Rapid Assembly and Cloning of Zinc Finger Proteins with Multiple Finger Modules

So-Young Park and Yang-Gyun Kim[*]
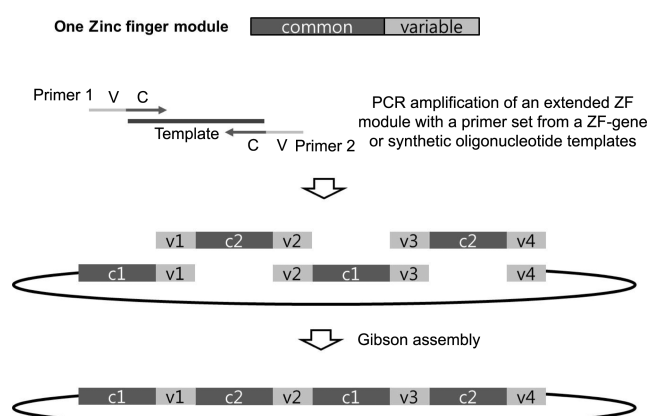
Department of Chemistry, Sungkyunkwan University, Suwon 440-746, Korea. [*]E-mail: ygkimmit@skku.edu
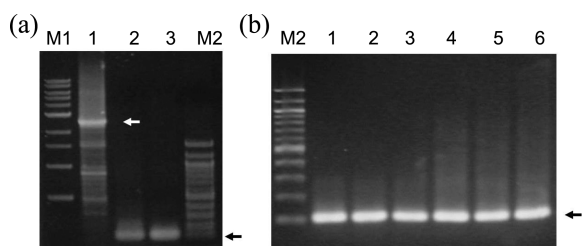
The C2H2-type zinc finger (ZF) is one of the most abundant structural motifs among eukaryotes, exceeding over 3% of the whole human genes.[1] Primarily, the ZF plays an important role in recognizing nucleic acid targets including DNA and RNA in a sequence-specific fashion. In-depth studies of the ZF domain from the transcription factor IIIA have provided vital informations about the ZF on the nucleic acid recognition principle.[2] A ZF recognizes DNA principally by one-to-one interaction between amino acids and DNA bases.[3-7] Amino acids at helical positions −1, +3, and +6 of the recognition helix form specific interactions with three consecutive DNA bases on one strand of the DNA, although there is an additional interaction from the position +2 to the other strand.[8] Many studies demonstrate that ZF can be modularly engineered to alter their sequence-specificity.[9-13] In particular, the phage-display is a powerful technique to screen and select novel sequence-specific ZF peptides that bind to various DNA sequences. Accumulating data in engineered ZFs demonstrates that almost any DNA sequences can be recognized by one-to-one interaction between ZFs and triplet bases.[14] Accordingly, the ZFs are an ideal frame for designing new DNA binding proteins to recognize any given sequence of DNA.
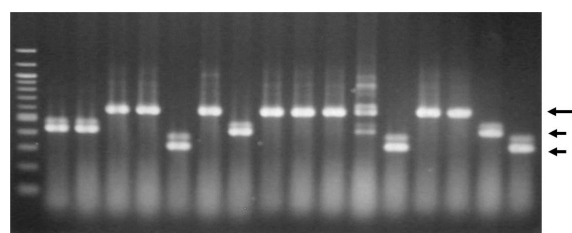

Scheme 1. Rapid assembly of zinc finger protein based on the Gibson assembly. PCR primers to amplify one-finger module have a variable region at the 5' end and a common region at the 3' end. The variable region encodes an α-helix region of the zinc finger, which is responsible for specific interaction with DNA bases. DNAs are amplified using a ZF-backbone templates. Each DNA contains common region (c1-c4) in the middle and variable regions (v1-v4) at both ends, which work as overlapping ends for the Gibson assembly.

