DOI: 10.1002/cbic.200900040

Artificial Restriction DNA Cutters as New Tools for Gene Manipulation

Hitoshi Katada and Makoto Komiyama*[a]

Two types of artificial tools that cut double-stranded DNA through hydrolysis of target phosphodiester linkages have been recently developed. One is the chemistry-based artificial restriction DNA cutter (ARCUT) that is composed of a Ce^{IV}–EDTA complex, which catalyses DNA hydrolysis, and a pair of pseudo-complementary peptide nucleic acid fragments for sequence recognition. Another type of DNA cutter, zinc finger nuclease (ZFN), is composed of the nuclease domain of naturally occurring FokI restriction endonuclease and a designed zinc finger DNA-binding domain. For both of these artificial tools, the scission site and specificity can be freely chosen according to our needs, so that even huge genomic DNA sequences can be selectively cut at the target site. In this article, the chemical structures, preparation, properties, and typical applications of these two man-made tools are described.

1. Importance of Artificial Restriction DNA Cutters

Construction of cloning vectors is one of the most fundamental processes for cell transformation in biotechnology and molecular biology. Plasmid vectors are usually manipulated by: 1) scission of the plasmid at predetermined sites by restriction enzymes, and 2) connection of the scission product with a foreign DNA fragment by using a DNA ligase. A considerable number of restriction enzymes have been already discovered, and most of them are commercially available. However, the versatility of their recognition sequences (mostly palindromes) is still limited so that we often encounter difficulty in finding an appropriate restriction enzyme for our aimed DNA manipulation. Accordingly, new tools to cut DNA at any site of choice, if available, should facilitate complicated gene manipulation. Furthermore, most naturally occurring restriction enzymes recognize only 4–6 bp long DNA sequences, and thus their digestion occurs at too many sites when the target molecules are large in size. With an adenovirus vector, which is composed of ~35 kbp, for example, the scission site of a 6 bp-recognizing restriction enzyme should appear, on average, at around 5–10 sites ($35\,000/4^6 = 8.5$). The situation is more critical when still larger genomic DNA molecules are digested. To cut large DNA molecules site selectively and to allow their precise manipulation, new DNA-cutting tools that recognize a predetermined longer sequence are required. If a 16 bp sequence is strictly recognized, for example, even the genome of human beings (composed of 3×10^8 bp) can be cut, at least in theory, at only one-site ($4^{16} > 3\times10^9$). Therefore, in these three decades, many laboratories have attempted the preparation of artificial DNA-cutting tools that recognize longer sequences.

In 1987, Dervan et al. attached a Fe^{II}–EDTA complex as a DNA-cleaving molecule to a triplex-forming oligonucleotide and successfully cleaved double-stranded DNA at the corresponding triplex site.[1] The conjugates of a 1,10-phenanthroline–Cu^{II} complex with triplex-forming oligonucleotides also showed sequence-specific scission.[2,3] In these studies, radical species, which were formed during catalysis by the metal complex, cleaved the neighboring ribose residues at the target site; this resulted in site-selective scission of the DNA. These elegant pioneering works were further extended to various types of site-selective DNA cleavers. Freedom of choice on cleavage site has been ensured by using poly(pyrrole-imidazole),[4–7] zinc finger proteins,[8–10] and other sequence-recognizing molecules[11] in place of the triplex-forming oligonucleotides. Many other metal complexes for DNA cleavage have also been reported.[12–21] These artificial cutters are so eminent in scission efficiency that they have been widely employed as new tools in nucleic acid chemistry, molecular biology, and many other relevant fields.

2. Artificial Tools for Site-Selective Hydrolysis of DNA at a Desired Site

Subsequently, two types of artificial tools that cut double-stranded DNA through hydrolysis of target phosphodiester linkages as naturally occurring nucleases do, have been developed and used for gene manipulation (Figure 1). They have been further applied to various biochemical and/or biological purposes. These hydrolytic DNA cutters are the main subject of this review. In this section, the molecular structure of these hydrolytic DNA cutters will be outlined (recent developments in engineering of naturally occurring, rare-cutting endonucleases will be briefly described in Section 5).

[a] *H. Katada, Prof. M. Komiyama*
Research Center for Advanced Science and Technology
The University of Tokyo, 4-6-1 Komaba, Meguro
Tokyo 153-8904 (Japan)
Fax: (+81)3-5452-5209
E-mail: komiyama@mkomi.rcast.u-tokyo.ac.jp

2.1 Chemistry-based artificial restriction DNA cutters (ARCUT)

As shown in Figure 1 A, this chemistry-based DNA cutter (ARCUT) is composed of: 1) a Ce^{IV}–EDTA complex (a catalyst for DNA hydrolysis), and 2) a pair of pseudo-complementary peptide nucleic acids (pcPNA; black lines) for sequence recognition.[22] In pcPNA, pseudo-complementary DNA bases (2,6-diaminopurine (D) and 2-thiouracil (U)) are bound to a poly(*N*-aminoethyl-L-glycine) backbone. Duplex formation of these two pcPNA strands is suppressed by steric repulsion between D and U so that two appropriately designed strands efficiently form a double-duplex invasion complex (note that they are complementary to each other).[23,24] The binding sites of two pcPNA strands in double-duplex invasion complexes in ARCUT are laterally shifted with respect to one another by several nucleobases, and thus single-stranded portions (the gray parts) are formed at predetermined sites in the DNA (the second structure from the top). These portions are preferentially hydrolyzed by the Ce^{IV}–EDTA complex, because this complex hydrolyzes only DNA in the single-stranded state, and double-stranded DNA is hardly hydrolyzed.[25] The scission site and site specificity of ARCUT can be tuned simply by changing the sequences and lengths of the pcPNAs. In fact, even huge DNA molecules (e.g., 4.6 Mbp genome of *E. coli*) can be selectively cut at the desired site, although it is absolutely impossible with conventional restriction enzymes.[26] Importantly, all of the DNA scission proceeds by hydrolysis of target phosphodiester linkages, exactly as with scissions by naturally occurring restriction enzymes.[27] Thus, fragments obtained by ARCUT scission can be combined with various DNA fragments by using DNA ligase. Thus, ARCUT can be applied to the construction of vectors of various sizes. Importantly, the scission site of ARCUT is straightforwardly predictable and determined, because it is governed only by Watson–Crick base pairings.

Makoto Komiyama graduated from the University of Tokyo in 1970, and received his Ph.D. from the same University in 1975. After spending four years at the Northwestern University (Evanston, IL, USA) as a postdoctoral fellow, he became an assistant professor at the University of Tokyo, and then an associate professor at the University of Tsukuba. Since 1991, he has been a professor at the University of Tokyo. His main research area is bioorganic and bioinorganic chemistry. He has received many honors and awards including from the Chemical Society of Japan, Japan IBM Science, the Rare Earth Society of Japan, Inoue Prize for Science, the Society of Polymer Science, Japan, and the Cyclodextrin Society of Japan. Prof. Komiyama is a member of the Editorial Advisory Board of *ChemBioChem*.

