Supplemental information

Zebrafish Olfactory Receptors ORAs Differentially Detect Bile Acids and Bile Salts

Xiaojing Cong,[2,†] Qian Zheng,[1,†] Wenwen Ren,[5,†] Jean-Baptiste Cheron,[2] Sébastien Fiorucci,[2] Tieqiao Wen,[1] Chunbo Zhang,[6] Hongmeng Yu,[3]* Jérôme Golebiowski[2,4]* and Yiqun Yu[1,3]*

From the [1] School of Life Sciences, Shanghai University, Shanghai 200444, People's Republic of China; [2] Université Côte d'Azur, CNRS, Institut de Chimie de Nice UMR7272, Nice 06108, France; [3] Department of Otolaryngology, Eye, Ear, Nose & Throat Hospital, Shanghai Key Clinical Disciplines of Otorhinolaryngology, Fudan University, Shanghai 200031, People's Republic of China; [4] Department of Brain and Cognitive Sciences, Daegu Gyeongbuk Institute of Science and Technology, Daegu 711-873, South Korea; [5] Institutes of Biomedical Sciences, Fudan University, Shanghai 200031, People's Republic of China; and [6] Department of Biological and Chemical Sciences, Illinois Institute of Technology, Chicago, IL 60616, USA

[†] These authors contributed equally to this work.

* To whom correspondence should be addressed: hongmengyush@fudan.edu.cn, jerome.golebiowski@unice.fr or yiqun_yu@shu.edu.cn

Supplemental information including:

Table S1. Structures of the steroid compounds tested by *in vitro* screening.

Physical properties of the ORA ligands identified (the top nine) were calculated by Dragon 6 [1]: LogP, partition coefficient; Sv, sum of atomic van der Waals volumes (scaled on Carbon atom).

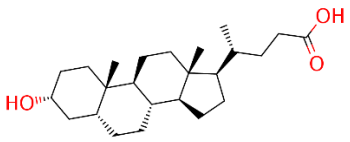
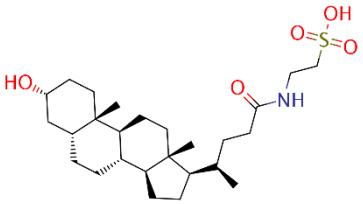
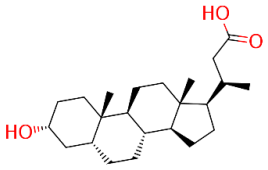
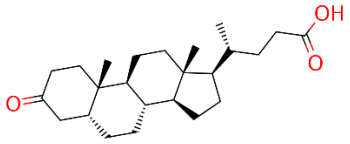
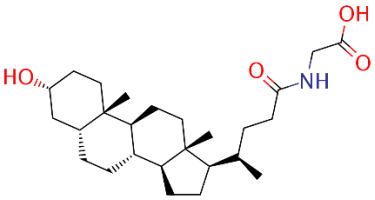
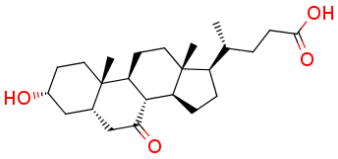
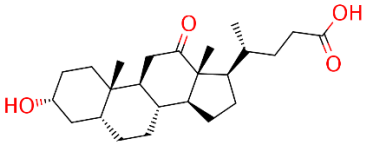
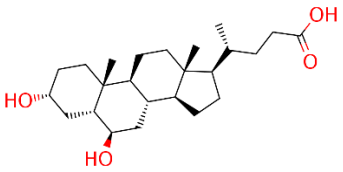
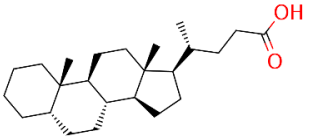
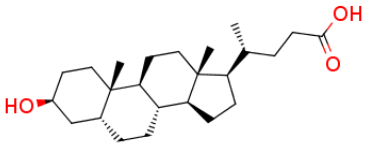
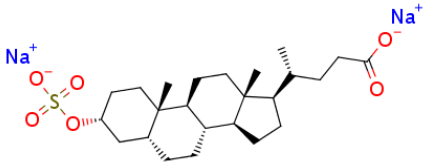
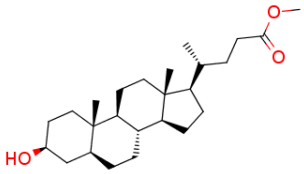
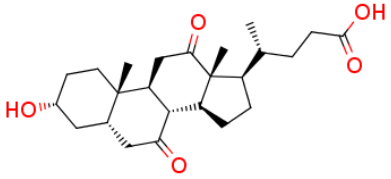
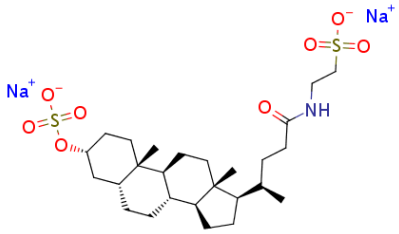
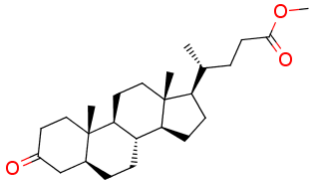
Active ORA ligands		
 Lithocholic acid (LCA), logP 4.9, Sv 37	 Glycolithocholic acid (GLCA), logP 4.0, Sv 41	 Norlithocholic acid (NLCA), logP 4.7, Sv 35
 3-ketolithocholic acid (3-KLCA), logP 4.8, Sv 36	 Taurolithocholic acid (TLCA), logP 3.5, Sv 43	 7-ketolithocholic acid (7-KLCA), logP 4.0, Sv 37
 12-ketolithocholic acid (12-KLCA), logP 4.0, Sv 37	 Murocholic acid (MCA), logP 4.1, Sv 37	 Cholanic acid (ChA), logP 5.7, Sv 36
Inactive steroid compounds		
 Isolithocholic acid (iso-LCA)	 lithocholic acid 3-sulfate disodium	 5α-cholanic acid-3β-ol
 7-12-ketolithocholic acid (7,12-KLCA)	 taurolithocholic acid 3-sulfate disodium	 5α-cholanic acid-3-one methyl ester

