International Journal of Science, Environment and Technology, Vol. 5, No 6, 2016, 3670 – 3677

ISSN 2278-3687 (O)
2277-663X (P)

MOLECULAR SCISSORS, THE WORK HORSES OF GENOME EDITING THEN AND NOW

Mashidur Rana*, Swaraj Biswas and Sudipta Bhat[1]
Division of Biochemistry, Indian Veterinary Research Institute,
Izatnagar, U.P-243122
[1]Division of Virology, Indian Veterinary Research Institute, Mukteshwar,
Uttarakhand-263138
E-mail: mailmashidur@gmail.com (**Corresponding Author*)

Abstract: The molecular scissors comprises of the restriction endonucleases, homing endonucleases, ZFNs (zinc finger nucleases), TALENs (Transcription Activator Like Effector Nucleases) and CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (CRISPR associated) system. Restriction endonucleases are enzymes which cleave DNA at particular site within the site of recognition. Homing Endonucleases are either RNA (group I and group II introns) coded or inteins which mediate site specific cleavages. Whereas ZFNs and TALENs are chimeric proteins which cleave DNA after binding to a specific site by their DNA binding domain. On the other hand CRISPR/Cas system does cleave DNA by gRNA guided method. Here we will discuss the attributes of all molecular scissors available to the researches at hand to precisely perform genome editing.
Keywords: restriction endonuclease, homing endonuclease, group I and group II introns, inteins, ZFNs, TALENs, CRISPR/Cas, Guide RNA (gRNA), genome editing.
Abbreviations: ZFNs: Zinc Finger Nucleases, TALEN: Transcription Activator Like Effector Nucleases, CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats, Cas: CRISPR associated.

Introduction

In pursuit of precise, accurate and user customize scission; molecular scissors had developed in leaps and bounds since their discovery. Genome editing has now become an integrative part of personalized medicine. This necessity plays a pivotal role in innovation of reprogrammable endonucleases. Predecessors like restriction endonuclease, homing endonuclease are now replaced with zinc-finger nuclease, TALE endonucleases. Eventually with the dawn of 21st century the CRISPR/Cas system came to limelight and started taking upper hand. Age old type II restriction endonucleases has cut sites of 8bps maximum thus their probability to cut eukaryotic genome only once is sufficiently less. With modular enzymes using the endonuclease domain of Fok1 like in ZFNs and TALENs the cleavage site is considerably long, imparting marked precision. Further with the advent of CRISPR/Cas system, RNA guided DNA endonuclease came into the arena of gene editing. Here we will

*Received Sep 27, 2016 * Published Dec 2, 2016 * www.ijset.net*

try to review molecular scissors starting from restriction endonuclease and homing endonucleases through ZFNs and TALENs to CRISPR/Cas system. We will highlight their attributes which accord their specificity.

Restriction Enzymes

Restriction Enzymes are the first generation molecular scissors, their discovery and characterization was recognized by 1978 Noble Prize for Physiology and Medicine to Arber, Smith and Nathans. The restriction enzymes are endonucleases serving as a part and parcel of restriction modification system imparting innate immunity to bacteria (Kruger and Bickle, 1983). Pioneering studies in the field of restriction enzymes were done in the laboratories of Luria and Bertani (Luria and Human, 1952). It was shown that for a λ phage which can grow flourishingly in one *Escherichia coli* strain when allowed to grow in another reduced significantly (Bertani and Wegle, 1953). Thus the second host cell in some way or the other is restricting the λ phage. In 1960, it was revealed by Arber and Meselson that this restriction is imposed due to enzymatic cleavage of phage DNA and was therefore named restriction endonuclease (Meselson and Yuan, 1968). These enzymes do cleave foreign DNA at particular sites; in contrast the host genome is shielded by methylation of those key bases (Kobayashi, 2001). In 1970, Smith and Kelly discovered the first TypeII restriction enzyme called *Hind*II (Kelly and Smith, 1970). Besides Type II enzymes which serve as an indispensible tool for genetic engineering there are Type I, Type III and Type IV enzymes. The chief challenge of any endonuclease is to attain a high degree of specificity. Within the enzyme-substrate complex, the DNA substrate is distorted to generate a magnesium binding site between the enzyme and DNA. The magnesium ion then binds and activates a water molecule, which attacks the phosphor-di-ester bond (Berg *et.al.,* 2002). Thousands of restriction enzyme has been discovered till date, REBASE is a database of restriction enzyme database maintained by New England Biolabs.

