

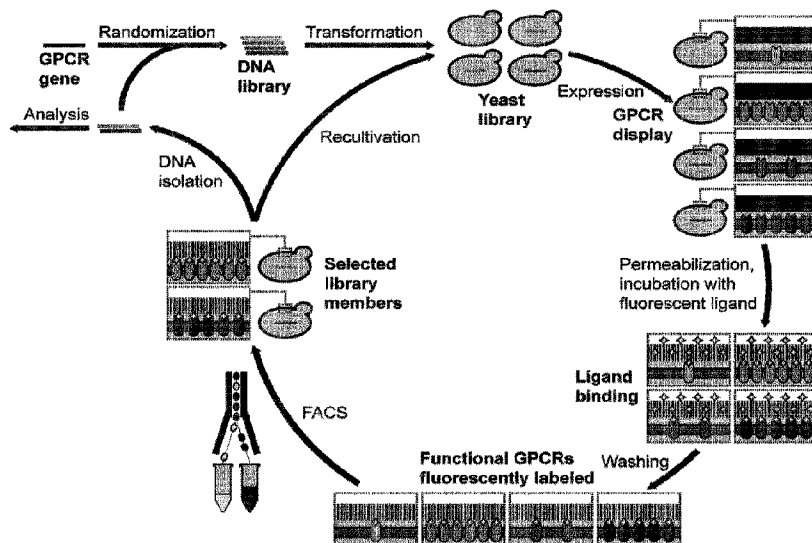


(86) Date de dépôt PCT/PCT Filing Date: 2015/12/21
(87) Date publication PCT/PCT Publication Date: 2016/06/30
(45) Date de délivrance/Issue Date: 2023/05/23
(85) Entrée phase nationale/National Entry: 2017/06/15
(86) N° demande PCT/PCT Application No.: EP 2015/080858
(87) N° publication PCT/PCT Publication No.: 2016/102508
(30) Priorité/Priority: 2014/12/22 (EP14199729.6)

(51) Cl.Int./Int.Cl. *C40B 20/04* (2006.01),
C12N 15/00 (2006.01), *C12Q 1/68* (2018.01),
C40B 20/00 (2006.01), *C40B 30/00* (2006.01),
C40B 40/02 (2006.01)
(72) Inventeurs/Inventors:
SCHUTZ, MARCO, CH;
PLUCKTHUN, ANDREAS, CH
(73) Propriétaire/Owner:
UNIVERSITÄT ZÜRICH, CH
(74) Agent: LAVERY, DE BILLY, LLP

(54) Titre : EVOLUTION DIRIGÉE DE PROTEINES MEMBRANAIRES DANS DES CELLULES EUCARYOTES A PAROI CELLULAIRE

(54) Title: DIRECTED EVOLUTION OF MEMBRANE PROTEINS IN EUKARYOTIC CELLS WITH A CELL WALL



(57) Abrégé/Abstract:

The invention relates to a method for selecting an expressed sequence from a library, comprising the following steps: Each of a plurality of eukaryotic cells comprising a cell wall comprises a nucleic acid sequence member of a library, which is expressed as a target membrane protein in said eukaryotic cells. The cell wall of the cells is permeabilized. The permeabilized cells are labeled with a ligand capable of binding to the target membrane protein. The ligand bears a detectable label. A subset of the labelled cells is selected as a function of detectable label present. Finally, an expressed nucleic acid sequence is isolated from said selection of cells in an isolation step.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
30 June 2016 (30.06.2016)



(10) International Publication Number
WO 2016/102508 A1

- (51) **International Patent Classification:**
G01N 33/566 (2006.01) *G01N 15/00* (2006.01)
C12N 15/00 (2006.01)
- (21) **International Application Number:**
PCT/EP2015/080858
- (22) **International Filing Date:**
21 December 2015 (21.12.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
14199729.6 22 December 2014 (22.12.2014) EP
- (71) **Applicant:** UNIVERSITÄT ZÜRICH [CH/CH]; Rämistrasse 71, 8006 Zürich (CH).
- (72) **Inventors:** SCHÜTZ, Marco; Zschokkestrasse 30, CH-8037 Zürich (CH). PLÜCKTHUN, Andreas; Möhrlistrasse 97, CH-8006 Zürich (CH).
- (74) **Agent:** SCHULZ JUNGHANS PATENTANWÄLTE PARTGMBB; Großbeerenstraße 71, 10963 Berlin (DE).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

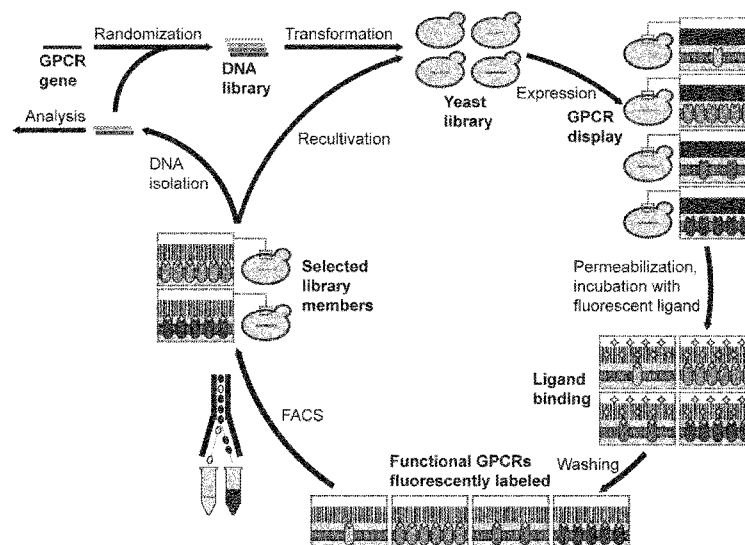
- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) **Title:** DIRECTED EVOLUTION OF MEMBRANE PROTEINS IN EUKARYOTIC CELLS WITH A CELL WALL

Fig. 1



(57) **Abstract:** The invention relates to a method for selecting an expressed sequence from a library, comprising the following steps: Each of a plurality of eukaryotic cells comprising a cell wall comprises a nucleic acid sequence member of a library, which is expressed as a target membrane protein in said eukaryotic cells. The cell wall of the cells is permeabilized. The permeabilized cells are labeled with a ligand capable of binding to the target membrane protein. The ligand bears a detectable label. A subset of the labelled cells is selected as a function of detectable label present. Finally, an expressed nucleic acid sequence is isolated from said selection of cells in an isolation step.

Directed evolution of membrane proteins in eukaryotic cells with a cell wall

Description

Structural and detailed biochemical investigation of eukaryotic transmembrane receptors is still hampered despite technical advances in expression, purification and crystallization strategies. Important difficulties are the lacking ability to produce the desired receptors in sufficient amounts and their inherent instability.

A method to increase functional expression levels of GPCRs in *Escherichia coli* by directed evolution is known in the art (Sarkar et al. (2008), PNAS 105:14808–14813; Dodevski & Plückthun (2011), J Mol Biol 408:599–615). Several high-expression variants of GPCRs have been isolated with this method. However, since *E. coli* is not able to perform post-translational modifications, lacks the secretory quality control and translocation machinery for eukaryotic membrane proteins, and differs in membrane composition, evolution of eukaryotic transmembrane receptors in *E. coli* is limited to receptors which inherently express in a prokaryotic system above the threshold required for successful evolution, thus not having strict requirements for eukaryotic processing or membrane composition.

Therefore the establishment of a fast, efficient and robust method to increase functional expression of transmembrane receptors in a eukaryotic host is required. In principle the ideal eukaryotic host would be cells that are used for the production of the evolved high-expression variants. However, these mammalian or insect cells are not easily transformed with libraries, which is essential for any type of directed evolution approach. The usage of “lower” eukaryotes such as the yeast *Saccharomyces cerevisiae* is common for other procedures involving libraries. However, yeast cells comprise a cell wall, which limits functionality assays by restricting access of administered ligands to expressed receptors located in the plasma membrane. For soluble proteins, a common solution to this problem is the use of methods such as yeast display, which expresses the protein of interest as fusion protein that is presented outside of the cell wall. However, this approach is not suitable for selection of functional high-expression variants of insoluble transmembrane receptors, which must be located in the plasma membrane, and thus below the cell wall, and cannot be expressed as fusion proteins attached to the cell wall.

The problem underlying the present invention is to provide novel efficient methods to facilitate and accelerate research and drug design in transmembrane receptors, particularly by identifying transmembrane protein variants resulting in optimized functional expression in eukaryotic cells comprising a cell wall. This problem is solved by the subject-matter of the independent claims.

Terms and definitions

In the context of the present specification, the term *alkaline pH* is used in its meaning known in the art of chemistry; it refers to a measure of the basicity of an aqueous solution. A pH greater than 7 indicates alkaline or basic solutions and a pH less than 7 indicates an acidic solution.

In the context of the present specification, the term *reducing agent* is used in its meaning known in the art of chemistry and biochemistry; it refers to an element or compound that donates an electron to another chemical substance in a redox reaction. In this process the reducing agent is oxidized.

In the context of the present specification, the terms *fluorescent cell sorting* or *fluorescence-activated cell sorting (FACS)* are used in their meaning known in the art of cell biology; they refer to a method for cell sorting, wherein cells are suspended in a stream of fluid and pass a detection device. Every passing cell can be directed into different compartments according to the intensity of a specific fluorescent signal of the passing cell.

In the context of the present specification, the term *random mutagenesis* is used in its meaning known in the art of cell biology and molecular biology; it refers to a method wherein DNA mutations are randomly introduced to produce mutant genes and proteins. A multitude of these mutant genes can then be compiled into a library. Non-limiting examples for random mutagenesis methods are error-prone PCR, UV radiation and chemical mutagens.

In the context of the present specification, the term *library* is used in its meaning known in the art of cell biology and molecular biology; it refers to a collection of nucleic acid fragments. One particular type of library is a randomized mutant library, generated by random mutagenesis. Another example would be a designed (or synthetic) library, which comprises specifically engineered DNA fragments.

In the context of the present specification, the term *expression level* is used in its meaning known in the art of cell biology and molecular biology; it refers to the level of transcription and/or translation of a DNA fragment and the derived mRNA, respectively. In certain embodiments an expression level is deemed to be *high* if the expression of a mutant gene is higher than a control gene or wild-type gene. In certain embodiments the expression of the mutant gene is at least two-fold higher than that of the control or wild-type gene to be deemed *high*.

According to a first aspect of the invention a method is provided for selecting a sequence from a library of expressed nucleic acid sequences, wherein the sequence is selected according to its expression level. The method comprises the following steps.

- a) A plurality of eukaryotic cells comprising a cell wall, particularly a plurality of yeast cells, is provided, wherein each of the eukaryotic cells comprises a nucleic acid sequence member of the library. The nucleic acid sequence member is a transgene to the cell, expressed under the control of a promoter sequence operable in the cell, as a target membrane protein in the plurality of eukaryotic cells comprising a cell wall.
- b) The cell wall of the plurality of eukaryotic cells comprising a cell wall is permeabilized in a permeabilization step, yielding a plurality of viable permeabilized cells.
- c) The plurality of viable permeabilized cells is contacted with a ligand specifically capable of binding to the target membrane protein in a labeling step. The ligand comprises a detectable label, yielding a plurality of viable permeabilized cells.
- d) The plurality of viable permeabilized cells is washed in a washing step, thereby removing most of, if not essentially all of, any ligand and detectable label not having bound specifically to the target membrane protein.
- e) The presence of the detectable label for each of the plurality of viable labelled cells is detected and a subset of the plurality of viable labelled cells is selected as a function of detectable label present in the plurality of viable labelled cells in a selection step. In other words, the cells that show any label, or a label quantity above a certain threshold is selected, yielding a selection of viable cells.
- f) The expressed nucleic acid sequence from the selection of cells is isolated in an isolation step.

In other words, the method according to the first aspect of the invention allows the expression of a library of transmembrane receptors in eukaryotic cells comprising a cell wall and allows for the selection of functional expression of these receptors. This is made possible by the permeabilization of the cell wall in a way that still maintains viable and structurally stable cells, which is different from other methods of permeabilization of the cell wall such as the preparation of spheroplasts. The permeabilization procedure of this invention allows even the binding of large ligands to the transmembrane receptors. Selection of the cells by the amount of bound ligand enables selection of the cells according to the amount of functional receptor on the plasma membrane and not by total receptor amount, which would also include non-functional receptors (e.g. receptor present in intracellular membranes).

In certain embodiments the method additionally comprises the following steps:

- i. after the selection step e) the selection of viable cells is expanded in an expansion step, yielding an expanded selection of viable cells. The expanded selection of viable cells is subjected to the steps b) to e) in this sequential order, and

- ii. step i. is performed at least 1, 2, 3, 4, 5, 6 or 7 times, finally followed by the isolation step f).

In certain embodiments the selection step comprises a multitude of selection procedures. After a first selection the selection of cells is used again for the selection of a subset of these cells as a function of detectable label present in these cells. Selection of cells is repeated at least 1, 2, 3, 4 or 5 times.

In certain embodiments the library of expressed nucleic acid sequences is obtained by amplification of a nucleic acid sequence encoding the target membrane protein by a process introducing mutations into the amplified sequence.

In certain embodiments the method additionally comprises the following steps:

- i. the expressed nucleic acid obtained in isolation step f) is introduced and expressed in a plurality of eukaryotic cells comprising a cell wall,
- ii. the plurality of eukaryotic cells comprising a cell wall is subjected, in this sequential order, to steps b) to f) according to the first aspect of the invention, and
- iii. steps i. and ii. are performed at least 1, 2, 3, 4, 5, 6 or 7 times.

In certain embodiments the expressed nucleic acid sequences obtained by the method of the invention are characterized by high expression levels and/or high thermodynamic stability of the encoded target membrane protein.

In certain embodiments according to all aspects of the invention in the selection step e), the selection of viable cells comprise the top 0.1% to 5% of the most fluorescent cells.

In certain embodiments the method additionally comprises the following steps:

- i. The expressed nucleic acid obtained in the isolation step f) is amplified by a process introducing mutations into the amplified sequence, yielding a second library of nucleic acid sequences.
- ii. This second library of nucleic acid sequences is transferred to the plurality of eukaryotic cells comprising a cell wall.
- iii. The plurality of eukaryotic cells comprising a cell wall, now comprising a nucleic acid sequence member of the second library, is submitted to the steps b) to f), in this sequential order, of the method according to the first aspect of this invention.

In certain embodiments the eukaryotic cells comprising a cell wall are yeast cells.

In certain embodiments the target membrane protein is a G-protein coupled receptor (GPCR).

In certain embodiments the permeabilization step comprises exposing the plurality of eukaryotic cells comprising a cell wall to an enzymatic and/or a chemical treatment.

In certain embodiments the enzymatic treatment in the permeabilization step comprises exposing the plurality of eukaryotic cells comprising a cell wall to enzymes or enzyme mixtures that permeabilize the cell wall. Non-limiting examples of such enzymes are glucanases, proteases, mannanases and/or sulfatases. Non-limiting examples of such enzyme mixtures are Zymolyase, Lyticase, and/or Glusulase.

In certain embodiments the chemical treatment in the permeabilization step comprises exposing the plurality of eukaryotic cells comprising a cell wall to a buffer of alkaline pH comprising lithium ions, a reducing agent and/or chelating agent.

In certain embodiments non-limiting examples of buffering agents contained in the buffer used in the permeabilization step are:

- Bicine (2-(Bis(2-hydroxyethyl)amino)acetic acid) or
- HEPES (2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethanesulfonic acid) or
- MOPS (3-morpholinopropane-1-sulfonic acid) or
- PIPES (1,4-Piperazinediethanesulfonic acid) or
- TAPS (3-[[1,3-Dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]propane-1-sulfonic acid) or
- TAPSO (3-[[1,3-Dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]-2-hydroxypropane-1-sulfonic acid) or
- TES (2-[[1,3-Dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid) or
- Tricine (*N*-(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine) or
- Tris (2-Amino-2-hydroxymethyl-propane-1,3-diol).

In certain embodiments the chelating agent used in the buffer of the permeabilization step is ethylenediaminetetraacetic acid (EDTA).

In certain embodiments the detectable label is a fluorescent dye and the selection step is accomplished by fluorescent cell sorting.

In certain embodiments non-limiting examples of reducing agent are:

- thiol-containing compounds, such as, dithiothreitol (DTT), dithioerythritol (DTE), mercaptoethanol or reduced glutathione or
- phosphine-containing compounds, such as tris-carboxyethyl-phosphine (TCEP).

In certain embodiments the permeabilization step comprises the following steps:

- a) incubating the yeast cells in TELi buffer comprising 50 mM Tris-HCl pH 9, 1 mM EDTA and 100 mM lithium acetate,
- b) incubating the yeast cells in TELi buffer additionally comprising 50 mM DTT for 30 min at 20°C,
- c) washing the yeast cells at least once in TELi buffer at 4°C.

In certain embodiments TELi buffer comprises 50 mM Tris-HCl pH 9, 1 mM EDTA and 100 mM lithium acetate.

In certain embodiments the buffer used in the permeabilization step comprises:

- i. lithium ions,
- ii. alkaline pH,
- iii. reducing agent, and/or
- iv. chelating agent.

In certain embodiments the labeling step comprises exposing the yeast cells to TELi buffer comprising the ligand at 4°C.

In certain embodiments the nucleic acid sequences for the library of expressed nucleic acid sequences are generated by random mutagenesis preferably by error-prone PCR.

In certain embodiments the library of expressed nucleic acid sequences are designed (synthetic) libraries.

In certain embodiments the library of expressed nucleic acid sequences comprises homologous sequences of at least 60% sequence identity with each other.

In certain embodiments the eukaryotic cell comprising a cell wall is a yeast cell, particularly *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*, *Candida boidinii*, or *Hansenula polymorpha*.

In certain embodiments the eukaryotic cells comprising a cell wall are a mutant or genetically engineered yeast strain not comprising a native cell wall.

In certain embodiments the eukaryotic cells comprising a cell wall is *Saccharomyces cerevisiae*, particularly the *S. cerevisiae* strain BY4741.

In certain embodiments the ligand specifically capable of binding to the target membrane protein is an agonist, antagonist or allosteric modulator.

In certain embodiments the ligand specifically capable of binding to the target membrane protein is an oligopeptide comprised of at least 3, 4, 5, 6, 8, 10, 14, 18 or 25 amino acids.

