

BMB Reports – Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-16-128

Title: CRISPR as a strong gene editing tool

Article Type: Mini Review

Keywords: CRISPR; gene editing; Non-homologous end joining; Homology directed repair

Corresponding Author: haihong Shen

Authors: shengfu shen¹, tiing jen loh¹, hongling shen¹, xuexiu zheng¹, haihong Shen^{1,*}

Institution: ¹Willston Northampton School,
²life sciences, gwangju institute of science and technology,

Title:

CRISPR as a strong gene editing tool

Shengfu Shen^a, Tiing Jen Loh^b, Hongling Shen^b, Xuexiu Zheng^b and Haihong Shen^b

^aWillston Northampton School, Easthampton, MA 01027, USA

^bSchool of life Sciences, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea

*Corresponding author: Haihong Shen, Phone: 82-62-715-2507, Fax: 82-62-715-2484, E-mail address: haihongshen@gist.ac.kr

*Corresponding author: Xuexiu Zheng, Phone: 82-62-715-2520, Fax: 82-62-715-2484, E-mail address: xuexiuzheng@gist.ac.kr

Running Title:

CRISPR as a strong gene editing tool

Keywords:

CRISPR, gene editing, Non-homologous end joining, Homology directed repair

ABSTRACT

Clustered regularly-interspaced short palindromic repeats (CRISPR) is a new and effective genetic editing tool. CRISPR was initially found in bacteria to protect it from virus invasions. In the first step, specific DNA strands of virus are identified by guide RNA that is composed of crRNA and tracrRNA. Then RNase III is required for producing crRNA from pre-crRNA. In the second step, a crRNA:tracrRNA:Cas9 complex guides RNase III to cleave target DNA. After cleavage of DNA by CRISPR-Cas9, DNA can be fixed by Non-Homologous End Joining (NHEJ) and Homology Directed Repair (HDR). Whereas NHEJ is simple and random, HDR is much more complex and accurate. Gene editing by CRISPR is able to be applied to various biological fields such as agriculture and treating genetic diseases in humans.

Introduction

Clustered regularly-interspaced short palindromic repeats (CRISPR) is a new and effective genetic editing tool. Because of the efficiency and affinity, CRISPR has been considered to be a revolutionary invention in the field of genetic-editing and molecular biology. Compared with other genetic-editing techniques, CRISPR has an irreplaceable advantage, which is its accessibility. The application of CRISPR does not require years long training and high laboratory cost. It only demands some basic training and less than 100 dollars for buying the CRISPR molecule, which enables researchers to study genetic editing and process an experiment. By developing CRISPR, numerous applications will be invented in the fields of agriculture, environmental science, biology, and public health care. Of course, there will be huge commercial interest in this field, which will boost the development of those applications. Even though CRISPR has shown us a bright and promising future, it also has its own challenges. One of the biggest challenges is the ethic problem.

How does CRISPR system correctly cut target DNA?

CRISPR is such a strong genetic editing tool is mainly caused by its unique structure and design. But, it might be hard to believe that the creation of CRISPR is based on an immune system of bacteria. At the beginning of the research, scientists found that some genes in the bacteria seem useless for constructing the cells and maintain survival of bacteria (1). But, as scientists discovered that this “useless” gene is commonly existed in bacteria and other micro-organisms, they started to unwind the mystery of the gene. This special gene in bacteria is coding for an interesting function for protecting the bacteria from the invasion by viruses (2). It is also the most primitive form of CRISPR. This defense system of bacteria can achieve elimination of the

invading virus by two steps. The first step is to target the invaders and a specific DNA strand of the virus. Then the next step is to cut the target DNA strand of the virus by using a nuclease. As described in figure 1, in step one, in order to identify the location of the viruses and target them, the defense system needs a help from the guide RNA. The guide RNA, which is composed of crRNA and tracrRNA, is responsible for locating and targeting the foreign invaders of the bacteria through identification of specific nucleotides of the invading viruses (3). But, guide RNA cannot just automatically find and match the specific nucleotides that bacteria looks for. It has to go through a complex procedure to develop the ability to locate a piece of DNA sequences in the viral genome. When Viruses start to invade the living cells of bacteria, they will release their own invading DNA strands into the bacteria cells for turning the invaded cells into factories that produce more Viruses. At this moment, the immune system of bacteria will make a copy of one DNA strand of the invading virus, and then inserts it into a group of genome of bacteria, which is called CRISPR array (4, 5).

