Published in final edited form as:

Methods Enzymol. 2015 ; 557: 27–43. doi:10.1016/bs.mie.2014.12.031.

A novel screening approach for optimal and functional fusion of T4 lysozyme in GPCRs

Elizabeth Mathew and Mark E Dumont¹

Department of Biochemistry and Biophysics, PO Box 712, University of Rochester School of Medicine and Dentistry, Rochester NY 14642

Abstract

Determination of high resolution, three dimensional structures of transmembrane proteins has, in many cases, only been accomplished through the use of stabilized variant forms of the proteins being studied. For the important G protein coupled receptor superfamily this has most often been achieved by inserting a stable soluble protein, such as T4 lysozyme (T4L) in an internal loop of a receptor. However, creation of such fusion proteins generally results in loss of the ability of receptors to activate their cognate cytoplasmic G proteins. Furthermore, the criteria for designing fusions that minimally perturb receptor structure are not well established. We describe here a method for creating a library of receptor variants containing T4L inserted into an internal loop at varying positions and as replacements for varying amounts of the original receptor sequence. We also describe methods for screening for variants displaying maximal expression levels, ligand binding capacity, and signaling function. When applied to the yeast α -factor receptor, Ste2p, this approach allowed recovery of well-expressed receptor variants containing internally fused T4L that retained nearly normal signaling function. The approach we describe can be readily adapted to creation of stabilized fusions of other transmembrane proteins expressed in yeast or other expression systems.

1. INTRODUCTION

Genes for transmembrane proteins (TMPs) make up 20–30 % of most genomes (Wallin and Heijne, 1998) and, based on their critical roles in cell physiology, transmembrane proteins are the targets of a large fraction of clinically useful drugs. Thus, determination of their structures and modes of action is of considerable importance. However, the number of available high resolution TMP structures remains relatively low (see web page of Stephen White; <http://blanco.biomol.uci.edu/mpstruc/>). Two explanations that are frequently invoked for the difficulty of obtaining well-diffracting crystals of TMPs for use in x-ray diffraction studies are: 1) TMPs are structurally dynamic with flexible regions that make it difficult for the protein to adopt the uniform conformation required for packing into a crystal; and 2) Surfaces of transmembrane regions of TMPs do not participate as readily as the surfaces of globular soluble proteins in protein-protein interactions that are required for crystal packing. Introducing a stable soluble protein into a surface-exposed region of a TMP provides a way of potentially circumventing both of these problems, since such an insertion at an internal

¹Corresponding author, Mark_Dumont@urmc.rochester.edu.

location in a TMP can reduce overall flexibility and since the presence of a readily crystallizable soluble domain can provide intermolecular contacts required for crystallization (Chun, et al., 2012, Engel, et al., 2002).

Fusion of T4 lysozyme (T4L) at internal locations in TMPs has, to date, provided a successful approach for determination of the structures of diverse members of the important GPCR superfamily (Cherezov, et al., 2007, Cherezov, et al., 2007, Chien, et al., 2010, Dore, et al., 2014, Fenalti, et al., 2014, Granier, et al., 2012, Haga, et al., 2012, Hanson, et al., 2012, Hollenstein, et al., 2013, Jaakola, et al., 2008, Kruse, et al., 2012, Manglik, et al., 2012, Rosenbaum, et al., 2007, Rosenbaum, et al., 2011, Shimamura, et al., 2011, Tan, et al., 2013, Wacker, et al., 2013, Wang, et al., 2013, White, et al., 2012, Wu, et al., 2010, Wu, et al., 2012, Xu, et al., 2011, Zhang, et al., 2012). Similar internal fusions of a thermally stabilized apocytochrome *b* (Wacker, et al., 2013, Wang, et al., 2013, Zhang, et al., 2014, Zhang, et al., 2014) and rubredoxin (Tan, et al., 2013) have also yielded GPCR structures. In addition, several GPCR structures have been obtained by crystallization of constructs in which the stabilizing protein partner was fused at the N-terminus of the receptor sequence, rather than at an internal position (Fenalti, et al., 2014, Rasmussen, et al., 2011, Siu, et al., 2013, Thompson, et al., 2012, Wu, et al., 2014), suggesting that the role of fusion proteins in facilitation of formation of intermolecular contacts may be more important for promoting crystallization than the reduction of TMP flexibility. Alternative approaches for obtaining GPCR structures that do not involve the use of fusion proteins have included co-crystallization with anti-receptor antibodies (Kruse, et al., 2013, Rasmussen, et al., 2007, Rasmussen, et al., 2011, Rasmussen, et al., 2011, Ring, et al., 2013) and the crystallization of GPCR variants that have been thermostabilized by the introduction of point mutations and deletions (Egloff, et al., 2014, Scott, et al., 2013, Tate, 2012). In fact, most fusion protein-containing GPCR variants used for successful structure determination have also incorporated point mutations and truncations designed to further enhance protein stability and reduce flexibility.