Using the engineered ZFs with novel sequence-specificity, chimeric proteins that fused to effector domains including transcription regulators,[15-17] nucleases[18] and recombinases[19,20] can be applied to regulating gene expression, gene targeting and gene modification. One prominent example is Zinc finger nuclease (ZFN) that is a chimeric endonuclease made of a C2H2-type ZFP for recognizing target DNA sequences and the *Fok*I nuclease domain (F_N) for target-specific DNA cleavage.[18] Although ZFP is a useful platform for designing and creating new DNA binding proteins with tailor-made sequence specificity, there is still room for improvement in the construction method that requires facile and rapid building strategy for creating multi-finger ZFPs. Recent methods for constructing ZFPs are the assembly from overlapping oligonucleotides by PCR,[21] the restriction digest-based modular assembly through consecutive restriction and ligation reactions,[22,23] and OLTA (OverLap extension PCR and TA-cloning) method.[24]

In this work, we utilized the Gibson assembly,[25] a multiple homologous sequences-based system, to synthesize multi-finger ZFPs that can be used to construct chimeric proteins such as ZFNs. As a proof of concept, we carried out the construction of Zif268, which is a well-known 3-finger ZFP.[2] As described in Scheme 1, we divided one ZF module into two regions, a variable region (bright grey) and a common region (dark grey). The variable region represents seven consecutive amino acids in the DNA recognition helix (−1 to +6) that are responsible for sequence-specific DNA binding, whereas the common region covers conserved amino acids and linker amino acids in the ZF2. In our design, all PCR primers for amplification of one ZF module are 42-nt long and their 5'-end regions contain different 21-nucleotide sequences that encode different variable regions, specific for distinct ZF modules. In addition, one pair of PCR primer set was used to amplify one ZF module from the plasmid containing the Zif268 gene. Consequently, the resulting PCR products should have two different variable regions at both ends and a common region in the middle−for example, the PCR product PCR-ZF12 amplified from Z268-ZF1U and Z268-ZF2-D primers (Supporting Information, Table S1) has variable regions of ZF1 and ZF2 at both ends, respectively. These variable regions of PCR-amplified DNAs were used for overlapping regions in the Gibson assembly reaction. Templates were the Zif268-gene containing plasmid for ZF modules and the ZFP-cloning vector DNA for the vector DNA. Each PCR-amplified ZF DNA was gel-purified (Figure

Figure 1. PCR-amplified DNAs used for the construction of ZFPs. (a) Gel electrophoresis of PCR-amplified DNAs for the 3-finger ZFP. M1, 1-kb size marker; 1, the vector-PCR DNA amplified from ZFP-cloning vector (white arrow); 2-3, 105-bp ZF encoding DNAs amplified from ZF-backbone templates (arrow); M2, 100-bp size marker. (b) Gel electrophoresis of PCR-amplified DNAs for the 4-finger ZFP. M2, 100-bp size marker; 1-3 and 4-6, 105-bp ZF encoding DNA amplified from a ZF-backbone template for CCR5-L and CCR5-R, respectively (arrow).

1) or used directly after the removal of residual primers if there is no noticeable non-specific band. The vector DNA was also amplified with an appropriate primer set containing overlapping regions with other ZF DNAs and gel-purified (Table S1 and Figure 1(a)). These amplified-DNA fragments were assembled and ligated by the Gibson assembly and used for transformation *E. coli* cells for cloning. Our screening with colony-PCR revealed that the efficiency of correctly assembled clones generated from the Gibson assembly was 100% (25 correct clones out of 25 screened colonies).
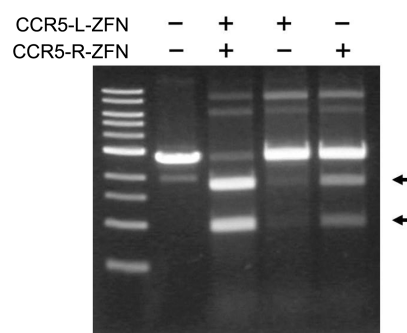
Encouraged by the 3-finger assembly result, next we attempted 4-finger construction using the same method. In the previous report,[26] a pair of CCR5-targeting ZFNs has reported that they accurately cleaved the internal target site (5'-<u>GTCATCCTCATC</u>CTGAT<u>AAACTGCAAAAG</u>-3') of the CCR5 gene. Using the almost identical amino acid sequences of the two reported CCR5-targeting ZFPs with few changes in the common regions, we carried out the construction of the CCR5-targeting ZFPs using PCR-amplified DNA fragments from synthetic templates and the ZFP-cloning vector DNA through the Gibson assembly. We used the essentially the same protocol for the 3-finger construction with the exception of the ZF module template. The synthetic 63-nt long ZF-template oligonucleotides (Table S1, TEMZ-HTK and TEMZ-HIR) that encode a common region of the typical ZF served as a template for the 4-finger construction. In fact, these synthetic templates can be used for any ZF module. At both ends, each DNA has a 21-nt long overlapping region that consists of DNA sequences encoding a variable region of different ZF modules. Thus, one vector DNA and three ZF module DNAs from PCR amplification produced a 4-finger ZFP clone in the vector (Scheme 1). Both CCR5-L and CCR5-R ZFP genes were constructed by the Gibson assembly described in the Experimental section. After the reaction was completed, DNAs were directly used for *E. coli* transformation. Plasmid DNAs from clones screened by the colony-PCR were isolated and analyzed by sequencing (Figure 2). The yield of correct clones that contain 4-finger ZFP gene was 11.9% for CCR5-L (5/42) and 20.2% for CCR5-R (24/119). We noticed that

