below each helix.
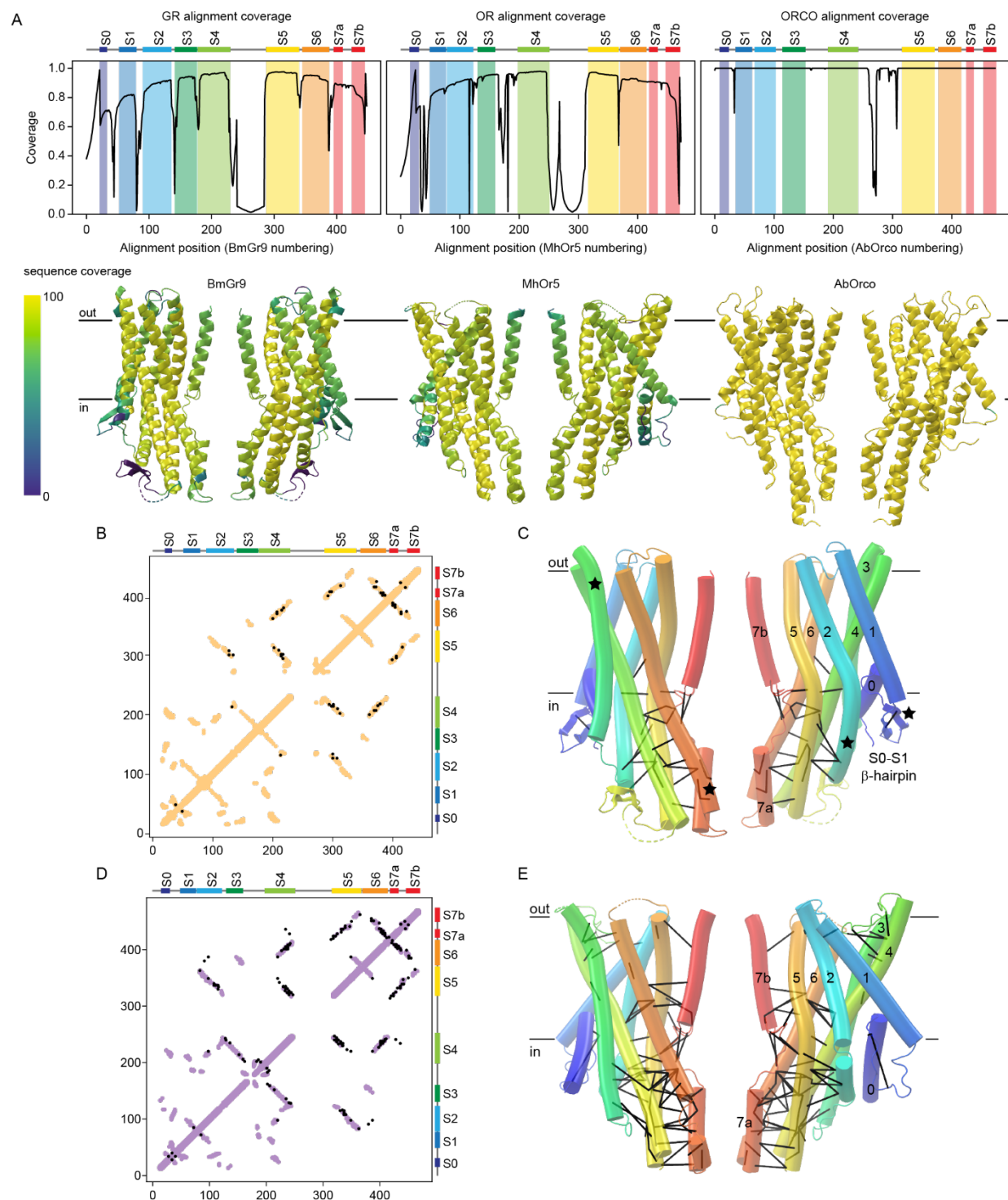
Figure S7. Sequence data supporting secondary structure and topology analyses, Related to Figure 2

(A) Sequence coverage for the three sequence alignments mapped onto the sequence (top) and structure (bottom) of the representative member, as follows: structure-driven alignment of 1895 GR sequences (includes 41 OR sequences) mapped to BmGr9; structure-driven alignment of 3885 OR sequences mapped to MhOr5; and previously published alignment of 176 ORCO sequences[25] mapped to AbOrco.

(B) Sequence covariation analysis using the structure-driven alignment of 1854 GR sequences mapped to BmGr9. Intrasubunit structural contacts in BmGr9 are in orange (8 Å Cα-to-Cα distance cutoff). The 28 evolutionarily coupled residue pairs above the 90% confidence threshold are indicated by black dots.

(C) The 28 evolutionarily coupled residues pairs mapped onto the BmGr9 structure as black Cα-to-Cα bonds. Two opposing subunits are shown in transparent cartoon representation colored in blue-to-red rainbow from N to C terminus. Nearly all top coupled pairs are in the intracellular anchor domain. Noteworthy structural features highlighted in Figure 2A are marked with black stars.

(D) Sequence covariation analysis using the structure-driven alignment of 3885 OR sequences mapped to MhOr5. Intrasubunit structural contacts MhOr5 are in purple (8 Å Cα-to-Cα distance cutoff). The 82 evolutionarily coupled residue pairs above the 90% confidence threshold are indicated by black dots.

