

# Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*

Pascal Egloff, Matthias Hillenbrand, Christoph Klenk, Alexander Batyuk, Philipp Heine, Stefanie Balada, Karola M. Schlinkmann, Daniel J. Scott, Marco Schütz, and Andreas Plückthun<sup>1</sup>

Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland

Edited by K. Christopher Garcia, Stanford University, Stanford, CA, and approved December 20, 2013 (received for review September 24, 2013)

Crystallography has advanced our understanding of G protein-coupled receptors, but low expression levels and instability in solution have limited structural insights to very few selected members of this large protein family. Using neurotensin receptor 1 (NTR1) as a proof of principle, we show that two directed evolution technologies that we recently developed have the potential to overcome these problems. We purified three neurotensin-bound NTR1 variants from *Escherichia coli* and determined their X-ray structures at up to 2.75 Å resolution using vapor diffusion crystallization experiments. A crystallized construct was pharmacologically characterized and exhibited ligand-dependent signaling, internalization, and wild-type-like agonist and antagonist affinities. Our structures are fully consistent with all biochemically defined ligand-contacting residues, and they represent an inactive NTR1 state at the cytosolic side. They exhibit significant differences to a previously determined NTR1 structure (Protein Data Bank ID code 4GRV) in the ligand-binding pocket and by the presence of the amphipathic helix 8. A comparison of helix 8 stability determinants between NTR1 and other crystallized G protein-coupled receptors suggests that the occupancy of the canonical position of the amphipathic helix is reduced to various extents in many receptors, and we have elucidated the sequence determinants for a stable helix 8. Our analysis also provides a structural rationale for the long-known effects of C-terminal palmitoylation reactions on G protein-coupled receptor signaling, receptor maturation, and desensitization.

membrane proteins | protein stability | protein engineering | detergents

Neurotensin is a 13-amino-acid peptide, which plays important roles in the pathogenesis of Parkinson's disease, schizophrenia, antinociception, and hypothermia and in lung cancer progression (1–4). It is expressed throughout the central nervous system and in the gut, where it binds to at least three different neurotensin receptors (NTRs). NTR1 and NTR2 are class A G protein-coupled receptors (GPCRs) (5, 6), whereas NTR3 belongs to the sortilin family. Most of the effects of neurotensin are mediated through NTR1, where the peptide acts as an agonist, leading to GDP/GTP exchange within heterotrimeric G proteins and subsequently to the activation of phospholipase C and adenylyl cyclase, which produce second messengers in the cytosol (5, 7). Activated NTR1 is rapidly phosphorylated and internalizes by a  $\beta$ -arrestin- and clathrin-mediated process (8), which is crucial for desensitizing the receptor (9). Several lines of evidence suggest that internalization is also linked to G protein-independent NTR1 signaling (10, 11). To improve our mechanistic understanding of NTR1 and to gain additional insight into GPCR features such as helix 8 (H8), we were interested in obtaining a structure of this receptor in a physiologically relevant state.

To date, by far the most successful strategy for GPCR structure determination requires the replacement of the intracellular loop 3 by a fusion protein, as the intracellular domain is otherwise too small to provide crystal contacts. The fusion protein approach has provided a wealth of valuable structural data on GPCRs, but as it renders the crystallized constructs signaling-

inactive, the most important functionality—the activation of G proteins—cannot be confirmed for these structures. This leads inevitably to a degree of uncertainty regarding the physiological relevance of intracellular structural aspects, and it also impedes the elucidation of signaling mechanisms, as functional assays and structure determination cannot be performed with the same GPCR constructs.

Crystallization in the absence of fusion proteins was so far mainly possible for rhodopsin (12), the A<sub>2A</sub> adenosine receptor (A<sub>2A</sub>AR) (13), and the  $\beta_1$ -adrenergic receptor (14). Together, they share a high stability, which is either given naturally (rhodopsin) or it is due to stabilizing mutations. High stability appeared to be crucial for crystallographic success, as it allowed the application of harsh short-chain detergents. These tend to form small micelles, which may explain why crystal contact formation can occur under these conditions despite the small extra- and intracellular domains of class A GPCRs.

Besides the stability requirement and/or the necessity of fusion proteins, structural studies of GPCRs have also been complicated by the need of eukaryotic expression systems [e.g., *Spodoptera frugiperda* (*Sf9*) insect cells], as prokaryotes exhibit generally low functional expression levels of wild-type GPCRs. However, prokaryotes such as *Escherichia coli* offer several advantages compared with insect cells, including quick genetic modification strategies, growth to high cell densities, fast doubling times, inexpensive media, absence of glycosylation, and robust handling. Furthermore, *E. coli* is well suited for producing fully

## Significance

Only a tiny fraction (<2%) of the unique structures in the protein database correspond to membrane proteins, and only a few of these are of eukaryotic origin, representing potential drug targets. The difficulties in structure determination of these proteins are due to two specific complications, which are unique for membrane proteins: first, low expression levels and, second, the necessity for detergent micelles, which are often destabilizing as they mimic the hydrophobic membrane environment only poorly. We prove that directed evolution has the potential to overcome these problems by determining several structures of evolved eukaryotic G protein-coupled receptor variants. High functional expression levels and superior receptor stability in harsh detergents allowed us to gain deeper insights into this important receptor family.

Author contributions: P.E., M.H., C.K., D.J.S., and A.P. designed research; P.E., M.H., C.K., A.B., P.H., S.B., K.M.S., D.J.S., and M.S. performed research; P.E., M.H., C.K., A.B., P.H., K.M.S., D.J.S., M.S., and A.P. analyzed data; and P.E., M.H., C.K., and A.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) [PDB ID codes 4BUO (TM86V- $\Delta$ IC3B), 3ZEV (TM86V- $\Delta$ IC3A), 4BV0 (OGG7- $\Delta$ IC3A), and 4BWB (HTGH4- $\Delta$ IC3)].

<sup>1</sup>To whom correspondence should be addressed. E-mail: [plueckthun@bioc.uzh.ch](mailto:plueckthun@bioc.uzh.ch).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317903111/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317903111/-DCSupplemental).

isotope-labeled proteins—a crucial requirement for many NMR studies, which are limited to date.

To exploit these advantages, we recently developed a directed evolution method for high functional GPCR expression levels in *E. coli* (15). In contrast to screening a few hundred mutants one by one, this strategy allows the simultaneous, competitive testing of  $>10^8$  different protein variants for highest prokaryotic expression and functionality. Briefly, diverse libraries of NTR1 variants were either obtained synthetically (16, 17) or by error-prone PCR on the wild-type sequence (15). The libraries were ligated to a plasmid encoding an inducible promoter, which was subsequently used to transform *E. coli*. Selection pressure for high functional expression levels was applied by incubating the induced cells with fluorescently labeled neurotensin, which allowed enrichment of the best expressing cells by fluorescence-activated cell sorting (FACS). The outlined procedure was performed in cycles, leading to a gradual adaptation of the NTR1 population toward high functional expression levels, and additionally, it gave rise to an increase in thermostability for certain variants.

In a second technology, termed CHES (cellular high-throughput encapsulation, solubilization and screening), we adapted this concept to directly evolve NTR1 variants for high thermostability in short-chain detergent micelles—a property that is not only beneficial for structural studies but also for in vitro drug screening (18). The crucial development of CHES was to surround, simultaneously, every *E. coli* cell by a semipermeable polysaccharide capsule. This allows us to solubilize the receptor mutants with harsh short-chain detergents, each mutant inside its own encapsulated cell, all at once and in the same test tube. Both the solubilized receptors and their encoding plasmids are maintained within the same capsules. Long-term incubation under these conditions followed by labeling of the encapsulated solubilized receptors with fluorescent neurotensin and rounds of FACS enrichment ensured a strong selection pressure and a gradual

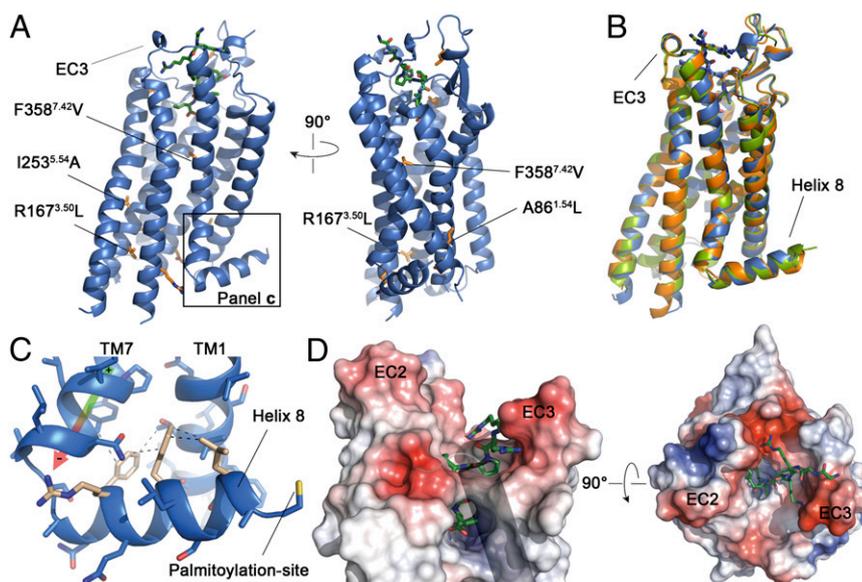
adaptation of the NTR1 population toward high stability in harsh short-chain detergents (18).

In this work, we present the crystal structures of three evolved NTR1 variants, which were either obtained by evolving high functional expression levels in *E. coli* or by directed evolution for stability in detergent micelles. In contrast to the majority of crystallized GPCRs, our NTR1 variants are devoid of bulky modifications at the cytoplasmic face and can thus remain signaling-active, which allows us to gain unique insights into the structure–function relationship of NTR1.

## Results

**Directed Evolution for High Expression Levels Enabled Structure Determination of NTR1-TM86V.** Directed evolution for high functional expression in *E. coli* yielded a population of well-expressed NTR1 mutants, which provided a basis for the identification of suitable variants for structural studies. We have chosen to use the variant NTR1-TM86V for crystallization, as it was the most thermostable mutant that was capable of catalyzing GDP/GTP exchange at the heterotrimeric G protein  $\alpha_1\beta_1\gamma_1$  ( $G_i$ ) in an agonist-dependent way (see Fig. 2 C and D) (17). NTR1-TM86V harbors 11 point mutations (A86L, H103D, H105Y, A161V, R167L, R213L, V234L, I253A, H305R, F358V, and S362A) that confer the high expression levels in *E. coli* and its stability in detergent solution (Table S1). We observed that the long and putatively flexible intracellular loop 3 and the receptor termini are not required for  $G_i$  signaling and hence shortened them to aid crystallization (TM86V- $\Delta$ IC3A). The protein could be purified to homogeneity (Fig. S1) in the short-chain detergent nonyl- $\beta$ -D-glucopyranoside, and it was crystallized by standard vapor diffusion techniques.

The crystal structure of TM86V- $\Delta$ IC3A at 3.26 Å [ $1/\sigma(I) = 2.0$ ] revealed a canonical GPCR fold (Fig. 1A and Table S2) with seven transmembrane helices (TMs) and the prototypical amphipathic



**Fig. 1.** Structures of three evolved NTR1 variants determined devoid of fusion proteins. (A) The signaling-competent NTR1-TM86V- $\Delta$ IC3A (blue) bound to its natural agonist neurotensin (green). All selected mutations for increased expression levels in *E. coli* and high stability in detergent solution are depicted (orange). (B) Superposition of NTR1-TM86V- $\Delta$ IC3A (blue), NTR1-OGG7- $\Delta$ IC3A (green), and NTR1-HTGH4- $\Delta$ IC3A (orange). (C) Close-up view of the H8 region in NTR1-TM86V- $\Delta$ IC3A. Certain hydrophobic contacts of amino acids of the semiconserved H8 motif (beige) are depicted by dashed lines for clarity. The helix dipole of TM7 is illustrated by an arrow. The first of the two palmitoylation sites adjacent to the H8 C terminus is indicated. Note the absence of the palmitoyl moiety due to the prokaryotic expression. (D) Vacuum-electrostatic surface representation (Pymol) of the neurotensin-binding pocket of TM86V- $\Delta$ IC3A. Parallel (Left) and perpendicular (Right) view to the membrane. TM5 is represented as a transparent tube in the Left panel for clarity. Neurotensin is a 13-amino-acid peptide in vivo, but only the C-terminal residues 8–13 were reported to be relevant for binding to NTR1. Strong electron density for these six amino acids was found and allowed us to model the ligand unambiguously (Fig. S2). In addition, relatively weak electron density for two N-terminal linker amino acids (Gly–Gly) of the peptide was observed in one complex of the asymmetric unit (modeled here).

H8 (Fig. 1C). We observed strong electron density for the agonist neurotensin, confirming that the GPCR produced in *E. coli* reaches a functional conformation despite the absence of the eukaryotic translation and membrane insertion machinery (Fig. 1D, Table S3, and Fig. S2). The resolution was subsequently further improved by a change of the intracellular loop 3 deletion (TM86V- $\Delta$ IC3B), which resulted in an additional crystal contact in the same space group. TM86V- $\Delta$ IC3B was overall identical to TM86V- $\Delta$ IC3A (RMSD<sub>C $\alpha$</sub>  = 0.3 Å) and could be refined to a resolution of 2.75 Å (Table S2).

