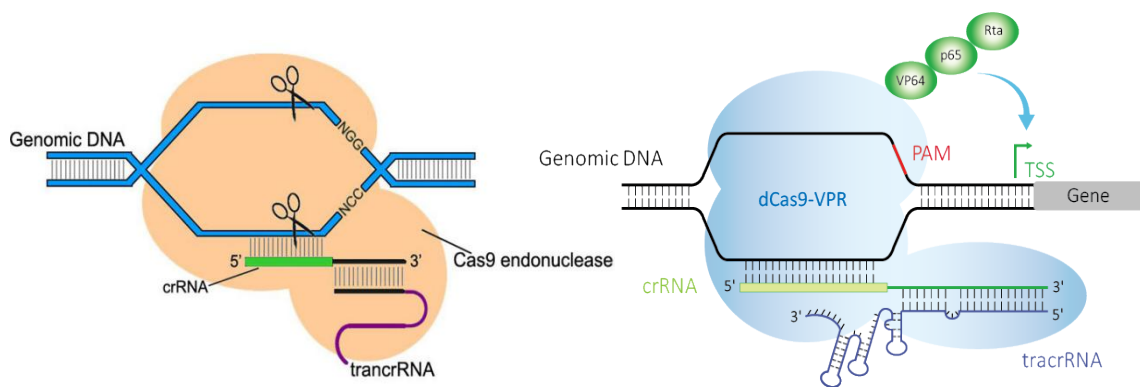


## DNA cleaving and gene repression using CRISPR/Cas



## Manual

Tianhe Wang, Yalan Yi, May, 2018

# Contents

1 Introduction	1
2 Theory	1
2.1 CRISPR system and Cas9 endonuclease.....	1
2.2 sg RNA and designing tools.....	3
2.5 Cell-free expression.....	4
2.4 Fluorescence and Microplate reader.....	5
2.4.1 Fluorescence and Jablonski diagram.....	5
2.4.2 Fluorophores and fluorescent proteins.....	6
2.4.3 Optical filters.....	6
2.4.4 Microplate Reader.....	7
2.5 Agarose electrophoresis.....	8
3 Experiment	8
3.1 Repression of gene expression <i>in vitro</i> .....	8
3.1.1 Goal of the experiment.....	8
3.1.2 Materials and experimental protocol.....	9
3.2 Guide RNA design.....	10
3.3.1 Goal of the experiment and procedures used.....	11
3.3 <i>In vitro</i> DNA digestion.....	12
3.3.1 Experiment purpose and procedures.....	12
3.3.2 Experiment protocol.....	13
4 Quiz	14
5 Reference	
6 Appendix(Optional)	

# 1 Introduction

Nowadays, targeted genome editing has become a very promising technique for efficiently modifying any gene of interest in a broad range of research and medical applications<sup>[1]</sup>. In the recent decade, programmable sequence-specific DNA nucleases have been developed to facilitate the precise editing of gene targeting or endogenous genomic loci in a wide range of prokaryotes and eukaryotes<sup>[2]</sup>.

A large number of genome editing technologies based on these DNA nucleases have been exploited in recent years, including zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the RNA-guided Cas nuclease from type II prokaryotic Cluster Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated system<sup>[3]</sup>. Unlike the previous two DNA nucleases based on sequence-specific DNA binding domain, Cas9 nuclease utilizes a guide RNA to target a specific DNA sequence and further induce DNA double-stranded breaks (DSBs) through Watson-Crick base pairing with target DNA. As RNA-guided nuclease, CRISPR-Cas9 can be engineered more easily and with less time consumption than ZFNs and TALENs, and it also is highly efficient for high-throughput multiplex genome editing in a variety of organisms. Besides, sgRNAs contain an approximately 20-base sequence specific to target genomic loci, which provides a larger sequence space that can be targeted. In addition, except for gene knockout and insertion, nuclease-deficient Cas9 (dead Cas9) has been applied into direct knockdown of the target gene transcription process without genetically altering the DNA sequence.

In this Praktikum, our learning goal includes the mechanism of CRISPR-Cas system, designing of sgRNA and experimental procedures in terms of Cas9-mediated manipulation for gene targeting and repressing. We are going to practice CRISPR-Cas system-mediated *in vitro* gene targeting and gene repression in solution and cell-free system (in vitro transcription and translation system). And we will also practice how to design sgRNAs based on our desired genomic loci by using NUPACK and CRISPRdirect.

## 2 Theory

### 2.1 CRISPR/Cas system and Cas9 endonuclease

CRISPR loci on the bacterial genome are composed of leader sequence, repeat sequences and spacer (Fig.1, a). It has been demonstrated that the CRISPR/Cas system works as a prokaryotic immune system and plays an important role in protecting microbes from the invasion by exogenous DNAs (such as virus genomes and plasmids) via transduction, conjugation, and transformation. The immune process of the CRISPR/Cas system functions through two general stages which are i) the immunization or adaptation stage and ii) the immunity or interference stage (Fig.1, b).

