Supplementary Information

for

Title: A comprehensive analysis of heterotrimeric G-protein complex diversity and their interactions with GPCRs in solution.

Authors: Matthias Hillenbrand, Christian Schori, Jendrik Schöppe and Andreas Plückthun[a]

Affiliations: [a]Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

Corresponding author: plueckthun@bioc.uzh.ch, Tel: +41-44-635 5570

Keywords: Heterotrimeric G proteins, G-protein-coupled receptor, Membrane protein, protein complex, protein-protein interaction

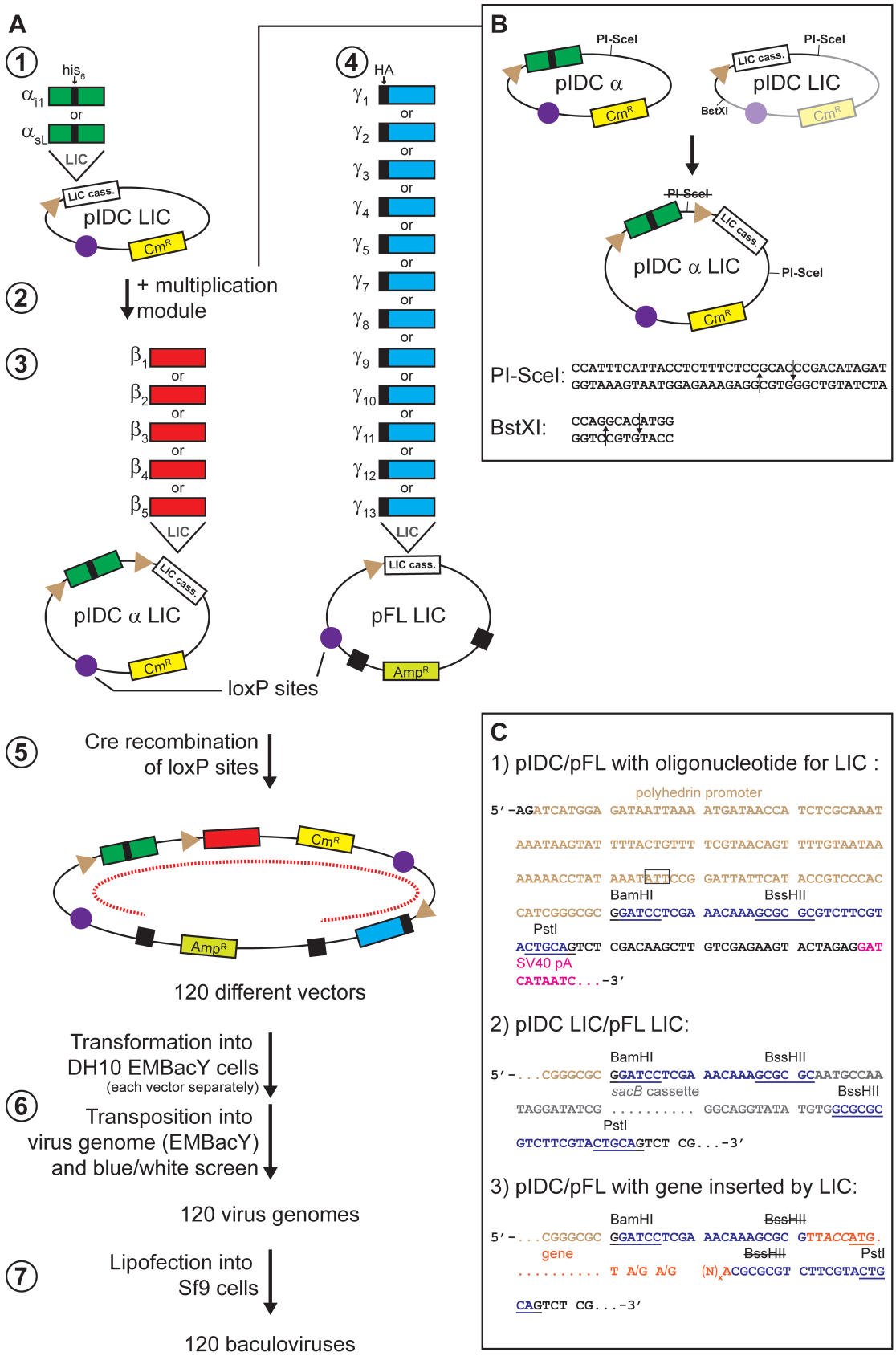


Fig. S1: Schematic representation of the assembly of Gαβγ subunits on a baculovirus. (A) (1) First one of the two α-subunits with an internal hexahistidine-tag are cloned into the donor vector "pIDC LIC" with the help of ligation independent cloning (LIC). (2) A new expression cassette is inserted with the help of the multiplication module (see panel (B) for details). (3)

