Unit Title

Running Title: Evolving aptamers with unnatural base pairs

Author(s): Michiko Kimoto[1], Ken-ichiro Matsunaga[1], Ichiro Hirao[1*]

Contact information: [1]Institute of Bioengineering and Nanotechnology, 31 Biopolis Way, The Nanos, #09-01, Singapore 138669, tel: +65-6824-7104, fax: +65-6478 9083, e-mail: ichiro@ibn.a-star.edu.sg

Significance Statement: A novel method to generate unnatural-base DNA aptamers that effectively bind to target proteins is described. To increase the affinity and specificity of DNA aptamers to targets, genetic alphabet expansion technology is utilized in the method, and the DNA aptamers contain hydrophobic unnatural bases as a fifth base. This protocol describes the preparation of a DNA library with a random sequence containing the unnatural bases, the isolation (selection) of DNA species that bind to target proteins from the library, the PCR amplification of the isolated DNA species, the separation of the single-stranded DNA for the next round of selection, the sequencing method of the enriched library, and the evaluation of the isolated DNA aptamers.

ABSTRACT: A novel technology, genetic alphabet expansion, has rapidly advanced through the successful creation of unnatural base pairs that function as a third base pair in replication. Recently, genetic alphabet expansion has been applied to some practical areas. Among them, the application to DNA aptamer generation is a good example of the broad utility of this technology. A hydrophobic unnatural base pair, Ds–Px, which exhibits high fidelity in replication as a third base pair, has been applied to an evolutionary engineering method called SELEX (Systematic Evolution of Ligands by EXponential enrichment) to generate DNA aptamers that bind to targets. A few Ds bases in DNA aptamers significantly increase the binding affinity to targets, enabling the use of DNA aptamers as an alternative to antibodies. This protocol describes the ExSELEX (genetic alphabet Expansion for SELEX) method to generate Ds-containing DNA aptamers.

Keywords: DNA aptamer, genetic alphabet expansion, SELEX, unnatural base pair, ExSELEX

INTRODUCTION

DNA aptamers are single-stranded DNA molecules that specifically bind to target molecules and materials, and are expected to function as an alternative to protein-based antibodies. Initially, DNA aptamers were generated by an evolutionary engineering method called SELEX, using a DNA library containing random base sequences (Ellington and Szostak, 1990; Tuerk and Gold, 1990). The DNA library is mixed with target molecules or materials, and the resulting DNA-target complexes are isolated (selection step). The isolated DNA species from the complexes are amplified by PCR, followed by the isolation of the single-stranded library from the amplified duplexes (amplification step). The selection and amplification steps are alternately iterated several times while gradually increasing the selection pressure (*e.g.*, reducing the library and target concentrations, adding competitors, intensifying the washing conditions of the

complexes). After several rounds of selection, each sequence in the enriched library is determined by an appropriate sequencing method. The DNA fragments of the top sequences are chemically synthesized and their binding affinities are evaluated to identify the best sequence meeting the desired criteria for use as the aptamer to the target.

One of the problematic issues with the conventional SELEX method results from the intrinsic limitation of the natural DNA molecules. DNA is highly hydrophilic and generally consists of only four different base components, A, G, C, and T, which have very similar physicochemical properties. Thus, the chemical and structural diversities of the conventional DNA library are much lower than those of proteins, composed of 20 standard amino acids, and the tight interaction between DNA and hydrophobic regions of targets is less favorable. To date, there have been many reports on modified SELEX methods to improve the affinity by increasing the chemical and structural diversities of DNA libraries (refer to the review papers, such as Bell and Micklefield, 2009; Diafa and Hollenstein, 2015; Houlihan et al., 2017; Keefe and Cload, 2008; Lapa et al., 2016; Y. Yang et al., 2011a).

Here, we describe one of the modified SELEX methods, called ExSELEX, which empolys genetic alphabet expansion using an artificial extra base pair (unnatural base pair), to generate unnatural-base DNA aptamers that bind to target proteins. Recently, several types of unnatural base pairs that function in replication with high fidelity have been developed (Benner et al., 2016; Hirao and Kimoto, 2012; Jung and Ellington, 2015; Kimoto et al., 2016a; Malyshev and Romesberg, 2015), allowing for their applications to SELEX (Biondi et al., 2016; Kimoto et al., 2013; Matsunaga et al., 2017; Sefah et al., 2014; L. Zhang et al., 2016; L. Zhang et al., 2015). Among them, the unnatural base pair between 7-(2-thienyl)imidazo[4,5-*b*]pyridine (Ds) and 2-nitro-4-propynylpyrrole (Px) (Figure 1) lacks a clear hydrogen bonding interaction between the pairing bases and exhibits high fidelity and efficiency in PCR amplification as a third base pair (Kimoto et al., 2009; Okamoto et al., 2016; Yamashige et al., 2012). For high fidelity PCR amplification, the Ds–Px pair system requires DNA polymerases with 3′→5′ exonuclease activity, such as Deep Vent or AccuPrime *Pfx* DNA polymerases. The amplification efficiency of the Ds–Px pair system depends on the sequence contexts around the Ds–Px pair. In general, pyrimidine-Ds-pyrimidine sequences are favorable relative to purine-Ds-purine sequences. Deep Vent DNA polymerase exhibits higher fidelity of Ds–Px pairing in PCR than AccuPrime *Pfx* DNA polymerase, but AccuPrime *Pfx* reduces the sequence dependency compared to Deep Vent (Yamashige et al., 2012). Thus, AccuPrime *Pfx* DNA polymerase is used for PCR amplification of the ExSELEX library. Due to the high hydrophobicity of the Ds base, only a few Ds bases in DNA aptamers significantly increase the affinity to targets (Kimoto et al., 2013; Matsunaga et al., 2017).

The ExSELEX scheme, depicted in Figure 2, consists of preparation of the library containing the Ds bases, DNA binding with targets and isolation of DNAs from the target-DNA complexes (selection step), PCR amplification of the isolated DNAs, isolation of the Ds strands from the amplified duplex DNAs (amplification step), repetition process of the selection and amplification steps, sequence determination of each DNA in the enriched library, evaluation of each DNA as an aptamer, and further optimization of the aptamer sequence. We will describe

the library preparation, selection, amplification, sequencing, and aptamer evaluation procedures.

STRATEGIC PLANNING

In ExSELEX, there are two types of DNA libraries containing Ds bases. One (Ds-predetermined library, Figures 2 and 3) is the mixture of several tens of sub-libraries, in which one to three Ds bases are embedded at defined positions in the random sequences consisting of the natural bases (Kimoto et al., 2013). Each sub-library contains Ds bases at different positions and has its specific bar-code sequence, consisting of a three natural-base sequence, to identify the Ds positions in the sub-library. The other (Ds-randomized library) is a library containing a randomized base sequence including the Ds base (Matsunaga et al., 2017). Here, we describe ExSELEX using the Ds-predetermined library.

The main issue of the library construction is how we determine the Ds positions in each sequence, from the enriched library after several rounds of selection-amplification steps. Currently, there is no method and sequencer that can directly determine the Ds positions in a mixture of different sequences. In order to employ current sequencing methods, the Ds bases in the enriched library must be replaced with natural bases by PCR in the absence of unnatural base substrates (replacement PCR) (Kimoto et al., 2013). Therefore, the method using the Ds-predetermined library is easier than that using the Ds-randomized library. However, the complexity (the number of different sequences in a library) of the Ds-predetermined library is theoretically lower than that of the Ds-randomized library, thus reducing the success rate of aptamer generation.

After the sequence determination of the enriched libraries, the aptamer sequence with the highest affinity is chosen by a binding assay, such as a gel-shift assay, filter-binding assay, or Surface Plasmon Resonance (SPR) analysis, of each sequence. The sequence and length of the selected aptamer can be optimized by a second round of ExSELEX using a doped library (Kimoto et al., 2013), in which the base sequence is partially randomized. In this protocol, we also describe the gel-shift assay, which is useful for the DNA aptamer selection and evaluation. For the optimization, refer to the literature (Kimoto et al., 2016a; Kimoto et al., 2016b; Kimoto et al., 2013; Matsunaga et al., 2015; Matsunaga et al., 2017).

BASIC PROTOCOL 1

PREPARATION OF DNA LIBRARY WITH A RANDOM SEQUENCE CONTAINING THE UNNATURAL BASES

First, we describe the preparation of the DNA library designated as N26Ds, a mixture of 28 sub-libraries, which was used in our previous Ds-aptamer selection targeting the vWF-A1 domain (Matsunaga et al., 2017). Within the central 26-nucleotide randomized region of each sub-library, two Ds bases are embedded, as shown in Figure 3. The two Ds bases in each sub-library

are separated by at least eight nucleotides, since the PCR efficiency is lower when the Ds bases are too close to each other. Each sub-library is chemically synthesized using the natural base and Ds phosphoramidite reagents by a conventional method (Hirao et al., 2006), deprotected and purified by denaturing PAGE, and mixed in equal amounts. This mixed 81-nucleotide library (N26Ds DNA library) is used for ExSELEX directly, as the initial Ds-containing DNA library. In the later ExSELEX rounds, the isolated library is amplified by PCR for further selection rounds, and their full lengths became 90-nucleotides (Figure 3).

Materials

8% (19:1) polyacrylamide gel containing 7 M urea (20 cm × 40 cm, 2-mm thick, 6-well) (see recipe)
1× TBE (see recipe)
Chemically synthesized Ds-containing DNA sub-libraries (200 nmol scale synthesis, 81-nucleotides): The sequence, 5'-CCGTTCTCTAATTTTGACGTTXXXACTNNNNNnnnnnnnnnnnnnnnNNNNNAGTYYYAGGGTCTGTATCGCAATAATT-3', consists of the 5'-primer part, 5'-stem (XXXACT), 26-nt randomized (N=A, G, C, or T; n = Ds or N) part, 3'-stem (AGTYYY), and 3'-primer regions (Figure 3). In each sub-library, the XXX and YYY sequences are used as a bar-code to identify the predetermined Ds positions.
Denaturing loading dye solution (see recipe)
Marker dye solution (see recipe)
3 M sodium acetate, pH 5.2 (NaOAc)
100% and 70% ethanol, -20°C

Polyacrylamide gel electrophoresis apparatus (BIO CRAFT Co., Ltd.)
Power supply (Bio-Rad PowerPac 3000 with a temperature probe)
Refrigerated microcentrifuge
75°C heating block
Spatula
Saran wrap
Thin-layer chromatography (TLC) plate containing a UV fluorescent indicator
Handheld UV lamp (254-nm)
Razor blade
50-ml conical tubes
Incubator with shaker
0.22-µm filter units (Steriflip, Millipore)
Vacuum pump
5-ml tubes
SpeedVac concentrator
DNA LoBind 1.5-ml tubes (Eppendorf)
UV spectrophotometer

Set up denaturing PAGE

Step 1. Remove the comb and the bottom seal from the gel plates.