In the first stage, the foreign DNA fragment which contains a protospacer sequence is recognized and integrated into CRISPR loci and becomes a new spacer on the CRISPR array after the leader sequence. Each spacer is subsequently transcribed from the leader region, yielding a long precursor CRISPR-RNA (crRNA) that is further processed into mature crRNAs by Cas proteins. Secondly, Cas proteins are bound to crRNAs and form Cas/crRNA complexes that can recognize invading DNAs via base-pairing between the protospacer and crRNA and further inactivate their targets. Furthermore, in type I and type II CRISPR systems, targeting and cleavage of invading DNAs require the participation of a protospacer adjacent motif (PAM) for successful interference.

By introducing single mutations on RuvC (10<sup>Spy</sup>) and HNH (841<sup>Spy</sup>) domains, dCas9 has been developed for non-mutagenic gene regulation and chromosome imaging in the prokaryotic and eukaryotic cell (Fig.3)<sup>[4]</sup>. These two single mutations on nuclease domains result in nuclease-deficiency of Cas9, which can no longer cleave target DNA but remains the capability to specifically bind to DNA when it is guided by a crRNA/tracrRNA complex (tracrRNA = trans-activating CRISPR-RNA) or a single guide RNA (sgRNA). In mammalian cells, specific binding to downstream of the promoter region will sterically block transcription elongation of RNA polymerase II, which is known as CRISPR interference. Moreover, fusing dCas9 with Red fluorescence protein (RFP) and transcriptional factors (VP64 or p65) can be used for live gene imaging on the chromosome and transcriptional activation.

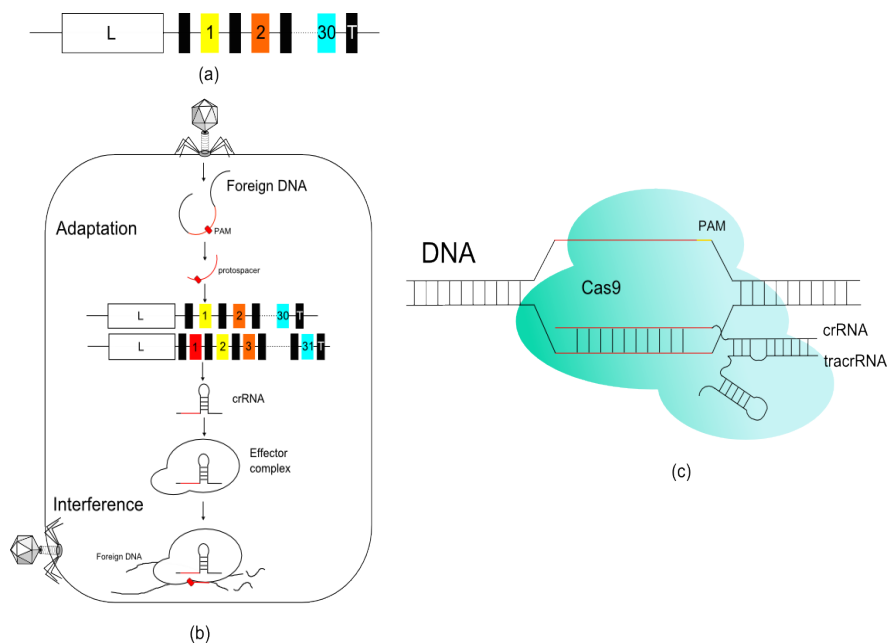


Figure.1 **a)**, Composition of CRISPR loci. The white rectangle indicates the leader which is present upstream of the CRISPR loci. Black rectangles indicate the repeat sequences. The colored boxes highlight the different spacer sequences of similar length found in a CRISPR locus. **b)**, Adaptation and interference stages of the CRISPR/Cas system. **c)**, Cas9/guide RNA complex binds to complementary target DNA sequence. PAM sequence is also bound by Cas9, which is necessary for cleaving activity.