(E) The 82 evolutionarily coupled residues pairs mapped onto the BmGr9 structure as black Cα-to-Cα bonds shown in a representation similar to the representation in C.
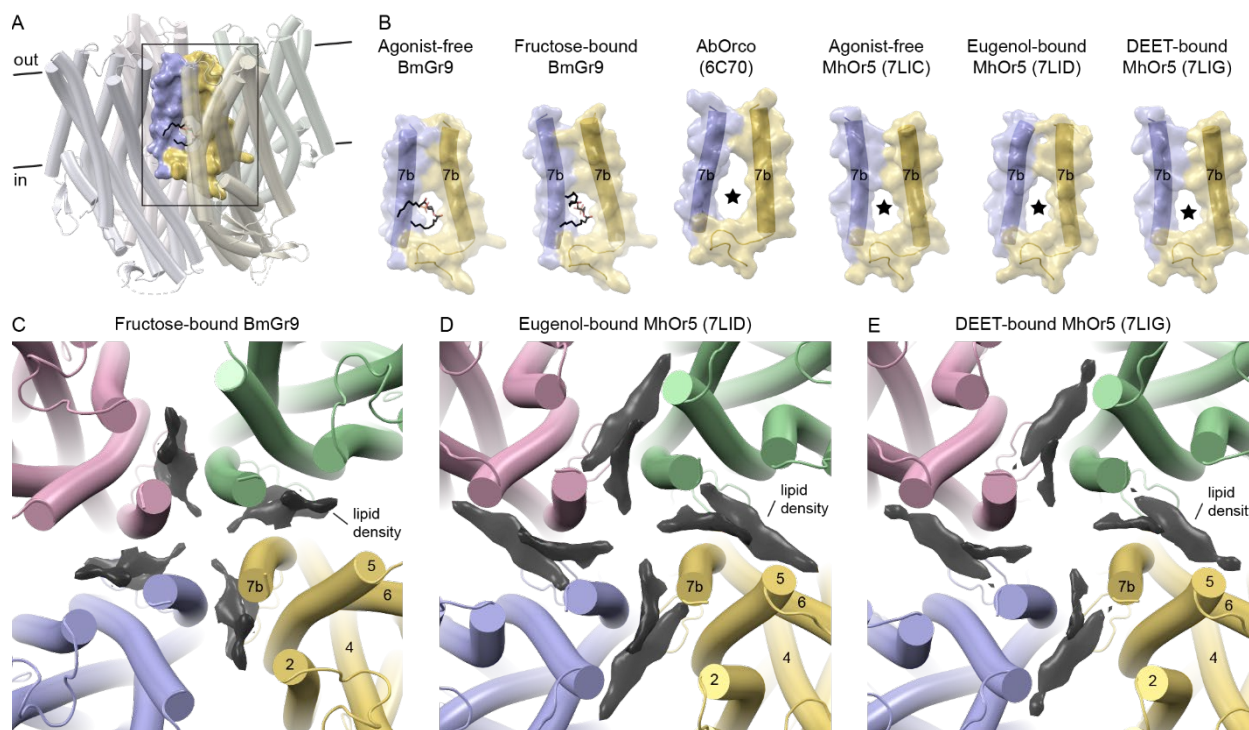
Figure S8. Density for a lipid penetrating the ion pore is observed in other GR and OR structures, Related to Figure 3

(A) Structure of agonist-free BmGr9 illustrating the position of the transmembrane fenestration between the blue and yellow subunits with the bound pore-penetrating lipid (black sticks). The boxed region is shown in B for this structure and other available structures of GRs and ORs.

(B) All GR and OR structures have fenestrations between pore helices in the membrane plane. For each available GR or OR structure, the transparent surface and cartoon representation of helix S7b from two adjacent subunits and the loop preceding helix S7b for the yellow subunit are illustrated. For the two BmGr9 structures, the modeled pore-penetrating lipid is shown in black sticks; for the other structures, the fenestration is marked by a black star.

(C-E) Transmembrane cross-sections of BrGr9 and MhOr5 structures viewed from the extracellular side of the membrane. Surfaces shown in black are densities in the respective cryo-EM map corresponding to potential pore-penetrating lipids. The fructose-bound BmGr9 structure and map contoured at 3.5σ are shown in (C); the eugenol-bound MhOr5 structure (PDB ID: 7LID) and map (EMDB ID: 23374) contoured at 7.5σ in (D); and the DEET-bound MhOr5 (PDB ID: 7LIG) and map (EMDB ID: 23375) contoured at 7.5σ in (E). Visible helices of the yellow subunit are labeled.
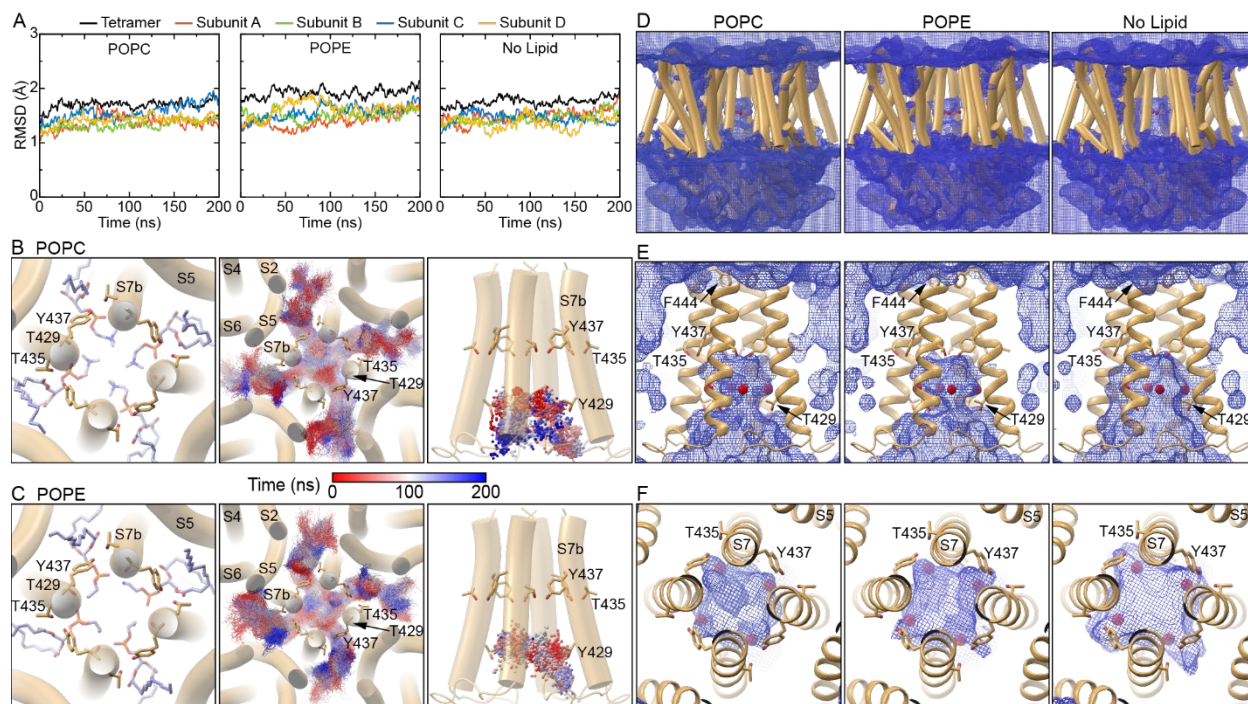
Figure S9. Molecular dynamics simulations of agonist-free BmGr9 in presence and absence of pore-penetrating lipids, Related to Figure 3