Step 2. Rinse the wells of the gel with water, to remove un-polymerized gel solution.

Step 3. Attach the gel plate to an electrophoresis apparatus, as suggested by the manufacturer.

Step 4. Fill the upper and lower reservoirs with 1× TBE as running buffer.

Step 5. Connect to a power supply and pre-run the gel for 30 to 60 min at a constant power of 50W, so that the temperature of the gel reaches around 50°C.

The use of a temperature probe is convenient to control the wattage for running the gel. If such temperature probes are not available, monitor and control the power setting manually, so that the temperature is not too high.

Prepare the loading samples

Step 6. Dissolve a chemically-synthesized Ds-containing DNA sub-library in 200 µl of water and mix with 200 µl of denaturing loading dye solution.

Step 7. Heat the solution at 75°C for 3 min and mix well.

Load the samples and run the gel

Step 8. Stop the power supply and rinse the wells with 1× TBE by flushing out the urea leached from the wells.

Step 9. Load the sample into the four central wells (100 µl each per well), and load 30 µl of the marker dye solution into each of the small wells on both sides of the gel.

Step 10. Run the gel until the marker dye reaches the predetermined position.

The migrations of bromophenol blue and xylene cyanol dyes in an 8% (19:1) denaturing polyacrylamide gel correspond to around 20-nts and 75-nts, respectively. After a run for around 4 hours, the xylene cyanol will have migrated to about one-third of the distance from the bottom of the gel.

Extract the DNA sub-library from the gel

Step 11. Terminate electrophoresis, remove the gel plate from the apparatus, and carefully detach one glass plate using a spatula.

Step 12. Cover the gel with Saran wrap, turn the glass plate over, and place the gel on a wrapped TLC plate. Remove the other glass plate from the gel.

Step 13. Identify the shadow bands corresponding to the full-length Ds-containing DNA, by illumination with a handheld UV lamp at 254 nm.

Under illumination at 254 nm, a fluorescent TLC plate will glow green. The DNA in the gel will absorb the UV light and will be detected as dark shadow bands, with intensities that correlate with the amount of DNA. Make sure to illuminate with the UV lamp oriented in the vertical direction, so that the shadow overlaps with the same position of the DNA in the gel.

Step 14. Excise the bands with a clean razor blade and transfer the gel slices into a 50-ml conical tube.

Step 16. Crush the slices, add 5.4 to 6 ml of water, and incubate the tube at 37°C with gentle agitation for at least 8 hours.

To easily crush gel slices, a disposable pipette or sterilized spatula can be used. If available, a disposable BioMasher stir bar is convenient. The amount of water required for elution is adjusted depending on the amount of gel slices.

Precipitate the sub-library DNA

Step 17. Pass the eluted DNA solution through a 0.22-µm Steriflip filter under vacuum, to remove the gel debris.

The recovered eluted solution will be around 4.8 ml.

Step 18. Transfer 1.2 ml of the filtrate into each of four 5-ml tubes, add 120 µl (1/10 vol.) of 3M NaOAc, and then mix with 3 ml (2 to 2.5 vol.) of 100% ethanol by inverting the tubes several times and by vortexing.

Step 19. Store each tube at -20°C for at least 30 min.

Step 20. Centrifuge the tubes at 15,300 × g for 40 min at 4°C.

Step 21. Carefully remove and discard the supernatant, and then add 1 ml of pre-chilled 70% ethanol to each tube.

Step 22. Centrifuge the tubes at 15,300 × g for 10 min at 4°C.

Step 23. Carefully remove and discard the supernatant, and evaporate the residual ethanol with a SpeedVac concentrator.

Step 24. Add 0.1 ml water to each tube and incubate at 75°C for 3 min, to dissolve the precipitated DNAs.

Step 25. Combine the solutions in a single DNA LoBind 1.5-ml tube and mix them well.

Step 28. Measure the UV absorbance at 260 nm and determine the DNA concentration by the Beer-Lambert Law: $Abs_{260} = E_{260} \times C \times L$, where Abs_{260} is the UV absorbance, E_{260} is a molecular coefficient (L/mol/cm), C is the DNA concentration (mol/L), and L is the length of the light path (cm).

The average molecular coefficient values for DNA fragments can be calculated by using the free DNA calculators on the internet, such as OligoAnalyzer 3.1 (IDT), or the software equipped with the UV spectrophotometer.

Make a Ds-containing DNA library (N26Ds) for the first round of ExSELEX

Step 29. Mix equal amounts of each sub-library in a 5-ml tube.

Step 30. Dilute to a 10 µM working solution and store 200-µl aliquots at -20°C.

BASIC PROTOCOL 2

SELECTION OF DNA SPECIES THAT BIND TO A TARGET PROTEIN

In the selection step, the DNA library is incubated with a target molecule or material. The DNA-target complex is then separated from the unbound DNA species, and the bound DNA species is recovered from the complex. To isolate the DNA-target complex, several conventional methods can be used, such as gel-shift and filter-binding methods, as well as pull-down methods with immobilization of the target to resins or magnetic beads via a tag. As an example, this protocol describes the selection method used in ExSELEX targeting the human vWF-A1 domain (Matsunaga et al., 2017), via biotinylation of the target protein, followed by a pull-down isolation with streptavidin-coated magnetic beads. In the later rounds of ExSELEX, the selection pressure is changed by reducing the target and library concentrations, shortening the incubation time, or washing in harsher conditions, as exemplified in Table 1.

Materials

N26Ds DNA library (see Basic Protocol 1)
10× PBS, pH 7.4 (Gibco)
0.5% (w/v) Nonidet P-40
Recombinant vWF-A1 domain, human, residues 1261–1468 (U-Protein Express BV)
EZ-Link Sulfo-NHS-LC-Biotin (Thermo Fisher Scientific)
2 M Glycine (see recipe)
1× Binding buffer (see recipe)
Hydrophilic streptavidin magnetic beads, 4 mg/ml (New England Biolabs)
50 mM NaOH
Dowex (H+) slurry (see recipe)
1 M Tris-HCl, pH 7.4

DNA LoBind 1.5-ml tubes (Eppendorf)
95°C heating block
Amicon Ultra-0.5 Centrifugal Filter Units with Ultracel-30 membrane (Merck)
Magnetic separation rack
Incubator with shaker
Tube rotator

0.22-µm Ultrafree-MC filter units (Merck)

Incubate the DNA library with the target protein

Step 1. Combine the following reagents (1 ml) in DNA Lobind 1.5-ml tubes (six tubes in total).

895 µl water
5 µl 10 µM Ds-containing DNA library
100 µl 10× PBS

Step 2. Heat the tubes containing the above mixture at 95°C for 5 min, snap-cool by placing the tubes on ice for 10 min, keep the tubes at room temperature for 10 min, and then add 10 µl of 0.5% (w/v) Nonidet P-40 to each tube (final concentration: 0.005% (w/v)).

The volume and final concentration of the DNA library can be adjusted depending on the selection conditions of each round.

Step 3. Add the vWF-A1 solution (25 pmol, final concentration: 50 nM) to each tube and incubate for 30 min at 25°C.

The final concentration is adjusted depending on the selection conditions of each round, and the incubation temperature depends on the protein stability.

Biotinylation of the DNA–protein complex

Step 4. Prepare 5 mM biotinylation reagent, freshly before use: weigh EZ-Link Sulfo-NHS-LC-Biotin (around 3 mg) and make a 5 mM solution by adding water (360 µl water per 1 mg of the compound for 5 mM).

The reagent is moisture-sensitive and should be carefully handled. The reagent powder must be stored desiccated at -20°C. Equilibrate the vial to room temperature before opening, to avoid condensation.

Step 5. Add 10 µl (1/100 vol.) of 5 mM biotinylation reagent to each DNA–protein solution in step 3, and incubate for 15 min at 25°C.

Biotinylation conditions can be adjusted, depending on the target's properties; i.e., the number of Lys residues on the target's surface and its reactivity.

Step 6. Add 53 µl of 2 M glycine (final concentration: 100 mM) to each tube, to stop the biotinylation.

Step 7. Perform diafiltration of the solution with Amicon Ultra-0.5 Centrifugal Filter Units with an Ultracel-30 membrane (for 1-ml volume, use one filter unit). Add 450 µl of the solution to the pre-rinsed filter device, concentrate by centrifugation at 13,000 × g for 5 min at 25°C, and remove the filtrate. Repeat several times, until all of the solution has been loaded. Remove the

excess biotinylation reagents, as follows. Add 450 µl of 1× Binding buffer, centrifuge at 13,000 × g for 5 min at 25°, and then remove the filtrate. Repeat this procedure two more times.

Pre-rinse the device with 1× Binding buffer before use, and do not allow the membrane to dry out once wet. If you are not using the device immediately after pre-rinsing, leave fluid on the membrane until the device is used. Selection of the cutoff (Nominal Molecular Weight Limit, NMWL) will be dependent on the molecular weights of the target and the DNA-target complex. In this case, the molecular weight of the vWF-A1 domain is 27.4 kDa. The molecular weight of the DNA is around 26.7 kDa, and the cutoff of the unit used is 30 kDa.

Step 8. Recover the DNA-target complex. Separate the filter device from the bottom tube, turn the filter device upside down in a new collection tube, and then centrifuge at 1,000 × g for 2 min at 25°C. Transfer each recovered solution into a DNA LoBind 1.5-ml tube.

Immobilization of biotinylated DNA–target complex to streptavidin magnetic beads

Step 9. Prepare the hydrophilic streptavidin magnetic beads for immobilization. In a 1.5 ml-tube, add 240 µl of 4 mg/ml bead slurry to each tube, and wash the beads with 200 µl of 1× Binding buffer three times.

For 25 pmol of target protein, 160 µg of beads would be used. Depending on the protein size, the amount of the beads can be adjusted.

Step 10. Mix the washed beads with each recovered solution in step 8, and incubate for 10 min at 25°C with gentle mixing (around 400 rpm).

Step 11. Centrifuge the tube for 10 sec in the microcentrifuge, put it on the magnetic rack for 1 min, and then remove the supernatant.

Washing

Step 12. Add 1 ml of 1× Binding buffer and mix the solution on a tube rotator for 5 min at 25°C (around 10 rpm).

Washing buffer conditions can be changed to increase the stringency for washing in later rounds of ExSELEX.

Step 13. Centrifuge the tube for 10 sec in the microcentrifuge, put it on the magnetic rack for 1 min, and then remove the supernatant.

Step 14. Repeat steps 12 and 13 four times (Total five washes).

Elution of DNA from the immobilized DNA-target complex

Step 15. To the magnet beads, add 200 µl of 50 mM NaOH and incubate for 5 min at 25°C.

Step 16. Centrifuge the tube for 10 sec in the microcentrifuge, put it on the magnetic rack for 1 min, and then transfer the supernatant into a new DNA LoBind 1.5-ml tube.