One of the five β-subunits is cloned into the α_{i1}- or α_{sL}-containing "pIDC α LIC" derivative again via LIC, resulting in 10 different Gαβ combinations. (4) Each one of twelve N-terminally HA-tagged γ-subunits is cloned into the acceptor vector "pFL LIC" by LIC. (5) With the help of loxP sites (violet circles) an *in vitro* Cre recombination is performed. This results in 120 different Gαβγ combinations with a subunit gene stoichiometry of 1:1:1 on one transfer vector, where each of the subunit genes is under the control of a polyhedrin promoter (light brown triangle). (6) Transfer vectors are separately transformed into *E. coli* DH10 EMBacY cells. There, a transposition with Tn7, expressed in these cells, of the expression cassettes into the baculovirus genome takes place (sequences between the two recognition sites (black squares); part indicated by red dotted line). (7) The virus genomes are isolated and then transfected into Sf9 cells, leading to production of the first virus generation. The vector pIDC contains a chloramphenicol resistance gene (Cm^R), pFL an ampicillin resistance gene (Amp^R). (B) Use of the multiplication module to place two (or more) genes on one pIDC vector. In the cloning strategy presented here the pIDC vector containing the α-subunit ("pIDC α") is linearized by the action of the homing endonuclease PI-SceI. Of the second vector ("pIDC LIC") the "empty" expression module (promoter, LIC site and polyadenylation site (not shown for clarity)) is excised by PI-SceI and BstXI (recognition sequences shown at the bottom of the panel). As BstXI-generated overhangs have been designed (general recognition sequence: 5'-CCANNNNNNTGG-3') to be compatible to PI-SceI overhangs, the expression cassette can be inserted into linearized "pIDC α" yielding "pIDC α LIC". As a result of the cloning the original PI-SceI site of the recipient vector is destroyed, while a new PI-SceI site is generated downstream of the newly inserted expression module, which can again be utilized to integrate another module. (C) DNA sequences describing promoter region, cloning site/gene and polyadenylation site of the used constructs. 1) Sequence of vector pIDC or pFL after inserting oligonucleotides between BamHI and PstI containing a designed LIC cassette (blue) without any insert. The full sequence of the polyhedrin promoter (light brown) and part of the SV40 polyadenylation site (SV40 pA; pink) are shown. Note that the boxed ATT in the polyhedrin promoter represents the original start-ATG of the polyhedrin gene (mutated to ATT; see ref. (1)), which is not used, since the target gene will bring its own ATG start codon. 2) Sequences of the vectors pIDC LIC and pFL LIC used in panel (A), after a *sacB* expression cassette has been introduced into the BssHII site as a negative selection marker to prepare for LIC. 3) pIDC and pFL vectors carrying a gene and accompanying sequences (orange) introduced by LIC after treating the vectors in 2 with BssHII (removes *sacB* cassette). For details on LIC see "Materials and Methods" in the main text. The gene starting with the start-ATG (underlined) and ending with a stop codon (T A/G A/G), in some constructs followed by additional nucleotides, as was present on the pcDNA3.1+ vectors, is preceded by a Kozak sequence (ACC, italic) and the sequence "TT". Whereas the first "T" is important for generation of a defined overhang during the T4 polymerase treatment (part of the LIC reaction), the second "T" had been introduced to avoid a placement of the gene in-frame to the mutated polyhedrin start-ATG (boxed ATT in 1) and following two ATT's (see ref. (2)). For some constructs additional sequences ("$(N)_x$") after the stop codon have been carried over for reasons of primer design. This is followed by an "A" that fulfills the same function in the LIC reaction as the first "T".

Additional note: In order to subsequently introduce the β-subunit into "pIDC α LIC" the vector has to be opened by BssHII. As α_{i1} contains one recognition sequence for this enzyme, the site in α_{i1} has been removed by silent point mutation in codon no. 31 (numbering includes start ATG) GC**G** to GC**T.**