Table S1. Continued.

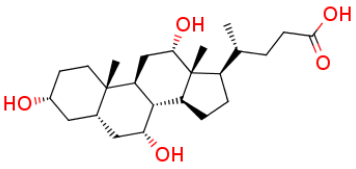
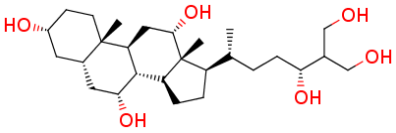
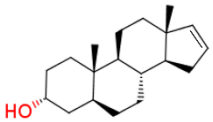
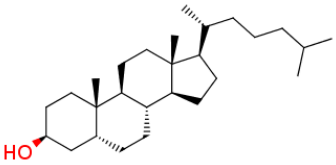
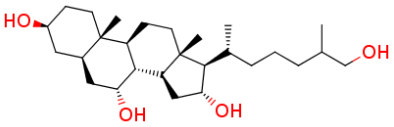
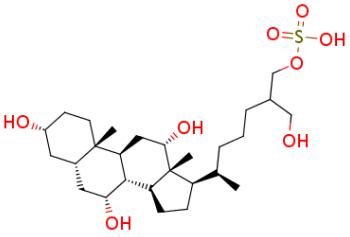
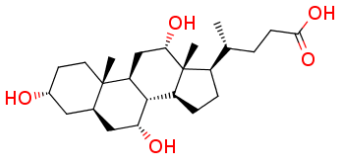
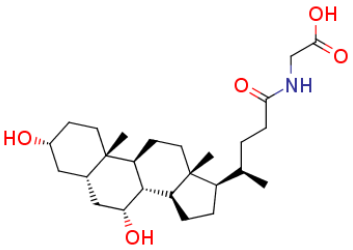
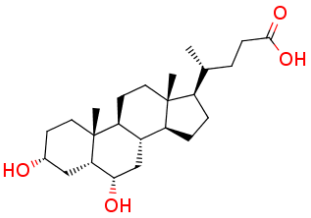
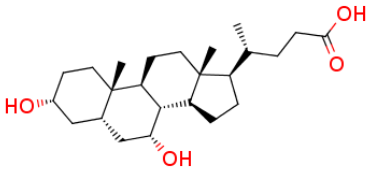
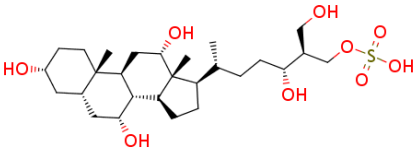
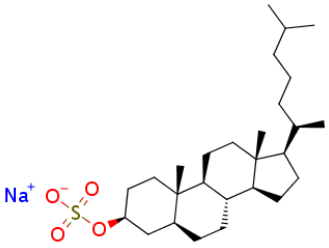
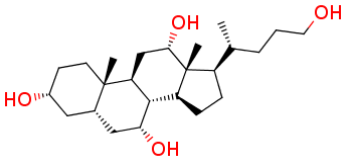
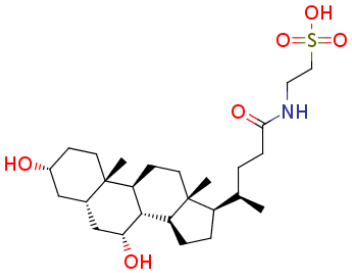
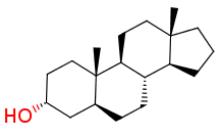
 Cholic acid	 5β-scymnol	 5α-androsten-3α-ol
 5β-cholestan-3b-ol	 5α-myxinol	 5β-cyprinol sulfate
 5α-cholic acid	 Glycochenodeoxycholic acid	 Hyodeoxycholic acid
 Chenodeoxycholic acid	 5β-scymnol sulfate	 Cholestanol sulfate sodium
 5β-petromyzonol	 Taurochendeoxycholic acid	 5α-androstan-3α-ol