Step 17. Mix with 20 µl of DOWEX (H+) slurry and vortex for 30 sec.

Step 18. Filter the solution by transferring the whole solution to an Ultrafree-MC filter unit, and centrifuge at 14,000 × g for 1 min at 25°C.

Step 19. Add 2 µl of 1 M Tris-HCl (pH 7.4) to the recovered filtrate, to neutralize the pH of the solution.

BASIC PROTOCOL 3

PCR AMPLIFICATION OF THE ISOLATED Ds-CONTAINING DNA

The isolated Ds-containing DNA species from the DNA-target complex are amplified by large-scale PCR in the presence of dDsTP and dPxTP. To determine the number of PCR cycles for the large-scale PCR, qPCR can be performed using a small amount of the isolated library, to check the amplification efficiency, in the second and later rounds of ExSELEX. In the first round of ExSELEX, qPCR should be omitted so the entire isolated library can be used for the large-scale PCR, since the number of each DNA with the same sequence is theoretically only one in the initial library. A 5'-Dss-labeled forward primer is used for easy detection of the Ds-strand in the denaturing PAGE process. Dss (7-(2,2'-bithien-5-yl)-imidazo[4,5-*b*]pyridine) (Figure 1), an analogue of Ds, is highly fluorescent (Kimoto et al., 2010). The Ds-strand (90 nucleotides) is shorter than the Px-strand (105 nucleotides), since the Ds-strand extension stops at the position opposite the linker in the reverse primer. Purification of the Ds-strands by denaturing PAGE secures the full-length DNA library.

Materials

2.5 U/µL AccuPrime *Pfx* DNA polymerase with 10× AccuPrime *Pfx* Reaction mix and 50 mM $MgSO_4$ (Life Technologies)
20 µM Dss-forward primer: 5'-Dss-ACGACCGTTCTCTAATTTTGACGTT-3'
20 µM linker-reverse primer: 5' -TTTTTTTTTTTTTTT-L-ACCAAATTATTGCGATACAGACCCT-3' (L: Spacer C12 Phosphoramidite)
10 mM each dNTP mix (see recipe)
0.5 mM dDsTP
0.5 mM dPxTP
1/7,500-diluted SYBR Green I (ThermoFisher) (see recipe)
Ds-containing DNA solution recovered from the complex (around 200 µl) (see Basic Protocol 2)
4% Agarose gel containing ethidium bromide
1× TBE (see recipe)
Agarose gel loading buffer (see recipe)
25-bp DNA step ladder (Promega)
8% (19:1) polyacrylamide gel containing 7 M urea (16 cm x 16 cm, 2-mm thick) (see recipe)
Denaturing loading dye solution (see recipe)
Marker dye solution (see recipe)

3 M sodium acetate, pH 5.2 (NaOAc)
100% and 70% ethanol, -20°C

Thin-wall 0.2-ml PCR tubes
CFX96 qPCR instrument (Bio Rad) or any other qPCR instrument that can detect SYBR Green I signals
Thermal cycler
Electrophoresis apparatus for agarose gel
Power supply
LAS4000 bio-imager, or a similar gel imaging system capable of DNA-stained gel detection
Electrophoresis apparatus for denaturing polyacrylamide gel
75°C heating block
Spatula
Saran wrap
Thin-layer chromatography (TLC) plate containing a UV fluorescent indicator
Handheld dual-UV lamp (254 nm and 365 nm)
Razor blade
2-ml tubes
Incubator with shaker
0.22-µm Ultrafree-CL filter unit (Millipore)
Vacuum pump
5-ml tubes
Refrigerated microcentrifuge
SpeedVac concentrator
DNA LoBind 1.5-ml tubes (Eppendorf)
UV spectrophotometer

Real-time PCR to determine the number of PCR cycles for large-scale PCR

Step 1. In a PCR tube, prepare the following PCR master mix, for 2-tube PCR reactions on a 25 µl scale:

5 µl 10× AccuPrime *Pfx* Reaction mix
2.5 µl 20 µM Dss-forward primer
2.5 µl 20 µM reverse-linker primer
2 µl 2.5 mM each dNTP mix (4-fold dilution of 10 mM each dNTP mix)
2 µl 12.5 mM $MgSO_4$ (4-fold dilution of 50 mM $MgSO_4$)
5 µl 0.5 mM dDsTP
5 µl 0.5 mM dPxTP
5 µl 1/7,500-diluted SYBR Green I
1 µl 2.5 U/µl AccuPrime *Pfx* DNA polymerase

Note that 1× AccuPrime Pfx Reaction mix includes final concentrations of 0.3 mM each dNTP and 1 mM $MgSO_4$. The final concentrations in PCR reactions are 0.4 m each dNTP, 0.05 mM each dDsTP and dPxTP, 1.5 mM $MgSO_4$, 1/75,000-diluted SYBR Green I, and 0.05 U/ µl AccuPrime Pfx DNA polymerase.

Step 2. In a new PCR tube, transfer a 15 µl portion of the master mix and add 10 µl of the Ds-containing DNA solution (designated as “DNA plus”).

Step 3. To the remaining PCR master mix, add 10 µl of water (designated as Non-Template-Control, “NTC”).

Step 4. Place the “DNA plus” and “NTC” reaction tubes in a real-time PCR instrument, and perform PCR with a program consisting of 94°C for 2 min (initial denaturation), 25 cycles at 94°C for 15 sec, and 65°C for 3 min 30 sec (fluorescence signal reading at the end of the elongation step in each cycle), followed by monitoring of the melting profile: 94°C for 15 sec, 55°C for 5 sec, and 95°C for 5 sec (fluorescent signal reading from 55°C to 95°C).

qPCR is useful not only to determine the number of PCR cycles required for the large-scale PCR, but also to assess the level of contamination in the PCR process from the amplification data of “NTC” . A comparison of the melting curve profiles in each round can also be used as an indicator of the enrichment process.

Large-scale PCR to amplify Ds-containing DNA

Step 5. Prepare the following PCR master mix in a 1.5-ml tube, for 9-tube PCR reactions on a 50 µl scale:

72 µl water
45 µl 10× AccuPrime *Pfx* Reaction mix
22.5 µl 20 µM Dss-forward primer
22.5 µl 20 µM linker-reverse primer
4.5 µl 10 mM each dNTP mix
4.5 µl 50 mM $MgSO_4$
45 µl 0.5 mM dDsTP
45 µl 0.5 mM dPxTP
9 µl 2.5 U/µl AccuPrime *Pfx* DNA polymerase

Step 6. In a PCR tube, transfer 30 µl from the master mix, add 20 µl of water, and mix the solution (designated as Non-Template-Control, “NTC”).

Step 7. To the remaining PCR master mix (240 µl), add 160 µl of the Ds-containing DNA and mix the solution. Aliquot 50 µl volumes from the solution into eight PCR tubes (designated as “DNA plus”).

Step 8. Place the “DNA plus” and “NTC” reaction tubes in a thermal cycler, and perform PCR with a program consisting of 94°C for 2 min (for polymerase activation), N cycles at 94°C for 30 sec, and 65°C for 3 min 30 sec.

In general, we stop the large-scale PCR after 10 extra cycles, relative to the cycle number when the qPCR amplification began to be detected in the log-phase mode in step 4.

Step 9. Analyze the PCR-amplified products by electrophoresis on a 4% agarose gel with 1× TBE as running buffer. Mix 5 µl each of the "DNA plus" and "NTC" reactions with 2 µl of agarose gel loading buffer, and load onto a 4% agarose gel. Load 1 µl of a 25-bp DNA step ladder in one lane as a marker. Run the gel at 150 V for 15 min, and detect the ds DNA bands with a bio-imager.

If the PCR cycles are not sufficient, you can perform additional PCR cycles. However, fewer than 25 cycles of PCR are recommended.

Step 10. Combine the "DNA plus" PCR reactions into a single 1.5-ml tube.

Run a denaturing gel to purify the PCR-amplified Ds-containing DNA library

Step 11. Set up denaturing PAGE, similarly to Basic Protocol 1. Attach the gel plate to the electrophoresis apparatus, fill the upper and lower chambers with 1× TBE, and connect to a power supply. Prior to loading, pre-run the gel for about 15 to 30 min until the temperature of the gel reaches around 50°C.

Step 12. Prepare the loading sample. Add 200 µl of the denaturing loading dye to the remaining "DNA plus" solution (total about 400 µl, prepared in step 10) and heat at 75°C for 3 min.

Step 13. Run the gel. Stop the power supply, and rinse the wells with 1× TBE to flush out the urea leached from the wells. Load the prepared samples into two wells (300 µl per well). Run the gel for around 1 hour, until the marker dye (xylene cyanol) reaches one-third of distance from the bottom of the gel.

It is important to run the gel until the Ds strand (90-nts) and the Px strand (105-nts) are sufficiently separated. While the gel is running, the Dss-labeled Ds-strand band can be monitored even through the glass plate, by irradiation with the 365-nm UV light.

Extract the amplified Ds-library from the gel

Step 14. Terminate electrophoresis, remove the gel plate from the apparatus, and carefully detach one glass plate using a spatula.

Step 15. Cover the gel with Saran wrap, turn the glass plate over, remove the other glass plate from the gel, and place the gel on a TLC plate.

Step 16. Check the DNA bands corresponding to the full length Dss-labeled Ds-strand, by illumination with a handheld UV lamp at 254 nm and at 365 nm, respectively.

Under the 254-nm UV light, two DNA bands (both Ds- and Px-strands) can be detected and the shadow corresponding to the Dss-labeled Ds-strand looks bluish, due to Dss fluorescence. Under the 365-nm UV light, only the Dss-labeled DNA band (Ds-strand, shorter than the Px-strand) can be detected (see Figure 4).

Step 17. Excise the Ds-strand bands with a clean razor blade and transfer gel slices to a 1.5-ml tube (one tube per one-well DNA band).

In comparison with UV-shadowing, Dss fluorescence is more easily detected, facilitating the purification process even when detection by UV shadowing is difficult due to low PCR amplification yield.

Step 18. Crush the gel slices and add 1 ml of water to each tube.

Step 19. Incubate the tubes at 37°C with gentle agitation for at least 8 hours.

The elution status can be monitored by irradiation of the tubes with 365-nm UV light. After the elution, the Dss fluorescence should mainly be observed in the solution part, rather than the gel part.

Precipitate the amplified Ds-containing DNA

Step 20. Pass the eluted solution through a 0.22-µm Ultrafree-CL filter unit by centrifugation at 4,400 × g for 10 min.

The recovery volume will be less than 2 ml, due to some water absorption to the gel slices.

Step 21. In each 5-ml tube, mix around 0.9 ml of the filtrate and 90 µl 3M NaOAc, and then add 2.5 ml ethanol (total two tubes). After mixing the solution by vortexing and inverting the tubes several times, store the tubes at -20°C for 1 hr.