Figure 1. *Restriction Endonucleases*. Hind III restriction endonuclease which recognizes GA/ATTC i.e a hexacutter with 5' overhang

Homing Endonucleases

"Homing" is lateral high frequency site specific gene transfer event of an intervening genetic sequence. Besides group I and group II introns the intervening sequence comprise of inteins (Chevaliar and Stoddard, 2001). These sequences are copied and transferred to a specific

insertion locus within a cognate allele devoid of the sequence. Eventually leads to duplication of the mobile elements within the genome. Homing is initiated by site specific endonucleases that are coded by ORFs within the sequences (Dujon.B, 1980). This extra-Mendelian genetic manoeuvre of mobile was first reported in "*ω*" genetic marker of *Saccharomyces cerevisiae* (Bos *et.al.,* 1978). This protein is rechristened as I-*Sec*I, and serves as the oldest representative of intron encoded homing endonucleases. Group I homing endonuclease are most exhaustively studied superfamily comprising of three families namely LAGLIDADG, GIY-YIG, His-Cys box and HNH (Belforst and Roberts, 1997). Two salient features of homing endonucleases make them to stand apart from other site specific nucleases. Firstly, despite their small size (<40 kDa) they bind to exceptionally long (14-40 bp) homing sites which ensures its precise specificity. Secondly, they tolerate subtle mutations in homing site sequences (Jurica and Stoddard, 1999). Cognate site recognition of these enzymes is attributed by β sheets that fold to juxtapose, size and curvature of major groove (Phillips, 1994). Sequence specificity is due to hydrogen bonding with the unique pattern of hydrogen bond donors and acceptors presented by bases in the major groove. Basic mechanism opted to cleave the covalent bond between phosphate and oxygen is by classical S_N2 mechanism (Chevaliar and Stoddard, 2001).

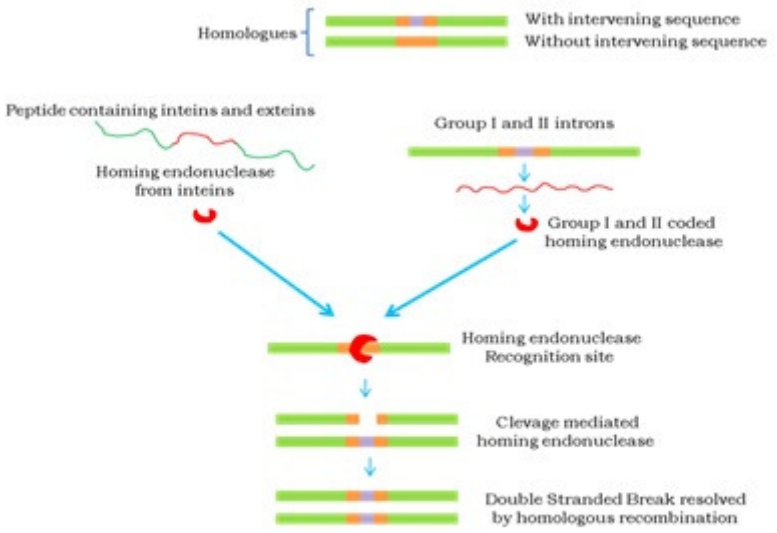

Figure 2. *Homing*. Homing of group I , group II and inteins which leads to incorporation of intervening sequence into the allele not containing the sequence earlier. The homing endonuclease cleaves the specific site and the double stranded break is repaired by homologous recombination.

ZincFinger Nucleases

To serve as a work horse of gene editing an endonuclease should be able to recognize long target sequence. By recognizing long target sequence it ensures sequences long enough to occur only once in the genome. ZFNs optimally meet this specification by harnessing DNA binding domain of eukaryotic transcription factor containing Zinc and the nuclease of domain of Fok1 restriction enzyme (Caroll, 2011). Both DNA-binding and catalytic domain can be re-programmed to customize as per user's choice. The Cys_2-His_2 zinc- finger domain is the most frequently appearing DNA- binding motif of eukaryotes. It is also the second most commonly encoded protein domain in human genome (Thomas *et.al*., 2013). Individual zinc-finger is a conserved ββα configuration of 30 residues (Beerli *et.al.,*2002). Diverse amino acids on the surface of the α helices comes in contact with three bases in the major groove of DNA.

Figure 3. *Zinc Finger Nucleases*. The zinc finger domain of DNA binding proteins has been fused with nuclease domain of endonucleases. Green: DNA binding domain, Purple: Nuclease, Blue: Linker.