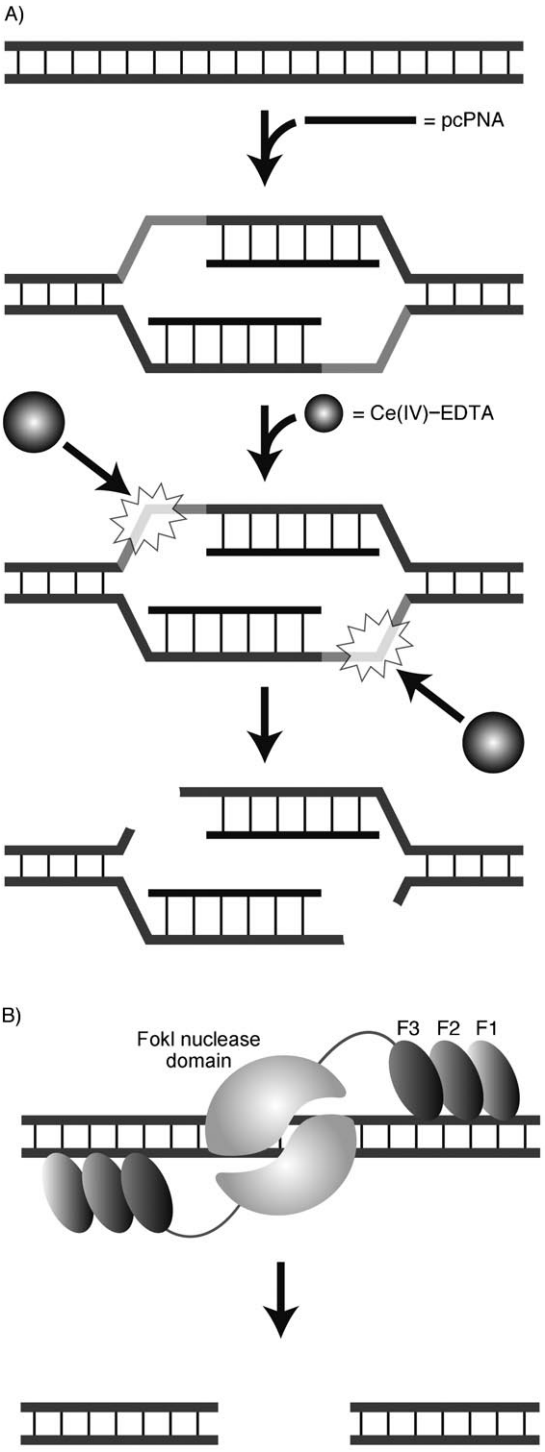

Figure 1. Artificial cutters for site-selective hydrolytic scission of double-stranded DNA. A) Chemistry-based artificial restriction DNA cutter (ARCUT) and B) zinc finger nuclease (ZFN); F1, F2, F3: zinc finger DNA-binding domains.

2.2 Zinc finger nuclease (ZFN)

Zinc finger nucleases (ZFNs) were first reported by Chandrasegaran and co-workers,[8] and are composed of: 1) the nuclease domain of naturally occurring FokI restriction endonuclease, and 2) zinc finger DNA-binding domains (F1, F2, F3; Figure 1 B).

The zinc finger domain, which is typically three or four tandem-arrayed Cys_2His_2 zinc finger proteins, binds to the target site in the substrate DNA, and puts the FokI nuclease domain in place. It is noteworthy that the nuclease domain of FokI becomes catalytically active only when two of them are placed in close proximity to form a dimer.[10,28] DNA scission by the resultant dimeric FokI nuclease occurs without any site specificity. During scission of double-stranded DNA, two conjugates of the nuclease domain and the zinc finger domain bind simultaneously to the DNA, as depicted in Figure 1B. The nuclease domain dimer is formed in situ and hydrolyzes the region between the binding sites of the two zinc finger domains.[10] The zinc finger portions in these two conjugates need not be identical. Rather, two different conjugates involving a zinc finger domain, which binds to its corresponding unique DNA sequence, can also be combined and thus even nonpalindromic sequences can be targeted for site-selective scission.

Each of the zinc finger proteins, which is composed of about 30 amino acid residues and is folded into a $\beta\beta\alpha$ structure in the presence of zinc ion, primarily recognizes 3 bp in DNA through hydrogen bonds.[29] There, the amino acid residues at positions -1 to $+6$—relative to the start of the α helix—make the predominant contributions to DNA sequence recognition. Accordingly, the zinc finger domain (in homo- or heterodimer conjugate(s)) involves about 200–300 amino acid residues in total and recognizes 18 or 24 bp, which is long enough to target only one sequence in human genomic DNA. At present, de novo design of a zinc finger protein that binds to a target binding site is rather difficult (rational design has only been successful to some extent).[30] Instead, a required protein is usually selected from libraries with either the phage display method[31–34] or a bacterial cell based two-hybrid system.[35,36] The nuclease domain that is derived from FokI is composed of 196 amino acid residues.

In the following sections, DNA manipulation by using either ARCUT or ZFN is described to show their usefulness and limitations.

3. DNA Manipulation by Artificial Restriction DNA Cutter (ARCUT)

3.1 Scission of vectors at desired sites by using ARCUT

The sequences and lengths of the pcPNAs that are used for ARCUT are almost freely chosen according to our needs. In principle, any kind of vector, irrespective of its size and the scission site, can be selectively cut by appropriately designed ARCUT. To cut pBR322 plasmid at 1830 bp site, for example, a pair of pcPNAs, $pcPNA^1$ and $pcPNA^2$, were designed (Figure 2A). These additives are complementary to C1826–A1840 in the upper strand of the plasmid DNA, and A1821–G1835 in the lower strand, respectively. Upon incubating these pcPNAs with the pBR322 plasmid, a double-duplex invasion complex is formed in which both T1821–T1825 in the upper strand and G1836–T1840 in the lower strand are kept single stranded (Figure 2A, underlined sequences). These single-stranded portions

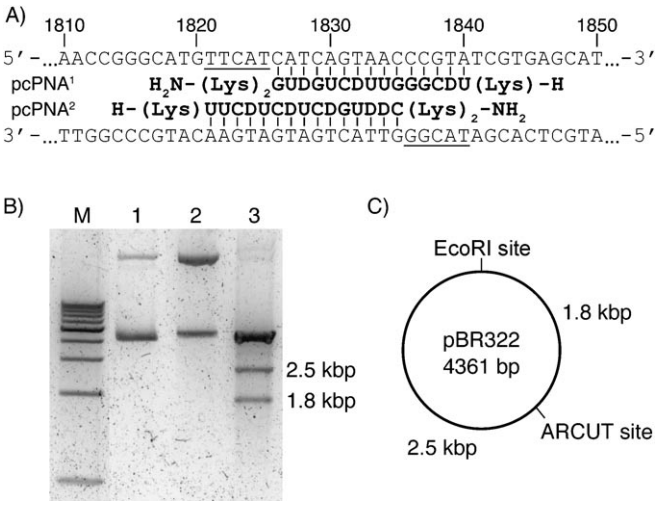

Figure 2. A) Sequences of the target site and pcPNA additives for site-selective hydrolysis of pBR322 plasmid. The underlined single-stranded portions are preferentially hydrolyzed by the Ce^{IV}–EDTA complex. In the pcPNA strands, 2,6-diaminopurine (D) and 2-thiouracil (U) were used in place of conventional DNA bases A and T, respectively.[23,24]. B) Agarose gel electrophoresis patterns; lane M: 1 kbp DNA ladder; lane 1: no treatment; lane 2: after ARCUT scission; lane 3: EcoRI digest of lane 2. C) Vector map of pBR322. The lengths of the fragments obtained by consecutive scissions by ARCUT and EcoRI are also shown.

were selectively hydrolyzed by the Ce^{IV}–EDTA complex at the target site to produce the linear, form III DNA (4361 bp length; Figure 2B, lower band in lane 2). Although there exist several scission products in the reaction mixtures, only the desired product is selectively ligated with foreign DNA by using the appropriate joint oligonucleotide and used for the following DNA manipulation (section 3.3B). To cut the vector at two sites and provide two truncated portions, when necessary, ARCUT can be combined with appropriate naturally occurring restriction enzymes (or with another set of ARCUT). In lane 3, for example, the primary scission product of pBR322 by the ARCUT (lane 2) was further treated with EcoRI, the recognition site of which is located at the 4359 bp site of this plasmid. As expected, two fragments of 1.8 and 2.5 kbp were formed, and both fragments can be used to prepare desired vectors.