Step 22. Centrifuge the tubes at 15,300 × g for at 40 min at 4°C.

Step 23. Remove and discard the supernatant.

Step 24. Add 1 ml of pre-chilled 70% ethanol to each tube, and centrifuge the tubes again at 15,300 × g for 10 min at 4°C.

Step 25. Remove and discard the supernatant.

Step 26. Evaporate the residual ethanol with a SpeedVac concentrator.

Step 27. Add 0.1 ml water to each tube and incubate at 75°C for 3 min, to dissolve the precipitated DNAs.

Step 28. Combine the solutions in a DNA LoBind 1.5-ml tube and mix them.

Step 29. Measure the UV absorbance at 260 nm and determine the DNA concentration according to the following formula: 1 Abs_{260} = 33 ng/µl = 33 mg / (330 × 90 g/mol) = 1.11 µM.

The amplified Ds-containing DNA library thus obtained can be used in the next round of ExSELEX (see Basic Protocol 2), in the binding analysis (see Basic Protocol 4), and in the replacement PCR for DNA sequencing (see Basic Protocol 5).

BASIC PROTOCOL 4

BINDING ANALYSIS OF THE ENRICHED DNA LIBRARY BY GEL-SHIFT ASSAYS

The enrichment process of the binding DNA species in each round can be monitored by a simple gel-shift assay. Using various gel-shift conditions, as well as in comparison with a conventional natural-base DNA aptamer (if available), the binding affinity of each DNA species is examined to estimate the enrichment of the library. As an example, this protocol describes the evaluation of DNA libraries after each round of selection targeting vWF-A1, using a conventional anti-vWF DNA aptamer, ARC1172, as the control (Figure 5).

Materials

1 µM each of PCR-amplified and gel-purified Ds-DNA libraries after each round of selection (see Basic Protocols 1 and 3)
1 µM ARC1172: 5'-GGCGTGCAGTGCCTTCGGCCGTGCGGTGCCTCCGTCACGCC-3'
8% (29:1) polyacrylamide-5% glycerol gel with or without urea (16 cm × 16 cm, 1-mm thick) (see recipe)
0.5 × TBE (see recipe)
10× PBS, pH 7.4 (Gibco)
0.05% (w/v) Nonidet P-40 in 1×PBS
200 nM vWF-A1 solution (diluted with 1× Binding buffer before use)
1× Binding buffer (see recipe)
Gel-shift loading buffer (see recipe)
Diluted SYBR Gold solution (see recipe)

95°C heating block
Incubator
Polyacrylamide gel electrophoresis apparatus (BIO CRAFT Co., Ltd.)
Power supply
Spatula
Shaker
LAS4000 bio-imager, or a similar gel imaging system capable of DNA-stained gel detection

Binding of each DNA library or control DNA to a target protein

Step 1. Combine the following reagents (27 µl in total) in each tube.

18.3 µl water
6 µl Ds-containing DNA library or ARC1172 (6 pmol)
2.7 µl 10× PBS

Step 2. Heat the mixture at 95°C for 5 min, snap-cool by placing the tubes on ice for 10 min, keep them at room temperature for 10 min, and then add 3 µl of 0.05%(w/v) Nonidet P-40 in 1× PBS to each tube (final concentration: 0.005% (w/v) Nonidet P-40).

Step 3. Mix 5 µl of each DNA solution (200 nM) with 5 µl of 200 nM vWF-A1 solution or 1× Binding buffer (without protein), and incubate for 30 min at 25°C or 37°C.

Perform gel electrophoresis and detect the DNA band patterns

Step 4. Set up PAGE, similarly to Basic Protocol 1. Attach the gel plate set to the electrophoresis apparatus, fill the upper and lower chambers with 0.5× TBE, and connect to the power supply. Prior to loading the samples, pre-run the gel for about 15 min at constant 300V.

Step 5. To the tubes in step 3, add 2.5 µl of gel-shift loading buffer and mix well by pipetting.

Step 6. Run the gel. Stop the power supply, and rinse the wells with 0.5× TBE. Load the samples prepared in step 5 into each well. Run the gel for around 40 min.

Step 7. Terminate electrophoresis, remove the gel plate from the apparatus, and carefully detach one glass plate using a spatula. Soak the gel in the diluted SYBR Gold solution.

Step 8. Stain the gel with gentle agitation by using a shaker, for 5 to 10 min.

Step 9. Analyze the DNA band patterns with a bio-imager.

BASIC PROTOCOL 5

PREPARTION OF DNA TEMPLATES FOR SEQUENCING

To determine the sequence of each DNA species in the enriched library**,** replacement PCR, where the Ds bases in the DNA library are replaced by natural bases, is required. In the replacement PCR, a few cycles of amplification are performed in the presence of dPa’TP (Hirao et al., 2006) (Figure 1), but without dDsTP and dPxTP. The addition of dPa’TP facilitates the replacement of the Ds bases with the natural bases (mainly A or T) via Ds−Pa’ pairing, followed by the A−Pa’ or T−Pa’ pairing (Kimoto et al., 2013). The DNA fragments after replacement PCR can be subjected to conventional DNA sequence determination methods, such as cloning and deep sequencing. The predetermined Ds positions, which are replaced with the natural bases in the sequencing data, can be estimated from the bar-code sequence in each original sub-library. Here, we describe the replacement PCR procedure, in which the Ds bases are replaced with the natural bases, using a conventional primer set.

Materials

2.5 U/µL AccuPrime *Pfx* DNA polymerase with 10× AccuPrime *Pfx* Reaction mix and 50 mM $MgSO_4$ (Life Technologies)
20 µM forward primer: 5'-ACGACCGTTCTCTAATTTTGACGTT-3'

20 µM reverse primer: 5'-ACCAAATTATTGCGATACAGACCCT-3'
10 mM each dNTP mix (see recipe)
0.5 mM dPa'TP
100 nM Ds-containing DNA library (see Basic Protocol 3)
4% Agarose gel containing ethidium bromide
1× TBE (see recipe)
Agarose gel loading buffer (see recipe)
25-bp DNA step ladder (Promega)

Thin-wall PCR tubes
Thermal cycler
Electrophoresis apparatus for agarose gel
Power supply
LAS4000 bio-imager, or a similar gel imaging system capable of DNA-stained gel detection
QIAquick Gel Extraction Kit (QIAGEN)
UV spectrophotometer

Replacement PCR

Step 1. Prepare the following PCR mix in a PCR tube (50 µl scale):

19 µl water
5 µl 10× AccuPrime *Pfx* Reaction mix
2.5 µl 20 µM forward primer
2.5 µl 20 µM reverse primer
2.5 µl 2 mM each dNTP mix
2.5 µl 10 mM $MgSO_4$
5 µl 0.5 mM dPa'TP
1 µl 2.5 U/µl AccuPrime *Pfx* DNA polymerase

Step 2. To the solution in the PCR tube, add 10 µl of 100 nM Ds-containing DNA library and mix the solution.

Step 3. Place the reaction tube in a thermal cycler, and perform PCR with a program consisting of 94°C for 2 min (initial denaturation), 8 cycles at 94°C for 30 sec, and 65°C for 3 min 30 sec.

Step 4. Analyze the PCR products on a 4% agarose gel (See Basic Protocol 3).

Step 5. Purify the PCR products by using a QIAquick Gel Extraction Kit, according to the manufacturer's instructions.

Step 6. Determine the DNA concentrations from the UV absorbance.

The samples can be subjected to any type of sequencing procedure (cloning or next generation sequencing).

REAGENTS AND SOLUTIONS

Use nuclease-free, deionized, sterile or ultra-pure water in all recipes and protocol steps.

Agarose gel loading buffer

12 ml Glycerol
10 mg Bromophenol blue
28 ml Nuclease-free water
Sterilize the solution using a 0.22-µm filtering apparatus. Store aliquots at -20°C for up to one year.

1× Binding buffer

100 ml 10×PBS pH7.4 (GIBCO)
1 ml 5% (w/v) Nonidet P-40
Bring the volume to 1,000 ml with water, and sterilize the solution using a 0.22-µm filtering apparatus. Store at room temperature for up to two months.

Denaturing loading dye

120 g Urea
50 mg Bromophenol blue
20 ml 10× TBE
Bring the volume to 200 ml with water and sterilize the solution using a 0.22-µm filtering apparatus. Store aliquots at -20°C for up to 1 year; while in use, store at room temperature for up to one month.

Diluted SYBR Gold solution

5 µl SYBR Gold (Thermo Fisher Scientific)
100 ml 0.5× TBE
Prepare before use. Store at room temperature for up to one day with protection from light.

1/30-diluted SYBR Green I

2 µl SYBR Green I (LONZA)
58 µl 10 mM Tris-HCl pH7.4
Store at 4°C for up five days with protection from light.

1/7,500-diluted SYBR Green I

4 µl 1/30-diluted SYBR Green I
996 µl Nuclease-free water
Store aliquots at -20° for up to one year.

Dowex (H+) slurry

20 g Dowex 50WX8 hydrogen form 100-200 mesh (Sigma)
Put the resin into a 50-ml tube, and add water up to 50 ml. Mix the solution gently, and then discard the supernatant. Repeat this washing procedure five times. Transfer the slurry into a disposable column and wash three times with 15 ml of water, followed by four washes with 15 ml of 4 N HCl. After checking the pH of the flow through (< pH 1), wash eight times with 15 ml of water. After checking the pH of the flow though (around pH 7), transfer the slurry to a new 50-ml tube. Store at 4°C for up to one year.

10 mM each dNTP mix

100 µl 100 mM dATP
100 µl 100 mM dCTP
100 µl 100 mM dGTP
100 µl 100 mM dTTP
600 µl Nuclease-free water
Store aliquots at -20° for up to one year.

Gel-shift loading buffer

250 µl Glycerol
100 µl 10× PBS pH 7.4 (GIBCO)
100 µl 0.05% Nonidet P-40
550 µl Nuclease-free water
Store at room temperature for up to one month.

2 mM glycine

7.51 g Glycine
Add water to 50 ml, and sterilize the solution using a 0.22-µm filtering apparatus. Store aliquots at -20°C for up to one year.

10× TBE (around pH 8.3)

54 g Tris base
27.5 g Boric acid
20 ml 0.5 M EDTA-2Na, pH 8

Add water to 1 l, and filter the solution to sterilize. Store at room temperature for up to 6 months.

Marker dye solution

120 g Urea
50 mg Bromophenol blue
50 mg Xylene cyanol
20 ml 10× TBE
Add water to 200 ml, and sterilize the solution using a 0.22-µm filtering apparatus. Store aliquots at -20°C for up to 1 year; while in use, store at room temperature for up to one month.

