Can pseudocomplementary peptide nucleic acid nucleases (pcPNANs) be a new tool for genetic engineering?

Penghui Shi

Email: shiph150@sina.cn
Gansu Key Lab of Crop Improvement & Germplasm Enhancement, Gansu Provincial Key Laboratory of Aridland Crop Sciences, Gansu Agricultural University, Lanzhou 730070, China
Gansu Agricultural University, Anning District, Lanzhou City, Gansu Province, China

Abstract: Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) comprise a powerful class of tools that are redefining the boundaries of biological research. Although these technologies have begun to enable targeted genome modifications, there remains a need for new technologies that are scalable, affordable, and easy to engineer. In this paper, we propose a new tool for genetic engineering, the pseudocomplementary peptide nucleic acid nucleases (pcPNANs), which are composed of a pseudocomplementary PNA (pcPNA) specific for a DNA target sequence, a FokI nuclease cleavage domain and a nuclear localization signal. pcPNANs may induce targeted DNA double-strand breaks that activate DNA damage response pathways and enable custom alterations. Their cleavage-site is determined by simple Watson-Crick rule, and thus pcPNANs for aimed cleavage of genomes can be straightforwardly designed and synthesized without any selection procedure. Accordingly, the cleavage-site and site-specificity are freely chosen by changing the sequences and the lengths of pcPNA strands. We believe that the potentiality of pcPNAN as a new tool for genetic engineering will be confirmed in the future.

Key-Words: genetic engineering, pseudocomplementary peptide nucleic acid nucleases, DNA double strand break, homologous recombination, error-prone nonhomologous end joining.

Introduction

The generation of genetically modified organisms with gene-targeted modifications or deletions is a powerful tool to analyze gene function and study disease. Until recently, zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) were developed and used for the generation of organisms with gene-targeted deletions and/or modifications when combined with homologous recombination (HR). ZFNs and TALENs enable efficient and precise genetic modifications by inducing targeted DNA double-strand breaks (DSBs) that stimulate the cellular DNA repair mechanisms, including error-prone nonhomologous end joining (NHEJ) and homology-directed repair (HDR) (**Figure.1**). NHEJ generates

short insertions or deletions at the cleavage site. Repair by HR using a DNA template results in gene knock-ins that are either a perfect repair or, if a modified template is introduced, sequence replacement[1,2]. These chimeric nucleases are composed of programmable, sequence-specific DNA-binding modules (ZFs and TALEs) linked to a nonspecific FokI nuclease cleavage domain. FokI requires dimerization to cut DNA. The binding of two heterodimers of designed ZFN-FokI hybrid molecules to two contiguous target sequences in each DNA strand separated by a 6 base-pair cleavage site results in FokI dimerization and subsequent DNA cleavage. The specificity of ZFNs is determined by their polymeric zinc finger domains, the DNA binding properties of which are generated through modular assembly of individual zinc fingers. In contrast to ZFs, the DNA-binding domain of each TALE typically consists of tandem 33-35-amino acid repeat modules rearranged according to a simple cipher to target new DNA sequences. TALE specificity is determined by two hypervariable amino acids that are known as the repeat-variable di-residues (RVDs)[3,4].