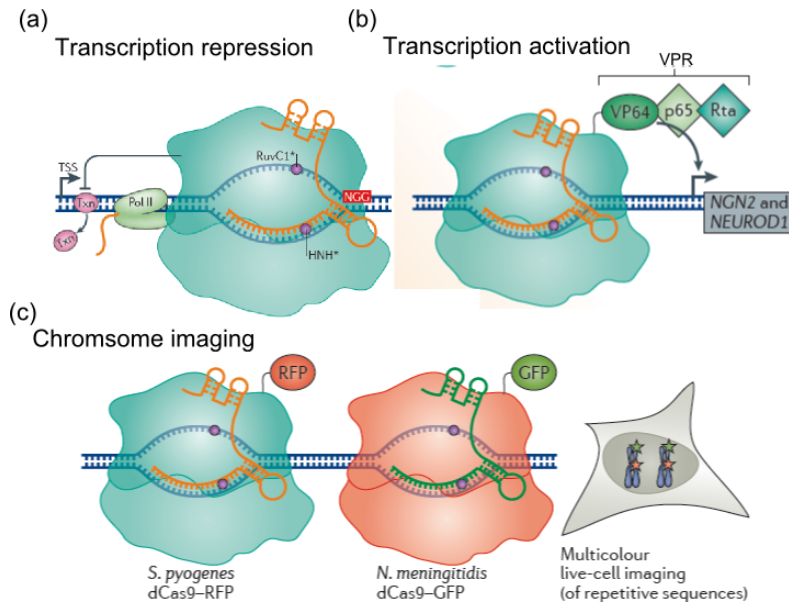


Figure.2 Overview of precision genome regulation and imaging using the dCas9 protein. **a)**, dCas9-mediated transcription repression in eukaryotic cells, transcription process is inhibited by blocking RNA polymerase elongation. **b)**, dCas9-mediated transcription activation in mammalian cells, transcription initiation results from fused VP64 domain on dCas9 C or N terminal. **c)**, dCas9-mediated chromosome imaging, target genes are labeled on chromosome by fused fluorescence protein on dCas9 C or N terminal.

## 2.2 sg RNA and design

Cas9 proteins rely on the guidance of crRNA/tracrRNA duplex to target the desired DNA sequence, thus, this duplex has been replaced by a synthetic sgRNA chimera that packages these two RNAs together and mimics the crRNA/tracrRNA duplex in order to simplify the RNA-guided targeting process. This sgRNA can be ligated together with the Cas9 gene and brought into one single plasmid for further transformation or transfection (the process of deliberately introducing exogenous nucleic acids into prokaryotes and eukaryotic cells). The chimeric sgRNA sequence is composed of a 20-25 nt base-pairing region which is complementary to the spacer sequence, a 42 nt double strand handle which mimics the crRNA/tracrRNA complex, a linker sequence (Nexus) and a 40 nt terminator sequence (Fig. 4)<sup>[5]</sup>. In addition, the base-pairing region consists of the 12 nt seed region which is adjacent to the double stranded handle on the sgRNA and the PAM on target DNA. It has been proven that mismatches between the seed region and its complementary spacer sequence decrease target DNA binding affinity, which is important for effective gene regulation. For sgRNA design and base-pairing sequence selection, there are plenty of available online design tools. In this lab course, we introduce NUPACK (<https://www.nupack.org>) and CRISPRdirect (<http://crispr.dbcls.jp/>) for screening sgRNA secondary structure and target sites on the genome, respectively. NUPACK is a free software suite for the analysis and design of nucleic acid structures. For the interested reader, further information on NUPACK and the thermodynamics of nucleic acid secondary structure can be found in the Appendix.

CRISPRdirect is a web server which provides functions for screening rational CRISPR/Cas targets from your input sequence, which provides an efficient selection of CRISPR/Cas target sites on many sequenced eukaryotic and also prokaryotic genomes with low numbers of potential off-target candidates (unintended mutations). This software helps to investigate the entire genome for perfect matches with each candidate target sequence including their seed sequence and the PAM. It will also provide a detailed list of potential off-target sites that are partially complementary to the target DNA sequence. The operating procedures in detail will be introduced in chapter 3, experiment procedure.

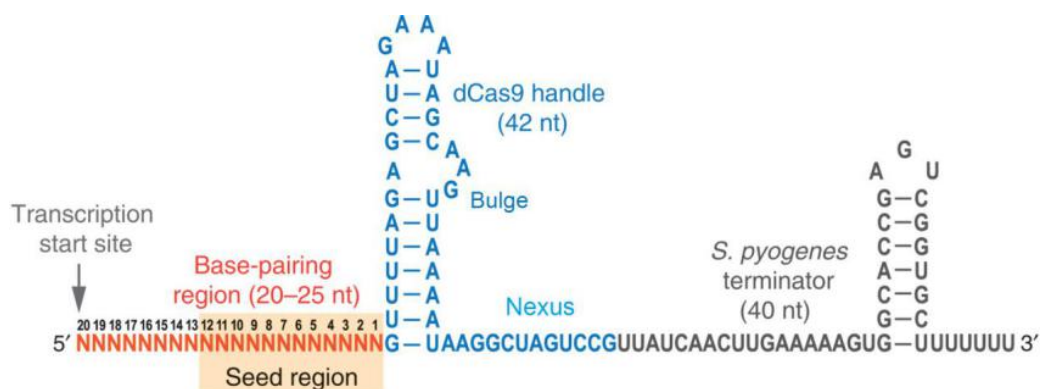


Figure.3 Overview of the structure of each module of a sgRNA of the SpyCas9. The sgRNA includes three regions: a 20-25 nt-long base-pairing region for specific DNA binding, a 42-nt-long dCas9 handle hairpin for Cas9 protein binding and a 40-nt-long transcription terminator hairpin derived from *S.pyogenes*. The 12-nt seed region is shaded in orange.

### 2.3 Cell-free expression system

Cell-free expression systems are a very useful *in vitro* tools for one-pot transcription and translation of target DNA templates into RNA and proteins, which are widely used to synthesize large amounts of recombinant proteins. *E.coli*-based cell-free systems normally include crude cytoplasmic bacterial extract that contains basic transcriptional and translational components such as RNA polymerases, ribosomes and cofactors, as well as a reaction buffer providing chemical energy, e.g. ATP, GTP for biosynthesis processes, and an amino acids mix.