Insertion of stable soluble proteins into TMPs for the purpose of promoting crystallization has generally been performed as an insertion or replacement for residues in the third intracellular (IC3) loop of GPCRs. This is based on the expectation that this region in receptors is flexible and may control the relative motion of surrounding helices (Rosenbaum, et al., 2007), as well as reducing the likelihood that insertion of the fusion partner will disrupt the overall GPCR structure (Chun, et al., 2012). While such insertions often do not disrupt the overall structures of receptors (based on retention of at least some ligand binding affinity by the fusion constructs), they do generally result in loss of the ability to activate the cognate trimeric G protein (Rosenbaum, et al., 2007), which is not surprising since the IC3 loop is often the site of functional interactions between GPCRs and the trimeric G proteins that are their immediate downstream effectors. Furthermore, the criteria for establishing particular junction sites for insertion of fusion proteins and for determining the amount of receptor sequence that they replace are not well-established. A successful fusion must provide an optimal match between the spacing and orientation of the N- and C-termini of the inserted protein and the junction sites in the GPCR (Chun, et al., 2012, Engel, et al., 2002). Thus, creation of a useful fusion may require synthesis or construction of multiple versions of fused genes (Rosenbaum, et al., 2007).

We describe here a strategy for generating and screening a library of variant forms of a receptor containing insertions of T4L at different points in the third intracellular loop (IC3) as a replacement for different numbers of amino acid residues in the loop sequence (Mathew, et al., 2013). By expressing receptors in yeast, it is possible to create randomized libraries of variants with lysozyme insertions, and to directly screen for constructs with maximal overall expression levels, numbers of ligand binding sites at the cell surface, and even receptor function, assayed based on activation of the cognate G protein. The approaches were successful in identifying receptors with insertions in the IC3 loop that retain nearly full signaling function. This suggests that, when attached via appropriate linking sequences, the T4L molecule can be inserted into this loop without preventing access of the G protein to relevant surfaces of the activated receptor.

The described protocols were developed using Ste2p, an endogenous GPCR of the yeast *Saccharomyces cerevisiae* used as a tractable initial system for which no three dimensional structure is yet available. However, similar approaches could be used to identify insertion-containing variants of the numerous mammalian GPCRs that have been reported to be capable of activating the yeast pheromone response pathway (Brown, et al., 2000, Dowell and Brown, 2009, King, et al., 1990, Pausch, 1997), or to other TMPs for which it is possible to assay for function in intact cells. Furthermore, yeast species have been used as expression systems for structure determination of GPCRs and other TMPs (Clark, et al., 2010), including for the determination of the crystal structure of the human H1 histamine receptor (Shimamura, et al., 2011, Shiroishi, et al., 2011). Ste2p is the receptor for the yeast mating pheromone α -factor, a 13-residue peptide. Binding of this ligand, results in activation of a cytoplasmic heterotrimeric G protein, which, in turn, activates a MAP kinase pathway, ultimately leading to transcriptional responses, changes in cell shape, cell cycle arrest, and fusion with yeast cells of the opposite mating type.

2. Overall Strategy

A library of mutant receptors was created in which T4L was inserted at different points in the predicted IC3 loop, accompanied by different extents of deletion of residues in this loop. (We refer to the region targeted for insertion and/or replacement by T4L as the “replacement region.”) As shown in Fig. 1, this was accomplished by initially creating an allele of the receptor-encoding gene *STE2* in which most of the IC3 loop was deleted and replaced by a unique restriction site. The receptor-encoding gene was also fused at its C-terminus to GFP. A cysteine-free version of T4L was PCR amplified, introducing restriction sites at the 5' and 3' ends of the coding region. A family of double stranded DNA linkers corresponding to different lengths of sequence from the receptor third loop was prepared from synthetic oligonucleotides. The linkers contained overhangs that allow them to be annealed and ligated to the 5' and 3' ends of the amplified T4L gene that had been restriction digested to create complementary overhangs. The ligated products containing T4L flanked by various lengths of sequence corresponding to the IC3 loop can be conveniently inserted into the IC3 loop-deleted receptor allele via Ligation-Independent Cloning (LIC cloning) (Alexandrov, et al., 2004, Aslanidis and de Jong, 1990, Dieckman, et al., 2002). Following transformation into bacteria, a library of plasmids is isolated and used to transform a yeast strain lacking the normal *STE2* gene, so that the T4L-containing variants are the only α -factor receptors

present in the cells. Screening for overall expression is accomplished by flow cytometry of the library of transformed cells, detecting expression of the Ste2p-GFP fusion. Screening for ligand binding capability is also conducted using flow cytometry to detect binding of a fluorescently labeled ligand. Screening for receptors signaling function is conducted based on expression of a *FUS1-HIS3* reporter gene that undergoes transcriptional induction in response to pheromone pathway activation, allowing growth of the host *his3⁻* strain on medium lacking histidine. Further characterization of receptor expression levels and signaling function can be conducted using quantitative assays of the binding of fluorescent ligands and of reporter gene induction as well as by immunoblotting of individual isolates.