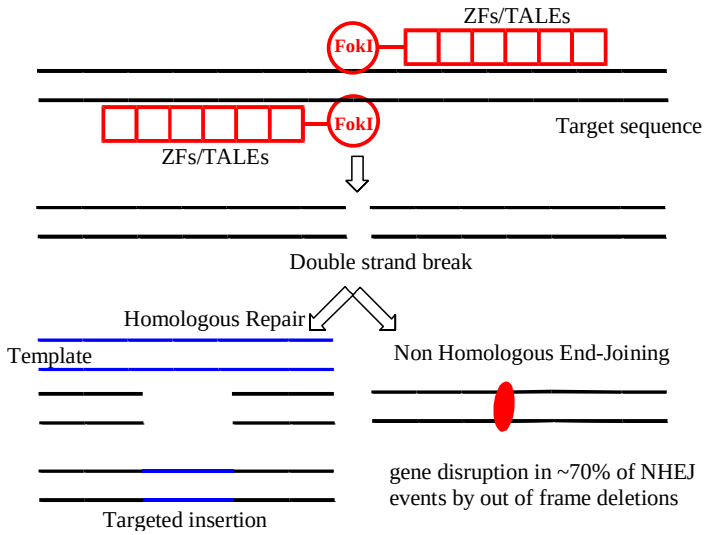

Figure.1 The artificial nucleases (ZFNs and TALENs) mediated targeted genome modification relies on cellular DNA repair pathways. The cellular DNA repair the break by non-homologous end joining, which is imperfect and generates in ~70% of the cases gene deletions due to out of frame deletions of variable lengths, or by homologous recombination when a repair sequence is introduced into the cell together with these artificial nucleases (Fig. modified from Rémy, 2010)[3].

Although these genome-editing technologies (ZFNs and TALENs) have begun to enable targeted genome modifications, for each target site, a new ZFN or TALEN chimeric protein needs to be engineered to recognize the target. This has been a major hurdle in the wide use of these two gene-editing systems because engineering a new protein is no trivial task. In addition, ZFNs and TALENs are complex proteins that must be expressed in cells from viral or plasmid vectors and that can produce variable levels of non-specific, off-target nuclease activity[5,6].

Peptide nucleic acids (PNAs) are DNA analogues in which the naturally occurring

nucleobases are attached via methylene carbonyl linkages to an achiral pseudopeptide backbone of N-(2-aminoethyl)-glycine units[7] (**Figure.2A**). PNAs can bind with complementary DNA or RNA by Watson-Crick bases pairing and have many advantages when used in vivo (**Figure.2B**). Because the backbone of PNAs is neutral, complementary PNA/DNA binds tighter than a comparable DNA/DNA duplex due to the lack of charge repulsion between the two sequences. Single base mismatches in PNA/DNA duplex are more destabilized than equivalent one in DNA/DNA. In addition, PNAs are shown to resist nucleases and proteases in vivo and capable of inhibiting translation in vivo[8].

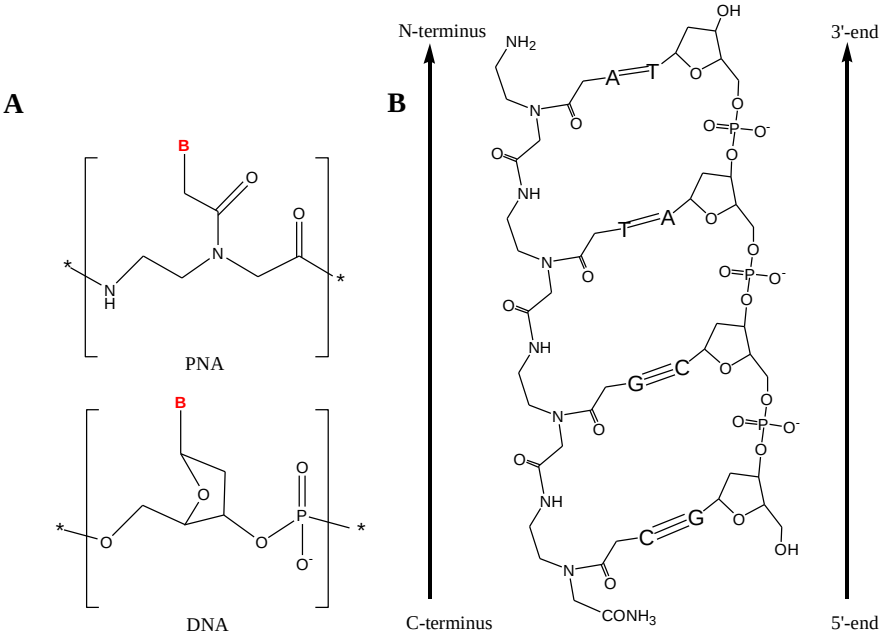

Figure.2 (A) **Chemical structures of PNA and DNA.** (B) **Preferred orientation of PNA:DNA duplexes.**

