Supporting Information

Structural architecture of a dimeric class C GPCR based on co-trafficking of sweet taste receptor subunits

Jihye Park[1], Balaji Selvam[2], Keisuke Sanematsu[3,4], Noriatsu Shigemura[3,4], Diwakar Shukla[2], and Erik Procko[1,*]

Table S1-S4
Figure S1-S12

Table S1. Residues at predicted contact sites.

Sites	Involved residues
I	V108, L112, L122, I124, Y128, S129, Y131, F153, L156, F157, L158, W418
II	H511, V512
III	K506, P507, I518, C520, L521, P522, T524, Y533, E534, C535, W543
IV	P594, F 743, A746, Y747, K750, E751, N755, Y756, K760, F761, T763, L764, T767, F770, T771, V774, T778, F779, S781

Table S2. Residues restrained within proximity of the central axis during ROSETTA relaxation.

T1R2 residues (Cα atoms)	T1R3 residues (Cα atoms)
Phe153	Phe156
Val512	Ser516
Thr763	Thr765
Ala782	Asn784

Table S3. Restrained residues during MD simulation (10-20 ns).

Receptor	Restrained residues (Cα atoms)
T1R2	Extracellular region: 145-156, 214-229
	Transmembrane region: 757-783
T1R3	Extracellular region: 148-160, 217-234
	Transmembrane region: 759-784

Table S4. Residue pairs with distance constraints in the MD simulation (20-50 ns).

Domain	Distance restraints (T1R2 - T1R3)
LBD	Phe53 - Trp424
	Leu156 - Met110
	Gln221- Ser224
CRD	His551 - His515
	Leu521- Glu525
TMD	Tyr783 - Asn784
	Phe779 - Leu781
	Thr771 - Ile773
	Thr763 - Phe766

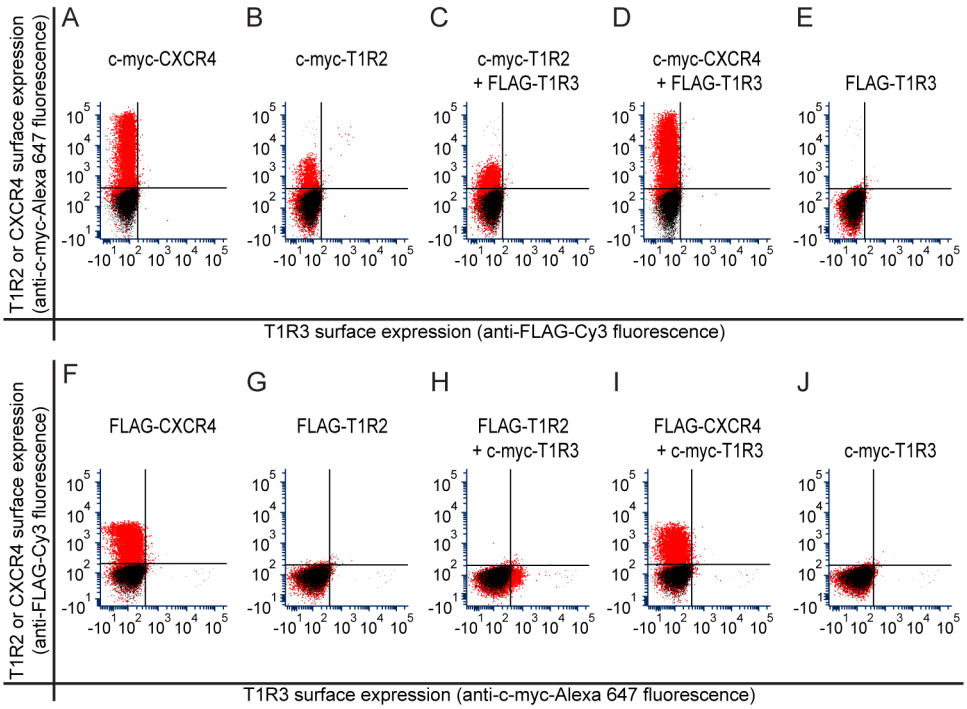

Figure S1. Surface expression of T1R2 and T1R3 with various N-terminal epitope tags detected with fluorophore-conjugated antibodies. Flow cytometry analysis of cells stained with anti-c-myc-Alexa 647 and anti-FLAG-Cy3. All the T1R2 and T1R3 constructs are cloned in pcDNA3.1(+). Cells transfected with an empty vector are shown in black, while cells transfected with expression plasmids are red. The anti-c-myc-Alexa 647 antibody gives greater signal-to-noise (c.f. panels A and F expressing N-terminally tagged CXCR4 as a positive control) and shows that T1R2 can escape to the cell surface by itself (B). T1R3 is not detected on the cell surface (E and J) unless co-transfected with T1R2 (H).