In other embodiments, the ligand specifically capable of binding to the receptor is an antibody, antibody fragment or another binding protein, e.g., a scaffold protein from the non-limiting list of examples of DARPin, affibodies, anticalins, nanobodies, affilins, fibronectin-derived scaffolds and other scaffolds.

According to an alternative to this first aspect of the invention a method is provided for selecting a sequence from a library of expressed nucleic acid sequences, wherein the

sequence is selected according to its expression level. The method comprises the following steps.

- a) A plurality of eukaryotic cells comprising a cell wall, particularly a plurality of yeast cells, is provided, wherein each of the eukaryotic cells comprises a nucleic acid sequence member of the library. The nucleic acid sequence member is a transgene to the cell, expressed under the control of a promoter sequence operable in the cell, as a target membrane protein in the plurality of eukaryotic cells comprising a cell wall.
- b) The plurality of eukaryotic cells comprising a cell wall is contacted in a labelling step with a ligand specifically capable of binding to the target membrane protein. The ligand comprises a detectable label, yielding a plurality of labelled cells.
- c) The plurality of labelled cells is washed in a washing step, thereby removing most of or any ligand and detectable label not having bound specifically to the target membrane protein.
- d) The presence of said detectable label for each of the plurality of labelled cells is detected and a subset of the plurality of labelled cells is selected as a function of detectable label present in the plurality of labelled cells in a selection step. In other words, the cells that show any label, or a label quantity above a certain threshold is selected, yielding a selection of viable cells.
- e) The expressed nucleic acid sequence from the selection of cells is isolated in an isolation step.

In other words, the method according to this alternative aspect of the first aspect of the invention allows the expression of a library of transmembrane receptors in eukaryotic cells comprising a cell wall and allows for the selection of functional expression of these receptors. Selection of the cells by the amount of bound ligand enables selection of the cells according to the amount of functional receptor on the plasma membrane and not by total receptor amount, which would also include non-functional receptors (e.g. receptor present in intracellular membranes). This alternative aspect of the first aspect of the invention might be advantageous for the use of small ligands.

According to a second aspect of the invention a method for the selection of an adapted yeast cell with the ability for high expression levels of functional membrane proteins is provided. The method comprises the following steps:

- a. A plurality of yeast cells is provided, wherein each of the yeast cells comprises a nucleic acid sequence member of a library of expressed nucleic acid sequences. The nucleic acid sequence member is expressed as a target membrane protein in the plurality of yeast cells.

- b. The cell wall of the plurality of yeast cells is permeabilized in a permeabilization step. This yields a plurality of viable permeabilized cells.
- c. The plurality of viable permeabilized cells is contacted in a labelling step with a ligand capable of binding to the target membrane protein. The ligand comprises a detectable label and thereby yields a plurality of viable labelled cells.
- d. The plurality of viable labelled cells is washed in a washing step.
- e. A subset of the plurality of viable labelled cells is selected as a function of detectable label present in the plurality of viable labelled cells in a selection step, yielding a selection of viable cells. In other words, the cells that show any label, or a label quantity above a certain threshold is selected, yielding a selection of viable cells.
- f. The selection of viable cells is expanded in an expansion step, yielding an expanded selection of viable cells.
- g. The expanded selection of viable cells is submitted to steps b. to f. at least 1, 2, 3, 4, 5, 6 or 7 times.
- h. The expanded selection of viable cells is submitted to steps b. to e.
- i. A subset of the selection of viable cells is selected as a function of detectable label present in the plurality of viable labelled cells. This yields the adapted yeast cell with the ability for high expression levels of functional membrane proteins from the expanded selection of viable cells.

Wherever alternatives for single separable features such as, for example, a cell strain or permeabilization buffers, sorting method or library type are laid out herein as "embodiments", it is to be understood that such alternatives may be combined freely to form discrete embodiments of the invention disclosed herein.

In certain embodiments, the present invention further provides the following items:

1. A method for selecting a sequence from a library of expressed nucleic acid sequences, wherein the sequence is selected according to its expression level, the method comprises the steps of
 - a) providing a plurality of eukaryotic cells comprising a cell wall, wherein each of said eukaryotic cells comprises a nucleic acid sequence member of said library, and said nucleic acid sequence member is expressed as a G protein-coupled receptor (GPCR) in the plasma membrane in said plurality of eukaryotic cells,
 - b) permeabilizing said cell wall of said plurality of eukaryotic cells in a permeabilization step, wherein the permeabilization step comprises exposing the plurality of eukaryotic cells to a non-enzymatic chemical treatment, yielding a plurality of viable permeabilized cells,

- c) contacting said plurality of viable permeabilized cells in a labelling step with a ligand capable of binding to said GPCR, wherein said ligand comprises a detectable label, yielding a plurality of labelled cells,
- d) washing said plurality of labelled cells in a washing step,
- e) selecting a subset of said plurality of labelled cells as a function of detectable label present in said plurality of labelled cells in a selection step, yielding a selection of cells, and
- f) isolating an expressed nucleic acid sequence from said selection of cells in an isolation step,

wherein the expressed nucleic acid sequence obtained in step f) is characterized by high expression levels of the encoded GPCR;

and wherein said detectable label is a fluorescent dye and said selection step is accomplished by fluorescent cell sorting.

2. The method according to item 1, wherein:

- i. after said selection step e) said selection of viable cells is expanded in an expansion step, yielding an expanded selection of viable cells and said expanded selection of viable cells is subjected to said steps b) to e),
- ii. said step i. is performed at least 1, 2, 3, 4, 5, 6 or 7 times, finally followed by said isolation step f).

3. The method according to item 1, wherein

- i. the expressed nucleic acid obtained in said isolation step f) is introduced and expressed in a plurality of eukaryotic cells comprising a cell wall,
- ii. said plurality of eukaryotic cells comprising a cell wall is subjected to said steps b) to f) according to item 1, and
- iii. steps i. and ii. are performed at least 1, 2, 3, 4, 5, 6 or 7 times.

4. The method according to any one of items 1 to 3, wherein said expressed nucleic acid obtained in said isolation step f) is amplified by a process introducing mutations into the amplified sequence, yielding a second library of nucleic acid sequences and

- i. said second library of nucleic acid sequences is transferred to said plurality of eukaryotic cells comprising a cell wall, and
- ii. said plurality of eukaryotic cells comprising a cell wall is submitted to the method according to any one of items 1 to 3.

5. The method according to any one of items 1 to 4, wherein the library of expressed nucleic acid sequences is obtained by amplification of a nucleic acid sequence encoding said GPCR by a process introducing mutations into the amplified sequence.

6. The method according to any one of items 1 to 5, wherein said chemical treatment comprises a buffer of alkaline pH comprising lithium ions, a reducing agent and/or a chelating agent.
7. The method according to any one of items 1 to 6, wherein said plurality of eukaryotic cells is a plurality of yeast cells.
8. A method for the selection of an adapted yeast cell with the ability for high expression levels of functional G protein-coupled receptors (GPCRs), comprising the steps of
 - a. providing a plurality of yeast cells, wherein each of said yeast cells comprises a nucleic acid sequence member of a library of expressed nucleic acid sequences, and said nucleic acid sequence member is expressed as a GPCR in the plasma membrane in said plurality of yeast cells,
 - b. permeabilizing the cell wall of said plurality of yeast cells in a permeabilization step, wherein the permeabilization step comprises exposing the plurality of eukaryotic cells to a non-enzymatic chemical treatment, yielding a plurality of viable permeabilized cells,
 - c. contacting said plurality of viable permeabilized cells in a labelling step with a ligand capable of binding to said GPCR, wherein said ligand comprises a detectable label, yielding a plurality of labelled cells,
 - d. washing said plurality of labelled cells in a washing step,
 - e. selecting a subset of said plurality of labelled cells as a function of detectable label present in said plurality of labelled cells in a selection step, yielding a selection of cells,
 - f. expanding said selection of cells in an expansion step, yielding an expanded selection of cells
 - g. submitting said expanded selection of cells to steps b. to f. at least 1, 2, 3, 4, 5, 6 or 7 times,
 - h. submitting said expanded selection of cells to steps b. to e., and
 - i. selecting a subset of said expanded selection of cells as a function of detectable label present in said plurality of labelled cells, yielding said adapted yeast cell with the ability for high expression levels of functional GPCRs from said expanded selection of cells;

wherein said detectable label is a fluorescent dye and said selection step is accomplished by fluorescent cell sorting.

The invention is further illustrated by the following examples and figures, from which further embodiments and advantages can be drawn. These examples are meant to illustrate the invention but not to limit its scope.

Short description of the figures

Fig. 1 shows the workflow for directed evolution of GPCRs in yeast. As a first step, the wild-type GPCR gene is randomized by error-prone PCR in order to create a DNA library. The DNA library is then combined with linearized yeast expression vector and the mixture is used for transformation of yeast. Insert DNA and vector backbone are assembled in vivo by homologous recombination. The obtained yeast library is cultivated and expression is induced. After expression, cells are permeabilized and incubated with fluorescent ligand, which binds exclusively to functional GPCRs in the plasma membrane. Subsequently, unbound ligand is removed by washing and cells are subjected to selection by FACS. Selection of cells exhibiting high fluorescence, correspondingly expressing GPCRs at high

functional level, allows isolation of desired high-expression GPCR variants. During FACS, yeast cells are directly sorted into growth medium for subsequent propagation. Selection by FACS is repetitively performed in order to obtain a strong enrichment of cells harboring the best expressing GPCRs. Whenever desired, plasmid DNA can be isolated from selected cells for analysis of individual variants or additional diversity can be introduced by random mutagenesis for another round of evolution.

Fig. 2 shows (A), (C), (E) histogram plots of ligand-binding FC data of NTR1, NK1R, and KOR1 variants. Functional expression in yeast of wild-type GPCRs (left panels), library pools obtained after the second round of evolution (middle panel) and variants evolved in the *E. coli*-based system (right panels). Total signal was obtained by binding of fluorescent ligand to receptors in yeast cells expressing the corresponding GPCR variants (black curves). Nonspecific binding was measured in the presence of excess unlabeled ligand (gray, tinted). For the wild-type GPCRs no specific signal is detected due to the low functional expression levels at the surface. The selected library pools show high specific signals (functional receptors at the surface per cell) in expressing cells with a small fraction of cells not expressing any functional receptor at the surface (note double peak of total signal). Variants previously evolved in the *E. coli*-based system (NTR1-D03, NK1R-E11) show a specific signal, however not reaching the levels obtained for the yeast library pools. Furthermore, for a significantly higher fraction of cells no functional GPCR expression is detected at the surface. For instance, only about 50% of all cells measured show a specific signal for NTR1-D03 expression. (B), (D), (F) Measurement of average total functional receptors per cell by RLBA of NTR1, NK1R, and KOR1 variants. Expression levels were quantified with [³H]-radioligands. Non-specific signal was measured in the presence of [³H]-radioligands with excess of unlabeled ligand and was subtracted from the total signal. Wild-type NTR1 and KOR1 show low expression levels, while for NK1R no functional expression at all is detected. NTR1-D03 shows moderate receptor-per-cell levels, whereas for NK1R-E11 expression levels are low. The selected library pools show functional production levels of 100,000 – 150,000 receptors per cell, an increase of 25 – 50-fold and 5 – 20-fold compared to the wild-type GPCRs and the variants evolved in *E. coli*, respectively. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 3 shows functional expression of selected NTR1 variants (NTR1-Y01 – NTR1-Y07) in non-adapted yeast strains. (A) Histogram plots of ligand-binding FC data with total signal (black curves) and nonspecific signal (gray, tinted) are shown. Compared to wild-type NTR1 all selected variants show an increase in functional surface expression (cf. Fig. 2 A). In comparison with the library pool NTR1 2.5, the individual variants depict an increase of the subpopulation not expressing functional receptor at the surface as well as lower specific signals (receptors per cell) in cells with surface expression of active receptor (cf. Fig. 2 A).

(B) Measurement of average total functional receptors per cell by RLBA. All variants show an increase in functional expression compared to wild-type NTR1, but do not reach the expression levels observed for the NTR1 2.5 pool (cf. Fig. 2 B). Non-specific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 4 shows quantitative Western blot analysis of HA-tagged NTR1 variants expressed in non-adapted and adapted yeast strains. Whole cell lysate protein extraction was performed with equal numbers of cells for each sample after expression. As a loading control, actin was used. GPCRs were detected via their HA tag, with main bands corresponding to monomeric GPCRs running slightly below the expected molecular size (44 kDa) and bands of higher molecular weight, most likely representing GPCR dimers not disintegrated under the conditions used. For quantification (bar chart) intensities of all defined bands were accounted for and signals were normalized to the GPCR and actin intensities obtained for wild-type NTR1. In the non-adapted strains, the total GPCR produced increases from wild-type NTR1 (lane 1) to the evolved variant (lane 2) by a factor of two. While total receptor levels of NTR1-Y06 also slightly increase in the adapted strain (lane 3) compared to the non-adapted strain (lane 2), the relative increase of total receptor produced is much lower than the increase in functional receptor (cf. Fig 13 B). For a negative control (lane 4), cells expressing NTR1-Y06 without HA tag were used.

Fig. 5 shows confocal fluorescence microscopy studies of NTR1 variants with C-terminal fusion to mCh expressed in non-adapted and adapted yeast strains. Fluorescence intensities obtained by binding of fluorescent ligand (top row) or from mCh (middle row) as well as bright-field microscopy overlays (bottom row) are shown. For expression of wild-type NTR1 (first column), no cells with a ligand-binding signal at the cell surface are detected, while some cells show distinct mCh signals, mainly located in the cell interior, and thus reflecting intracellularly retained receptors. Additionally, many cells lacking any fluorescent signal are detected, representing a non-expressing subpopulation. In contrast, expression of NTR1-Y06 in non-adapted cells (second column) allows visualization of functional receptor at the cell surface by ligand binding. These cells show also strong signals for mCh, however still to a large extent localized in the cell interior. Similar to wild-type NTR1, NTR1-Y06 expression in the non-adapted strain gives rise to non-expressing cells, depicting neither a signal for ligand binding nor for mCh. For expression of NTR1-Y06 in the adapted strain (third column), significantly fewer non-expressing cells are detected and expressing cells show at the surface a strong signal for both ligand-binding and mCh, with only little mCh detected in the cell interior. For a negative control (fourth column) cells expressing NTR1-Y06 without a mCh fusion were incubated with fluorescently labeled neurotensin in excess of non-labeled neurotensin. Representative pictures are shown.

Fig. 6 shows FC and RLBA analysis of non-adapted and adapted strains expressing NTR1 variants with a C-terminal mCh fusion. (A) FC data of non-adapted strains expressing NTR1 (left panel), NTR1-Y06 (middle panel), and the adapted strain expressing NTR1-Y06 (right panel) are shown. In the histogram plots of ligand-binding experiments (top row), the surface expression of active receptor variants are compared with total signal (black curves) and nonspecific signal (gray, tinted) shown. NTR1 shows only little active receptor at the surface, whereas surface expression is significantly increased for NTR1-Y06 in the non-adapted and the adapted strain (shift of the specific signal towards higher fluorescence intensity). Compared to the non-adapted strain expressing NTR1-Y06 with a mCh fusion, adaptation further leads to a substantial decrease of the fraction of cells showing no surface expression of active receptor. In the histograms of mCh expression (middle row) total receptor produced is quantified. Black curves depict the mCh signal, whereas autofluorescence of cells expressing NTR1-Y06 without mCh-fusion incubated with fluorescent ligand represents the background (gray, tinted). NTR1 and NTR1-Y06 expressions in non-adapted strains have very similar mCh signal profiles. Since for expression of NTR1 in ligand-binding FC experiments only little active receptor at the surface is detected, the strong mCh signal means that most of NTR1 must be intracellularly retained. A subpopulation of cells not expressing any receptor at all is seen for all variants (note double peak), but this non-expressing fraction is significantly decreased in the adapted strain. In correlation analysis (bottom row) mCh fluorescence intensity (*total* receptor produced) is compared to ligand-binding fluorescence intensity (*functional* receptor at the surface). For expression of wild-type NTR1, functional receptor at the surface does not correlate well with total receptor produced. This correlation is better for NTR1-Y06 expressed in the non-adapted strain, and if NTR1-Y06 is expressed in the adapted strain, functional receptor at the surface correlates well with total receptor produced. (B) Measurement of average total functional receptors per cell by RLBA. Functional expression levels of receptors with a mCh-fusion increase from wild-type receptor to the evolved variant by a factor of three in the non-adapted strain. Compared to expression of NTR1-Y06 in the non-adapted strain, adaptation leads to a further increase in average total functional expression by a factor of 6. Non-specific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 7 shows functional expression levels of evolved GPCRs in Sf9 insect cells and signaling activity of NK1R variants. (A), (B), (C) Measurement of average total functional receptors per cell by RLBA of NTR1, NK1R, and KOR1 variants. Compared to wild-type GPCRs all evolved variants show a significant increase of average receptor-per-cell levels. NTR1-Y06 shows a fivefold increase compared to wild-type NTR1, NK1R-Y09 shows a fourfold and twofold increase compared to NK1R and NK1R-ΔC, respectively, and the

strongest increase (27-fold) is detected for KOR1-Y05 compared to wild-type KOR1. For each GPCR two independent expression experiments were performed (separate bars). Nonspecific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates. (D) Measurement of signaling activity of NK1 R variants by [³⁵S]-GTPγS binding. Equal amounts of active GPCR and reconstituted G protein were assayed in the presence (grey) and absence (black) of substance P. All variants show a low basal activity without agonist stimulation. Upon addition of agonist, signaling is detected by [³⁵S]-GTPγS binding. NK1R and NK1R-AC have identical signaling activities, and signaling of NK1R-Y09 remains similar to wild-type receptor. For each GPCR variant two independent signaling assays from two independent expression experiments were performed (separate bars). Error bars indicate standard deviations from triplicates.

Fig. 8: Scheme of GPCR topology depicting the different regions harboring mutations (see Table 1).