One of the biggest advantages of ARCUT over naturally occurring restriction enzymes is its extraordinarily high aptitude for recognizing long sequences. Thus, even large vectors that are too big to be manipulated by conventional restriction enzymes can be satisfactorily cut at one target site. As demonstrated in Figure 3, adenovirus vector pAd/PL-DEST-BFP (ca. 35 kbp) is specifically cut at site 1490 bp by ARCUT (the sequences of target site and pcPNA additives are shown in Figure 3A). As a result of the ARCUT scission, form I DNA (supercoiled DNA in lane 1) was converted to form III (linear ~35 kbp DNA; Figure 3B, lane 3). Upon post-treatment of the ARCUT product with PmeI, a 10.5 kbp fragment was formed as expected (lane 4). The site specificity of the ARCUT scission was further confirmed (another 24.4 kbp fragment could not be separated from form III DNA under the conditions used).

Note that the direct scission of this adenovirus vector by EcoRI, which recognizes 6 bp, provides five fragments (Fig-

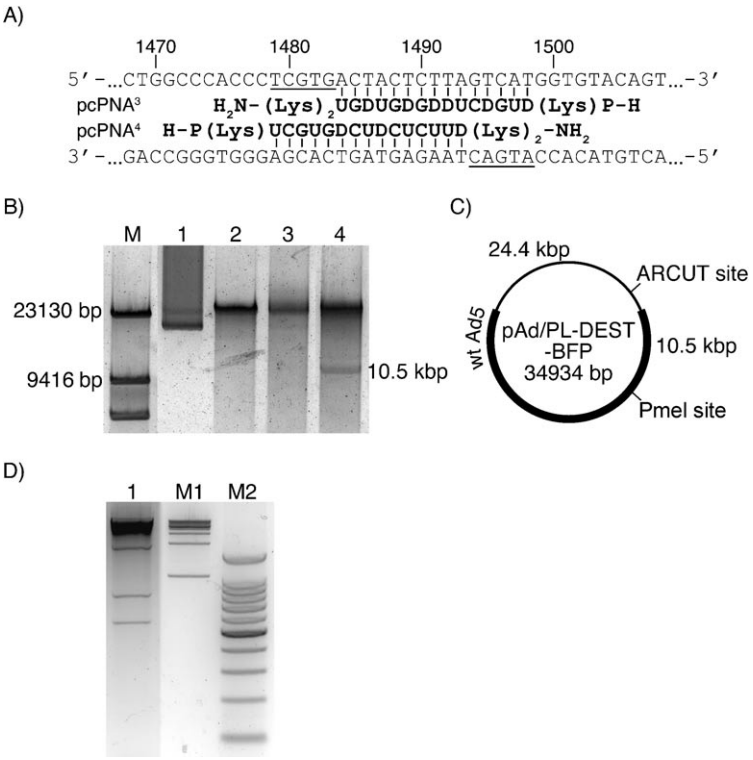

Figure 3. Site-selective hydrolysis of adenovirus vector, pAd/PL-DEST-BFP (~35 kbp). A) Sequences of target site and pcPNA additives; phosphoserine (P) was attached to the N terminus of the pcPNAs to promote scission activity. B) Agarose gel electrophoresis patterns; lane 1: no treatment; lane 2: PmeI digest; lane 3: ARCUT scission product; lane 4: PmeI digest of the product in lane 3; lane M: λ/HindIII marker. C) Vector map of pAd/PL-DEST-BFP. The lengths of the fragments obtained by consecutive scissions by ARCUT and PmeI (in lane 4) are also shown. D) Agarose gel electrophoresis patterns obtained after direct scission of pAd/PL-DEST-BFP by EcoRI; lane 1: EcoRI digests of pAd/PL-DEST-BFP; lane M1: 1 kbp DNA ladder; lane M2: 100 bp DNA ladder.

ure 3D). The necessity of ARCUT is, therefore, evident. In a typical ARCUT, two pcPNA strands involving 15 nucleic acid bases are employed. According to a systematic study, this ARCUT strictly recognizes 16–20 bp at the target scission site.[37] This specificity is sufficiently high for site-selective scission of human genomic DNA.

3.2 Preparation of inserts for ARCUT-based DNA manipulation

Because the site-selective DNA scission by ARCUT proceeds by simple hydrolysis of target phosphodiester linkages, the resultant fragments are directly subject to catalysis by DNA ligase and other enzymes. Hence, all of the inserts used for conventional DNA manipulation (e.g., PCR products and restriction enzyme digests) can be directly employed as the inserts for ARCUT-based manipulation.

It is noteworthy that ARCUT can be used to prepare inserts that are otherwise hardly obtainable. Two sets of ARCUT are used simultaneously (or stepwise) on genomic DNA, and this huge DNA is selectively hydrolyzed at two predetermined sites. Alternatively, one set of ARCUT can be combined with a naturally occurring restriction enzyme. By these two-site scissions, a gene of interest can be excised and used as an insert for vector preparation. The sequence-specificity of ARCUT is high enough to cut one target site selectively even in huge genomic DNA molecules. Importantly, epigenetic information, such as DNA methylation, is kept intact during the excision, which is in marked contrast with its complete loss during PCR amplification of the target. Therefore, this technique promises to be extremely useful for various applications, such as gene analysis and diagnosis. Figure 4 shows a typical example. A pair of pcPNA strands for each of these two sets of ARCUT bind to the corresponding target sequences in the DNA, and Ce^{IV}–EDTA promotes hydrolysis at this site. The desired fragment (1010 bp) can thus be successfully excised (Figure 4B, lane 1).

3.3 Ligation of the vector and insert obtained by using ARCUT

The vectors and the inserts prepared by using ARCUT have unique termini structures that involve overhangs of 15–20 nucleotides long. They are the results of the Ce^{IV}–EDTA-catalyzed hydrolytic scission of the single-stranded portions, which are formed in the double-duplex invasion complex (Figure 1A). These ARCUT products cannot be directly ligated with the inserts obtained by PCR and/or restriction enzyme digestions (they have either conventional cohesive ends or blunt ends). To solve this problem, two methods have been developed. One method is to convert the unique termini of ARCUT products (and cohesive end of the insert, if any) into blunt ends by using single-stranded DNA specific endonucleases (Figure 5A). Another is to add a joint oligonucleotide to ligation mixtures and fill the gap between these two kinds of termini (Figure 5B). As a result, the "complementary" structure is temporarily formed and successfully ligated.

A. Formation of blunt ends by single-stranded DNA-specific endonucleases. ARCUT products are treated with single-stranded DNA-specific endonucleases, and their unique protruding termini are converted to blunt ends. An insert fragment that has blunt ends in both termini is independently prepared. The ligation of these two fragments by DNA ligase is straightforward. Because of the feature of ARCUT, the insertion site in the vector can be freely chosen. This is an important advantage, although the direction of the insert in the vector and the trimming length by the single-stranded specific nuclease cannot be precisely controlled.