8% (19:1) polyacrylamide gel – 7 M urea stock solution

200 ml 40% acrylamide/bis (19:1) solution (Bio Rad)
80 ml 10 x TBE
336 g Urea
Add water to 800 ml, and sterilize the solution using a 0.22-µm filtering apparatus. Store at 4°C for up to three months.

For the preparation of a 20 cm x 40 cm size gel (2-mm thick), mix 140 ml of the gel stock solution with 140 µl of TEMED. Just before pouring, mix the solution with 700 µl of 10% (w/v) ammonium persulfate (APS) by gentle swirling, and immediately pour the solution into the gel plate assembly. For a 16 cm x 16 cm size gel (2-mm thick), use 40 ml of the stock solution mixed with 40 µl of TEMED and 200 µl of 10% APS. When preparing a denaturing polyacrylamide gel for purification purposes, making the gel the day before it will be used is recommended to ensure complete polymerization, although the gel generally polymerizes within 30 min.

8% (29:1) polyacrylamide-5% glycerol gel stock solution

100 ml 40% acrylamide/bis (29:1) solution (Bio-Rad)
25 ml 10 x TBE
25 ml Glycerol
30 or 90 g Urea (1 M or 3 M final)
Add water to 500 ml, and sterilize the solution using a 0.22-µm filtering apparatus. Store at 4°C for up to three months.

For the preparation of a 16 cm x 16 cm size gel (1-mm thick), mix 20 ml of the gel stock solution with 20 µl of TEMED. Just before pouring, mix the solution with 100 µl of 10% APS by gentle swirling, and immediately pour the solution into the gel plate assembly. The gel generally polymerizes within 30 min.

COMMENTARY

Background Information

The initial idea of genetic alphabet expansion was mentioned in a review paper written by Alexander Rich in 1962, in which an unnatural base pair between isoguanine (iG) and isocytosine (iC) was designed as an alternative pair with a different hydrogen bonding pattern from those of the A–T and G–C pairs (Rich, 1962). In 1989, the iG–iC pair was examined by Benner's group and shown to be replicable in DNA duplexes by polymerase reactions (C. Switzer et al., 1989). However, the fidelity and efficiency of the iG–iC pairing in replication was not high, because of the mispairing with natural bases by the tautomerization of iG (Lutz et al., 1998; Robinson et al., 1998; C. Y. Switzer et al., 1993), as well as the chemical instability of the iC nucleoside (Benner, 2004). However, the iG–iC pair was applied to real-time quantitative PCR for multiplex PCR, called the Plexor system (Sherrill et al., 2004).

In 1998, Kool's group chemically synthesized surrogates of the natural A–T and G–C base pairs, such as Z–F, in which the hydrogen bonding moieties were removed, and demonstrated that non-hydrogen-bonded base pairs can function in replication, revealing the importance of the shape complementarity between pairing bases for polymerase recognition (Morales and Kool, 1998). Furthermore, in 1999, Romesberg's group presented that hydrophobic bases, called PICS, selectively formed self-pairs in polymerase reactions (McMinn et al., 1999). Although these Z–F and PICS–PICS pairs could not be used as a third base pair in replication, these studies expanded and supported the design concepts of unnatural base pairs.

In 2002, we developed the s–y pair by combining the concepts of the different hydrogen bonding pattern and the shape complementarity (Hirao et al., 2002). The s–y pair can be used in transcription and translation, and employed for the site-specific incorporation of an unnatural amino acid into a protein. In 2006, we newly developed a hydrophobic unnatural base pair, Ds–Pa, by removing the hydrophilic residues from the s–y pair, and the Ds–Pa pair displayed high specificity as a third base pair in replication and transcription (Hirao et al., 2006). By fine-tuning the Ds–Pa pair, the Ds–Px pair, which exhibits high fidelity (>99.9% selectivity per duplication) as a third base pair in PCR amplification was developed in 2009–2012 (Kimoto et al., 2009; Yamashige et al., 2012). Thus, in 2013, we applied the Ds–Px pair to DNA aptamer generation by SELEX, as described here (Kimoto et al., 2013).

In 2007–2011, Benner's group also developed an unnatural hydrogen-bonded base pair, Z–P, in which the chemical structure of Z is different from that of Kool's Z base, by addressing the problems of the iG–iC pair (Z. Yang et al., 2011b; Z. Yang et al., 2007). The Z–P pair showed high selectivity (~99.8% per PCR cycle) in PCR amplification, and was applied to SELEX targeting cancer cells in 2014 (Sefah et al., 2014).

In 2009–2012, Romesberg's group developed a series of hydrophobic unnatural base pairs, such as NaM–5SICS and MMO2–5SICS, with high fidelity (99.0–99.99% per duplication) in replication, as well as transcription, by screening a huge number of hydrophobic base analogues (Malyshev et al., 2012; Malyshev et al., 2009). Furthermore, in 2013, they reported other hydrophobic base pairs, such as NaM–TPT3, for efficient PCR amplification (Li et al.,

2014). They successfully created a semi-synthetic organism, using the combination of the NaM–5SICS and NaM–TPT3 pairs (Malyshev et al., 2014; Y. Zhang et al., 2017).

Critical Parameters and Troubleshooting

Design of the Ds-predetermined DNA library (Basic Protocol 1)

To maximize the yield of the DNA chemical synthesis, we usually design the initial DNA libraries with several-nucleotide shorter versions, by trimming the 5′- and 3′-ends of both primer parts. Note that the direct use of short primer sequences in our ExSELEX is precluded, as they would be digested by the proofreading activity of AccuPrime *Pfx* DNA polymerase, and thus the use of longer primers (25 nucleotides or longer) is recommended.

Amplification of DNA containing the Ds–Px pair by AccuPrime *Pfx* DNA polymerase can reduce the sequence dependency in PCR amplification, in comparison with that by Deep Vent DNA polymerase, but some dependency still exists. To relieve the undesired PCR amplification bias in the repetitive rounds of selection and amplification, the two Ds positions in each sub-library should be separated by at least five nucleotides (Kimoto et al., 2009). In addition to the design of a Ds-predetermined DNA library containing two Ds bases, a library containing three Ds bases is also acceptable, since our anti-vWF aptamers contain two or three Ds bases in the enriched sequences (Matsunaga et al., 2017). However, the possible number of combinations of Ds positions for each sub-library increases, and choosing the suitable positions of Ds in each sub-library is still a challenging issue.

In this protocol, we described the DNA library design with the complementary stem sequences outside of the randomized region, and part of the stem sequences is used as a bar-code. However, the design of a DNA library without a stem, and with just a two- or three-nucleotide bar-code sequence, is also acceptable depending on the targets.

Initial library size used in ExSELEX (Basic Protocol 2)

As one of the common issues in SELEX, and not simply restricted to ExSELEX, the initial DNA library diversity affects the success rates of the high-affinity DNA aptamer generation. In Basic Protocol 2, we described a selection procedure using a 300-pmol initial DNA library (around 1.8×10^{14} species). In our previous ExSELEX method targeting vWF-A1, we performed six sets of the described selection simultaneously (corresponding to around 1.1×10^{15} species) to increase the library diversity. Theoretically, all of the sequences cannot be covered in one batch of ExSELEX, due to the experimental limitation of the reasonable size of the library, when using a library containing more than 25 bases in the randomized region. Thus, the second ExSELEX using a doped library could optimize the sequences. Furthermore, the doped selection provides secondary structure information (Kimoto et al., 2013), which is useful for further optimization, such as the elimination of some sequence parts that are unnecessary for binding to the target and stabilization by implementing a stable mini-hairpin DNA in a stem-loop structure (Kimoto et al., 2016b; Matsunaga et al., 2015; Matsunaga et al., 2017). To avoid cross-contamination issues

by repeating ExSELEX to the same target, the use of a forward primer with a different sequence is recommended.

Selection conditions (Basic Protocol 3)

Setting up proper selection conditions is important for the successful generation of high-affinity DNA aptamers. Stringent selection conditions are essential to efficiently select the tightest binders through several rounds of selection, but are not recommended in the early rounds of selection since the copy number of such tight binders still might not be high enough to isolate them from the DNA-protein complex. In particular, the selection conditions are largely dependent on the properties of the target proteins or materials, and thus the binding buffer, binding temperature, and incubation time should be optimized in each target's case. For the washing process in the selection step, increasing the temperature, increasing the salt concentrations, and/or adding urea (final 3 M concentration) might be employed, since we found that our high-affinity Ds-containing DNA aptamers still bound to the target under such conditions. We usually employ 50 mM NaOH for eluting DNA species from the DNA-target complexes, because such tight-binding aptamers are regenerated in SPR with 50 mM NaOH. The repetitive rounds of selection in the same method might also easily increase the background species. From the second round, a pre-counter selection can be applied. After several rounds, a post-counter selection is also useful. Another strategy to reduce the background species would be to use different selection methods alternatively. As an example, the selection conditions used in our previous ExSELEX targeting vWF-A1 are shown in Table 1 (Matsunaga et al., 2017).

PCR mutations involving unnatural bases (Basic Protocols 3 and 5)

Through repetitive rounds of ExSELEX, the number of PCR cycles in the total ExSELEX rounds should be less than 150 cycles. Using longer PCR cycles, the possibility of a mutation from unnatural to natural bases might increase. In addition, the mutation from natural to unnatural bases also occurs, as seen in our ExSELEX targeting vWF-A1 (Matsunaga et al., 2017), although the possibility is very low. This is very critical when using the predetermined library, and thus, careful examinations including further sequencing analysis of the generated aptamers are required (Matsunaga et al., 2017).

Evaluation of binding of enriched DNA libraries by the gel-shift assay (Basic Protocol 4)

To evaluate the binding affinity and specificity of the enriched DNA library, the initial library should also be used as a negative control. In addition, if available, a conventional natural-base DNA aptamer reported previously should be used as a positive control, to check the reliability of the experiments. Depending on the target's size and properties, the native PAGE conditions in the gel-shift assay should be optimized, such as the percentage of the polyacrylamide (*i.e.*, lower percentages, ~6% or 4%, for larger sized targets), the inclusion of additional salts (potassium, sodium, or magnesium, if required for the binding), and the compositions and pH values of the running buffer and gel. If a target protein is highly basic, it might be difficult for it

to enter the gel without binding to its DNA aptamer during gel electrophoresis. When the aptamer tightly binds to the target, the DNA-target complex can enter the gel, due to the negative charge of the DNA. To confirm the protein migration in the gel-shift assay, CBB staining of the gel after SYBR Gold detection is useful. When the binding affinities of aptamers are too high, the evaluation of the binding capability from the band patterns would be difficult. In such a case, increasing the electrophoresis temperature, adding 1 M or 3 M urea in the gel, or reducing the DNA and target concentrations for the incubation might be effective (Matsunaga et al., 2017). If the gel-shift assays work well, a gel-shift method can also be employed to separate the DNA-target complex, instead of the pull-down method using magnetic beads described in Basic Protocol 2, as shown in Table 1.