3. Plasmids, Yeast Strains, and Media

3.1 Yeast Strains

A575 – (Mata *ste2⁻ bar1⁻ far1 cry1^R ade2-1 his4-580 lys2_{oc} trp1_{am} tyr1_{oc}SUP4-3^{ts} leu2 ura3 FUS1::p[FUS1-lacZ TRP1]* (Sommers, et al., 2000)

A2953 – Mata *ste2⁻ bar1⁻ far1⁻ leu 2⁻ his3⁻ 200 ura3⁻ FUS1-HIS3 FUS1-lacZ* (Taslimi, et al., 2012)

BJ5460 – (ATCC® 208285) Mata *ura3-52 trp1 lys2-801 leu2 1 his3 200 pep4::HIS3 prb1 1.6R can1 GAL* (Jones, 1991)

3.2 Plasmids

pMD1167 – Full length *STE2* with a C-terminal GFP fusion, under the control of the *STE2* promoter (confers ampicillin-resistance in *E.coli* and uracil auxotrophy in *S.cerevisiae*) (Gehret, et al., 2006).

pMD1859 – Created by site directed mutagenesis of plasmid pMD1167, replacing the IC3 loop of *STE2* with nucleotide sequences that code for two BsaBI restriction sites flanking a SmaI site (confers ampicillin-resistance in *E.coli* and uracil auxotrophy in *S.cerevisiae*) (Mathew, et al., 2013).

pSGP46 – Expression vector that contains the alcohol dehydrogenase (*ADH2*) promoter (confers ampicillin-resistance in *E.coli* and uracil auxotrophy in *S.cerevisiae*) (Clark, et al., 2010).

3.3 Media

Yeast media were prepared as described by Sherman (Sherman, 2002).

4. Library Construction

4.1 Generation of T4 Lysozyme-containing cassettes with variable lengths of flanking regions

1. Design and purchase a set of synthetic oligonucleotides corresponding to the replacement region of the receptor according to the following criteria:

5. Ligate the mixed linkers to the amplified T4L gene. Mix 0.1 pmole of each annealed linker such that the final concentration of total mixed linkers is 0.1 pmole/ μ l. Incubate 9 μ l (0.9 pmole) of this annealed linker mix with 0.3 pmole of digested amplified T4L gene in the presence of 1 mM ATP and 700 units T4 DNA ligase (New England Biolabs) in a total volume of 35 μ l annealing buffer. Incubate overnight at 16°C
6. Purify the ligation products by electrophoresis on an agarose gel, excising a gel slice encompassing the size range of the expected products. Extract DNA from the gel using a commercially available kit.

4.2 Insertion of cassettes containing the T4 lysozyme gene into the replacement region of the receptor

1. Create an allele of the receptor gene that containing a deletion of the replacement region (the region of insertion or replacement by T4L) (Fig. 2). A unique restriction site should be inserted as a replacement for the deleted region. This can be performed by standard techniques of site-directed mutagenesis. If ligation-independent cloning is to be used to insert the T4L-containing cassettes, this restriction site should be flanked by sequences that contain three of the four types of nucleotide bases but lacking the fourth type of base. If screening for overall expression levels is planned, the starting allele should be constructed so as to encode a receptor fused at its C-terminal to green fluorescent protein (GFP). It is anticipated that the GFP will subsequently be removed from forms of the receptor to be used for protein purification and crystallization.
2. If possible, the deletion should introduce a frame shift such that undesired self-ligation of vectors without insert will not result in read-through of the C-terminal region of the receptor. This will reduce background from plasmids in the library that lack an inserted T4L.
3. Perform ligation-independent cloning essentially as described by Clark et al. (Clark, et al., 2010). The individual steps are as follows:
 - i. Digest the plasmid that has the deletion-containing receptor gene with the restriction enzyme at the introduced site and gel purify the linearized plasmid.
 - ii. Use the exonuclease activity of T4 DNA polymerase to digest the 3' ends of the cut vector in the presence of only a single dNTP corresponding to the nucleotide that is missing from the junction sequence. Under these conditions (listed below), the polymerase removes nucleotides until it reaches a base complementary to the one present in the reaction, creating 5' overhangs.

Restriction enzyme-digested vector	400 ng
single dNTP	2.5 mM

Dithiothreitol	5 mM
T4 DNA polymerase buffer (Novagen)	final dilution 1X
T4 DNA polymerase (LIC-qualified; Novagen)	2 Units
Total volume	40 μ l

↓
22°C/40 min

75°C/15 min (to inactivate enzyme)