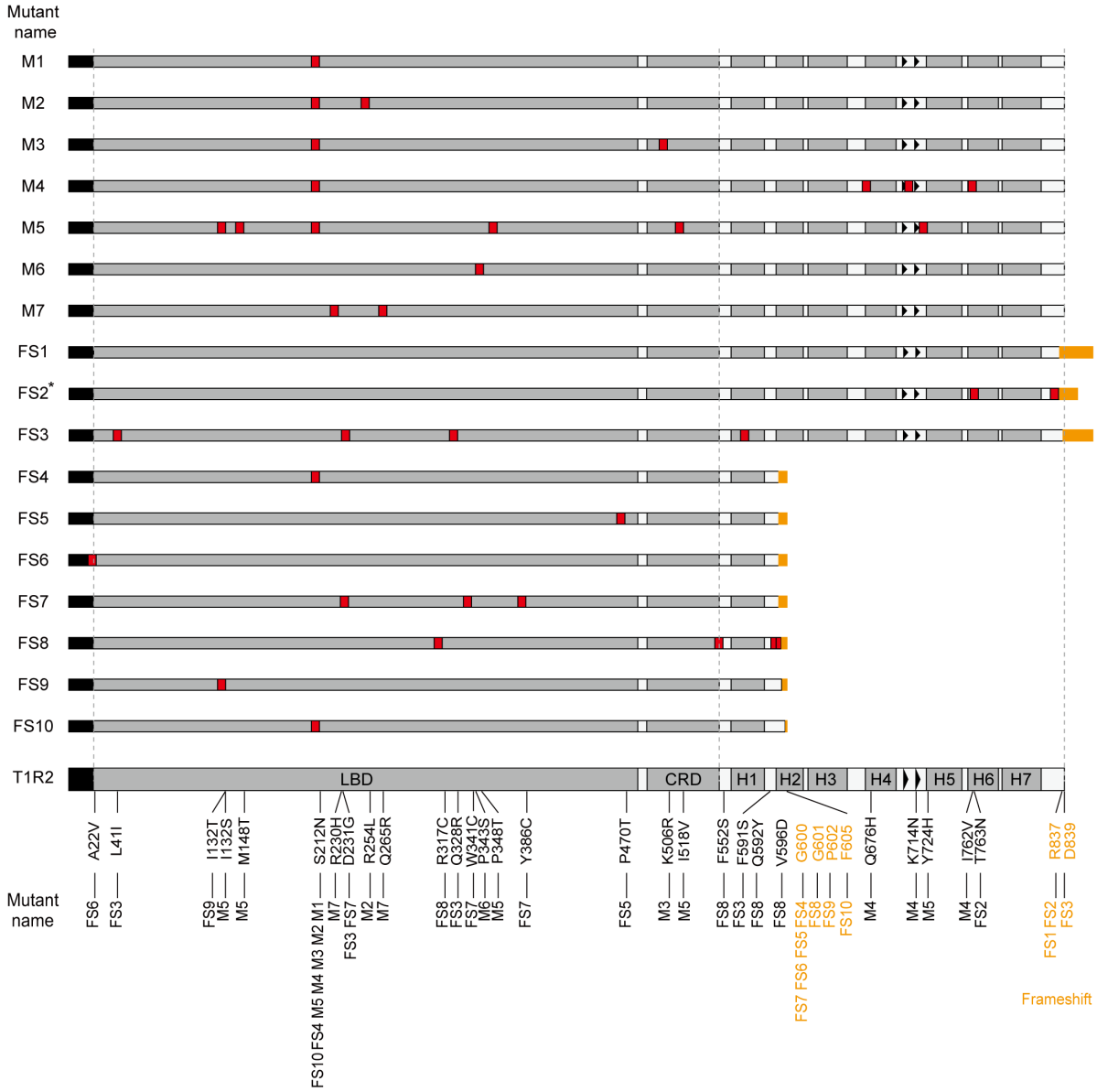

Figure S2. Summary of T1R2 mutants with enhanced surface expression. Schematic diagram showing mutations present in 17 T1R2 clones isolated from a selection for enhanced surface expression. Missense mutations are shown as red boxes. Random sequences generated by frameshifts are shown as orange boxes. *, one missense mutation in the signal peptide.