There are also several commercial cell-free reaction kits available such as expressway® and purexpress®, which have several advantages for cell-free production of recombinant proteins in the laboratory. Proteins can be produced within 4 to 5 hours, they are easily detected and purified after the reaction and are then ready for use in downstream applications. Also, the cell-free system has a better tolerance towards toxic compounds or protein products than living cells. Moreover, the reaction components and conditions such as temperature, pH value can be better manipulated and controlled. For the lab course we are going to use a home-made cell extract to test repression of gene expression by dCas9 guided by sgRNA, which can be read out on a microplate reader (see next section) via the production of a fluorescent protein within 4-5 hours.

## 2.4 Fluorescence and Microplate reader

### 2.4.1 Fluorescence and Jablonski diagram

Fluorescence is the emission of light of a substance which is activated by absorbed light or other electromagnetic radiation. The absorption process first causes an electron to transition from the ground state ( $S_0$ ) to the excited singlet state ( $S_1$ ), where it non-radiatively relaxes to its lowest vibrational energy level. The following relaxation to the ground state ( $S_1 \rightarrow S_0$  transition) is therefore accompanied by the emission of light with a longer wavelength than the absorbed light (Stokes-Shift,  $\lambda_{2\text{emission}} > \lambda_{1\text{absorbance}}$ ).

A Jablonski diagram can be used to illustrate the electronic energy states of a molecule and the transitions between them (Fig.4a). Next to fluorescence, other radiative transitions with longer lifetimes can occur (phosphorescence) and also non-radiative transitions (including vibrational relaxation, internal conversion and intersystem crossing). The average time for a fluorophore spends in the excited state before returning to the ground state by emitting a photon is called the fluorescence lifetime  $\tau$ . It is defined as  $\tau = (k_f + k_{nr})^{-1}$  and thus depends on the intrinsic transition rate  $k_f$  characteristic of the given molecule and an environmental contribution  $k_{nr}$ . As a function of time, the fluorescence decay of a sample, which is excited at  $t = 0$  is given by  $I(t) = I(0)\exp(-t/\tau)$ . Typical lifetimes of excited states/fluorescence lifetimes are on the order of  $10^{-7} \sim 10^{-9}$ s.

Another important quantity for the characterization of fluorescence processes is the quantum yield (QY), which is the ratio between the number emitted photons and the number of absorbed photons. The presence of non-radiative de-excitation processes generally decrease the QY. "Good" fluorophores can have QYs above 90%.

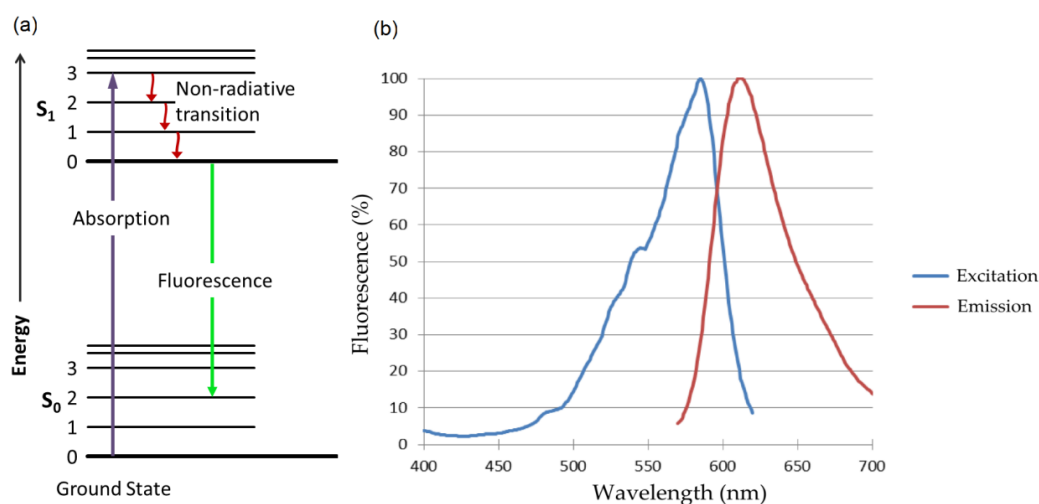


Figure. 4 Jablonski diagram and RFP fluorescence spectrum. **a)** The Jablonski diagram schematically shows vibrational energy levels, absorption and non-radiative transitions, and fluorescence. **b)** RFP excitation and emission spectrum with maximum EX/EM wavelengths at 584nm and 601nm, respectively.

Fluorescence spectra can be measured by either exciting the sample at a single wavelength and collecting the fluorescence emission (emission spectra), or by measuring the fluorescence at a fixed wavelength and varying the excitation (excitation spectra). For instance, the maximum excitation and emission wavelength of red fluorescent protein (RFP) are 584nm and 601nm respectively (Fig.4b), corresponding to the emission light having a lower energy than the absorbed radiation.

