CRISPR/Cas9: The 3G of Gene Editing

# Introduction

**Gene editing** is a technique developed for modifying genetic sequence of an organism by insertions, deletions or base editions. Many diseases, called genetic diseases, are caused by mutations in a single gene. Gene editing technology is a method to control those diseases at genetic level to provide a reliable and permanent cure.

**CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) / Cas (CRISPR Associated Protein) – 9** is the third generation of developments in gene editing technology. The first two generations of gene editing tools include **Zinc Finger Nucleases (ZNFs)** and **transcription activator-like effector nucleases (TALENs**). The CRISPR is derived from a process naturally occurring in bacteria which acts as a defence mechanism from bacteriophages and plasmid infestations, and was first observed in 1987.

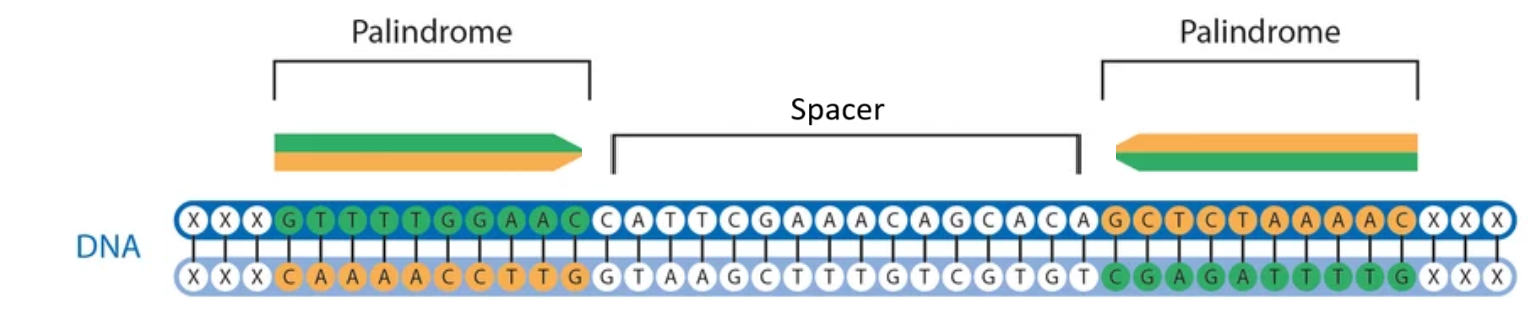
# Discovery of CRIPSR

The CRISPR sequence was first discovered in a bacteria *E. coli*. While it was being studied, an unusual sequence was found in its genome which included 5 homologous palindromic sequence containing 29 nucleotides separated by 32 nucleotides and over next decade, this sequence was detected in many bacteria. This sequence was named **Clustered Regularly Interspaced Short Palindromic Repeats** or the **CRISPR**. In addition, many CRISPR Associated Proteins (Cas) were also discovered.

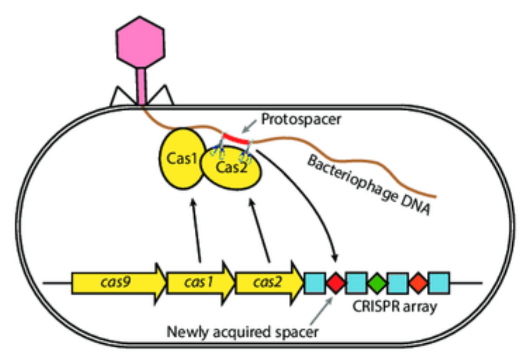
# The Defence Mechanism

## CRISPR Structure

The CRISPR consists of a repeated palindromic sequence of nucleotides, called **Palindromic Repeats** and between those repeats were seemingly random sequences of DNA called **Spacers.** In 2005, it was observed that most of the spacer sequences were derived from exogenous DNA, i.e. form the outside world. And also, that viruses were less likely to infect bacteria having CRISPR sequence. It was conjectured that CRISPR is involved in defence mechanism of the bacteria, which was confirmed 2 years later.