As illustrated in Figure 6, enhanced green fluorescent protein (EGFP) was inserted as a cassette into pBR322. The linearized pBR322, which was prepared by ARCUT scission at 1830 bp site (section 3.1), was treated with Mung bean nuclease and Klenow fragment to convert the end structures to blunt ones. An insert containing the EGFP gene was independently prepared by PCR and its 5′ termini were phosphorylated with T4 polynucleotide kinase. The linearized plasmid and the

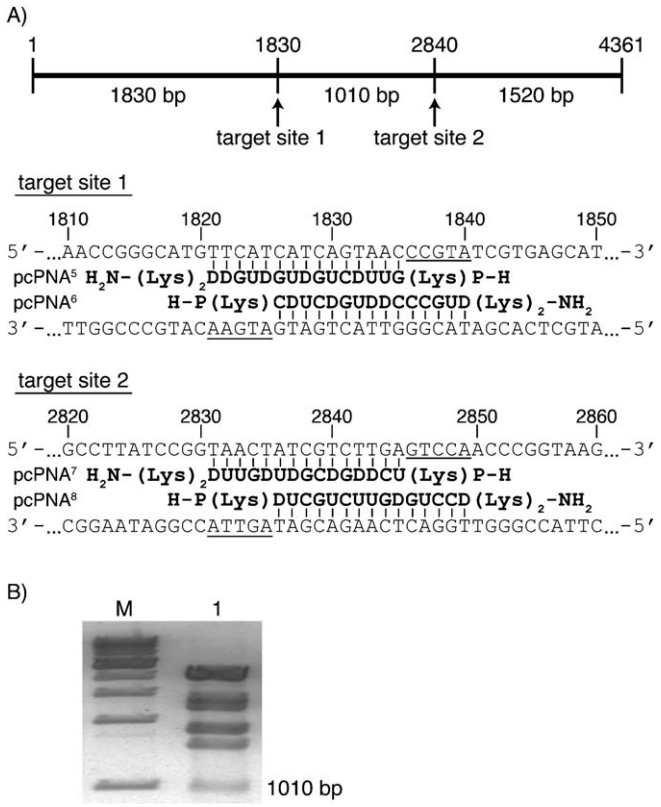

Figure 4. Excision of a desired fragment by using two sets of ARCUT. A) Sequences of target scission site 1, target scission site 2, and pcPNA additives used for double digestion of linearized pBR322. The sequences hydrolyzed by the Ce^{IV}–EDTA complex are underlined. B) Agarose gel electrophoresis patterns obtained after digestion; lane M: 1 kbp DNA ladder; lane 2: after two-site scission by two sets of ARCUT. Note that the desired 1010 bp fragment is observed in lane 1, together with the products formed by one-site scission.

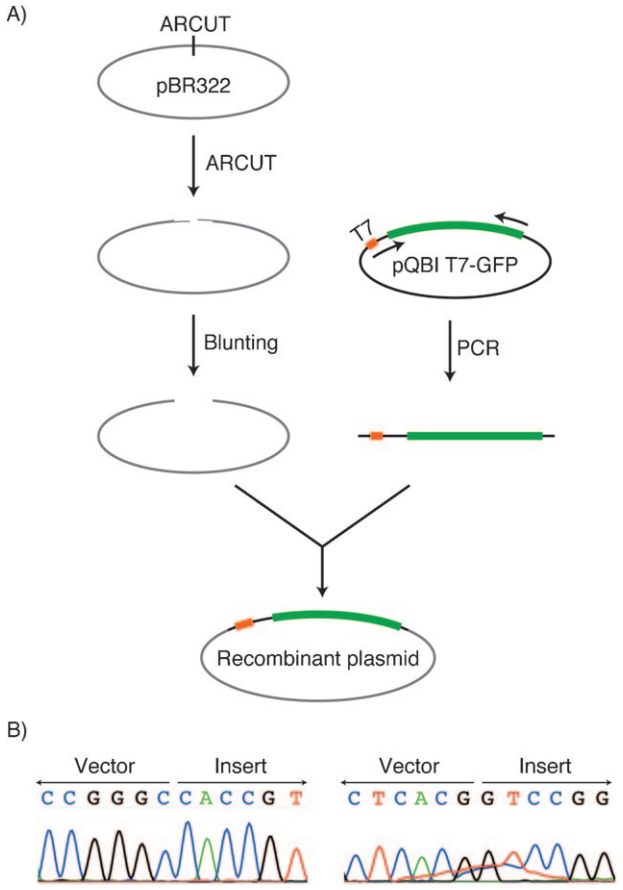

Figure 6. Gene cassette insertion into a vector prepared by ARCUT through blunt-end ligation. A) Schematic outline of the procedure (see also Figure 5A). The terminal phosphates of the ARCUT product were removed with alkaline phosphatase after blunting to prevent self-ligation of the vector in the subsequent ligation step. B) Sequence analysis of one of the recombinant vectors in which the EGFP gene was inserted in the forward (left) and reverse (right) conjunction sites.

insert were incubated with DNA ligase under conventional conditions. By using this simple procedure, the EGFP gene cassette was successfully inserted into the target site of the vector in high efficiency, and was confirmed by sequencing experiments (Figure 6B).

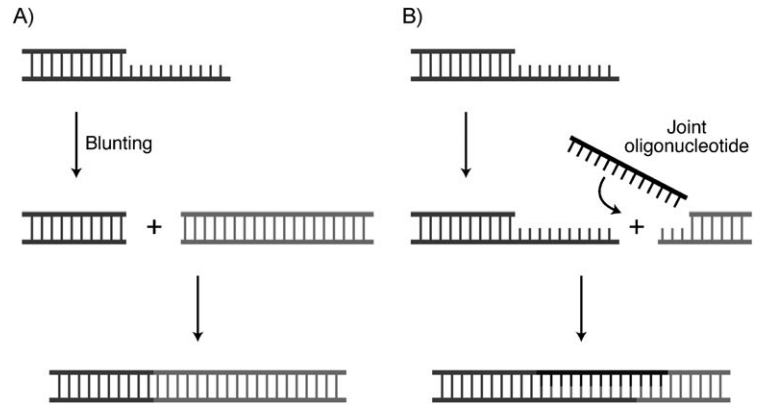

Figure 5. Two strategies for the connection of a vector prepared by ARCUT to an insert prepared by conventional methods. The ARCUT product and foreign fragment are in dark- and light-gray, respectively. A) The ARCUT-treated vector and the insert with cohesive ends, if any, are first treated with single-stranded DNA-specific endonuclease, and then the resultant blunt ends are ligated. B) A "joint oligonucleotide" (black) is added to fill the gap between the vector and the insert, and the resultant temporarily "complementary" end structure is ligated.

B. Addition of joint oligonucleotide to form temporarily complementary end structures. When one should like to incorporate an insert into a vector precisely in a predetermined position (without even 1 bp misplacement), the vector should be cut by both ARCUT and an appropriate restriction enzyme to provide two different termini to the product (the vector can be alternatively cut by two sets of ARCUT). The insert can be prepared by using restriction enzymes or ARCUT. In the schematic example shown in Figure 7A, the ligation in the left-hand side of the insert is straightforward because both ends are prepared by using the same restriction enzyme. To connect the ARCUT formed end of the vector with the restriction enzyme formed end of the insert (the right-hand side), however, the strategy shown in Figure 5B must be used, and an appropriate joint oligonucleotide should be added (Figure 7A, black line). As depicted in Figure 7B, a part of this oligonucleotide is comple-