Figure 2. Gel electrophoresis for a typical colony-PCR screening. With the 100-bp size marker, DNA products from the colony-PCR reactions for the 4-finger construction were electrophoresed to screen correct-size inserts (595-bp, Long arrow). Reactions from incorrect clones often gave smaller PCR products in size (small arrows).

there was significant variation in the success rate for different batches of the experiments, which may result from the quality of PCR products.

To demonstrate target specificity of CCR5-L and CCR5-R ZFPs to the CCR5 gene sequence, we used ZFNs containing these ZFPs as a DNA recognition domain. To construct CCR5-targeting ZFNs, the DNA fragments of the CCR5-targeting ZFP genes made by our method were inserted into the ZFN expression vector to fuse to the *Fok*I nuclease domain (FN). The expression of CCR5-targeting ZFNs (CCR5-L-ZFN and CCR5-R-ZFN) is under the control of T7 promoter. Using the *in vitro* transcription/translation system containing T7 RNA polymerase, these ZFN proteins were produced *in vitro*. The sequence-specific DNA cleavage of the CCR5-targeting ZFNs was confirmed by the digestion of the *Sca*I-linearized plasmid DNA containing CCR5 target sequences in the multicloning site of pUC19. As shown in Figure 3, two DNA fragments, 0.9 kb and 1.8 kb, would be produced if target-specific DNA cleavages occurred by the ZFNs. The control reaction incubated with only either of the CCR5-targeting ZFNs produced weak or almost no DNA cleavage products (Figure 3), because the dimerization of both CCR5-targeting ZFNs that recognize either side of the target sequence is required to have efficient dsDNA cleavage at the target DNA. In contrast, the reaction with two CCR5-targeting ZFNs showed almost 100% target-specific

Figure 3. Site-specific DNA cleavage experiment. The 2.7-kb *Sca*I-linearized CCR5 plasmid DNA substrate was incubated with *in-vitro* produced ZFNs. 1, no ZFN; 2, CCR5-L ZFN + CCR5-R ZFN; 3, CCR5-L ZFN; 4, CCR5-R ZFN. Arrows indicate sequence-specific cleavage of the 2.7-kb *Sca*I-linearized substrate DNA into 1.8 kb and 0.9 kb fragments.

DNA cleavages as expected.

In our design, the overlapped region among DNA fragments for the Gibson assembly is the variable region of the ZF. For each zinc finger, there is a DNA recognition helix region responsible for sequence-specificity to DNA, whereas highly conserved region consists of the structural backbone. The former consists of 7 amino acids that are highly diverse among ZFs. Accordingly, the variable region can be used as a disparate sequence region in the PCR primers to generate a distinct DNA fragment sharing the same overlapping region at one end with other DNA fragment to be used for the Gibson assembly. Taking an advantage of this, a DNA with two specific matching regions can be generated PCR and assembled by the Gibson assembly reaction. In addition, the conserved region and the linker region in ZF are used as a template frame for PCR. 63-nt templates and 42-nt primers are readily synthesized in high quality by the conventional phosphoramidite method[27] that used by most of suppliers. Another advantage to use oligonucleotides as a backbone PCR template is that variations in amino acids, if necessary, can be easily introduced.

