

GPCR Catalog: A novel method to construct a cDNA / protein library enriched in a family of functional protein

Tai Kubo

National Institute of Advanced Industrial Science and Technology (AIST), Neuroscience Research Institute, Tsukuba, Japan

G-protein coupled receptors (GPCRs) are present in the cell membrane and mediate essential processes in cell signaling; receiving a specific signal molecule from outside of the cell and transducing the signal inside triggering intracellular cascade reactions. As the genome projects are reaching completion, we will soon get a complete list of GPCRs. The GPCR catalog with molecular entities, however, is an acute demand for drug screening, and for studying cell-signal networks.

We have developed a new method to construct a cDNA library, in which representative cDNAs encoding a family of a functional protein, such as G protein coupled receptor, ion channel, transporter, proteinase, proteinase inhibitor, growth factor, or so on are highly enriched. The cDNAs are transcribed and translated to express a family of the functional protein, thus suitable for activity-based screening of a target protein for a drug, a ligand, a modulator or an enzyme substrate. Here I describe the method for preparing GPCR-concentrated cDNA libraries as examples.

Outline of the method

Figure 1 presents a schematic flow of the patented technology (1). One of the key elements of the technology is designing antisense oligonucleotides that may maximally recognize a family of a functional protein. The oligonucleotide is termed here as “ID primer”. cDNAs encoding N-terminal half of the protein family are synthesized by extension from the ID primer. The C-terminal half of the protein, on the other hand, is made up by joining the corresponding portion of certain cDNA belongs to the family. The resulting chimera cDNAs in the library, which cover the entire protein coding region, are transcribed and translated in vitro.

(a) Designing oligonucleotide primers

Antisense oligonucleotide primers are designed on the basis of amino acid sequences conserved among a family of a functional protein of interest. The region should be chosen between a functional region and a carboxy terminus of the protein. cDNAs are synthesized using poly(A)⁺ RNA as a template and the

oligonucleotides as a primer.

(b) Construction of a chimera cDNA library

Ligate the primer-extended cDNAs together with a cDNA containing a carboxy terminal region of a known functional protein of the same family. The reading frame for translation should be the same before and after the joint, so that the chimera cDNAs could be functionally expressed.