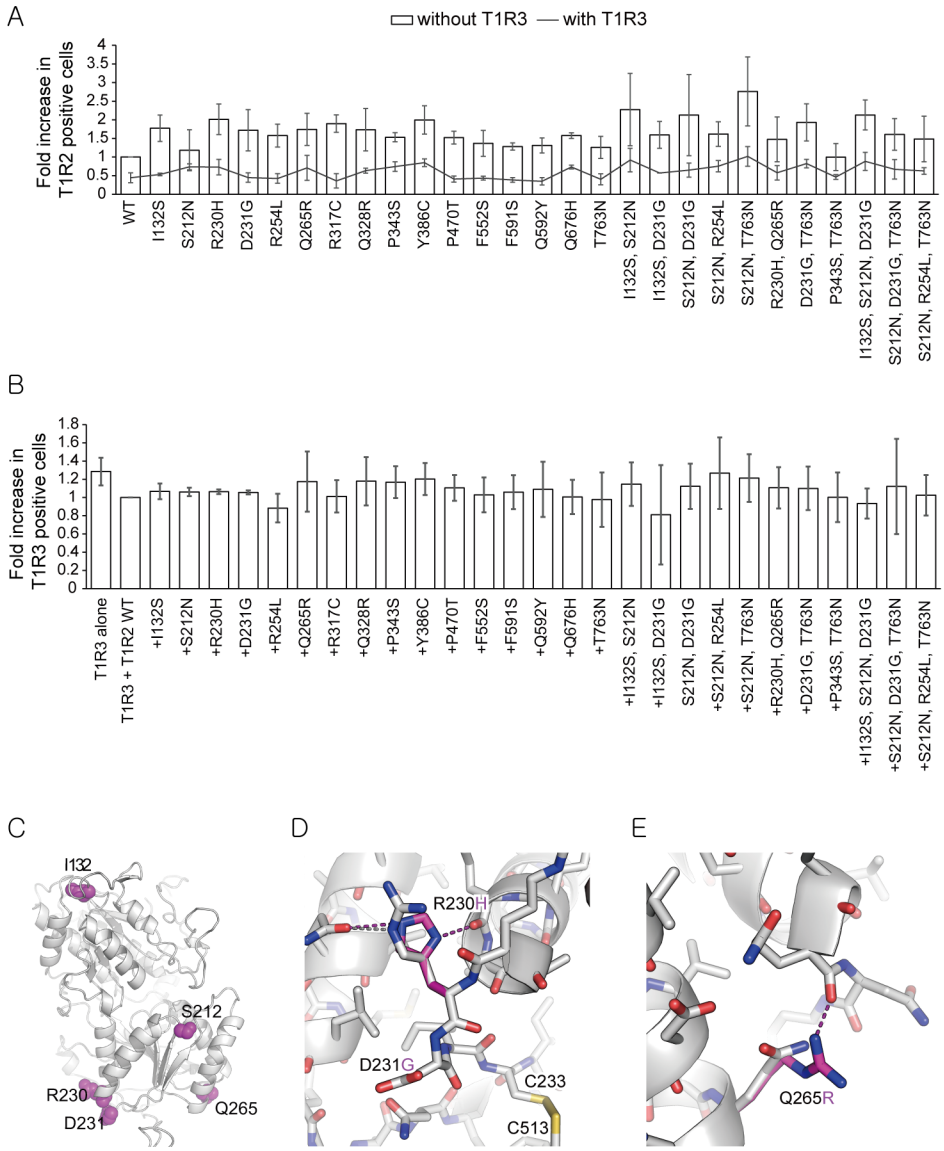

Figure S3. Total cell expression of individual T1R2 mutants. (A) Quantified flow cytometry data of total T1R2, either when T1R2 is expressed alone (white bars) or together with T1R3 (black line). The gate for T1R2 (anti-FLAG-Cy3 fluorescence) positive cells was set at 0.5% of control cells. The percentage of positive cells was normalized to the percentage of cells expressing wildtype T1R2 alone. Data are averaged over at least three experiments (± SD). (B) Quantified flow cytometry data of total T1R3 expression when co-expressed with T1R2. The gate for T1R3 (anti-c-myc-Alexa 647 fluorescence) positive cells was set at 0.5% of control cells. The percentage of positive cells was normalized to the percentage of positive cells co-expressing wildtype T1R2 and T1R3. Data are averaged over at least three experiments (± SD). (C) Structural model of T1R2 (grey cartoon) showing the positions of five residues that when mutated enhance surface expression of the T1R2-T1R3 sweet taste receptor. The five residues are shown as magenta spheres. (D) T1R2 mutations R230H and D231G may favor a loop conformation. The modeled mutant structure is overlaid in magenta. (E) T1R2 (grey) mutation Q265R (mutant side chain in magenta) may stabilize the backbone conformation of a nearby loop.