Table S2. Potency of the steroid compounds for zebrafish ORAs in Hana3A cells.

logEC50 [a]	ORA1	ORA2	ORA3	ORA4	ORA5	ORA6
LCA	-4.7 ± 0.1	-4.3 ± 0.5	-4.0 ± 32.4	-3.0 ± 7.9	-4.7 ± 0.03	-4.3 ± 0.2
NLCA	-4.8 ± 0.1	-5.0 ± 7.4	-4.9 ± 0.1	-4.0 ± 0.3	-4.5 ± 1.0	-4.3 ± 0.1
3-KLCA	n.s.	-4.1 ± 0.9	n.s.	n.s.	n.s.	n.s.
7-KLCA	n.s.	n.s.	-2.9 ± 53.7	n.s.	n.s.	n.s.
12-KLCA	n.s.	n.s.	-4.2 ± 2.2	n.s.	n.s.	n.s.
TLCA	-4.7 ± 0.4	-3.6 ± 3.6	-4.5 ± 0.5	-3.5 ± 1.4	-4.3 ± 0.3	n.s.
MCA	-4.3 ± 0.3	-4.4 ± 0.4	n.s.	n.s.	n.s.	n.s.
GLCA	-3.4 ± 2.9	-4.2 ± 5.1	-4.2 ± 8.7	n.s.	-4.1 ± 1.0	n.s.
ChA	n.s.	n.s.	n.s.	n.s.	n.s.	-4.0 ± 1.4

[a] Mean ± SEM, n = 3–9. (n.s.) non-significant response up to 100 µM.

Table S3. Zebrafish behavioral responses to ORA ligands.

Stimulus	*n*	Displacement [a] (% of tank length)	*p* value [b]	Average velocity [c] (% of no stimulus)	*p* value
No stimulus	-	-8.0 ± 1.5	-	100.0 ± 33.2	-
Tank water (negative control)	5	-11.6 ± 1.2	0.5944	n.a.	n.a.
Food (positive control)	6	25.4 ± 4.3	<0.0001	92.9 ± 36.2	0.7174
LCA	12	28.0 ± 2.6	<0.0001	61.9 ± 41.1	0.1315
ChA	11	9.0 ± 4.6	0.0110	68.1 ± 26.2	0.1409

[a] Mean ± SEM, measured as average distance to initial fish position, expressed in percentage of tank length after stimulus addition.
[b] Tested with one-way ANOVA and corrected with Holm-Sidak multiple comparison test. (n.a.) not available.
[c] Mean ± SEM, fish track length summed over measurement time (30 frames/second during 5 minutes each phase).

Table S4. ORA sequence identity and similarity.

Pair-wise percentage sequence identity (lower triangular) and similarity (upper triangular) of the predicted transmembrane domain among the ORAs and the template.

(%)	hAGTR1	ORA1	ORA2	ORA3	ORA4	ORA5	ORA6
hAGTR1	100	32	32	37	37	30	39
ORA1	16	100	56	36	37	37	39
ORA2	13	38	100	42	38	40	47
ORA3	19	19	24	100	44	35	38
ORA4	16	20	22	27	100	31	33
ORA5	16	17	19	18	16	100	43
ORA6	22	16	23	15	17	25	100

Table S5. Effects of point mutations on the ORAs' response to ligands.

