In Vitro Selection of High Affinity DNA-Binding Protein Based on Plasmid Display Technology

CHOI, YOO SEONG¹, HYUN JOO², AND YOUNG JE YOO*¹

¹School of Chemical and Biological Engineering, Seoul National University, Seoul 151-742, Korea
²Department of Molecular Physiology, College of Medicine, Inje University, Busan 633-165, Korea

Received: January 24, 2005
Accepted: February 25, 2005

Abstract Based on plasmid display technology by the complexes of fusion protein and the encoding plasmid DNA, an in vitro selection method for high affinity DNA-binding protein was developed and experimentally demonstrated. The GAL4 DNA-binding domain (GAL4 DBD) was selected as a model DNA-binding protein, and enhanced green fluorescent protein (EGFP) was used as an expression reporter for the selection of target proteins. Error prone PCR was conducted to construct a mutant library of the model. Based on the affinity decrease with increased salt concentration, mutants of GAL4 DBD having high affinity were selected from the mutant protein library of protein-encoding plasmid complex by this method. Two mutants of (Lys33Glu, Arg123Lys, Ile127Lys) and (Ser47Pro, Ser85Pro) having high affinity were obtained from the first generation mutants. This method can be used for rapid in vitro selection of high affinity DNA-binding proteins, and has high potential for the screening of high affinity DNA-binding proteins in a sequence-specific manner.

Key words: Plasmid display, in vitro selection, GAL4, DNA-binding protein, protein screening

DNA-binding proteins have a central role in several different biological activities such as replication, regulation of transcription, and repair of damaged DNA. The change of DNA-binding properties affects the protein DNA recognition, and induces different biological activities. The DNA-binding properties of proteins appear to be governed by the nonbonded atomic interactions between nucleotides and amino acids, van der Waals interactions, and hydrogen bonds, and some substitutions of the amino acid residues have an effect on the binding property and stability [6]. A phage display system was used to alter the sequence specificity of the zinc finger motif [18]. Affinity selective isolation of ligands from peptide libraries through display on a lac repressor and affinity selection of single chain Cro by phage display has also been constructed to isolate novel DNA-binding proteins [3, 15]. The enrichment of DNA-binding proteins from a genomic or cDNA library has been achieved in a sequence-specific manner [8]. However, in vitro selection of high affinity DNA-binding mutant proteins on a target protein has not yet been reported.

In vitro selection of biological macromolecules by display technologies has become a tool for enriching molecular diversity and producing novel types of proteins [13, 19]. In vitro selection of proteins has been achieved, based on the linkage between each protein and its encoding nucleic acid. There are currently a number of technologies available for discovery, and they directed evolution of new molecules in biological and nonbiological aspects such as cell surface display, phage display, ribosome display, mRNA display, and plasmid display [12, 14, 16]. In particular, plasmid display is an in vitro selection approach, based on protein-plasmid complexes, in which proteins are expressed and folded in vivo. Plasmid display neatly avoids the potential difficulties of other display systems such as cDNA size limitation, mRNA stability, and high throughput readiness [20]. The fusion proteins, including DNA-binding domain, are expressed in vivo, and the proteins bind to the specific DNA sequence on the encoding plasmids by sequence specific protein-DNA interaction. The complexes of fusion proteins and the encoding plasmid DNA can be used for in vitro selection from protein library on cell lysis. This technique was first validated experimentally by Cull et al. [4] who fused short peptides to the carboxyl-terminus of the lac repressor protein, and was also constructed to enrich a target protein from genomic library by Speight et al. [20] using DNA-binding protein nuclear factor κB (NF-κB). However, the success of plasmid display relies on the maintenance of the protein-DNA interaction, and the high

*Corresponding author
Phone: 82-2-880-7411; Fax: 82-2-887-1659;
E-mail: yjyoo@snu.ac.kr
binding affinity of DNA-binding proteins in the display is very important.

Here, we describe a simple in vitro selection method of high affinity DNA-binding proteins, based on a plasmid display system, in which high affinity DNA-binding mutants of a target protein were screened (Fig. 1). Because GAL4 DBD binds the 17-bp upstream activating sequence in a sequence-specific manner and has a homodimeric structure like many other trans-acting DNA-binding factors [21], the GAL4 DNA-binding domain (N-terminal 147 amino acid residues) was selected as a model DNA-binding protein, and enhanced green fluorescent protein (EGFP) was used as an expression reporter of target proteins [22]. These results suggest that the method is sufficiently sensitive to select high affinity GAL4 DBD mutants, and the technology can be used to discover high affinity DNA-binding proteins from large libraries by relatively simple steps.