Anticipated Results

The recovery yields of single-stranded DNA fragments by denaturing polyacrylamide gel purification, using the gel-crushing and soaking method described here, are around 30% to 60%. In Basic Protocol 1, a portion of each sub-library, around 2 to 5 nmol each, is generally used for the preparation of the mixed library and storage in 2 nmol aliquots (10 µM, 0.2 ml) is useful for experiments.

The recovery yields of DNA species in Basic Protocol 2 are directly correlated to the number of PCR cycles. To prepare a sufficient amount of DNA libraries for the following round of ExSELEX, around 15 to 25 cycles of PCR are required. Performing ExSELEX **without** the target would be useful to know the levels of background noise and contamination by other DNA species. In general, the enrichment of the library can be confirmed by decreasing the number of PCR cycles in later rounds, especially under the same selection conditions used for the consecutive rounds.

Typical gel band patterns of PCR-amplified DNA libraries in denaturing PAGE (Basic Protocol 3) are shown in Figure 4. After PCR amplification using the Dss-labeled forward primer, only the Ds-strand DNAs (shorter fragments) can be detected under the 365-nm UV light, due to the Dss-fluorescence. Currently, the 5'-end fluorescent labeling of DNA libraries is also commonly used for the characterization of their binding properties. The detection limit of the Dss fluorescence is higher than that of UV shadowing.

Figure 5 shows examples of the gel-shift assay, in which the enrichment of the DNA libraries in each round can be monitored. From the band density of the complex, the binding affinity, especially related to k_{off}, of the isolated DNA library in each round can be assessed. If the interaction between the DNA and the target is easily dissociated (*i.e.*, weak binding), then the shifted bands corresponding to the complex formation would not be detected clearly, sometimes providing only smearing band patterns. Experimentally, when shifted bands can be detected in the presence of the target in the gel-shift assay, their Kd values are generally less than several nM order at most, with 100 nM of each DNA and the target concentration. Note that the dissociation rate largely dominates the shifted band patterns in the gel-shift assay, and thus to obtain association rate information, other types of assays, such as SPR and Bio-Layer Interferometry, are recommended.

Time Considerations

The design and chemical synthesis of the sub-libraries require one to three weeks, depending on the number of sub-libraries. Purification, including denaturing PAGE, elution, and ethanol precipitation, of each chemically synthesized sub-library takes 3 days, but can be done simultaneously for several sub-libraries. The selection process described here took around 3 to 5 hours, and the qPCR and large-scale PCR require around 3 hours. Purification of the PCR-amplified DNA library requires about 2 hours for PAGE. The DNA elution from the gels takes about 9 hours (or overnight), and around 3 hours are needed for the recovery of the eluted DNAs by ethanol precipitation. Completion of ExSELEX to obtain enriched libraries requires 2 to 4 weeks, although it depends on the number of ExSELEX rounds. Gel-shift assays will take around 3 hours, and replacement PCR can be completed in about 1 hour.

ACKNOWLEDGEMENT

This work was supported by the Institute of Bioengineering and Nanotechnology (Biomedical Research Council, Agency for Science, Technology and Research, Singapore).

LITERATURE CITED

Bell, N. M., and Micklefield, J. 2009. Chemical modification of oligonucleotides for therapeutic, bioanalytical and other applications. *Chembiochem, 10*, 2691-2703.

Benner, S. A. 2004. Understanding nucleic acids using synthetic chemistry. *Acc. Chem. Res., 37*, 784-797.

Benner, S. A., Karalkar, N. B., Hoshika, S., Laos, R., Shaw, R. W., Matsuura, M., Fajardo, D., and Moussatche, P. 2016. Alternative Watson-Crick Synthetic Genetic Systems. *Cold Spring Harb. Perspect Biol., 8*.

Biondi, E., Lane, J. D., Das, D., Dasgupta, S., Piccirilli, J. A., Hoshika, S., Bradley, K. M., Krantz, B. A., and Benner, S. A. 2016. Laboratory evolution of artificially expanded DNA gives redesignable aptamers that target the toxic form of anthrax protective antigen. *Nucleic Acids Res., 44*, 9565-9577.

Diafa, S., and Hollenstein, M. 2015. Generation of Aptamers with an Expanded Chemical Repertoire. *Molecules, 20*, 16643-16671.

Ellington, A. D., and Szostak, J. W. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature, 346*, 818-822.

Hirao, I., and Kimoto, M. 2012. Unnatural base pair systems toward the expansion of the genetic alphabet in the central dogma. *Proc. Jpn. Acad. Ser. B, Phys. Biol. Sci., 88*, 345-367.

Hirao, I., Kimoto, M., Mitsui, T., Fujiwara, T., Kawai, R., Sato, A., Harada, Y., and Yokoyama, S. 2006. An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods, 3*, 729-735.

Hirao, I., Ohtsuki, T., Fujiwara, T., Mitsui, T., Yokogawa, T., Okuni, T., Nakayama, H., Takio, K., Yabuki, T., Kigawa, T., Kodama, K., Yokogawa, T., Nishikawa, K., and Yokoyama, S. 2002. An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol., 20*, 177-182.

Houlihan, G., Arangundy-Franklin, S., and Holliger, P. 2017. Engineering and application of polymerases for synthetic genetics. *Curr. Opin. Biotechnol., 48*, 168-179.

Jung, C., and Ellington, A. D. 2015. Synthetic biology: Six pack and stack. *Nat. Chem., 7*, 617-619.

Keefe, A. D., and Cload, S. T. 2008. SELEX with modified nucleotides. *Curr. Opin. Chem. Biol., 12*, 448-456.

Kimoto, M., Kawai, R., Mitsui, T., Yokoyama, S., and Hirao, I. 2009. An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules. *Nucleic Acids Res., 37*, e14.

Kimoto, M., Matsunaga, K., and Hirao, I. 2016a. DNA Aptamer Generation by Genetic Alphabet Expansion SELEX (ExSELEX) Using an Unnatural Base Pair System. *Methods Mol. Biol., 1380*, 47-60.

Kimoto, M., Mitsui, T., Yamashige, R., Sato, A., Yokoyama, S., and Hirao, I. 2010. A new unnatural base pair system between fluorophore and quencher base analogues for nucleic acid-based imaging technology. *J. Am. Chem. Soc., 132*, 15418-15426.

Kimoto, M., Nakamura, M., and Hirao, I. 2016b. Post-ExSELEX stabilization of an unnatural-base DNA aptamer targeting VEGF165 toward pharmaceutical applications. *Nucleic Acids Res., 44*, 7487-7494.

Kimoto, M., Yamashige, R., Matsunaga, K., Yokoyama, S., and Hirao, I. 2013. Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nat. Biotechnol., 31*, 453-457.

Lapa, S. A., Chudinov, A. V., and Timofeev, E. N. 2016. The Toolbox for Modified Aptamers. *Mol. Biotechnol., 58*, 79-92.

Li, L., Degardin, M., Lavergne, T., Malyshev, D. A., Dhami, K., Ordoukhanian, P., and Romesberg, F. E. 2014. Natural-like replication of an unnatural base pair for the expansion of the genetic alphabet and biotechnology applications. *J. Am. Chem. Soc., 136*, 826-829.

Lutz, M. J., Horlacher, J., and Benner, S. A. 1998. Recognition of a non-standard base pair by thermostable DNA polymerases. *Bioorg. Med. Chem. Lett., 8*, 1149-1152.

Malyshev, D. A., Dhami, K., Lavergne, T., Chen, T., Dai, N., Foster, J. M., Correa, I. R., Jr., and Romesberg, F. E. 2014. A semi-synthetic organism with an expanded genetic alphabet. *Nature, 509*, 385-388.

Malyshev, D. A., Dhami, K., Quach, H. T., Lavergne, T., Ordoukhanian, P., Torkamani, A., and Romesberg, F. E. 2012. Efficient and sequence-independent replication of DNA containing a third base pair establishes a functional six-letter genetic alphabet. *Proc. Nat. Acad. Sci. USA, 109*, 12005-12010.

Malyshev, D. A., and Romesberg, F. E. 2015. The Expanded Genetic Alphabet. *Angew. Chem. Int. Ed. Engl., 54*, 11930-11944.

Malyshev, D. A., Seo, Y. J., Ordoukhanian, P., and Romesberg, F. E. 2009. PCR with an expanded genetic alphabet. *J. Am. Chem. Soc., 131*, 14620-14621.

Matsunaga, K., Kimoto, M., Hanson, C., Sanford, M., Young, H. A., and Hirao, I. 2015. Architecture of high-affinity unnatural-base DNA aptamers toward pharmaceutical applications. *Sci. Rep., 5*, 18478.

Matsunaga, K., Kimoto, M., and Hirao, I. 2017. High-Affinity DNA Aptamer Generation Targeting von Willebrand Factor A1-Domain by Genetic Alphabet Expansion for Systematic Evolution of Ligands by Exponential Enrichment Using Two Types of Libraries Composed of Five Different Bases. *J. Am. Chem. Soc., 139*, 324-334.

McMinn, D. L., Ogawa, A. K., Wu, Y., Liu, J., Schultz, P. G., and Romesberg, F. E. 1999. Efforts toward expansion of the genetic alphabet: DNA polymerase recognition of a highly stable, self-pairing hydrophobic base. *J. Am. Chem. Soc., 121*, 11585-11586.

Morales, J. C., and Kool, E. T. 1998. Efficient replication between non-hydrogen-bonded nucleoside shape analogs. *Nat. Struct. Biol., 5*, 950-954.

Okamoto, I., Miyatake, Y., Kimoto, M., and Hirao, I. 2016. High Fidelity, Efficiency and Functionalization of Ds-Px Unnatural Base Pairs in PCR Amplification for a Genetic Alphabet Expansion System. *ACS Synth. Biol., 5*, 1220-1230.

Rich, A. (1962). Problems of evolution and biochemical information transfer. In M. Kasha & B. Pullman (Eds.), *Horizons Biochem.* (pp. 103-126): Academic Press.

Robinson, H., Gao, Y. G., Bauer, C., Roberts, C., Switzer, C., and Wang, A. H. 1998. 2'-Deoxyisoguanosine adopts more than one tautomer to form base pairs with thymidine observed by high-resolution crystal structure analysis. *Biochemistry, 37*, 10897-10905.

Sefah, K., Yang, Z., Bradley, K. M., Hoshika, S., Jimenez, E., Zhang, L., Zhu, G., Shanker, S., Yu, F., Turek, D., Tan, W., and Benner, S. A. 2014. In vitro selection with artificial expanded genetic information systems. *Proc. Natl. Acad. Sci. U S A, 111*, 1449-1454.

Sherrill, C. B., Marshall, D. J., Moser, M. J., Larsen, C. A., Daude-Snow, L., Jurczyk, S., Shapiro, G., and Prudent, J. R. 2004. Nucleic Acid Analysis Using an Expanded Genetic Alphabet to Quench Fluorescence. *J. Am. Chem. Soc., 126*, 4550-4556.