Although polypyrimidine PNA can recognize double-stranded DNA (dsDNA), the recognition sites are restricted to continuous polypurine/polypyrimidine sequences. For this reason, a new modification of PNAs has been introduced for sequence-unrestricted targeting of double-stranded DNA. Along with ordinary guanines (G) and cytosines (C), these PNAs, dubbed pseudocomplementary PNAs (pcPNAs), carry 2,6-diaminopurines (D) and 2-thiouracils (sU) instead of adenines (A) and thymines (T), respectively (**Figure.3A**). The extra NH_2 group in D and substitution of oxygen by sulfur in sU do not interfere with the formation of hydrogen bonds between T-D and A-sU while there is a steric clash between the amino group in D and thiol group in sU. This makes the formation of pcPNA/pcPNA duplex disfavored. Only DNA/DNA and pcPNA/DNA can be formed. pcPNAs have been prepared and are shown to bind with high specificity and efficiency to complementary targets in double stranded DNA by a mechanism termed “double duplex invasion” in which the duplex is unwound and both DNA strands are targeted simultaneously, each by one of the two pcPNAs (**Figure.3B**). This novel mode of pcPNA-mediated DNA

recognition substantially extends the range of possible DNA targets for pcPNAs, since almost any chosen mixed-base site in duplex DNA can be targeted with pcPNAs (A+ T ≥40%)[9].

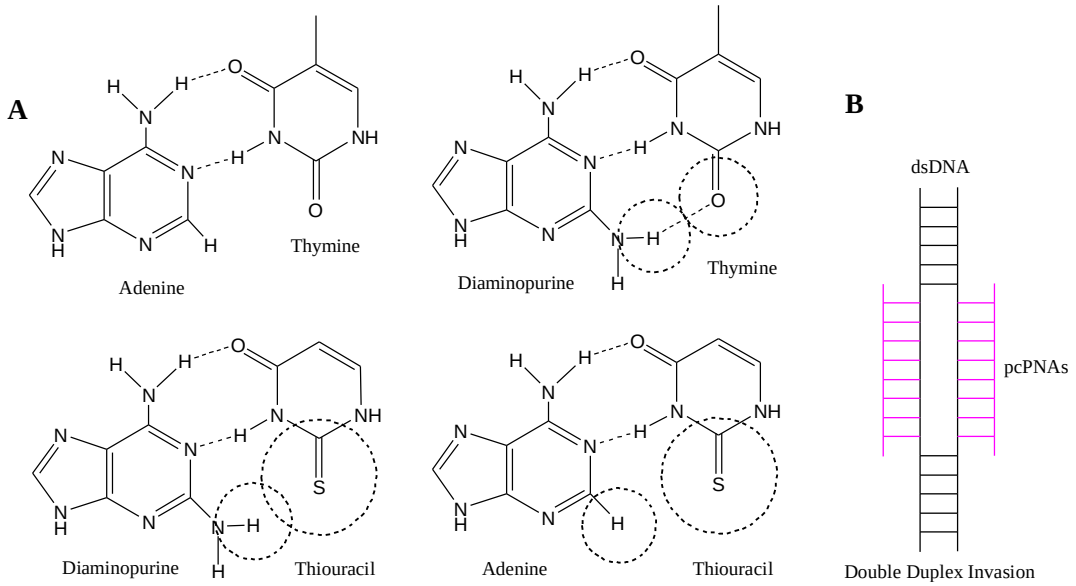

Figure.3 (A) **Base pairing between 2,6-diaminopurine:thymine and between adenine: 2-thiouracil in comparison to the canonical A:T base pair.** Steric hindrance prevents similar base pairing between 2,6-diaminopurine and 2-thiouracil. (B) **Diagram of the double duplex strand invasion complex formed by a pair of pcPNAs on complementary double stranded DNA**[9].

Many studies show that the pcPNAs have the potential to serve as highly specific gene targeting agents. For example, Some researches showed that pcPNAs can block access of RNA polymerase to the corresponding promoter site in vitro thus inhibiting transcription initiation[10,11]. Later it was shown that a pair of psoralen-conjugated pcPNAs can direct the formation of targeted psoralen photoadducts on duplex plasmid DNA in vitro as well as at a chromosomal site in living cells, leading to the production of site-specific mutations with high specificity and efficiency. Psoralen conjugated with pcPNA induced targeted mutagenesis at a mixed sequence with both episomal and chromosomal DNA without obvious off-target modification. The induced mutations and deletions were at the predicted pcPNA binding site[12,13]. Recently, an artificial restriction DNA cutter (ARCUT) has been developed as a chemical tool to cut double-stranded DNA at a pre-determined site (**Figure.4**). These cutters are composed of Ce(IV)/EDTA complex as molecular scissors and two strands of pcPNA as a sequence-recognizing moiety[15]. It was shown that the site-selective scission by ARCUT is satisfactorily recognized by the repair systems in human cells, but the method is also relatively inefficient. There remains a need for new technologies that are scalable, affordable, and easy to engineer.