The DNA of the virus will be stored in the spacers between two palindromic repeats in the CRISPR array (6). Then bacteria will transcribe a portion of CRISPR array to make a pre-crRNA strand (7) that also includes multiple complementary spacers and palindromic repeats. In the process of turning pre-crRNA to crRNA, a tracrRNA and a double-stranded RNA-specific ribonuclease are required (8). TracrRNA will bind to one of the palindromic repeats of the pre-crRNA (9), and the ribonuclease III will cut the tracrRNA (8, 10) . Finally this separated pre-crRNA is matured and called Cr-RNA. Now the crRNA and tracrRNA will bind to the cas9 nuclease to form a crRNA:tracrRNA:Cas9 complex. CrRNA in the complex is completely complementary to a DNA sequence of the invading virus, so that crRNA guide the cas9 nuclease to the viral DNA sequence that are complementary to the crRNA (11). Importantly, a protospacer

adjacent motifs (PAM) is required during this process. PAM is a sequence of nucleotides that are located close to the sequences complementary to the crRNA (12), and acts as an initiation of attaching crRNA to its complementary strand in virus (13). After crRNA:tracrRNA:Cas9 complex is launched to the invading virus's DNA strand, the location of the virus is clearly identified. The tool that kills the virus is the Cas9 nuclease. By binding crRNA:tracrRNA:Cas9 complex to the targeted strand, the Cas9 nuclease will cut the DNA strands of the virus near the PAM (13), so that the virus will have incomplete DNA strands, which will defeat the viruses.

Non-Homologous End Joining (NHEJ) and (Homology Directed Repair) HDR.

CRISPR's capability of cutting not only viruses but also genomes of every living life leads the further development of genetic editing in other organisms. Inside bacteria, CRISPR is able to silence invading viruses by tracing and cutting their DNA strands, due to the existence of Cas9 nuclease and crRNA:tracrRNA:Cas9 complex (11). This tracing and cutting mechanism can also be used in cutting the DNA strand of other organism. Once CRISPR is inserted into a specific site of a DNA strand, it can perform the function of cutting DNA strand by introducing a DNA strands breaks (DSB) to the site, which is just like how the CRISPR kills the invading viruses of the bacteria. But using CRISPR to cut the DNA strand is not the point of doing the CRISPR research. Instead, CRISPR is used to introduce the DSB to those DNA strands for a chance of changing the genetic sequence of the nucleotides on that DNA strand (14). However, CRISPR will not provide such a function of changing the specific nucleotide, because the only function it can provide is cutting the DNA strand. Thus, a method of fixing and editing the broken DNA strands is necessary for the DNA to perform its function again and introducing changes to the targeted DNA sequences. There are two different mechanisms to edit and fix the damaged DNA

sequence, which are Non-Homologous End Joining (NHEJ) and Homology Directed Repair (HDR) (15).

NHEJ is a natural methods of healing damages of DNA strands in most organisms (16). For example, when UV lights damages the DNA strands of the skin cells, our body can use NHEJ to reconnect two broken DNA strands, so they are still functioning. Generally, this method is relatively simple and effective in terms of fixing damages of genes, because this fixing pathway does not require a homologous template to fix the DNA. As described in figure 2, in NHEJ, the Ku protein will bind to the two ends of the broken DNA strands and form a Ku DNA end complex (17). This Ku DNA end complex will combine with DNA-PKcs complex to chop away overhangs of nucleotides near the end of the two broken strands (18). Then, with the help from the XLF:XRCC4 DNA ligase IV, the two broken ends of DNA strands will join together and reconnect to each other via ligation (19). Due to the introduction of DSB and some nucleotide deletions during repairing, the original DNA sequence is permanently changed after NHEJ. Thus, eventually, repairing DNA strands via using NHEJ is easy to introduce some unexpected mutations to the targeted DNA strand, because people cannot control what nucleotide is knocked out and what is not. Importantly, due to the simplicity and efficiency of this mechanism, it may have a bright future in the field of non-human genetic editing, like using CRISPR in the field of agriculture and fixing eco-system.