**TM86V- $\Delta$ IC3A Exhibits the Functional Characteristics of a Typical GPCR.** To verify the physiological relevance of the initial structure of TM86V- $\Delta$ IC3A, we characterized the crystallized construct regarding ligand affinities, G protein activation, and neurotensin-dependent internalization. Ligand-binding assays on whole *E. coli* cells revealed that TM86V- $\Delta$ IC3A exhibits an apparent dissociation constant of  $2.3 \pm 0.4$  nM for the agonist neurotensin (cf. wild-type NTR1,  $2.8 \pm 0.3$  nM). In contrast to the agonist, the antagonist SR142948 had never been used as a ligand during directed evolution, but we still observed only a moderate increase in IC<sub>50</sub> for TM86V- $\Delta$ IC3A ( $30 \pm 2.4$  nM) compared with the wild-type receptor ( $8.4 \pm 0.9$  nM) (Fig. 2A and B and Fig. S3), which may be attributable to the point mutation F358V in NTR1-TM86V—a residue that was shown to be specifically involved in antagonist (but not agonist) binding (19).

To confirm interactions with G proteins, we measured GDP/GTP exchange in membranes containing TM86V- $\Delta$ IC3A and the reconstituted heterotrimeric G protein  $\alpha_1\beta_1\gamma_1$  (G<sub>i</sub>) (Fig. 2C and Fig. S4). The crystallized GPCR construct exhibited a slightly increased basal GDP/GTP exchange catalysis at G<sub>i</sub> compared with wild-type NTR1, which was further stimulated by the addition of agonist. Even though the maximal signaling level is reduced compared with wild-type NTR1, it appears that the crystallized construct is indeed able to bind to and activate G<sub>i</sub>. To confirm these observations, we also demonstrated specific G<sub>i</sub> binding of detergent-solubilized TM86V- $\Delta$ IC3A in a pull-down experiment using immobilized G protein on magnetic beads (Fig. 2D). Basal and agonist-dependent signaling of the crystallized construct TM86V- $\Delta$ IC3A was further increased by reverting the mutation R167<sup>3.50</sup>L [superscript according to Ballesteros-Weinstein (20)] in the highly conserved D/ERY motif (Fig. 2C and Fig. S5). Even though the reintroduction of R167<sup>3.50</sup> resulted in significantly reduced expression levels in *Sf9* insect

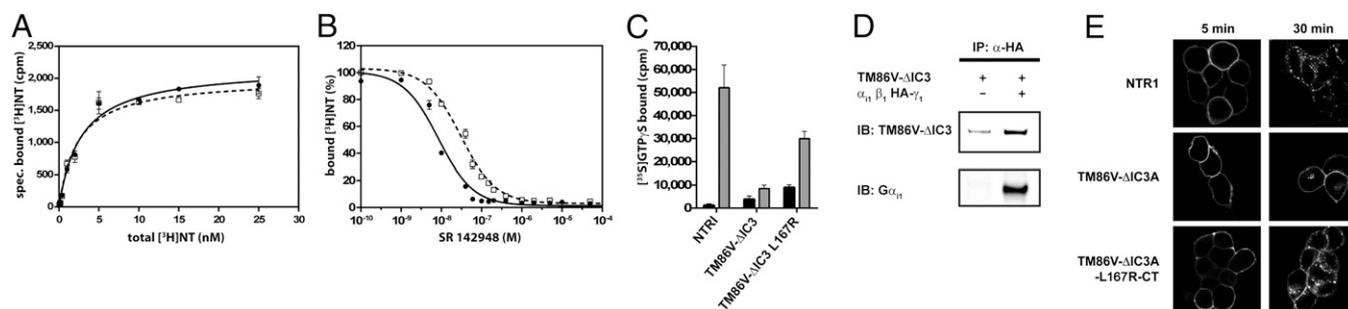
cells (Fig. S4A), the thermostability remained almost unperturbed (Fig. S4D).

We also investigated  $\beta$ -arrestin2-dependent desensitization behaviors by confocal microscopy on living HEK293T cells, which coexpressed TM86V- $\Delta$ IC3A and  $\beta$ -arrestin2-YFP. Despite the lacking C terminus in the crystallized construct, we observed a weak internalization when bound to fluorescent neurotensin (Fig. 2E and Fig. S6). Furthermore, after reconstituting R167<sup>3.50</sup> and the receptor C terminus, a pronounced cointernalization of  $\beta$ -arrestin and fluorescent neurotensin was observed, suggesting that this mutant can indeed interact with  $\beta$ -arrestin2 in a fashion similar to wild type (Fig. S6).

In summary, our pharmacological data clearly suggest that the crystallized NTR1 construct TM86V- $\Delta$ IC3A exhibits all essential core functions of a GPCR. Considering the simplicity of expression and genetic modification strategies in *E. coli* and the high stability of TM86V- $\Delta$ IC3A, the protein will likely serve as a valuable model system for future structural and functional studies.

**Two Structures of Stability-Evolved NTR1 Variants.** NTR1-TM86V was obtained by evolving high functional expression in *E. coli* and subsequently by choosing and recombining the most thermostable mutations (16, 17). In contrast, the CHESSE technology can directly generate detergent-stable NTR1 variants by an evolutionary process (18). As a proof of this principle, we were interested in confirming the structural integrity of these variants as well. NTR1-OGG7 and NTR1-HTGH4 were generated by CHESSE and represent the most thermostable mutants obtained so far. They crystallized readily under various conditions, and the structures were refined to 3.1 Å (OGG7- $\Delta$ IC3A) and 3.57 Å (HTGH4- $\Delta$ IC3A), respectively. Despite significant sequence variations, OGG7- $\Delta$ IC3A and HTGH4- $\Delta$ IC3A are structurally nearly identical to TM86V- $\Delta$ IC3A (TM86V- $\Delta$ IC3A/OGG7- $\Delta$ IC3A RMSD<sub>C $\alpha$</sub>  = 0.4 Å; TM86V- $\Delta$ IC3A/HTGH4- $\Delta$ IC3A RMSD<sub>C $\alpha$</sub>  = 0.4 Å) (Fig. 1B and Table S1). This suggests that the ligand-guided selection pressure has favored or preserved the same conformational state in these evolved variants, independent of the particular kind of directed evolution (for functional expression or high stability in detergents).

Because TM86V- $\Delta$ IC3A is signaling-active (Fig. 2C and D) and exhibiting wild-type-like ligand affinities (Fig. 2A and B) and also desensitization characteristics (Fig. 2E and Fig. S6), it is likely that all our structures represent a naturally occurring conformation of NTR1. Taken together, our four structures of



**Fig. 2.** Pharmacological characterizations of the crystallized NTR1 construct TM86V- $\Delta$ IC3A. (A) Neurotensin saturation-binding assay of wild-type NTR1 (circles) and TM86V- $\Delta$ IC3A (open squares). Note that  $B_{\max}$  levels are not representative for the expression levels of the different mutants, as 10-fold more cells were used for wild-type NTR1 to obtain a similar signal-to-noise ratio—that is, the normalized  $B_{\max}$  would be about 10-fold lower. (B) SR142948 antagonist competition binding experiment using wild-type NTR1 and TM86V- $\Delta$ IC3A. (C) GDP/[<sup>35</sup>S]GTP- $\gamma$ S signaling assays of wild-type NTR1, TM86V- $\Delta$ IC3A, and TM86V- $\Delta$ IC3A L167<sup>3.50</sup>R in insect cell membranes. Equivalent amounts of active GPCR and reconstituted G<sub>i</sub> were assayed in the presence (gray) or absence (black) of neurotensin. The signals correspond to the average of two signaling assays performed in parallel from two independent GPCR expressions, and the error bars represent SDs. (D) Pull-down experiment using immobilized G<sub>i</sub> and solubilized GPCR from *E. coli* membranes. (E) Confocal imaging of living HEK293T cells expressing NTR1, TM86V- $\Delta$ IC3A, or TM86V- $\Delta$ IC3A L167<sup>3.50</sup>R-CT (reconstituted D/ERY motif and C terminus) after stimulation with fluorescent neurotensin8-13-HL647 for the indicated times.

the three different NTR1 variants exemplify that two directed evolution methods in *E. coli*, which we have recently developed, are valuable tools for structural studies of GPCRs. Our technologies have been applied successfully to a number of other receptors (21), underlining the potential of Darwinian evolution in protein research.

**Improved Interhelical Surface Complementarity May Contribute to Increased Thermostability in NTR1-TM86V.** We were interested in identifying the molecular causes of different thermostability characteristics among NTR1 mutants. When comparing the thermostabilities of NTR1-TM86V with one of its precursors, termed NTR1-D03 (15), we uncovered a pattern that sheds light on this issue. NTR1-D03 harbors all NTR1-TM86V mutations except A86<sup>L54L</sup>, I253<sup>S54A</sup>, and F358<sup>T42V</sup>. Despite only three amino acid differences, NTR1-D03 exhibited a very low thermostability in the short-chain detergent octyl- $\beta$ -D-glucopyranoside, whereas NTR1-TM86V exhibited a high thermal denaturation point of 38 °C under these particularly harsh detergent conditions (Fig. 3A). Interestingly, the mutations cause only the replacement of hydrophobic amino acids with other hydrophobic residues. It is striking that the bulky wild-type amino acids at positions 253 and 358, where directed evolution favored a shortening of the hydrophobic side chains, would lead to obvious clashes for all common rotamers in silico (Fig. 3C and D). At position 86, where the longer leucine was preferred over the shorter alanine, the in silico back-mutation would cause a loss of favorable van der

Waals contacts between TM1 and TM2 (Fig. 3B). These observations suggest that improved interhelical surface complementarity contributes significantly to the high thermostability of NTR1-TM86V, and conversely, it may be speculated that optimal helix packing is not required for this particular state of wild-type NTR1 in nature.

#### NTR1 Can Adopt a Prototypical Inactive State at the Cytosolic Domain.

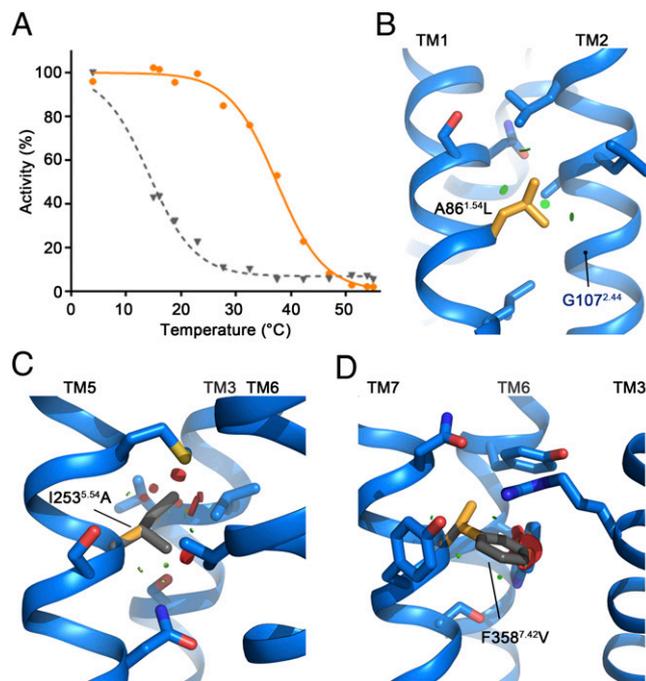
In activated GPCR states, the cytosolic ends of TM5 and TM6 were described to be tilted outward relative to their inactive state (22). This is observed in the most prominent way in the structure of the  $\beta_2$ -adrenergic receptor bound to  $G\alpha_s\beta_1\gamma_2$  (23). Even though TM86V- $\Delta$ IC3A is bound to its natural agonist and capable of triggering GDP/GTP exchange at  $G_i$ , the conformations of TM5 and TM6 that were trapped in the crystal are highly similar to dark-state rhodopsin, which represents an inactive or “closed” state (Fig. 4, Fig. S5, and Fig. S7). Our finding is in agreement with other agonist-bound GPCR structures that were crystallized in inactive states, and it provides further evidence that fully active states require the G protein for stabilization. Several structural studies on rhodopsin and on the  $\beta_2$ -adrenergic receptor suggest that the observed closed conformation would occlude the G protein-binding site (22–25). Nevertheless, TM86V- $\Delta$ IC3A is able to functionally couple to G proteins to a certain degree (Fig. 2C), suggesting that the crystallized construct exhibits structural flexibility and allows a conformational change when bound to agonist. The evolved NTR1 thus shows characteristics consistent with a conformational equilibrium typical for GPCRs: In the absence of a G protein, energetically the most favorable arrangement at the intracellular side of TM86V- $\Delta$ IC3A is likely the inactive conformation that was trapped in the crystal. This may also be the case for wild-type NTR1 in the apo-state, as it exhibits very low basal signaling activity toward  $G_i$  (Fig. 2C).