Besides modular assembly several other strategies viz. OPEN, bacterial one-hybrid etc. have been developed to identify zinc finger proteins with desired specificity and precision for use in genetic engineering. Of late by incorporating more fingers the enzymes are manipulated to cleave longer and rarer targets (Fyodor *et.al*., 2010).

TALE Nucleases

Insight into plant pathogenic bacteria of *Xanthomonas* genus led to the discovery of TALE proteins. Nature Methods, in the year 2011 accredited precise genome editing as method of the year counting TALEN as a preeminent tool. TALE proteins constitute a DNA binding domain, a nuclear localization signal and domain that activates the target gene (Moscou and Bogdanobe, 2009). DNA binding domain contains monomers comprising tandem repeats of 34 residues required to bind to a single nucleotide in the target sequence. Within the tandem repeats two residues located at position 12 and 13 are highly variable, called repeat variable di-residue, RVD. These RVDs are responsible for single nucleotide recognition (Mak *et.al,* 2012). Investigations revealed that the 5'end of the target sequence should contain a thymidine monomer. The intelligibility of one monomer one nucleotide recognition has fascinated researchers across the globe. This eccentric attribute has been harnessed in

expressing chimeric proteins containing different RVDs and the catalytic domain of type IIs restriction endonuclease FokI (Christian, *et al*., 2010). TALENs work in pairs and their cognate sites are so chosen to be located in the opposite strand separated by spacer.

Figure 4. *Transcription Actiavtor Like Effector Nucleases*. The TALE protein DNA binding domain has been fused with nuclease domain of endonucleases. Orange: DNA binding domain, Green: Nuclease.

The genetic construct essentially comprises of a nuclear localization signal besides the requisite RVD and catalytic site. Upon entering the nucleus artificial nucleases bind to cognate sites the FokI catalytic domains dimerize and yield a double stranded break (Nemuydri *et al*., 2014).

CRISPR/Cas System

In late 80s queer repeats were discovered in bacterial genome (Ishino *et al*., 1987). Genome sequencing unveiled this repeats have unique feature of short regions of unique DNA spacers separated from each other by palindromic repeats. This distinctive feature was crucial in naming them as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). CRISPR cassettes are located in close approximation with Cas genes which express proteins having helicase and nuclease activity (Haft *et al*., 2005). In 2007, it was reported that certain *Streptococcus thermophilus* cells were resistant to a particular bacteriophage. This exceptional attribute was due to a spacer sequence within the CRISPR locus that is complementary to the bacteriophage genomic DNA fragment (Barrangau *et al*., 2007). Thus it became perceptible that besides restriction modification system, CRISPR/Cas system does also serve as limiter of horizontal genetic transfer thus providing acquired immunity. In due course with better understanding of CRISPR/Cas system, it emerged as a programmable RNA-guided DNA endonucleases. The spacers transcribe and process into CRISPR RNA (crRNA) containing protospacers, complementary to the foreign DNA sequence. These crRNAs hybridize by Watson-Crick pairing to trans-activating crRNAs (tracrRNAs) to administer sequence specific cleavage and silencing of foreign DNA bearing the protospacer (Nemuydri *et al.,* 2014). In 2015, the nuclease Cpf1 of CRISPR/Cpf1 system of bacteria *Francisella novicidia* was brought to limelight. Cpf1 unlike Cas9 brings about staggered cut instead of non-cohesive blunt cut. Nuclease activity of Cpf1 depends on T rich protospacer adjacent motif (PAM) requiring only crRNA for successful execution (Fonfara *et al*., 2016).

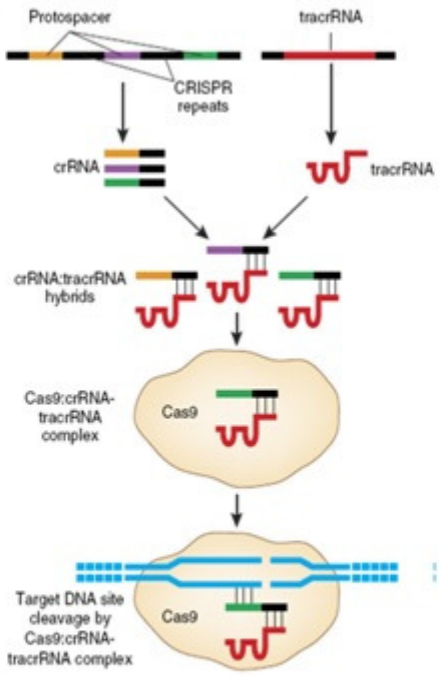

Figure 5. *CRISPR/Cas*. CRISPR systems incorporate foreign DNA into CRISPR arrays, which then produce crRNAs bearing "protospacer" regions that are complementary to the foreign DNA site. crRNAs hybridize to tracrRNAs and this pair of RNAs in association with Cas9 nuclease recognize and cleave foreign DNAs bearing the protospacer sequences *[adapted from Sander, J.D and Joung J.K, 2014.Nat. Biotechnol. (32) pp.349]*