(A) Plots showing 1-ns running averages of protein backbone RMSD during 200-ns molecular dynamics simulations of agonist-free BmGr9 in presence of pore-penetrating POPC lipids (left panel), POPE lipids (middle panel), and in the absence of pore-penetrating lipid (right panel). Black lines indicate the RMSD of the entire tetramer while colored lines indicate the RMSD of individual subunits.

(B-C) Behavior of the POPC (B) and POPE (C) pore-penetrating lipids during molecular simulations. The left-hand panels show an extracellular view of the central pore for the simulation systems after minimization. The pore-penetrating lipids and lipid-interacting protein residues T429, T435, and Y437 shown as sticks. The middle panels show snapshots along the 200-ns long simulations of the pore-penetrating lipids shown as lines, colored to indicate the simulation time from 0 ns (red) to 200 ns (blue). The right-hand panels show a view of the S7b-lined central pore from the membrane plane with snapshots along the 200-ns simulations of the lipid phosphorus atoms marked as dots with the color scheme similar to the middle panel.

(D-F) Views of the three simulations systems, with the averaged density of water molecules from the 200-ns long simulations shown as blue mesh. From left to right are the systems with pore-penetrating POPC lipids, pore-penetrating lipids POPE lipids, and no pore-penetrating lipids. The positions of each phosphorus atom of the pore-penetrating lipids model in the agonist-free BmGr9 structure shown as red spheres for reference.

(D) Views from the membrane plane of the whole BmGr9 tetramer in each system. Both the central pore vestibule open to the cytosol and the ligand-binding pockets open to the extracellular solvent are well hydrated.

(E) Close-ups showing the average water density in and around the pore-forming S7b helices. Notably, the outer pore region, with its two hydrophobic gates closed, is not hydrated.

(F) Close-up views of the central pore from the extracellular side. The right-hand panel shows increased water density in the intersubunit fenestrations in the absence of pore-penetrating lipids.
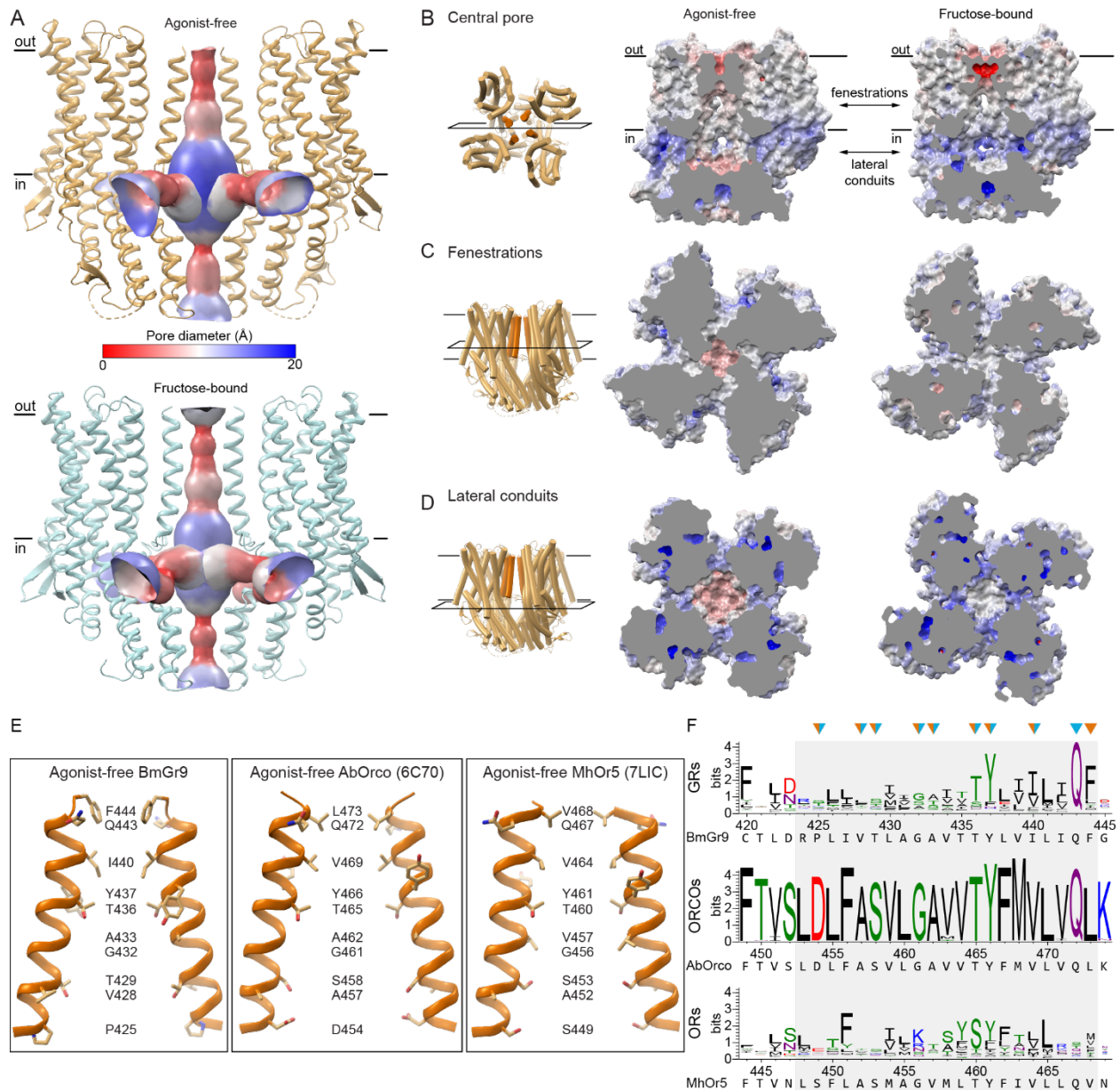
Figure S10. Comparisons of pore between BmGr9 and AbOrco and MhOr5, Related to Figure 4