Switzer, C., Moroney, S. E., and Benner, S. A. 1989. Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc., 111*, 8322-8323.

Switzer, C. Y., Moroney, S. E., and Benner, S. A. 1993. Enzymatic recognition of the base pair between isocytidine and isoguanosine. *Biochemistry, 32*, 10489-10496.

Tuerk, C., and Gold, L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science, 249*, 505-510.

Yamashige, R., Kimoto, M., Takezawa, Y., Sato, A., Mitsui, T., Yokoyama, S., and Hirao, I. 2012. Highly specific unnatural base pair systems as a third base pair for PCR amplification. *Nucleic Acids Res., 40*, 2793-2806.

Yang, Y., Ren, X., Schluesener, H. J., and Zhang, Z. 2011a. Aptamers: selection, modification and application to nervous system diseases. *Curr. Med. Chem., 18*, 4159-4168.

Yang, Z., Chen, F., Alvarado, J. B., and Benner, S. A. 2011b. Amplification, mutation, and sequencing of a six-letter synthetic genetic system. *J. Am. Chem. Soc., 133*, 15105-15112.

Yang, Z., Sismour, A. M., Sheng, P., Puskar, N. L., and Benner, S. A. 2007. Enzymatic incorporation of a third nucleobase pair. *Nucleic Acids Res., 35*, 4238-4249.

Zhang, L., Yang, Z., Le Trinh, T., Teng, I. T., Wang, S., Bradley, K. M., Hoshika, S., Wu, Q., Cansiz, S., Rowold, D. J., McLendon, C., Kim, M. S., Wu, Y., Cui, C., Liu, Y., Hou, W., Stewart, K., Wan, S., Liu, C., Benner, S. A., and Tan, W. 2016. Aptamers against Cells Overexpressing Glypican 3 from Expanded Genetic Systems Combined with Cell Engineering and Laboratory Evolution. *Angew. Chem. Int. Ed. Engl., 55*, 12372-12375.

Zhang, L., Yang, Z., Sefah, K., Bradley, K. M., Hoshika, S., Kim, M. J., Kim, H. J., Zhu, G., Jimenez, E., Cansiz, S., Teng, I. T., Champanhac, C., McLendon, C., Liu, C., Zhang, W., Gerloff, D. L., Huang, Z., Tan, W., and Benner, S. A. 2015. Evolution of functional six-nucleotide DNA. *J. Am. Chem. Soc., 137*, 6734-6737.

Zhang, Y., Lamb, B. M., Feldman, A. W., Zhou, A. X., Lavergne, T., Li, L., and Romesberg, F. E. 2017. A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proc. Natl. Acad. Sci. USA, 114*, 1317-1322.

FIGURE LEGENDS

Figure 1. Chemical structures of unnatural Ds–Px and natural A–T and G–C pairs, Dss phosphoramidite, and dPa'TP.

Figure 2. Overview of ExSELEX using the Ds-predetermined DNA library. The Ds-predetermined DNA library consists of a mixture of sub-libraries. Each sub-library contains a bar-code sequence

in the constant region, indicating the two Ds positions in the randomized region (Basic Protocol 1 and Figure 3). The 5'-end of the DNA library is fluorescently-labeled with an unnatural fluorescent Dss base, by performing PCR with a Dss-labeled forward primer. The reverse primer used for PCR has a spacer (wavy line) and a linker (oligo T_{15} sequence) at the 5'-end, and elongation of the Ds-strand is stopped at the position opposite the spacer, which allows Ds-strand separation in denaturing PAGE, as described in Basic Protocol 3. The enriched library was subjected to PCR amplification without dDsTP and dPxTP (replacement PCR, Basic Protocol 5), for further processing with conventional DNA sequencing methods.

Figure 3. Schematic illustration of the N26Ds library, a mixture of 28 sub-libraries, focusing on each bar-code sequence and the Ds positions.

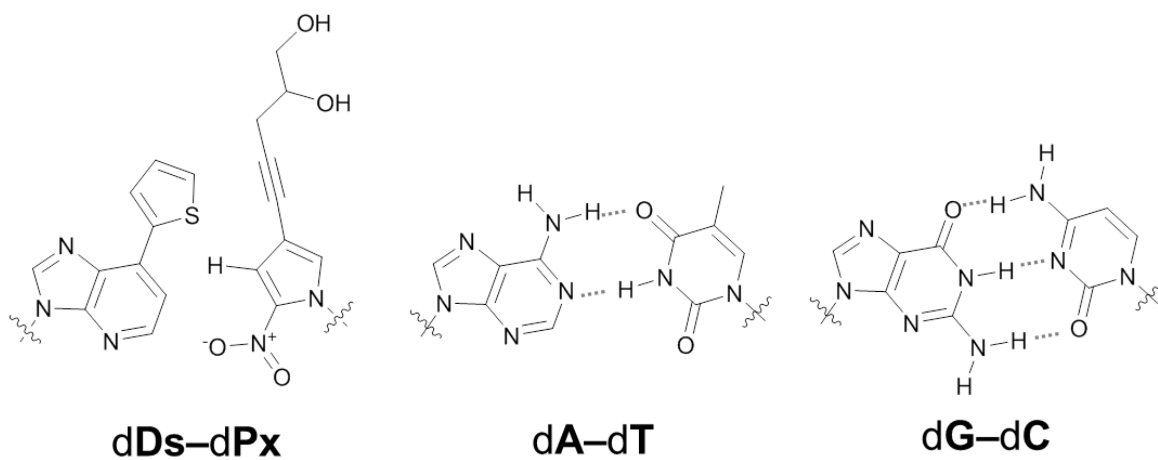
Figure 4. Band patterns of the PCR-amplified DNA library on a denaturing PAGE. DNA bands were detected with UV shadowing (left) and Dss fluorescence (right). In each detection method, the left and right images show the gel before and after the excision of the gel band corresponding to the Ds-strand (DNA library).

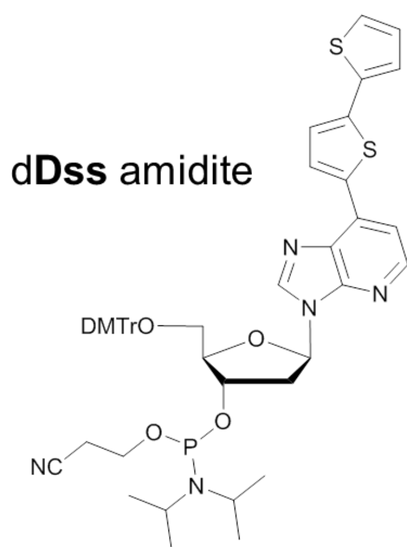
Figure 5. Binding analysis of DNA libraries by gel-shift assays. (A) The DNA library in each round (100 nM) and the target protein, vWF-A1 (100 nM), was incubated at 37°C for 30 min in binding buffer, and the complexes were separated from the free DNA on a native 8% polyacrylamide gel (upper panel) or an 8% polyacrylamide gel in the presence of 3 M urea. (B) The DNA library in each round (100 nM) was incubated at 25°C for 30 min in binding buffer, in the presence (100 nM) and absence of vWF-A1, and then analyzed on an 8% polyacrylamide gel in the presence of 1 M urea (upper panel) or 3 M urea (lower panel). The DNA bands on the gels were stained with SYBR Gold, and their band patterns were detected with a bio-imaging analyzer (Fuji Film LAS4000).

Table 1. Conditions for ExSELEX targeting vWF-A1

Round	Selection method	DNA		vWF-A1	Number of washes		PCR cycles
		pmol	(nM)	(nM)	Without Urea	With 3 M Urea	
1	Pull-down	300	50	25	5	-	25
2	Pull-down	50	25	10	5	-	23
3	Gel-shift 4°C (PAGE, native)	25	500	2000	-	-	12
4	Gel-shift at 4°C (PAGE, native)	5	250	250	-	-	16
5	Gel-shift at 4°C (PAGE, native)	5	125	125	-	-	16
6	Pull-down	3	1	1	5	-	25
7	Pull-down	3	1	1	3	2	17
8	Gel-shift at 4°C (PAGE, 3 M urea)	5	100	100	-	-	26