#### Structural Comparison of the Evolved NTR1 Variants to NTR1-GW5.

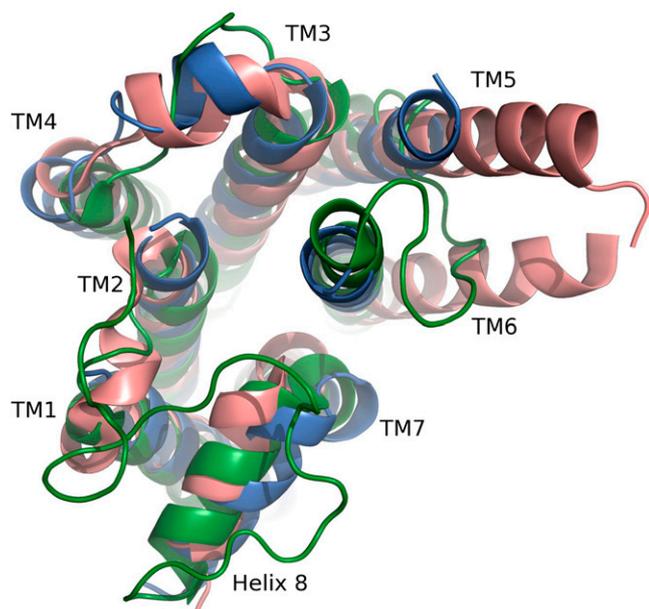
The observation of a prototypical inactive NTR1 state represents one of the unique features that distinguishes the structures presented in this work from the structure of the NTR1 variant GW5 [Protein Data Bank (PDB) ID code 4GRV] (Fig. 5A) (26). The mutations present in NTR1-GW5 were identified by alanine-scanning mutagenesis, and the protein required expression in *Sf9* insect cells, fusion to T4 lysozyme replacing intracellular loop 3 (GW5-T4L), and crystallization in the presence of ligand in lipidic cubic phase (26). The crystallized construct GW5-T4L exhibits a 200-fold increased  $IC_{50}$  value for the antagonist SR48692, and it is signaling-inactive in the presence and absence of the fusion protein. The authors suggested nevertheless that the structure represents an active-like conformation, based on a partial outward tilt of the intracellular end of TM6 and on the observation of a hydrogen bond between R167<sup>S50</sup> and N257<sup>S58</sup>. Similar features had previously been found in other GPCR structures, which represent most likely active states (27).

White et al. (26) suggested that an unusual elongation of TM7 may cause the observed lack of signaling. Indeed, a comparison with our structures reveals that TM7 is extended in GW5-T4L by a peptide segment that corresponds to the amphipathic H8 (Fig. 5A and B). A canonical H8 would clash into a neighboring molecule in the lipidic cubic phase crystal lattice of GW5-T4L. Instead, the H8 segment resides at the center of the cytosolic domain, where it blocks the prototypical inactive position of TM6 and thus also the putative G protein-binding pocket—an arrangement that has never been observed for other GPCRs.

In contrast, all our structures suggest a canonical H8 (Fig. 1B), and one of the two TM86V- $\Delta$ IC3A molecules in the asymmetric unit exhibits no crystal contacts at H8. Furthermore, as described above, TM6 is positioned in a prototypical inactive conformation when bound to agonist. The observed outward tilt of TM6 in



**Fig. 3.** Improved interhelical surface complementarity correlates with increased thermostability. (A) Thermostability assays of NTR1-D03 (gray) and NTR1-TM86V (orange) bound to neurotensin in the harsh detergent octyl- $\beta$ -D-glucopyranoside. Note that the low stability of NTR1-D03 in this detergent did not permit an accurate determination of its thermal denaturation transition point. NTR1-D03 and NTR1-TM86V are identical except for three additional mutations in NTR1-TM86V, which must confer this thermostability difference. (B–D) The structure of TM86V- $\Delta$ IC3B illustrates the 3-dimensional context at these positions. In silico back-mutating the selected residues (orange) to the wild-type amino acids (gray) would either cause a reduction of favorable van der Waals contacts (green circles in B), or it would lead to steric clashes (red circles in C and D). For the wild-type residues in C and D, the most common rotamers based on the library of PYMOL are shown. (See Fig. S8 for additional rotamers.)



**Fig. 4.** View from the cytosol onto the superposition of TM86V- $\Delta$ IC3A (blue), dark-state bovine rhodopsin (green, PDB ID code 1U19), and  $\beta_2$ -adrenergic receptor bound to  $G\alpha_s\beta_1\gamma_2$  (salmon, PDB ID code 3SN6;  $G\alpha_s\beta_1\gamma_2$  is omitted).

GW5-T4L could thus alternatively be explained by the unusual contacts of TM7 to TM6.

Although the cytosolic regions of the evolved NTR1 variants described here are very different from GW5-T4L, at the extracellular side, only one major discrepancy can be observed (Fig. 5 C and D). The  $2F_o - F_c$  omit map of TM86V- $\Delta$ IC3B suggests a single  $\alpha$ -helical turn of ECL3 with several ligand contacts (including a salt bridge between D336 and R9 of neurotensin). The same arrangement was found in all structures of the evolved mutants, and it is in agreement with published mutagenesis data (28, 29). In GW5-T4L, on the other hand, the loop contains no secondary structural element and it was modeled significantly more distant to neurotensin with an unusual *cis*-peptide bond following D336 (Fig. 5D).

**NTR1-Specific Determinants of Reduced H8 Stability.** Practically all high-resolution GPCR structures exhibited an amphipathic H8 following TM7. Its presence thus appeared to be a general feature of GPCRs, but surprisingly, the recently determined structures of proteinase-activated receptor 1 (PAR1), chemokine receptor 4 (CXCR4), and NTR1 (GW5-T4L) do not exhibit H8. How relevant are these findings physiologically?

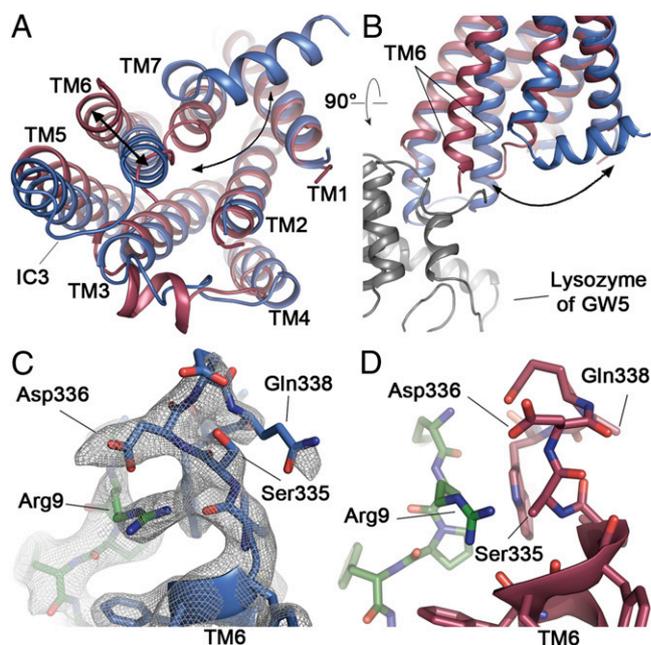
Class A GPCRs exhibit relatively small intracellular domains, and it is thus apparent that the presence or absence of the complete H8 is a major factor determining the characteristics of the cytosolic interface. Multiple lines of evidence suggest an important functional role for this protein segment including G protein coupling and  $\beta$ -arrestin activation (30–36). Our finding of a canonical H8 in the evolved NTR1 variants now shows that a GPCR, which was previously crystallized without H8 being formed, can exhibit a canonical H8 structure (Fig. 1C).

To understand structural key features that are critical for the presence (or absence) of the canonical H8, we compared our structures to A2AR (PDB ID code 4E1Y) (Fig. 6 A and B): A2AR likely exemplifies one of the most stable H8 arrangements, as the amphipathic helix shows large contacts to TM1, TM2, IC1, and TM7 and because all reported A2AR structures exhibited a canonical H8 irrespective of the presence or absence

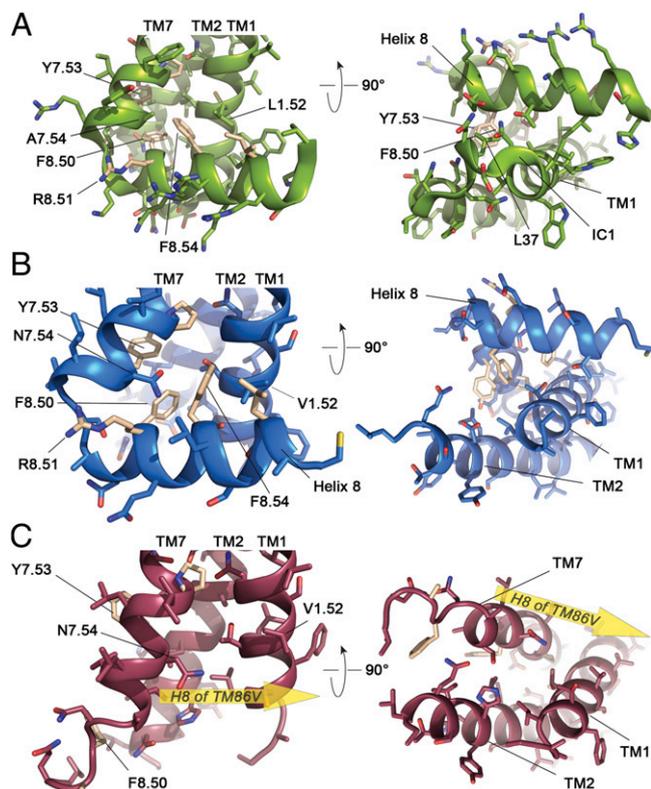
of crystal contacts in this region and despite a variety of crystallization conditions.

Both A2AR and NTR1 encode the semiconserved H8 motif F(R/K)xx(F/L)xxx(L/F) (Fig. 7). A common feature of our structures and of A2AR is the location of the positively charged guanidinium group of R<sup>8.51</sup> (Ballesteros–Weinstein numbering, 8.50 = F376 in NTR1) at the negative helix-dipole at the C terminus of TM7 (Fig. 6 A and B). This interaction likely contributes to the stabilization of the helix break between TM7 and H8, which is not encoded per se, as helix-destabilizing residues (prolines or glycines) are absent in the peptide segment connecting the two helices. Nonetheless, R(K)<sup>8.51</sup> is conserved among class A GPCRs (81%) and similar interactions can be found in the majority of published GPCR structures. Another similarity between the crystal structures of A2AR and NTR1 is the absence of palmitoylation membrane anchors, in the case of A2AR because of the absence of cysteines adjacent to the H8 C terminus and in the case of NTR1-TM86V because of expression in a prokaryotic system.

Clearly distinct interactions in NTR1 and A2AR are observed for the most conserved residue F<sup>8.50</sup> of the H8 motif. In A2AR, F<sup>8.50</sup> is entirely surrounded by a hydrophobic pocket (Fig. 6A). Most prominently, the conserved Y<sup>7.53</sup> of the NPxxY motif at H7 exhibits  $\pi$ - $\pi$  stacking interactions to F<sup>8.50</sup>, typical for a receptor in the inactive state (13). On the opposite side of its aromatic side chain, F<sup>8.50</sup> engages in van der Waals contacts to the poorly conserved L37 in IC1 (sequential numbering of A2AR used for this residue). The aromatic ring of F<sup>8.50</sup> is thus sandwiched between Y<sup>7.53</sup> and L37. Furthermore, several hydrophobic residues



**Fig. 5.** Comparison of neurotensin-bound TM86V- $\Delta$ IC3B and GW5-T4L. (A) Superposition of TM86V- $\Delta$ IC3B (blue) and GW5-T4L (red), view from the intracellular side. The fused T4 lysozyme of GW5 replacing IC3 is omitted for clarity. Black arrows highlight the two different C-terminal conformations and the alternative states of TM6. (B) View along the inner leaflet of the membrane, including a part of the fused T4 lysozyme of GW5. (C and D) Comparison of the ligand-binding pockets, focusing on the interactions of EC3 with neurotensin (green). (C) The  $2F_o - F_c$  omit map of TM86V- $\Delta$ IC3B (contoured at a  $\sigma$  level of 1.2) suggests a single  $\alpha$ -helical turn of ECL3 in close proximity to the ligand. (D) In GW5-T4L (26) the loop contains no secondary structural element and it was modeled more distant to the peptide agonist with a *cis*-peptide bond following Asp336. Side chains of Ser335 and Gln338 were modeled up to C $\beta$  only.



**Fig. 6.** Key interactions of the H8 region in A2AR and NTR1. (A–C) Depicted are the cytosolic ends of TM1, TM2, TM7, and H8 of A2AR (A; PDB ID code 4E1Y), TM86V-ΔIC3B (B), and GW5-T4L (C; PDB ID code 4GRV) viewed parallel to the membrane (Left) or from the intracellular side (Right). The yellow arrow in the GW5-T4L structure corresponds to the approximate position of H8 in TM86V-ΔIC3B.

of TM1, TM2, TM7, and H8 contact the CH groups of the aromatic system of F<sup>8.50</sup> (not depicted in Fig. 7), and in addition, F<sup>8.54</sup> of the H8 motif covers the hydrophobic pocket of F<sup>8.50</sup>. F<sup>8.54</sup> is accommodated between the poorly conserved A<sup>7.54</sup> and L<sup>1.52</sup>.