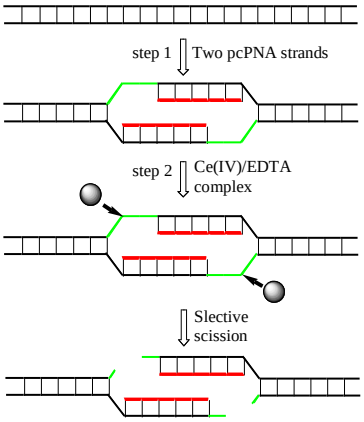

Figure.4 Site-selective scission of double-stranded DNA by ARCUT. Two pcPNA strands (the red lines) invade DNA substrate to form single-stranded portions (the green lines) in both DNA strands (step 1).The phosphodiester linkages in the single-stranded portions are hydrolyzed by Ce(IV)/EDTA (step 2) (Fig. modified from Katada, 2009)[14].

pcPNANs: a new tool for genetic engineering?

On the basis of the mentioned analysis, we propose a novel tool for genetic engineering, the pseudocomplementary peptide nucleic acid nucleases (pcPNANs), which are composed of a pcPNA (≥16bp), a FokI nuclease cleavage domain and a nuclear localization signal (NLS) (**Figure.5**). The specificity of pcPNANs is determined by pcPNAs, the DNA binding properties of which are generated through Watson-Crick bases pairing. The FokI nuclease cleavage domain requires dimerization to create an active endonuclease complex able to cut dsDNA. NLS is a short stretch of amino acids that mediates the transport of the nuclear proteins into the nucleus. In theory, the binding of two heterodimers of designed pcPNAN hybrid molecules to the DNA target sequences results in FokI dimerization and subsequent DNA cleavage, and induction of a DNA double strand break ultimately leads to the activation of a cellular response known as the DNA damage response. Repair by HDR using a DNA template results in gene knock-ins at the targeted site, NHEJ-mediated repair leads to the introduction of small insertions or deletions at the targeted site, resulting in knockout of gene function via frameshift mutations (**Figure.5**).

The Synthesis of pcPNANs mainly includes two steps (**Figure.6**). The first step is the synthesis of pcPNA-peptide conjugates (pcPNA-NLS hybrid molecules). Boc-protected PNA monomers of 2,6-diaminopurine and 2-thiouracil are synthesized according to published methods[9]. These monomers are used together with commercially available Boc-protected G and C PNA monomers. The synthesis of pcPNA-peptide conjugates involved the solid-phase synthesis of the different peptides, then the growth of the pcPNAs chains on the support functionalized with the peptide. The second step is cross-linking of FokI endonucleases and pcPNA-peptide conjugates. The FokI restriction endonucleases are prepared by gene engineering. Then the pcPNA-peptide conjugate are covalently attached to the FokI restriction endonuclease by chemical cross-linking.