In contrast to the simplicity and randomness of NHEJ, HDR is a much more complex and accurate tool of fixing and editing the genetic sequence of the targeted organism. The main factor that contributes to the greater accuracy of editing is the introduction of a specific template strand (20, 21). NHEJ can repair and edit DNA sequence through deletions of nucleotides to the broken site. Thus, the NHEJ does not have a promising accuracy on processing genetic editing. But

HDR is able to edit the DNA sequence with great accuracy by using a prepared DNA strands (22). When cr-RNA:tracrRNA:Cas9 complex launches to the PAM site and DSB is introduced, one of the broken DNA strand will lose some nucleotides until it has a 3' end at the broken site (23). This DNA strand will unzip the prepared DNA strands, and begin to launch at 5' end of one DNA strand (24, 25). This launching site will become a primer of DNA replication (24). When the invading strand is properly launched, DNA polymerase will pull some nearby nucleotides to create a complementary strand of DNA of the prepared DNA strand from the primer (26). Thus, one of the broken strands successively repairs the broken site and obtains desirable DNA sequences from the prepared strand. Then, as the broken site obtains the desired DNA strand from the provided DNA template, the rest of the missing nucleotides can be reconstructed based on the nucleotides that are available. Eventually, after repairing, the broken site of DNA strands obtain designed genetic sequences from the prepared DNA strands. With assistance of fixing and repairing from NHEJ and HDR, CRISPR is able to edit any gene in any organism.

Applications of CRISPR in gene editing

Theoretically, human obtains the ability to change the genome sequences using CRISPR. This possibility sounds very attractive to the fields that are related to biology, because there are numerous opportunities that can be offered by CRISPR.

For agriculture, CRISPR may be used to edit the genome of crops and plants to offer them phenotypes that provide resistance to bacteria, so that less pesticide will be used to treat it, which will possibly save money for farmers and lower the damages from overuse pesticides. CRISPR can also be used in treating genetic diseases that are caused by some mutations in DNA strands.

For example, Huntington's disease is caused by the mutation on HTT genes (27), which can be

seen as having incorrect nucleotide pairs on the genome. If CRISPR can be used to replace the incorrect nucleotides on HTT gene with correct ones, Huntington's disease can be easily treated (28). Although CRISPR has infinite applications, it still faces two challenges. The first challenge is the off-target problem. Recently, a Chinese team tried to use CRISPR to edit the genetic information of human embryos to code for a mutation on CCR5 gene for getting resistance to HIV. HIV needs a very specific launching spot on normal white blood cells to invade them, but a mutation on CCR5 gene can cause a change on that spot and HIV will lose the ability to invade those cells (29). Thus, by introducing a mutation to CCR5 gene, there is a hope to develop resistance to HIV. However, when scientists tried to introduce this mutation, they discovered a high off-target rate for CRISPR (30). These mismatched nucleotides or unexpected mutations can all lead to serious problems.

Another problem that CRISPR has to face before real applications is the ethics. The ethical problems of CRISPR does not come from editing the DNA sequences of organism like what happened last century when people still concerned about side effect of genetic editing. Most of countries now offer permissions for processing genetic editing on organisms for improving agriculture or other biological use. However, even if permission of editing organisms for agricultural or biological use is granted, editing human genome is still very controversial. There are two main reasons that cause this ethical problem. First, CRISPR is not a completely developed and mature method yet to process precise genetic editing, especially when the target is human genome. If mistakes or off-target effect takes place during genetic editing, patients may result in serious genetic mutation, which will not only affect the health condition of patients but also the future generations of patients. Thus, before scientists develop more advanced version of CRISPR, it is better not to use it on human. Another ethical problem for using CRISPR is not