- iii. Incubate the T4-polymerase-treated vector (final concentration ~ 2.5 ng/ μ l) and T4L-containing cassettes (final concentration of assembled cassettes is ~ 0.5 ng/ μ l) in a total volume of 10 μ l annealing buffer and anneal for 5 – 60 min at 22°C. Add 2.5 μ l of 25 mM EDTA and continue to incubate at 22°C for an additional 5 minutes.
 - iv. Use 3 μ l of the vector/insert mix to transform 25 μ l competent NOVABlue cells following the manufacturer's instructions (Novagen). Enzymes present in the recipient cells carry out ligation of the vector and insert and can also fill in gaps formed in annealing overhangs that are not fully complementary. Plate the cells on an appropriate selection medium for the vector being used. (Typically, 6000 – 7500 transformants are obtained, compared with 4 – 6 in the absence of insert).
4. Perform DNA sequencing on plasmids isolated from randomly selected *E. coli* transformants to evaluate the frequency of plasmids with inserts and the diversity of the library. Of the 40 plasmids that were sequenced from the *STE2* library, all contained inserted T4L with variable flanking regions.
 5. Pool the bacterial transformants by adding 3 – 5 ml LB to the plate and gently scraping the surface of the agar with a sterile glass spreader. Pipette off as much of the LB plus cells as possible and transfer to a sterile microfuge tube. Culture the pooled colonies and libraries in LB supplemented with 100 μ g/ml ampicillin and isolate plasmid DNA. This DNA library can now be introduced into the test cells of choice (bacteria, yeast, tissue culture, etc.) and screened as needed for function and expression. For the *STE2* library, the pooled plasmids were used to transform a yeast strain containing a chromosomal *STE2* deletion (A2953), creating a library of 2400 yeast transformants.

5. Expression Screening and Functional assays

1. Flow cytometry is used to screen and quantitate the overall expression of the Ste2p-lysozyme fusions by detecting the emission from the C-terminally fused GFP.

2. Screening for fusions that are correctly folded and trafficked to the cell surface is conducted based on binding of a fluorescent derivative of the agonist α -factor essentially as described previously (Bajaj, et al., 2004, Bajaj, et al., 2007, Mathew, et al., 2011). The α -factor analog, [K⁷(bodipy-TR),Nle¹²] α -factor (provided by Dr. Fred Naider of the College of Staten Island, City University of New York), with a red shifted fluorophore, bodipy-TR (6-(((4-(4,4-difluoro-5-(2-thienyl)-4-bora-3a,4a-diaza-s-indacene-3-yl)phenoxy)acetyl)amino)hexanoic acid), is used to minimize interference from the GFP fused to the receptor. Cells grown to an OD₆₀₀ of ~1 in SD-Ura (synthetic dextrose medium lacking uracil)(Sherman, 2002) are incubated at a density of ~3 X 10⁶ cells/ml in 0.4 ml 20 mM acetate buffer (pH 4.6) with 100 nM fluorescent ligand, on ice (to prevent internalization), for 45 minutes. Fluorescence is then determined using a flow cytometer (*i.e.* a LSRII from BD Biosciences). The bodipy fluorophore was excited at 532 nm and detected with a 610/20 detector.
3. Screening for functional receptors is conducted based on induction of the pheromone responsive *FUS1-HIS3* reporter gene, which is integrated into the chromosome of strain A2953. The library transformants are replica-plated onto solid SD-ura and SD-ura-his medium with and without ~300 nM α -factor. There should be no growth when histidine and α -factor are both absent.
4. Clones containing fusions with the desired expression levels and functional properties should be cultured for the re-isolation of fusion-encoding plasmids. This allows for the determination of the sequences of the T4-lysozyme-containing genes. It is also important to confirm apparent phenotypes of particular alleles by retransforming into a fresh host strain and re-evaluating expression, folding, and signaling function to verify that the observed properties are plasmid-dependent. Plasmids are isolated from *S. cerevisiae* using a modification of the manufacturer's protocol for the "Wizard® Plus SV Minipreps DNA Purification System" (Promega). To disrupt the yeast cell wall before plasmid isolation, approximately 0.5 ml of 0.5 mm zirconia/silica beads (BioSpec Products, Inc) are added to cells pelleted from 1.5 ml of fresh culture grown overnight to an OD₆₀₀ > 1.5. "Resuspension" and "Cell Lysis" solutions (250 μ l each) are added, and the cells are vortexed for 5 minutes and centrifuged at 12,000 X g for 10 min to pellet the beads and unbroken cells. The supernatant from this spin is then pipetted into 350 μ l "Neutralization" buffer. The manufacturer's instructions are followed from this point except for the final elution volume, which is reduced from 100 to 30 μ l. Since plasmid yield is generally poor, the entire miniprep is used to transform competent *E.coli*, then plasmid is isolated from the bacteria for further manipulation.
5. To quantitate signaling function of T4L-containing receptors, plasmids of interest are used to transform the *ste2*- yeast strain, A575 containing a chromosomally integrated *Fus1-LacZ* reporter (Sommers and Dumont,

1999). β -galactosidase assays were performed as described previously (Sommers and Dumont, 1999).