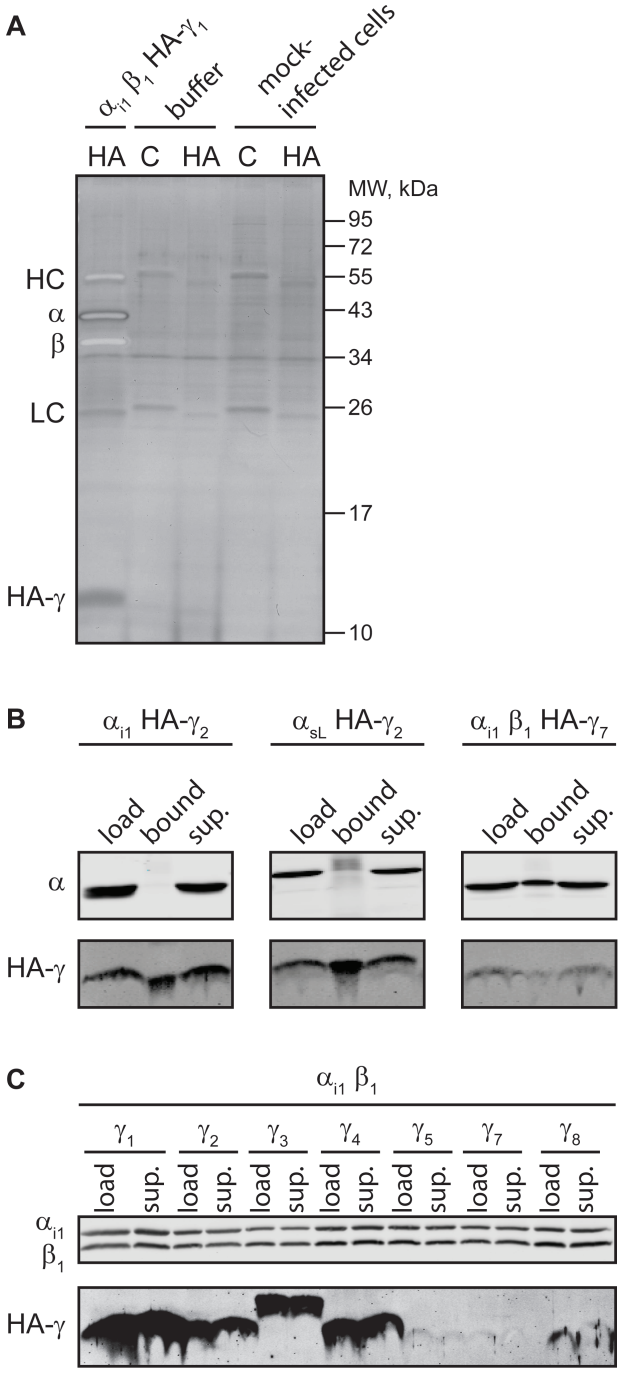


Fig. S2: Control experiments for the co-immunoprecipitation of Gαβγ heterotrimers. (A) Anti-HA ("HA") or control ("C") beads were incubated either with solubilized G protein (α_{i1} β_1 HA-γ_1), with buffer or a solubilizate of mock-infected cells. Proteins bound to the beads were visualized by silver gel. Only in the case of the G protein are three bands visible, which correspond to the α-, β- and γ-subunits. An additional band below the β-subunit (at 34 kDa) can be seen in all lanes. The heavy ("HC") and light ("LC") chain of the antibodies used for the pulldown are also visible. (B) α_{i1} or α_{sL} were expressed together with HA-γ_2, but without a β-subunit. As a reference, the full heterotrimeric G protein $\alpha_{i1}\beta_1$HA-γ_7 was expressed. Those three combinations were solubilized and subjected to a co-immunoprecipitation by anti-HA

beads. The western blots detecting the α- and γ-subunit show that both subunits were expressed (“load”), but only in the case of the full G protein the α-subunit is co-immunoprecipitated (“bound”), whereas the γ-subunit is present in all cases, due to its HA-tag. It should be noted that the light shadow visible above the expected α_{sL} band in the “bound” lane is due to a crossreactivity of the secondary antibody with the antibody used for the immunoprecipitation. This becomes visible, because the exposure time for the α_{sL}-subunit was higher than for the α_{i1}-subunit. The supernatant of the beads after incubation with the anti-HA beads is also shown (“sup.”). (C) Comparison of the amount of α-, β- and γ-subunits before (“load”) and after (“sup.”) incubation with anti-HA beads. Equal amounts were loaded and analyzed by western blot. No decrease in signal strength after incubation with beads is visible, indicating a saturation of the beads.