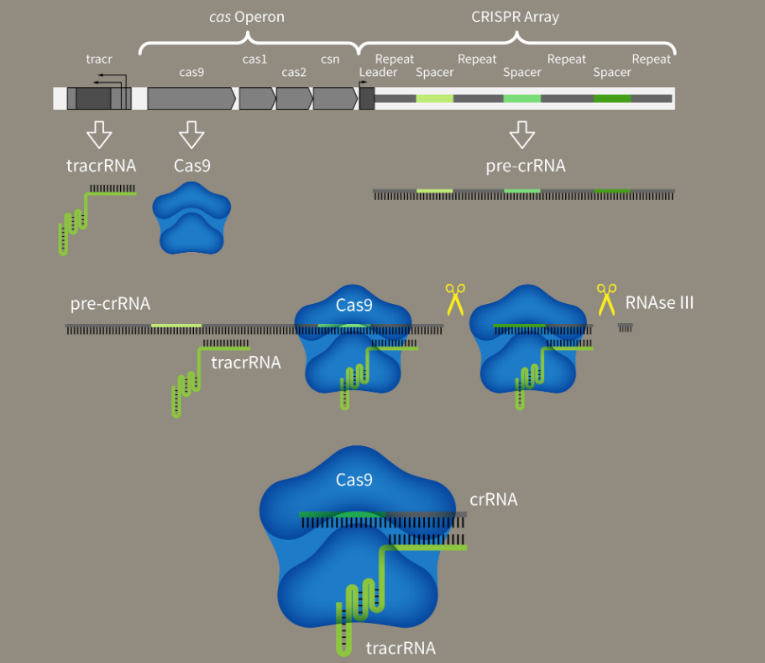


## Memorising

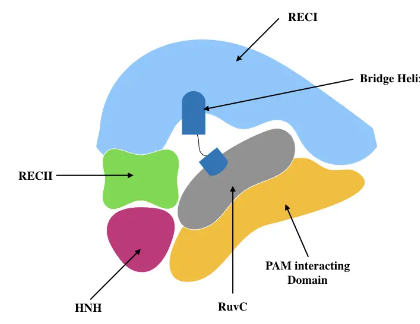
When the bacterium is first confronted with a phage or a plasmid infestation, the enzymes Cas-1 and Cas-2 cut out a region of the foreign genome called the **protospacer**. The cut is made upstream to a **Protospacer Adjacent Motif (PAM)**. The PAM is a sequence consisting 3 bases, often in the form of **NGG,** where N can be any nucleotide and G stands for guanine. The PAM is crucial in recognising the cleavage site for Cas-9.

The protospacer is turned into spacer and inserted into 5’ end of the bacterial CRIPSR array and a new palindromic repeat is generated as well. The CRIPSR array is flexible in size, as the bacterium is infected by more and more viruses, it gains more spacers to keep a memory of the infection.

## Preparation

After transcription of the CRISPR array, the RNA is obtained called **pre-CRISPR RNA (pre-crRNA),** which contains transcribed states of the repeats and the spacer. Another RNA called unprocessed tracer RNA, which contains regions complementary to the pre-crRNA bind with it and forms double strands. The RNAase III enzyme then cuts through this RNA complex to give a piece of RNA containing fragments of pre-cr and unprocessed tracer RNA. This fragment is now called **cr:tracrRNA.** This RNA molecule is then combined to the protein **Cas-9** and is called **guide RNA (gRNA)**.

Cas-9 protein is made of a single polypeptide chain but it has different parts, or domains, in it.

* **PAM Interacting Domain:** Recognises the PAM sequence (NGG)
* **Nuclease Domains (HNH and RuvC):** Cuts the DNA
* Other domains to hold the gRNA and attach to the target DNA.

The tracrRNA holds and attaches the crRNA to the Cas-9 protein.

## Counterattack

When infected again with the same virus, the Cas-9 protein recognises the PAM sequence and then tries to match the gRNA sequence on the target genome. If the sequence matches, a cut is made on both the strands of the DNA. Since the spacer on CRISPR doesn’t precedes PAM sequence, Cas-9 doesn’t cut through the bacterial DNA.