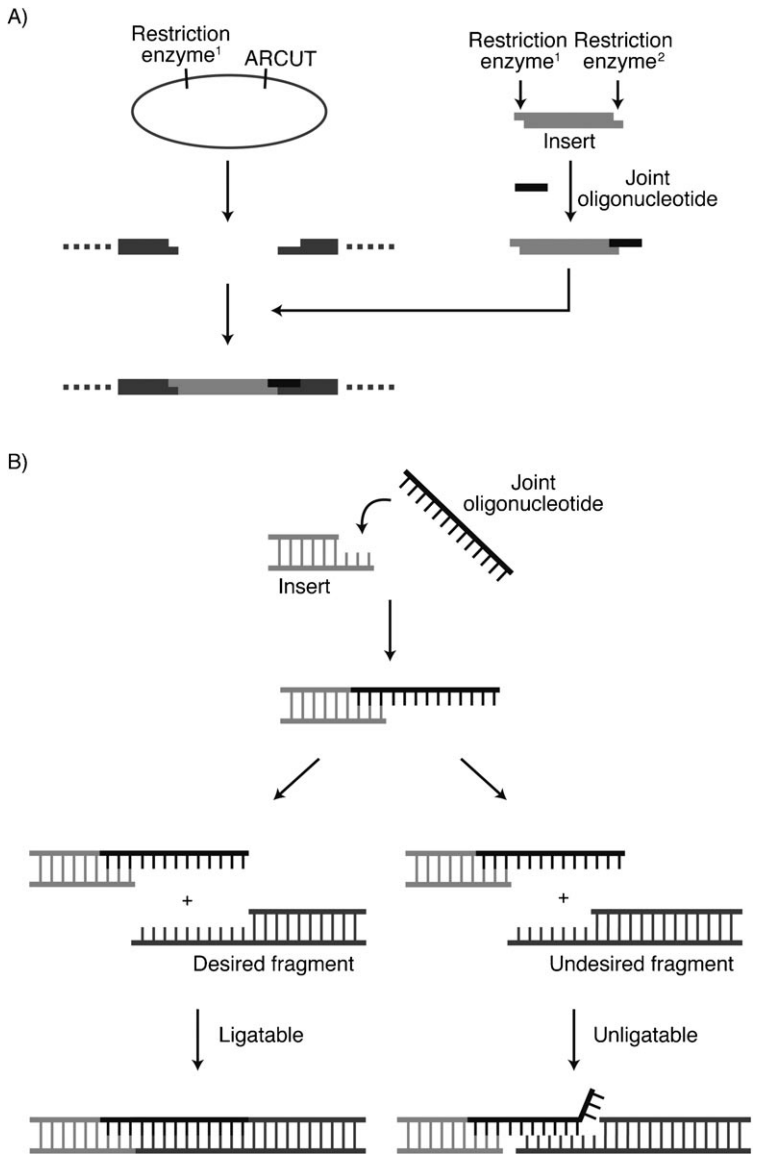

Figure 7. A) Schematic example of precise gene manipulation by using a joint oligonucleotide (see also Figure 5B). B) In the ligation step that uses a joint oligonucleotide, only the desired fragment is selectively recognized, and undesired fragments are left unchanged in the solution.

mentary with the protruding portion of the ARCUT product (in dark gray), and the rest is complementary with that of the restriction enzyme end (in light gray). Thus, in its presence, the ARCUT terminus is temporarily made "complementary" with the terminus of the insert. Under conventional ligation conditions, only the desired vector fragment is selectively picked up and ligated with the insert (Figure 7B, the left-hand pathway). All other ARCUT products that exist in the solution are left unchanged (the right-hand pathway). Note that scission by ARCUT occurs at several phosphodiester linkages in the single-stranded portions. The reading frame of the resultant recombinant gene is also successfully adjusted to allow the expression of the recombinant protein. This method is suitable for precise DNA manipulation, such as for introduction of point mutations, preparation of fusion proteins, and others.

This method was used to prepare mutants of blue fluorescent protein (BFP) through a selection procedure (Figure 8). Ser65 and His66 in the chromophore of BFP were changed to other amino acids. First, the BFP gene in plasmid pQE60-BFP was cut in the middle by both ARCUT and restriction enzyme SpeI, and a 60 bp fragment was removed from the chromophore region. Then, an insert of the same length in which the 6 bp portion that contained amino acids 65 and 66 was randomized and incorporated into the vector. In this ligation step, the joint oligonucleotide, $Oligo^{joint}$ (Figure 8C), was added as the key component for precise ligation. The ligation product was transformed into *E. coli* strain DH5α and colonies that generated bright fluorescence were observed. As shown in Figure 8D, two DNA fragments were successfully connected in the correct reading frames. Three fluorescent proteins other than BFP were obtained, and emitted fluorescence at different wavelengths (Figure 8E).

3.4 PCR-free and restriction enzyme-free vector construction by using ARCUT

In the experiment shown in Figure 9A, both vector and insert were prepared by ARCUT. Because neither PCR nor naturally occurring restriction enzymes are necessary, we can obtain DNA vectors that are completely free of polymerases, restriction enzymes, and their contaminants. Plasmid pET-28b was linearized by ARCUT and treated with nuclease S1 to provide a vector with blunt ends (section 3.3). Separately, an insert containing the EGFP gene (720 bp) was removed from plasmid pQBI T7-GFP by using two sets of ARCUT, as outlined in section 3.2. This EGFP gene was also treated with nuclease S1 to form blunt ends. These vectors and inserts were connected by using DNA ligase, and transformed into *E. coli* strain JM109. The colonies containing the correct open reading frame were screened (Figure 9B) and transformed into *E. coli* strain BL21-Gold (DE3). The expressed EGFP successfully emitted strong green fluorescence (Figure 9C); this confirms that the DNA was kept intact during the ARCUT-based manipulation.

4. DNA Manipulation with Zinc Finger Nucleases (ZFNs)

In contrast to the completely chemistry-based ARCUT described in section 3, ZFNs are conjugates of two naturally occurring peptides (the nuclease domain of FokI enzyme and zinc finger proteins; Figure 1B). The linker portions between these two domains are also conventional peptides (e.g., $(Gly_4Ser)_3$). Accordingly, ZFNs are usually prepared in cells by using the corresponding expression vectors, which can be directly introduced into the cells. ZFNs formed in situ can be conveniently used for various applications in vivo.