(A) The quadrivial pore of BmGr9 in the agonist-free (top) and fructose-bound structures illustrated as a surface colored according to its diameter. The cartoon representation of two opposing protein subunits is included, and black lines make the membrane boundaries.

(B-D) Electrostatics surface representation of BmGr9 in the agonist-free (left) and fructose-bound (right) structures, in three cross-sections as indicated by the cartoon representations and planes on the left: (B) Vertical cross-sections through the ion pore. Black lines mark the membrane boundaries, and arrows point to the hydrophobic lateral fenestrations filled with the pore-penetrating lipids and the lateral conduits of the ion pore. (C) Horizontal cross-sections through the hydrophobic fenestrations. (D) Horizontal cross-sections through the lateral conduits. The electrostatics potentials were calculated using APBS in PyMOL and colored as a range from -20 kcal/mol (red) to 20 kcal/mol (blue).

(E) Comparison of the central ion pore of the agonist-free structures of BmGr9 (left), AbOrco (middle; PDB ID: 6C70), and MhOr5 (right; PDB ID: 7LIC). Helix S7b from two opposing protein subunits are shown in cartoon representation, and the sidechains of residues that line the pore walls are shown in sticks and labeled.

(F) Sequence logos of the helix S7b positions (grey box) of alignments of 1854 insect GR sequences, 176 ORCO sequences, and 3885 insect OR sequences, with the reference sequences of BmGr9, AbOrco, and MhOr5,

respectively, indicated below the logo. Residues lining the central pore are marked with arrowheads colored orange (agonist-free structure) or blue (fructose-bound structure) or both colors for sidechains that are pore-lining in both structures.
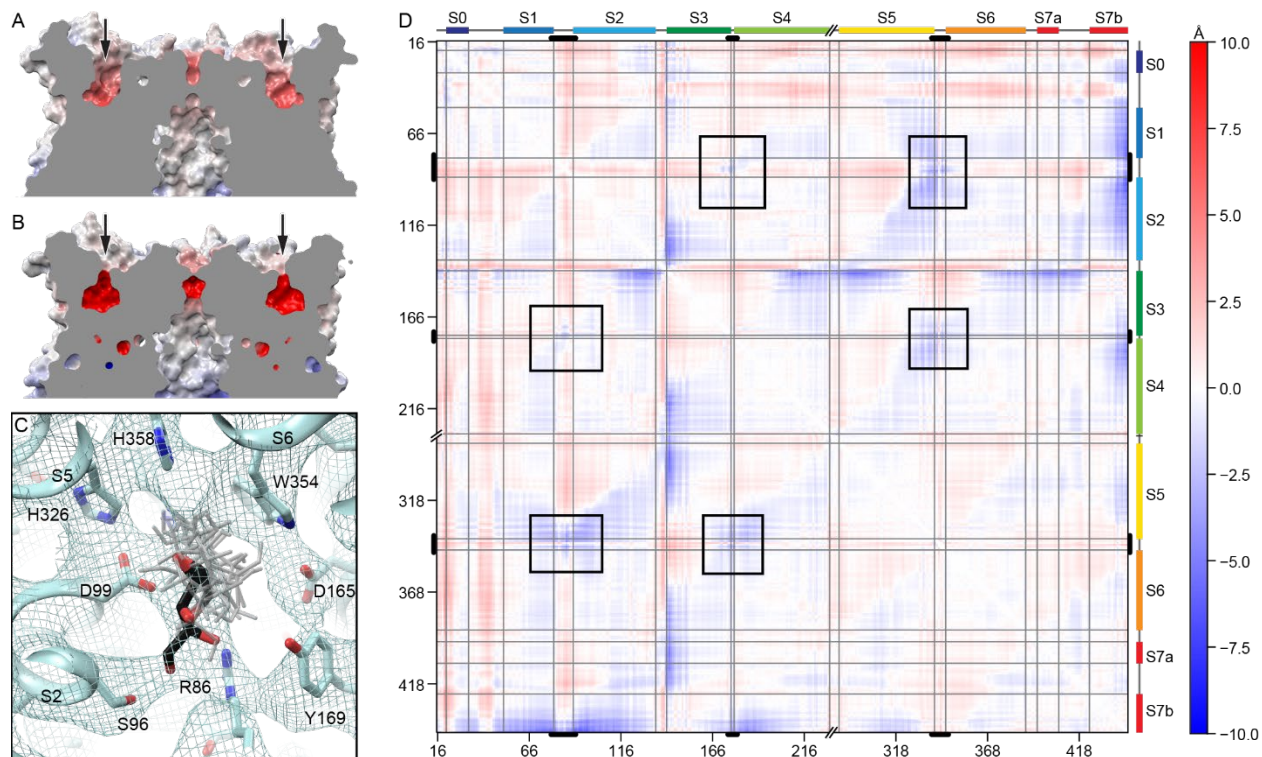
Figure S11. Conformational changes in and fructose docking into the BmGr9 ligand-binding pocket, Related to Figure 5