In our NTR1 structures, F<sup>8.50</sup> mediates only weak interhelical interactions and no  $\pi$ - $\pi$  stacking to Y<sup>7.53</sup>, as it is only partially inserted into the pocket between TM1, TM2, and TM7 (Fig. 6B). This is due to the following three reasons: First, F<sup>8.54</sup> fails to cover the pocket, because it cannot obtain an analogous rotamer conformation to A2AR; it would clash into the longer side chain of N<sup>7.54</sup> (A<sup>7.54</sup> in A2AR), and V<sup>1.52</sup> would not be long enough to stabilize the A2AR-like rotamer (L<sup>1.52</sup> in A2AR). Second, the pocket itself is considerably different, as the weakly conserved residues at TM1, TM2, and TM7 contacting the CH groups of

the aromatic system of F<sup>8.50</sup> in A2AR are less hydrophobic in NTR1 (not depicted in Fig. 6B). And third, because IC1 of NTR1 is longer and presumably flexible (disordered in all our structures and in GW5-T4L), it does not provide a hydrophobic residue like L37 in A2AR to sandwich F<sup>8.50</sup> from the intracellular side (Fig. 6B and C). Additionally, the absence of a structured IC1 in NTR1 causes also another lack of interactions to H8, as the loop mediates not only the L37 to F<sup>8.50</sup> contacts in A2AR but also extensive interactions with non-conserved H8 residues (Fig. 6A).

In summary, the shape complementarity (37) between H8 and the receptor is significantly worse in NTR1 ( $S_c = 0.642$ ) than in A2AR ( $S_c = 0.81$ ) and the buried surface area is strongly reduced (NTR1, 222 Å<sup>2</sup>; A2AR, 303 Å<sup>2</sup>). As described above, these differences are due to alternative amino acids at poorly conserved positions in IC1, TM1, TM2, and TM7, including the residues constituting the pocket of F<sup>8.50</sup>, and they imply that the canonical H8 arrangement in NTR1 is significantly less stable than in A2AR.

## Discussion

### Ligand and Palmitoylation Dependence of the Canonical H8 State.

In this work, we present the agonist-bound structures of the three NTR1 variants TM86V, OGG7, and HTGH4, which were generated by directed evolution for high functional expression and for stability in short-chain detergents. In contrast to most other crystallized GPCR constructs so far, TM86V-ΔIC3A not only exhibited wild-type-like ligand-binding properties, it was also able to signal to G<sub>i</sub> to some extent. Moreover, when expressed in eukaryotic cells, the classical features of receptor desensitization and internalization were detected, suggesting that the structure derived from this construct resembles a physiologically relevant state. The canonical inactive-like positioning of TM6 is distinct from the outward tilted helix in GW5-T4L, and it is only permitted because of the presence of a canonical H8 that does not occlude this position and the putative G protein-binding pocket. We observed elevated B-factors in the H8 region for all our structures (Fig. S9), and in addition, we found comparatively weak contacts to TM1 and TM7 and a lack of interactions to IC1 and TM2 (Fig. 6B). Considering these observations and the absence of H8 in GW5-T4L, it is tempting to speculate that the canonical H8 of NTR1 is of lower stability than that of the prototypical version in A2AR and/or only partially occupied under certain conditions.

PAR1 and CXCR4 may represent even more extreme cases in this regard, as none of their crystal structures exhibited H8 (38, 39). These absences can be explained by the fact that they are not only different at the nonconserved positions, which cause the reduced H8 stability in NTR, but also by the observation that these two receptors lack parts of the rather conserved H8 motif (Fig. 7). This correlation points to a sequence-specific origin, and thus a naturally evolved feature of physiological relevance.

	TM7							H8																		
	7.49	7.50	7.51	7.52	7.53	7.54	7.55	7.56	8.50	8.51	8.52	8.53	8.54	8.55	8.56	8.57	8.58	8.59	8.60	8.61	8.62	8.63	8.64			
A <sub>2</sub> AR	N	P	F	I	Y	A	Y	R	I	R	E	F	R	Q	T	F	R	K	I	I	R	S	H	V	L	R
NTR1	N	P	I	L	Y	N	L	V	S	A	N	F	R	Q	V	F	L	S	T	L	A	C	L	C	P	G
PAR1	D	P	L	I	Y	Y	A	S	S	E	C	Q	R	Y	V	Y	S	I	L	C	C	K	E	S	S	
CXCR4	N	P	I	L	Y	A	F	L	G	A	K	F	K	T	S	A	Q	H	A	L	T	S	V	S	R	G
	N	P	X	X	Y							F	R	X	X	F	X	X	X	L						

**Fig. 7.** Sequence alignment representing the end of TM7 and H8. The sequences are numbered according to Ballesteros–Weinstein (residue 8.50 chosen as F376 of NTR1). The NPxxY and F(R/K)xx(F/L)xxx(L/F) motifs are highlighted (green) and putative palmitoylation sites [experimentally confirmed in NTR1 (44, 49)] are depicted (yellow).

H8 dynamics have previously been investigated for a number of GPCRs (34, 40, 41). Among all published GPCR structures that include the amphipathic helix, it can be observed that alterations of poorly conserved residues cause a variety of subtly deviating H8 arrangements. The resulting stability differences of H8 likely reflect an evolutionary adaptation of every receptor to the particular requirement on its amphipathic helix. Importantly, the most conserved interaction (the stacking of F<sup>8.50</sup> and Y<sup>7.53</sup>) was described to be disrupted upon agonist binding, as Y<sup>7.53</sup> flips toward the G protein-binding cavity upon receptor activation (42). The current body of high-resolution structural data therefore suggests that a reduction of forces that keeps H8 in its canonical arrangement is a common theme of GPCR activation.

Besides intramolecular interactions, another important parameter influences the stability of the amphipathic helix: the number and positions of palmitoyl anchors at its C terminus. Palmitoylation is known to be a reversible and dynamic protein modification that can be cell-cycle-dependent (43) and developmentally regulated (44). In GPCRs, the palmitoylation state was reported to affect G protein signaling, receptor maturation, membrane delivery, phosphorylation efficiency, and desensitization (44, 45). Our finding of an unstable canonical H8 in nonpalmitoylated NTR1 implies that the occupancy of the canonical H8 state of this receptor—and potentially also of other GPCRs—may depend crucially on the palmitoylation state. Considering that the presence or absence of H8 certainly represents an important source of binding specificity to cytosolic interaction partners, it is conceivable that palmitoylation/depalmitoylation events exert their physiological effects in many cases via modulating the stability and dynamics of H8.

**Potential of Directed Evolution for Membrane Protein Research.** Most membrane proteins are unsuitable for high-resolution structure determination, because of difficulties in overexpression, instability in detergent solution, or both. To date, the most successful approaches to circumvent these problems rely on trial-and-error procedures, like homology screens or alanine scans, which involve expression, stability, and purification tests of individual proteins in high-throughput formats. Miniaturization has indeed advanced membrane protein structural biology significantly in recent years, but given the resources it takes and the still striking underrepresentation of structural data in the PDB, it is apparent that alternative approaches are needed.

Loss of functionality and low sequence identity to the protein of interest (e.g., by using a bacterial homolog) are frequently accepted as necessary evils on the way to the structure. We have shown in this work that a fundamentally different approach was successfully applied to generate several crystallizable GPCR variants with high sequence identity to the protein of interest (93.2–97.5%) (15–18). Instead of screening mutants or homologs one by one, our method exploited the power of evolution on populations of more than a hundred million GPCR variants at once. Analogous to natural evolution, directed evolution amplified favorable GPCR traits through the alternation of random mutagenesis and selection pressure, allowing a gradual adaptation of the characteristics of the whole GPCR population toward the selected phenotype—it

tailored an array of GPCR variants with suitable properties for structural biology independent of previous structural knowledge.

Importantly, the evolutionary system allowed us to determine structures of GPCRs produced in *E. coli*, thus establishing a prokaryote as a novel and robust host for quantitative, functional, and very rapid GPCR overexpression (15, 21). As *E. coli* is well suited for producing isotope-labeled proteins, we also provide the basis for an array of NMR studies that were not feasible for this class of membrane proteins so far. Furthermore, the high stability of functional GPCRs generated by directed evolution will facilitate high-throughput ligand screening in vitro, and thus likely contribute to the discovery of new drugs.

## Materials and Methods

**Construct Design and Expression for Crystallization.** All NTR1 variants were expressed in *E. coli* using an isopropyl- $\beta$ -D-thiogalactopyranoside-inducible pBR322-derived vector, which was derived from a plasmid originally obtained as a kind gift from R. Grishammer (National Institute of Neurological Disorders and Stroke, National Institutes of Health, Rockville, MD) (46–48). They were N- and C-terminally truncated at G50 and G390, respectively, and linked via human rhinovirus 3C protease sites to maltose-binding protein (N terminal) and thioredoxin (C terminal). Amino acids V280-I295 were deleted in the constructs  $\Delta$ IC3A and E273-T290 in  $\Delta$ IC3B. Directed evolution of NTR1 was performed as previously described (17, 18). Full details are given in *SI Text*, and a list of all evolved mutations is given in *Table S1*.

**Purification and Crystallization.** Whole *E. coli* cells were solubilized in 50 mM Hepes pH 8, 10% (vol/vol) glycerol, 200 mM NaCl, protease inhibitor tablets (Roche), and 0.6% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS), 0.12% cholesteryl hemisuccinate tris salt (CHS), and 1.6% (wt/vol) decyl- $\beta$ -D-maltopyranoside. All NTR1 variants were purified based on ligand affinity, cation exchange, and size exclusion in nonyl- $\beta$ -D-glucopyranoside, and they were crystallized in standard vapor diffusion experiments using various mixtures of glucoside detergents and cholesterol hemisuccinate as additives. (See *SI Text* for details.) The reservoir solutions of the different NTR1 crystallization conditions varied significantly regarding buffering compound (acetate pH 5.5 or glycine pH 9.4), salt (500 mM or 2 M NaCl or 0.2 M CaCl<sub>2</sub>), and PEG 600 concentrations [20% (vol/vol)–26% (vol/vol)]. (See *SI Text* for details.) Diffraction data were collected from one single crystal per protein at the Swiss Light Source, and the structure was determined by molecular replacement. (See *SI Text* for details.)

**Functional Assays.** Ligand affinity measurements were performed on whole *E. coli* cells using either <sup>3</sup>H-neurotensin or <sup>3</sup>H-neurotensin and the unlabeled NTR1 antagonist SR142948 for competition experiments. (See *SI Text* for details.) Signaling assays were performed with purified G<sub>i</sub> protein (expression in Sf9 insect cells) composed of G $\alpha_{i1}$ , G $\beta_1$ , and G $\gamma_1$  and a defined amount of active NTR1, TM86V- $\Delta$ IC3A, or TM86V- $\Delta$ IC3A-L167<sup>3.50R</sup> on urea-washed membranes. (See *SI Text* for details.) Pull-down experiments were performed with purified G<sub>i</sub> and solubilized *E. coli* membranes containing the expressed TM86V- $\Delta$ IC3A. (See *SI Text* for details.) Fluorescence microscopy was performed on living HEK293T cells that were transiently transfected with NTR1 variants and  $\beta$ -arrestin2-YFP. (See *SI Text* for details.)

**ACKNOWLEDGMENTS.** We thank Prof. Raimund Dutzler for comments on crystallization procedures, Peer Mittl for support during structure determination, Mattia De Luigi and Christian Schori for insights regarding GPCR and G protein purification, and Beat Blattmann and Céline Stutz [National Center of Competence in Research (NCCR) crystallization facility] for their efforts during initial crystallization screening. This work was funded by the NCCR Structural Biology (Schweizerischer Nationalfonds).