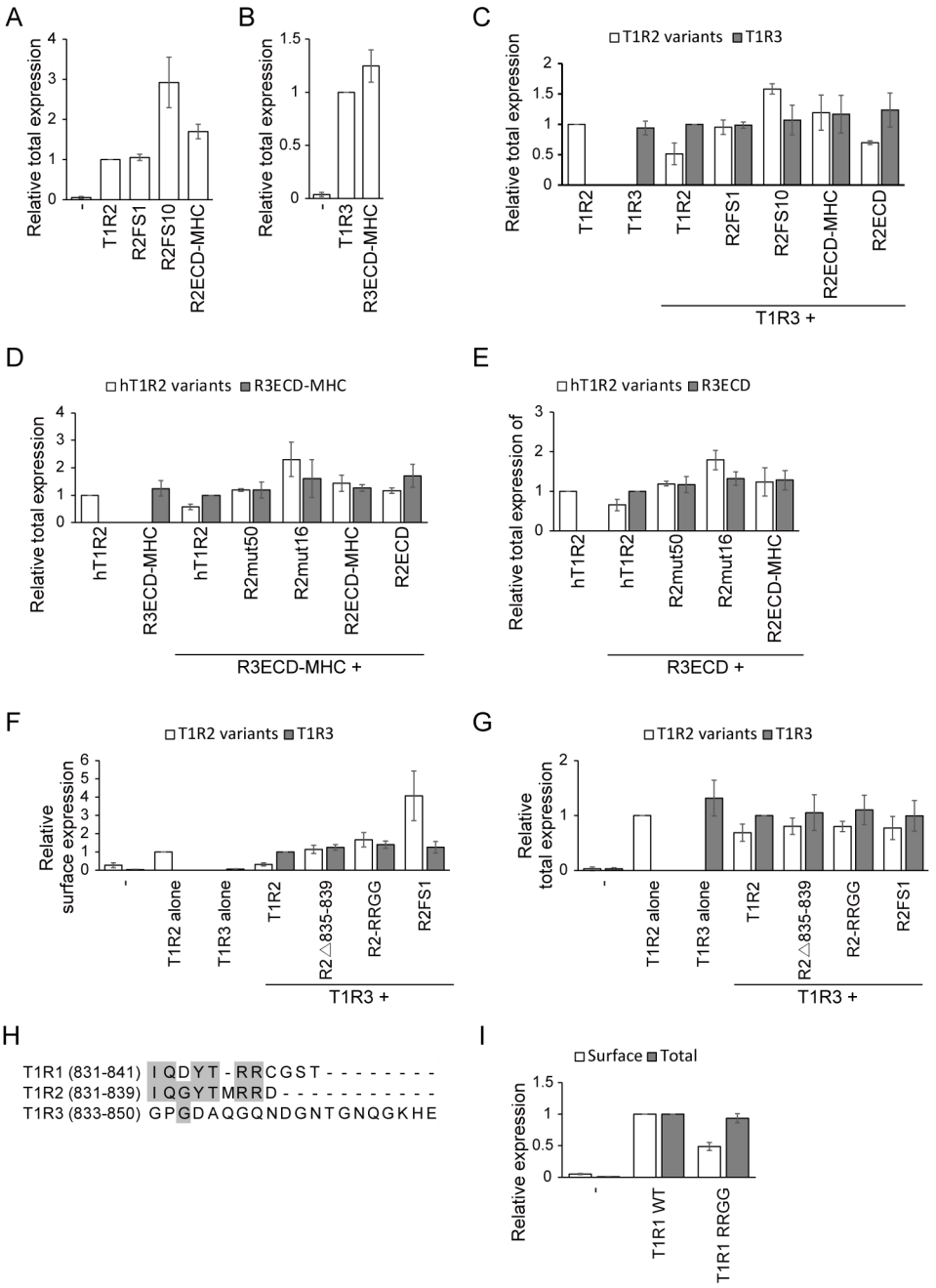

Figure S4. Total expression of T1R1, T1R2, and T1R3 constructs. (A and B) Quantified flow cytometry data under permeabilizing conditions to measure total expression of FLAG-T1R2 (A) and c-myc-T1R3 constructs (B) expressed individually. The gates for T1R2 (anti-FLAG-Cy3) and T1R3 (anti-c-myc-Alexa 647) positive cells were set at 0.5% of control cells. The percentage of gated cells for each construct was normalized to the percent positive cells expressing wildtype protein. Mean ± SD, n = 3. (C and E) Quantified flow cytometry data of total expression of FLAG-T1R2 constructs (white bars) co-expressed with, in grey bars, full-length T1R3 (C), R3ECD-MHC (D), and R3ECD (E). Gates for positive cells were set at 0.5% of negative control cells. The percent positive cells for each construct was normalized to the percent positive cells expressing T1R2 alone (for T1R2 expression) or co-expressing wildtype T1R2 with the relevant wildtype T1R3 construct (for T1R3 expression). Mean ± SD, n = 3. (F and G) Quantified flow cytometry data of surface (F) and total (G) expression of T1R2 clones co-expressed with c-myc-T1R3. Gates for T1R2 (anti-FLAG-Cy3) and T1R3 (anti-c-myc-Alexa 647) positive cells were set at 0.5% of control cells. Percentage of gated cells was normalized to the percent positive cells expressing T1R2 alone (for T1R2, white) or to cells

co-expressing T1R2 and T1R3 (for T1R3, grey). Data are averaged over at least ten experiments (± SD). (H) Sequence alignment of the C-terminal tails of human T1R1, T1R2, and T1R3. (I) Quantified flow cytometry data of surface (white) and total (grey) expression of c-myc-T1R1 constructs. Gate for T1R1 (anti-c-myc-Alexa 647 fluorescence) positive cells was set on 0.5% of negative control cells. Percentage of positive cells was normalized to the wildtype T1R1 sample. T1R1 RRGG has the two terminal arginines substituted to glycines. Mean ± SD, n = 3.