Conclusion

Poised with delivering precise site specific nuclease, scientists have discovered user customized enzymes. Undoubtedly these cutting edge discoveries have bestowed remarkable success in manipulating genes and treating genetic disorders. But bottleneck still remains regarding optimum delivery methods such as adenovirus dependent system can deliver ZFNs fairly but cannot accommodate large size TALENs, sequence specificity of PAM region of CRISPR/Cas system is also an important constraint. In a nutshell, the toddling from predetermined readymade restriction endonuclease to site specific nuclease having tailored seek and destroy property has brought revolution in the field of genome editing.

References:

[1] A. A. Nemudryi, K. R. Valetdinova, S. P. Medvedev, and S. M. Zakian., 2014. TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery. Acta Naturae. (6): 19-40.

[2] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P.,2007. Science (315): 1709-12.

[3] Beerli, R.R. and Barbas, C.F.III. 2002. Engineering polydactyl zinc-finger transcription factors. Nat. Biotechnol. (20): 135-41.

[4] Belforst, M. and Roberts, R.,J. 1997. Homing endonucleases: keeping the house in order. Nucleic Acids Res. (25): 3379-88.

[5] Berg JM, Tymoczko JL, Stryer L. Biochemistry. 5th edition. New York: W H Freeman; 2002. Section 9.3, Restriction Enzymes: Performing Highly Specific DNA-Cleavage Reactions.

[6] Bertani, G and Weigle, J.J 1953. Host controlled variation in bacterial viruses. J. Bacteriol (65): 113-21.

[7] Bos, J.L., Heyting, C., Borst, P., Arnberg, A.C., and Van Bruggen, E.F. 1978. An insert in the single gene for the large ribosomal RNA in yeast mitochondrial DNA. Nature (275): 336-38.

[8] Carroll, D. 2011. Genome engineering with zinc-finger nucleases. Genetics (188): 773-82.

[9] Chevalier, B.S and Stoddard, B.L. 2001. Homing Endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucleic Acids Res. (29): 3757-74.

[10] Christian, M., Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF., Genetics (186):757-61

[11] Dujon, B. 1980. Sequence of intron and flanking exons of the mitochondrial 21S rRNA gene of yeast strains having different alleles at omega and rib-1 loci. Cell (20): 185-97.

[12] Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. 2016. The CRISPR-associated DNA-cleaving enzyme cpf1 also processes precursor CRISPR RNA. Nature (532): 517-21

[13] Haft, D.H., Selengut, J., Mongodin, E.F., Nelson, K.E. 2005. A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes Plos Comput. Biol. (6): e60

[14] Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A.,1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J. Bacteriol. (169): 5429-33.

[15] Jurica, M.S., and Stoddard, B.L. 1999. Homing endonucleases: structure, function and evolution. Cell. Mol. Life Sci. (55): 1304-26.

[16] Kelly, T.J and Smith H.O. 1970. A restriction enzyme from Hemophilus influenzae. II. J. mol. biol. (51):393-09.

[17] Kobayashi, I. 2001. Beheviour of restriction-modification systems of selfish mobile elements and their impact on genome evolution. Nucleic Acids Res.(29):3742-56.

[18] Kruger, D.H. and Bickle, T.A. 1983. Bacteriphage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. Micrbiol. Rev. (47):345-60

[19] Luria, S.E and Human, H.L. 1952. A nonhereiditary, host-induced variation of bacterial viruses. J. Bacteriol. (64): 557-69

[20] Mak, A.N. Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL., 2012. The crystal structure of TAL effector PthXo 1bound to its DNA target. Sceince (335): 716-19.

[21] Meselson, M and Yuan, R. 1968. DNA restriction enzyme from E. coli. Nature (217):1110-4.

[22] Moscou, M.J. and Bogdanove, A.J. 2009. A simple cipher governs DNA recognition by TAL effcetors. Science (326): 1501.

[23] Phillips, S.E. 1994. The beta-ribbon DNA recognition motif. Annu. Rev. Biophys. Biomol. Struct. (23). 671-01.

[24] Sander JD, Joung JK. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. (32): 347-55

[25] Thomas, G., Charles, A.G., Carlos, F.B. 2013. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. (31): 397-405.

[26] Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. 2010. Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. (11): 636-46.