(A-B) Electrostatics surface representation of agonist-free (A) and fructose-bound (B) BmGr9, sliced through the two ligand-binding pockets of opposing subunits in the tetramer. The two pockets are indicated with black arrows. The electrostatics potentials were calculated using APBS in PyMOL and colored as a range from -20 kcal/mol (red) to 20 kcal/mol (blue).

(C) Cryo-EM density for the fructose-bound map contoured at 4.5σ around the fructose-binding site, with nearby sidechains shown as sticks. The five lowest-energy poses of docked β-D-fructopyranose are illustrated in grey sticks, and the β-D-fructopyranose pose after real-space refinement of the lowest-energy pose against the cryo-EM density is shown in black sticks.

(D) Distance difference matrix of Cα-to-Cα distances for agonist-free versus fructose-bound BmGr9. The breaks in the axes mark missing residues 230-283. Thick black lines on the four axes mark the extracellular loops, and grey lines mark the helix boundaries. Black boxes mark the relative distances between the three pocket-forming regions of BmGr9, highlighting that the extracellular regions generally move closer together upon fructose binding.
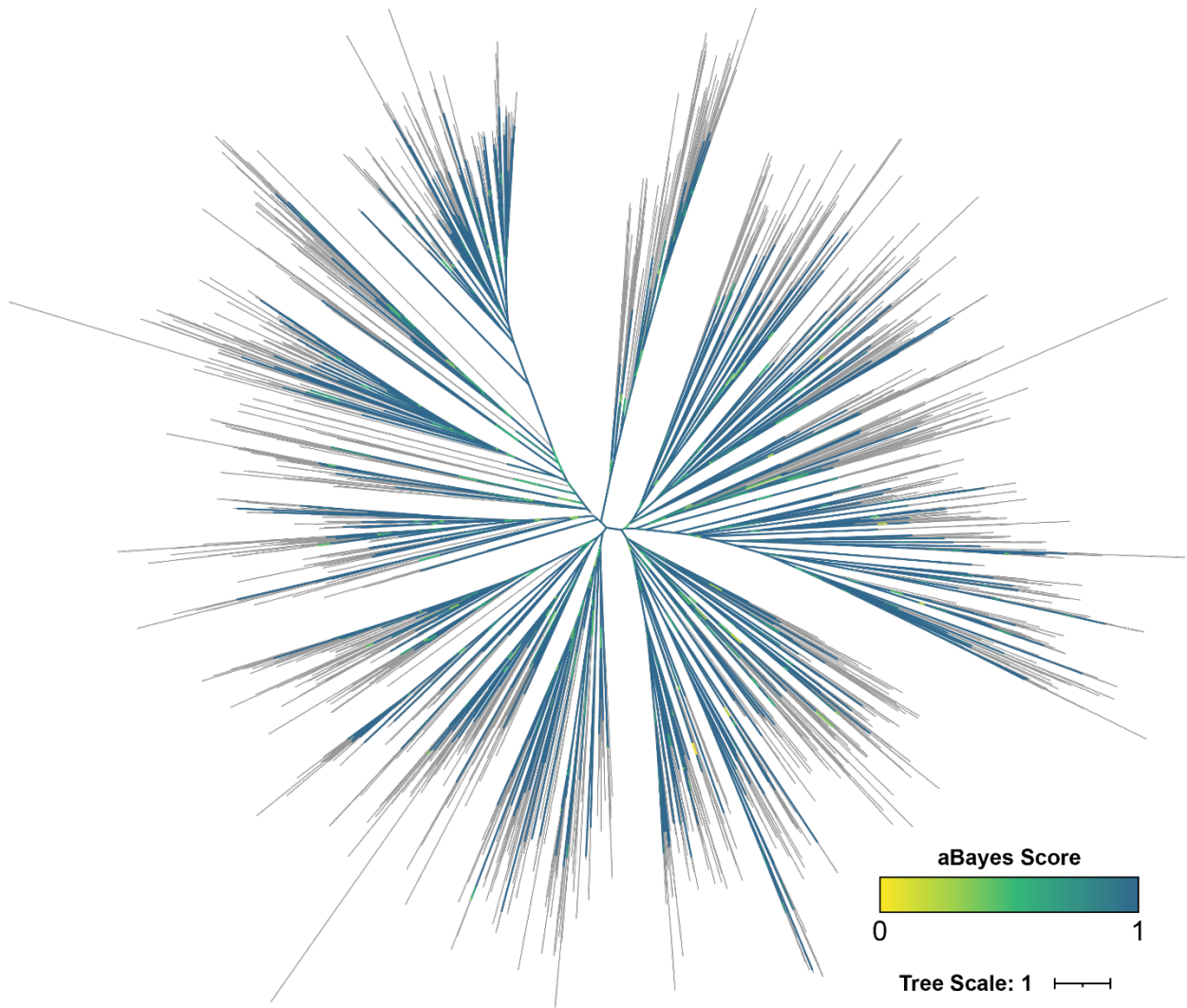
Figure S12. Posterior probability branch supports for the maximum likelihood phylogenetic tree of insect GR sequences, Related to Figure 6