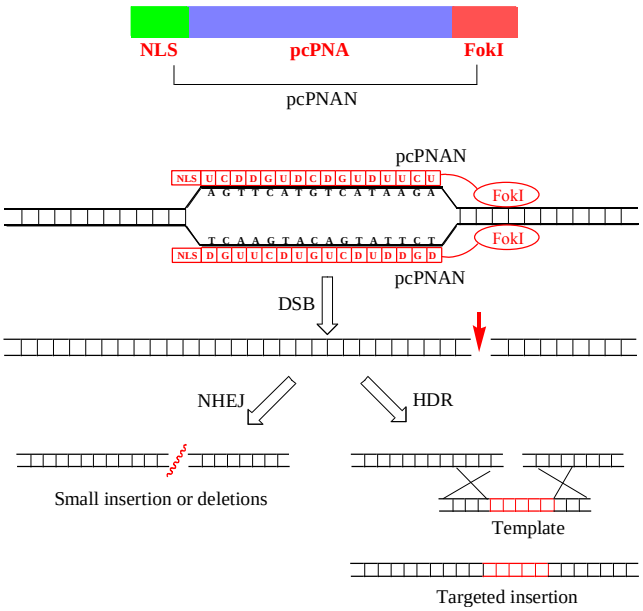

Figure.5 The mechanism of action of pcPNANs for genome modification. The binding of two pcPNA strands of a pair of pcPNANs to DNA substrate results in FokI dimerization and subsequent DNA cleavage at the DNA target site. DNA double-strand breaks (DSBs) can be repaired by error-prone nonhomologous end joining (NHEJ) or homology directed repair (HDR): (a) NHEJ-mediated repair leads to the introduction of small insertions or deletions at the targeted site, resulting in knockout of gene function via frameshift mutations. (b) In the presence of donor DNA with extended homology arms, HDR can lead to the introduction of single or multiple transgenes to correct or replace existing genes. A diagram of the pcPNAN is shown above.

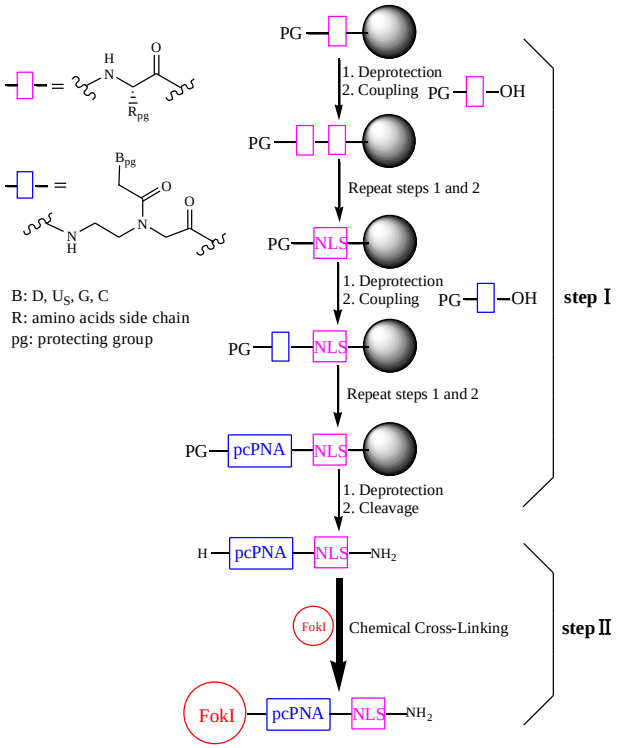

Figure.6 A Strategy for the Synthesis of pcPNANs (Fig. modified from Avitabile, 2011)[16].

pcPNANs can be introduced into the cell by electroporation or microinjection. In addition, Gaj *et al.*[17] recently developed a simple alternative based on the direct delivery of purified ZFN proteins into cells. Inspired by this, we deduce that pcPNANs may also penetrate the cell in the absence of additional modification.

One of the most significant advantages of this tool is that the site of selective cleavage is a priori determined by Watson-Crick base pairings between the pcPNA strands and the DNA substrate. Thus, pcPNAN for aimed cleavage of genomes can be straightforwardly designed and synthesized without any selection procedure. The site specificity is high enough to cut one site in human genome, because it was confirmed that ARCUT strictly distinguishes the target site from highly analogous sequences and cuts it[18,19]. From statistical viewpoints, 16 base-pairs in the target sequence are recognized by this pcPNANs in an almost completely on-off way. Because the cleavage site of this pcPNANs should appear at every $4^{16}(=4.29\times10^{9})$ bp, which is larger than the size of human genome (3.0×10^{9}bp). Of course, when necessary, the site-specificity can be further increased by using still longer pcPNAs.