- Bissette G, Nemeroff CB, Loosen PT, Prange AJ, Jr., Lipton MA (1976) Hypothermia and intolerance to cold induced by intracisternal administration of the hypothalamic peptide neurotensin. *Nature* 262(5569):607–609.
- Schimppf RM, et al. (2001) Increased plasma neurotensin concentrations in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 70(6):784–786.
- Alifano M, et al. (2010) Neurotensin receptor 1 determines the outcome of non-small cell lung cancer. *Clin Cancer Res* 16(17):4401–4410.
- Griebel G, Holsboer F (2012) Neuropeptide receptor ligands as drugs for psychiatric diseases: The end of the beginning? *Nat Rev Drug Discov* 11(6):462–478.
- Tanaka K, Masu M, Nakanishi S (1990) Structure and functional expression of the cloned rat neurotensin receptor. *Neuron* 4(6):847–854.
- Chalon P, et al. (1996) Molecular cloning of a levocabastine-sensitive neurotensin binding site. *FEBS Lett* 386(2-3):91–94.
- Skrzydelski D, et al. (2003) Differential involvement of intracellular domains of the rat NTS1 neurotensin receptor in coupling to G proteins: A molecular basis for agonist-directed trafficking of receptor stimulus. *Mol Pharmacol* 64(2):421–429.
- Savdie C, Ferguson SS, Vincent J, Beaudet A, Stroh T (2006) Cell-type-specific pathways of neurotensin endocytosis. *Cell Tissue Res* 324(1):69–85.
- Ferguson SS (2001) Evolving concepts in G protein-coupled receptor endocytosis: The role in receptor desensitization and signaling. *Pharmacol Rev* 53(1):1–24.
- Souazé F, Rostène W, Forgez P (1997) Neurotensin agonist induces differential regulation of neurotensin receptor mRNA. Identification of distinct transcriptional and post-transcriptional mechanisms. *J Biol Chem* 272(15):10087–10094.
- Toy-Miou-Leong M, Cortes CL, Beaudet A, Rostène W, Forgez P (2004) Receptor trafficking via the perinuclear recycling compartment accompanied by cell division is

- necessary for permanent neurotensin cell sensitization and leads to chronic mitogen-activated protein kinase activation. *J Biol Chem* 279(13):12636–12646.
12. Okada T, et al. (2000) X-Ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles. *J Struct Biol* 130(1):73–80.
  13. Lebon G, et al. (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* 474(7352):521–525.
  14. Warne T, et al. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454(7203):486–491.
  15. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* 105(39):14808–14813.
  16. Schlinkmann KM, et al. (2012) Critical features for biosynthesis, stability, and functionality of a G protein-coupled receptor uncovered by all-versus-all mutations. *Proc Natl Acad Sci USA* 109(25):9810–9815.
  17. Schlinkmann KM, et al. (2012) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3):414–428.
  18. Scott DJ, Plückthun A (2013) Direct molecular evolution of detergent-stable G protein-coupled receptors using polymer encapsulated cells. *J Mol Biol* 425(3):662–677.
  19. Kitabgi P (2006) Functional domains of the subtype 1 neurotensin receptor (NTS1). *Peptides* 27(10):2461–2468.
  20. Ballesteros JA, Weinstein H (1992) Analysis and refinement of criteria for predicting the structure and relative orientations of transmembrane helical domains. *Biophys J* 62(1):107–109.
  21. Dodevski I, Plückthun A (2011) Evolution of three human GPCRs for higher expression and stability. *J Mol Biol* 408(4):599–615.
  22. Choe HW, et al. (2011) Crystal structure of metarhodopsin II. *Nature* 471(7340):651–655.
  23. Rasmussen SG, et al. (2011) Crystal structure of the  $\beta_2$  adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.
  24. Standfuss J, et al. (2011) The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* 471(7340):656–660.
  25. Scheerer P, et al. (2008) Crystal structure of opsin in its G-protein-interacting conformation. *Nature* 455(7212):497–502.
  26. White JF, et al. (2012) Structure of the agonist-bound neurotensin receptor. *Nature* 490(7421):508–513.
  27. Deupi X, Standfuss J (2011) Structural insights into agonist-induced activation of G-protein-coupled receptors. *Curr Opin Struct Biol* 21(4):541–551.
  28. Botto JM, et al. (1997) Identification in the rat neurotensin receptor of amino-acid residues critical for the binding of neurotensin. *Brain Res Mol Brain Res* 46(1-2):311–317.
  29. Barroso S, et al. (2000) Identification of residues involved in neurotensin binding and modeling of the agonist binding site in neurotensin receptor 1. *J Biol Chem* 275(1):328–336.
  30. Huynh J, Thomas WG, Aguilar MI, Pattenden LK (2009) Role of helix 8 in G protein-coupled receptors based on structure-function studies on the type 1 angiotensin receptor. *Mol Cell Endocrinol* 302(2):118–127.
  31. Delos Santos NM, Gardner LA, White SW, Bahouth SW (2006) Characterization of the residues in helix 8 of the human beta1-adrenergic receptor that are involved in coupling the receptor to G proteins. *J Biol Chem* 281(18):12896–12907.
  32. Feng GJ, et al. (2003) Selective interactions between helix VIII of the human mu-opioid receptors and the C terminus of periplakin disrupt G protein activation. *J Biol Chem* 278(35):33400–33407.
  33. Zhong M, Navratil AM, Clay C, Sanborn BM (2004) Residues in the hydrophilic face of putative helix 8 of oxytocin receptor are important for receptor function. *Biochemistry* 43(12):3490–3498.
  34. Kirchberg K, et al. (2011) Conformational dynamics of helix 8 in the GPCR rhodopsin controls arrestin activation in the desensitization process. *Proc Natl Acad Sci USA* 108(46):18690–18695.
  35. Lodowski DT, et al. (2007) Crystal packing analysis of Rhodopsin crystals. *J Struct Biol* 158(3):455–462.
  36. Faussner A, et al. (2005) The role of helix 8 and of the cytosolic C-termini in the internalization and signal transduction of B(1) and B(2) bradykinin receptors. *FEBS J* 272(1):129–140.
  37. Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* 234(4):946–950.
  38. Wu B, et al. (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* 330(6007):1066–1071.
  39. Zhang C, et al. (2012) High-resolution crystal structure of human protease-activated receptor 1. *Nature* 492(7429):387–392.
  40. Kim TY, Schlieter T, Haase S, Alexiev U (2012) Activation and molecular recognition of the GPCR rhodopsin—Insights from time-resolved fluorescence depolarisation and single molecule experiments. *Eur J Cell Biol* 91(4):300–310.
  41. Bruno A, Costantino G, de Fabritiis G, Pastor M, Selent J (2012) Membrane-sensitive conformational states of helix 8 in the metabotropic Glu2 receptor, a class C GPCR. *PLoS ONE* 7(8):e42023.
  42. Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP (2008) Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* 454(7201):183–187.
  43. Mundy DI, Warren G (1992) Mitosis and inhibition of intracellular transport stimulate palmitoylation of a 62-kD protein. *J Cell Biol* 116(1):135–146.
  44. Qanbar R, Bouvier M (2003) Role of palmitoylation/depalmitoylation reactions in G-protein-coupled receptor function. *Pharmacol Ther* 97(1):1–33.
  45. Chini B, Parenti M (2009) G-protein-coupled receptors, cholesterol and palmitoylation: Facts about fats. *J Mol Endocrinol* 42(5):371–379.
  46. Tucker J, Grishammer R (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem J* 317(Pt 3):891–899.
  47. White JF, Trinh LB, Shiloach J, Grishammer R (2004) Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett* 564(3):289–293.
  48. Grishammer R, White JF, Trinh LB, Shiloach J (2005) Large-scale expression and purification of a G-protein-coupled receptor for structure determination—An overview. *J Struct Funct Genomics* 6(2-3):159–163.
  49. Heakal Y, et al. (2011) Neurotensin receptor-1 inducible palmitoylation is required for efficient receptor-mediated mitogenic-signaling within structured membrane microdomains. *Cancer Biol Ther* 12(5):427–435.

# Supporting Information

Egloff et al. 10.1073/pnas.1317903111

## SI Text

**Materials.** The tritiated agonist [<sup>3</sup>H] neurotensin (NT) ([β,11-tyrosyl-3,5-<sup>3</sup>H(N)]-pyroGlu-Leu-Tyr-Glu-Asn-Lys-Pro-Arg-Arg-Pro-Tyr-Ile-Leu) and [<sup>35</sup>S]GTPγS (1250 Ci/mmol) were purchased from Perkin-Elmer. HyLite Fluor 647-labeled NT8-13 (NT8-13-HL647) was purchased from Anawa. Unlabeled NT8-13 (Arg-Arg-Pro-Tyr-Ile-Leu) and NT1 were purchased from Anaspec. N-decyl-β-D-maltopyranoside (DM), 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS), n-nonyl-β-D-glucopyranoside (NG), n-decyl-β-D-glucopyranoside (DG), n-dodecyl-β-D-maltopyranoside (DDM), n-octyl-β-D-glucopyranoside (OG), and n-dodecyl-β-D-glucopyranoside (DDG) were obtained from Anatrace. Cholesteryl hemisuccinate tris salt (CHS) and lysozyme were purchased from Sigma. Empty PD10 columns, N-hydroxysuccinimide (NHS)-activated Sepharose, sulphopropyl (SP)-Sepharose, and Superdex 200 10/300 GL were obtained from GE Healthcare. Complete EDTA-free protease inhibitor mixture and DNase I were purchased from Roche. The neurotensin receptor 1 (NTR1) antagonist SR142948 was obtained from Axon Medchem. Sf-900 II serum-free media and Protein G Dynabeads were purchased from Life Technologies. Nickel-nitrilotriacetic acid (Ni-NTA) resin was obtained from Qiagen. Amicon Ultra concentrators, glass fiber filtration plates (MultiScreen-FB), and nitrocellulose filtration plates (MultiScreen<sub>HTS</sub>-HA plates) were purchased from Millipore.

**Construct Design.** All NTR1 variants were expressed in a pBR322-derived vector, which was constructed from a plasmid originally obtained as a kind gift from R. Grishammer (National Institute of Neurological Disorders and Stroke, National Institutes of Health, Rockville, MD) (1, 2) and was modified to encode now an N-terminal maltose-binding protein (MBP) linked via a hexahistidine tag and a human rhinovirus (HRV) 3C protease site to G50 (sequential NTR1 numbering). The variants were C-terminally truncated at G390 and linked via a HRV 3C protease site, a pentasparagine linker, and a di-glycine-serine linker to thioredoxin A (TrxA), which is followed by a deca-histidine tag. Amino acids V280-I295 were deleted in the constructs ΔIC3A and E273-T290 in ΔIC3B. Directed evolution of NTR1 was performed as previously described (3, 4). For mammalian expression, all constructs were subcloned into the vector pcDNA3 (Life Technologies) containing an N-terminal FLAG epitope.

**Ligand-Binding Experiments.** Whole-cell radioligand-binding assays (RLBAs) in *Escherichia coli* were used to perform affinity measurements. Receptors were expressed in BL21 Tuner cells (Novagen) in 50 mL 2YT medium with 0.3% glucose and 100 μg/mL ampicillin. Cultures were inoculated to OD<sub>600</sub> = 0.05 and grown at 37 °C to OD<sub>600</sub> = 0.5. G protein-coupled receptor (GPCR) expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 20 h at 24 °C or for 16 h at 28 °C, respectively. From the expression cultures, 80 μL (TM86V-ΔIC3A) or 800 μL (NTR1 wild-type) samples were centrifuged, and the pellet was resuspended in 3 mL binding buffer (50 mM Tris-HCl pH 7.4, 0.1% BSA, 1 mM EDTA, and 40 μg/mL bacitracin). We added 20 μL of resuspended cells to 160 μL binding buffer containing [<sup>3</sup>H]NT at various concentrations for agonist saturation-binding experiments. Antagonist competition experiments were performed using a binding buffer with constant [<sup>3</sup>H]NT concentrations (10 nM) and various concentrations of the antagonist SR142948.

For determination of expression levels of NTR1 wild-type and TM86V-ΔIC3A in *Spodoptera frugiperda* (Sf9) insect cells, membranes were incubated in 200 μL binding buffer containing 10 nM [<sup>3</sup>H]NT. NTR1 expression in Sf9 insect cells and membrane preparation is described below.

For both RLBA on *E. coli* and Sf9 insect cells, all binding reaction mixtures were incubated for 2 h at 4 °C in a 96-well plate on a plate shaker. Note that not all binding reactions for TM86V-ΔIC3A may have completely reached equilibrium after 2 h incubation due to slow on-rates or off-rates. Nonspecific binding was determined in the presence of 10 μM unlabeled NT. The unbound ligand was separated by vacuum filtration on 96-well glass fiber filtration plates (Millipore) that were pretreated with polyethylenimine. Filters were washed four times by 200 μL ice-cold 50 mM Tris-HCl pH 7.4, and they were dried for 1 h at 65 °C. Subsequently, they were dissolved in 200 μL OptiPhase Super-Mix (PerkinElmer) per well. Scintillation was counted on a Microbeta 1450 Plus liquid scintillation counter (Wallac). Data were analyzed by nonlinear regression using GraphPad5 Prism software and fit to one-site binding equations: for agonist saturation-binding experiments after subtraction of background,  $y = B_{\max} * L / (L + K_d)$ ; for antagonist competition binding experiments,  $y = B + (T - B) / (1 + 10^{(x - \log(IC_{50}))})$ , where B describes the minimum level of the signal and T the maximum. Less than 10% of free ligand was bound by the receptors in all reactions.