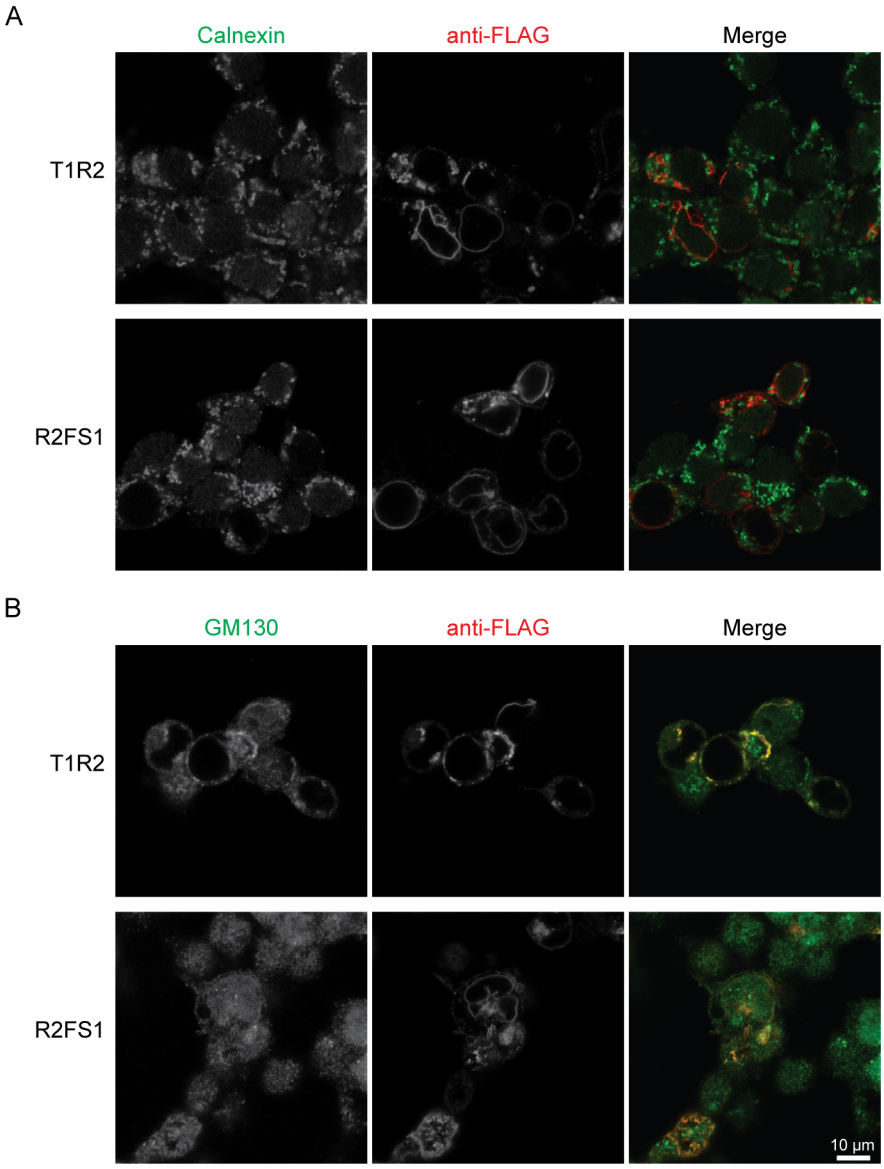

Figure S5. Subcellular localization of T1R2 and R2FS1. HEK293T cells transiently transfected with FLAG-tagged T1R2 or R2FS1 were fixed, permeabilized, and co-stained with anti-FLAG-Cy3 and with ER marker anti-Calnexin (A) or Golgi marker anti-GM130 (B). In the merged images, T1R2 and R2FS1 are red, and ER and Golgi markers are green. T1R2 and R2FS1 colocalize with GM130.

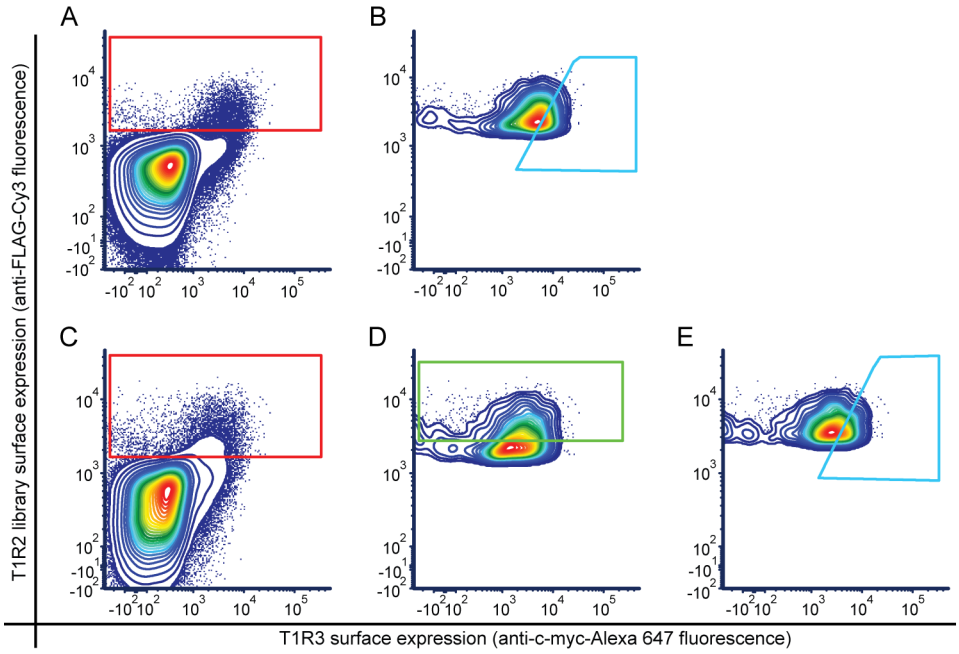

Figure S6. Sorting of T1R2 SSM libraries for surface expression and heterodimerization. An Expi293F cell line stably expressing c-myc-tagged T1R3 was transfected with FLAG-T1R2 SSM libraries, and stained with anti-FLAG-Cy3 and anti-c-myc-Alexa 647. Cells were first gated by FSC-SSC to include the main singlet population, and propidium iodide-positive dead cells were excluded. To sort cells based on T1R2 surface expression, gates were set on the total Cy3-positive population for ECD libraries (A) or on 55% of the Cy3-positive population for the TMD library (C shows total Cy3-positive cells with a red gate, and D shows in green 55% within the red gate). To sort for cells co-expressing T1R3, the previous gates were used for T1R2 expression, plus additional gates were set to sort for the top 20% of Alexa 647-positive cells (B, ECD libraries; and F, TMD library).