ORA5-LCA	**Efficacy (% of *wt*) [a]**	***p* value [b]**	**ORA6-LCA**	**Efficacy (% of *wt*)**	***p* value**
$R12^{1.39}A$	125 ± 22.2	n.s.	$R48^{1.39}A$	55.6 ± 14.6	0.0288
$P62^{2.60}Y$	44.2 ± 23.3	0.0293	$Y97^{2.60}A$	60.8 ± 12.9	0.0288
$D87^{3.33}A$	24.1 ± 19.0	0.0022	$V121^{3.33}D$	150 ±16.1	0.0288
$Y177^{5.42}A$	24.2 ± 19.0	0.0022	$K215^{5.39}A$	55.2 ± 14.6	0.0288
$H252^{6.51}A$	53.2 ± 19.0	0.0293	$Y286^{6.51}H$	65.8 ± 14.6	0.0484
$R256^{6.55}H$	29.6 ± 21.2	0.0064	$F306^{7.39}A$	46.2 ± 14.6	0.0065
$K269^{7.36}A$	25.6 ± 20.0	0.0033	$T309^{7.42}A$	47.1 ± 16.1	0.0288
$N272^{7.39}A$	19.0 ± 23.3	0.0055	**ORA5-ChA**	**Efficacy (% of *wt*)**	***p* value**
$N272^{7.39}F$	238 ± 21.2	<0.0001	$P62^{2.60}Y$	343 ± 96.4	0.0417
ORA2-LCA	**Efficacy (% of *wt*)**	***p* value**	$E90^{3.36}A$	395 ± 102	0.0291
$R61^{2.57}A$	33.8 ± 16.2	0.0355	$Y177^{5.42}F$	450 ± 110	0.0218
			$N272^{7.39}F$	344 ± 96.4	0.0417

[a] Mean ± SEM at 100 μM concentration, n = 3–8.

[b] Tested with one-way ANOVA and corrected with Holm-Sidak multiple comparison test.

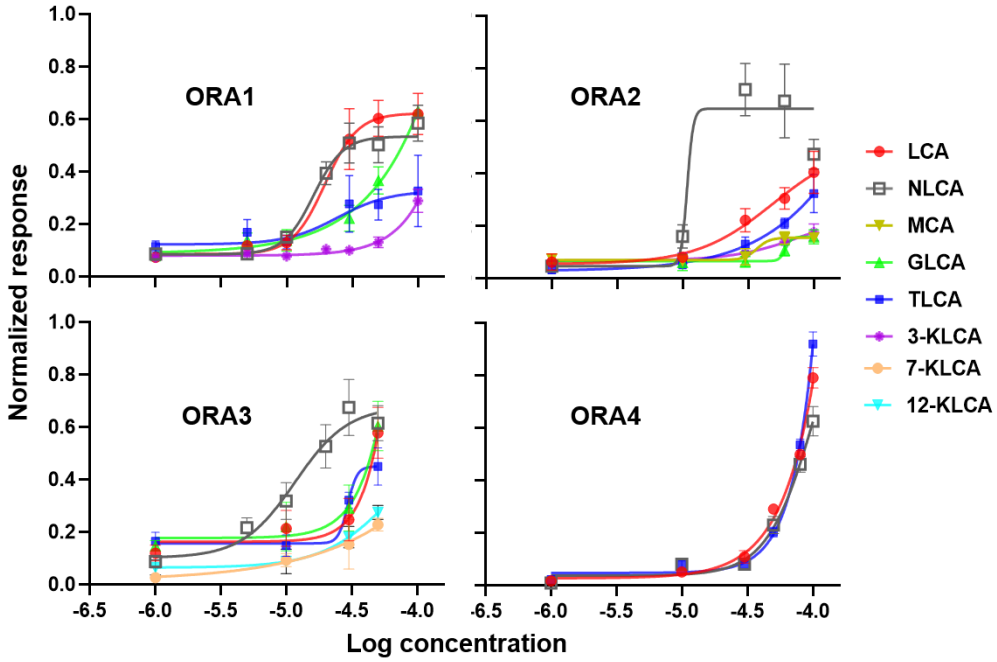


Fig. S1 Dose-response curves for the ORA ligands identified.