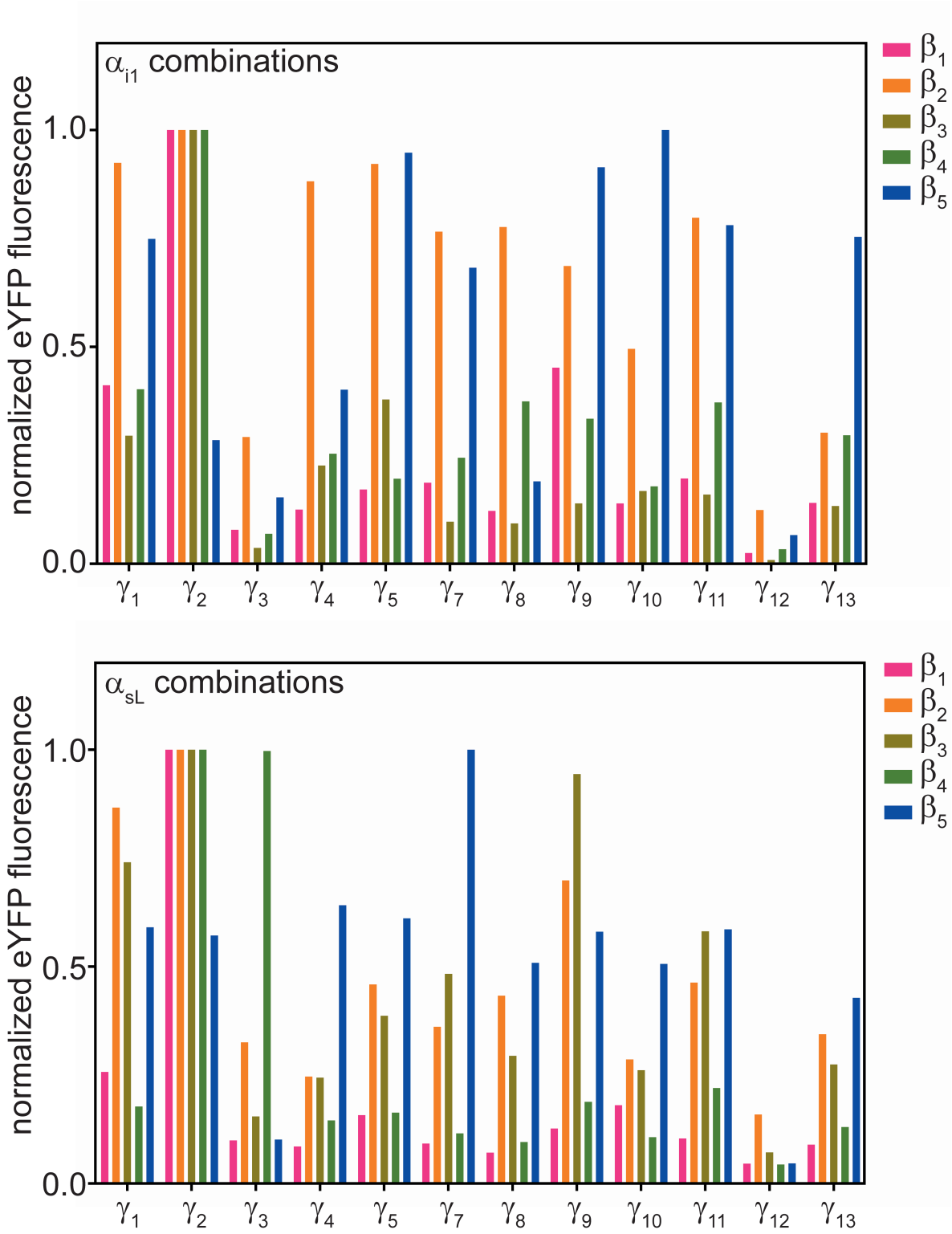


Fig. S3: eYFP fluorescence measured in the cleared supernatants of solubilized G protein expressing cells, which were used for incubation with the antibody-coated beads (co-immunoprecipitation). Since eYFP is co-expressed with the heterotrimeric G-protein complex (see ”Material and Methods”), it can aid as an indicator for general protein expression level of each virus. Interestingly, eYFP fluorescence is generally lowest for G-protein combinations containing γ_{12} (and to some extend also γ_3). This is an indication that expression of this γ-subunit(s) pose a burden for the expression machinery of the cell. This finding may explain

why expression levels of G-protein α- and β-subunits, co-expressed especially with γ_{12}, are low in Fig. 1 C and D. The generation of new viruses or the use of more virus for expression did not resolve the found differences of expression levels shown in Fig. 1 C and D. The eYFP fluorescence signals were normalized by setting the highest signal to 1 for each group of 12, consisting of defined α- and β-subunit and variable γ-subunit ($\gamma_{1\text{-}5,7\text{-}13}$); i.e., normalization took place within bars of the same color of each chart.