A representation of the phylogenetic tree in Figure 6A showing aBayes support values for each branch. Terminal branches are shown in grey.
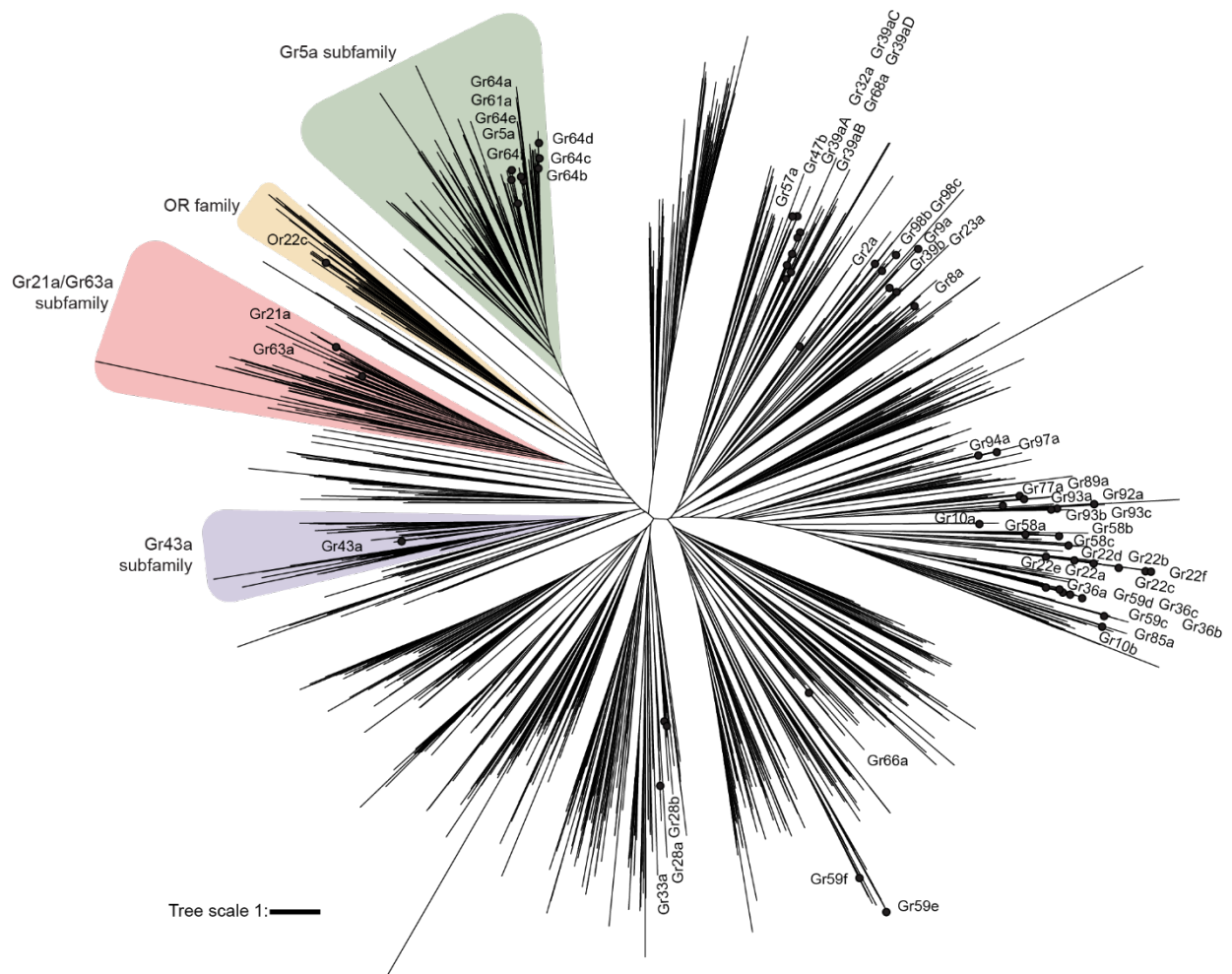
Figure S13. Annotated phylogenetic tree of insect GR sequences, Related to Figure 6