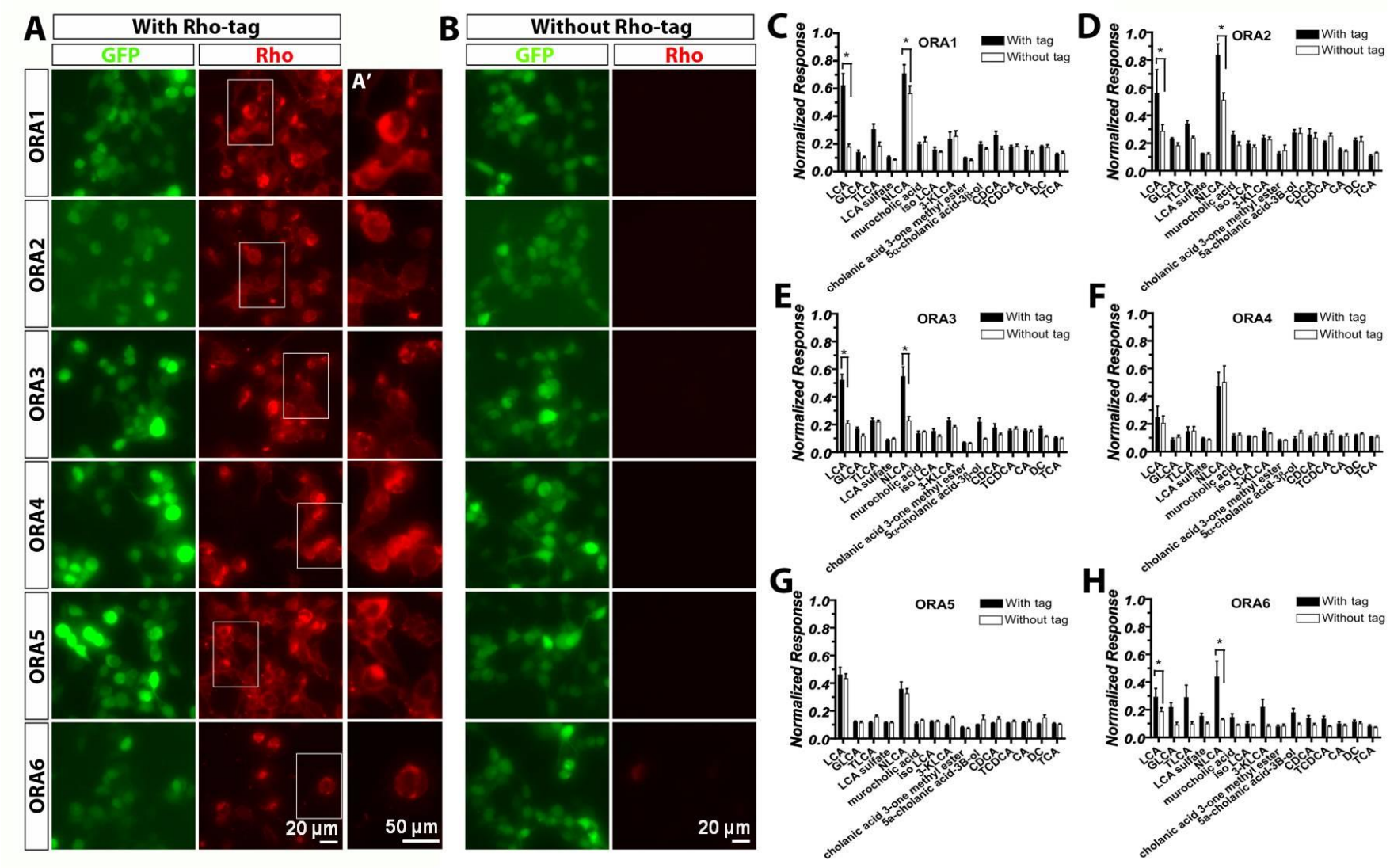


Fig. S2 Rhodopsin tag facilitated surface expression of zebrafish ORAs in Hana3A cells. (A, B) Cells were transfected with Rho-tagged ORAs and green fluorescent protein (GFP, green) and underwent live cell immunostaining with Rho-antibody (red). The square areas in the Rho staining images are magnified in (A'), showing the surface localization of the Rho-tagged receptors. **(C-H)** Responses of Rho-tagged and untagged ORAs to the compounds at 30 μM concentration.