6. As a first step towards high level production of fusion proteins for use in crystallization, the *STE2* genes encoding fusion proteins of interest (but without the C-terminally fused GFP) are PCR amplified and inserted via ligation-independent cloning into the expression vector pSGP46 under the control the inducible *ADH2* promoter, as described previously (Clark, et al., 2010). Reading frames are expressed from this vector with C-terminal ZZ (protein A) and RGSHis₁₀ tags that are cleavable by rhinovirus 3C protease. Plasmids encoding the desired constructs are isolated from *E. coli* and transformed into the protease-deficient expression strain BJ5460 (Jones, 1991).
7. For immunoblotting, yeast expressing receptor-T4L fusions from plasmid pSGP46 are cultured in 10 ml SD-ura medium to an OD₆₀₀ of ~3, then diluted 4-fold in YPD (total volume is 40 ml) and re-cultured to an OD₆₀₀ of ~10, depleting glucose in order to induce the *ADH2* promoter. Cells are harvested by centrifuging at 1000 × g for 5 minutes at 4°C, washed with 10 ml water, and resuspended in 300 µl of 20 mM sodium acetate (pH 4.6), 0.1 mM EDTA and 1 mM PMSF. They are then lysed by vortexing at 4 °C together with 0.5 mm zirconia/silica beads (BioSpec Products, Inc; use a volume of beads corresponding to ~250 µl in a microfuge tube) in 10 cycles of 20 seconds each, separated by 2 min of cooling on ice to prevent overheating. After centrifugation at 12,000g for 2 minutes to pellet the beads, the supernatant is transferred to a fresh tube. A total of 8 µl of lysate is then mixed with 22 µl of loading buffer (40 mM Tris-HCl (pH 6.8), 0.1 mM EDTA, 9 M urea, 5% sodium dodecyl sulfate, 0.02 mg/ml bromophenol blue), and incubated at 37°C for 5 min. To prevent irreversible aggregation of TMPs, it is important to avoid overheating the samples at this stage. Following SDS-PAGE, proteins are transferred onto nitrocellulose and probed with an HRP-conjugated anti-RGS-His antibody (Qiagen) or an anti-enolase antibody (for visualizing enolase as a loading control; a gift of Michael Holland of the University of California, Davis) followed by an HRP-conjugated anti-rabbit secondary antibody (Sigma)). Protein bands are visualized using SuperSignal reagent (Pierce) and exposed to X-ray film (CL-X Posure Film by Thermo Scientific).

6. RESULTS FOR THE YEAST α -FACTOR RECEPTOR STE2P

Application of the described procedures for creating a library of variant forms of the yeast α -factor receptor containing T4L inserted into the IC3 loop allowed efficient recovery of the full range of possible receptor variants, with some bias against variants containing insertions with the shortest loop sequences (Mathew, et al., 2013). Based on GFP fluorescence, many of the variant receptors were expressed at levels that were similar to the levels of receptors with no insertions. Cells expressing T4L-containing receptors exhibited up to 80% of the

ligand capacity of cells expressing normal receptors, with similar ligand affinities. Surprisingly, some receptor variants containing the inserted protein exhibited up to 66% of the maximal signaling response of normal receptors, based on induction of the *FUS1-lacZ* reporter, with similar EC₅₀s for activation by pheromone.

A screen was also conducted searching for insertions of T4L into the third loop of receptors with truncations of their cytoplasmic C-terminal tails. C-terminally-truncated receptors are generally fully functional and present at the cell surface in greater numbers than similarly-expressed full-length receptors. While random screening uncovered some C-terminally truncated Ste2p-T4L fusions that bound up to 18% of the high levels of ligand bound by similarly-expressed truncated receptors with no insertion, none of the truncated receptors containing the T4L indertion exhibited more than ~5% of the signaling activity of the unfused receptors. The C-terminal tail may, therefore, provide stabilization to receptors containing perturbations such as inserted proteins (Mathew, et al., 2013). This is an important point to be considered in view of the fact that GPCR constructs used for crystallography involve truncations of the C-terminal to remove regions presumed to have excess flexibility.

The most activeT4L-containing variants of full length Ste2p had contained the inserted protein flanked by nearly complete IC3 loop sequences that were duplicated at both the N- and C-terminal sides of the insertion. We confirmed that duplication of the IC3 loop, even in the absence of any inserted T4L, is not detrimental to receptor signaling. Thus, it seems likely that the presence of extended sequences on both sides of the inserted protein provides the flexibility needed to move T4L out of the way so that the receptor can participate in the interactions with the the cytoplasmic G protein that are required fpr signaling. The extended sequences present in the active constructs may not provide significant rigidity to the surrounding transmembrane helices, but recent successes in structure determination of GPCRs with fusions only at their extreme N-termini (Fenalti, et al., 2014, Siu, et al., 2013, Thompson, et al., 2012, Wu, et al., 2014, Zou, et al., 2012) suggest that such rigidification may not be an important aspect of the fusion protein approach to stabilization.

Acknowledgments

We thank Dr. Fred Naider (CUNY, Staten Island) for providing fluorescent ligand and many helpful discussions. We also thank the University of Rochester Flow Cytometry Core facility for assistance with cell sorting and flow cytometry. This work was supported by National Institutes of Health grants R01GM084083 and U54GM09461 to M.E.D.