In summary, we developed the simple and facile method to construct ZFPs that recognize target DNA sequences in a sequence-specific manner. Multi-finger ZFPs can be generated by PCR from the template oligonucleotide and PCR primers containing disparate sequence at the 5'-end for corresponding the DNA recognition helix of the ZF and the following the Gibson assembly reaction. The usefulness of ZFP as a tailor-made DNA binding protein with desired sequence-specificity leads to many and diverse applications. Regulation of gene expression and the modification of genetic information are very attractive goals in the field of medical and biological sciences. Chimeric proteins of the ZFP protein with one of various effector domains are effective tools to achieve such goals. Thus, our method developed here certainly contributes and facilitates these researches.

Experimental

PCR Amplification. For convenience and efficiency of the vector PCR, the vector template DNA was constructed into pJHU-1 vector (a gift from professor Srinivasan Chandrasegaran at department of Environmental Health Sciences of the Johns Hopkins University), a 2-kb minimal size cloning vector. Two ZF template oligonucleotides and all PCR primers were purchased from Bioneer Cooperation. For a typical PCR reaction, the reaction volume was 50 μL containing 5 μL of the 10X *Pfu* buffer, 4 μL of dNTP (2.5 mM each), 1 μL of 10 μM primer mix, 1 μL of 50 ng/μL template DNA and 1 μL of *Pfu* polymerase (5 U/μL) (Labo Pass). For vector PCR, the reaction mixture was incubated at 94 °C 5 min, and 30 cycles of denaturation (94 °C for 1 min), annealing (57 °C for 1 min) and extension (72 °C for 4 min) were performed. And, the reaction mixture was kept at 72 °C for 10 min. For the ZF PCR, the reaction condition was same except the extension time, which was 1 min. After PCR, reactions were subjected to *Dpn*I digestion at 37 °C for 4 h to remove template plasmid DNAs, if necessary. Each PCR product was subjected to electrophoresis and size-checked in agarose gel.

Gibson Assembly and Cloning. Vector DNA fragments were gel-isolated by MEGA quick-spin™ total kit (iNtRON). The gel-isolated vector fragments and two (or three) PCR products of ZF modules for 3 (or 4)-finger ZFP construction. For the Gibson assembly, Gibson Assembly™ master mix was used (New England Biolab). In this assay, the reaction volume was 20 μL containing 10 μL of the 2X Gibson Assembly Master Mix and PCR-amplified DNA fragments, 0.2 pmole of vector fragment and 0.6 pmole of each ZF module fragment. The reaction was incubated for 1 h at 50 °C. After incubation, the DNA was used for transformation by electroporation. For expression of CCR5-L and -R ZFN, the constructed ZFP genes was transferred into pET28a vector (Novagen) using *Nco*I and *Hind*III sites.

DNA Cleavage Activity Assay CCR5-L-ZFN and CCR5-R-ZFN cloning vectors were used for *in-vitro* translation reactions by TNT® T7 Quick-Coupled Transcription/Translation System (Promega). In this assay, reaction volumes were 20 μL containing 2 μL of the 10X reaction buffer (50 mM potassium acetate, 20 mM Tris-acetate, pH 7.9, 10 mM magnesium acetate, 1 mM DTT), 500 ng of the *Sca*I-linearized target DNA, and 2 μL of each expressed either CCR5-L or -L- ZFN. The reaction was incubated for 12 h at 22 °C. After the DNA cleavage, 10 μg of RNase A (GENE ALL) was added into the reactions and incubated for 30 min at 22 °C to remove RNAs. For removal of proteins in the reaction, 10 μg of Protease K (New England Biolab) was added to the samples and incubated for 2 h at 55 °C. The reaction was subjected to agarose gel electrophoresis and stained with ethidium bromide. And DNAs were visualized under UV.