Maximum likelihood phylogenetic tree of 1895 aligned insect GR sequences. Black dots indicate positions of *D. melanogaster* GRs on the tree. The alignment included 41 OR sequences, marking the branch point of the OR family within the more ancestral GR sequences (orange). The three subfamilies analyzed in this work are highlighted: Gr5a subfamily in green, Gr63a subfamily in red, and Gr43a subfamily in purple.
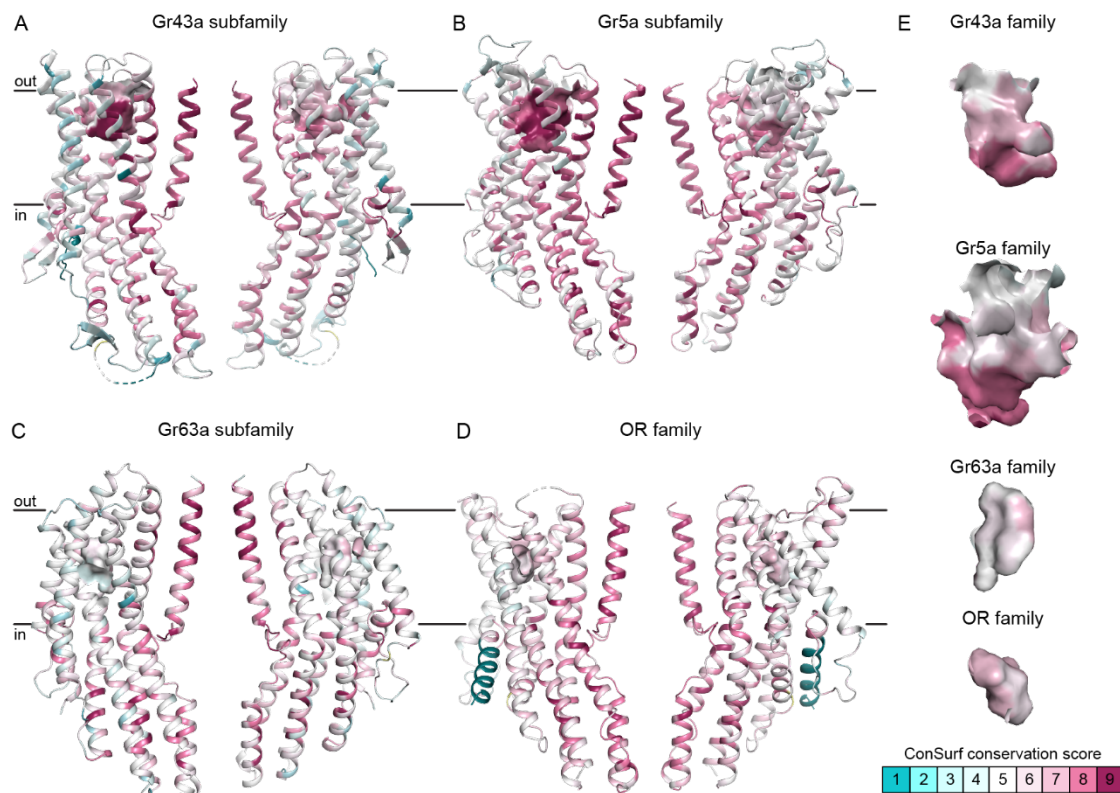
Figure S14. Side views of predicted ligand-binding pockets in representative GR subfamily members, Related to Figure 6

(A-E) Cartoon representation of two opposing subunits from representative subfamily members viewed from the membrane plane, with surface representation of the atoms forming the predicted ligand-binding pockets to illustrate the relative location and size of the binding pockets. The cartoons and surfaces are colored by ConSurf conservation score based on the sequence alignment of the corresponding GR subfamily. As in Figure 6, the following subunits and corresponding protein (sub)families are illustrated: (A) Gr43a subfamily (74 sequences) on the agonist-free BmGr9 structure; (B) Gr5a subfamily (251 sequences) on the AlphaFold2 model of Gr5a (UniProt ID: Q9W497); (C) Gr63a subfamily (107 sequences) on the AlphaFold2 model of Gr63a (UniProt ID: Q9VZL7); and (D) the OR family (3885 sequences) on the agonist-free MhOr5 structure (PDB ID: 7LIC). For the AlphaFold2 models, tetramers were built for visualization purposes by superposition of four subunit models on the BmGr9 tetramer.

(E) Zoomed-in views of the predicted ligand-binding pockets from the right-hand subunit of panels A-D to highlight relative size of the pockets across (sub)families. Some stray surfaces not corresponding to the pocket volumes were removed for clarity.
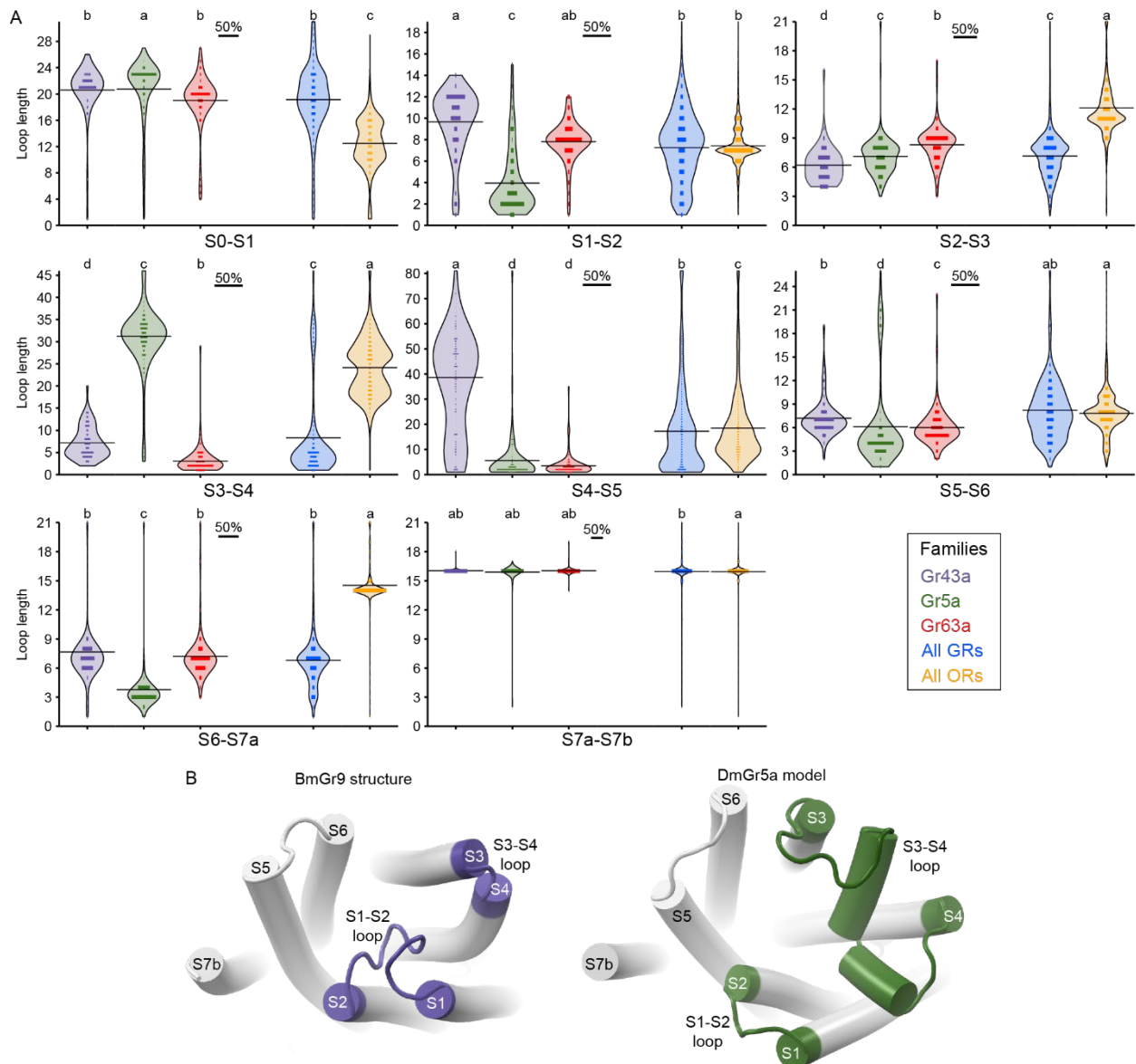
Figure S15. Distribution of loop lengths across GR subfamilies, the GR family and the OR family, Related to Figure 6