Fig. 9 shows functional expression of selected NK1R and KOR1 variants in non-adapted yeast strains. (A), (B) Histogram plots of ligand binding FC data of NK1R and KOR1 variants, respectively, with total signal (black curves) and nonspecific signal (gray, tinted) shown. Nonspecific signal was obtained in the presence of an excess of unlabeled ligand. All variants show an increase in functional expression compared to the corresponding wild-type GPCRs (NK1R and KOR1, cf. Fig. 2 C, E). While the specific signal of the expressing cells of most variants is similar or slightly lower compared to the corresponding library pools (NK1R 2.5 and KOR1 2.5, cf. Fig. 2 C, E), a higher fraction of cells not expressing any functional receptor at the surface is observed for the individual variants. (C), (D) Measurement of average total functional receptors per cell by RLBA. While all variants show an increase in functional expression compared to the corresponding wild-type GPCRs, for most variants the average numbers of receptors per cell measured by RLBA are lower than for the corresponding library pools (cf. Fig. 2 D, F). Non-specific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 10 shows functional expression levels of NTR1-Y06 in a yeast strain directly isolated from the NTR1 2.5 pool and a newly transformed strain subsequently adapted by phenotypic selection with FACS. (A), (B) Histogram plots of ligand-binding FC data with total signal (black curves) and nonspecific signal (gray, tinted) are shown. Non-specific signal was obtained in the presence of an excess of unlabeled ligand. The isolated strain shows a

similar expression profile as the NTR1 2.5 library pool (cf. Fig. 2 A), in contrast to the non-adapted retransformed strain expressing NTR1-Y06 (cf. Fig. 3 A). In the subsequently adapted strain shown in (B), obtained after five rounds of phenotypic selection by FACS with the newly transformed strain, high surface expression levels are reconstituted. (C) Measurement of average total functional receptors per cell by RLBA. Compared to the retransformed strain expressing NTR1-Y06 (cf. Fig. 3 B), high functional NTR1-Y06 expression in the isolated strain as well as reconstitution of this phenotype in the adapted strain is detected. Non-specific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 11 shows expression profiles of NTR1-Y06 after individual sorts during phenotypic selection for host adaptation. Histogram plots of ligand binding FC data with total signal (black curves) and nonspecific signal (gray, tinted) are shown. Non-specific signal was obtained in the presence of an excess of unlabeled ligand. In the course of the phenotypic selection in which the cells with the highest NTR1-Y06 expression were sorted (gating of the top 1% of the most fluorescent cells), the population of cells not expressing any functional receptor at the surface gradually decreases. At the same time, the specific signal of the expressing cells shifts to higher levels during the selection, indicating an increase of functional receptors per cell at the surface of these cells. Note that two sorts are required to induce the adaptation. While the expression profile is not significantly changed after the first two sorts, a clear decrease of the subpopulation with no surface expression of active NTR1-Y06 as well as a shift of the specific signal within the expressing subpopulation is observed after sort 3. This trend is continued in subsequent sorts, as seen in the expression profile after sort 4.

Fig. 12 shows functional expression of NTR1-Y06 in non-adapted, adapted, and adapted strain previously cultivated under non-expressing conditions in *S. cerevisiae*. (A) Histogram plots of ligand binding FC data with total signal (black curves) and nonspecific signal (gray, tinted) are shown. Non-specific signal was obtained in the presence of an excess of unlabeled ligand. Adaptation of newly transformed strain expressing NTR1-Y06 leads to a higher specific signal (receptors per cell) of the expressing cell subpopulation and a decrease of the subpopulation not showing any surface expression of active receptor. If the adapted strain is repetitively cultivated under non-expressing conditions, the average expression level decreases again, depicted by a drop of the specific signal (receptors per cell) in the expressing subpopulation and an increase of the fraction of cells with no surface expression of active NTR1-Y06. (B) Measurement of average total functional receptors per cell by RLBA. Results from RLBA show the significant increase of the number of average receptors per cell from the non-adapted to the adapted strain, which in turn drops again by about 30% if the adapted strain is repetitively cultivated under non-expressing conditions

prior to inducing expression. The lack of complete reversion to the original lower expression level of the non-adapted strain may be attributed to leaky expression, sufficient to partially retain the high-expression phenotype even under non-inducing conditions. Non-specific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 13 shows functional expression of NK1R-Y09 and KOR1-Y05 in non-adapted and adapted *S. cerevisiae* strain. (A), (C) Histogram plots of ligand binding FC data with total signal (red curves) and nonspecific signal (green, tinted) are shown. Non-specific signal was obtained in the presence of an excess of unlabeled ligand. While the specific expression signals (receptors per cell) of expressing cells in the adapted strains slightly increase compared to the non-adapted strains, a significant decrease of the subpopulation with no surface expression of active receptor is observed. (B), (D) Measurement of average total functional receptors per cell by RLBA. The decrease of the fraction of cells which show no surface expression of active receptor in the adapted strains, shown in the FC data, leads to an increase of the average functional expression measured with RLBA. Non-specific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 14 shows functional expression of HA-tagged NTR1-Y06 in non-adapted and adapted *S. cerevisiae* strain. (A) Histogram plots of ligand-binding FC data with total signal (black curves) and nonspecific signal (gray, tinted) are shown. Non-specific signal was obtained in the presence of an excess of unlabeled ligand. Adaptation of the strain expressing NTR1-Y06 with a C-terminal HA-tag leads to a decrease of the fraction of cells showing no surface expression of active receptor and an increase of active receptors per cell at the surface of cells with surface expression (shift of specific signal towards higher fluorescence intensity). (B) Measurement of average total functional receptors per cell by RLBA. Functional expression levels increase from wild-type receptor to the evolved variant by a factor of four in the non-adapted strain. Compared to expression of HA-tagged NTR1-Y06 in the non-adapted strain, adaptation leads to a further increase in average total functional expression by a factor of 10. Non-specific signal was subtracted from total signal. Error bars indicate standard deviations from triplicates of a representative expression experiment.

Fig. 15 shows purification of NK1R-Y09. (A) SEC profile of purified receptor after IMAC. Two peaks are obtained of which peak 1 most likely corresponds to defined higher oligomeric states of NK1R-Y09, while peak 2 represents the monomeric receptor fraction. (B) Analysis of purified NK1R-Y09 by SDS-PAGE under reducing conditions. Equal amounts of total protein were loaded in each lane. After IMAC (lane 1) pure NK1R-Y09 (38.5 kDa) is obtained, with detection of weak additional protein bands at higher molecular weight. Such bands represent most likely oligomers of NK1R-Y09, which are not disintegrated under

conditions used. Next to a strong band for monomeric NK1 R-Y09, the bands of higher molecular weight are also detected in the fraction of peak 1 (lane 2), while in the fraction of peak 2 (lane 3) monomeric NK1 R-Y09 is the sole species. Note that NK1 R-Y09 was not further engineered and expressed in insect cells as it was obtained from the selection in yeast. For purifications with the aim to perform crystallization trials, some further engineering will be required, for instance removal of potential glycosylation sites, N-terminal truncations, loop deletions, and/or introduction of fusion proteins (e.g. T4 lysozyme or thermostabilized apocytochrome b₅₆₂RIL (Chun et al. (2012), Structure 20:967-976)). Such measures potentially further improve purification of the receptor by reducing heterogeneity, restricting conformational flexibility, and increasing stability (Maeda & Schertler (2013), Curr Opin Struct Biol 23:381-392).

Examples

Example 1: Generation of high-expressing functional GPCRs by directed evolution in yeast

With about 800 different members in the human genome, G protein-coupled receptors (GPCRs) comprise the largest superfamily of cell surface receptors. GPCRs evolved to highly versatile signaling mediators in eukaryotic life, responding to a wide variety of ligands and transducing signals via heterotrimeric G proteins as well as in a G protein-independent fashion. The pivotal role of GPCRs is reflected by the great number of human diseases linked to aberrant GPCR signaling, for instance obesity, diabetes, cardiovascular diseases, osteoporosis, immunological disorders, neurodegenerative diseases, and cancer. As a consequence, GPCRs represent highly relevant drug targets for the pharmaceutical industry. About 30 - 50% of marketed drugs act on GPCRs or GPCR-associated mechanisms, among them many top-selling drugs.

The major challenges in structural and detailed biochemical investigation of GPCRs are the difficulties to produce the desired receptors in sufficient amounts as well as their inherent instability and flexibility. Advances in expression and crystallization strategies, like the use of the baculovirus/insect cell expression system, or the establishment of integral membrane protein crystallization in lipidic cubic phases (LCP), lead to breakthroughs in determination of three-dimensional receptor structures by X-ray crystallography. While the available structural datasets provided the scientific community with insights towards the mechanisms of GPCR function at the atomic level and might begin to enable rational structure-based drug design, it is clear that detailed understanding of receptor dynamics and conformational changes has still remained rather incomplete.

Furthermore, the 26 unique GPCR structures deposited in the Protein Data Bank to date still represent only a minor fraction of all receptors (< 4%),

reflecting that the bottlenecks for structural investigations of GPCRs persist. For instance, while insect cells were successfully used for production of recombinant receptors for about 85% of all GPCR structures determined so far (source: PDB; excluding structures obtained from protein extracted from native tissues), this expression system does not provide a generic solution. Even in insect cells several members of the GPCR superfamily are expressed at low yields, suggesting fundamental issues with the biosynthesis and membrane insertion of these receptors also in eukaryotic cells. This problem cannot be solved by simple optimization of expression conditions, especially since there seems to be no consistency of optimal conditions for individual GPCRs. Thus, to facilitate and accelerate GPCR research, novel approaches are required.

Recently, the inventors developed a method to increase functional expression levels of GPCRs in *Escherichia coli* by directed evolution (Sarkar et al. (2008), Proc Natl Acad Sci USA 105:14808–14813; Dodevski & Plückthun (2011), J Mol Biol 408:599–615). Random mutagenesis of wild-type GPCRs followed by selection with fluorescence-activated cell sorting (FACS) allowed isolation of high-expression variants from randomized libraries of four different GPCRs, namely neurotensin receptor 1 (NTR1), NK-1 receptor (NK1R, also termed tachykinin receptor 1 or substance-P receptor) and of the alpha-1A and alpha-1B adrenergic receptors. In successive studies for NTR1, the obtained first-generation variants built the basis for creation of second-generation mutants by extensive selection (Schlinkmann et al. (2012), Proc Natl Acad Sci USA 109:9810–9815) and exhaustive recombination (Schlinkmann et al. (2012), J Mol Biol 422:414–428). The synthetic libraries created in these studies were also used to generate NTR1 variants with improved stability in short-chain detergents by a complementary directed evolution approach termed CHESS (Scott & Plückthun (2013), J Mol Biol 425:662–677; Scott et al. (2014), Biochim Biophys Acta 1838:2817–2824). Ultimately, these efforts resulted in the structures of three different variants of agonist-bound NTR1 from material produced in *E. coli* (Egloff et al. (2014), Proc Natl Acad Sci USA 111:E655–62), underlining the power of directed evolution in membrane protein engineering.

The successful results obtained with evolution in *E. coli* and the proven effectiveness of eukaryotic expression hosts demanded to further develop this approach towards high functional GPCR expression specifically in eukaryotic expression systems. The inventors hypothesized that it might be advantageous to perform evolution of GPCRs directly in a eukaryotic system in order to achieve specific sequence adaptation and thereby improved functional production in eukaryotes. Furthermore, compared to eukaryotic hosts, *E. coli* is not able to perform post-translational modifications, lacks the secretory quality control and translocation machinery for eukaryotic membrane proteins, and differs in membrane composition. Therefore, directed evolution of GPCRs in *E. coli* is limited to receptors which

inherently express in a prokaryotic system above the threshold required for successful evolution, thus not having strict requirements for eukaryotic processing or membrane composition.

Here the inventors disclose the establishment of a fast, efficient and robust method to increase functional expression of GPCRs in eukaryotic hosts by directed evolution in the yeast *Saccharomyces cerevisiae*. The method is generally applicable as demonstrated with three different GPCRs, leading, with only two rounds of evolution, to receptor variants which show high functional expression in both yeast and insect cells. In addition, functional expression levels in yeast can be further increased reproducibly by induced host adaptation.

Directed evolution of three different GPCRs in yeast results in high functional expression in S. cerevisiae

The general approach of this method is depicted in Fig. 1. First, the wild-type GPCR gene is randomized by error-prone PCR and the resulting DNA library is used for transformation of yeast. The insert DNA and yeast expression vector backbone are assembled *in vivo* via designed homologous recombination sites. Next, expression in the obtained yeast library is induced, and subsequently cells are treated with an optimized buffer for permeabilization of the yeast cell wall. Permeabilization is necessary for allowing access and thus binding of administered fluorescent ligand to the functionally expressed receptors in the plasma membrane. After incubation with saturating concentrations of fluorescent ligand, unbound ligand is removed by washing, and cells are subjected to selection with FACS. Yeast cells expressing variants with the largest number of functional receptor molecules at the surface correspondingly exhibit the highest fluorescence. These cells are isolated by gating the top 0.5 – 1% of the most fluorescent yeast cells, and sorting them directly into growth medium for subsequent propagation. In order to achieve strong enrichment of cells expressing GPCR variants with the desired phenotype, several repetitive cycles of expression, incubation with fluorescent ligand, and FACS are performed. Plasmid DNA can be isolated from selected cells after FACS for analysis or introduction of further diversity by additional random mutagenesis, thereby starting the next round of evolution.

The inventors aimed to evolve three different GPCRs in parallel: (i) rat NTR1, (ii) human NK1R, and (iii) human kappa-type opioid receptor (KOR1). NTR1 has been shown to be readily evolvable in the *E. coli*-based system in several studies, and thus was considered a positive control. NK1R has been successfully subjected to directed evolution towards higher expression in *E. coli* as well (Dodevski & Plückthun (2011), J Mol Biol 408:599–615). However, despite strong relative improvements, due to the very low expression levels of wild-type NK1R in the prokaryote, the evolved receptor variants expressed still at only moderate absolute levels compared to the other receptors evolved in *E. coli* (Sarkar et al. (2008), Proc

Natl Acad Sci USA 105:14808–14813; Dodevski & Plückthun (2011), J Mol Biol 408:599–615). Moreover, expression levels of these evolved NK1R variants were also low in *S. cerevisiae* (see below), suggesting that further improvement of functional production might be possible, which may benefit future studies on this receptor. Concerning KOR1, no attempts to evolve this receptor have been undertaken so far and it represents a challenging example regarding heterologous expression.

Following the procedure outlined above, for each of the three receptors only two rounds of evolution – each round consisting of one randomization by error-prone PCR and five subsequent selections by FACS – were sufficient to strongly increase functional expression levels. Fig. 2 shows a comparison of the functional expression levels in *S. cerevisiae* of wild-type GPCRs, NTR1 and NK1R variants previously evolved in the *E. coli*-based system (NTR1-D03 and NK1R-E11, respectively), and the library pools obtained after the second round of evolution, namely NTR1 2.5, NK1R 2.5, and KOR1 2.5 (library nomenclature: GPCR *a.b*, where *a* is the number of total randomizations, and *b* is the number of total FACS selections). Functional expression levels were measured either with flow cytometry (FC) or radioligand binding assays (RLBAs). Note that FC ligand-binding experiments determine functional receptors exclusively in the plasma membrane at the surface of intact individual cells, whereas RLBAs account for the total amount of functional receptors averaged across an entire population of lysed cells, and thus also detect functional GPCRs in intracellular membranes. For the wild-type GPCRs no specific signal is obtained in FC experiments, thus no active receptor in the plasma membrane at the surface is detected (Fig. 2 A, C, E). By contrast, in RLBAs for NTR1 and KOR1 low expression levels of functional receptor are measured, indicating that a small amount of active receptor is present, however, most likely retained in intracellular membranes (Fig. 2 B, F). For NK1R, no functional receptor is detected in any of the methods used (Fig. 2 C, D), consistent with what has been reported before (Butz et al. (2003), Biotechnol Bioeng 84:292–304).

Strikingly, the expression levels of the selected libraries are not only much higher than the wild-type receptors but also higher than the corresponding variants evolved in *E. coli* (NTR1-D03, NK1R-E11), illustrated by FC and RLBA analysis. 100,000 – 150,000 total functional receptors per cell are measured in RLBA for the selected yeast libraries, representing a 25 – 50-fold and 5 – 20-fold increase compared to the corresponding wild-type GPCRs and *E. coli*-evolved variants, respectively (Fig. 2 B, D, F). The FC data confirm active receptor at the surface by showing a distinct specific signal for both evolved variants from *E. coli* and selected yeast libraries (Fig. 2 A, C, E). However, the specific signal of the selected yeast libraries is shifted towards higher fluorescence intensity compared to the *E. coli*-evolved variants, reflecting an increase in active receptors at the surface. Furthermore, a significantly

smaller subpopulation of cells not expressing any functional receptor at the surface is detected in the selected yeast libraries.

Such a division into two subpopulations was reproducibly observed in FC experiments upon expression of all GPCRs tested in both yeast single clones and libraries. Hence, this effect appears to be an intrinsic feature of GPCR expression in *S. cerevisiae*, in which functional expression at the surface is lacking in a substantial fraction of cells. For instance, depicted by the FC data of the *E. coli*-evolved variants, active surface expression of NTR1-D03 is obtained in about 50% of all cells, and for the low-expressing variant NKR1-E11 only a minority of cells show active surface expression (Fig. 2 A, C). It remains unclear why some cells of the population grown from one single clone do not produce any functional receptor at the surface, but a loss of the expression vector in these cells can be excluded since all cultivations are performed exclusively in selective minimal medium.

Improved expression levels of enriched receptor variants in yeast are further increased by host adaptation occurring concurrently with selection.

In order to identify enriched clones, plasmid DNA was isolated from the library pools after the first (stage 1.5 libraries) and second round (stage 2.5 libraries) of evolution for DNA sequencing. A summary of the selected clones depicting the different mutations of each clone is given in Table 1, with a scheme of GPCR topology depicting the different regions harboring mutations shown in Fig. 8.

Table 1: Overview of the most highly enriched clones during evolution for each GPCR. Mutations of each clone are indicated and the corresponding wild-type amino acids are given on top of the mutation data. Positions of mutations are indicated by structural regions, Ballesteros-Weinstein numbering (3), and sequential amino acid numbering. Section A: NTR1 variants. Section B: NK1R variants. Section C: KOR1 variants. Scheme of GPCR topology depicting the different regions harboring mutations is shown in Fig. 8.