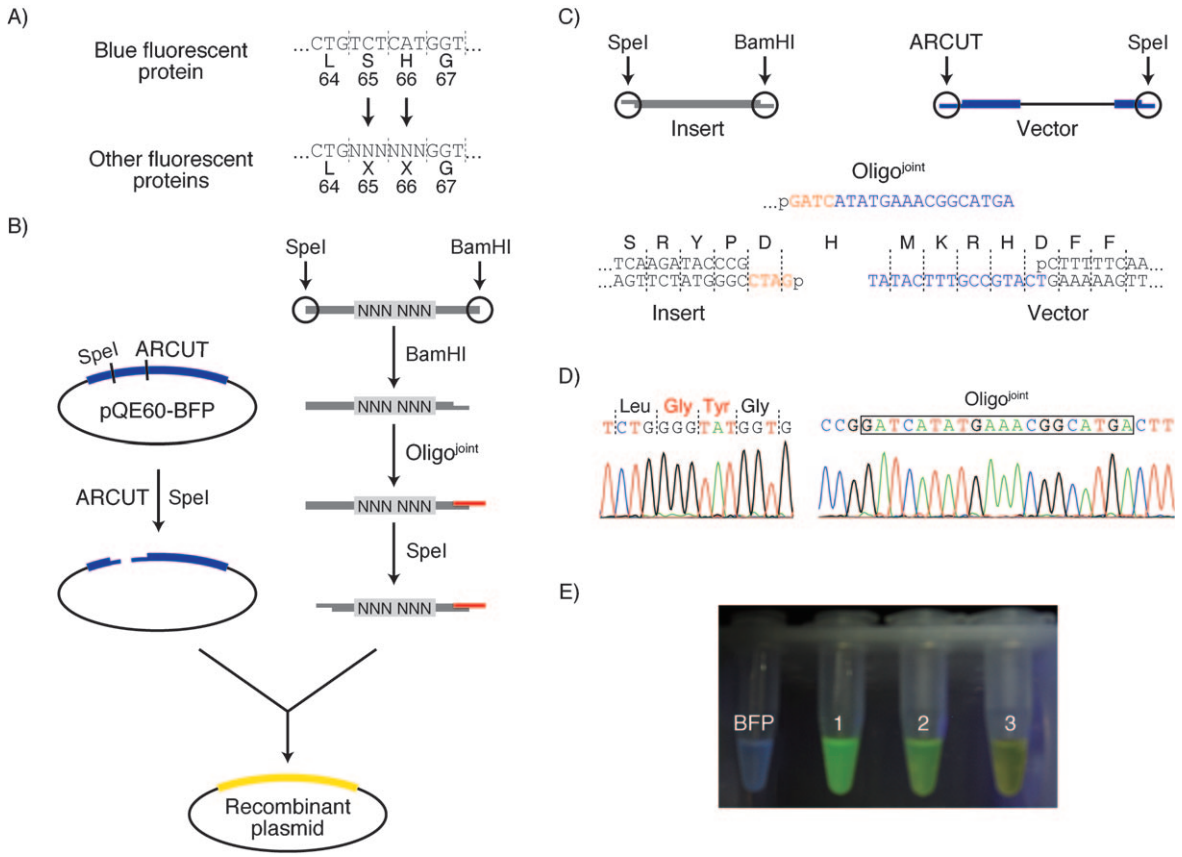

Figure 8. A) Selection of fluorescent proteins by randomizing the chromophore of BFP. The genetic sequence and amino acid residues near the chromophore are shown. B) Outline of gene recombination for BFP. C) Ligation of the vector with ARCUT termini to the insert with BamHI termini. D) Sequence analyses of cloned recombinant vector at the mutation site (left) and Oligo^{joint} conjunction (right). E) Fluorescence emission of proteins obtained in the selection. The 65th and 66th amino acid residues are as follows: 1) Cys65 and Tyr66, 2) Ala65 and Tyr66, and 3) Gly65 and Tyr66. The recombinant DNA, which was obtained as described in the text, was further cloned into pET19b vector and over-expressed in *E. coli* strain BL21-Gold (DE3).

4.1 Promotion of homologous recombination

One of the most important applications of ZFNs is to cut genomic DNA in cells and promote homologous recombination. Such genetic recombination is an evolutionarily and physiologically essential process, in which two similar (or identical) strands of DNA are exchanged by intracellular biological systems. At the same time, homologous recombination provides a very useful means to introduce desired DNA sequences into various organisms. For example, a target gene can be selectively corrected and/or modified to create recombinant DNA and genetically modified organisms. In mammalian cells, however, this recombination occurs with very low frequency; this hampers the practical applications of this unique bioprocess.[38]

It is well known that homologous recombination is remarkably enhanced when both strands of genomic DNA are cut near the target recombination site (double-strand break; DSB).[39–41] This DSB-induced enhancement of repair is an intrinsic property of living cells and causes the reparation of harmful DSB DNA damage, which can be formed in the genome by chemicals, X-ray radiation, or enzymes. In 2003, Porteus and Baltimore first used ZFNs to induce DSB at a target site in human cells, and in fact notably enhanced the efficiency of targeting gene correction.[42] As shown in Figure 10, a defective GFP gene that was stably integrated into the human genome was employed as a reporter gene. The recognition sequence of the ZFN used was within this defective GFP gene. When the expression vector of this ZFN was introduced into the cells together with a donor plasmid coding the correct sequence of GFP, the number of GFP-positive cells enormously increased. Apparently, a DSB was formed by the ZFN at the target site in the defective GFP gene, and the defect in this gene was corrected and the DSB was repaired with the use of the donor plasmid as template. This pioneering work was further extended by Porteus and Holmes et al. to gene correction for therapeutic purposes.[43] The target gene corrected was *IL2Rγ*, which is associated with X-linked severe combined immunodeficiency. Two ZFNs, each of which contained four zinc finger proteins, were assembled to recognize 24 bp at the target site. Exactly as designed, a high efficiency of targeted gene correction (up to 20%) was successfully accomplished. Gene modification by ZFNs in plants has also been reported.[44,45]

4.2 Knockout of a target gene

ZFNs have also been used to knockout a target gene. Although homologous recombination (section 4.1) is achieved in the presence of a donor DNA, it is unnecessary in targeting a

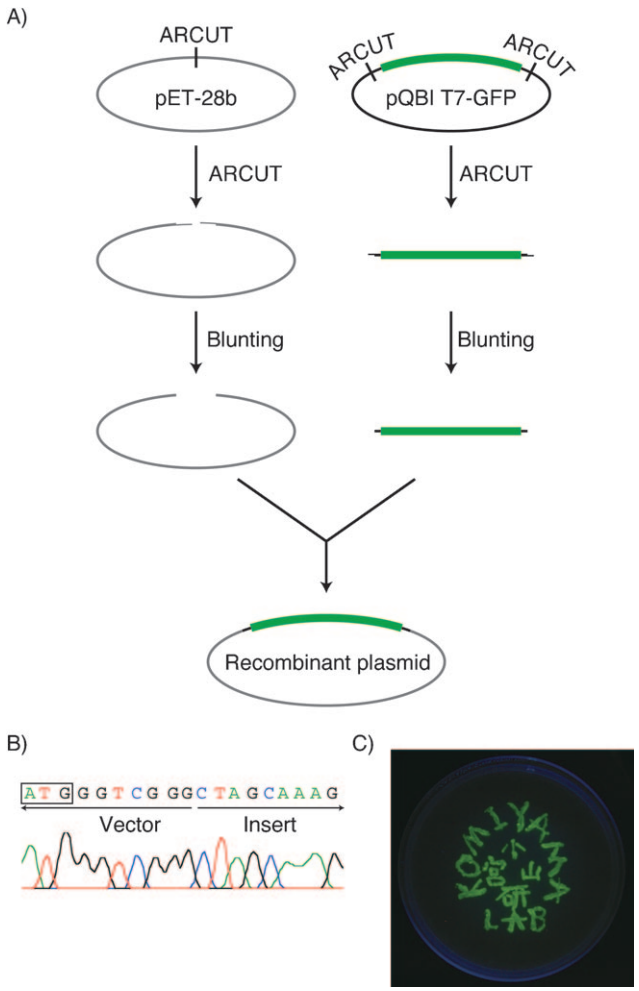

Figure 9. Use of ARCUT for gene manipulation without either PCR or restriction enzymes. A) The outline of procedure. B) Sequence analysis of recombinant plasmid. The initiation codon from the pET-28b vector is boxed. C) Fluorescence emission of EGFP expressed from a recombinant plasmid constructed by using this method.

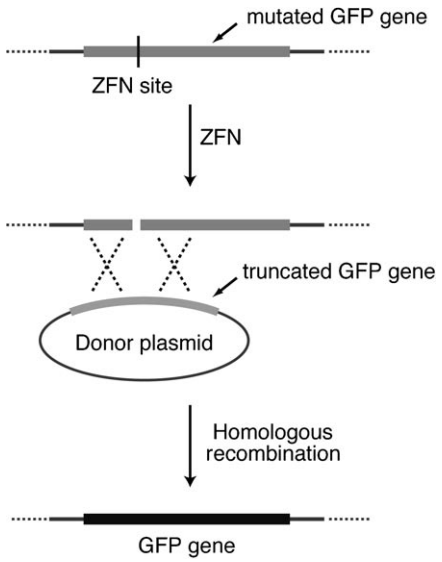

Figure 10. Outline of ZFN-induced gene correction in a human cell line. The mutated GFP gene that was stably integrated into the genome was cleaved site-specifically by ZFNs and corrected through the homologous recombination pathway by using the donor plasmid template, which possesses the truncated GFP gene.

knockout. When a DSB is formed in a gene in cells under these conditions, it is repaired through another pathway, namely nonhomologous end joining. This repair process proceeds without any template, which is in contrast to homologous recombination. As a result of this type of repair, insertions or deletions take place in the target gene and lead to its disruption. When the *CCR5* gene, which codes for the major coreceptor for HIV-1 entry, was targeted by using the corresponding ZFNs, mutations occurred in the gene with high frequencies, and *CCR5* was knocked out in approximately 80% of cases.[46]