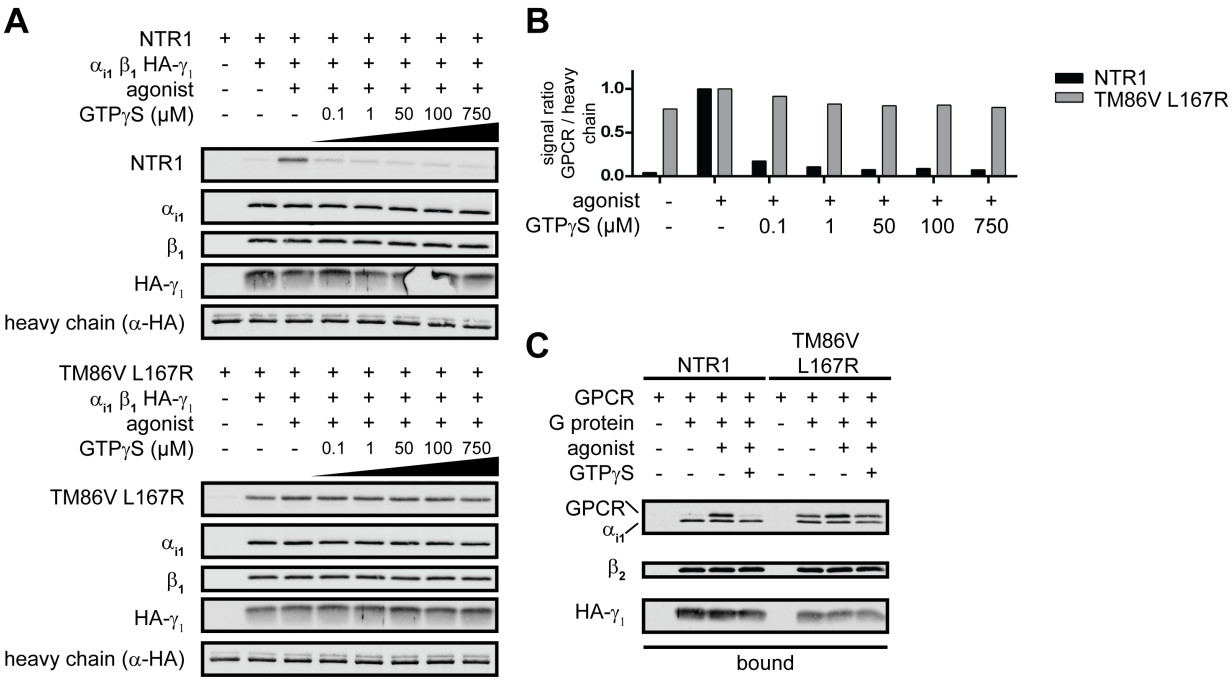
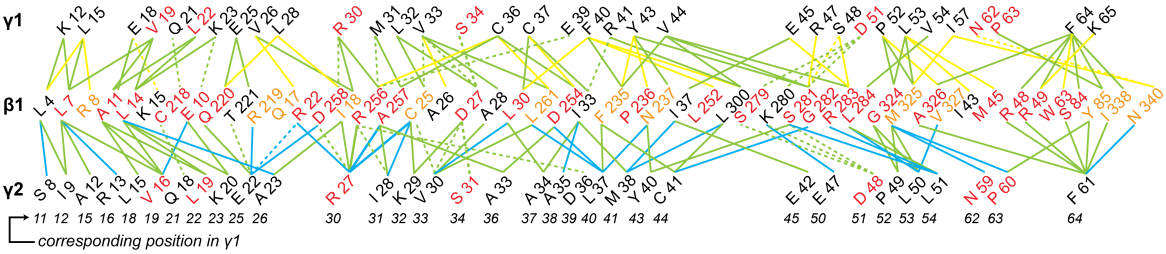


Fig. S4: Results of co-immunoprecipitation experiments. (A) and (B) GTPγS titration experiments. (A) Anti-HA (α-HA) beads were incubated first with (+) or without (-) G protein ($\alpha_{i1}\beta_1$HA-γ_1). In a second incubation the beads (theoretical concentration of G protein is around 50 nM) were incubated with solubilized GPCR (NTR1 or TM86V L167R) in the presence (+) or absence (-) of agonist (20 μM neurotensin) and the indicated amount of GTPγS (0-750 μM). The protein bound to the beads after washing is visualized by western blot. As an internal loading control the heavy chain of the antibody used for the pull down is shown. The effects of the agonist and GTPγS on the binding of GPCR to the G protein are visualized in (B) by the signal ratio of the GPCR band and the heavy chain of the antibody. Titration of GTPγS reveals that already at 100 nM GTPγS a significant amount of GPCR has been dissociated from the G protein. However, for both GPCRs analyzed the effect was reaching a plateau between 1 and 50 μM GTPγS. We decided to use the rather high GTPγS concentration of 750 μM to secure maximal dissociation in the following experiments. GPCR/G-protein complexes on beads (in a similar concentration range) have been disassembled by the use of a comparable high (100 μM) GTPγS concentration in ref. (3). (C) Co-immunoprecipitation of G protein ($\alpha_{i1}\beta_2$HA-γ_1) and NTR1 wild-type or TM86V L167R. After incubation of the solubilized G protein with anti-HA beads, the beads were incubated with solubilized GPCR. The presence of GPCR/G protein bound to the beads as a function of the presence of G protein, agonist (20 μM neurotensin) and GTPγS (750 μM) is shown by western blot ("bound"). The blot shows the same behavior of $\alpha_{i1}\beta_2$HA-γ_1 regarding GPCR interaction as is seen for $\alpha_{i1}\beta_1$HA-γ_1, indicating a functional coupling of this newly reported G-protein combination, which was previously assumed to be non-existent.