Position	N										C3										C										Total mutations
	2.44	2.50	4.49	5.38	5.49	5.58	5.62	5.69			6.45	7.40	7.42	7.47	7.54	8.53															
	49	107	113	193	237	248	257	261	268	280	284	292	318	356	358	363	370	379	386	392	398										
NTR1	A	G	D	I	V	F	N	A	V	V	N	N	V	A	F	A	N	V	C	R	R										
NTR1-Y01'	P		N					S						S																4	
NTR1-Y02''	P		N			L		S		A				S																6	
NTR1-Y03'''	P	C	N					S						S																5	
NTR1-Y04'''	P		N					S						S																5	
NTR1-Y05'''	P		N					S						S					G	C										6	
NTR1-Y06'''	P		N				S	S			D			T	L	T	I			C										10	
NTR1-Y07'''	P		N	L	A			S	G			D	A	T	L		I	A		C										13	
Isolated from library: *NTR1 1.5																															
**NTR12.5																															

Position	H1										H5										H7										Total mutations
	1.40			3.36	3.49			5.59	5.64	5.66	6.51					7.34	7.47														
	40	62	95	116	129	138	181	217	222	224	264	274	278	286	299	322	331	335	336	337											
NK1R	T	R	H	V	D	L	M	T	T	W	F	N	Y	V	M	C	Y	E	M	K											
NK1R-ΔC																		K	*											2	
NK1R-Y01'								S						A					*											3	
NK1R-Y02''												Y							*											2	
NK1R-Y03'''																	*													1	
NK1R-Y04'''				I			K			R									*											4	
NK1R-Y05'''				I			K			R						W		K	*											6	
NK1R-Y06'''	A			I			K			R						W		K	*											7	
NK1R-Y07'''				I			K									W		K	*											5	
NK1R-Y08'''	A	S									Y				T				*											5	
NK1R-Y09'''			R			H	K			R			F			W		K	*											8	
NK1R-Y10'''					V		K		R		Y					W		K	*											7	
Isolated from library: *NK1R 1.5																															
**NK1R 2.5																															

Position	N										H3										H5										Total mutations
	1.28	2.45	2.50	2.65	2.63	3.31	3.35	3.47	3.50		6.42	7.44																			
	8	10	24	55	100	105	110	118	137	141	153	156	203	281	301	321	353	377													
KOR1	F	G	P	S	N	D	T	V	I	N	S	R	E	A	S	T	E	N													
KOR1-Y01'							A	D		S	C					I	K													6	
KOR1-Y02''		R			S	G		D		S	C					G	I	K												9	
KOR1-Y03'''					S	G		D		S	C						I	K												7	
KOR1-Y04'''			R		S	G		D		S	C						I	K												8	
KOR1-Y05'''					S	G		D		S	C				V		I	K												8	
KOR1-Y06'''					S	G		D		S	C	H	D				I	K												9	
KOR1-Y07'''	I			P	S	G		D	T	S	C						I	K	D											11	
Isolated from library: *KOR1 1.5																															
**KOR1 2.5																															

Interestingly, whereas for NTR1 and KOR1 only full-length variants were selected, all selected NK1R variants contain a stop codon at a position within a narrow range of seven amino acids after the presumed helix 8 and palmitoylation site, resulting in a shortened C-terminus. This indicates that a full-length C-terminus is strongly selected against and a truncation leads to a significant increase in expression for this receptor. Since no functional wild-type NK1R can be detected in yeast, the inventors constructed an alternative reference variant, termed NK1R-ΔC, with the most prominently selected C-terminal truncation combined with the wild-type sequence.

This provided the inventors with a reference for the selected mutations and also allowed them to quantify how much a C-terminal truncation contributes to improved functional expression for NK1R.

Individual clones were analyzed after retransformation and revealed that all enriched receptor variants were indeed significantly better expressed in yeast than the corresponding wild-type receptors (Fig. 3 and Fig. 9). This indicates that variants have indeed been obtained that are better adapted to the biosynthesis in this eukaryotic host.

While only a small fraction of cells not expressing any functional receptor at the surface was detected in the selected GPCR libraries (Fig. 2 A, C, E), a larger subpopulation of such cells was observed when the individual variants had been retransformed into fresh host cells (Fig. 3 A and Fig. 9 A, B). Furthermore, lower receptor-per-cell levels of individual NTR1 variants in cells with active surface expression after retransformation were also detected, as

indicated by the shift of the specific FC signal towards lower fluorescence intensity (Fig. 3 A). Consistent with these observations of a combined effect in FC, significantly lower average RLBA readings were obtained in retransformed strains expressing individual NTR1 variants (Fig. 3 B), compared to the selected GPCR library (Fig. 2 B).

The inventors hypothesized that in the course of selection an adaptation of the expression host occurred which led to a further increased level of membrane protein expression in the selected libraries. To test this hypothesis, a yeast clone, expressing a strongly selected variant termed NTR1-Y06, was directly isolated from the NTR1 2.5 library pool. As expected, the isolated strain expresses NTR1-Y06 at high functional levels, with a similar expression profile as the NTR1 2.5 library pool (Fig. 10 A, C). Sequencing of whole plasmid DNA isolated from both the retransformed and the isolated strain expressing NTR1-Y06 revealed no difference. Thus, the inventors concluded that the actual procedure of repetitive expression and sorting of the best expressing clones by FACS induced an adaptation of yeast towards improved GPCR production. To test this hypothesis, five subsequent mock selections (without any randomization of the sequence) with FACS with the retransformed single clone expressing NTR1-Y06 were performed, in which the cells with the highest NTR1-Y06 expression were isolated by gating the top 1.0% of the most fluorescent cells. Indeed, in the course of this phenotypic selections the strain adapted to higher expression of NTR1-Y06 by a gradual decrease of cells with no active surface expression and an increase of functional receptors per cell in the subpopulation with surface expression (Fig. 11). Finally after adaptation, NTR1-Y06 is expressed at total functional levels as high as the expression levels obtained in the selected NTR1 2.5 pool with about 160,000 average functional receptors per cell and a similar expression profile (Fig. 10 B, C). However, by repetitive cultivation of the adapted strain under non-expressing conditions, a partial reversion of the adaptive effect by a 30% drop of the expression level was observed (Fig. 12).

In order to test whether the phenotypic adaptation might even have been the main feature of the selection, the inventors also tried to adapt the strain expressing wild-type NTR1. However, after FACS, the sorted cells did not propagate anymore, preventing repetitive selections. This observation clearly shows that the selection of mutations, conferring improved properties to the receptors, is a strict prerequisite of the host adaptation, and that the sequence change is actually the key evolutionary event. The failure to induce adaptation in cells expressing wild-type NTR1 may be explained by the combined stressful effects of expression of the toxic wild-type GPCR followed by selection, together limiting cell survival.

Hence, adaptation induced by repetitive sorting appears to be restricted to evolved GPCR variants, which intrinsically are less toxic for the cells. Indeed, adaptation of strains expressing NK1R-Y09 and KOR1-Y05, evolved NK1R and KOR1 variants, respectively, was possible as well. Both variants were strongly enriched during selection and newly

transformed cells expressing NK1R-Y09 or KOR1-Y05 showed lower expression levels than the corresponding library pools. For each of the variants five rounds of phenotypic selection by FACS, identical to the procedure used to adapt the strain expressing NTR1-Y06, were performed. For both variants host adaptation towards higher expression levels was observed, gradually arising in the course of selection (Fig. 13). Even though the host adaptation effect for NK1R-Y09 and KOR1-Y05 is less pronounced than for NTR1-Y06, these results suggest that the adaptation is not receptor-specific, and can be reproducibly induced, if evolved receptors are expressed.

The fraction of functional receptor is increased in adapted strains, while the amount of total receptor produced is not significantly higher.

Since host adaptation was most pronounced for expression of NTR1-Y06, this variant was used to further study the effect. While ligand-binding FC analysis and RLBA permit the detection of active receptor, neither of these methods allow quantification of total receptor produced regardless of ligand-binding activity. Therefore, to detect total receptor protein produced in different strains by immunoblotting, NTR1 and NTR1-Y06 expression constructs with a hemagglutinin (HA) tag at the C-terminus of the receptors were created.

The strain expressing the HA-tagged NTR1-Y06 was adapted towards higher expression by five rounds of phenotypic selection with FACS (Fig. 14 A). Subsequently, total receptor production for the different strains expressing wild-type NTR1 and NTR1-Y06, the latter in non-adapted and adapted strain, was measured by HA tag detection in quantitative Western blots from whole cell lysates, using equal numbers of yeast cells after expression (Fig. 4).

In non-adapted strains, the amount of *total* receptor produced increases from wild-type NTR1 to the evolved variants by a factor of two (Fig. 4). The observed increase of *functional* expression by RLBA data is by a factor of four (Fig. 14 B). This indicates that the ratio of functional receptor to total receptor produced is just slightly increased from wild-type to the evolved receptor.

However, when comparing the adapted to the non-adapted strain, based on RLBA data, there is a tenfold increase in functional expression of NTR1-Y06 (Fig. 14 B) but there is less than a twofold increase of total receptor production (Fig. 4). Since the amount of total receptor produced increases only slightly, it can be deduced that in the adapted strain the fraction of functional receptor is significantly increased.

Adaptation increases surface expression of active receptor, reduces the number of cells exclusively expressing intracellularly retained GPCRs, and reduces the non-expressing cell subpopulation.

To further describe the host adaptation effect by assessing cellular localization and functionality of NTR1 variants in non-adapted and adapted strains, NTR1 and NTR1-Y06 constructs with a C-terminal mCherry (mCh) fusion were used for confocal microscopy studies and FC experiments: while total receptor is quantified by mCh, only functional receptors located at the cell surface in the plasma membrane will exhibit a signal from fluorescent ligand binding.

Identical to the HA-tagged variant, the strain expressing NTR1-Y06 with a mCh fusion was first adapted by phenotypic selection. Next, the non-adapted strains expressing NTR1 and NTR1-Y06, as well as the adapted strain expressing NTR1-Y06 were exposed to fluorescent ligand to perform co-localization studies with confocal fluorescence microscopy (Fig. 5).

As seen in ligand-binding FC experiments (Fig. 6 A), levels of functional wild-type NTR1 in the plasma membrane are very low, and consequently no ligand-binding signal is detected in microscopy (Fig. 5). In contrast, clear signals of mCh are obtained for some cells, however exclusively located in the cell interior, thus reflecting cells expressing intracellularly retained NTR1-mCh fusion receptors. According to the low RLBA readings obtained for wild-type NTR1 (Fig. 6 B), most of the intracellular receptor must be inactive. Additionally, many other cells are detected without any mCh signal at all, representing a non-expressing subpopulation (Fig. 5). Notably, both of these subpopulations, cells exclusively expressing intracellularly retained receptor as well as non-expressing cells, cannot be discriminated in ligand-binding experiments with FC. Thus, these two subpopulations together constitute the fraction of cells not showing any surface expression in FC ligand-binding data.

Evolved NTR1-Y06 expressed in the non-adapted strain shows detectable ligand binding at the surface as well as a mCh signal (Fig. 5). However, the signal of mCh at the plasma membrane is still rather weak with a strong mCh signal detected in the cell interior, indicating that, similar to wild-type NTR1, a substantial amount of receptor is retained in intracellular membranes. Furthermore, non-expressing cells without any signal, neither for ligand binding nor for mCh, are frequently detected as well. Hence, also for the evolved receptor, the subpopulation of cells with no surface expression detected in FC ligand-binding experiments comprises two types of cells: those without any receptor expression, and those cells in which expressed receptor is not translocated to the plasma membrane.

The situation is different for NTR1-Y06 expressed in the adapted strain, in which strong signals for ligand binding and mCh are detected in the plasma membrane (Fig. 5). Even

though some signal is still seen in the cell interior, mCh is mainly localized at the surface. Furthermore, significantly fewer non-expressing cells are observed.

Further investigations by FC experiments (Fig. 6 A), comparing the ligand-binding to the mCh signal, confirmed the results obtained in the confocal microscopy studies: signals for mCh in the expressing subpopulation are detected for all variants, however, for wild-type NTR1 the mCh fluorescence intensity does not correlate well with ligand-binding signal intensity. This is reflecting the fact that most of the wild-type receptor is not translocating to the cell surface. The finding that fluorescence intensity of wild-type GPCRs fused to an autofluorescent protein (e.g. GFP or mCh) correlates only weakly with functional receptor levels has been described before. Interestingly, this correlation is better for NTR1-Y06 expression in the non-adapted strain, and for NTR1-Y06 expressed in the adapted strain, total receptor produced and functional receptor correlate well. In this case, the receptor is mostly transported to the cell surface, reflecting the decrease of cells with exclusively intracellular receptor expression. Interestingly, the subpopulation showing not any mCh signal (non-expressing cells) is also decreased in the adapted strain, compared to the non-adapted strains, further leading to an overall increase of functional GPCR production levels in the adapted strain.

In summary, host adaptation of *S. cerevisiae* for GPCR expression increases surface expression of active receptor without increasing the total amount of receptor produced, which is equivalent to a higher percentage of active receptor produced. In contrast to non-adapted cells, for which a large fraction of cells does not show any functional surface expression, adapted cells show only small amounts of intracellularly retained receptor. Since a significant part of intracellular receptor is inactive, host adaptation leads to an overall increase of active receptor. Additionally, the subpopulation of cells not expressing any receptor at all is also decreased in the adapted strains compared to the non-adapted ones, contributing further to higher average functional production levels.

Evolved variants of all three GPCRs show high functional expression levels in Spodoptera frugiperda (Sf9) insect cells.

Ultimately, the inventor's goal was to increase functional GPCR expression not only in yeast but in insect cells as well. Thus, the functional expression levels of the evolved GPCRs NTR1-Y06, NK1R-Y09, and KOR1-Y05 were determined in Sf9 insect cells and compared to the corresponding wild-type receptors. For all GPCRs identical standard expression conditions (non-optimized) were used. Indeed, a significant increase of functional expression of the evolved variants was observed. For NTR1-Y06 compared to wild-type NTR1, functional expression is increased 5-fold (Fig. 7 A), for NK1R-Y09 a 4-fold increase is measured over wild type (Fig. 7 B), and for KOR1-Y05 – notably being the most challenging example regarding expression studied in this work – a 27-fold increase of functional

expression is detected (Fig. 7 C). The measured average functional receptor-per-cell levels ($4.1 \times 10^6 - 5.5 \times 10^6$ receptors per cell) are high and surpass expression levels of first- and even second-generation NTR1 mutants generated in the *E. coli*-based directed evolution system when those are expressed in Sf9 insect cells (Schlinkmann et al. (2012), J Mol Biol 422:414–428).

With such high expression levels, as obtained for all of our three GPCR mutants tested, expression cultures in liter scales are sufficient to produce enough material for structural investigations. To test this, purifications of NK1R-Y09 from insect cell expression cultures were performed, reproducibly yielding 4 – 6 mg/L pure protein. Fig. 15 shows the analysis of purified NK1R-Y09 by size-exclusion (SEC) and SDS-PAGE.

Given the very high purification yields, NK1R-Y09 represents an interesting and promising candidate for structural investigations. Thus, the functionality of NK1R-Y09 was characterized by signaling assays based on [^{35} S]-GTPyS binding (Fig. 7 D). NK1R-Y09 shows similar activity as wild-type NK1R and the C-terminally truncated variant NK1R- Δ C, both regarding [^{35} S]-GTPyS binding without stimulation (basal activity) and upon agonist stimulation, further underlining the potential of our approach.

Summary

Production of recombinant GPCRs in functional form still remains a demanding, laborious, and cost-intensive task. Given that most of the recently determined three-dimensional GPCR structures were obtained from receptors produced in eukaryotic expression hosts and encouraged by the success of protein engineering methods for increased expression of GPCRs in *E. coli*, the inventors aimed to transfer the directed evolution approach to *S. cerevisiae* to specifically improve functional GPCR production in eukaryotes.

The yeast *S. cerevisiae*, characterized by fast growth, cost-effective cultivation, and ease of genetic manipulations, is an ideal eukaryotic host for protein engineering and directed evolution for several additional reasons: First, yeast surface display has become a very powerful technology in a range of different applications. Second, high-diversity libraries are nowadays easily obtained with yeast. Third, the yeast cell is equipped with the cellular machinery required for functional GPCR production, since *S. cerevisiae* endogenously expresses two different GPCRs with signaling pathways similar to higher eukaryotes. Fourth, yeast has already been extensively used for heterologous expression of several different GPCRs. And fifth, based on a widely used GPCR assay which couples heterologous GPCRs to the yeast endogenous signaling pathway, directed evolution approaches on GPCRs have been successfully used before with the aim to generate designer receptors exclusively activated by designer drugs (DREADDs) activated by new ligands (Armbruster et al. (2007), Proc Natl Acad Sci USA 104:5163–5168).

So far, functional production yields of GPCRs in yeast have generally remained low (Sarramegna et al. (2003), Cell Mol Life Sci 60:1529–1546). Strategies to increase GPCR expression in *S. cerevisiae* included optimization of expression conditions, co-expression of molecular chaperones, screening for high-expressing host clones, screening of engineered GPCR variants, or host engineering. In this specification the inventors disclose the general applicability of this invention to directly obtain GPCR variants with increased functional GPCR expression levels by evolving three different GPCRs. Remarkably, only two rounds of evolution – meaning two randomizations each followed by selection with FACS – were required to obtain highly increased expression levels, and thus it was possible to evolve all three GPCRs efficiently in a short time. In the *E. coli*-based system, usually four rounds of evolution were required. Furthermore, the inventors disclose that the yeast host cell can be adapted towards increased GPCR production. This is an inherent feature of the selection system, which happens concurrently with the selection of improved receptor variants. It is important to reiterate that the directed evolution *did* lead to sequence changes, whose improved expression phenotype could be transferred to insect cells, and that the host adaptation is only an additional factor occurring in yeast during the selection. Moreover, the inventors found that expression of an evolved receptor is a prerequisite for inducing the host adaptation, as attempts to induce host adaptation in cells expressing a wild-type GPCR failed. The host adaptation of yeast can reproducibly be induced by repetitive phenotypic selection with FACS. This usually requires 4 – 5 sorts, and the effect gradually increases during the selection.