related to the technical problem. CRISPR can be considered as the strongest genetic editing tool that is invented in human history, because it enables scientists to edit any genome in any organisms. Thus, some ethicists concern about the consequences of CRISPR overuse in editing human genome. If CRISPR is commonly used to modify human genome, huge social conflicts can be possibly generated. For example, people may be able to design their children based on their interests and some people can possibly have a much longer life span than the others due to the use of CRISPR. These possible applications sound very attractive to us, but they may cause very serious social conflicts in terms of the increasing social inequality and other problems. Thus, ethicists think that CRISPR still needs a long period of time to be allowed to edit human genome, because society has to gradually adapt to the changes that are brought by CRISPR.

As a summary, CRISPR is a young and very powerful genetic editing tool. Its mechanism allows scientists to edit any DNA sequences in any organisms, which is historical for genetic editing. The breakthrough of CRISPR will possibly lead to a golden age of genetic editing. In the future, CRISPR and genetic editing will be applied to almost every corner of our life.

Acknowledgements

This work was supported by the NRF-2015R1A2A1A15054247 grant to Haihong Shen, the NRF-2016R1A2B1007135 grant to Xuexiu Zheng and Cell Logistics Research center (2016R1A5A1007318) funded by the National Research Foundation (NRF) of Korea, and an integrative aging research grant at the Gwangju Institute of Science and Technology (GIST).

Figure legends

Figure 1. How does CRISPR system correctly cut target DNA? (left panel) The Procedures of producing crRNA from pre-crRNA, binding of cr-RNA:tracrRNA:Cas9 complex to the target sequence near the PAM and the cleavage of target DNA are described. (right panel) The components of figures are described at right panel.

Figure 2. Non-Homologous End Joining (NHEJ). (left panel) procedures of NHEJ is described. (right panel) components of the figure is described.

Figure 3. Homology Directed Repair (HDR). (left panel) procedures of HDR is described.

References

1. Ishino Y, Shinagawa H, Makino K, Amemura M and 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, 5429-5433
2. Barrangou R, Fremaux C, Deveau H et al (2007) Crispr provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709-1712
3. Hwang WY, Fu Y, Reyon D et al (2013) Efficient genome editing in zebrafish using a crispr-cas system. *Nat Biotechnol* 31, 227-229
4. Deveau H, Garneau JE and Moineau S (2010) Crispr/cas system and its role in phage-bacteria interactions. *Annu Rev Microbiol* 64, 475-493
5. Deveau H, Barrangou R, Garneau JE et al (2008) Phage response to crispr-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* 190, 1390-1400
6. Bhaya D, Davison M and Barrangou R (2011) Crispr-cas systems in bacteria and archaea: Versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* 45, 273-297
7. Brouns SJ, Jore MM, Lundgren M et al (2008) Small crispr RNAs guide antiviral defense in prokaryotes. *Science* 321, 960-964
8. Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P and Siksnys V (2013) CrRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol* 10, 841-851
9. Carte J, Christopher RT, Smith JT et al (2014) The three major types of crispr-cas systems function independently in crispr RNA biogenesis in *Streptococcus thermophilus*. *Mol Microbiol* 93, 98-112
10. Deltcheva E, Chylinski K, Sharma CM et al (2011) Crispr RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607
11. Gasiunas G, Barrangou R, Horvath P and Siksnys V (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* 109, E2579-2586
12. Swarts DC, Mosterd C, van Passel MW and Brouns SJ (2012) Crispr interference directs strand specific spacer acquisition. *PLoS ONE* 7, e35888
13. Cencic R, Miura H, Malina A et al (2014) Protospacer adjacent motif (PAM)-distal sequences engage crispr Cas9 DNA target cleavage. *PLoS ONE* 9, e109213
14. Jiang W, Bikard D, Cox D, Zhang F and Marraffini LA (2013) RNA-guided editing of bacterial genomes using crispr-cas systems. *Nat Biotechnol* 31, 233-239
15. Iliakis G, Wang H, Perrault AR et al (2004) Mechanisms of DNA double strand break repair and chromosome aberration formation. *Cytogenet Genome Res* 104, 14-20
16. Gu J and Lieber MR (2008) Mechanistic flexibility as a conserved theme across 3 billion years of nonhomologous DNA end-joining. *Genes Dev* 22, 411-415
17. Aravind L and Koonin EV (2001) Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res* 11, 1365-1374
18. Zhu S and Peng A (2016) Non-homologous end joining repair in *Xenopus* egg extract. *Sci Rep* 6, 27797
19. Riballo E, Woodbine L, Stiff T, Walker SA, Goodarzi AA and Jeggo PA (2009) Xif-cernunnos promotes DNA ligase IV-Xrcc4 re-adenylation following ligation. *Nucleic Acids Res* 37, 482-492
20. Cong L, Ran FA, Cox D et al (2013) Multiplex genome engineering using crispr/cas systems. *Science* 339, 819-823
21. Zu Y, Tong X, Wang Z et al (2013) Talen-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods* 10, 329-331