**G Protein Purification.** G protein composed of Gα<sub>i1</sub>, Gβ<sub>1</sub>, and Gγ<sub>1</sub> was expressed in Sf9 cells using a single baculovirus encoding all three subunits, which was generated by using the MultiBac system (5, 6). To purify the heterotrimeric G protein, a 3C-protease-cleavable deca-histidine tag had been fused to the N terminus of Gβ<sub>1</sub>, whereas an HA tag had been fused to the N terminus of Gγ<sub>1</sub> for immobilization on magnetic beads. Sf9 cells grown in Sf-900 II serum-free media were infected at a density of  $7 \times 10^6$  cells/mL with a multiplicity of infection (MOI) of 5 with the G protein-encoding virus. After 72 h incubation, the infected cells were harvested by centrifugation and resuspended in 30 mL lysis buffer (50 mM Hepes pH 8.0, 50 mM NaCl, 1 mM MgCl<sub>2</sub>, 10 μM GDP, 5 mM β-mercaptoethanol, complete protease inhibitor mixture) and lysed by sonication. The lysate was centrifuged for 5 min at 500 × g to remove cell debris, and the resulting supernatant was ultracentrifuged for 40 min at 108,000 × g to collect the membranes. The G protein was purified from the membranes following the procedure described by Rasmussen et al. (7).

**Pull-Down Experiments of GPCR by Immobilized G Protein.** TM86V-ΔIC3A, as N-terminal MBP and C-terminal TrxA fusion, was expressed in *E. coli* as described above. We incubated  $3.1 \times 10^9$  cells in PBS containing 1.25 mg/mL lysozyme, 20 μg/mL DNaseI, 1 mM MgCl<sub>2</sub>, and complete protease inhibitor mixture for 45 min at 4 °C. After the addition of 5 mM EDTA, cells were disrupted by sonication. Cell debris was removed by a low speed centrifugation (10 min at 5,000 × g, 4 °C), and cell membranes were pelleted by ultracentrifugation at 118,000 × g for 1 h (4 °C). The membrane pellet was resuspended in 1.5 mL 20 mM Hepes pH 8.0, 50 mM NaCl, 30% (vol/vol) glycerol, and flash-frozen for storage at -80 °C.

For the pull-down experiment, 20 μL *E. coli* membranes and 25 μg purified G protein were incubated for 2 h and 40 min, respectively, at 4 °C in 500 μL solubilization buffer [20 mM Hepes pH 8.0, 100 mM NaCl, 2 mM MgCl<sub>2</sub>, 20 μM NT8-13, 30% (vol/vol) glycerol, 0.6% CHAPS, 0.12% CHS, 1% DDM, complete protease

inhibitor mixture]. Before the binding of the HA-tagged G protein to the anti-HA-antibody-coated (Sigma, H9658) Protein G Dynabeads (5  $\mu$ g antibody for 900  $\mu$ g beads/condition), a centrifugation at 20,000  $\times$  g for 30 min was performed to remove potential protein aggregates. Before incubation of the GPCR with the beads for 1 h, nonsolubilized material was removed by ultracentrifugation at 86,000  $\times$  g for 30 min. Protein bound to the beads was eluted with standard 1 $\times$  SDS loading buffer. Between all incubation and elution steps, the beads were washed extensively with solubilization buffer. Eluted proteins were analyzed by Western blot with antibodies against thioredoxin (Sigma, T0803) and G $\alpha_{i1}$  (Lifespan Biosciences, LS-B4007).

**Isolation of Sf9 Membranes.** The GPCRs, tagged N-terminally with tobacco etch virus (TEV) protease-cleavable FLAG and decahistidine tags, were expressed in Sf9 cells by infection with GPCR virus at an MOI of 5. At 60 h after infection, cells were lysed by incubation in lysis buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA, 5  $\mu$ g/mL Leupeptin, 0.1 mM Pefabloc SC, 1  $\mu$ g/mL Pepstatin) for 30 min at 4  $^{\circ}$ C and subsequently forcing them several times through a 27G1/4 needle. After a low-speed centrifugation at 1,000  $\times$  g, membranes were collected at 20,000  $\times$  g and incubated for 30 min in buffer A (50 mM Tris-HCl pH 7.4, 1 mM EDTA) containing 7 M urea to remove peripherally bound proteins. The urea concentration was then reduced to 3.5 M by adding buffer A, and the membranes were collected again by centrifugation. The membranes were washed once with buffer A and were flash-frozen for storage at  $-80^{\circ}$  C in buffer A containing 20% (wt/vol) sucrose. For ligand-binding experiments (i.e., Fig. S4), the urea wash step was omitted.

**[ $^{35}$ S]GTP $\gamma$ S Assay.** The [ $^{35}$ S]GTP $\gamma$ S assay was performed with slight changes from a previously described protocol (8). Briefly, urea-washed membranes containing 1 nM of GPCR, which were mixed with 100 nM purified G protein ( $\alpha_{i1}\beta\gamma_1$ ) and 200  $\mu$ M NT or no ligand, were incubated in a total volume of 50  $\mu$ L of assay buffer (50 mM Tris-HCl pH 7.4, 1 mM EDTA, 100 mM NaCl, 1 mM DTT, 3 mM MgSO $_4$ , 0.3% BSA, 2  $\mu$ M GDP, 4 nM [ $^{35}$ S]GTP $\gamma$ S) for 20 min at 25  $^{\circ}$ C. The reaction was stopped by filtration over nitrocellulose filters (Multiscreen<sub>HTS</sub>-HA plates) and washing four times with ice-cold wash buffer (50 mM Tris-HCl pH 7.4, 1 mM EDTA). Counts arising from G protein alone (non-GPCR-stimulated GTP $\gamma$ S binding) had been subtracted from all reactions. The counts for GPCR-containing urea-washed membranes—that is, without G protein—were all in the same range (Fig. S4). Thus, the higher basal counts (compared with NTR1) observed for TM86V- $\Delta$ IC3A and the TM86V- $\Delta$ IC3A L167R back-mutant are due to intrinsic properties of the GPCRs, and not due to potentially different membrane preparations. Furthermore, the amounts of membrane used for the mutants had been less than for NTR1.

**Large-Scale Expression of NTR1 Variants.** A 1 L preculture of *E. coli* BL21 Tuner cells harboring the expression plasmid was grown overnight at 37  $^{\circ}$ C in 2YT medium, containing 1% glucose and 50  $\mu$ g/mL ampicillin. A 50 L fermenter (Bioengineering) containing 2YT medium, 0.6% glucose, and 50  $\mu$ g/mL ampicillin, was inoculated by the complete preculture and grown to an OD $_{600}$  of 2.5 at 37  $^{\circ}$ C, followed by induction with 1 mM IPTG at 29  $^{\circ}$ C overnight. Cells were harvested using a continuous-flow centrifuge.

**Large-Scale Purification of NTR1 Variants.** In a typical purification, 50 g of cell pellet (corresponding to a 7% aliquot of a fermenter run) were resuspended in 100 mL 2 $\times$  solubilization buffer containing 100 mM Hepes pH 8, 20% (vol/vol) glycerol, 400 mM NaCl, and five tablets of complete EDTA-free protease inhibitor mixture. All subsequent steps were carried out at 4  $^{\circ}$ C. We added 0.5 mL of 1 M MgCl $_2$ , a spatula tip of DNase I, 200 mg lysozyme,

20 mL of CHAPS [6% (wt/vol)]/CHS [1.2% (wt/vol)], and 34 mL of 10% (wt/vol) DM to the resuspended cells while stirring. The mixture was incubated for 15 min. Sonication was performed for 30 min in an ice-water bath. Subsequently, 4 mL of 0.5 M EDTA was added, and the mixture was incubated for another 30 min while stirring. The suspension was centrifuged for 30 min at 15,000 rpm (SLA 1500 rotor), and the supernatant was decanted and mixed to 5 mL slurry pD-NT ligand resin followed by incubation overnight. The mixture was centrifuged at 400  $\times$  g for 10 min, and the supernatant was decanted. The pelleted resin was packed into an empty PD10 column, and it was washed on a bench-top pressure-flow device with 75 mL of wash buffer 1, containing 25 mM Hepes pH 8, 10% (vol/vol) glycerol, 600 mM NaCl, and 0.5% DM. It was then washed with 40 mL of wash buffer 2, containing 25 mM Hepes pH 7, 10% (vol/vol) glycerol, 150 mM NaCl, 4 mM DTT, and 0.4% NG. The resin was resuspended in a small volume of wash buffer 2, containing 0.7 mg of HRV 3C protease (produced in house), followed by incubation for 2 h. The eluted protein (10 mL) was diluted three times with SP binding buffer, containing 10 mM Hepes pH 7, 10% (vol/vol) glycerol, 4 mM DTT, and 0.3% NG, and it was subjected to another PD10 column containing 5 mL SP Sepharose, which had been preequilibrated with SP binding buffer. The SP resin was washed with 10 mL SP binding buffer, followed by 25 mL SP wash buffer containing 10 mM Hepes pH 7.7, 10% (vol/vol) glycerol, 35 mM NaCl, 4 mM DTT, and 0.3% NG, followed by another 3 mL SP binding buffer. Elution was carried out by  $\sim$ 15 mL SP elution buffer containing 10 mM Hepes pH 7, 10% (vol/vol) glycerol, 350 mM NaCl, 4 mM DTT, 0.3% NG, and 500 nM NT1. The NTR1 variants were concentrated by an Amicon-15 Ultra concentrator with 50 kDa cutoff to less than 500  $\mu$ L and loaded on a Superdex 200 10/300 GL column preequilibrated in a buffer containing 10 mM Hepes pH 8.0, 150 mM NaCl, 4 mM DTT, 0.28% NG, and 100 nM NT1. Peak fractions were pooled (1.5 mL) and concentrated by an Amicon-4 Ultra concentrator (50 kDa cutoff) to 3–20 mg/mL.

**Vapor Diffusion Crystallization.** For TM86V- $\Delta$ IC3A, initial crystals were observed at 5 mg/mL protein concentrations in a sitting drop vapor diffusion experiment. The best diffraction quality crystals were obtained, when before crystallization 10% (vol/vol) of X-mix solution 1, containing 10% (wt/vol) NG, 5% (wt/vol) DG, 0.1% DDG, and 1% CHS, was added to the concentrated protein. Crystals were observed after 24 h and grew for about 1 wk in sitting drops containing 1  $\mu$ L of protein/X-mix solution and 1  $\mu$ L reservoir solution [50 mM glycine pH 9.4, 1 M NaCl, 26% (vol/vol) PEG 600] at 4  $^{\circ}$ C. Crystals of TM86V- $\Delta$ IC3B and OGG7- $\Delta$ IC3A were analogously prepared by using as reservoir a solution of 21.5% (vol/vol) PEG 600, 2 M NaCl, 50 mM glycine pH 9.4, or 22.5% (vol/vol) PEG 600, 0.5 M NaCl, 50 mM glycine pH 9.4, respectively. For HTGH4- $\Delta$ IC3A crystallization, the purified protein was mixed with 10% (vol/vol) of X-mix solution 2, containing 10% (wt/vol) OG, 5% (wt/vol) DG, 0.1% DDG, and 1% CHS, and the reservoir solution was 20% (vol/vol) PEG 600, 0.2 M CaCl $_2$ , 50 mM NaAcetate pH 5.5. Crystals were harvested by 0.3–0.4 mm loops and cryoprotected by incubation for 3  $\times$  1 min in a solution containing 50 mM glycine pH 9.4, 1 M NaCl, 36% (vol/vol) PEG 600, and 0.3% NG for TM86V- $\Delta$ IC3A, TM86V- $\Delta$ IC3B, and OGG7- $\Delta$ IC3A, or in 36% (vol/vol) PEG 600, 0.2 M CaCl $_2$ , 50 mM NaAcetate pH 5.5 for HTGH4- $\Delta$ IC3A, respectively. Crystals were flash-frozen in liquid propane.

**Data Collection and Structure Determination.** All crystals belonged to the space group P2 $_1$ 2 $_1$ 2 $_1$ , with two protein molecules in the asymmetric unit. Data were collected from a single, cryocooled (100 K) crystal at the beamline  $\times$ 06SA at the Swiss Light Source with a PILATUS 6 M high-resolution diffractometer and at

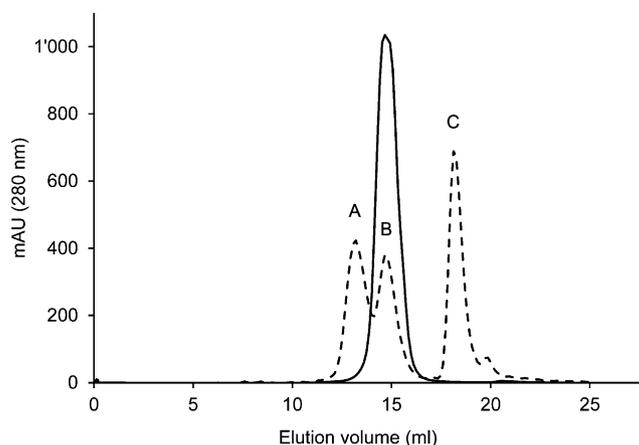
the wavelength  $\lambda = 1 \text{ \AA}$ . Data were processed and scaled with the X-Ray Detector Software (XDS) package (9).