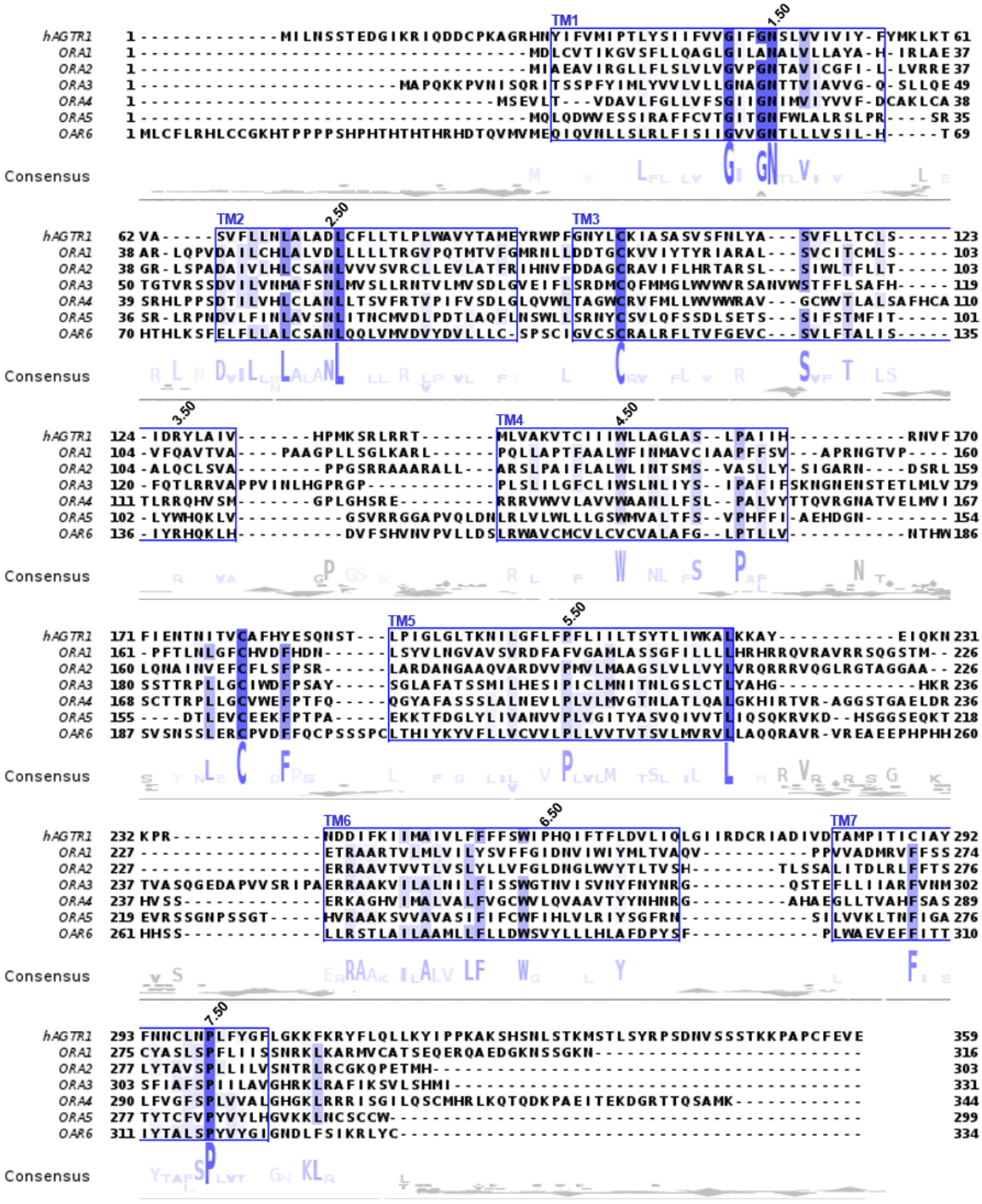


Fig. S3 Sequence alignment of zebrafish ORAs with hAGTR1. The TM domains are boxed. The most conserved position in each TM helix is labeled with the Ballesteros-Weinstein number.