**Materials and Methods**

**Strains and Cultivation**

*E. coli* DH5α and *E. coli* BL21 (DE3) were used as host cells for gene cloning and expression experiments, respectively. Luria-Bertani (LB) medium and LB medium, which were supplemented with 50 µg/ml ampicillin and 50 µg/ml kanamycin into LB medium, respectively, were used for the growth and selection of *E. coli* transformants. All *E. coli* strains were grown at 37°C and 200 rpm.

**Genetic Manipulation**

Standard molecular biology techniques were used, unless otherwise stated. PCR amplification was performed with Taq DNA polymerase (Stratagene, LaJolla, CA, U.S.A.), and purification of all DNA fragments was conducted with QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) and QIAEX II Gel Extraction Kit (Qiagen). All ligations were carried out with T4 DNA ligase (Bioneer, Daejon, Korea), and all primers and oligonucleotides were obtained from Bioneer, Korea.

The plasmid vector pEGU for this study was prepared as follows. Genes encoding GAL4 DBD and EGFP were obtained from plasmid pGBK7 (Clontech, Heidelberg, Germany) and pEGFP-1 (Clontech) by PCR amplification, respectively. *NdeI* and *EcoRI* restriction sites were introduced into the 5’ and 3’ of the GAL4 DBD fragment using the following primers: 5’-GCC GCC CAT ATG GCT AGC AAT CTA CTG TCT TCT A TC-3’ and 5’-GGC GGC CGA ATT CGA TAC AGT CAA GGG CGA GGA GCT -3’, respectively. *EcoRI* and *HindIII* restriction sites were introduced into the EGFP fragment by the primers: 5’-GCC GGC GAA TTC GGT GAG CAA GGG CGA GGA GCT-3’ and 5’-GGC GGC AAC CTT CTT G T A CAG CTC G TC CA T GCC-3’, respectively. The amplified sequences were identified by sequencing.

To insert 17-bp GAL4 binding sequence (5’-CGGG TG-ACAGCCCTCCG-3’) with the vector, two nucleotides of sense and antisense were prepared as follows: 5’-GAT CIT CGA AAT TAA TAC GAC TCA CTA TAG GGC GGG TGA CAG CCC CCC GAA CCC CT-3’ and 5’-CTA GAG GGG TTC GGA GGG CTG TCA CCC GCA CTA TAG TGA GTC GIA TTA ATT TCG CA-3’. Two fragments were mixed to 20 pmol/µl in annealing buffer (200 mM Tris buffer, pH 8.0, 0.1 M EDTA, and 400 mM NaCl) for annealing of the oligonucleotides. The solution was heated to 90°C for 3 min and slowly cooled to below 45°C over 2 h. The annealed oligonucleotides had BglII

![Fig. 1. Schematic representation of plasmid display library construction and in vitro selection.](image-url)
and XbaI restriction sites with the ends of the sequence, respectively.

DNA fragments of origin of replication and ampicillin were obtained by cutting pUC19 vector by PstI and SspI restriction enzymes, and multiple cloning sites and promoter region were also obtained by cutting pET23b vector by the same enzymes. A plasmid vector (pEU23) that had pUC replication origin, ampicillin resistance, and the cloning/expression region of pET23b was constructed through the ligation of two DNA fragments. The annealed oligonucleotide, GAL4 DBD gene, and EGFP gene were introduced to the vector by the restriction and ligation with BglII/XbaI, NdeI/EcoRI, and EcoRI/HindIII sites, respectively. Finally, the vector, pEGU, for this study was obtained.

Protein Expression and Purification
The recombinant plasmid was introduced into E. coli BL21 (DE3), and the E. coli cells bearing the plasmid were grown in LA medium at 37°C and 200 rpm and induced with 0.4 mM IPTG and 20 μM ZnSO₄. The cells were harvested by centrifugation and resuspended in a lysis buffer (100 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 20 μM ZnSO₄, 0.05% Tween-20, 10 mM Imidazole) after 20°C, 200 rpm, and 20 h post-induction. The cells were then disrupted by sonication, and the supernatants were prepared for protein purification. The cell supernatant was applied to the Ni-NTA resin (Qiagen) and equilibrated in the lysis buffer at 4°C for 1 h. The lysate-Ni-NTA mixture was loaded into a mini-column, and washed with 4 column volumes of a buffer (100 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 20 μM ZnSO₄, 0.05% Tween-20, 20 mM imidazole). The fusion protein was eluted with the elution buffer (100 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 20 μM ZnSO₄, 0.05% Tween-20, 250 mM imidazole). The final purity of the proteins was assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), and protein concentrations were determined using the modified Lowry reagent (Sigma, St. Louis, MO, U.S.A.).