Adapted strains show a significant increase of surface-expressed active receptor, leading to a much higher fraction of *functional* receptors in the plasma membrane, with only a small increase of *total* receptor produced. On the contrary, in non-adapted strains, and especially for expression of wild-type receptors, a large amount of receptors remain in intracellular compartments, of which a significant amount represents inactive misfolded protein. The overall higher average functional production yields in expression cultures of adapted cells can be explained by a decrease of the subpopulation of cells expressing intracellularly retained inactive receptor as well as a decrease of the non-expressing subpopulation.

As the inventors did not wish to be limited to *S. cerevisiae* as a production host, it was explored whether the obtained GPCR variants would permit higher expression in other eukaryotic hosts as well. Indeed, functional expression of GPCR variants evolved in yeast was also increased in insect cells. Notably, variants evolved in the *E. coli*-based system show less improved functional expression when moved to eukaryotic hosts like insect cells, compared to the variants generated in yeast (Schlinkmann et al. (2012), J Mol Biol 422:414–428). The high expression levels in Sf9 cells obtained with the novel yeast-evolved variants are sufficient for crystallographic investigations with expression cultures in liter scales under

standard conditions. More specifically, purification of the NK1R variant NK1R-Y09 yielded very large quantities of pure protein. Furthermore, NK1R-Y09 showed signaling activities similar to wild-type NK1R, illustrating that the disclosed method can give rise to biologically active variants able to perform their naturally intended function, which consequently can be studied conveniently with these variants.

In summary, the method for directed evolution of GPCRs in *S. cerevisiae* disclosed herein allows efficient generation of GPCR variants with high functional expression in eukaryotic expression hosts in a short time. Using *S. cerevisiae* as evolution host resulted in GPCR variants with high functional expression levels in both yeast and insect cells, in the latter system by up to a factor 27. Thus, the mutations are transferable and selected variants showed a greater increase in expression in yeast as well as in insect cells than when evolution had been performed in *E. coli*.

The transferability of the beneficial effects of the receptor mutations to insect cells further promotes Sf9 cells as expression host for large-scale production of the generated mutants. Nonetheless, the fact that high expression levels in *S. cerevisiae* can now also be obtained as a result of GPCR evolution, combined with induced host adaptation, implies a potential of yeast as a large-scale production host. From a practical point of view, the fact that in adapted strains the ratio of functional receptor to total receptor produced is improved is a great advantage for purification strategies like IMAC, which cannot discriminate between non-functional and functional protein. This is of special interest for purifications of receptors for which no ligand affinity column, for instance like the cleavable ligand column as established for purification of NTR1 (Egloff et al. (2014), Protein Expr Purif, *in press*), is available.

The versatility of *S. cerevisiae* broadens the portfolio of GPCRs amenable to our approach, allowing evolution of GPCRs for which the previously established system in *E. coli* is inadequate. For instance, for GPCRs like KOR1, which show very low functional expression levels even in eukaryotes, expression levels in *E. coli* may be below the threshold needed for successful evolution. Furthermore, while neither prokaryotes nor native yeast do produce cholesterol – on which activity of some GPCRs relies – cholesterol-producing *S. cerevisiae* strains have been engineered. In principle, the disclosed method should be easily transferable to such or any other engineered yeast strains.

Yeast strain, vectors, cultivation, and expression.

For all experiments *S. cerevisiae* strain BY4741 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*), obtained from EUROSCARF, was used. Standard expression was performed with pMS03het, derived from p415 GAL1. For more details on different yeast expression vectors see below. BY4741 transformed with pMS03het vectors were cultivated at 30°C in SDD-Leu⁻ medium (6.9 g/L yeast nitrogen base without amino acids (Formedium), 690 mg/L complete

supplement mixture without leucine (Formedium), 20 g/L glucose, 35 mM sodium citrate tribasic, 35 mM citric acid). For expression, yeast cells in the logarithmic growth phase grown in SDD-Leu⁻ medium at 30°C were centrifuged and subsequently resuspended in SDG-Leu⁻ medium (identical to SDD-Leu⁻ but with 20 g/L galactose instead of glucose). Initial OD₆₀₀ was always chosen to be 1.0 after resuspension in SDG-Leu⁻ and expression was performed at 20°C for 24 h.

Yeast expression vectors

To obtain pMS03het, the α -mating factor prepro sequence was cloned from pPICZ α A (Life Technologies) into the multiple cloning site (*XhoI/SpeI*) of p415 GAL1. pMS03het contains *NheI/BamHI* restriction sites which allow efficient vector linearization for high-efficiency transformation or in-frame cloning of genes preceded by the α -mating factor prepro sequence. For expression of GPCRs with a C-terminal HA tag or fusion to mCherry, vectors pMS03het_HA or pMS03het_mCh were used, respectively. To obtain pMS03het_HA and pMS03het_mCh, sequences coding for HA tag or mCherry were cloned via *BamHI* into pMS03het.

Yeast library construction and transformation.

The wild-type gene of rat NTR1 (N-terminally truncated from amino acids 1 - 42) was a kind gift from Reinhard Grisshammer (National Institutes of Health). Wild-type cDNA of human NK1 R and human KOR1 was obtained from the Missouri S&T cDNA Resource Center. All wild-type GPCR genes were cloned into pMS03het (*NheI/BamHI*). DNA library construction was performed by amplification of wild-type genes or isolated DNA after the first round of evolution with error-prone PCR using the GeneMorph™ II Random Mutagenesis Kit (Agilent Technologies) according to the manufacturer's protocol. Transformation of BY4741 with DNA libraries was performed by square wave electroporation on a GenePulser Xcell™ electroporator (Bio-Rad). On average, libraries with a diversity of 5×10^7 - 1×10^8 were obtained. For more details on library construction and high-efficiency transformation see below.

DNA library construction and high-efficiency transformation

DNA libraries of wild-type GPCRs or isolated versions after the first round of evolution were generated by performing two error-prone PCRs (epPCRs), one epPCR with 20 and one with 25 cycles, of which the obtained products were subsequently pooled. Primers used for epPCR introduced sites homologous to linearized pMS03het (digested with *NheI/BamHI*) at each end of the gene.

Forward primer:

5'-CTAAAGAAGAAGGGGTATCTCTCGAGAAACGTGAGGCGGAAGCGGCTAGC-3';

Reverse primer:

5'-ATTACATGACTCGACTCGATGCCGACGAGAGCGGCCGCCTATTAGGATCC-3'.

The obtained epPCR products were further amplified by standard PCR with the identical primers in order to obtain enough DNA material for transformation.

High-efficiency transformation by square wave electroporation was performed analogously to a previously published method (Van Deventer & Wittrup (2014), *Methods Mol Biol* 1 131:151-181) with some minor adaptations. Yeast cells were grown in 60 mL YPD at 30°C to an $OD_{600} = 1.8 - 2.0$. As soon as this cell density was reached, 50 mL of culture were centrifuged, the medium aspirated, and cells were treated in 25 mL conditioning solution (100 mM lithium acetate, 10 mM DTT) at 30°C for 15 min. Subsequently, cells were pelleted, washed in 25 mL cold ddH₂O, pelleted again, and resuspended in cold ddH₂O to a total volume of 500 μ L. Henceforward, cells were always kept at 4°C. For one transformation, 250 μ L of yeast cells were mixed with 4 μ g of linearized pMS03het and 12 μ g PCR product and the transformation mixture was transferred to a 2 mm electroporation cuvette. Square wave electroporation was performed with one pulse with a voltage of 500 V and a pulse length of 15 ms. After electroporation, cells were allowed to recover in 5 mL YPD without shaking at 30°C for 1 h. Finally, recovered cells were pelleted, transferred to 500 mL SDD-Leu^r for selective growth at 30°C for 20 - 24 h, and stored in glycerol stocks at -80°C. To obtain high-diversity libraries, always two transformations per library were performed.

Permeabilization of yeast cells and binding of fluorescent ligand.

After expression, cultures were centrifuged, medium was aspirated, and cells were resuspended in TELi buffer (50 mM Tris-HCl pH 9.0 (at 4°C), 1 mM EDTA, 100 mM lithium acetate) at RT. Next, cells were incubated in TELi Buffer supplemented with 50 mM DTT at 20°C for 30 min and subsequently washed twice in cold TELi Buffer. Henceforward, cells were always kept at 4°C. For fluorescent ligand binding, permeabilized cells were incubated with ligand labeled with HiLyte Fluor™ 488 (AnaSpec) (NTR1 variants: 25 nM fluorescent neurotensin (8-13); NK1 R variants: 20 nM fluorescent substance P; KOR1 variants: 10 nM fluorescent dynorphin A (1-11)) in TELi buffer at 4°C without exposure to light for 2 h. After incubation, cells were washed once in TELi buffer prior to measurements. Non-specific binding was determined in the presence of a 1000-fold excess of unlabeled ligand (NTR1 variants: 25 μ M neurotensin (8-13) (AnaSpec); NK1 R variants: 20 μ M substance P (AnaSpec); KOR1 variants: 10 μ M dynorphin A (1-11) (GenScript)). For more details on fluorescent ligands see below.

Fluorescent ligands

All fluorescent ligands were labeled with HiLyte Fluor™ 488 (Anaspec). Neurotensin (8-13) (KKPYIL) was covalently labeled at the N-terminal amino group. Substance P (RPPKQQFFGLM) was covalently labeled at the amino group of lysine-3. Dynorphin A (1-11) (YGGFLRRIRPK) was covalently labeled at the amino group of lysine-11.

Flow cytometry and FACS.

Cells fluorescently labeled by ligand binding were kept in TELi buffer for measurements. Flow cytometry was performed on a BD FACSCanto™ II cytometer (BD Biosciences) or on a BD LSRFortessa™ cell analyzer (BD Biosciences) and FACS was performed on a BD FACSAria™ III sorter (BD Biosciences). For analytical measurements always 50,000 events were recorded. In FACS, 3×10^5 - 5×10^5 of the 0.5 - 1.0% most fluorescent cells were sorted into SDD-Leu-medium for subsequent cultivation at 30°C for 24 h. For all samples identical acquisition settings were used in order to allow comparative analysis. Data were analyzed with FlowJo vX.0.7.

Radioligand binding assays.

RLBAs on whole yeast cells were performed as follows: First, 1×10^8 cells (assuming $OD_{600} = 1.0$ corresponds to 10^7 cells/mL) were harvested after expression and treated by consecutive washing first in ddH₂O, then SPH1 buffer (1 M sorbitol, 25 mM EDTA, 50 mM DTT, pH 8.0), and finally in 1 M sorbitol. Next, cells were resuspended in SPH2 buffer (1 M sorbitol, 1 mM EDTA, 10 mM potassium citrate tribasic, pH 5.8) and cell wall digestion was performed by addition of 6 U/mL Zymolyase 20T (AMS Biotechnology) followed by incubation at 30°C for 30 min. Henceforward, cells were kept at 4°C. Subsequently, cells were incubated at 4°C for 2 h in 50 mM Tris-HCl pH 7.4 (at 4°C) containing [³H]-labeled ligand (NTR1 variants: 20 nM [3,11-Tyrosyl-3,5-³H(N)]-neurotensin (Perkin Elmer); NK1 R variants: 15 nM [Leucyl-3,4,5-³H(N)]-substance P (Perkin Elmer); KOR1 variants: 15 nM [15, 16-³H]-diprenorphine (Perkin Elmer)). Nonspecific binding was determined in the presence of a 1000-fold excess of unlabeled ligand (NTR1 variants: 20 μM neurotensin (8-13) (AnaSpec); NK1 R variants: 15 μM substance P (AnaSpec); KOR1 variants: 15 μM diprenorphine (Tocris Bioscience)). After incubation, cells were filtered on MultiScreen™ filter plates (Merck Millipore) with a vacuum manifold, filters were washed four times with cold 50 mM Tris-HCl pH 7.4 (at 4°C), transferred to Isoplate-96™ scintillation plates (Perkin Elmer), dried at 65°C for 2 h, and Optiphase Supermix™ scintillation cocktail (Perkin Elmer) was added. RLBA measurements were performed on a 1450 MicroBeta™ Plus liquid scintillation counter (Wallac).

RLBAs on whole Sf9 cells were performed as described previously (Schlinkmann et al. (2012), J Mol Biol 422:414-428; Egloff et al. (2014), Proc Natl Acad Sci USA 111:E655-62.). Whole cells were incubated in binding buffer (50 mM Tris-HCl pH 7.4 (at 4°C), 1 mM EDTA, 0.1 % (w/v) BSA and 40 pg/mL bacitracin) containing 15 nM [³H]-labeled ligand. Nonspecific binding was determined in the presence of 15 μM unlabeled ligand. The same ligands as for RLBAs with yeast cells were used.

Quantitative Western blot.

After expression, 4×10^7 cells (assuming $OD_{600} = 1.0$ corresponds to 10^7 cells/mL) for each sample were centrifuged and whole cell protein extraction was performed according to a previously published protocol (Zhang et al. (2011), Yeast 28:795-798.). Protein detection was performed with the primary antibodies rabbit anti-HA (Sigma-Aldrich, H6908) and mouse anti-

actin (Abcam, ab8224) and the secondary antibodies goat anti-rabbit conjugated to Alexa Fluor™ 680 (Life Technologies, A-21076) and donkey anti-mouse conjugated to IRDye™ 800 (Rockland Immunochemicals, 610-732-124). Image acquisition was performed on an Odyssey™ system (LI-COR Biosciences) and quantification was performed with Image Studio Lite™ version 3.1.4 (LI-COR Biosciences). For more details see below.

Yeast whole cell protein extraction, SDS-PAGE, and Western blot

For whole cell protein extraction, samples after expression were centrifuged, medium was aspirated, and pelleted cells were resuspended in 500 µL 2 M lithium acetate for incubation on ice for 5 min. Next, cells were pelleted, the supernatant was removed, 100 µL of 0.4 M NaOH were added, and samples were incubated on ice for 5 min. After incubation, the samples were centrifuged, the supernatant was removed and the cell pellets were resuspended in 200 µL reducing NuPAGE™ LDS sample buffer (Life Technologies). Samples were incubated at 20°C for 15 min, centrifuged, and 5 µL of each sample were run on a NuPAGE™ Novex™ 4-12% Bis-Tris protein gel (Life Technologies) in NuPAGE™ MES SDS running buffer (Life Technologies). Wet blotting was performed onto Immobilon™-FL membranes (Merck Millipore). Blocking of membranes was performed in 1 x Casein blocking buffer (Sigma-Aldrich) in PBS at RT for 20 min. Antibody binding was performed in 1 x Casein blocking buffer in PBST (PBS, 0.05% (v/v) Tween-20) at RT for 1 h, and PBST was used for all membrane washing steps. Primary rabbit anti-HA antibody was used at a dilution of 1:5,000, primary mouse anti-actin antibody at a dilution of 1:1,000, and the secondary antibodies (goat anti-rabbit conjugated to Alexa Fluor™ 680 and donkey anti-mouse conjugated to IRDye™ 800) both at a dilution of 1:10,000.

Confocal fluorescence microscopy.

Yeast cells were permeabilized and binding of fluorescent neurotensin was performed as described above. After washing, cells were transferred into Nunc™ Lab-Tek™ II chambered coverglasses (Thermo Scientific) and confocal microscopy was performed on a Leica™ TCS SP5 microscope (Leica Microsystems). For all samples magnification was 630-fold and identical acquisition settings were used in order to allow comparative analysis.

Spodoptera frugiperda (Sf9) vectors and expression.

Wild-type and evolved receptor constructs were amplified by PCR from yeast expression vector pMS03het and cloned via SLIC into a modified MultiBac™ pFL vector. The vector designated as pFL_mFLAG_His₁₀_TEV_SLIC contains an expression cassette with an N-terminal melittin signal sequence followed by a FLAG tag, a deca-histidine tag, a TEV protease cleavage site, and a SLIC cloning site. *E. coli* DH 10 EM BacY cells were transformed with pFL vectors containing the different receptor genes and the resulting baculovirus genome was isolated. For details on generation of recombinant baculovirus and baculovirus-infected insect cell stocks (BICs) see below. Expression was performed in Sf-900™ II SFM medium (Life Technologies) by infection of

Sf9 cells at a density of 3×10^6 cells/mL with 100-fold diluted BMC stocks and cultivation at 27°C for 4 d. After expression, cells were harvested by centrifugation, washed in cold PBS and stored at -80°C until use.

Generation of recombinant baculovirus and baculovirus-infected insect cell stocks (BIIcs).

Recombinant baculovirus was generated by transfecting 8×10^5 *Sf9* cells in 2 mL of Sf-900™ II SFM medium (Life Technologies) using 8 µL Cellfectin™ II reagent (Life Technologies). After 4 h of incubation in a humidified incubator at 27°C, the transfection medium was removed and replaced by 2 mL of fresh Sf-900™ II SFM. V0 viral stock was harvested after 5 d at 27°C and used to generate V1 high-titer virus stock (10^8 - 10^9 viral particles per mL). V1 virus stock was then used to generate BIIcs. Briefly, *Sf9* cells at a density of 10^6 cells/mL were infected with a multiplicity of infection (MOI) of 5, incubated for 24 h in suspension, harvested and frozen at -80°C in aliquots in Sf-900™ II SFM containing Penicillin-Streptomycin (Life Technologies) and 10% (v/v) DMSO.

[³⁵S]-GTPγS binding assay.

Membranes used for [³⁵S]-GTPγS binding assays were isolated as previously described (Egloff et al. (2014), Proc Natl Acad Sci USA 111:E655-62). Briefly, cells were disrupted by shear force after osmotic shock. In order to reduce unspecific [³⁵S]-GTPγS binding, membranes were washed in a urea containing buffer. The isolated membranes were frozen in aliquots and stored at -80°C. Receptor levels on urea-washed membranes were determined by radioligand binding assays as described above.

The [³⁵S]-GTPγS binding assay was performed as previously described (Egloff et al. (2014), Proc Natl Acad Sci USA 111:E655-62). Briefly, 1 nM GPCR in urea-washed membranes and 100 nM G protein (Gaii3iYi, purified according to (Rasmussen et al. (2011), Nature 477:549-555.)) in assay buffer (50 mM Tris-HCl pH 7.4 (at 4°C), 1 mM EDTA, 100 mM NaCl, 1 mM DTT, 3 mM MgSO₄, 0.3% (w/v) BSA, 2 µM GDP, 4 nM [³⁵S]-GTPγS (Perkin Elmer)) were incubated at 25°C for 20 min in the presence or absence of 200 µM substance P (AnaSpec).