A



B

Interface residues found for β1 that are not conserved throughout β1-β4

β1	β2	β3	β4
L 4	conserved	M	conserved
K 15	R	conserved	R
A 26 (Cα)	G	conserved	N
A 28 (O)	S	V	conserved
I 33	conserved	L	conserved
I 37	L	L	M
I 43	conserved	V	conserved
T 221	conserved	conserved	S
K 280 (N)	R	L	conserved
L 300	M	M	conserved

C

```
                                  |10        |20        |30        |40        |50        |60        |70
                                  .     :  *:.*: :    *   *:   ::  :        * ::       **:       *
Bovine     γ1       --MPVINIEDLTEKDKLKMEVDQLKKEVTLERMLVSKCCEEFRDYVEERSGEDPLVKGIPEDKNPFKELKGGCVIS  74
Human      γ1       --MPVINIEDLTEKDKLKMEVDQLKKEVTLERMLVSKCCEEVRDYVEERSGEDPLVKGIPEDKNPFKELKGGCVIS  74
Human[bov.] γ2       -----MASNNTASIAQARKLVEQLKMEANIDRIKVSKAAADLMAYCEAHAKEDPLLTPVPASENPFREKKFFCAIL  71
Human[bov.] γ3       MDKGETPVNSTMSIGQARKMVEQLKIEASLCRIKVSKAAADLMTYCDAHACEDPLITPVPTSENPFREKKFFCALL  76
Human      γ4       -MKEGMSNNSTTSISQARKAVEQLKMEACMDRVKVSQAAADLLAYCEAHVREDPLIIPVPASENPFREKKFFCTIL  75
Human[bov.] γ5       -------MSGSSSVAAMKKVVQQLRLEAGLNRVKVSQAAADLKQFCLQNAQHDPLLTGVSSSTNPFRPQKV-CSFL  68
Human      γ7       -------MSATNNIAQARKLVEQLRIEAGIERIKVSKAASDLMSYCEQHARNDPLLVGVPASENPFKDKKP-CIIL  68
Human      γ8       ------MSNNMAKIAEARKTVEQLKLEVNIDRMKVSQAAAELLAFCETHAKDDPLVTPVPAAENPFRDKRLFCVLL  70
Human      γ9       ------MAQDLSEKDLLKMEVEQLKKEVKNTRIPISKAGKEIKEYVEAQAGNDPFLKGIPEDKNPFKE-KGGCLIS  69
Human      γ10      -------MSSGASASALQRLVEQLKLEAGVERIKVSQAAAELQQYCMQNACKDALLVGVPAGSNPFREPRS-CALL  68
Human      γ11      --MPALHIEDLPEKEKLKMEVEQLRKEVKLQRQQVSKCSEEIKNYIEERSGEDPLVKGIPEDKNPFKE-KGSCVIS  73
Human      γ12      ---MSSKTASTNNIAQARRTVQQLRLEASIERIKVSKASADLMSYCEEHARSDPLLIGIPTSENPFKDKKT-CIIL  72
Human      γ13      --------MEEWDVPQMKKEVESLKYQLAFQREMASKTIPELLKWIEDGIPKDPFLNPDLMKNNPWVEKGK-CTIL  67
```

[bov.]: human and bovine sequence are 100% identical.