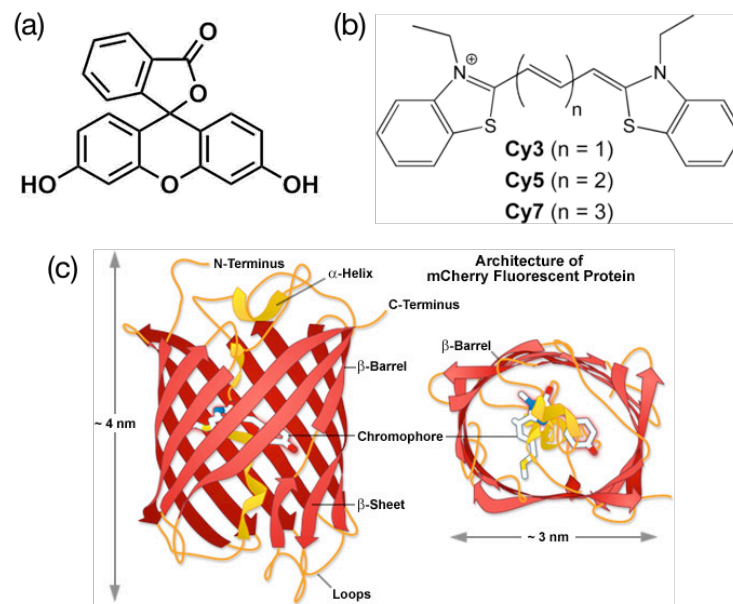


Figure. 5 Structure of typical fluorophores and RFP. (a) Fluorescein, (b) Cyanine dyes, (c) Red fluorescent protein (here mCherry). The chromophore is formed by chemical reaction between Tyr67 and either Met, Gln, Thr, Cys, or Glu at position 66.

## 2.4.2 Fluorophores and fluorescent proteins

Figure 5 shows the molecular structures of two typical fluorophores (fluorescein and Cy3) and also the structure of RFP. In fluorescent proteins, the fluorophore is formed after folding of the  $\beta$  barrel protein by a chemical reaction between several amino acid residues pointing into the interior of the barrel. Depending on the efficiency of this process, fluorescent proteins become fluorescent with a delay after translation (the maturation time  $\tau_{mat}$ ). Fluorescein has a QY of 95% and Cy dyes have QYs in the range 15 – 30%. RFP has a QY of 35% and a maturation half-time of about 20 minutes (these values differ widely among the FPs).

## 2.4.3 Optical filters

Optical filters selectively transmit or reject a wavelength or a range of wavelengths (Fig. 5). They find applications in optical instruments such as fluorescence microscopes and spectrophotometers. There are several different types of optical filters for different applications. For example, interference filters are made by coating a glass substrate with a series of optical coatings, in which interfaces between the



layers of different refractive index produce phased reflections, selectively enhancing certain wavelengths of light and reflecting with other wavelengths. In addition, bandpass filters only transmit a certain wavelength band, while rejecting other wavelengths. The center wavelength and bandwidth are two important characteristics, the former one can be chosen by the user, the latter can be calculated at 50% of peak transmission.

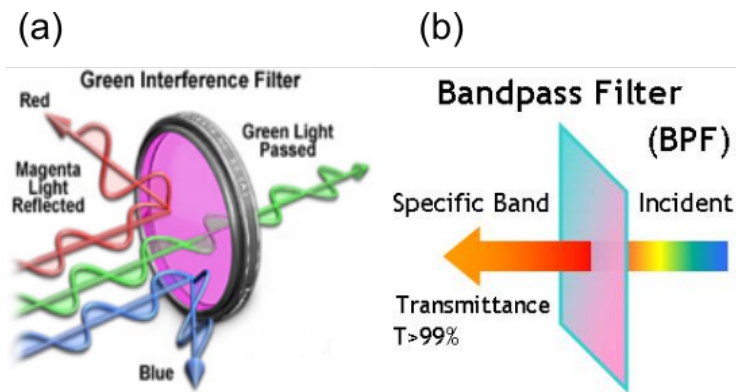


Figure. 6 Operating principle of Optical filters. a), Interference filter and b), Bandpass filter.

#### 2.4.4 Microplate reader

Microplate readers also known as Microplate spectrophotometers, can detect biological, chemical reactions including absorbance, fluorescence and luminescence intensity. It allows researchers to analyze multiple samples in parallel as compared to single sample measurements. Filter-based multimode microplate readers are the most common type of microplate readers (Fig. 6). They usually contain excitation and emission filter wheels which have several sets of optical filters for different wavelengths. The samples are excited by light of the desired wavelength. This emitted fluorescence light is passed through an emission filter that transmits it to a photomultiplier detector, which converts it into a readable output signal. In this lab course we will use the microplate reader to detect RFP fluorescence signals during our dCas9 gene repression experiment.