22. Thompson LH and Schild D (2001) Homologous recombinational repair of DNA ensures mammalian chromosome stability. *Mutat Res* 477, 131-153
23. Lengsfeld BM, Rattray AJ, Bhaskara V, Ghirlando R and Paull TT (2007) Sae2 is an endonuclease that processes hairpin DNA cooperatively with the mre11/rad50/xrs2 complex. *Mol Cell* 28, 638-651
24. Heyer WD, Ehmsen KT and Liu J (2010) Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* 44, 113-139
25. Motamedi MR, Szigety SK and Rosenberg SM (1999) Double-strand-break repair recombination in *escherichia coli*: Physical evidence for a DNA replication mechanism in vivo. *Genes Dev* 13, 2889-2903
26. Szostak JW, Orr-Weaver TL, Rothstein RJ and Stahl FW (1983) The double-strand-break repair model for recombination. *Cell* 33, 25-35
27. Walker FO (2007) Huntington's disease. *Lancet* 369, 218-228
28. Trager U, Andre R, Lahiri N et al (2014) Htt-lowering reverses huntington's disease immune dysfunction caused by nfkappab pathway dysregulation. *Brain* 137, 819-833
29. Lopalco L (2010) Ccr5: From natural resistance to a new anti-hiv strategy. *Viruses* 2, 574-600
30. Kang X, He W, Huang Y et al (2016) Introducing precise genetic modifications into human 3pn embryos by crispr/cas-mediated genome editing. *J Assist Reprod Genet* 33, 581-588

Figure 1.