(A) Violin plots depict the distribution of lengths for each loop connecting two adjacent helices as indicated in each plot. Each plot shows the loop length (in number of residues) on the y axis and receptor (sub)family types on the x axis. The length of the lines inside the violins is proportional to the fraction of sequences in each (sub)family with the given loop lengths, with the 50% scale bars indicated at the top of each plot. Each plot shows five different violins: Gr43a subfamily (violet), Gr5a subfamily (green), Gr63a subfamily (red), all GRs (blue), and all ORs (orange). Distinct letters above each violin denote statistically distinct classes, listed alphabetically from highest median value to lowest such that adjacent categories bear adjacent letters based on a Steel-Dwass test, p<0.01.

(B) Top views of a single subunit from the agonist-free BmGr9 structure (left) and the DmGr5a AlphaFold2 model (right) showing relative differences in length and structure of their S1-S2 and S3-S4 loops.

Table S1. RMSD values in Å for superpositions of AbOrco, MhOr5, BmGr9, Gr43a, and Gr64a structures, with the number of Cα atoms used in each alignment in parentheses.

	AbOrco (6C70)	MhOr5 (7LIC)	Eugenol-bound MhOr5 (7LID)	DEET-bound MhOr5 (7LIG)	Agonist-free BmGr9 (8VC1)	Fructose-bound BmGr9 (8VC2)	Agonist-free Gr43a (8JM9)	Fructose-bound Gr43a (8JMA)	Agonist-free Gr43a I418A (8X82)	Fructose-bound Gr43a I418A (8X83)	Fructose+Ca^{2+}-bound Gr43a I418A (8X84)	Agonist-free Gr64a (8JME)	Sucrose-bound Gr64a (8JMH)	Maltose-bound Gr64a (8JMI)
AbOrco (6C70)	-	3.4 (360)	3.8 (360)	3.8 (360)	5.1 (312)	4.6 (328)	4.7 (320)	4.0 (336)	4.1 (328)	5.0 (320)	4.3 (328)	4.0 (352)	4.2 (352)	4.0 (352)
MhOr5 (7LIC)		-	1.2 (376)	1.2 (376)	4.2 (336)	4.1 (336)	4.2 (336)	4.0 (344)	4.3 (344)	4.5 (344)	4.5 (344)	3.9 (352)	3.9 (352)	3.9 (352)
Eugenol-bound MhOr5 (7LID)			-	0.7 (376)	4.4 (344)	4.4 (336)	4.5 (344)	4.1 (344)	5.1 (336)	5.1 (336)	5.1 (336)	4.2 (360)	4.2 (360)	4.2 (360)
DEET-bound MhOr5 (7LIG)				-	4.7 (336)	4.2 (344)	4.2 (336)	4.2 (336)	4.2 (336)	4.2 (336)	4.2 (336)	4.2 (360)	4.2 (360)	4.5 (360)
Agonist-free BmGr9 (8VC1)					-	1.5 (376)	2.6 (360)	2.6 (360)	2.6 (360)	2.5 (360)	2.5 (360)	3.6 (344)	3.6 (344)	3.7 (344)
Fructose-bound BmGr9 (8VC2)						-	2.5 (368)	2.3 (368)	2.2 (368)	2.3 (368)	2.2 (368)	3.8 (352)	4.0 (352)	3.5 (344)
Agonist-free Gr43a (8JM9)							-	0.5 (376)	0.5 (376)	0.6 (376)	0.6 (376)	4.0 (352)	4.0 (352)	4.0 (352)
Fructose-bound Gr43a (8JMA)								-	0.3 (376)	0.4 (376)	0.4 (376)	4.0 (352)	4.0 (352)	4.0 (352)
Agonist-free Gr43a I418A (8X82)									-	0.4 (376)	0.3 (376)	4.0 (352)	4.0 (352)	4.0 (352)
Fructose-bound Gr43a I418A (8X83)										-	0.1 (376)	4.0 (352)	4.0 (352)	4.0 (352)
Fructose+Ca^{2+}-bound Gr43a I418A (8X84)											-	4.0 (352)	4.0 (352)	4.0 (352)
Agonist-free Gr64a (8JME)												-	0.2 (400)	0.2 (400)
Sucrose-bound Gr64a (8JMH)													-	0.2 (400)
Maltose-bound Gr64a (8JMI)														-