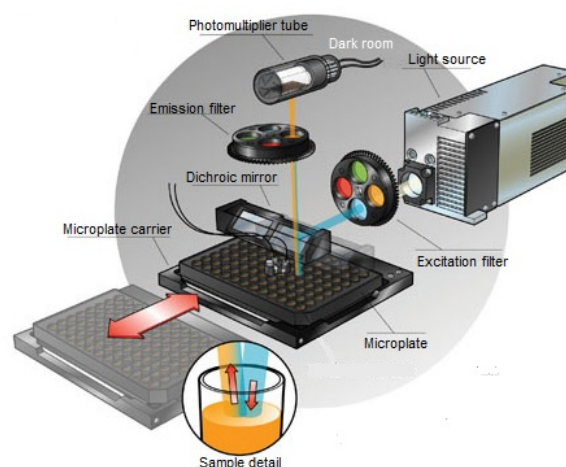


Figure.6 Overview of filter-based multimode microplate reader with its individual components.

## 2.5 Agarose gel electrophoresis

Agarose gel electrophoresis is a common method for separating a mixed population of macromolecules such as DNA, RNA and proteins, which has been widely used in biochemistry, molecular biology, genetics research. The agarose gel is a common type of stabilizing medium which serves as a molecular sieve with channels and pores through which the sample molecules travel. More highly concentrated gel networks have internal pores and channels of smaller size, which typically provides a better resolution for the separation process. The phosphate backbone of DNA and RNA is negatively charged at neutral pH. When they are placed in an electric field, nucleic acid fragments therefore migrate towards the positive electrode in a process called electrophoresis (Fig. 7a). Shorter DNA fragments as well as supercoiled DNA molecules migrate through the pores in the agarose matrix more easily compared to higher molecular weight and linear DNA fragments and therefore have a correspondingly higher mobility (Fig. 7b). In our experiment, we will use agarose gel electrophoresis to check the cleavage of a DNA plasmid by Cas9-sgRNA *in vitro*.

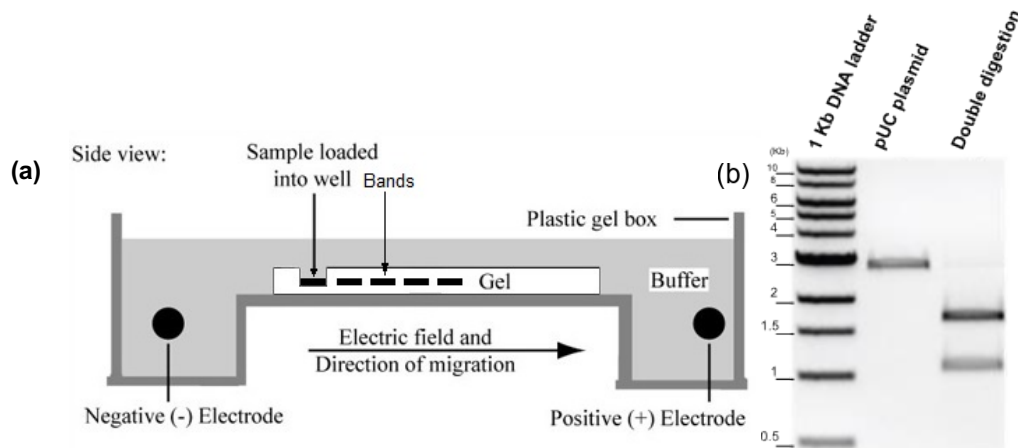


Figure.7 Agarose gel electrophoresis. **a)** Schematic representation, **b)** Agarose gel showing the product of double digestion of pUC plasmid.

## 3 Experiment

### 3.1 Repression of gene expression *in vitro*

#### 3.1.1 Goal of the experiment

In the first experiment, we will express the dCas9 protein and a reporter protein, i.e. red fluorescent protein (RFP), from two different plasmids (details in appendix) in the cell extract to observe the gene repression effect of dCas9 protein upon addition of a specific sgRNA (Fig. 7). dCas9 protein is expressed from a T7 promoter, while RFP is expressed from a constitutive promoter. SgRNA complementary to the RFP coding sequence is added to the cell free mix containing plasmids encoding dCas9 and RFP along with T7 RNA polymerase. Repression is observed when the complex of dCas9

and sgRNA binds to the coding region of RFP on the plasmid.

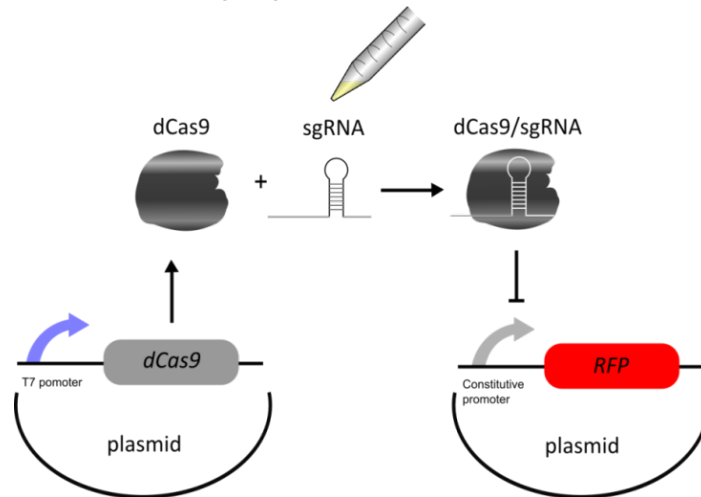


Figure 8. Overview of the gene repression by dCas9/sgRNA complex.