The initial structure of TM86V- $\Delta$ IC3A was determined by molecular replacement in Phaser (10) using the nociceptin/orphanin FQ receptor structure [Protein Data Bank (PDB) ID code 4EA3] as a search model. The search model was chosen based on the BLAST search of the PDB results (on June 25, 2012) with E-value 4.3087e-19, % Identity = 26.91 and % Positives = 46.48, followed by model preparation in Sculptor (11). Initial maps after molecular replacement were improved by the mr\_rosetta protocol (12, 13) with 5 Rosetta (14) models in Phenix (15). Model building was carried out in Coot (16), and refinement was performed with REFMAC5 (17) and phenix.refine (18). The structures of TM86V- $\Delta$ IC3B, HTGH4- $\Delta$ IC3A, and OGG7- $\Delta$ IC3A

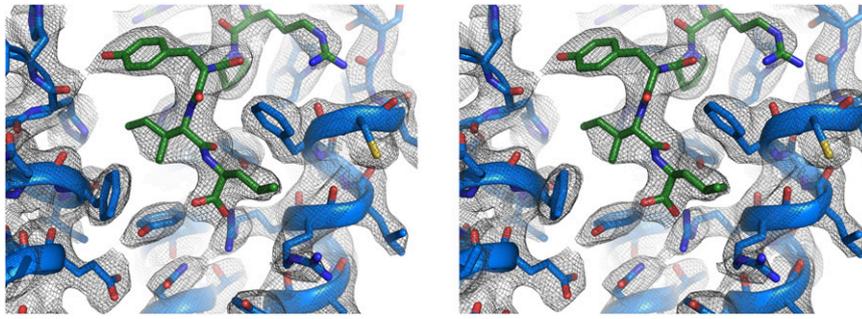
were determined by molecular replacement using TM86V- $\Delta$ IC3A as a search model. The crystallographic statistics (Table S2) were generated with Phenix Graphic Tools (19) and adjusted manually.

**Life Cell Microscopy.** HEK293T cells were cultured on 18-mm coverslips and transiently transfected with GPCRs and  $\beta$ -arrestin2-YFP. At 48 h after transfection, coverslips were mounted in an imaging chamber and maintained in Tyrode buffer (137 mM NaCl, 5.4 mM KCl, 2 mM  $\text{CaCl}_2$ , 1 mM  $\text{MgCl}_2$ , 10 mM Hepes pH 7.3) at 37 °C. Cells were stimulated with 50 nM NT8-13-HyLite647, and confocal images were taken at indicated time points on a Leica TCS SP5 laser scanning microscope using single-line excitation at 514 nm and 633 nm and an emission bandwidth of 525–580 and 645–720 nm.

1. Tucker J, Grishammer R (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem J* 317(Pt 3):891–899.
2. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* 105(39):14808–14813.
3. Schlinkmann KM, et al. (2012) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3):414–428.
4. Scott DJ, Plückthun A (2013) Direct molecular evolution of detergent-stable G protein-coupled receptors using polymer encapsulated cells. *J Mol Biol* 425(3):662–677.
5. Fitzgerald DJ, et al. (2006) Protein complex expression by using multigene baculoviral vectors. *Nat Methods* 3(12):1021–1032.
6. Bieniossek C, Imasaki T, Takagi Y, Berger I (2012) MultiBac: Expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci* 37(2):49–57.
7. Rasmussen SG, et al. (2011) Crystal structure of the  $\beta_2$  adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.
8. Shibata Y, et al. (2009) Thermostabilization of the neurotensin receptor NTS1. *J Mol Biol* 390(2):262–277.
9. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125–132.
10. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
11. Bunkóczi G, Read RJ (2011) Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):303–312.
12. DiMaio F, et al. (2011) Improved molecular replacement by density- and energy-guided protein structure optimization. *Nature* 473(7348):540–543.
13. DiMaio F, Tyka MD, Baker ML, Chiu W, Baker D (2009) Refinement of protein structures into low-resolution density maps using rosetta. *J Mol Biol* 392(1):181–190.
14. Leaver-Fay A, et al. (2011) ROSETTA3: An object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* 487:545–574.
15. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
16. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
17. Murshudov GN, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–367.
18. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352–367.
19. Echols N, et al. (2012) Graphical tools for macromolecular crystallography in PHENIX. *J Appl Cryst* 45(Pt 3):581–586.

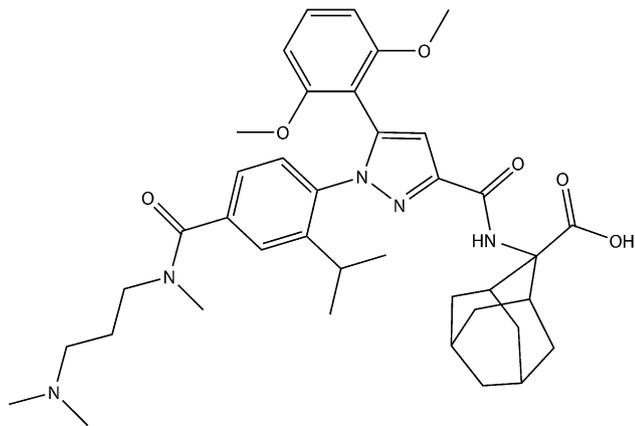


**Fig. S1.** Preparative size exclusion chromatogram of TM86V- $\Delta$ IC3A (solid line) in the short-chain detergent NG. A monodisperse symmetric peak was observed despite the harsh nature of the detergent. The last step of a representative large-scale purification from whole *E. coli* cells with a final yield of about 4 mg purified TM86V- $\Delta$ IC3A (from 3.5 L *E. coli* culture) is depicted. Wild-type NTR1 could not be purified under these conditions due to its low stability when solubilized. The harsh short-chain detergents were chosen for crystallization because they are expected to form small micelles around the GPCR. Therefore, they occlude the small hydrophilic domains of GPCRs less than long-chain detergents, and as a result, they allow crystal contact formation.  $\beta$ -amylase (A, 200 kDa), BSA (B, 66 kDa), and cytochrome c (C, 12.7 kDa) were used as reference markers (dashed line). The size exclusion runs were performed by a Superdex 200 10/300 GL column from GE-Healthcare (25 mL column volume).



**Fig. S2.** Quality of the  $\sigma_A$ -weighted  $2F_o-F_c$  electron density map contoured at  $1.2\sigma$ . Stereoview on the five C-terminal neurotensin residues Arg9 to L13 (green) and on parts of transmembrane helix 2 (TM2) (Left), TM6 (Right), and TM7 (back) of TM86V- $\Delta$ IC3B (blue). TM1 and TM3–TM5 (front) were omitted for clarity.

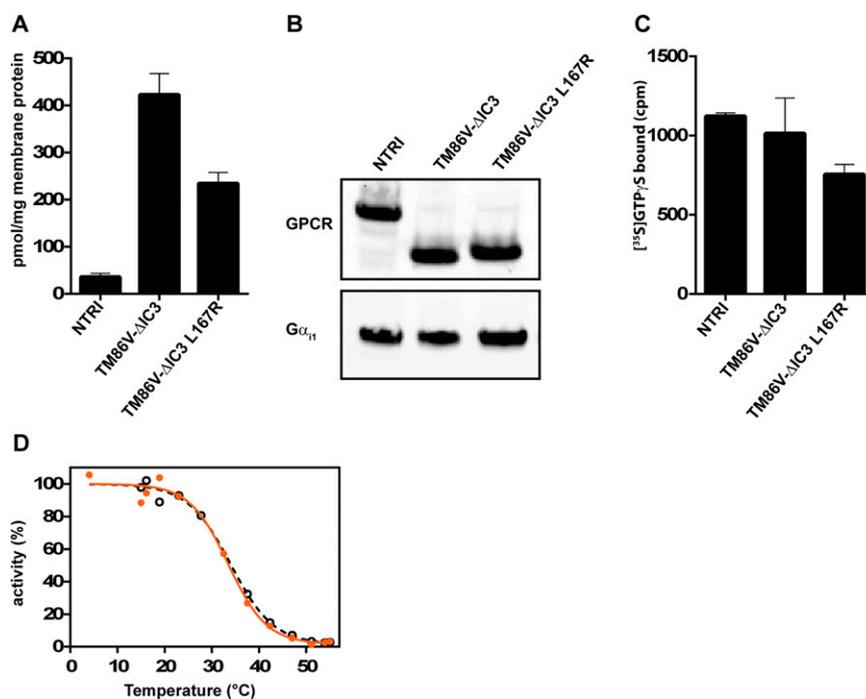
**A**



**B**

	TM86V- $\Delta$ IC3A	NTR1 wild-type
$K_d^{app}$ NT	$2.3 \pm 0.4$ nM	$2.8 \pm 0.3$ nM
$IC_{50}^{app}$ SR142948	$30.0 \pm 2.4$ nM	$8.4 \pm 0.9$ nM

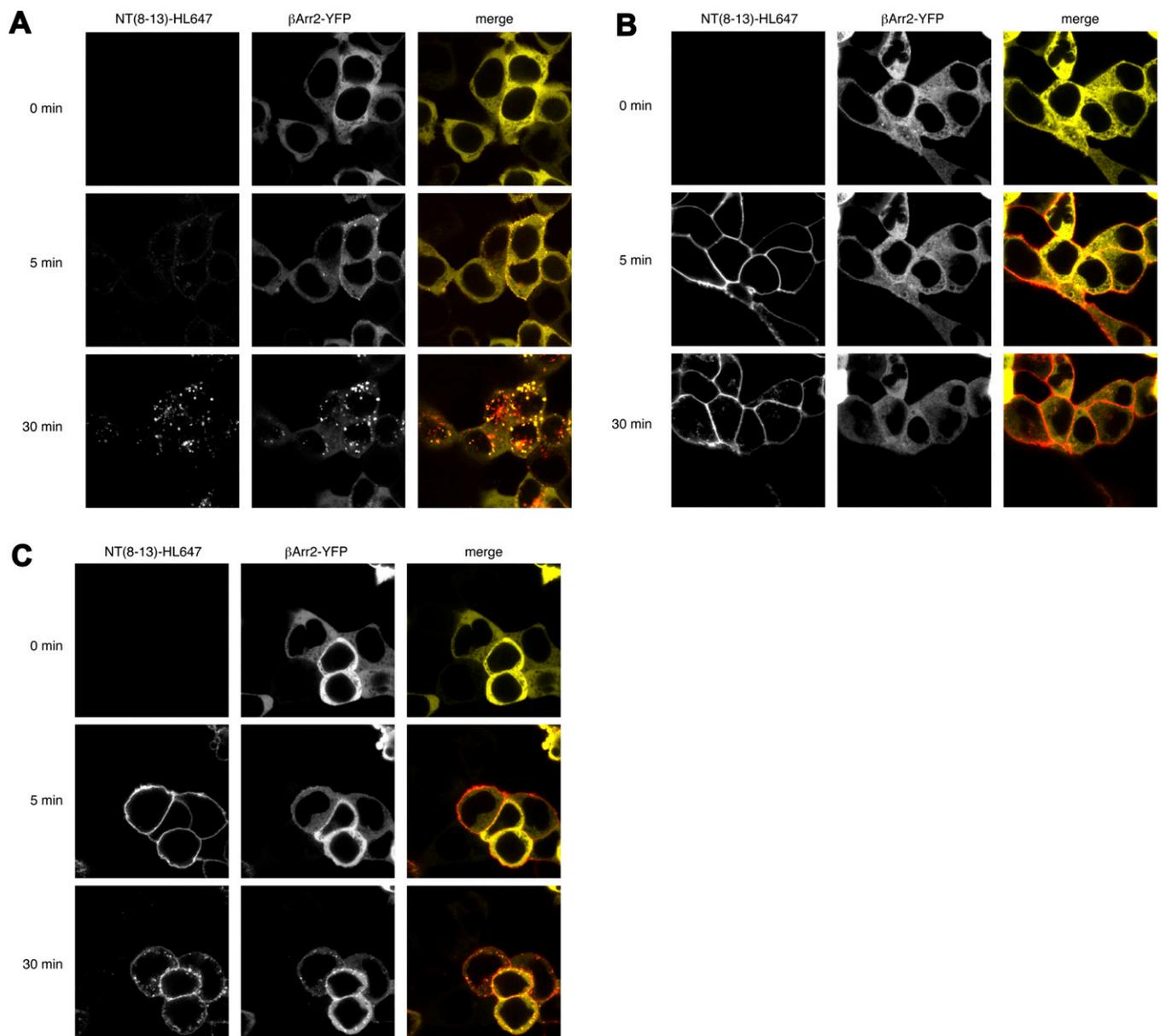
**Fig. S3.** Antagonist SR142948 (A) and summary table for neurotensin saturation-binding and antagonist competition experiments (B). Identical  $^3H$ -NT concentrations were applied for wild-type NTR1 and TM86V- $\Delta$ IC3A in competition experiments.