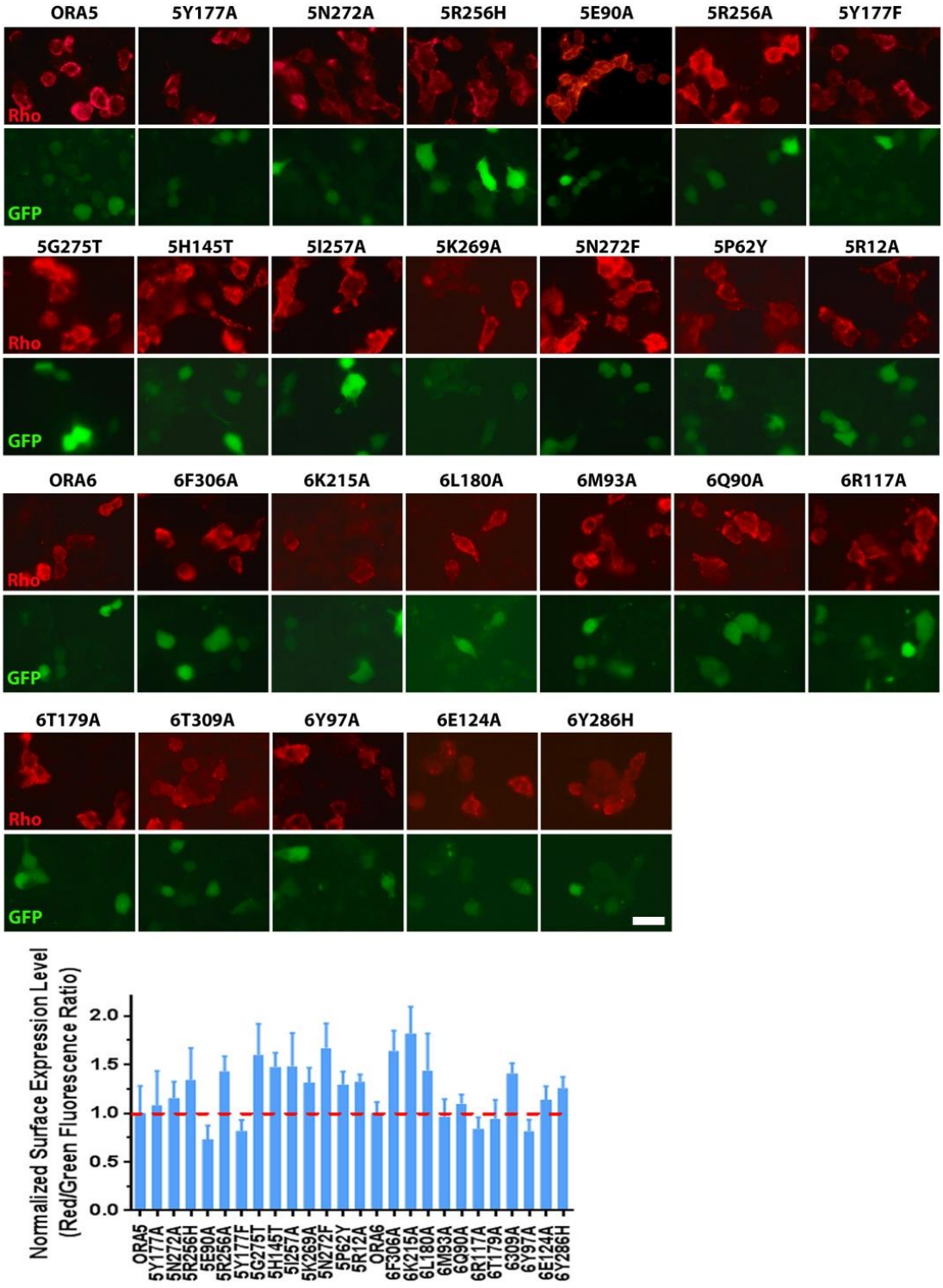


Fig. S4 Surface expression of ORA5 and ORA6 *wt* and mutants verified by fluorescence activated cell sorting (FACS). The ratio of red (Rho^+)/green (GFP^+) fluorescence intensity (normalized to that in *wt* ORA5 or ORA6) was used to evaluate the surface expression level of the mutants. Scale bar: 20 μm.