Electrophoretic Mobility Shift Assay (EMSA)
A 50-bp binding site probe that included the 17-bp (5’-CGGGTGACAGCCCTCCG-3’) ATP and T4 polynucleotide kinase (Promega). Labeled probe was purified by ProbeQuant™ G-50 Micro Columns (Amersham Biosciences). EMSA reaction mixture contained 50 mM HEPES (pH 7.5), 50 mM NaCl, 5 mM MgCl₂, 20 μM ZnSO₄, 10% (v/v) glycerol, and 2 μg/ml salmon sperm DNA, and the reactions were carried out at 4°C for 20 min and at room temperature for 30 min, based on the method described elsewhere [10]. The protein-DNA complexes were resolved on a 5% nondenaturing polyacrylamide gel, containing 0.5 × triborate (TB), at 150 V for 2 h, visualized with Typhoon 8600 image analyzer and Phoretix™ 2D software (Nonlinear Dynamics Ltd.).

Mutant Library Preparation
Mutant genes of GAL4 DBD were prepared by error-prone PCR with the following primers: 5’-TCC CGC GAA ATT AAT ACG AC-3’ and 5’-TGG TGC AGA TGA TCA GG-3’. Mutagenesis reaction was conducted by Diversity™ PCR random mutagenesis Kit (Clontech) according to the manufacturer’s instruction. Mutant genes were prepared by the restriction at NdeI and EcoRI sites. Plasmid vectors for mutants were constructed by the insertion of mutant genes to the N-end of EGFP of NdeI-EcoRI-digested pEGU.

Fluorescence Spectrometry
Fluorescence of GAL4 DBD/EGFP fusion protein was measured by spectrofluorometry (Shimazu, RF-5301PC). Solution including proteins was excited at 488 nm and emission was collected at 508 nm.

Protein Library Preparation and In Vitro Selection of Target Protein
All mutant proteins were expressed in E. coli BL21 (DE3) and induced with 0.4 mM IPTG, 20 μM ZnSO₄ at 20°C, and 200 rpm for 20 h. The cells were harvested by centrifugation and resuspended in a binding buffer (20 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 20 μM ZnSO₄, 0.05% Tween 20), and lysozyme was added to 0.2 mg/ml and the suspension was incubated at 37°C and 200 rpm for 20 min. The cell supernatant was applied to the Ni-NTA resin (Qiagen) for recovering the protein-DNA complexes and equilibrated in the lysis buffer at 4°C for 1 h. The lysate-Ni-NTA mixture was washed five times with the binding buffer of four-fold volume of the supernatant. Then, additional washings were carried out by increasing the NaCl concentration of the binding buffer by nonlinear gradient from 200 mM NaCl to 1 M NaCl. Wash fractions were collected, and the genes in the fractions were transformed into E. coli BL21 (DE3). When proteins in colonies were expressed, colonies with green light were selected and proteins in the green light colonies were purified by the Ni-NTA affinity purification method described above.

Results and Discussion
Firstly, plasmid for expressing fusion protein of wild-type GAL4 DBD (1-147) and EGFP under the control of T7 promoter was constructed with a C-terminal hexa-histidine sequence in order to select mutant proteins with high affinity in vitro (Fig. 2). The vector allowed us to insert the GAL4 DNA-binding domain at the N-end of EGFP, and pUC replication origin was introduced to obtain plasmid DNA of high copy number. A 17-bp GAL4 upstream binding
sequence (5'-CGGG TGACAGCCCTCCG-3') was incorporated into the vector in order to create a phenotype-genotype linkage for individual fusion proteins, which was not shown in the other location of the plasmid.

The fusion protein was expressed in *E. coli* strain BL21 (DE3). To efficiently produce the fusion protein, the expression was optimized with various temperatures, IPTG concentrations, and induction times (data not shown). The fusion protein was expressed at 20°C, 0.4 mM IPTG, and 20 h induction, and was purified by affinity purification on Ni-NTA resin. When the final purity of the proteins was assessed by SDS-PAGE, the proteins were more than 90% homogeneous (Fig. 3A). We also attempted to purify the fusion protein by 1 M NaCl-containing washing buffer. However, no elution of the proteins by high salt concentration of 1 M NaCl was detected in the washing step, and the proteins were obtained only by the elution buffer (Fig 3B). The result also demonstrated that the affinity between C-terminal hexa-histidine of the fusion protein and Ni-NTA resin was conserved at high salt concentration.

In order to examine the ability of the fusion protein to bind the DNA, electrophoretic mobility shift assay (EMSA) was performed using a probe that contained the 17-bp GAL4 DBD-binding sequence. Thus, the fusion protein was purified from DNaseI-treated cell lysate. Before EMSA analysis, the concentration of purified protein was first determined using the modified Lowry reagent (Sigma), and was controlled by fluorescence of EGFP in detail.

Titration of binding reaction of increasing amount of the fusion protein with 20 fmol/µl DNA fragment was conducted. As expected, the fusion protein bound to the probe (Fig. 4).

In order to select high affinity mutants of GAL4 DBD *in vitro*, a mutant library of GAL4 DBD was constructed using error-prone PCR according to the strategy suggested in this study. The error rate in the PCR amplification was found to be three to five changes per 1,000 bases, based on the sequence information on the first generation variants. The plasmids including mutant genes were transformed to *E. coli* BL21 (DE3) with the first generation gene library to generate a mutant library. Because it has been reported that the affinity decreases with increasing salt concentration [2, 7, 9], enrichment experiments of high affinity mutants were performed by increasing the NaCl concentration in the washing solution after Ni-NTA affinity binding of expressed proteins. After additional washings were carried out by increasing the NaCl concentration, the wash fractions...
were collected, and the genes in the fractions were transformed into *E. coli* BL21 (DE3). When the colony number in each fraction was counted, the number decreased in isocratic elution fractions as the number of times of washing increased. However, when the buffer was changed with higher NaCl concentration, the colony number increased because of the loss of plasmid caused by weakening of the protein-DNA interaction, and then was also decreased by the isocratic elution (data not shown). Finally, 20 colonies from the fraction of 1 M NaCl concentration were chosen, and proteins in colonies were respectively expressed. Five colonies, which have whole cell fluorescence by EGFP, out of 20 colonies of *E. coli* cells were selected as potential targets, based on the consensus that the fluorescence of *E. coli* cells expressing such GFP fusions is related to the correct folding of the upstream protein domains expressed alone [22]. Proteins in the green light colonies were then purified by Ni-NTA affinity chromatography. Finally, (Lys33Glu, Arg123Lys, Ile127Lys) and (Ser47Pro, Ser85Pro) mutants were obtained, and confirmed by sequencing of the selected colonies. The affinity of wild-type GAL4 DBD fusion and the mutants was compared by titration of increasing amounts of three proteins with the 50-bp binding site probe (Fig. 5). The dissociation constant of wild-type GAL4 DBD was found in the nanomolar (10^-9) range [18, 21]. When the fractions of protein-DNA complex were determined by quantitation of the gel using image analysis software, the results showed that the mutant (Lys33Glu, Arg123Lys, Ile127Lys) had about two-fold increase in binding affinity (Fig. 5B), and the mutant (Ser47Pro, Ser85Pro) had a slightly higher affinity (Fig. 5C) than wild-type GAL4 DBD (Fig. 5A). High affinity GAL4 DBD mutants from the first generation mutants could be selected directly via functional selection by the method proposed in this study.

In this study, we have developed a new method for in vitro selection of high affinity DNA-binding proteins, based on the plasmid display system by a linkage of the phenotype to its encoding DNA, which included in vivo protein expression and in vitro target selection based on protein-DNA-complex formation. Even though the display has been applied for in vitro mutant protein selection of GAL4 DBD, based on salt dependence of DNA-binding affinity [2, 7, 9], many other DNA-binding proteins could also be used as targets for the screening of high affinity mutants. This method is technically simple and includes advantages of plasmid display. It does not require high downstream efforts, and it can easily select and identify the high affinity DNA-binding mutants from a library, compared with over-colony screening approaches [1, 11]. Because the genetic information of high affinity mutants is carried on plasmid, faster in vitro evolution can also be allowed.

Based on the selection strategy based on the plasmid display technique, novel DNA-binding proteins with high affinity can be generated. These proteins can function in living cells as transcriptional activators or repressors; therefore, it is possible to find broad application in highly specific control of gene expression and future gene therapy strategies. Moreover, considering that a key in designing a plasmid display system is the fact that the integrity of the
protein-DNA interaction must be maintained during selection processes [20], this strategy could be possible to easily screen adapters for plasmid display, DNA-binding proteins, which have DNA-binding property of high affinity in soluble form. In addition, the high affinity DNA-binding proteins could also be used for a protein microarray, based on a DNA chip platform by specific high affinity of protein-DNA interaction [3].

In conclusion, a method based on the plasmid display system was used for the selection of DNA-binding mutants of high affinity, and the results demonstrate that this tool is useful and has high potential for the selection of DNA-binding proteins of high affinity in a sequence-specific manner. Therefore, it could be a powerful tool for screening DNA-binding protein, based on display technology.

Acknowledgment

This work was supported by a grant No. R01-2003-000-10261-0 from the Basic Research Program of the Korea Science and Engineering Foundation.

REFERENCES