4.3 Improvements of ZFN for further applications

Occasionally, ZFNs cut DNA at sequences other than the target site.[42,47–49] This off-target cleavage, which is crucial for practical applications, is mainly attributed to: 1) insufficient sequence recognition by the zinc finger DNA-binding domains, and/or 2) unspecific dimerization of the FokI nuclease domains. To solve the first point, various approaches to improve the fidelity of recognition of zinc finger proteins have been attempted. Typical methods employed for this purpose include construction of chimeric motifs,[50] the use of designed linker peptides,[51,52] and improvements of selection methods.[34,53]

In the second case the off-target cleavage is associated with the formation of undesired FokI dimers in the reaction mixture. Note that the two conjugates used for ZFNs (the conjugates of FokI domain and zinc finger protein domain) should be different from each other, unless the scission site has a palindromic sequence (Figure 11 A). The dimerized FokI in such a heterodimer, which is formed from these two different conjugates, is the active species for the required site-specific cleavage (Figure 11 A, top, left-hand side). In principle, however, homodimers of two identical conjugates can also be formed in situ, and if they are formed, they should bind the nontarget sites (palindromes); this results in off-target cleavage (Figure 11 A, second and third rows on the left-hand side). To solve this problem, point mutations were introduced at appropriate positions in the dimerization interface of the FokI nuclease domain (Figure 11 A, the right-hand side). For example, negatively charged amino acids (Glu) were introduced by site-selective mutagenesis at the interface of one of the two different conjugates, and positively charged ones (Lys) were introduced at the interface of another conjugate. Because of electrostatic attraction between the charges of different signs, the dimeric FokI in the heterodimer (for the target scission) was efficiently formed (Figure 11 A, top, right-hand side). On the other hand, the dimeric FokI in the homodimers, which is responsible for off-target scissions, was destabilized by electrostatic repulsion between the charges (Figure 11 A, second and third rows on the right-hand side). The validity of these arguments was investigated in terms of the cytotoxicity of the parent and modified ZFNs (Figure 11 B). The pink signals correspond to DNA that is

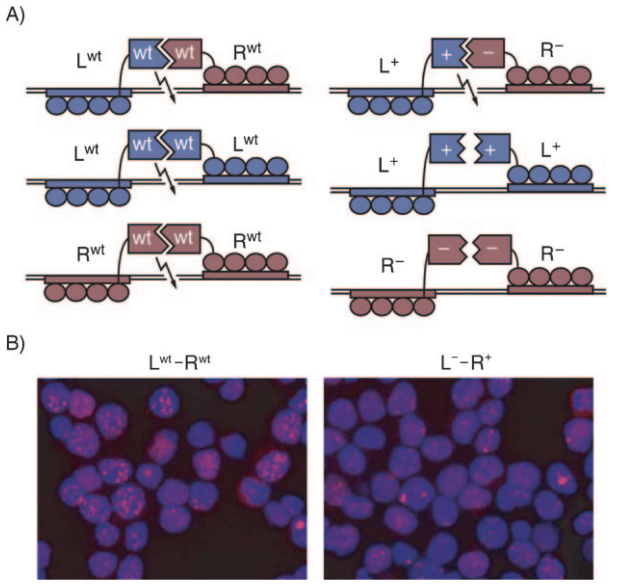

Figure 11. Suppression of cytotoxicity of ZFN by introducing point mutations into the FokI dimerization interface. A) Left: sketch of off-target scission by wild-type ZFN. Undesired homodimers (L^{wt}–L^{wt} and R^{wt}–R^{wt}) are formed along with heterodimer (L^{wt}–R^{wt}); this results in off-target scission. Right: shows the point mutations at the FokI dimerization interface. The formation of undesired homodimers (L^{+}–L^{+} and R^{-}–R^{-}) were suppressed by the electrostatic repulsion between the charges introduced. B) Representative images of cells transfected with the wild-type (left) and mutant ZFN (right). Note that cells transfected with the mutant ZFN show a much reduced pink staining with the antibody against 53BP1 protein, which localizes at the site of DNA damage. This figure was reprinted from ref. [47] with permission. Copyright: Nature Publishing Group, 2007.

noticeably damaged (DSBs are formed by the ZFNs and are not repaired). Exactly as designed, the modified ZFN induced much smaller undesired DNA damage than the parent ZFN (compare the two pictures in Figure 11 B). Apparently, the off-target scissions by the ZFN and the resultant cytotoxicity were greatly reduced by introducing charges to the dimerization interface. Promising future applications of these modified ZFNs have been indicated.[47,48]

5. Homing Endonucleases that Recognize Long DNA Sequences

In nature, there exists a family of homing endonucleases that recognize substantially longer DNA sequences (usually 15–40 base pairs) than conventional restriction enzymes.[54] In fact, these enzymes can induce DSBs at target sites in genomic DNA and stimulate homologous recombination in cells.[40] However, sequences recognized by naturally occurring homing endonucleases are limited in number. To solve this problem, many attempts have been made to alter the recognition sequence by protein engineering and to create variants with tailor-made specificities. These approaches provide promising possibilities for applications of these enzymes. However, their detailed description is beyond the scope of this review, and should be referred to other reviews[55,56] and literature.[57–60]

6. Summary and Outlook of Gene Manipulation by ARCUT and ZFN

With the use of ARCUT or ZFN, target phosphodiester linkages in substrate DNA can be selectively hydrolyzed irrespective of the size of DNA and the position of scission. Even huge genomic DNA molecules can be satisfactorily manipulated. To date, ARCUT has been mainly used as a tool for in vitro manipulation. Even when a small DNA, such as plasmid DNA, is manipulated, freedom of scission site is a big advantage. PCR-free and restriction-enzyme-free manipulation is also possible. One of the most important features of ARCUT is that both the scission site and specificity can be directly predicted by Watson–Crick base-pairings between the DNA and pcPNA strands. Hence, when one should like to cut DNA at a predetermined position, the required ARCUT (the sequences and lengths of the pcPNA strands used) can be straightforwardly designed and synthesized. No trial-and-error procedure is necessary.

On the other hand, ZFNs are more advantageous for in vivo applications. They are directly prepared in cells by using the corresponding expression vectors. Consequently, they hold great promise for many applications in therapy and diagnosis, among others. Progress in research on both ARCUT and ZFN has been remarkable so that these new tools, possibly in still more advanced versions, should pave the way to further developments in biotechnology and molecular biology.

Acknowledgements

The authors should like to thank the member of our laboratory who made great contributions to the works presented here. This work was partially supported by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Science, Sports, Culture and Technology, Japan (18001001) and by the Global COE Program for Chemistry Innovation.

Keywords: artificial restriction DNA cutters · DNA cleavage · gene technology · zinc finger nucleases

[1] H. E. Moser, P. B. Dervan, *Science* **1987**, *238*, 645–650.
[2] J. C. Francois, T. Saison-Behmoaras, C. Barbier, M. Chassignol, N. T. Thuong, C. Helene, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 9702–9706.
[3] J. C. Francois, T. Saison-Behmoaras, M. Chassignol, N. T. Thuong, C. Helene, *J. Biol. Chem.* **1989**, *264*, 5891–5898.
[4] J. M. Turner, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1997**, *119*, 7636–7644.
[5] S. White, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1997**, *119*, 8756–8765.
[6] J. W. Trauger, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1998**, *120*, 3534–3535.
[7] A. Y. Chang, P. B. Dervan, *J. Am. Chem. Soc.* **2000**, *122*, 4856–4864.
[8] Y. G. Kim, J. Cha, S. Chandrasegaran, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1156–1160.
[9] J. Smith, J. M. Berg, S. Chandrasegaran, *Nucleic Acids Res.* **1999**, *27*, 674–681.
[10] J. Smith, M. Bibikova, F. G. Whitby, A. R. Reddy, S. Chandrasegaran, D. Carroll, *Nucleic Acids Res.* **2000**, *28*, 3361–3369.
[11] V. Demidov, M. D. Frank-Kamenetskii, M. Egholm, O. Buchardt, P. E. Nielsen, *Nucleic Acids Res.* **1993**, *21*, 2103–2107.
[12] M. Komiyama, J. Sumaoka, *Curr. Opin. Chem. Biol.* **1998**, *2*, 751–757.