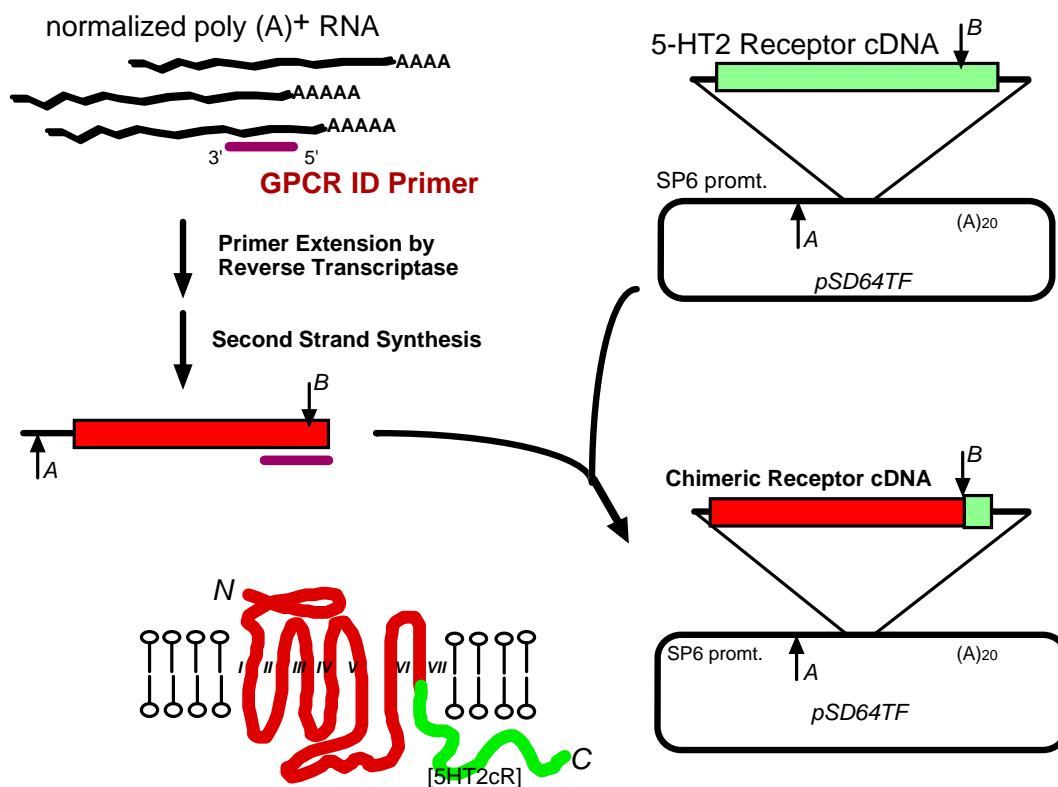


Figure 1 Schematic method flow to construct a GPCR-enriched expression cDNA library. ID primer is designed to anneal with most of the GPCR RNAs. Red represents a library of N-terminal half of the GPCRs, and Green, C-terminal portion of one of the GPCRs, are ligated in frame for translation so that the resulting cDNA to express functional GPCR proteins.

Examples

GPCR chimera cDNA libraries were prepared by setting a GPCR ID primer, characteristic to and conserved among GPCRs, within the transmembrane segment 7. DNA sequence corresponding to the carboxy terminus was introduced from those of the 5HT_{2c} cDNA. The cDNA library thus prepared was transformed and evaluated by functional expression or by sequencing after picking up colonies at random. The libraries contained a series of 5HT receptor subtypes and two orphan receptors. This

GPCR library could be used to identify a receptor that interacts with a specific drug, chemical substance, taste or odorant molecule, a peptide ligand, toxin, and so on.

1. Serotonin receptor library

1.1 Oligonucleotide primers common to serotonin receptors

The amino acid sequences for serotonin (5HT) receptor subtypes, 5HT_{1B} (DDBJ accession number, M89478), 5HT_{1C} (M81778), 5HT_{1D} (M89955), 5HT_{1E} (M91467), 5HT₂ (X57830), 5HT_{2B} (X77307), 5HT_{4B} (Y12505), 5HT₄ (Y08756), 5HT₆ (L41147) and 5HT₇ (L21195), were aligned. A region on the C-terminal side in the seventh predicted transmembrane segment with high homology among these sequences, NP(I/L/F)(I/L/V)Y(T/A/P), was selected, and a mixture of degenerate oligonucleotide primers GPCR-AS for 5HT receptors was synthesized.

GPCR-AS: 5'-gttgactagtGT(A/G)TA(G/A)ATIA(T/A)NGG(A/G)TT-3'

A restriction enzyme SpeI site (underlined lowercase letters) to facilitate the following subcloning procedures and a protective sequence thereof (gttg on the 5' side) were introduced. N represents a mixture of A, T, G and C, and I represents inosine.

1.2 Preparation and normalization of poly(A)⁺ RNA from rat cerebrum

Total RNA was extracted from rat cerebrum using TRIZOL (Invitrogen) and then poly(A)⁺ RNA was prepared using Oligotex-dT30 (Takara). The first strand cDNA synthesis and normalization of poly(A)⁺ RNA were principally followed as described (2).

- (a) 200 µg of poly(A)⁺ RNA and 25 mg of Oligotex-dT30 were annealed.
- (b) The first strand cDNAs were synthesized on the latex by reverse transcriptase.
- (c) 20 µg of poly(A)⁺ RNA was annealed with single-strand cDNA on the latex in the step (b) for 15 minutes at 55 °C. A first supernatant obtained by centrifugation was subjected to a second annealing. A pellet, the latex with single-strand cDNA::poly(A)⁺ RNA of abundant species, was regenerated by heating for 5 minutes at 70 °C. These procedures, annealing, separation and regeneration, were repeated until the fourth supernatant was obtained.
- (d) Normalized poly(A)⁺ RNA in the fourth supernatant was collected by ethanol precipitation and subjected to the following cDNA synthesis.

1.3 cDNA preparation by primer extension

Using normalized poly(A)⁺ RNA as a template and 0.1 nmol of GPCR-AS as a primer, cDNA was synthesized (Stratagene). The cDNA was ligated with a BamHI adaptor (Takara), phosphorylated with polynucleotide kinase, and digested with SpeI.

The resulting cDNA fragments were separated in agarose gel and those of the length around 1 to 2 kb were extracted (Qiagen).

1.4 Amplification of cDNA encoding C-terminus of serotonin 2C receptor

We chose 5HT_{2C} receptor to make up for the missing C-terminal half of the GPCR protein. Referring to human 5HT_{2C} receptor cDNA sequence (M81778), an upstream primer 5HT_{2c}-S1

5HT_{2c}-S1: 5'-GGTGTATACaCTagTCAACAAAA-3

containing a SpeI recognition site (underlined) and a downstream primer 5HT_{2c}-AS1

5HT_{2c}-AS1: 5'-GCTgGTACCGTAGGAAAAGACTG-3'

containing a KpnI recognition site (underlined) were synthesized; lowercase letters are substituted nucleotides to generate SpeI and KpnI recognition sites, respectively.