Supporting Information. Supplementary data associated with this article can be found in the online version.

References

1. Bateman, A.; Coin, L.; Durbin, R.; Finn, R. D.; Hollich, V.; Griffiths-Jones, S.; Khanna, A.; Marshall, M.; Moxon, S.; Sonnhammer, E. L.; Studholme, D. J.; Yeats, C.; Eddy, S. R. *Nucleic Acids Res.* **2002**, *30*, 276-280.
2. Klug, A. *Annu. Rev. Biochem.* **2010**, *79*, 213-231.
3. Berg J. M. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 99-102.
4. Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635-637.
5. Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809-817.
6. Nakaseko, Y.; Neuhaus, D.; Klug, A.; Rhodes, D. *J. Mol. Biol.* **1992**, *228*, 619-636.
7. Neuhaus, D.; Nakaseko, Y.; Schwabe, J. W.; Klug, A. *J. Mol. Biol.* **1992**, *228*, 637-651.
8. Fairall, L.; Schwabe, J. W.; Chapman, L.; Finch, J. T.; Rhodes, D. *Nature* **1993**, *366*, 483-487.

9. Desjarlais, J. R.; Berg, J. M. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2256-2260.

10. Rebar, E. J.; Pabo, C. O. *Science* **1994**, *263*, 671-673.

11. Jamieson, A. C.; Kim, S. H.; Wells, J. A. *Biochemistry* **1994**, *33*, 5689-5695.

12. Choo, Y.; Klug, A. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11168-11172.

13. Wu, H.; Yang, W. P.; Barbas, C. F. 3rd. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 344-348.

14. Mandell, J. G.; Barbas, C. F. 3rd. *Nucleic Acids Res.* **2006**, *34*, W516-W523.

15. Choo, Y.; Sánchez-García, I.; Klug, A. *Nature* **1994**, *372*, 642-645.

16. Blancafort, P.; Chen, E. I.; Gonzalez, B.; Bergquist, S.; Zijlstra, A.; Guthy, D.; Brachat, A.; Brakenhoff, R. H.; Quigley, J. P.; Erdmann, D.; Barbas, C. F. 3rd. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11716-11721.

17. Lee, D. K.; Seol, W.; Kim, J. S. *Curr. Top. Med. Chem.* **2003**, *3*, 645-657.

18. Kim, Y.-G.; Cha, J.; Chandrasegaran, S. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1156-1160.

19. Tan, W.; Zhu, K.; Segal, D. J.; Barbas, C. F. 3rd.; Chow, S. A. *J. Virol.* **2004**, *78*, 1301-1313.

20. Gordley, R. M.; Smith, J. D.; Gräslund, T.; Barbas, C. F. 3rd. *J. Mol. Biol.* **2007**, *367*, 802-813.

21. Carroll, D.; Morton, J. J.; Beumer, K. J.; Segal, D. J. *Nat. Protoc.* **2006**, *1*, 1329-1341.

22. Wright, D. A.; Thibodeau-Beganny, S.; Sander, J. D.; Winfrey, R. J.; Hirsh, A. S.; Eichtinger, M.; Fu, F.; Porteus, M. H.; Dobbs, D.; Voytas, D. F.; Joung, J. K. *Nat. Protoc.* **2006**, *1*, 1637-1652.

23. Gonzalez, B.; Schwimmer, L. J.; Fuller, R. P.; Ye, Y.; Asawapornmongkol, L.; Barbas, C. F. 3rd. *Nat. Protoc.* **2010**, *5*, 791-810.

24. Fujii, W.; Kano, K.; Sugiura, K.; Naito, K. *PLoS One* **2013**, *8*, e59801.

25. Lombardo, A.; Genovese, P.; Beausejour, C. M.; Colleoni, S.; Lee, Y. L.; Kim, K. A.; Ando, D.; Urnov, F. D.; Galli, C.; Gregory, P. D.; Holmes, M. C.; Naldini, L. *Nat. Biotechnol.* **2007**, *25*, 1298-1306.

26. Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A. 3rd.; Smith, H. *Nat. Methods* **2009**, *6*, 343-345.

27. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1992**, *48*, 2223-2311.