References

- Alexandrov A, Vignali M, LaCount DJ, Quartley E, de Vries C, De Rosa D, Babulski J, Mitchell SF, Schoenfeld LW, Fields S, Hol WG, Dumont ME, Phizicky EM, Grayhack EJ. A facile method for high-throughput co-expression of protein pairs. *Molecular & cellular proteomics* : MCP. 2004; 3(9): 934–938. [PubMed: 15240823]
- Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic acids research*. 1990; 18(20):6069–6074. [PubMed: 2235490]
- Bajaj A, eli A, Ding F-X, Naider F, Becker JM, Dumont ME. A Fluorescent α -Factor Analogue Exhibits Multi ple Steps on Binding to Its G Protein Coupled Receptor in Yeast†. *Biochemistry*. 2004; 43(42):13564–13578. [PubMed: 15491163]

- Bajaj A, Connelly SM, Gehret AU, Naider F, Dumont ME. Role of extracellular charged amino acids in the yeast α -factor receptor. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. 2007; 1773(6):707–717. [PubMed: 17433461]
- Brown AJ, Dyos SL, Whiteway MS, White JHM, Watson M-AEA, Marzioch M, Clare JJ, Cousens DJ, Paddon C, Plumptre C, Romanos MA, Dowell SJ. Functional coupling of mammalian receptors to the yeast mating pathway using novel yeast/mammalian G protein α -subunit chimeras. *Yeast*. 2000; 16(1):11–22. [PubMed: 10620771]
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi H-J, Kuhn P, Weis WI, Kobilka BK, Stevens RC. High-Resolution Crystal Structure of an Engineered Human 2-Adrenergic G Protein Coupled Receptor. *Science*. 2007; 318(5854):1258–1265. [PubMed: 17962520]
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi H-J, Kuhn P, Weis WI, Kobilka BK, Stevens RC. High-Resolution Crystal Structure of an Engineered Human β 2-Adrenergic G Protein Coupled Receptor. *Science*. 2007; 318(5854):1258–1265. [PubMed: 17962520]
- Chien EYT, Liu W, Zhao Q, Katritch V, Won Han G, Hanson MA, Shi L, Newman AH, Javitch JA, Cherezov V, Stevens RC. Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. *Science*. 2010; 330(6007):1091–1095. [PubMed: 21097933]
- Chun E, Thompson AA, Liu W, Roth CB, Griffith MT, Katritch V, Kunken J, Xu F, Cherezov V, Hanson MA, Stevens RC. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure (London, England: 1993)*. 2012; 20(6):967–976.
- Clark KM, Fedoriw N, Robinson K, Connelly SM, Randles J, Malkowski MG, DeTitta GT, Dumont ME. Purification of transmembrane proteins from *Saccharomyces cerevisiae* for X-ray crystallography. *Protein expression and purification*. 2010; 71(2):207–223. [PubMed: 20045057]
- Dieckman L, Gu M, Stols L, Donnelly MI, Collart FR. High throughput methods for gene cloning and expression. *Protein expression and purification*. 2002; 25(1):1–7. [PubMed: 12071692]
- Dore AS, Okrasa K, Patel JC, Serrano-Vega M, Bennett K, Cooke RM, Errey JC, Jazayeri A, Khan S, Tehan B, Weir M, Wiggin GR, Marshall FH. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature*. 2014; 511(7511):557–562. [PubMed: 25042998]
- Dowell, SJ.; Brown, AJ. G Protein-Coupled Receptors in Drug Discovery. Leifert, WR., editor. Humana Press; 2009. p. 213-229.
- Egloff P, Hillenbrand M, Klenk C, Batyuk A, Heine P, Balada S, Schlinkmann KM, Scott DJ, Schutz M, Pluckthun A. Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(6):E655–662. [PubMed: 24453215]
- Engel CK, Chen L, Privé GG. Insertion of carrier proteins into hydrophilic loops of the *Escherichia coli* lactose permease. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2002; 1564(1):38–46. [PubMed: 12100994]
- Fenalti G, Giguere PM, Katritch V, Huang XP, Thompson AA, Cherezov V, Roth BL, Stevens RC. Molecular control of delta-opioid receptor signalling. *Nature*. 2014
- Gehret AU, Bajaj A, Naider F, Dumont ME. Oligomerization of the Yeast α -Factor Receptor. *Journal of Biological Chemistry*. 2006; 281(30):20698–20714. [PubMed: 16709573]
- Granier S, Manglik A, Kruse AC, Kobilka TS, Thian FS, Weis WI, Kobilka BK. Structure of the δ -opioid receptor bound to naltrindole. *Nature*. 2012; 485(7398):400–404. [PubMed: 22596164]
- Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, Zhang C, Weis WI, Okada T, Kobilka BK, Haga T, Kobayashi T. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature*. 2012; 482(7386):547–551. [PubMed: 22278061]
- Hanson MA, Roth CB, Jo E, Griffith MT, Scott FL, Reinhart G, Desale H, Clemons B, Cahalan SM, Schuerer SC, Sanna MG, Han GW, Kuhn P, Rosen H, Stevens RC. Crystal Structure of a Lipid G Protein-Coupled Receptor. *Science*. 2012; 335(6070):851–855. [PubMed: 22344443]
- Hollenstein K, Kean J, Bortolato A, Cheng RK, Dore AS, Jazayeri A, Cooke RM, Weir M, Marshall FH. Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature*. 2013; 499(7459):438–443. [PubMed: 23863939]