D

```
                                  |10        |20        |30        |40        |50        |60        |70        |80
                                : .*:.***.*: :: : *    *  * ::.  :: :*:. *:*****:**  *: .*.*  *.: :**:*****:*:
Human[bov.] β1       --------MSELDQLRQEAEQLKNQIRDARKACADATLSQITNNIDPVGRIQMRTRRTLRGHLAKIYAMHWGTDSRLLVSASQDGKLII  81
Human[bov.] β2       --------MSELEQLRQEAEQLRNQIRDARKACGDSTLTQITAGLDPVGRIQMRTRRTLRGHLAKIYAMHWGTDSRLLVSASQDGKLII  81
Human      β3       --------MGEMEQLRQEAEQLKKQIADARKACADVTLAELVSGLEVVGRVQMRTRRTLRGHLAKIYAMHWATDSKLLVSASQDGKLIV  81
Human      β4       --------MSELEQLRQEAEQLRNQIQDARKACNDATLVQITSNMDSVGRIQMRTRRTLRGHLAKIYAMHWGYDSRLLVSASQDGKLII  81
Human      β5       MATEGLHENETLASLKSEAESLKGKLEEERAKLHDVELHQVAERVEALGQFVMKTRRTLKGHGNKVLCMDWCKDKRRIVSSSQDGKVIV  89

                                  |90        |100       |110       |120       |130       |140       |150       |160
                                ***:**** **: : .:***:******* :******* **:* *.  : :  .  .:.:  ** ***.* * : * :*:*:*** ***
Human[bov.] β1       WDSYTTNKVHAIPLRSSWVMTCAYAPSGNYVACGGLDNICSIYNLKTREGN--VRVSRELAGHTGYLSCCRFLD-DNQIVTSSGDTTCA  167
Human[bov.] β2       WDSYTTNKVHAIPLRSSWVMTCAYAPSGNFVACGGLDNICSIYSLKTREGN--VRVSRELPGHTGYLSCCRFLD-DNQIITSSGDTTCA  167
Human      β3       WDSYTTNKVHAIPLRSSWVMTCAYAPSGNFVACGGLDNMCSIYNLKSREGN--VKVSRELSAHTGYLSCCRFLD-DNNIVTSSGDTTCA  167
Human      β4       WDSYTTNKMHAIPLRSSWVMTCAYAPSGNYVACGGLDNICSIYNLKTREGN--VRVSRELPGHTGYLSCCRFLD-DSQIVTSSGDTTCA  167
Human      β5       WDSFTTNKEHAVTMPCTWVMACAYAPSGCAIACGGLDNKCSVYPLTFDKNENMAAKKKSVAMHTNYLSACSFTNSDMQILTASGDGTCA  178

                                  |170       |180       |190       |200       |210       |220       |230       |240       |250
                                ***:*:.*    * ** .* ::* ::*.   . *:**.** .  :**:*.. * *:*  * ****:: ::*.* *: :*****:***:*
Human[bov.] β1       LWDIETGQQTTTFTGHTGDVMSLSLAPD--TRLFVSGACDASAKLWDVREGMCRQTFTGHESDINAICFFPNGNAFATGSDDATCRLFD  254
Human[bov.] β2       LWDIETGQQTVGFAGHSGDVMSLSLAPD--GRTFVSGACDASIKLWDVRDSMCRQTFIGHESDINAVAFFPNGYAFTTGSDDATCRLFD  254
Human      β3       LWDIETGQQKTVFVGHTGDCMSLAVSPD--FNLFISGACDASAKLWDVREGTCRQTFTGHESDINAICFFPNGEAICTGSDDASCRLFD  254
Human      β4       LWDIETAQQTTTFTGHSGDVMSLSLSPD--MRTFVSGACDASSKLWDIRDGMCRQSFTGHVSDINAVSFFPNGYAFATGSDDATCRLFD  254
Human      β5       LWDVESGQLLQSFHGHGADVLCLDLAPSETGNTFVSGGCDKKAMVWDMRSGQCVQAFETHESDINSVRYYPSGDAFASGSDDATCRLYD  267

                                  |260       |270       |280       |290       |300       |310       |320       |330       |340
                                ****:*:  :*::.** * :** ** *****:***:*:. *:** :*..*..:* **:**** * *: ** *..:****  *::*
Human[bov.] β1       LRADQELMTYSHDNIICGITSVSFSKSGRLLLAGYDDFNCNVWDALKADRAGVLAGHDNRVSCLGVTDDGMAVATGSWDSFLKIWN  340
Human[bov.] β2       LRADQELLMYSHDNIICGITSVAFSRSGRLLLAGYDDFNCNIWDAMKGDRAGVLAGHDNRVSCLGVTDDGMAVATGSWDSFLKIWN  340
Human      β3       LRADQELICFSHESIICGITSVAFSLSGRLLFAGYDDFNCNVWDSMKSERVGILSGHDNRVSCLGVTADGMAVATGSWDSFLKIWN  340
Human      β4       LRADQELLLYSHDNIICGITSVAFSKSGRLLLAGYDDFNCNVWDTLKGDRAGVLAGHDNRVSCLGVTDDGMAVATGSWDSFLRIWN  340
Human      β5       LRADREVAIYSKESIIFGASSVDFSLSGRLLFAGYNDYTINVWDVLKGSRVSILFGHENRVSTLRVSPDGTAFCSGSWDHTLRVWA  353
```

[bov.]: human and bovine sequence are 100% identical.

E

```
                                :** **: *****:**:***.  *  :**. :*:::*:* ::*:**::************ **.*:**
Human      γ9       ----MAQDLSEKDLLKMEVEQLKKEVKNTRIPISKAGKEIKEYVEAQAGNDPFKE-KGGCLIS   69
Human      γ1       MPVINIEDLTEKDKLKMEVDQLKKEVTLERMLVSKCCEEVRDYVEERSGEDPLVKGIPEDKNPFKELKGGCVIS   74
Human      γ11      MPALHIEDLPEKEKLKMEVEQLRKEVKLQRQQVSKCSEEIKNYIEERSGEDPLVKGIPEDKNPFKE-KGSCVIS   73
```

Fig. S5: Interaction interfaces and multiple sequence alignments of β- and γ-subunits. (A) Interaction interface of the bovine $\beta_1\gamma_1$ (PDB ID: 1TBG) and bovine $\beta_1\gamma_2$ (1GP2) dimer. Amino acids (one letter code and position number) involved in interaction between β- and γ-subunits are given for each subunit. For a better comparison between γ_1 and γ_2 the corresponding positions of γ_1 are indicated below the amino acids of γ_2. Amino acids colored in *red* indicate a full conservation throughout all β- or γ-subunits, *orange* indicates a full conservation throughout β_1-β_4, i.e. excluding β_5, and *black* indicates more variable positions. Interactions between amino acids of β- and γ-subunits are indicated by *solid* lines for hydrophobic interactions and *dashed* lines for hydrogen bonds. *Green* lines indicate that the interaction between the two positions in β and γ is found in both structures ($\beta_1\gamma_1$ and $\beta_1\gamma_2$), whereas *yellow* or *blue* lines indicate unique interactions found only in $\beta_1\gamma_1$ or $\beta_1\gamma_2$, respectively.