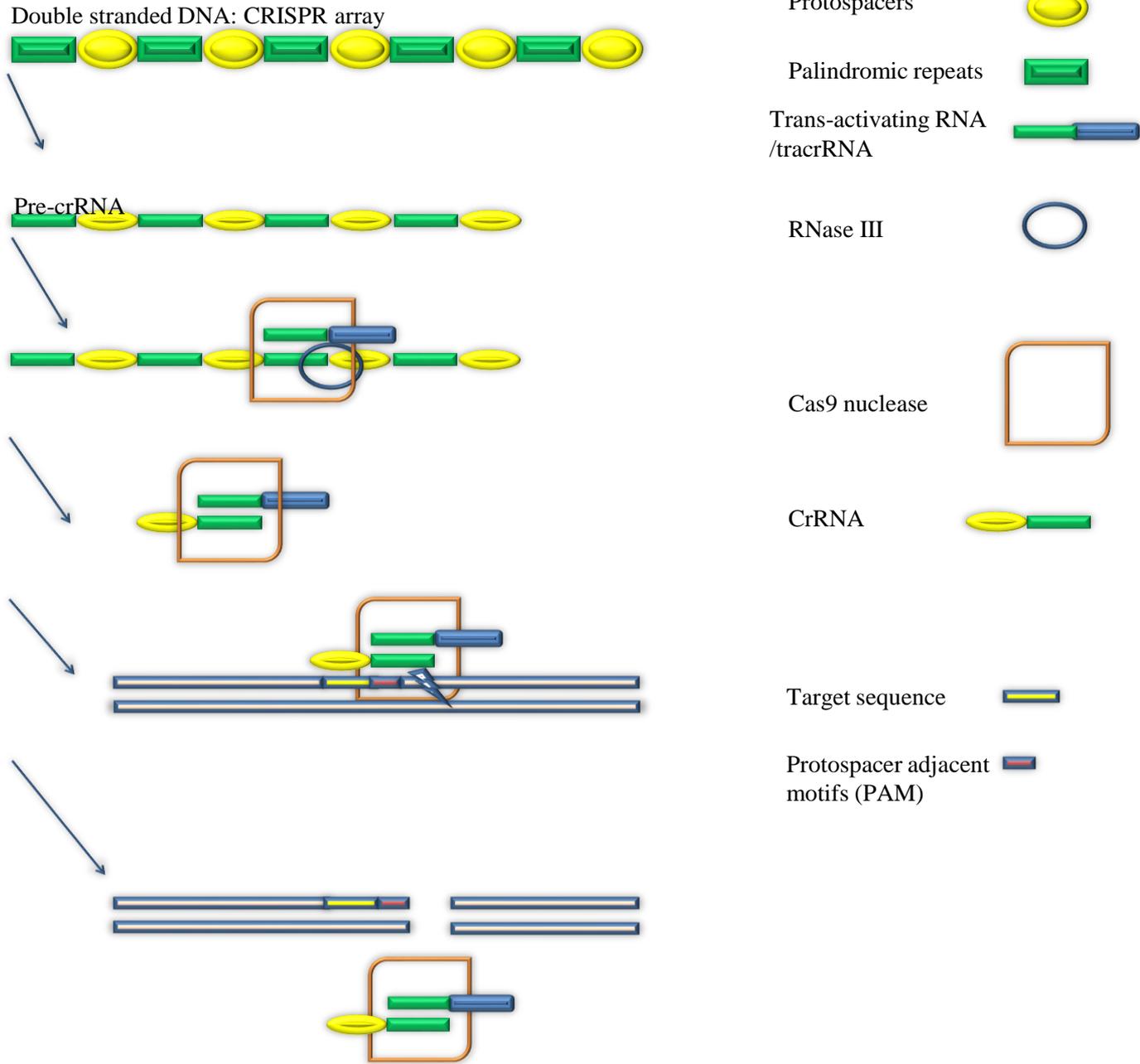
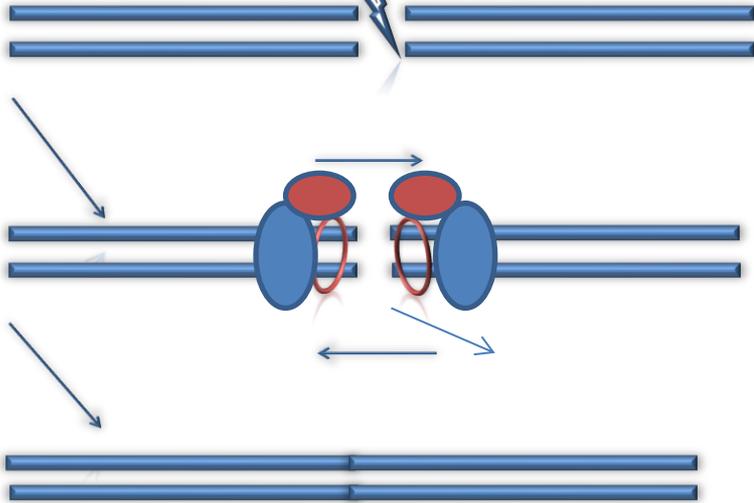


Figure 2.

NHEJ

Double stranded break



Ku proteun



PKcs



XLF:XRCC4 DNA ligase IV



Figure 3.

HDR