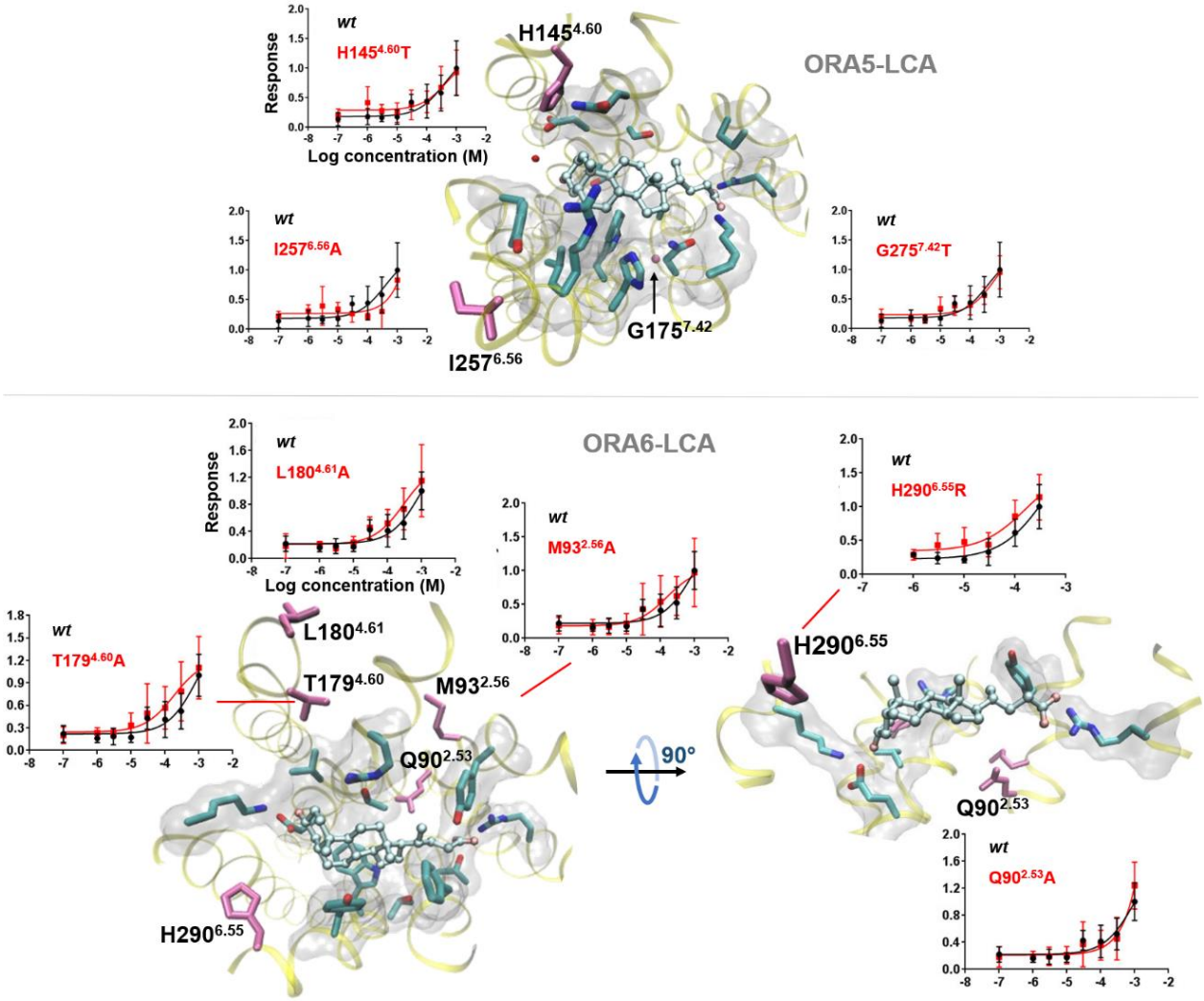


Fig. S5 Control mutations that have marginal effects on the receptors' response to LCA. The mutation sites are in magenta sticks.

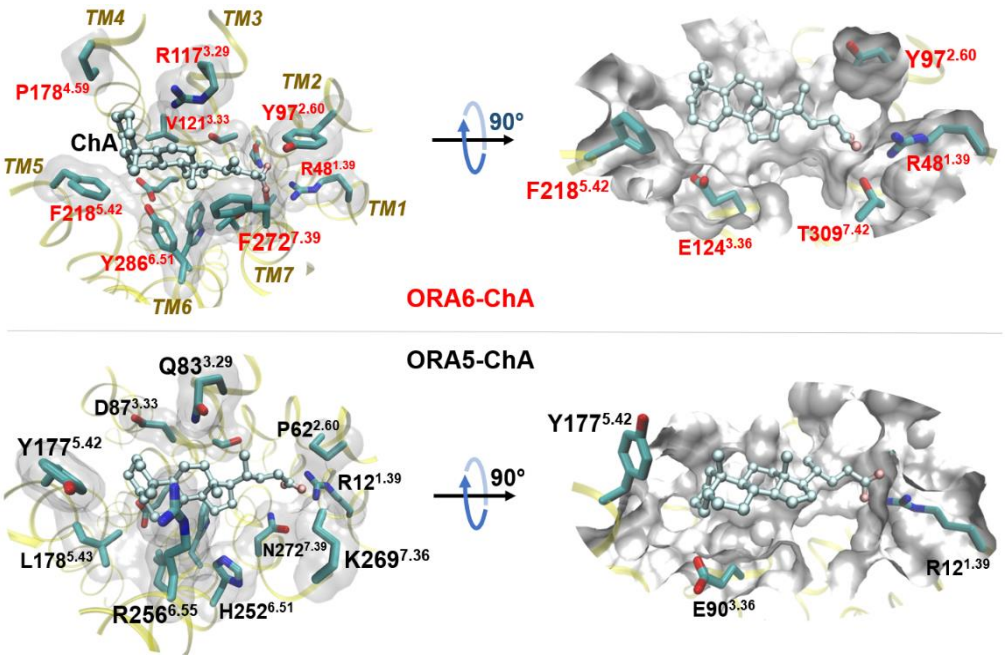


Fig. S6 Predicted ChA binding mode in *wt* ORA5 and ORA6.

Supplementary reference

1. Todeschini R, Consonni V (2009) Molecular Descriptors for Chemoinformatics, vol I/II. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany