Rapid Communication

J. Biochem. **135**, 285–288 (2004)
DOI: 10.1093/jb/mvh034

DNA Display of Biologically Active Proteins for *In Vitro* Protein Selection

Masato Yonezawa[1,2], Nobuhide Doi[1], Toru Higashinakagawa[2] and Hiroshi Yanagawa[*,1]

[1]Department of Biosciences and Informatics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522; [2]Department of Biology, School of Education, Waseda University, 1-6-1 Nishi-waseda, Shinjuku-ku, Tokyo 169-8050

Received December 20, 2003; accepted January 9, 2004

***In vitro* display technologies are powerful tools for screening peptides with desired functions. We previously proposed a DNA display system in which streptavidin-fused peptides are linked with their encoding DNAs *via* biotin labels in emulsion compartments and successfully applied it to the screening of random peptide libraries. Here we describe its application to functional and folded proteins. By introducing peptide linkers between streptavidin and fused proteins, we achieved highly efficient (>95%) formation of DNA-protein conjugates. Furthermore, we successfully enriched a glutathione-*S*-transferase gene by a factor of 20–30-fold per round on glutathione-coupled beads. Thus, DNA display should be useful for rapidly screening or evolving proteins based on affinity selection.**

Key words: biotin, DNA display, *in vitro* translation, protein engineering, streptavidin.

Abbreviations: aa, amino acid; SA, streptavidin; GST, glutathione-*S*-transferase; BLIP, β-lactamase inhibitory protein; RYBP, Ring1 and YY1 binding protein.

Determining the functions of proteins is an important task in the process of applying data from the human genome project to achieve new advances in therapeutics, diagnosis, and medicine. One approach is to search for interaction partners, including ones for protein-protein, small molecule-protein, and nucleic acid-protein interactions. So far, several methods have been used. The yeast two-hybrid approach has been a leading technology for identifying protein interactions (reviewed in Ref. *1*). However, it has significant limitations that arise because the interactions take place in a living system. Mass spectrometric identification of proteins after affinity chromatography is also a powerful technology (reviewed in Ref. *2*). However, this approach often fails after laborious and time-consuming efforts to afford sufficient amounts of purified proteins due to low-level expression or inadequate purification. Protein microarrays (Ref. *3* and references therein) are another alternative for systematically identifying protein-protein interactions and drug receptors. However, they require the preparation of a large number of proteins and the use of special equipment such as microarrayers.

Display technologies that link genotype and phenotype molecules comprise a powerful alternative [reviewed in (*4–6*)]. Large libraries can be screened iteratively by amplifying selected genes in host cells or PCR, so even very low copy number proteins can be identified. Phage display (*7*) is the most widely used display technology, though its use has been significantly hampered by the limitations of producing libraries in a living system. As proteins are expressed as fusions to phage coat proteins in bacterial cells, those that are toxic or that abrogate the assembly of virion particles are not displayed. Indeed, only proteins containing less than 200 amino acids (aa) have been screened from phage-displayed cDNA libraries (*8–11*). Previously, we developed an *in vitro* DNA display system called 'STABLE' (*12*), which allows the complete *in vitro* construction of DNA-displayed peptide libraries. Recently, we showed that short linear peptides that bind to a monoclonal antibody could be screened from a random peptide library (*13*). In the present study, we have displayed a folded, functional protein on linear DNA and selectively enriched a glutathione-*S*-transferase (GST) gene by affinity purification with glutathione-immobilized beads.

In vitro compartmentalization utilizing water-in-oil emulsions was originally developed by Tawfik and Griffiths to cage a single gene per micelle and to select catalytically active proteins (*14*). In our DNA display strategy (Fig. 1), proteins are expressed as streptavidin (SA)-fusion proteins in order to conjugate the proteins to their encoding DNAs with biotin labels in emulsion compartments. The resulting DNA-protein conjugates are selected by affinity purification followed by PCR amplification.

Since the conjugation of a protein to its encoding DNA is a crucial step in DNA display, we examined whether or not folded proteins fused to SA can be efficiently conjugated to a biotinylated DNA (Fig. 2). When a BLIP gene (encoding a 165-aa protein, Ref. *15*) was directly fused to the SA gene and expressed as a fusion protein, about half of the input DNA was not conjugated to the protein (Fig. 2B, lanes 1 and 2). To reduce steric hindrance, which may affect the formation of DNA-protein conjugates, peptide linkers were introduced between SA and BLIP (Fig. 2A). We chose a ~22-aa linker containing a hemagglutinin-tag (HA), a ~50-aa linker containing a hemagglutinin-tag and a glycine/serine-rich sequence (GS), which is often

[*]To whom correspondence should be addressed. Tel: +81-45-566-1775, Fax: +81-45-566-1440, E-mail: hyana@bio.keio.ac.jp

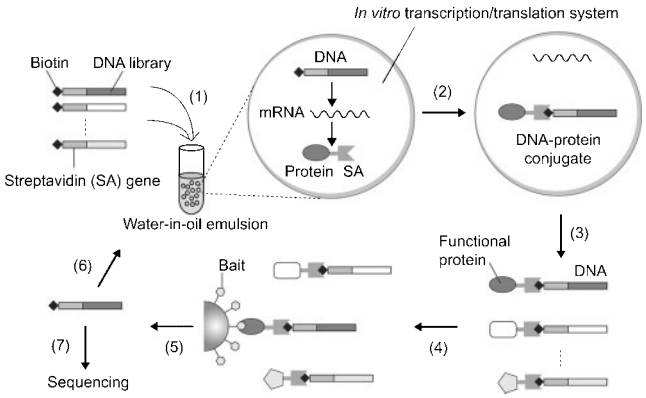

Fig. 1. **A schematic representation of the DNA display selection procedure.** (1) A DNA library encoding SA-fusion proteins is labeled with biotin and compartmentalized in a water-in-oil emulsion containing an *in vitro* transcription/translation system. (2) In each compartment, SA-fusion proteins are synthesized and attached to the template DNA *via* biotin labels. (3) DNA-protein conjugates are recovered from the emulsion and (4) subjected to affinity selection on an immobilized bait. (5) After washing and elution, the DNA portion of the bound molecules is amplified by PCR. (6) DNA is subjected to the next round of selection or (7) identified by sequencing. Although SA-fusion proteins form a tetramer and thus four copies of protein can be displayed, only one copy of the protein is shown in this figure for simplicity.

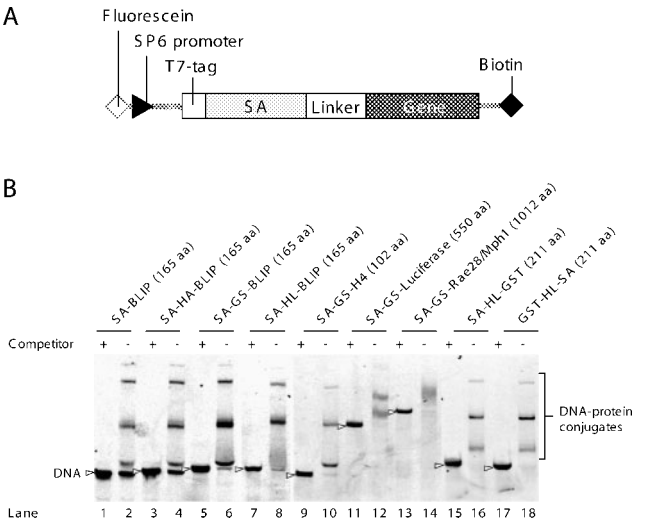

Fig. 2. **Formation of DNA-protein conjugates.** (A) A schematic representation of DNA constructs for *in vitro* transcription/translation. DNA constructs were labeled by PCR with fluorescein at the upstream ends and with biotin at the downstream ends using labeled primers as described (*13*). Plasmids were constructed by modifying pSta4 (*13*) using standard subcloning techniques (*28*). The translated open reading frame consists of sequences for a T7-tag, streptavidin (SA), a peptide linker, and a fused protein (gene). The amino acid sequences of the linker portions are YPGSQLYPYDVPDYASLGGHMA (22 residues, termed HA), YPGS$(GGGGS)_5$GGGRSQLYPYDVPDYASLGGHMA (52 aa, termed GS) or SA$(EAAAK)_4$ARSQLENLYFQGGS (36 residues, termed HL). DNA fragments coding for BLIP and Rae28/Mph1 were amplified by PCR from plasmids carrying these genes (*29*, *30*). GST and luciferase genes were amplified from pGEX-4T-3 (Amersham) and TNT control template (Promega), respectively. Histone H4 and RYBP genes were amplified from a mouse testis cDNA library (Clontech). All DNA sequences were confirmed with a CEQ2000 sequencer (Beckman Coulter). (B) *In vitro* transcription/translation was performed using a TNT-SP6 coupled wheat germ extract (Promega) with 10 nM DNA templates in the presence (+) or absence (–) of 0.1 mM free biotin competitor. The reaction products (3 μl) were separated in 2% agarose gels containing 0.1% SDS prepared with Seakem Gold agarose (BioWhittaker Molecular Applications). DNA (open arrowheads) and DNA-protein conjugates were detected *via* fluorescence using Molecular Imager FX (Bio-Rad). DNA-protein conjugates exhibited lower mobility than free DNA. The multiple bands of DNA-protein conjugates originate from binding of different numbers of DNA molecules to a tetramerized SA-fusion protein in non-emulsified reactions. The size of each displayed protein excluding the SA and linker portions is indicated in parenthesis.

used in ribosome display, and a ~40-aa linker containing EAAAK repeats (HL), which forms a helical structure and effectively separates the domains of bifunctional proteins (*16*). The insertion of peptide linkers improved the efficiency of formation of DNA-protein conjugates. The efficiency reached ~70% when HA was introduced (Fig. 2B, lanes 3 and 4) and >95% when GS or HL was introduced (Fig. 2B, lanes 5–8).

While the total yields of synthesized proteins were almost constant, the proportion of tetrameric protein relative to total protein increased when peptide linkers were introduced (data not shown). This effect was largely dependent on the length of the linkers. These results suggest that tetramerization of SA-fusion proteins, which is required for tight association with biotin (*17*), was enhanced when relatively long peptide linkers were introduced and that it allowed highly efficient formation of DNA-protein conjugates.

Furthermore, proteins other than BLIP (~20 kinds) whose sizes were in the range of 102–1012 aa were fused to SA through GS or HL. We also examined C-terminal as well as N-terminal SA fusions. All proteins showed the formation of DNA-protein conjugates with an efficiency of >95% (Fig. 2B, lanes 9–18, and unpublished data).

It is unclear how large proteins can be used for screening with other *in vitro* display technologies based on cell-free translation, such as ribosome display (*18–20*) and mRNA display (*21–23*). In ribosome display, scFv (single chain antibody fragments, ~300 aa) from large libraries and folded proteins of up to 400-aa in length in model systems have been selected (*19*, *20*, *24–26*). In mRNA display, the sizes of proteins screened so far have been restricted to ~100 aa (*21*, *22*). In our DNA display system, larger proteins can potentially be selected.

Next, we addressed whether or not the GST gene can be enriched with our DNA display strategy. DNA templates for SA-fusion of GST or RYBP, as a negative control, were mixed in a ratio of 1:10 or 1:1,000, and then used for affinity selection on glutathione-coupled beads. As shown in Fig. 3A, the DNA band corresponding to GST was enriched after selection. By conducting two rounds of selection from the 1:1,000 mixture, GST was enriched to a detectable level (Fig. 3A, lanes 3–5). To confirm the enrichment efficiency, each DNA sample was subjected to quantitative real-time PCR (Fig. 3B). From the 1:1,000 mixture, the GST gene was enriched ~700-fold after two selection rounds, which corresponds to 20–30-fold per round. These results indicate that our DNA display system is applicable to the selection of folded proteins.

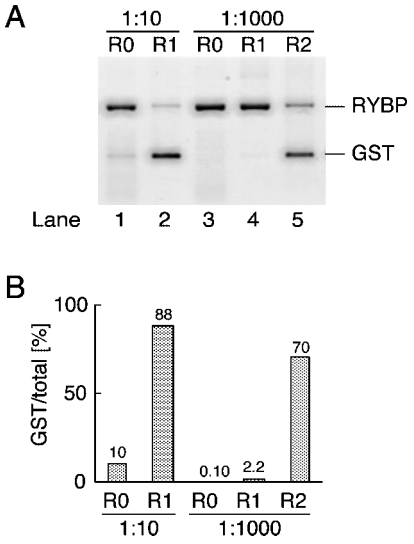

Fig. 3. **Affinity enrichment of GST.** (A) The DNA construct encoding GST-HL-SA (1,388 bp) was mixed with that for RYBP-HL-SA (1,406 bp) in a ratio of 1:10 or 1:1,000 to be subjected to the DNA display selection procedure. *In vitro* transcription/translation in an emulsion and recovery of DNA-protein conjugates were performed as described (*13*). The DNA-protein conjugates were incubated with 25 μl of glutathione-Sepharose 4B (Amersham) on a rotator for 1 h at 4°C. After extensive washing with PBST (137 mM NaCl, 8.1 mM NaH_2PO_4, 2.7 mM KCl, 1.5 mM KH_2PO_4, 0.5% Triton X-100, pH 7.4), bound molecules were eluted with 40 μl of elution buffer (10 mM glutathione, 50 mM Tris-HCl, 50 mM NaCl, pH 7.5). PCR amplification of the eluates was performed as described (*13*) with a PCR program consisting of 30 cycles of denaturation (95°C, 20 s), annealing (58°C, 20 s), and extension (68°C, 100 s). The amplified products were purified with QIAquick and used for the next round of selection or analyses. PCR products amplified from fractions before (R0) or after the first (R1) or second (R2) round of selection were digested with *Bcl*I to selectively cleave the GST gene, separated in 2% agarose gels, and detected *via* fluorescein present at the upstream ends of DNA fragments. (B) The content of the GST gene in each fraction. DNA (0.2 fmol) was subjected to quantitative real-time PCR with GST-specific primers (GCTGACAAGCACAACATGT and GCAATTCTCGAAACAC) using a LightCycler (Roche). The DNA template for GST was mixed with that for RYBP in a ratio of 1:0, 1:10, 1:100 or 1:1,000, and used as a standard (r = –1.00).

The DNA genotype offers several advantages compared with the RNA genotype in ribosome display and mRNA display. The selection procedure for DNA display does not require strictly RNase-free conditions because of the chemical stability of DNA. As the removal of stop codons is unnecessary, translation products of cDNA libraries can be easily displayed. In addition, DNA display does not require several complicated steps essential for mRNA display, such as transcription, ligation of puromycin to the 3′-end of mRNA, and reverse transcription. The simplicity of DNA display allows one or even two selection rounds to be conducted per day while the other methods mentioned above mostly require 1–3 days for each selection round.

Microbead display is also based on an *in vitro* compartmentalization system for linking a protein with DNA on microbeads (*27*). In this system, a library containing ~10^9 repertoires can be subjected to selection by means of flow cytometry. The selection process for DNA display is based on affinity purification and the library size can be easily extended.

In summary, we demonstrated that folded proteins of even larger than 1000 aa can be efficiently displayed on the corresponding biotinylated DNA templates by inserting peptide linkers between SA and fused proteins. The resulting DNA-protein conjugates can be affinity-enriched by a factor of 20–30-fold per round. The DNA display system should be useful for rapidly screening or evolving proteins based on affinity selection.

We thank N. Matsumura, H. Takashima and Y. Oishi for the discussions, experimental advice and plasmids, respectively. The plasmid carrying Rae28/Mph1 cDNA was a gift from K. Shimada. This work was supported in part by the Special Coordination Fund of the Ministry of Education, Science and Technology of Japan.

REFERENCES

1. Luban, J. and Goff, S.P. (1995) The yeast two-hybrid system for studying protein-protein interactions. *Curr. Opin. Biotechnol.* **1**, 59–64
2. Bauer, A. and Kuster, B. (2003) Affinity purification-mass spectrometry. Powerful tools for the characterization of protein complexes. *Eur. J. Biochem.* **270**, 570–578
3. Kawahashi, Y., Doi, N., Takashima, H., Tsuda, C., Oishi, Y., Oyama, R., Yonezawa, M., Miyamoto-Sato, E., and Yanagawa, H. (2003) *In vitro* protein microarrays for detecting protein-protein interactions: Application of a new method for fluorescence labeling of proteins. *Proteomics* **3**, 1236–1243
4. Amstutz, P., Forrer, P., Zahnd, C., and Plückthun, A. (2001) *In vitro* display technologies: novel developments and applications. *Curr. Opin. Biotechnol.* **12**, 400–405
5. Dower, W.J. and Mattheakis, L.C. (2002) *In vitro* selection as a powerful tool for the applied evolution of proteins and peptides. *Curr. Opin. Chem. Biol.* **6**, 390–398
6. Doi, N. and Yanagawa, H. (2001) Genotype-phenotype linkage for directed evolution and screening of combinatorial protein libraries. *Comb. Chem. High Throughput Screen.* **4**, 497–509
7. Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317
8. Zozulya, S., Lioubin, M., Hill, R.J., Abram, C., and Gishizky, M.L. (1999) Mapping signal transduction pathways by phage display. *Nat. Biotechnol.* **12**, 1193–1198
9. Danner, S. and Belasco, J.G. (2001) T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proc. Natl Acad. Sci. USA* **98**, 12954–12959
10. Cochrane, D., Webster, C., Masih, G., and McCafferty, J. (2000) Identification of natural ligands for SH2 domains from a phage display cDNA library. *J. Mol. Biol.* **297**, 89–97
11. Sche, P.P., McKenzie, K.M., White, J.D., and Austin, D.J. (1999) Display cloning: functional identification of natural product receptors using cDNA-phage display. *Chem. Biol.* **6**, 707–716
12. Doi, N. and Yanagawa, H. (1999) STABLE: protein-DNA fusion system for screening of combinatorial protein libraries *in vitro*. *FEBS Lett.* **457**, 227–230
13. Yonezawa, M., Doi, N., Kawahashi, Y., Higashinakagawa, T., and Yanagawa, H. (2003) DNA display for *in vitro* selection of diverse peptide libraries. *Nucleic Acids Res.* **31**, e118
14. Tawfik, D.S. and Griffiths, A.D. (1998) Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* **16**, 652–656
15. Strynadka, N.C., Jensen, S.E., Johns, K., Blanchard, H., Page, M., Matagne, A., Frere, J.M., and James, M.N. (1994) Structural and kinetic characterization of a beta-lactamase-inhibitor protein. *Nature* **368**, 657–660

16. Arai, R., Ueda, H., Kitayama, A., Kamiya, N., and Nagamune, T. (2001) Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng.* **14**, 529–532
17. Bayer, E.A., Ben-Hur, H., and Wilchek, M. (1990) Isolation and properties of streptavidin. *Methods Enzymol.* **184**, 80–89
18. Mattheakis, L.C., Bhatt, R.R., and Dower, W.J. (1994) An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl Acad. Sci. USA* **91**, 9022–9026
19. Hanes, J. and Plückthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl Acad. Sci. USA* **94**, 4937–4942
20. He, M. and Taussig, M.J. (1997) Antibody-ribosome-mRNA (ARM) complexes as efficient selection particles for *in vitro* display and evolution of antibody combining sites. *Nucleic Acids Res.* **25**, 5132–5134
21. Hammond, P.W., Alpin, J., Rise, C.E., Wright, M., and Kreider, B.L. (2001) *In vitro* selection and characterization of Bcl-X(L)-binding proteins from a mix of tissue-specific mRNA display libraries. *J. Biol. Chem.* **276**, 20898–20906
22. McPherson, M., Yang, Y., Hammond, P.W., and Kreider, B.L. (2002) Drug receptor identification from multiple tissues using cellular-derived mRNA display libraries. *Chem. Biol.* **6**, 691–698
23. Miyamoto-Sato, E., Takashima, H., Fuse, S., Sue, K., Ishizaka, M., Tateyama, S., Horisawa, K., Sawasaki, T., Endo, Y., and Yanagawa, H. (2003) Highly stable and efficient mRNA templates for mRNA-protein fusions and C-terminally labeled proteins. *Nucleic Acids Res.* **31**, e78
24. Amstutz, P., Pelletier, J.N., Guggisberg, A., Jermutus, L., Cesaro-Tadic, S., Zahnd, C. and Plückthun, A. (2002) *In vitro* selection for catalytic activity with ribosome display. *J. Amer. Chem. Soc.* **124**, 9396–9403
25. Takahashi, F., Ebihara, T., Mie, M., Yanagida, Y., Endo, Y., Kobatake, E., and Aizawa, M. (2002) Ribosome display for selection of active dihydrofolate reductase mutants using immobilized methotrexate on agarose beads. *FEBS Lett.* **514**, 106–110
26. Bieberich, E., Kapitonov, D., Tencomnao, T., and Yu, R.K. (2000) Protein-ribosome-mRNA display: affinity isolation of enzyme-ribosome-mRNA complexes and cDNA cloning in a single-tube reaction. *Anal. Biochem.* **287**, 294–298
27. Sepp, A., Tawfik, D.S., and Griffiths, A.D. (2002) Microbead display by *in vitro* compartmentalisation: selection for binding using flow cytometry. *FEBS Lett.* **532**, 455–458
28. Sambrook, J. and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY
29. Doi, N. and Yanagawa, H. (1999) Design of generic biosensors based on green fluorescent proteins with allosteric sites by directed evolution. *FEBS Lett.* **453**, 305–307
30. Nomura, M., Takihara, Y., and Shimada, K. (1994) Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cells: one of the early inducible clones encodes a novel protein sharing several highly homologous regions with a Drosophila polyhomeotic protein. *Differentiation* **57**, 39–50