**Fig. S4.** Functional and total expression levels of wild-type NTR1 and of evolved variants in Sf9 insect cells. (A) Radioligand-binding assays with [<sup>3</sup>H]-neurotensin were carried out to compare the relative functional expression levels of wild-type NTR1, TM86V-ΔIC3A, and TM86V-ΔIC3A L167<sup>3.50</sup>R (back-mutation). As previously reported (1), the higher functional expression levels of the evolved variants in *E. coli* also translate to higher values in eukaryotic expression systems. The back-mutation L167<sup>3.50</sup>R, which restores the highly conserved D/ERY motif in TM86V-ΔIC3A, causes a 40% drop of functional expression. (B) Although the functional expression levels of the three protein variants are different, quantitative Western blot experiments of the membranes used in A imply that the total expression level is similar. As a reference, G $\alpha_{11}$  is shown, which has been coexpressed. (C) Background binding of [<sup>35</sup>S]GTP $\gamma$ S to the membranes only (without added G protein), expressing the different receptors, under same the reaction conditions. Note that the background levels seen here are insignificant (compare cpm here to those in Fig. 2C) and that they do not correlate with the amount of membrane applied. This control is important, as for the [<sup>35</sup>S]GTP $\gamma$ S assay (Fig. 2A), we applied equal amounts of functional receptors, and due to the large differences in functional expression levels, different amounts of GPCR-containing membranes had to be used. (D) Thermal denaturation profiles of TM86V (orange) and TM86V-L167<sup>3.50</sup>R (black) in the detergent  $\beta$ -D-octyl glucopyranoside. Remaining agonist-binding activity ([<sup>3</sup>H]-neurotensin) was measured after thermal denaturation in the agonist-bound state for 20 min at the indicated temperature. TM86V displays an apparent  $T_M$  of 33.5 °C compared with 34 °C for TM86V-L167<sup>3.50</sup>R. Data of a representative measurement are shown. The method has been described previously (2, 3).

1. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* 105(39):14808–14813.
2. Schlinkmann KM, et al. (2012) Maximizing detergent stability and functional expression of a GPCR by exhaustive recombination and evolution. *J Mol Biol* 422(3):414–428.
3. Dodevski I, Plückthun A (2011) Evolution of three human GPCRs for higher expression and stability. *J Mol Biol* 408(4):599–615.

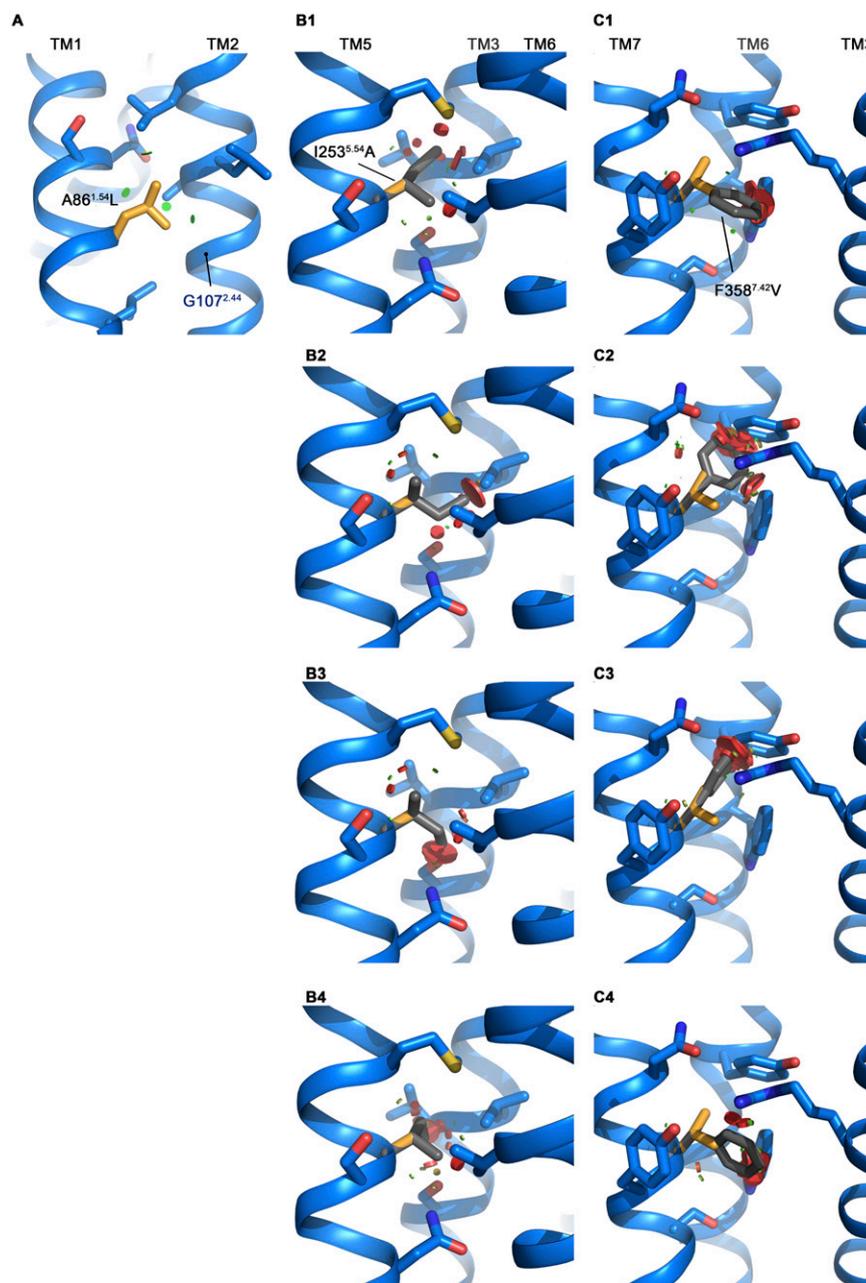




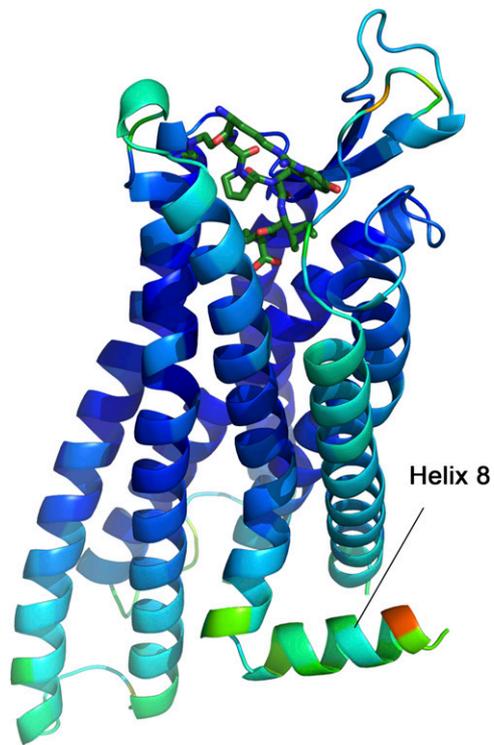
**Fig. 56.** Agonist binding and  $\beta$ -arrestin2-dependent internalization of NTR1. Confocal images of HEK293T cells coexpressing  $\beta$ -arrestin2-YFP (yellow) with either (A) wild-type NTR1, (B) TM86V- $\Delta$ IC3A, or (C) TM86V- $\Delta$ IC3A L167R-CT (reconstituted R167<sup>3,50</sup> and full C-terminal tail), after stimulation with 50 nM fluorescent neurotensin [NT8-13-HL647; red]. All receptor variants bound the agonist readily at the plasma membrane. Wild-type NTR1 strongly cointernalized with  $\beta$ -arrestin2. TM86V- $\Delta$ IC3A was internalized weakly, but no clear  $\beta$ -arrestin2 interaction was observed. However, when Arg167<sup>3,50</sup> of the conserved D/ERY motif as well as the C-terminal tail had been reconstituted (to allow receptor phosphorylation and  $\beta$ -arrestin2 binding), a clear cointernalization was visible. These findings suggest that TM86V- $\Delta$ IC3A L167R-CT can interact with  $\beta$ -arrestin2 in a fashion similar to wild type.







**Fig. S8.** Close-up view on three thermostabilizing mutations in NTR1-TM86V, completing Fig. 3. (A) The selected A86L mutation (orange) allows for additional van der Waals contacts in silico between TM1 and TM2 (green circles). (B) The I253A mutation appears to remove clashes (red circles) between residues of TM5, TM3, and TM6. The selected alanine is shown in orange, and the four most common rotamers (1–4) of the wild-type isoleucine are depicted in gray. (C) The selected F358V mutation appears to remove clashes between residues of TM5, TM3, and TM6. The selected valine is shown in orange, and the four most common rotamers (1–4) of the wild-type phenylalanine are depicted in gray.



**Fig. 59.** Relative B-factor coloring on the structure of TM86VΔIC3B. From high to low B-factors: red, yellow, green, and blue. The region with the highest B-factors corresponds to the amphipathic helix 8. Note that NT8-13 is shown independent of B-factors in dark green.



**Table S2. Data collection and refinement statistics**

	TM86V- $\Delta$ IC3B	TM86V- $\Delta$ IC3A	OGG7- $\Delta$ IC3A	HTGH4- $\Delta$ IC3A
<b>Data collection</b>				
Wavelength, Å	1.000	1.000	1.000	1.000
Resolution, Å	50–2.75 (2.84–2.75)*	50–3.0 (3.11–3.0)	50–3.1 (3.21–3.1)	50–3.56 (3.7–3.56)
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			
<b>Unit cell</b>				
a, b, c, Å	63.3, 89.4, 212.1	58.6, 90.2, 209.4	60.6, 91.6, 208.6	60.9, 90.9, 211.0
$\alpha$ , $\beta$ , $\gamma$ , °	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Multiplicity	13.1 (13.9)	6.9 (5.7)	6.3 (5.6)	7.2 (7.6)
Completeness, %	99.9% (99.8%)	99.8% (98.7%)	99.2% (97.0%)	99.7% (98.4%)
Mean $I/\sigma(I)$	19.85 (0.45)	8.68 (0.76)	15.55 (0.51)	8.88 (0.32)
$R_{\text{merge}}$	0.08062 (8.91)	0.1083 (2.821)	0.1026 (3.716)	0.108 (7.401)
$CC_{1/2}$	1 (0.232)	0.997 (0.294)	0.999 (0.126)	0.999 (0.227)
<b>Refinement</b>				
Resolution, Å	19.82–2.75	19.93–3.0	19.88–3.1	21.88–3.57
Total reflections	422,226 (43,292)	158,535 (12,678)	136,489 (11,754)	104,564 (10,617)
$R_{\text{work}}/R_{\text{free}}$	0.2478/0.2728	0.2418/0.2792	0.2840/0.3105	0.3092/0.3449
Number of atoms	5,049	4,956	4,865	4,838
Macromolecules	5,049	4,956	4,865	4,838
Ligands	14	5	0	0
Water	0	0	0	0
<b>rms deviations</b>				
Bond lengths, Å	0.003	0.003	0.003	0.003
Bond angles, °	0.65	0.74	0.65	0.78
Ramachandran favored, %	96	99	97	97
Ramachandran outliers, %	0	0	0	0.17
Average B-factor	82	125.7	82.6	189.7
Macromolecules	82	125.7	82.6	189.7
Ligands	91.89	136.28	n/a	n/a
Water	n/a	n/a	n/a	n/a

For each construct the data were collected from a single crystal. Signal to noise of  $I/\sigma(I) = 2.0$  was at 3.15 Å, 3.26 Å, 3.55 Å, and 4.14 Å for TM86V- $\Delta$ IC3B, TM86V- $\Delta$ IC3A, OGG7- $\Delta$ IC3A, and HTGH4- $\Delta$ IC3A, respectively.

\*Highest resolution shell is shown in parentheses.

**Table S3. Hydrogen bonds and salt bridges between neurotensin ligand and TM86V- $\Delta$ IC3A**

NTR1 atoms	Distance, Å	Ligand atoms
<b>Hydrogen bonds*</b>		
B:THR 226 [OG1]	2.91	D:TYR 11 [O]
B:TYR 347 [OH]	3.17	D:ILE 12 [O]
B:ARG 327 [NH1]	3.23	D:LEU 13 [O]
B:ARG 327 [NH2]	3.03	D:LEU 13 [O]
B:TYR 146 [OH]	2.70	D:LEU 13 [OXT]
B:GLU 337 [O]	3.40	D:GLY 6 [N*]
B:TRP 339 [O]	3.09	D:GLY 7 [N*]
B:ASP 54 [O]	3.18	D:ARG 8 [NH2]
B:ASP 56 [O]	2.56	D:ARG 8 [NH2]
B:ILE 334 [O]	2.21	D:ARG 9 [NH2]
B:LEU 55 [O]	2.86	D:TYR 11 [OH]
B:HIS 132 [O]	3.25	D:TYR 11 [OH]
<b>Salt bridges*</b>		
B:ARG 327 [NH1]	3.23	D:LEU 13 [O]
B:ARG 327 [NH2]	3.03	D:LEU 13 [O]
B:ASP 336 [OD1]	3.19	D:ARG 9 [NE]
B:ASP 336 [OD1]	3.43	D:ARG 9 [NH1]
B:ASP 336 [OD1]	3.07	D:ARG 9 [NH2]

\*Analysis performed using the PDBEPIA server. Atoms of linker amino acids of the in-house produced peptide agonist are labeled with an asterisk.