Inspection of the interaction network reveals that 65-70% of interactions are found in both structures. Interactions were defined by an interatomic distance of less than 4 Å and burying of ≥20% of the total solvent-accessible surface area upon interaction. Hydrogen bonds were inspected by eye. Structures were analyzed by the bioinformatic (web)tools ContPro (4), the EBI PDBe PISA web server (5) and PyMOL (Schrödinger).

(B) 10 position identified in (A) in the interaction interface of β_1 that are not conserved throughout β_1-β_4. Positions and their corresponding residue in β_2-β_4 are given. For residues that contact the γ-subunit solely via the peptide backbone, the relevant atoms are indicated in parentheses. Substitutions of β_1 residues to the corresponding residue in β_2-β_4 that potentially lead to steric clashes in the inspected structures (PDB IDs: 1TBG ($\beta_1\gamma_1$), 1GOT ($\alpha_{t/i1}\beta_1\gamma_1$), 1GP2 ($\alpha_{i1}\beta_1\gamma_2$), 3SN6 ($\alpha_s\beta_1\gamma_2/\beta_2$ adrenergic receptor complex) and 3AH8 ($\alpha_{i1/q}\beta_1\gamma_2$)) are indicated by a flash. A *green flash* indicates a clash that is found for structures containing $\beta_1\gamma_1$ and $\beta_1\gamma_2$, a *yellow flash* indicates clashes found for $\beta_1\gamma_1$ containing structures and a *blue flash* indicates clashes found for $\beta_1\gamma_2$ containing structures. Nevertheless, the clashes for position I37 could not be found in structures 1GOT and 3AH8. Steric clashes were analyzed by PyMOL (Schrödinger).

(C) Multiple sequence alignment of γ-subunits. Shown are all human γ-subunits and the bovine γ_1. Subunits that are 100% identical between human and bovine sequences are indicated with a bov. superscript. Numbering is according to the γ_1-subunit. Residues involved in interaction with β are highlighted in yellow or blue for γ_1 and γ_2, respectively. The last three amino acids (boxed red) are removed upon lipid modification of the preceding C-terminal cysteine and are therefore not present in the mature form of the γ-subunits. The point mutation found between human (L40) and bovine (F40) γ_1-subunit is boxed black. This point mutation has been described to be responsible for the incompatibility between β_2 and γ_1 (6).

(D) Multiple sequence alignment of all human β-subunits. Numbering is according to the β_1-subunit. Residues involved in interaction with γ are highlighted in *green*, *yellow* or *blue* if their interaction with γ has been found in structures of $\beta_1\gamma_1$ (1TBG) and $\beta_1\gamma_2$ (1GP2), only $\beta_1\gamma_1$ or only $\beta_1\gamma_2$, respectively.

(E) Multiple sequence alignment of γ_1, γ_9 and γ_{11}. The last three amino acids (boxed red) are removed upon lipid modification of the preceding C-terminal cysteine and are therefore not present in the mature form of the γ-subunits. All alignments were generated by Clustal O (version 1.2.1) at EMBL-EBI (7, 8). An asterisk indicates positions which are fully conserved. A colon indicates conservation between residues of strongly similar properties (scoring > 0.5 in the Gonnet PAM 250 matrix). A period indicates conservation between groups of weakly similar properties (scoring ≤ 0.5 in the Gonnet PAM 250 matrix).

References

1. Luckow VA, Summers MD (1989) High level expression of nonfused foreign genes with Autographa californica nuclear polyhedrosis virus expression vectors. *Virology* 170:31–39.

2. López-Ferber M, Sisk WP, Possee RD (1995) Baculovirus transfer vectors. *Methods Mol Biol* 39:25–63.

3. Wu Y et al. (2007) Rapid-mix flow cytometry measurements of subsecond regulation of G protein-coupled receptor ternary complex dynamics by guanine nucleotides. *Anal Biochem* 371:10–20.

4. Firoz A, Malik A, Afzal O, Jha V (2010) ContPro: A web tool for calculating amino acid contact distances in protein from 3D -structures at different distance threshold. *Bioinformation* 5:55–57.

5. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.

6. Lee C, Murakami T, Simonds WF (1995) Identification of a discrete region of the G protein gamma subunit conferring selectivity in beta gamma complex formation. *J Biol Chem* 270:8779–8784.

7. Sievers F et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.

8. Goujon M et al. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38:W695–699.