Human brain cDNA was prepared from the cDNA library (Clontech). The reaction solution is composed of 200 pM each of synthetic DNA primers (5HT_{2c}-S1 and 5HT_{2c}-AS1), 1 ng of template cDNA, 0.2 mM of dNTPs, 0.025 unit/μl of Taq DNA polymerase (ExTaq, Takara), and a buffer accompanying the enzyme. In thermal cycling, a thermal cycler (Perkin-Elmer) was used and a reaction at 94 °C for 2 minutes was followed by a cycle of 94 °C for one minute, 58 °C for one minute, and 72 °C for one minute repeated thirty times. The PCR product was confirmed by 1.2% agarose gel electrophoresis and ethidium bromide staining. Details were in accordance with the method described in (3). As a result,

about 300 bp of DNA was amplified specifically. The DNA product was digested with SpeI and KpnI and inserted into SpeI and KpnI sites of pSD64TF (BH) vector [a vector modified in order to enhance the expression efficiency in *Xenopus* oocyte by adding a multi-cloning site and 5' and 3' untranslated regions of a *Xenopus* β-globin gene to pSP64T (4), provided by Dr. T. Snutch] using

Ligation High (Toyobo)

(pSD64TF-5HT_{2C} R C-term: Fig. 2).

From transformed *Escherichia coli* XL1Blue (Stratagene), plasmid DNA

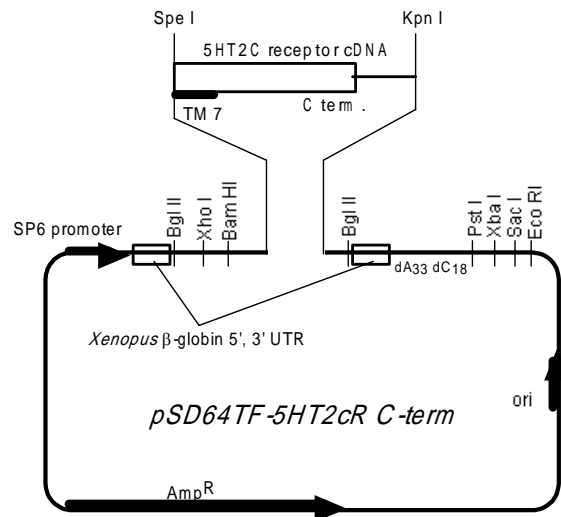


Figure 2 pSD64TF-5HT_{2C} R C-term; 5HT_{2C} receptor cDNA corresponding to the C terminal portion from the transmembrane 7 (TM7) and the 3' noncoding region is joined to the vector pSD64TF. Chimera GPCRs are generated by ligating the N terminal half of the GPCR cDNAs at the SpeI site.

was prepared (Wizard, Promega). The nucleotide sequence of the inserted DNA was analyzed by the reaction using BigDye Terminator Cycle Sequencing FS Ready Reaction Kit (PE Biosystems) and SP6 promoter primer (Promega) and sequencing using a fluorescent automatic sequencer (ABI Sequencer 377). The obtained information on nucleotide sequences was analyzed using Lasergene (DNASTAR), and it was confirmed that there was no difference between the nucleotide sequence of the inserted DNA fragment and a corresponding portion of known nucleotide sequence (M81778).

1.5 Construction of chimera cDNA library of serotonin receptor related protein

A BamHI-SpeI fragment of cDNA (1 to 2 kb), which was extended from the degenerate primer for 5HT receptor obtained in the step 1.3 above, was inserted into the BamHI and SpeI sites of the plasmid pSD64TF-5HT_{2C} R C-term obtained in 1.4. Thus, the 5HT receptor associated cDNA which was extended from the GPCR ID primer was connected to the cDNA encoding the C-terminal side of 5HT_{2C} receptor such that the transcripts produced from the cDNAs are translated in the same protein coding frame. *E. coli* XL1-Blue strain was transformed with the cDNAs to produce a chimera cDNA library of 5HT receptor related proteins.