### 3.1.2 Materials and experimental protocol

#### Materials:

DNA templates (plasmids, details in appendix).

sgRNA: RNA sequence and length will be explained in experiment 2.

Transcription-translation cell-free expression system (TX-TL) consisting of:

Crude cell-extract from *E. coli* Rosetta™2(DE3),

Reaction buffer (amino acid mix and energy supply)

T7 RNA polymerase: (New England Biolab®, Item No: M02512)

384-well plate (IBIDI)

#### Instrument:

Fluostar® microplate reader.

#### Protocol:

##### 1. Prepare cell extract mix

(1) Take one aliquot cell extract and one aliquot reaction buffer from -80°C freezer and put them on ice to completely melt.

(2) Mix 28.5 µL cell extract solution and 35.7 µL reaction buffer in an Eppendorf tube and keep it on ice.

##### 2. Mix DNA template with cell extract

(1) Pipette 5 nM (final concentration) of DNA plasmids for dCas9 gene and RFP respectively into a 0.2 mL small PCR tube, including positive control (without sgRNA). Add same amount of nuclease free water instead of RFP plasmid into negative control group. The volume of each DNA template depends on their concentration and can be calculated by the formula in appendix.

(2) Pipette 5nM (final concentration) of purified sgRNA into the experiment group with

DNA template. Add the same amount nuclease free water (nF H<sub>2</sub>O) into positive control (Appendix table.1). The volume of each sgRNA depends on their concentration can be calculated by the formula in appendix.

- (3) Pipette 0.6 µL T7 RNA polymerase (50 U/µl) into the same tube.
- (4) Pipette 11.25 µL cell extract mix into each tube and mix the solution.
- (5) Pipette additional nF H<sub>2</sub>O into each tube making the final volume to 15 µL.
- (6) Add each of the reaction (3 in total) to 384-wells plate.
- (7) Put the plate into Fluostar microplate reader and follow the measurement protocol (ask supervisor which protocol to use) (Fig.9)
- (8) Run the protocol to measure the repression of RFP and compare it to controls. Excitation and emission filters are 584 nm and 640 nm respectively. The fluorescence is measured every 5 minutes for 4 hours.

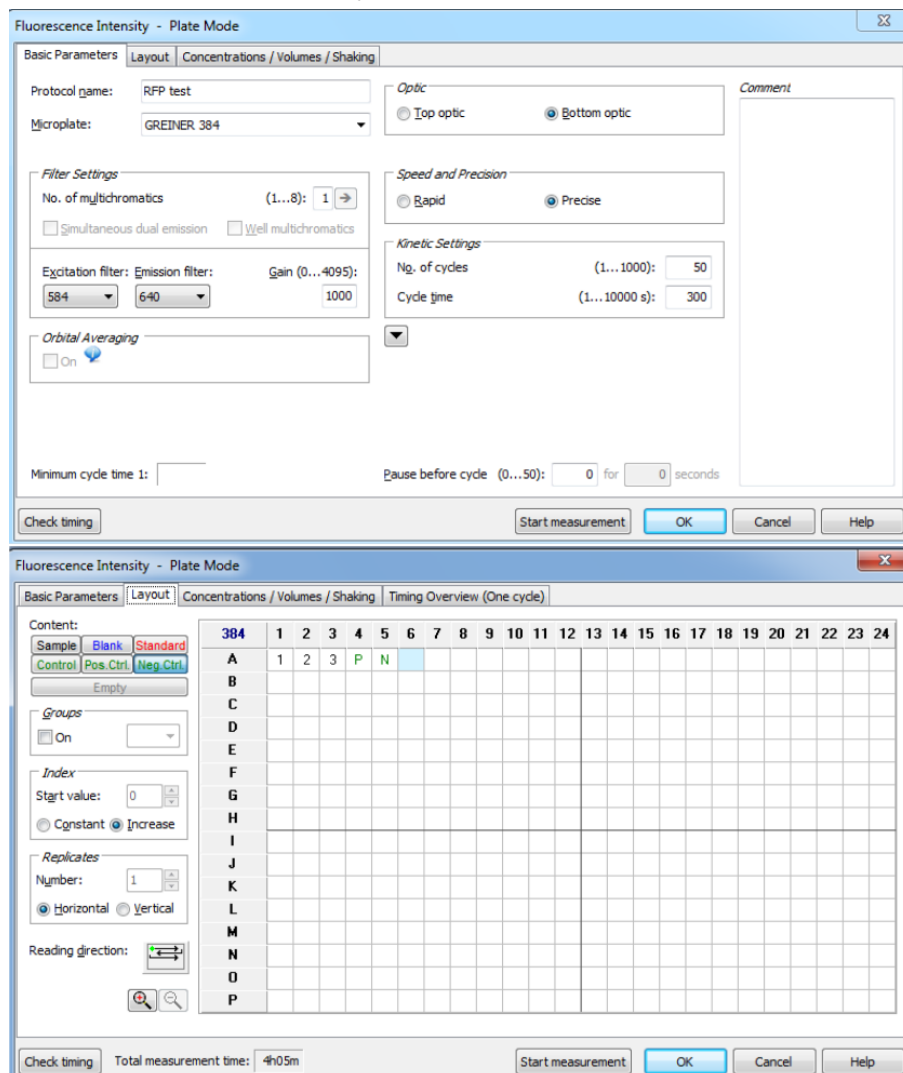


Figure. 9 RFP measurement protocol dialog box in Fluostar.