Conclusions

In this paper, we propose a new tool for genetic engineering: pseudocomplementary peptide nucleic acid nucleases (pcPNANs). pcPNANs are hybrid molecules composed of a pcPNA specific for a DNA target sequence, a FokI nuclease cleavage domain and a nuclear localization signal. pcPNANs may induce targeted DNA double-strand breaks (DSBs) that activate DNA damage response pathways and enable custom alterations. If the tool is proved feasible by an experiment, it will be very useful in molecular biology research. Because unlike other technologies, pcPNAN for aimed cleavage of genomes can be straightforwardly designed and synthesized without any selection procedure. Our future target is to implement this design proposal, we believe that some details may need to be revised, but the basic idea for such a design is reasonable, and the potentiality of pcPNAN as a new tool for genetic engineering will be confirmed. Therefore, more studies will be necessary to test its feasibility and efficiency.

References

1. Porteus MH, Carroll D: **Gene targeting using zinc finger nucleases**. *Nature biotechnology.* 2005; **23**(8): 967-973.

2. Miller JC, Tan S, Qiao G, *et al.*: **A TALE nuclease architecture for efficient genome editing**. *Nature biotechnology.* 2010; **29**(2): 143-148.

3. Rémy S, Tesson L, Ménoret S, *et al.*: **Zinc-finger nucleases: a powerful tool for genetic engineering of animals**. *Transgenic research*. 2010; **19**(3): 363-371.

4. Mussolino C, Morbitzer R, Lütge F, *et al.*: **A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity**. *Nucleic acids research*. 2011; **39**(21):

9283-9293.
5. Foley JE, Yeh JRJ, Maeder ML, *et al.*: **Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN)**. *PloS one.* 2009; **4**(2): e4348.
6. Geurts AM, Cost GJ, Freyvert Y, *et al.*: **Knockout rats via embryo microinjection of zinc-finger nucleases**. *Science*. 2009; **325**(5939): 433-433.
7. Nielsen P E, Egholm M, Berg RH, *et al.*: **Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide**. *Science*. 1991; **254**(5037): 1497-1500.
8. Dias N, Stein CA: **Antisense oligonucleotides: basic concepts and mechanisms**. *Molecular cancer therapeutics*. 2002; **1**(5): 347-355.
9. Lohse J, Dahl O, Nielsen PE: **Double duplex invasion by peptide nucleic acid: a general principle for sequence-specific targeting of double-stranded DNA**. *Proceedings of the National Academy of Sciences*. 1999; **96**(21): 11804-11808.
10. Nielsen PE, Egholm M, Buchardt O: **Sequence-specific transcription arrest by peptide nucleic acid bound to the DNA template strand**. *Gene*. 1994; **149**(1): 139-145.
11. Larsen HJ, Nielsen PE: **Transcription-mediated binding of peptide nucleic acid (PNA) to double-stranded DNA: sequence-specific suicide transcription**. *Nucleic acids research*. 1996; **24**(3): 458-463.
12. Kim KH, Nielsen PE, Glazer PM: **Site-directed gene mutation at mixed sequence targets by psoralen-conjugated pseudo-complementary peptide nucleic acids**. *Nucleic acids research*, 2007; **35**(22): 7604-7613.
13. Kim KH, Nielsen PE, Glazer PM: **Site-specific gene modification by PNAs conjugated to psoralen**. *Biochemistry*, 2006; **45**(1): 314-323.
14. Katada H, Komiyama M: **Artificial restriction DNA cutters as new tools for gene manipulation**. *ChemBioChem*. 2009; **10**(8): 1279-1288.
15. Aiba Y, Sumaoka J, Komiyama M: **Artificial DNA cutters for DNA manipulation and genome engineering**. *Chemical Society Reviews*. 2011; **40**(12): 5657-5668.
16. Avitabile C: **Synthesis and characterization of PNA based molecules able to interfere with gene regulation**. 2011.
17. Gaj T, Guo J, Kato Y, *et al.*: **Targeted gene knockout by direct delivery of zinc-finger nuclease proteins**. *Nature methods*. 2012; **9**(8): 805-807.
18. Katada H, Harumoto T, Shigi N, *et al.*: **Chemical and biological approaches to improve the efficiency of homologous recombination in human cells mediated by artificial restriction DNA cutter**. *Nucleic acids research*. 2012; **40**(11): e81-e81.
19. Komiyama M: **Cut-and-Paste of DNA Using an Artificial Restriction DNA Cutter**. *International journal of molecular sciences.* 2013; **14**(2): 3343-3357.