1.6 Expression screening for cDNA of serotonin-reactive functional proteins

The cDNA library was plated onto an agar medium and about 10,000 colonies were subdivided into 96 pools (about 100 colonies/pool). Plasmid DNA was prepared from each pool and cRNA synthesis was carried out using Cap analog (Pharmacia) and SP6 RNA polymerase (MEGA Script SP6 Kit: Ambion).

cRNA pool thus obtained was injected into the *Xenopus* oocyte and, two to three days later, the current response induced by serotonin was measured [these electrophysiological measuring methods are described in (5)]. Some serotonin-reaction positive pools were further subdivided and a similar procedure was repeated until a single clone was obtained. As a result, at least three types of serotonin-reaction positive clones were obtained. The nucleotide sequences of these clones were determined and analyzed in accordance with a method described in 1.4. As a result of homology search with databases, cDNAs corresponding to the rat 5HT receptor subtypes, i.e., 2A (M30705), 2B (X66842), and 2C (M21410), were isolated. This result indicates that the chimera cDNA library thus prepared could be used for screening a functional protein based on its physiological function.

2. G-protein coupled receptor library

2.1 Preparation of chimera cDNA library of G protein-coupled receptor

- (a) Poly(A)⁺ RNA was prepared from mouse brain in the same manner as described in 1.2 and, using 0.5 µg of poly(A)⁺ RNA as a template, cDNA was synthesized and amplified using SMART PCR cDNA synthesis kit (Clontech). In the first strand synthesis, the primer GPCR-AS shown in 1.1 was used. For amplification of cDNA, a cap site primer in the SMART PCR system and the primer GPCR-AS were used as an upstream and a downstream primer, respectively.
- (b) The amplified cDNA was blunt-ended by T4 DNA polymerase (Blunting Kit; Takara). Addition of BamHI adaptor and phosphorylation were carried out, and after digestion with SpeI, 1% agarose gel electrophoresis and ethidium bromide staining were performed thereby fractionating and extracting DNA of around 1 to 2 kb (same procedures as in 1.3).
- (c) The BamHI-SpeI cDNA fragment (1 to 2 kb) separated in (b) was inserted into BamHI and SpeI sites of vector pSD64TF-5HT_{2C} R C-term (Fig. 2) having C-terminal cDNA of human 5HT_{2C} receptor, using Ligation High. *E. coli* XL1Blue strain was transformed with the plasmid and a chimera cDNA library of mouse brain-derived GPCRs was constructed.

2.2 Isolation of orphan G protein-coupled receptor cDNA

The cDNA library prepared in 2.1 was inoculated onto an agar medium (LB; ampicillin, 75 µg/ml), and 36 colonies were randomly picked out to prepare plasmid DNA. The nucleotide sequence of the inserted DNA was determined in the same manner as used in 1.4. Based on this nucleotide sequence, homology search was carried out. As a result, it was demonstrated that one of the obtained cDNAs encoded a part of an orphan G protein-coupled receptor (M80481) wherein the *in vivo* ligand to the receptor was not known.

2.3 Preparation of cDNA on C-terminal side of orphan G protein-coupled receptor

cDNA was synthesized (BRL) using 2.7 µg of poly(A)⁺ RNA from mouse brain as a template and 162 ng of random nonamer pd(N)₉ (Takara) as a primer. After ligation with BamHI linker (TAKARA) and digestion with BamHI, the cDNA fragment was cloned into a BamHI site of a vector pBluescript II KS- (Stratagene). *E. coli* XL1-Blue was transformed with the cDNA and the resulting cDNA library was amplified on agarose medium. Colonies were collected and plasmid DNA was prepared from a portion of the *E. coli* suspension, and used as a template in the following PCR. As a

sense strand primer, the following S1 and S2 were designed referring to the nucleotide sequence of the orphan G protein-coupled receptor cDNA, identified in the step 2.2.

S1: 5'-TGCCACAGAGCAGTACCTC-3'

S2: 5'-CGTGAAGATGCTGGTGCTTG-3'

As an antisense strand primer, T7 (T7 promoter primer, Nippon Gene) in the vector pBluescript II KS was used. The reaction solution was composed of 200 pM each of S1 and T7, 150 ng of template cDNA, 0.2 mM of dNTPs, 0.025 unit/ μ l of Taq DNA polymerase (ExTaq), and a buffer accompanying the enzyme, and the total reaction volume was 20 μ l. In thermal cycling, a thermal cycler (Perkin-Elmer) was used and a reaction at 94 °C for 2 minutes was followed by a cycle of 94 °C for one minute, 58 °C for one minute, and 72 °C for one minute repeated five times. Thereafter, 2 μ l of the reaction solution was subjected to a second PCR using the nested primers S2 and T7. PCR was carried out under the same conditions as for the first PCR except that 200pM each of S2 and T7 were employed and the thermal cycling was repeated 25 times. In accordance with the instructions of TOPO TA cloning kit (Invitrogen), 1 μ l of the reaction solution was subcloned into PCR2.1-TOPO vector. The subclone was introduced into *E. coli* TOPO10 (Invitrogen) and the resulted transformant was subjected to sequencing as in 1.4. As a result, cDNA which overlapped with the orphan G protein-coupled receptor cDNA isolated in 2.2 and encoding its C-terminus was obtained. Thus, a cDNA nucleotide sequence containing the entire coding region of the orphan G protein-coupled receptor and the primary structure of the receptor protein deduced therefrom were clarified. The nucleotide sequence is identical to a known nucleotide sequence (M80481) except the following two positions; T → C at nucleotide residue 323 and G → A at the residue 475).