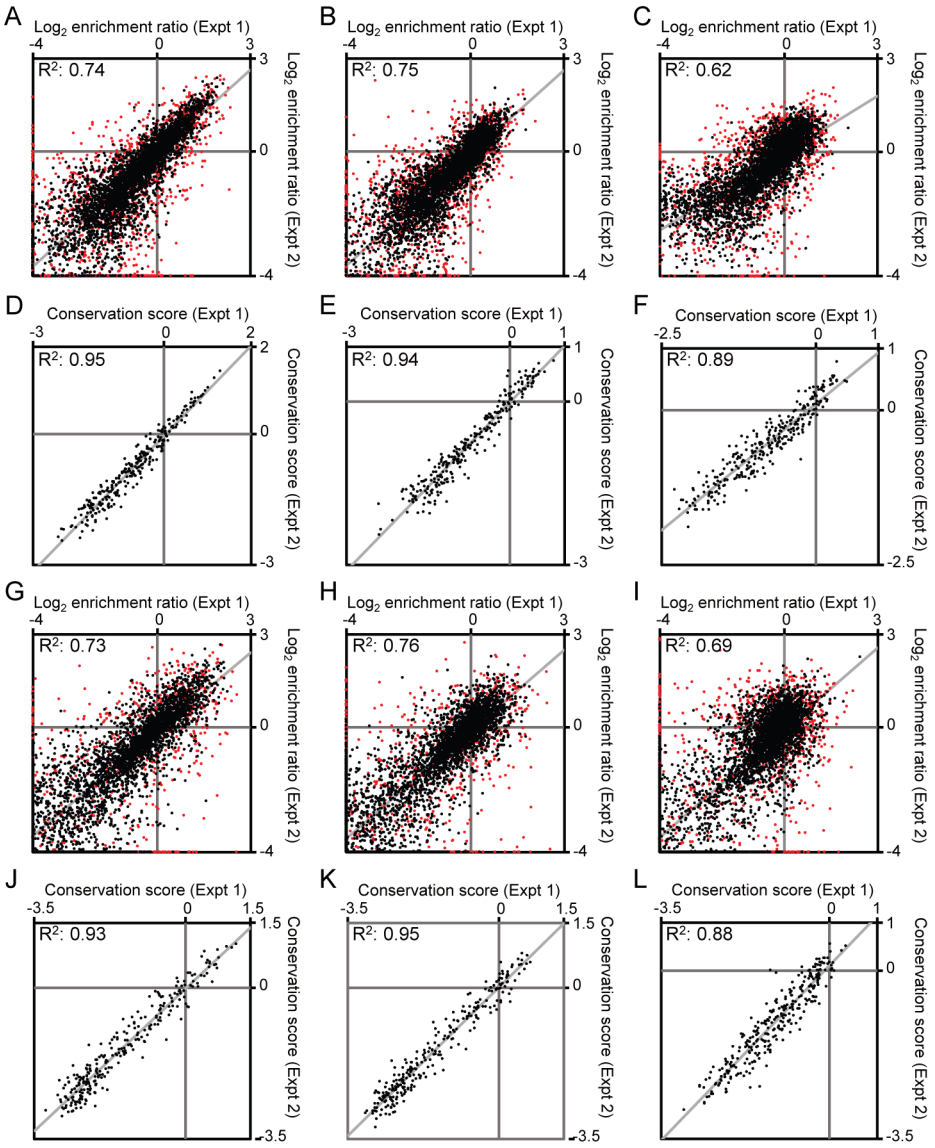

Figure S7. Agreement between replicate experiments. All in vitro selections were independently duplicated. Plots show correlations between replicates of Log_2 enrichment ratios for each amino acid substitution (A-C and G-I) and conservations scores (D-F and J-L). The T1R2 ECD1 (A, D, G, and J), ECD2 (B, E, H, and K), and TMD (C, F, I, and L) SSM libraries were sorted for T1R2 surface expression (A-F) and dimerization/co-trafficking with T1R3 (G-L). For correlation plots of Log_2 enrichment ratios, R^2 values are calculated for mutations that were present in the naïve libraries with frequencies $> 5 \times 10^{-5}$ (black). Rarer mutations in the naïve libraries (frequencies between 5×10^{-5} and 5×10^{-6}) are shown in red. Mutations with a frequency $< 5 \times 10^{-6}$ are considered absent from the naïve libraries and are not plotted.

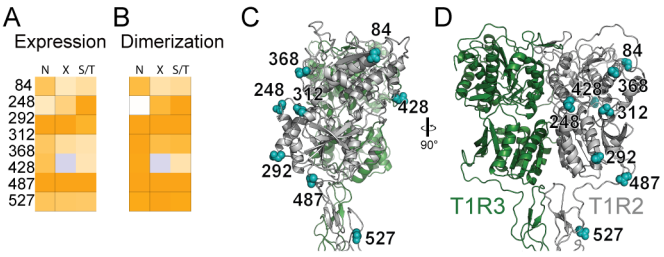

Figure S8. Putative N-linked glycosylation sites in T1R2 are generally intolerant of mutations. (A and B) Heatmaps plotting average conservation scores of residues within N-linked glycosylation motifs of T1R2 for surface expression (A) and co-trafficking/dimerization with T1R3 (B). Conservation scores are plotted from ≤ -1.5 (mutations are deleterious, orange) to 0 (mutations are on average neutral, white) to ≥ +1.5 (mutations are enriched, dark blue). Residue numbers are of asparagine residues. (C and D) Cartoon representations of T1R2 (gray) and T1R3 (green) ECD heterodimer model. Asparagine residues predicted to be glycosylated are shown as cyan spheres.