Background counts arising from buffer, GPCR and G protein alone have been taken into account and subtracted. Therefore, given counts represent the GPCR-induced [^{35} S]-GTP γ S binding to G protein in the presence and absence of agonist.

Purification of NK1R variant NK1R-Y09

All steps were performed at 4°C. Frozen Sf9 cells were thawed and swelled in a hypotonic buffer (10 mM HEPES pH 7.4, 20 mM KCl, 10 mM MgCl₂, cOmplete ULTRA EDTA-free tablets (Roche), 1 μ M substance P) for 1 h. Cell membranes were then disrupted by homogenization (Dounce homogenizer) and collected by centrifugation at 130,000 rcf. The isolated membranes were extensively washed twice by repeated homogenization (Dounce homogenizer) and centrifugation in hypotonic buffer, followed by three repetitions of this procedure using a hypertonic buffer (10 mM HEPES pH 7.4, 1 M NaCl, 20 mM KCl, 10 mM MgCl₂, cOmplete ULTRA EDTA-free tablets, 1 μ M substance P). Purified membranes were resuspended in solubilization buffer (30 mM HEPES pH 7.4, 150 mM NaCl, 10 mM MgCl₂, cOmplete ULTRA EDTA-free tablets, 1 μ M substance P, 2 mg/mL iodoacetamide (Sigma)) and stirred for 45 min. Membranes were then solubilized in 1.5% (w/v) n-dodecyl- β -D-maltopyranoside (DDM, Anatrace) and 0.3% (w/v) cholesteryl hemisuccinate (CHS, Sigma). After 3 h of stirring, non-solubilized material was removed by centrifugation (130,000 rcf, 40 min, 4°C). The supernatant was adjusted to contain an end concentration of 800 mM NaCl and 25 mM imidazole before being incubated overnight with 1.5 mL TALON Superflow resin (Clontech). The protein-bound resin was transferred into gravity flow columns and then washed with 10 column volumes (CV) each of Wash 1 buffer (25 mM HEPES pH 7.4, 800 mM NaCl, 10 mM MgCl₂, 25 mM imidazole, 10% (v/v) glycerol, 0.5 μ M substance P, 0.3%/0.06% DDM/CHS), Wash 2 buffer (25 mM HEPES pH 7.4, 400 mM NaCl, 10 mM MgCl₂, 40 mM imidazole, 10% (v/v) glycerol, 0.5 μ M substance P, 0.2%/0.04% DDM/CHS, 10 mM ATP), and Wash 3 buffer (25 mM HEPES pH 7.4, 200 mM NaCl, 40 mM imidazole, 10% (v/v) glycerol, 0.5 μ M substance P, 0.2%/0.04% DDM/CHS). NK1R-Y09 was eluted in 4 CV of elution buffer (25 mM HEPES pH 7.4, 200 mM NaCl, 300 mM imidazole, 10% (v/v) glycerol, 0.5 μ M substance P, 0.2%/0.04% DDM/CHS). Purified receptor was concentrated to 0.5 mL with 100 kDa molecular weight cut-off Vivaspins centrifuge concentrators (Sartorius Stedim Biotech). Size-exclusion chromatography (SEC) was performed on an Akta Pure FPLC system (GE Healthcare) with a Superdex S200 Increase 10/300 GL column (GE Healthcare) equilibrated with SEC buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 0.05%/0.01% DDM/CHS, 0.1 μ M substance P).

Example 2: Directed evolution of GPCRs in yeast

Directed evolution of GPCRs towards higher expression levels in *Escherichia coli* has been established previously (Sarkar et al. (2008), Proc Natl Acad Sci USA 105:14808–14813; Dodevski & Plückthun (2011), J Mol Biol 408:599–615). This invention discloses the analogous method for directed evolution in the yeast *Saccharomyces cerevisiae*. Several GPCRs have been evolved with this method in *S. cerevisiae*, leading to high functional expression in yeast as well as in insect cells and outperforming expression levels obtained from variants evolved in the *E. coli*-based system. Usually, only two rounds of evolution were required to obtain highly expressed variants.

Interestingly, the yeast strain used can also be adapted towards increased GPCR production. This is automatically induced during evolution by repetitive sorting with fluorescence-activated cell sorting (FACS) and gradually arises in the course of selection. Thus, obtained expression levels of selected libraries are a result of GPCR evolution combined with host adaption. If individual GPCR variants are expressed in newly transformed yeast cells, the average expression level will be lower than for selected libraries, since these cells were not adapted. However, simply by repetitive FACS selections the adaption can be induced for any evolved variant.

GPCR Expression Constructs

Several different GPCR expression constructs are available. All vectors are derived from p415 GAL1. The vectors are maintained as low-copy plasmids in *S. cerevisiae* providing a functional LEU2 gene for selective growth of strains auxotrophic for leucine and can be easily re-isolated from yeast after transformation. Being shuttle-vectors, they can be propagated as high-copy vectors in *E. coli*, in which correspondingly all cloning steps are performed.

Every vector encodes the α -mating factor prepro sequence preceding the cloning site for insertion of the GPCR gene. Cloning is done via *NheI* and *BamHI* restriction sites, allowing in-frame insertion of the gene of interest. Expression is under control of the inducible GAL1 promoter. The different vectors vary in the sequences after the cloning site. The standard vector used for evolution, pMS03het, does not encode any tag or protein after the cloning site, thus GPCRs without any fusions are produced with pMS03het. The other vectors encode constructs containing a C-terminal cleavable deca-histidine-tag (pMS03het_3CHis10), a HA-tag (pMS03het HA), a fusion to mCherry (pMS03het_mCh), or an AviTag™ (pMS03het_Avi).

Permeabilization of Yeast Cells for Ligand Binding

This section describes the permeabilization of yeast cells for ligand binding experiments. Conditioning with TELi buffer and DTT permeabilizes the cell wall and allows diffusion of

ligands to the GPCRs in the cell membrane. The advantage of this permeabilization method is that the cells remain viable and are not as fragile as spheroplasts. Therefore, this procedure can be used for analytical flow cytometry and FACS, as well as for any other application in which ligand binding is detected (e.g. confocal fluorescence microscopy, RLBA).

Protocol:

- Cells are collected after expression by centrifugation (5,000 rcf, 10 min, RT) and the supernatant is aspirated.
- Cells are resuspended in 1 mL TELi buffer at RT, centrifuged (4,000 rcf, 3 min, RT) and the supernatant is aspirated.
- Cells are resuspended in 900 μ L TELi buffer at RT and 100 μ L freshly prepared and filter-sterilized 0.5 M DTT (dissolved in 100 mM lithium acetate solution) is added. This gives a final concentration of 50 mM DTT.
- Cells are incubated at 20°C in a thermomixer, shaking with 700 rpm for 30 min.
- Cells are centrifuged (4,000 rcf, 3 min, 4°C) and the supernatant is aspirated.
- Cells must from now on always be kept on ice.
- Cells are resuspended in 1 mL cold TELi buffer, centrifuged (4,000 rcf, 3 min, 4°C) and the supernatant is aspirated.
- Cells are resuspended in cold TELi buffer (the volume is chosen according to the experimental needs).

Measurement of GPCR Expression with Flow Cytometry

Detection of functional GPCR expression at the cell surface can in principle be performed by fluorescent ligand binding measured in flow cytometry. Upon incubation with fluorescent ligand, only functionally expressed GPCRs located in the plasma membrane will bind the ligand. After removal of unbound ligand by washing, ligand binding is detected by measuring fluorescence intensity. However, the yeast cell wall acts as a protective barrier, not allowing larger molecules like peptide ligands to pass through. Therefore, yeast cells need to be permeabilized prior to incubation with fluorescent ligand. Next to analytical quantification of functional expression levels, fluorescent ligand binding is also used for selections with FACS. Cells expressing a variant with a high-expression phenotype will correspondingly exhibit a higher fluorescence intensity. By sorting the most fluorescent cells, highly expressing GPCR variants can be isolated.

While incubation with fluorescent ligand allows measurement of the total signal, it is important to determine the non-specific signal as well. This can be done by incubation of fluorescent ligand with an excess of non-labeled ligand in a competition binding experiment.

The obtained non-specific signal in such measurements depicts the background. Usually, GPCR expression in yeast leads to two subpopulation of cells. Whereas one fraction of cells shows surface expression of active GPCRs, there is also a subpopulation for which no active surface expression is observed. Typically, this fraction of non-expressing cells decreases in adapted strains.

Workflow of Selections with FACS

An overview of the workflow of selections is depicted in Fig. 1. For directed evolution of GPCRs the selection starts with the generation of a library. A new DNA library is created by random mutagenesis with error-prone PCR (epPCR) on a GPCR gene. The DNA library has to be amplified by PCR (ampPCR) with special primers generating an insert suitable for high-efficiency transformation. Once a yeast library has been created by transformation with the desired efficiency, selections with FACS can be performed. Typically, 5 rounds of FACS are performed. Once the selections with FACS are done, the DNA of the selected clones is isolated. At this step, the selected clones can be analyzed. If desired, another round of directed evolution can be performed by creating a new library from the selected clones by epPCR and subsequent selections with FACS. In previous directed evolutions of GPCRs, two rounds of evolution — meaning two randomizations, each followed by five selections with FACS — were required to obtain highly increased expression levels.

For adaptation of strains expressing an evolved GPCR variant, the workflow is almost identical. The only difference is that no library needs to be prepared. Strain adaptation starts with a single clone expressing the GPCR variant of choice which is subjected to 5 rounds of selection with FACS. After the selections, glycerol stocks of the adapted strains can be prepared.

Claims

1. A method for selecting a sequence from a library of expressed nucleic acid sequences, wherein the sequence is selected according to its expression level, the method comprises the steps of
 - a) providing a plurality of eukaryotic cells comprising a cell wall, wherein each of said eukaryotic cells comprises a nucleic acid sequence member of said library, and said nucleic acid sequence member is expressed as a G protein-coupled receptor (GPCR) in the plasma membrane in said plurality of eukaryotic cells,
 - b) permeabilizing said cell wall of said plurality of eukaryotic cells in a permeabilization step, wherein the permeabilization step comprises exposing the plurality of eukaryotic cells to a non-enzymatic chemical treatment, yielding a plurality of viable permeabilized cells,
 - c) contacting said plurality of viable permeabilized cells in a labelling step with a ligand capable of binding to said GPCR, wherein said ligand comprises a detectable label, yielding a plurality of labelled cells,
 - d) washing said plurality of labelled cells in a washing step,
 - e) selecting a subset of said plurality of labelled cells as a function of detectable label present in said plurality of labelled cells in a selection step, yielding a selection of cells, and
 - f) isolating an expressed nucleic acid sequence from said selection of cells in an isolation step,wherein the expressed nucleic acid sequence obtained in step f) is characterized by high expression levels of the encoded GPCR;
and wherein said detectable label is a fluorescent dye and said selection step is accomplished by fluorescent cell sorting.
2. The method according to claim 1, wherein:
 - i. after said selection step e) said selection of viable cells is expanded in an expansion step, yielding an expanded selection of viable cells and said expanded selection of viable cells is subjected to said steps b) to e),
 - ii. said step i. is performed at least 1, 2, 3, 4, 5, 6 or 7 times, finally followed by said isolation step f).
3. The method according to claim 1, wherein

- i. the expressed nucleic acid obtained in said isolation step f) is introduced and expressed in a plurality of eukaryotic cells comprising a cell wall,
 - ii. said plurality of eukaryotic cells comprising a cell wall is subjected to said steps b) to f) according to claim 1, and
 - iii. steps i. and ii. are performed at least 1, 2, 3, 4, 5, 6 or 7 times.
- 4. The method according to any one of claims 1 to 3, wherein said expressed nucleic acid obtained in said isolation step f) is amplified by a process introducing mutations into the amplified sequence, yielding a second library of nucleic acid sequences and
 - i. said second library of nucleic acid sequences is transferred to said plurality of eukaryotic cells comprising a cell wall, and
 - ii. said plurality of eukaryotic cells comprising a cell wall is submitted to the method according to any one of claims 1 to 3.
- 5. The method according to any one of claims 1 to 4, wherein the library of expressed nucleic acid sequences is obtained by amplification of a nucleic acid sequence encoding said GPCR by a process introducing mutations into the amplified sequence.
- 6. The method according to any one of claims 1 to 5, wherein said chemical treatment comprises a buffer of alkaline pH comprising lithium ions, a reducing agent and/or a chelating agent.
- 7. The method according to any one of claims 1 to 6, wherein said plurality of eukaryotic cells is a plurality of yeast cells.
- 8. A method for the selection of an adapted yeast cell with the ability for high expression levels of functional G protein-coupled receptors (GPCRs), comprising the steps of
 - a. providing a plurality of yeast cells, wherein each of said yeast cells comprises a nucleic acid sequence member of a library of expressed nucleic acid sequences, and said nucleic acid sequence member is expressed as a GPCR in the plasma membrane in said plurality of yeast cells,
 - b. permeabilizing the cell wall of said plurality of yeast cells in a permeabilization step, wherein the permeabilization step comprises exposing the plurality of eukaryotic cells to a non-enzymatic chemical treatment, yielding a plurality of viable permeabilized cells,
 - c. contacting said plurality of viable permeabilized cells in a labelling step with a ligand capable of binding to said GPCR, wherein said ligand comprises a detectable label, yielding a plurality of labelled cells,
 - d. washing said plurality of labelled cells in a washing step,

- e. selecting a subset of said plurality of labelled cells as a function of detectable label present in said plurality of labelled cells in a selection step, yielding a selection of cells,
- f. expanding said selection of cells in an expansion step, yielding an expanded selection of cells
- g. submitting said expanded selection of cells to steps b. to f. at least 1, 2, 3, 4, 5, 6 or 7 times,
- h. submitting said expanded selection of cells to steps b. to e., and
- i. selecting a subset of said expanded selection of cells as a function of detectable label present in said plurality of labelled cells, yielding said adapted yeast cell with the ability for high expression levels of functional GPCRs from said expanded selection of cells;

wherein said detectable label is a fluorescent dye and said selection step is accomplished by fluorescent cell sorting.

Fig. 1

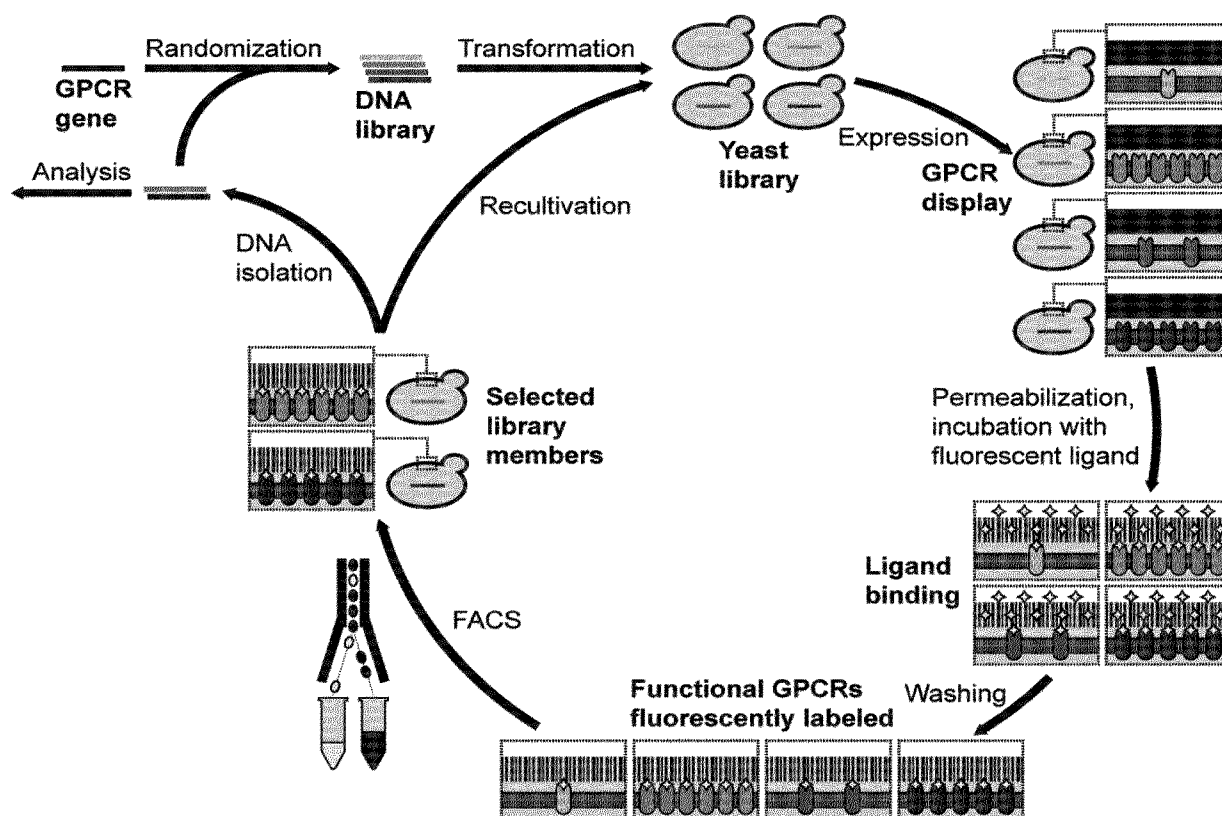
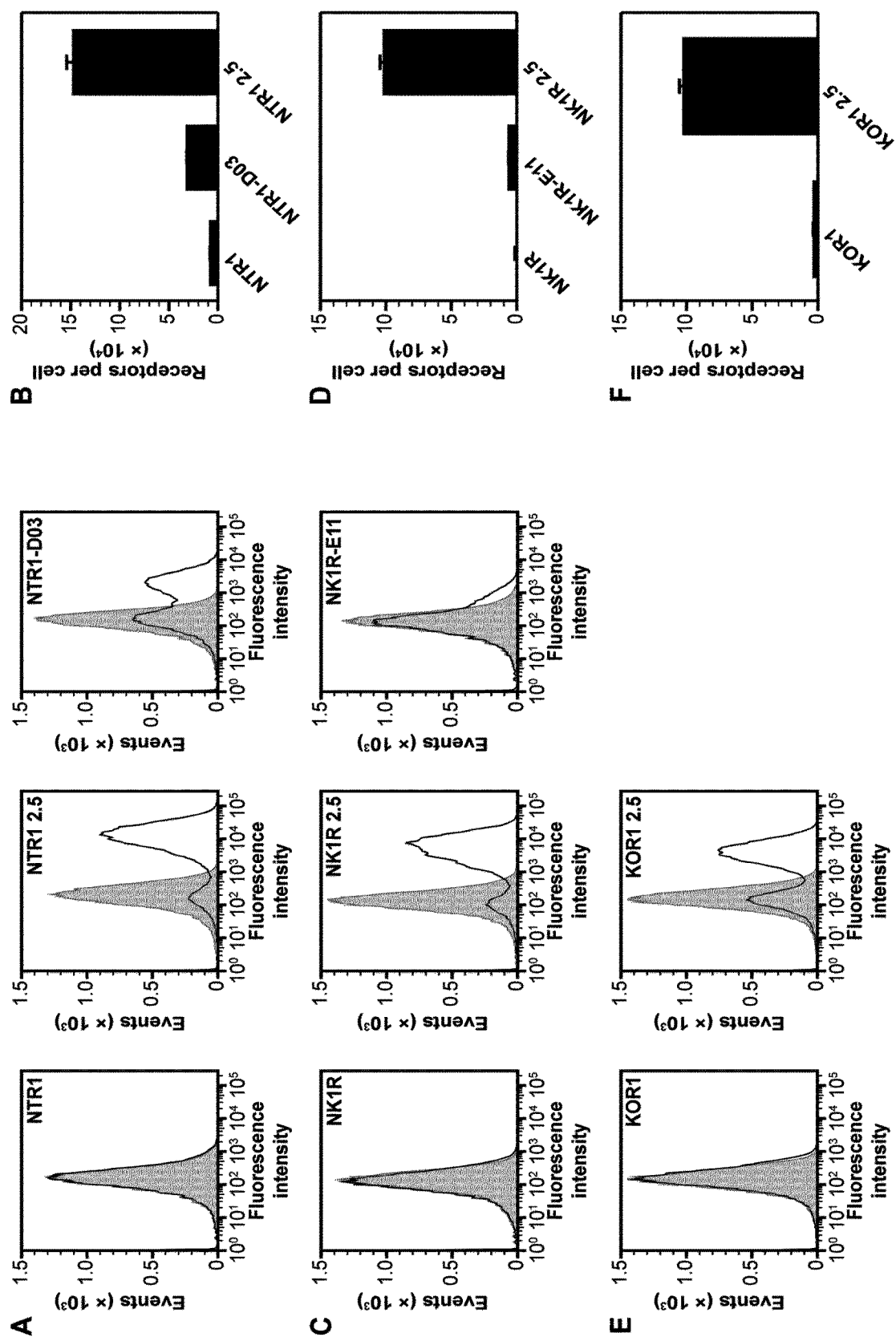