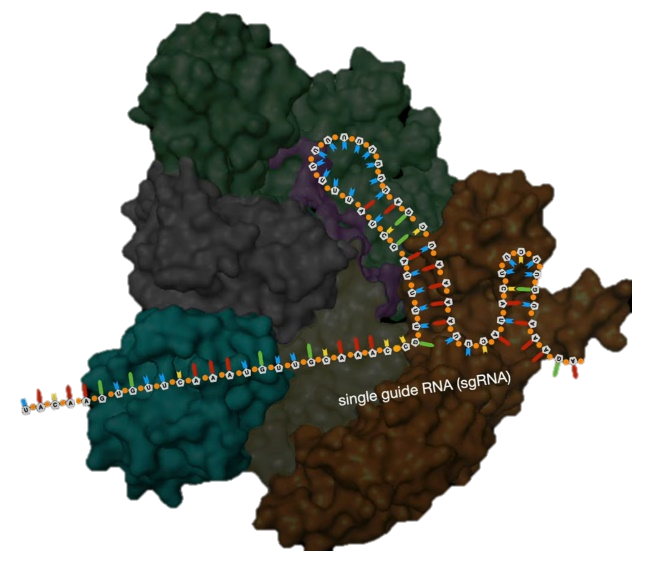
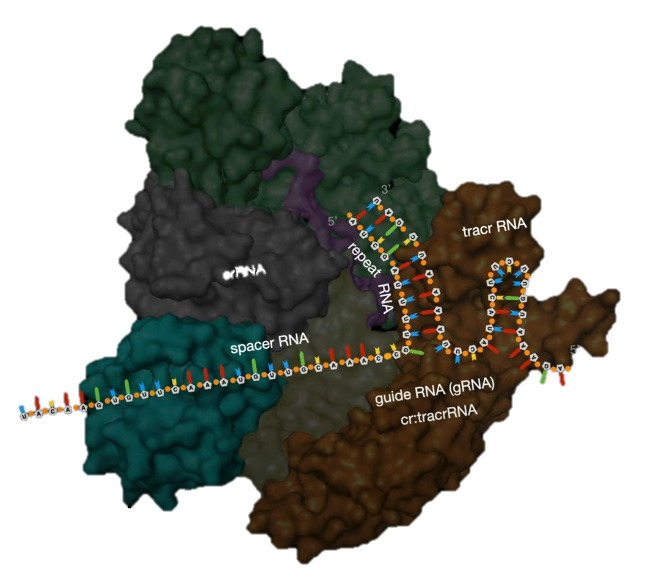
# Using CRISPR/Cas9 Mechanism for Gene Editing

## Transferring the System in Cell

Scientists constructed crRNA-tracrRNA fusion transcripts to form **single guide RNA (sgRNA)**, which essentially is a single polyribonucleotide chain comprising of both crRNA and tracrRNA parts constructed synthetically. The sgRNA is designed to make the cut at the precise location within the DNA. The CRISPR system is delivered into the cell using a vector. Plasmid DNA (pDNA) is an ideal vector for loading the CRISPR system, but the process is very tedious. Loading the system on mRNA greatly simplifies the process but mRNA is easily degraded and has low stability. Continuous innovations are being made in delivery vectors and exosomes are identified to be a promising approach to deliver Cas9 Ribonucleoproteins.

Guide RNA

Once the system enters the cell, it makes a cut at the desired location, 4 base pairs upstream to the PAM.