Figure 1





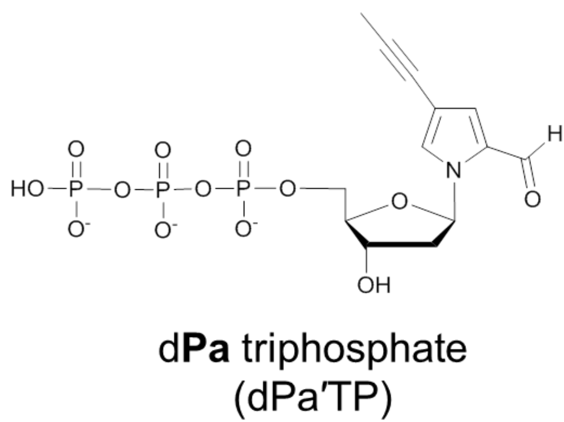


Figure 2

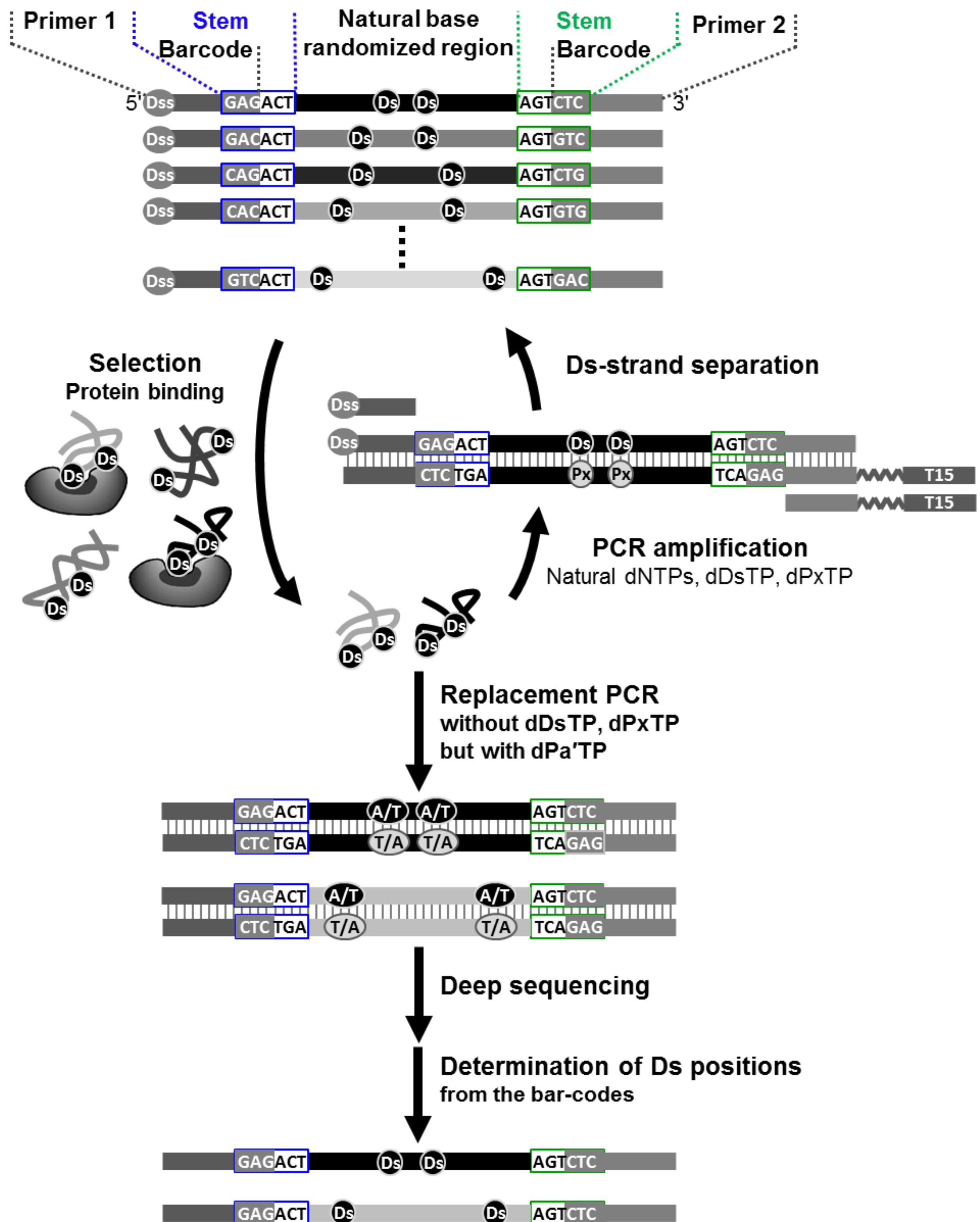


Figure 3

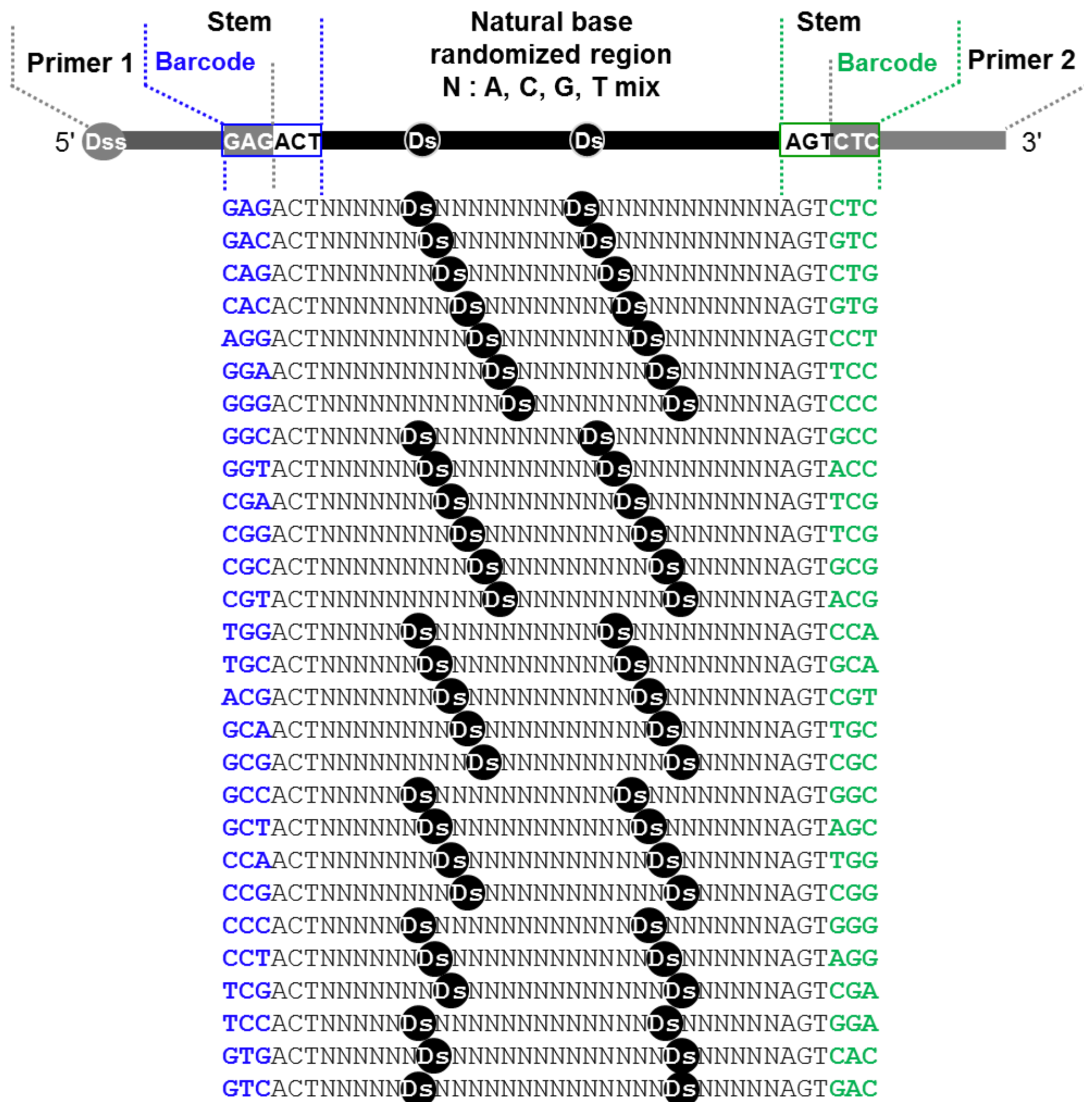


Figure 4

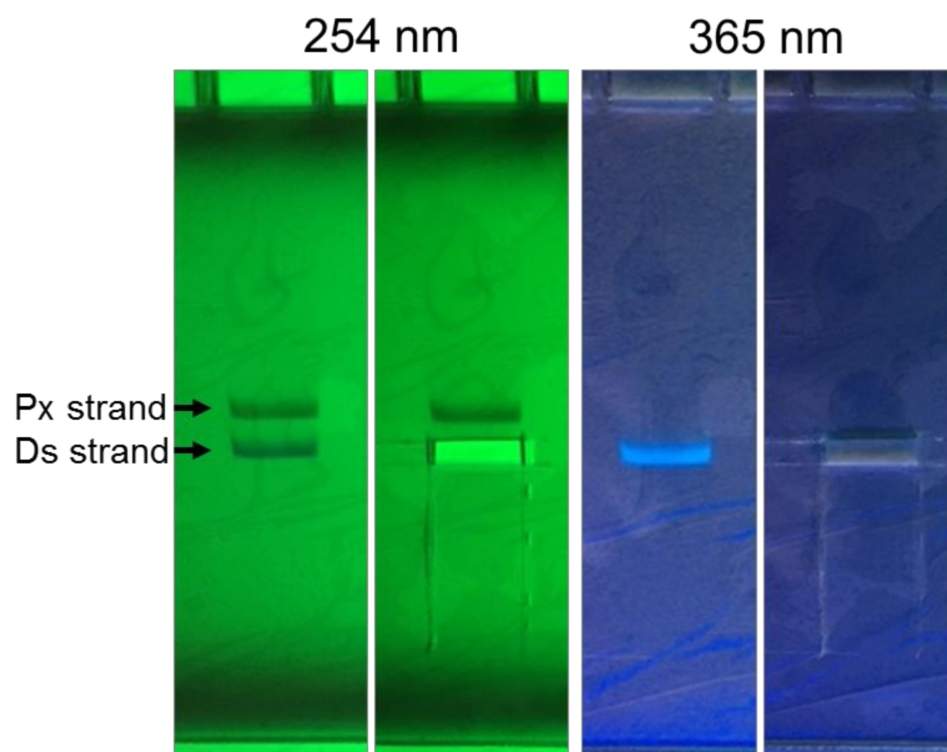


Figure 5A

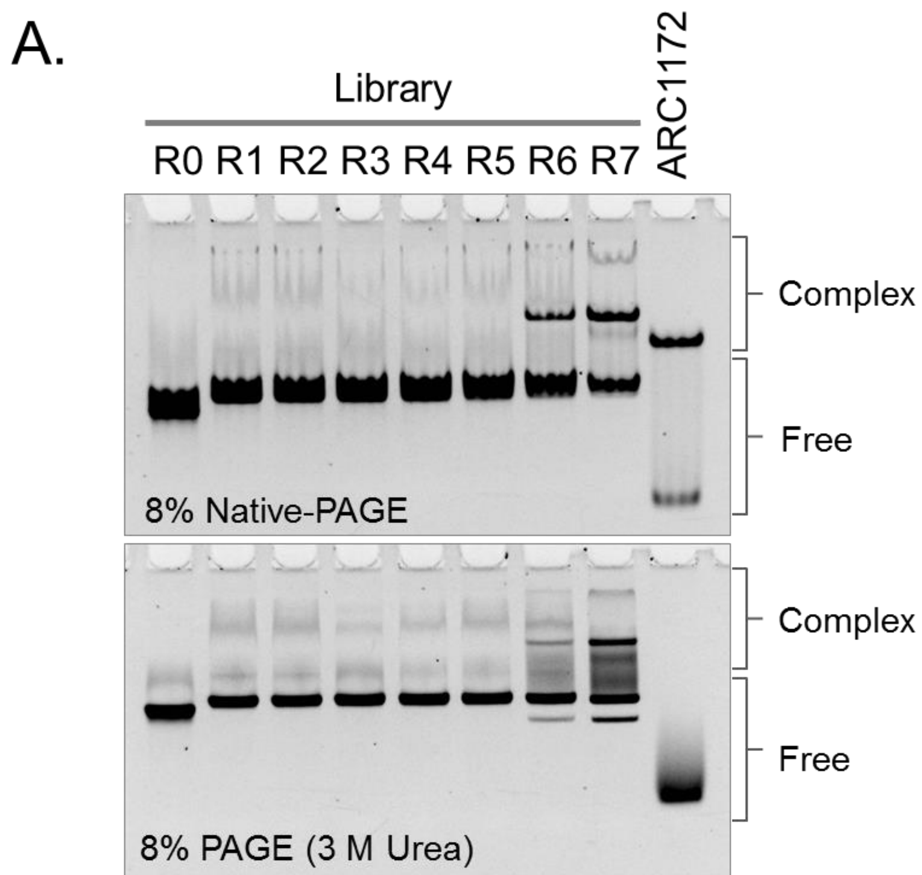


Figure 5B