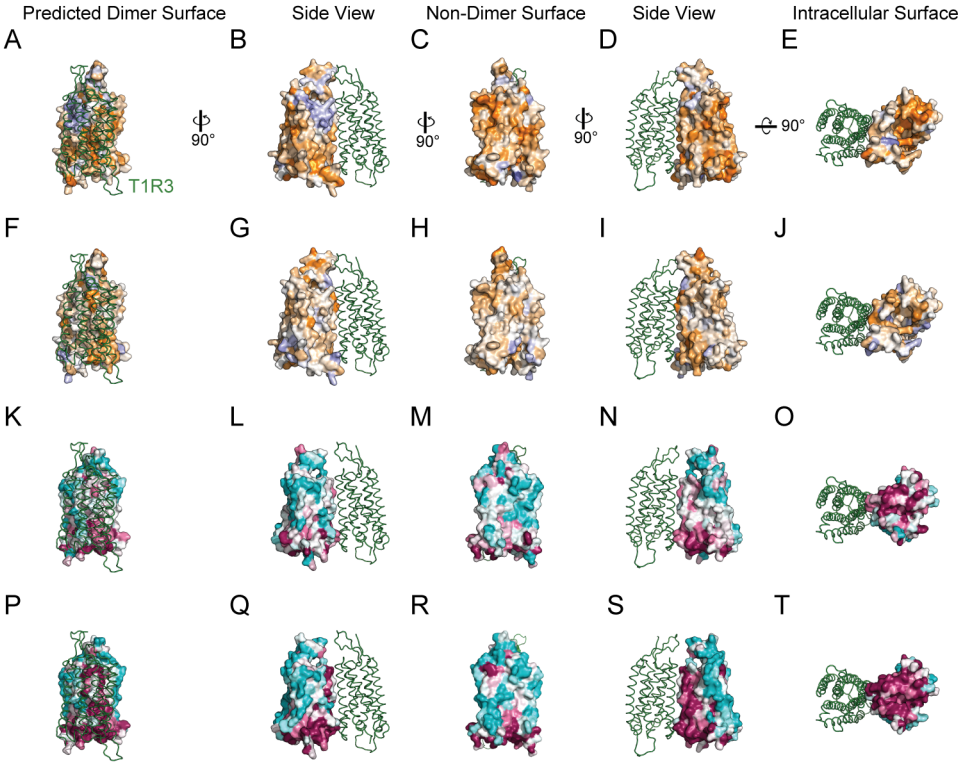

Figure S9. Comparison of sequence conservation in the TMD during in vitro selection and natural evolution. (A-J) Conservation scores from the deep mutational scan for T1R2 surface expression (A-E) and difference scores highlighting residues under preferential conservation for T1R2-T1R3 surface co-expression (F-J) are mapped to the structural surface of modeled T1R2 TMD. T1R2 expression conservation scores are colored from ≤ -2 (conserved, orange) to ≥ +2 (under selection for change, blue), and difference conservation scores are colored from ≤ -1.5 (more conserved for T1R2-T1R3 dimerization and co-expression, orange) to ≥ +1.5 (more conserved for T1R2 surface expression alone, blue). (K-T) Evolutionary conservation mapped to the structural surfaces of T1R2 (K-O) and mGluR1 (P-T; PDB 4OR2) TMD. The dimeric partner is shown as a green ribbon. For the dimer of the mGluR1 TMD, the second subunit is positioned based on our dimeric T1R2-T1R3 model, not based on the crystal structure which shows a different crystallographic dimer arrangement. Residues are colored from variable in turquoise to highly conserved in maroon. All panels are vertically aligned in the same orientations. There is a common conserved surface in both the deep mutational scan for T1R2-T1R3 co-trafficking and in natural history that is anticipated to form the dimerization interface.

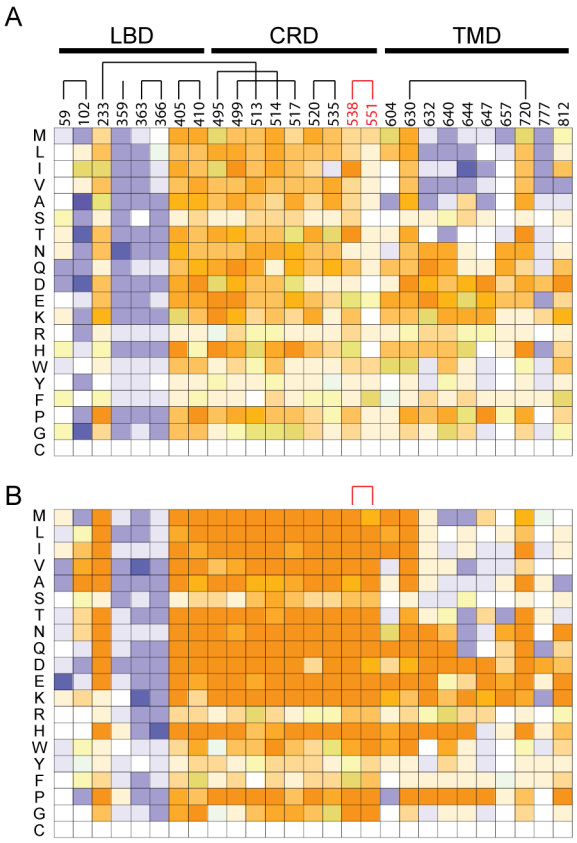

Figure S10. Sequence-activity landscapes of T1R2 cysteine residues. Sequence-activity landscapes of T1R2 cysteine residues based on selection for T1R2 surface expression (A) or co-trafficking with T1R3 (B). Log_2 enrichment ratios are plotted from ≤ -3 (depleted, orange) to 0 (neutral, white) to ≥ +3 (enriched, blue). Data are average from two independent selections.