Fig. 2



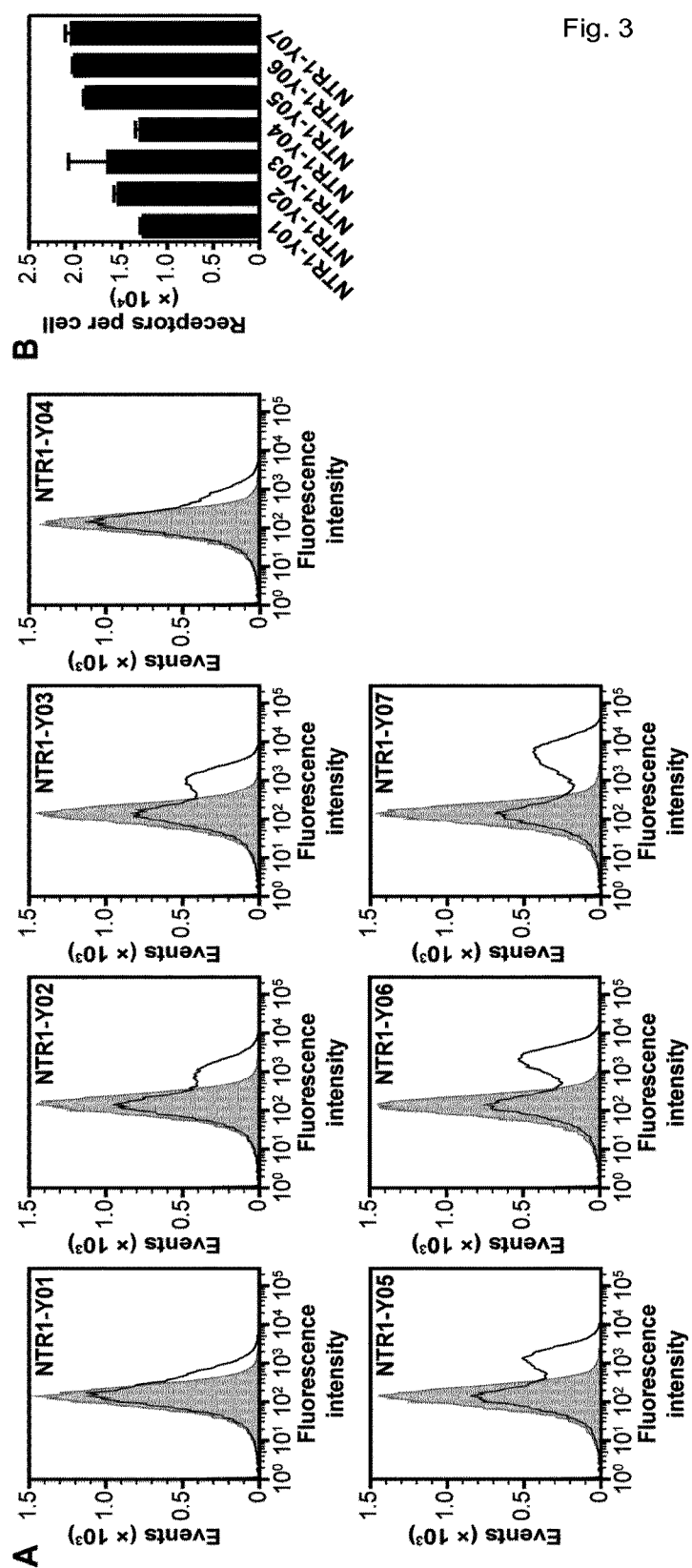


Fig. 4

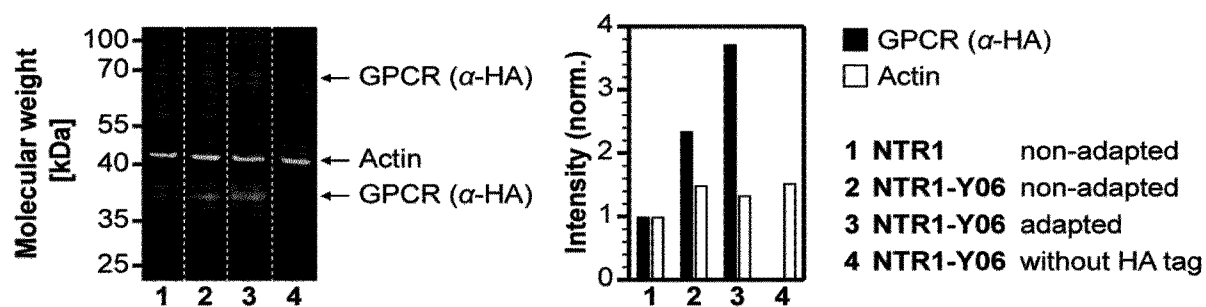
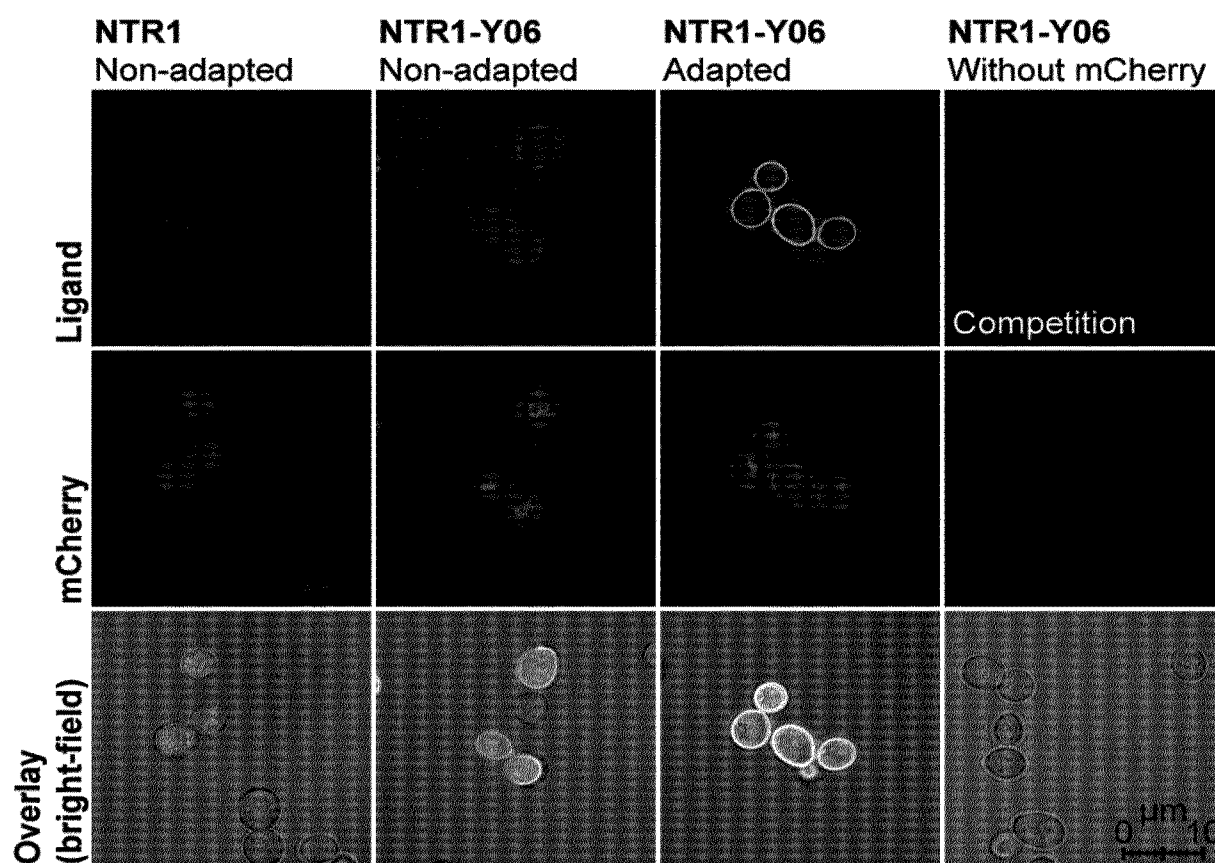


Fig. 5



WO 2016/102508

PCT/EP2015/080858

Fig. 6

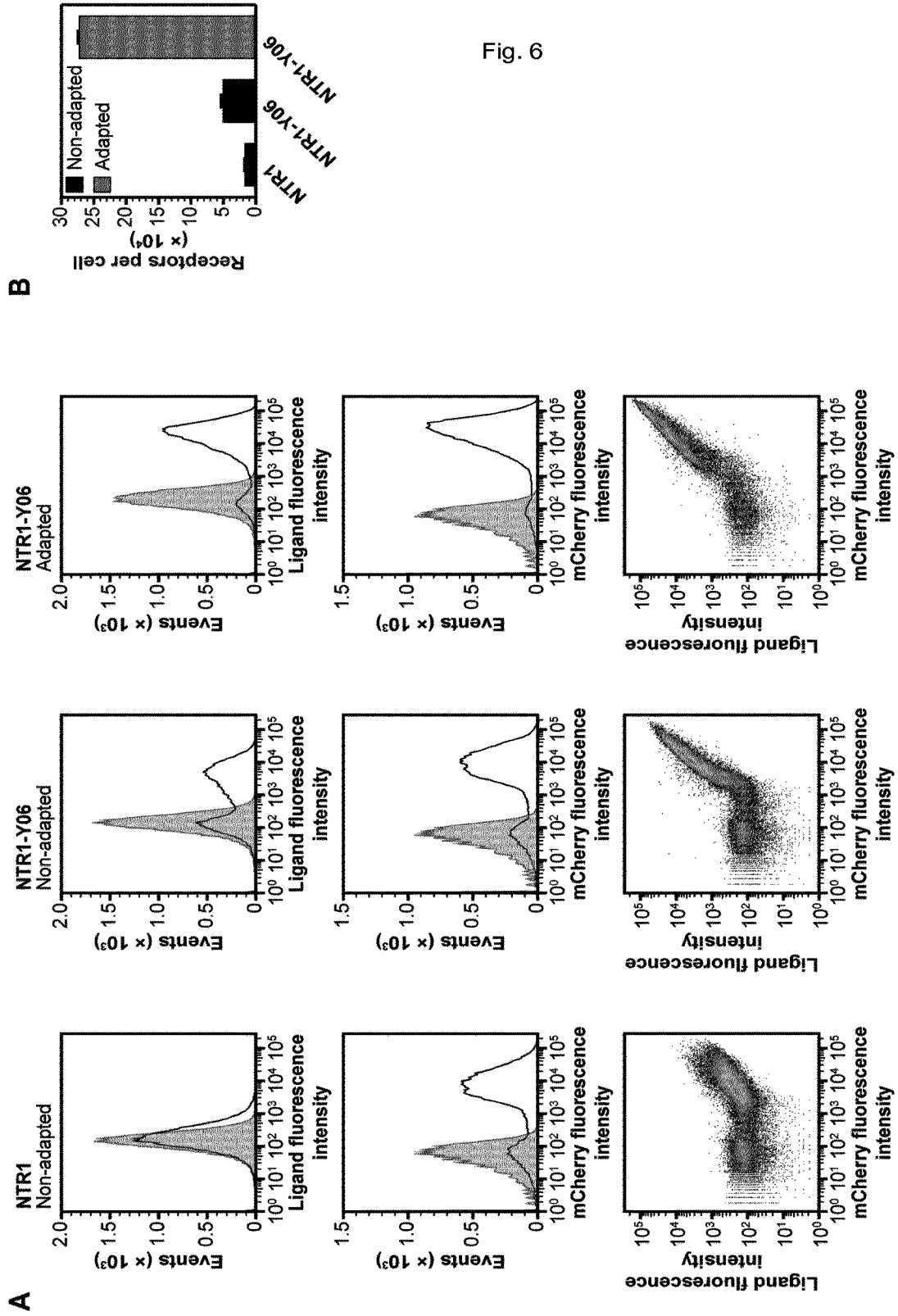


Fig. 7

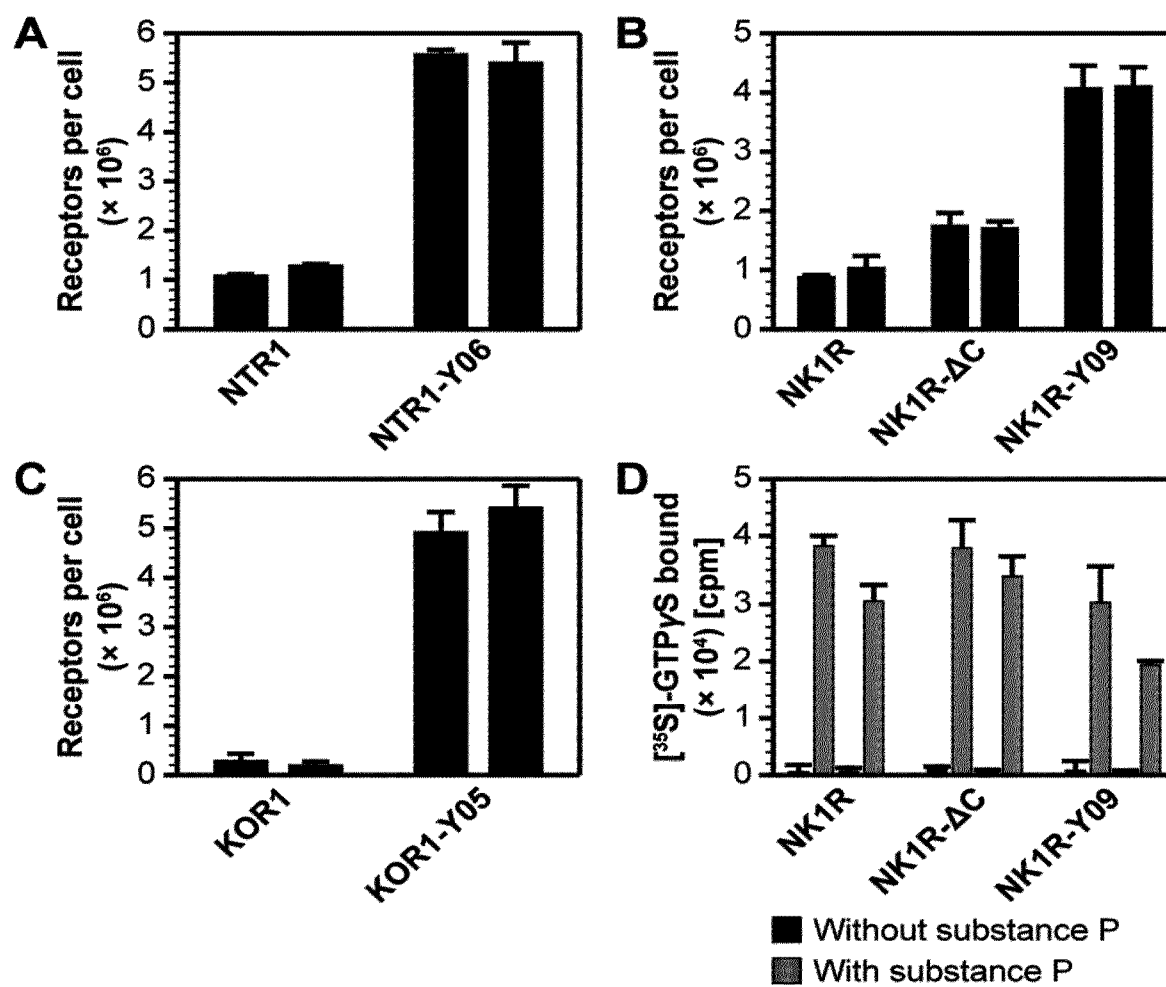
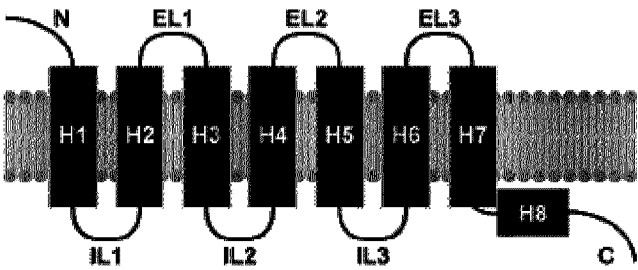