Utilization of the library

The libraries described above have the following advantages compared to those prepared by existing methods: The library is subjected to screening either by sequence or by biochemical and biophysical properties of protein, and screening efficiency is greatly improved. As the relevant proteins are highly enriched, the protein/peptide library, which is provided in solution, or fixed on a chip or micro/nano spheres, can be screened based on their biological and biophysical activities (Fig. 3). Furthermore, the library is normalized and highly enriched, and thus the RNA species that are too rare to be detected so far could be unveiled now.

The ID primer for a certain family of functional protein could be designed from conserved amino acid sequences, preferably located near the carboxy terminus. Once

the ID primers are accordingly prepared, the method to construct enriched cDNA libraries is applicable to any functional protein families. In fact the cDNA libraries for serine proteinase inhibitors, K⁺ channels and growth factors are available now.

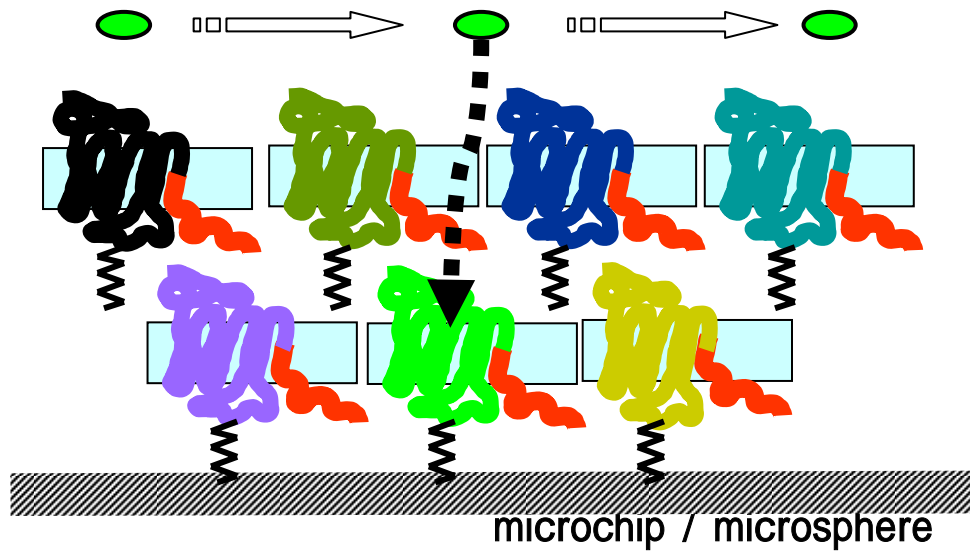


Figure 3 A GPCR catalogue embedded in biomembrane or lipid bilayer is displayed on microchip or microsphere. Ligand molecules to examine (green) are applied to the system and interactions with each receptor are monitored by biochemical or biophysical techniques.

REFERENCES

1. JPN Patent No.: 3243531 (Date of Patent: 10/26/2001, Filed 12/28/1999)
2. Sasaki, YF, Iwasaki, T, Kobayashi, H, Tsuji, S, Ayusawa, D, Oishi, M, "Construction of an equalized cDNA library from human brain by semi-solid self-hybridization system.", *DNA Research* 1: 91-96 (1994)
3. Sambrook, J, Russel, D., "*Molecular Cloning: A Laboratory Manual*", 3rd edn.", Cold Spring Harbor Laboratory Press (2001)
4. Krieg PA, Melton DA., "Functional messenger RNAs are produced by SP6 in vitro transcription of cloned cDNAs.", *Nucleic Acids Res.* 12: 7057-7070 (1984)
5. Julius D, MacDermott AB, Axel R, Jessell TM, "Molecular characterization of a functional cDNA encoding the serotonin 1c receptor.", *Science* 241: 558-564 (1988)