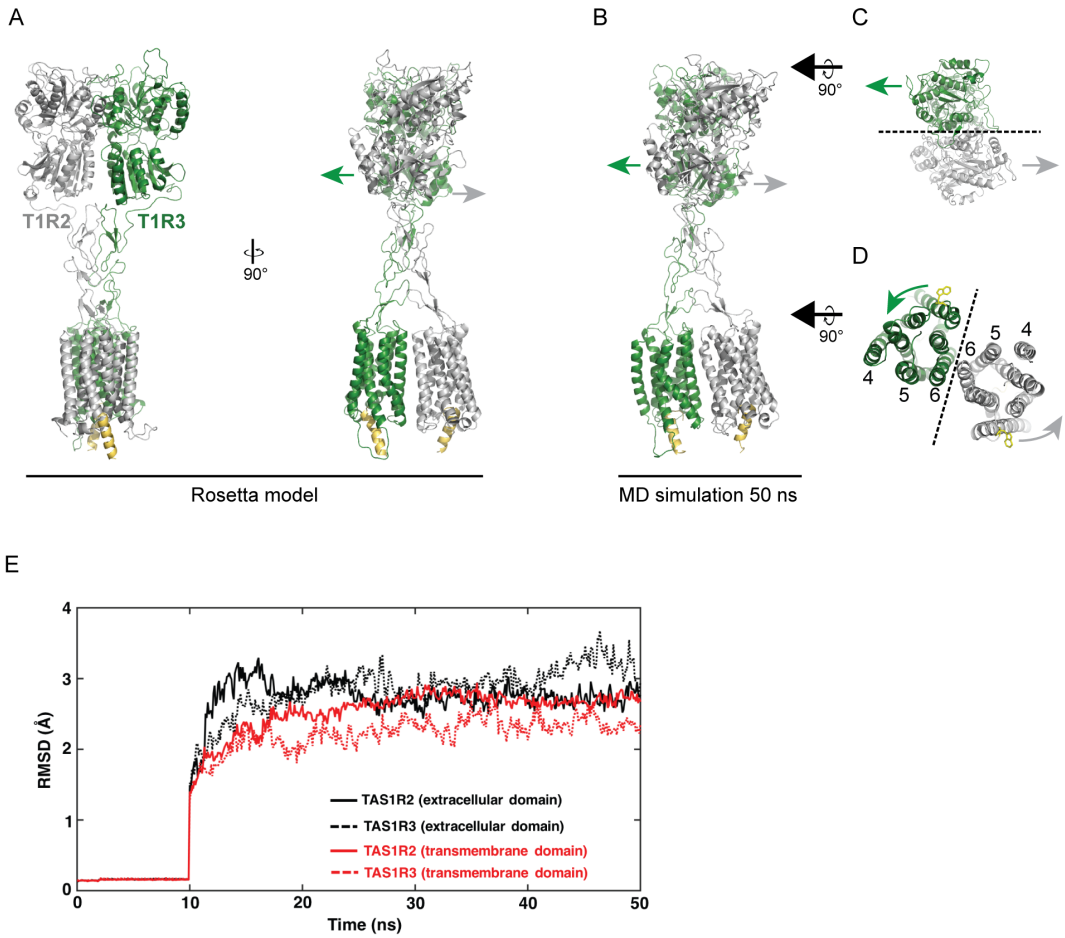

Figure S11. Modeling of the full-length T1R2-T1R3 heterodimer. Representative model structures based on Rosetta (A) and MD (B, C, and D) simulations. C-terminal peptides of Gα proteins are yellow. Arrows in A, B, and C show the translational movement of lobe 2 to the ligand-free state observed in ECD crystal structures of class C GPCRs. Arrows in d show the possible rotational movement of TMDs to the ligand-free state where TM4/5 helices are in proximity (37)(*37*). The CRD C-termini will swing outwards during LBD opening, but the orientation of the TMDs is restrained orthogonal to the plane of the membrane. Therefore as the LBDs open, the 'twisted' architecture of the receptor means the CRDs may push down on ECLs 2, promoting transmembrane domain rotation to accommodate steric strain. For spatial reference, conserved tryptophan residues near the N-terminal ends of TM1 are shown in yellow sticks. (E) RMSD analysis of extracellular and transmembrane domains during MD simulation. Domains converge with a mean RMSD of 2.5 Å, showing that the simulations are stable.

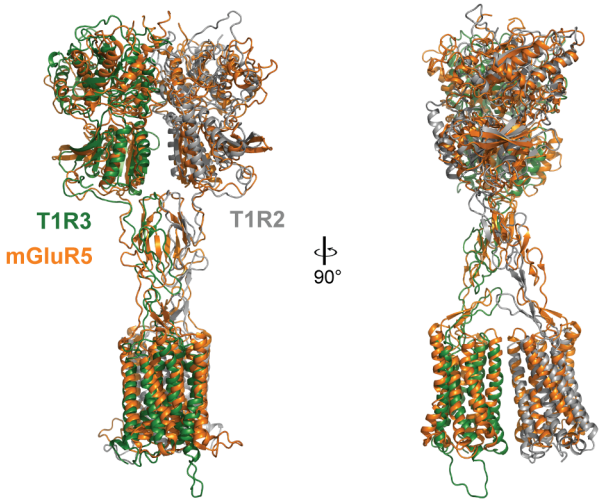

Figure S12. The modeled sweet taste receptor closely matches the cryo-EM structure of mGluR5 in an active conformation. Alignment of the modeled sweet taste receptor (grey and green ribbons) to the cryo-EM structure of mGluR5 (orange; PDB 6N51).