Fig. 8



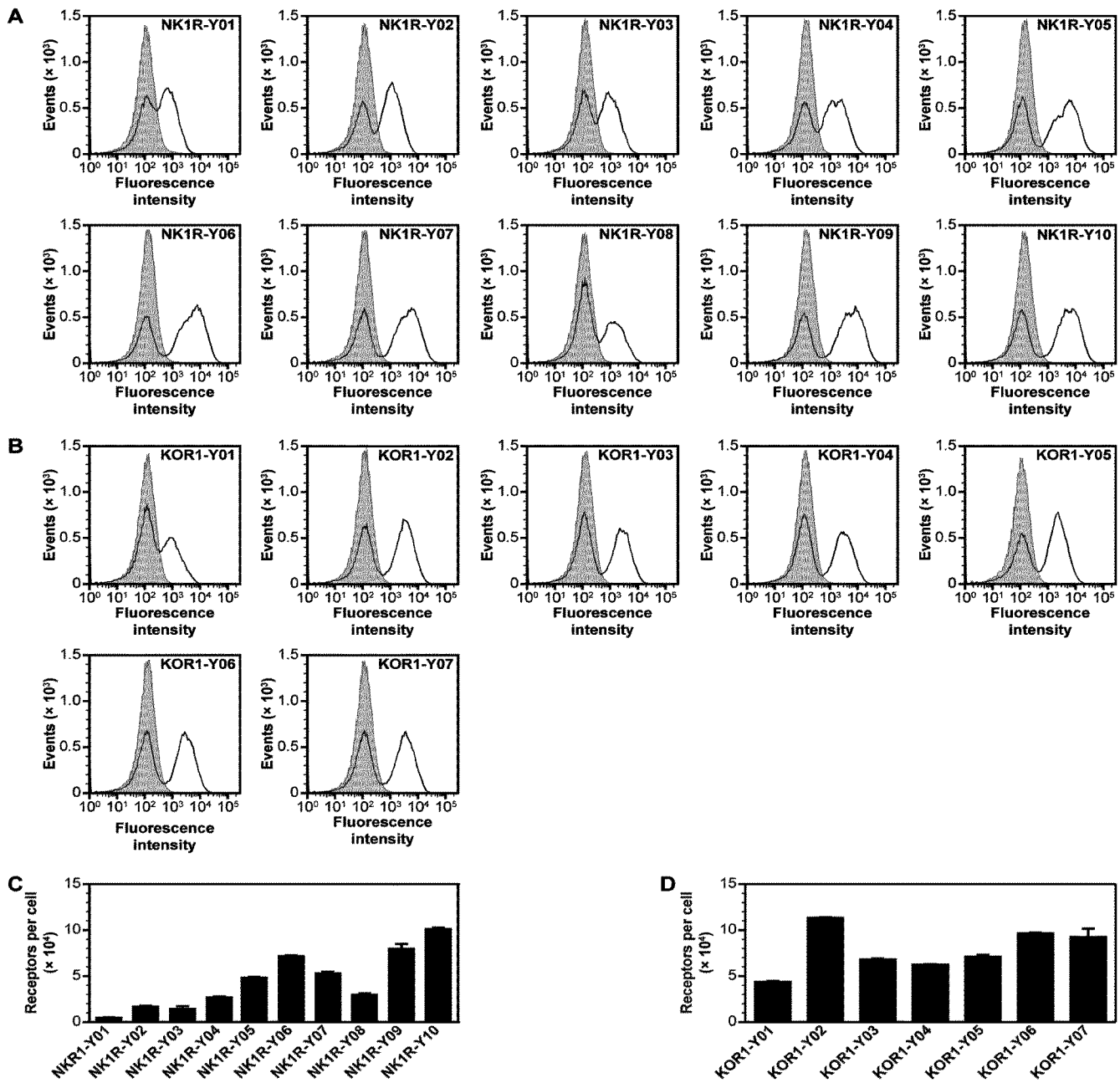


Fig. 9

Fig. 10

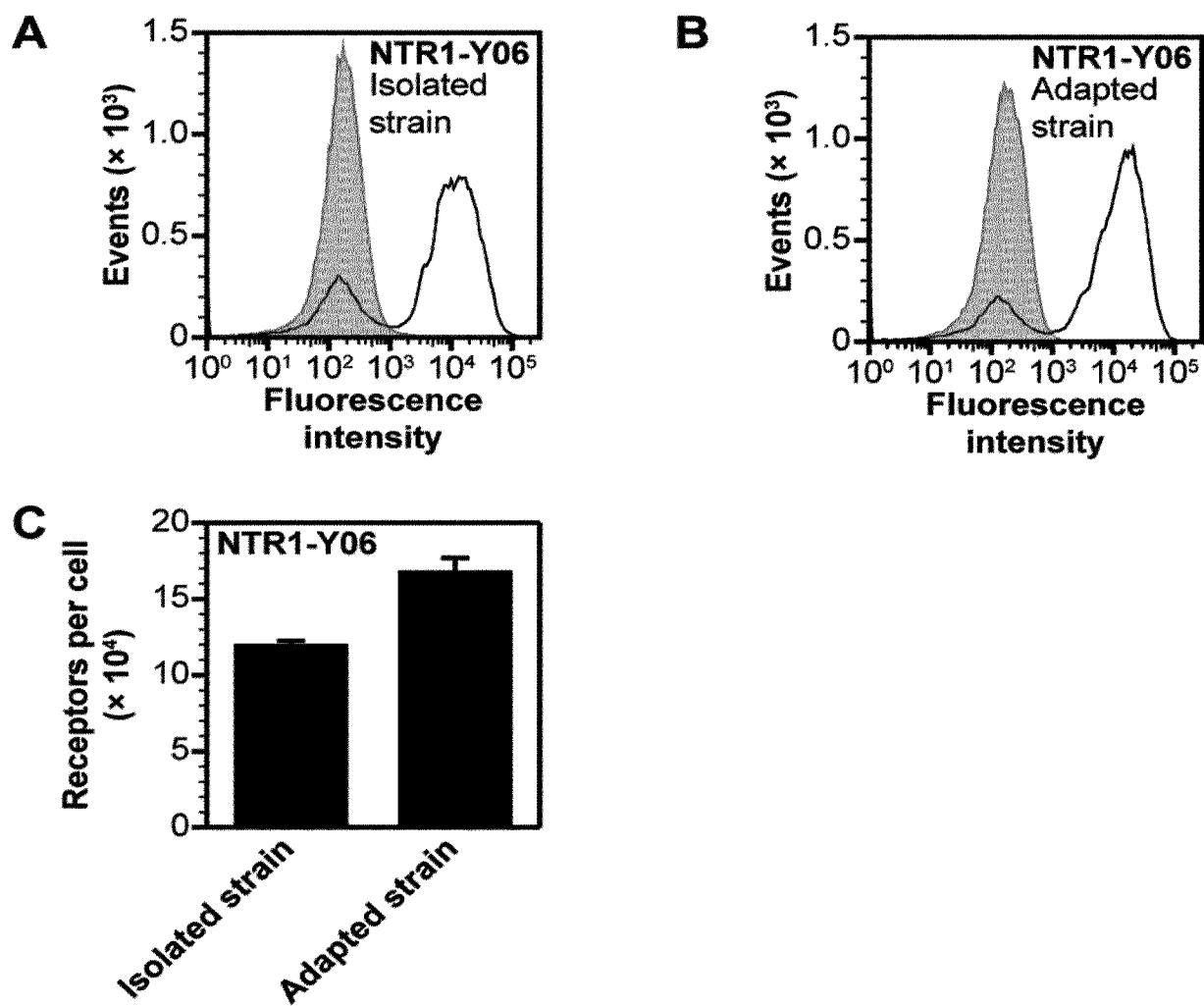


Fig. 11

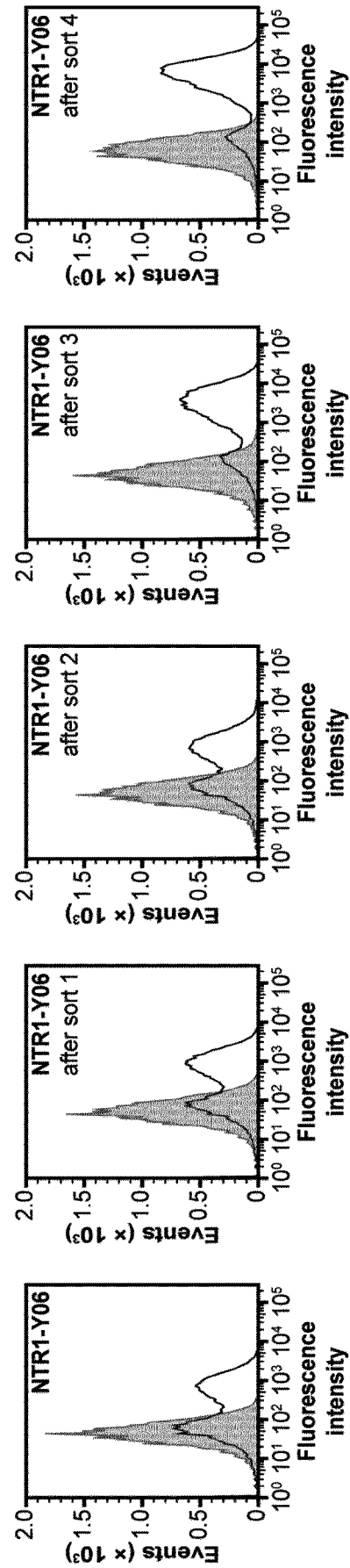


Fig. 12

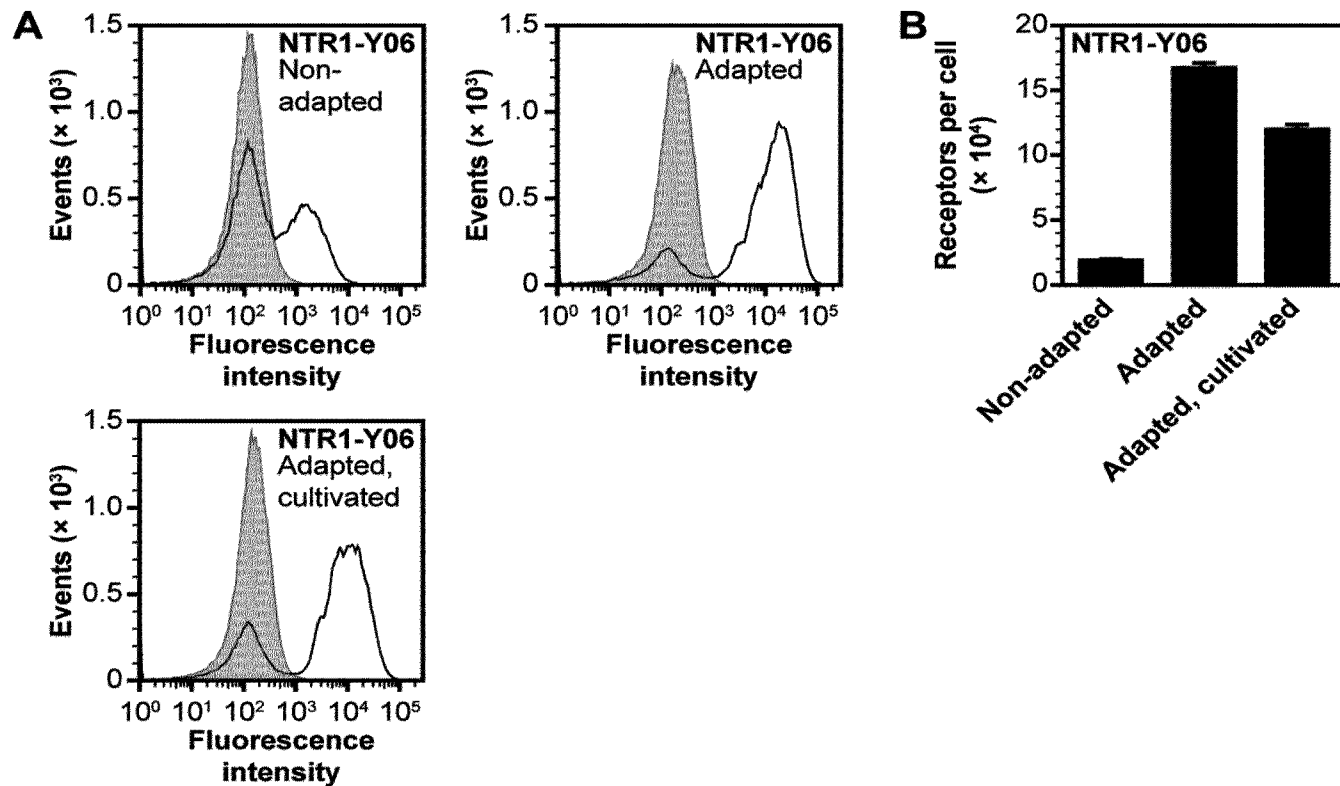


Fig. 13

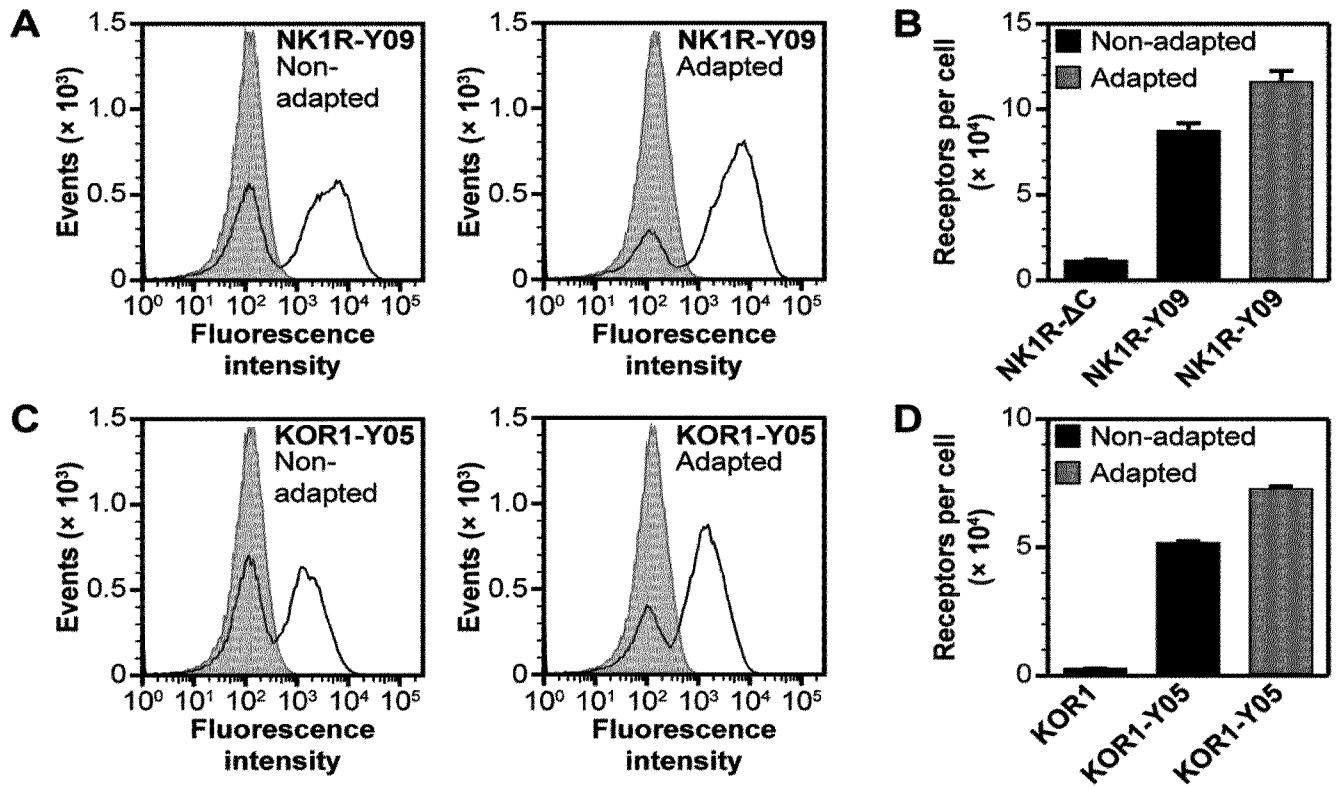


Fig. 14

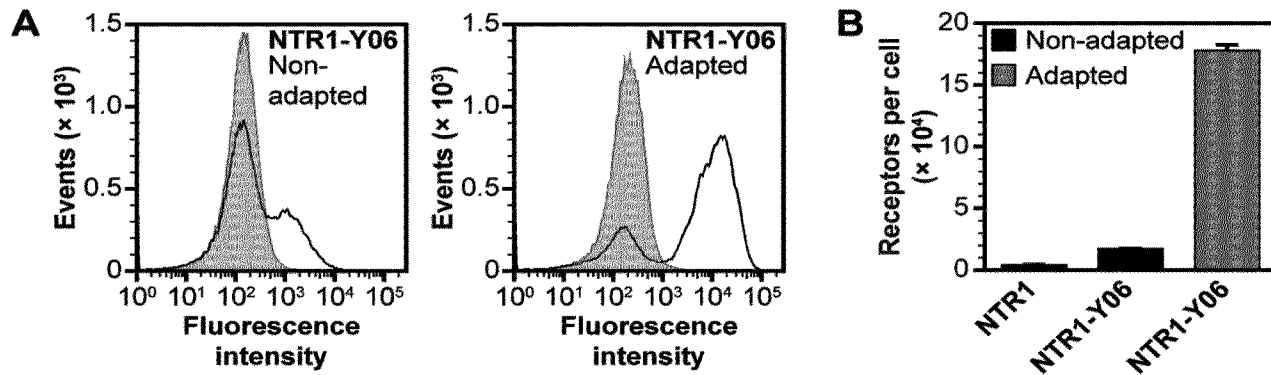


Fig. 15