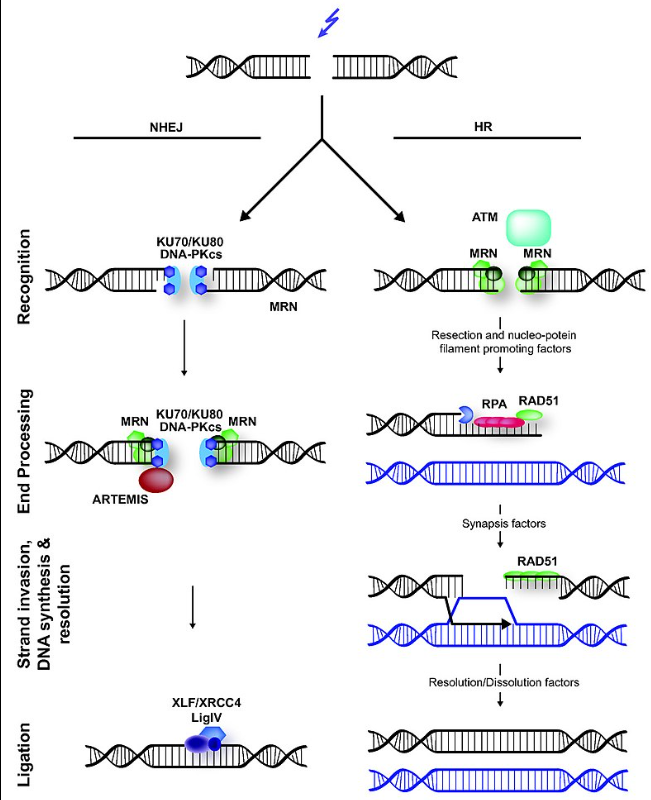


## Bringing out the Changes

Single Guide RNA

After the cut has been made, the cell uses built in repair techniques to join the DNA strands back. There are two methods for this.

* **Non-Homologous End Joining (NHEJ)**

****When a DNA strand is damaged or broken, a complex mechanism consisting of four different proteins ligases the DNA strands. But this ligation is not always perfect. It often adds or deletes bases causing mutation of the said gene. Hence, this method is very effective in silencing a particular gene. Causing mutations at promoter sequence is effective in gene silencing.

* **Homology-directed Repair (HR)**

HDR mechanism involves repairing the broken DNA strand using a template DNA. This mechanism has low chances of mutations as it requires a second copy of the DNA to be template. It uses proteins, DNA ligase and DNA polymerase to generate the exact copy from the template and repairs the cut DNA.

This method is used for adding foreign DNA fragments in the host genome. Along with loading the CRISPR system, we also deliver a suitable DNA fragment containing foreign DNA and the cell copies the foreign fragment in the host DNA while trying to repair cut DNA.

## Deactivated Cas9 (dCas9)

dCas9 protein was prepared by mutating the nuclease domains within the wild type Cas9 protein, which causes it to lose its cleavage activity. dCas9 bind to the target DNA but doesn’t cut it. This also causes gene silencing due to stearic blockage of RNA polymerase elongation.

dCas9 is also used in single base gene editing techniques. dCas9 when loaded with enzymes that can transform nitrogenous bases can be used to treat single point base mutations.

# Limitations and Challenges

## Limitations of CRISPR/Cas9

* **Off Target Effects:** Base mismatches between sgRNA and nontarget sequences may lead to off-target effects. The introduction of one or even multiple unknown mutations can lead to serious complications.
* **Effectivity:** Silencing/Mutating a gene does not necessarily have a therapeutic effect. The transcriptional activation efficiency of sgRNAs at each position is not the same but is strongly linked to the cell and gene.
* **Applicability:** CRISPR/Cas9 method is limited to PAM sequences which prevents Cas9 from reaching certain positions, specially while using base editing tools as the edited bases are located at specific relative position to PAM sites. The researchers have worked to modify Cas9 so that it is not restricted to recognising NGG by mutating Cas9 site or adding modified structural domains.
* **Chromosomal Disorientation:** When Cas9 cleaves double-stranded DNA, it activates NHEJ repair, usually resulting in the insertion or deletion of a small number of bases. This is anticipated, but sometimes large-scale base deletions and chromosomal structural translocations can happen, which causes significant complications.

## Limitations of Targeted Delivery

* **Deviation from the desired position:** Unmodified vectors can be captured by the body's metabolic organs. The CRISPR/Cas system retains its activity in nontarget cells, and can cause genetic modifications in healthy cells leading to unpredictable outcomes. To minimize the entry of gene drugs into nontarget cells, improved delivery vehicles are needed.
* **Compatibility:** The immune response resulting from delivery of the material into the body must also be considered when designing the system. Commonly used Cas9 proteins derived from *S. pyogenes* and *S. aureus* have been reported to trigger an immune response in humans