[13] E. L. Hegg, J. N. Burstyn, *Coord. Chem. Rev.* **1998**, *173*, 133–165.
[14] A. Blasko, T. C. Bruice, *Acc. Chem. Res.* **1999**, *32*, 475–484.
[15] N. H. Williams, B. Takasaki, M. Wall, J. Chin, *Acc. Chem. Res.* **1999**, *32*, 485–493.
[16] R. Ott, R. Krämer, *Appl. Microbiol. Biotechnol.* **1999**, *52*, 761–767.
[17] J. K. Bashkin, *Curr. Opin. Chem. Biol.* **1999**, *3*, 752–758.
[18] A. Sreedhara, J. A. Cowan, *J. Biol. Inorg. Chem.* **2001**, *6*, 337–347.
[19] S. J. Franklin, *Curr. Opin. Chem. Biol.* **2001**, *5*, 201–208.
[20] C. L. Liu, M. Wang, T. L. Zhang, H. Z. Sun, *Coord. Chem. Rev.* **2004**, *248*, 147–168.
[21] F. Mancin, P. Scrimin, P. Tecilla, U. Tonellato, *Chem. Commun.* **2005**, 2540–2548.
[22] M. Komiyama, Y. Aiba, Y. Yamamoto, J. Sumaoka, *Nat. Protoc.* **2008**, *3*, 655–662.
[23] J. Lohse, O. Dahl, P. E. Nielsen, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11804–11808.
[24] M. Komiyama, Y. Aiba, T. Ishizuka, J. Sumaoka, *Nat. Protoc.* **2008**, *3*, 646–654.
[25] Y. Kitamura, M. Komiyama, *Nucleic Acids Res.* **2002**, *30*, e102.
[26] Y. Yamamoto, M. Mori, Y. Aiba, T. Tomita, W. Chen, J. M. Zhou, A. Uehara, Y. Ren, Y. Kitamura, M. Komiyama, *Nucleic Acids Res.* **2007**, *35*, e53.
[27] T. Igawa, J. Sumaoka, M. Komiyama, *Chem. Lett.* **2000**, 356–357.
[28] J. Bitinaite, D. A. Wah, A. K. Aggarwal, I. Schildkraut, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10570–10575.
[29] N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809–817.
[30] Y. Mineta, T. Okamoto, K. Takenaka, N. Doi, Y. Aoyama, T. Sera, *Biochemistry* **2008**, *47*, 12257–12259.
[31] E. J. Rebar, C. O. Pabo, *Science* **1994**, *263*, 671–673.
[32] H. Wu, W. P. Yang, C. F. Barbas, 3rd, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 344–348.
[33] H. A. Greisman, C. O. Pabo, *Science* **1997**, *275*, 657–661.
[34] M. Isalan, A. Klug, Y. Choo, *Nat. Biotechnol.* **2001**, *19*, 656–660.
[35] J. K. Joung, E. I. Ramm, C. O. Pabo, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7382–7387.
[36] J. A. Hurt, S. A. Thibodeau, A. S. Hirsh, C. O. Pabo, J. K. Joung, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12271–12276.
[37] Y. Miyajima, T. Ishizuka, Y. Yamamoto, J. Sumaoka, M. Komiyama, *J. Am. Chem. Soc.* **2009**, *131*, 2657–2662.
[38] K. M. Vasquez, K. Marburger, Z. Intody, J. H. Wilson, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8403–8410.
[39] H. Puchta, B. Dujon, B. Hohn, *Nucleic Acids Res.* **1993**, *21*, 5034–5040.
[40] P. Rouet, F. Smih, M. Jasin, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6064–6068.
[41] A. Choulika, A. Perrin, B. Dujon, J. F. Nicolas, *Mol. Cell. Biol.* **1995**, *15*, 1968–1973.
[42] M. H. Porteus, D. Baltimore, *Science* **2003**, *300*, 763.
[43] F. D. Urnov, J. C. Miller, Y. L. Lee, C. M. Beausejour, J. M. Rock, S. Augustus, A. C. Jamieson, M. H. Porteus, P. D. Gregory, M. C. Holmes, *Nature* **2005**, *435*, 646–651.
[44] A. Lloyd, C. L. Plaisier, D. Carroll, G. N. Drews, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2232–2237.
[45] D. A. Wright, J. A. Townsend, R. J. Winfrey, Jr, P. A. Irwin, J. Rajagopal, P. M. Lonosky, B. D. Hall, M. D. Jondle, D. F. Voytas, *Plant J.* **2005**, *44*, 693–705.
[46] E. E. Perez, J. Wang, J. C. Miller, Y. Jouvenot, K. A. Kim, O. Liu, N. Wang, G. Lee, V. V. Bartsevich, Y. L. Lee, D. Y. Guschin, I. Rupniewski, A. J. Waite, C. Carpenito, R. G. Carroll, J. S. Orange, F. D. Urnov, E. J. Rebar, D. Ando, P. D. Gregory, J. L. Riley, M. C. Holmes, C. H. June, *Nat. Biotechnol.* **2008**, *26*, 808–816.
[47] J. C. Miller, M. C. Holmes, J. Wang, D. Y. Guschin, Y. L. Lee, I. Rupniewski, C. M. Beausejour, A. J. Waite, N. S. Wang, K. A. Kim, P. D. Gregory, C. O. Pabo, E. J. Rebar, *Nat. Biotechnol.* **2007**, *25*, 778–785.
[48] M. Szczepek, V. Brondani, J. Büchel, L. Serrano, D. J. Segal, T. Cathomen, *Nat. Biotechnol.* **2007**, *25*, 786–793.
[49] T. I. Cornu, S. Thibodeau-Beganny, E. Guhl, S. Alwin, M. Eichtinger, J. K. Joung, T. Cathomen, *Mol. Ther.* **2008**, *16*, 352–358.
[50] J. S. Kim, C. O. Pabo, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2812–2817.
[51] M. Moore, Y. Choo, A. Klug, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 1432–1436.
[52] M. Moore, A. Klug, Y. Choo, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 1437–1441.
[53] A. C. Jamieson, J. C. Miller, C. O. Pabo, *Nat. Rev. Drug Discovery* **2003**, *2*, 361–368.
[54] B. S. Chevalier, B. L. Stoddard, *Nucleic Acids Res.* **2001**, *29*, 3757–3774.
[55] B. L. Stoddard, *Q. Rev. Biophys.* **2005**, *38*, 49–95.
[56] F. Paques, P. Duchateau, *Curr. Gene Therapy* **2007**, *7*, 49–66.
[57] J. Ashworth, J. J. Havranek, C. M. Duarte, D. Sussman, R. J. Monnat, Jr, B. L. Stoddard, D. Baker, *Nature* **2006**, *441*, 656–659.
[58] J. B. Doyon, V. Pattanayak, C. B. Meyer, D. R. Liu, *J. Am. Chem. Soc.* **2006**, *128*, 2477–2484.
[59] H. Li, S. Pellenz, U. Ulge, B. L. Stoddard, R. J. Monnat, Jr, *Nucleic Acids Res.* **2009**, *37*, 1650–1662.
[60] Z. Chen, F. Wen, N. Sun, H. Zhao, *Protein Eng. Des. Sel.* **2009**, *22*, 249–256.

Received: January 26, 2009
Published online on April 24, 2009