- Jaakola V-P, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, Ijzerman AP, Stevens RC. The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science*. 2008; 322(5905):1211–1217. [PubMed: 18832607]
- Jones EW. Tackling the protease problem in *Saccharomyces cerevisiae*. *Methods in enzymology*. 1991; 194:428–453. [PubMed: 2005802]
- King K, Dohlman HG, Thorner J, Caron MG, Lefkowitz RJ. Control of yeast mating signal transduction by a mammalian beta 2-adrenergic receptor and Gs alpha subunit. *Science*. 1990; 250(4977):121–123. [PubMed: 2171146]
- Kruse AC, Hu J, Pan AC, Arlow DH, Rosenbaum DM, Rosemond E, Green HF, Liu T, Chae PS, Dror RO, Shaw DE, Weis WI, Wess J, Kobilka BK. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature*. 2012; 482(7386):552–556. [PubMed: 22358844]
- Kruse AC, Ring AM, Manglik A, Hu J, Hu K, Eitel K, Hubner H, Pardon E, Valant C, Sexton PM, Christopoulos A, Felder CC, Gmeiner P, Steyaert J, Weis WI, Garcia KC, Wess J, Kobilka BK. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*. 2013; 504(7478):101–106. [PubMed: 24256733]
- Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S. Crystal structure of the μ -opioid receptor bound to a morphinan antagonist. *Nature*. 2012; 485(7398):321–326. [PubMed: 22437502]
- Mathew E, Bajaj A, Connelly SM, Sargsyan H, Ding F-X, Hajduczuk AG, Naider F, Dumont ME. Differential Interactions of Fluorescent Agonists and Antagonists with the Yeast G Protein Coupled Receptor Ste2p. *Journal of Molecular Biology*. 2011; 409(4):513–528. [PubMed: 21477594]
- Mathew E, Ding F-X, Naider F, Dumont ME. Functional fusions of T4 lysozyme in the third intracellular loop of a G protein-coupled receptor identified by a random screening approach in yeast. *Protein Engineering Design and Selection*. 2013; 26(1):59–71.
- Pausch MH. G-protein-coupled receptors in *Saccharomyces cerevisiae*: high-throughput screening assays for drug discovery. *Trends in biotechnology*. 1997; 15(12):487–494. [PubMed: 9418303]
- Rasmussen SGF, Choi H-J, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VRP, Sanishvili R, Fischetti RF, Schertler GFX, Weis WI, Kobilka BK. Crystal structure of the human β 2 adrenergic G-protein-coupled receptor. *Nature*. 2007; 450(7168):383–387. [PubMed: 17952055]
- Rasmussen SGF, Choi H-J, Fung JJ, Pardon E, Casarosa P, Chae PS, DeVree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI, Kobilka BK. Structure of a nanobody-stabilized active state of the [bgr]2 adrenoceptor. *Nature*. 2011; 469(7329):175–180. [PubMed: 21228869]
- Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah STA, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK. Crystal structure of the [bgr]2 adrenergic receptor-Gs protein complex. *Nature*. 2011; 477(7366):549–555. [PubMed: 21772288]
- Ring AM, Manglik A, Kruse AC, Enos MD, Weis WI, Garcia KC, Kobilka BK. Adrenaline-activated structure of beta2-adrenoceptor stabilized by an engineered nanobody. *Nature*. 2013; 502(7472):575–579. [PubMed: 24056936]
- Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi H-J, Yao X-J, Weis WI, Stevens RC, Kobilka BK. GPCR Engineering Yields High-Resolution Structural Insights into β 2-Adrenergic Receptor Function. *Science*. 2007; 318(5854):1266–1273. [PubMed: 17962519]
- Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SGF, Choi H-J, DeVree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK. Structure and function of an irreversible agonist-[bgr]2 adrenoceptor complex. *Nature*. 2011; 469(7329):236–240. [PubMed: 21228876]
- Scott DJ, Kummer L, Tremmel D, Pluckthun A. Stabilizing membrane proteins through protein engineering. *Current opinion in chemical biology*. 2013; 17(3):427–435. [PubMed: 23639904]
- Sherman F. Getting started with yeast. *Methods in enzymology*. 2002; 350:3–41. [PubMed: 12073320]

- Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Stevens RC, Iwata S. Structure of the human histamine H1 receptor complex with doxepin. *Nature*. 2011; 475(7354):65–70. [PubMed: 21697825]
- Shiroishi M, Kobayashi T, Ogasawara S, Tsujimoto H, Ikeda-Suno C, Iwata S, Shimamura T. Production of the stable human histamine H1 receptor in *Pichia pastoris* for structural determination. *Methods*. 2011; 55(4):281–286. [PubMed: 21903167]
- Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS, Liu W, Lau J, Cherezov V, Katritch V, Wang MW, Stevens RC. Structure of the human glucagon class B G-protein-coupled receptor. *Nature*. 2013; 499(7459):444–449. [PubMed: 23863937]
- Sommers, CM.; Dumont, ME. Genetic approaches for studying the structure and function of G protein-coupled receptors in yeast. Wiley-Liss; New York: 1999.
- Sommers CM, Martin NP, Akal-Strader A, Becker JM, Naider F, Dumont ME. A Limited Spectrum of Mutations Causes Constitutive Activation of the Yeast α -Factor Receptor $\dagger$. *Biochemistry*. 2000; 39(23):6898–6909. [PubMed: 10841771]
- Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, Li T, Ma L, Fenalti G, Li J, Zhang W, Xie X, Yang H, Jiang H, Cherezov V, Liu H, Stevens RC, Zhao Q, Wu B. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*. 2013; 341(6152):1387–1390. [PubMed: 24030490]
- Taslami A, Mathew E, el A, Wessel S, Dumont ME. Identifying Functionally Important Conformational Changes in Proteins: Activation of the Yeast α -factor Receptor Ste2p. *Journal of Molecular Biology*. 2012; 418(5):367–378. [PubMed: 22387470]
- Tate CG. A crystal clear solution for determining G-protein-coupled receptor structures. *Trends Biochem Sci*. 2012; 37(9):343–352. [PubMed: 22784935]
- Thompson AA, Liu W, Chun E, Katritch V, Wu H, Vardy E, Huang X-P, Trapella C, Guerrini R, Calo G, Roth BL, Cherezov V, Stevens RC. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature*. 2012; 485(7398):395–399. [PubMed: 22596163]
- Wacker D, Wang C, Katritch V, Han GW, Huang X-P, Vardy E, McCorvy JD, Jiang Y, Chu M, Siu FY, Liu W, Xu HE, Cherezov V, Roth BL, Stevens RC. Structural Features for Functional Selectivity at Serotonin Receptors. *Science*. 2013; 340(6132):615–619. [PubMed: 23519215]
- Wallin E, Heijne GV. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*. 1998; 7(4):1029–1038. [PubMed: 9568909]
- Wang C, Jiang Y, Ma J, Wu H, Wacker D, Katritch V, Han GW, Liu W, Huang X-P, Vardy E, McCorvy JD, Gao X, Zhou XE, Melcher K, Zhang C, Bai F, Yang H, Yang L, Jiang H, Roth BL, Cherezov V, Stevens RC, Xu HE. Structural Basis for Molecular Recognition at Serotonin Receptors. *Science*. 2013; 340(6132):610–614. [PubMed: 23519210]
- White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, Grisshammer R. Structure of the agonist-bound neurotensin receptor. *Nature*. 2012; 490(7421):508–513. [PubMed: 23051748]
- Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC. Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science*. 2010; 330(6007):1066–1071. [PubMed: 20929726]
- Wu H, Wacker D, Mileni M, Katritch V, Han GW, Vardy E, Liu W, Thompson AA, Huang X-P, Carroll FI, Mascarella SW, Westkaemper RB, Mosier PD, Roth BL, Cherezov V, Stevens RC. Structure of the human κ -opioid receptor in complex with JDTic. *Nature*. 2012; 485(7398):327–332. [PubMed: 22437504]
- Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, Niswender CM, Katritch V, Meiler J, Cherezov V, Conn PJ, Stevens RC. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science*. 2014; 344(6179):58–64. [PubMed: 24603153]
- Xu F, Wu H, Katritch V, Han GW, Jacobson KA, Gao Z-G, Cherezov V, Stevens RC. Structure of an Agonist-Bound Human A2A Adenosine Receptor. *Science*. 2011; 332(6027):322–327. [PubMed: 21393508]

- Zhang C, Srinivasan Y, Arlow DH, Fung JJ, Palmer D, Zheng Y, Green HF, Pandey A, Dror RO, Shaw DE, Weis WI, Coughlin SR, Kobilka BK. High-resolution crystal structure of human protease-activated receptor 1. *Nature*. 2012; 492(7429):387–392. [PubMed: 23222541]
- Zhang J, Zhang K, Gao ZG, Paoletta S, Zhang D, Han GW, Li T, Ma L, Zhang W, Muller CE, Yang H, Jiang H, Cherezov V, Katritch V, Jacobson KA, Stevens RC, Wu B, Zhao Q. Agonist-bound structure of the human P2Y₁₂ receptor. *Nature*. 2014; 509(7498):119–122. [PubMed: 24784220]
- Zhang K, Zhang J, Gao ZG, Zhang D, Zhu L, Han GW, Moss SM, Paoletta S, Kiselev E, Lu W, Fenalti G, Zhang W, Muller CE, Yang H, Jiang H, Cherezov V, Katritch V, Jacobson KA, Stevens RC, Wu B, Zhao Q. Structure of the human P2Y₁₂ receptor in complex with an antithrombotic drug. *Nature*. 2014; 509(7498):115–118. [PubMed: 24670650]
- Zou Y, Weis WI, Kobilka BK. N-terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor. *PloS one*. 2012; 7(10):e46039. [PubMed: 23056231]

Fig. 1. Overall strategy for library construction

Sense and antisense oligonucleotides are annealed to form linkers that have progressive amino acid deletions of the GPCR replacement region. The N and C-terminal linkers have deletions from the ends that are ligated with the N and C-terminal ends of T4L, respectively. Ligation of the linkers and the restriction-digested T4L PCR product results in a library of T4L cassettes that have varied lengths of the GPCR replacement region sequence flanking the T4L gene. These cassettes are then introduced into the GPCR gene that is lacking the replacement region.

Fig. 2.
Construction of the GPCR vector lacking the replacement region.