### 3.2 guide RNA design

### 3.2.1 Goal of the experiment and procedures used

In this section, we will study how the sgRNA used in the previous experiment for targeting the coding sequence of *rfp* gene is designed. The optimal sgRNA should have the lowest free energy value and lowest amount of secondary structures. RFP coding sequence was selected from [http://parts.igem.org/Part:BBa\\_E1010](http://parts.igem.org/Part:BBa_E1010). It consists of 705 nt coding 235 aa. A 20-25 nt base-pairing region is selected from the coding sequence, including a seed region and the PAM. The conserved sequence of sgRNA of the SpyCas9 is kept as such (blue) and grey regions in Fig. 4 and the base-pairing region (red) are replaced by the *rfp* gene. Simulate several designed sgRNA candidates for the *rfp* gene in NUPACK and check the secondary structure of each sgRNA, using the following steps:

- (1) Go to NUPACK (<http://www.nupack.org/>) by clicking “Analysis”.
- (2) Select each parameter for your desired sequence, including the type of nucleic acid, reaction temperature and the number of strands.
- (3) Copy-paste your sequence in the strand form and click “Analyze”
- (4) The results page shows the MFE structure at 37°C along with the total free energy and pair probability.
- (5) Choose the sgRNA sequence with the lowest free energy and lowest amount of secondary structures.

In addition, we will also use CRISPRdirect to screen the potential base-pairing sequence for our desired sgRNA for *cd4* gene (important for T lymphocytes immune response) on the human genome.

- (1) Search and copy the human *cd4* gene sequence on NCBI website (<https://www.ncbi.nlm.nih.gov/gene/>).
- (2) There are three options for uploading your desired sequence onto the CRISPRdirect input page. Type in the genome location of your desired gene and the entire sequence text, or upload your sequence file in FASTA format.
- (3) Click the design button and find all available base-pairing sequences, sgRNA sequences and off-target sites on the result page.
- (4) Download the data from CRISPRdirect and analyze secondary structure and total free energy value for each sgRNA candidate.

Similarly, the optimal sgRNA for *cd4* gene is supposed to have the lowest free energy value and less off-target sites on the genome compared to other sgRNA candidates.

### 3.3 *In vitro* DNA digestion

### 3.3.1 Experiment purpose and procedures

In the third experiment, we are going to test the site-specific DNA cleavage activity of the Cas9 protein *in vitro* (Fig 10). Two different sgRNAs which can bind to the Cas9 protein and target the 5' and 3' end of our desired gene (*rfp*) on the plasmid DNA are designed, transcribed and purified from their respective DNA templates. Cas9/sgRNA duplexes then bind to their complementary sequence and create two double-strand breaks. Our desired fragment will finally be cut off from the plasmid DNA, which can be easily tested using agarose gel electrophoresis.

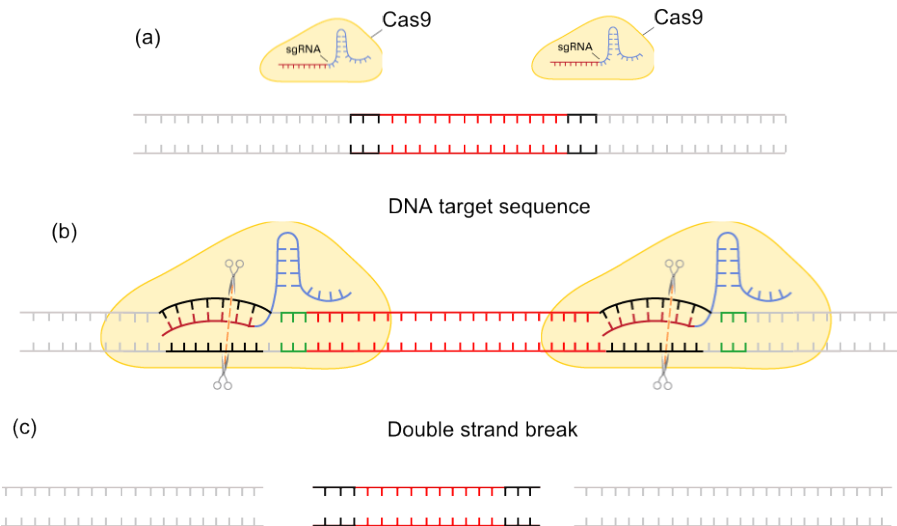


Figure.10 Cas9/sgRNA duplex targeting and cleavage process. (a), sgRNAs bind to Cas9 protein brings it to the target gene (red) located on the plasmid DNA. (b), Cas9/sgRNA duplexes hybridize each 20 nt DNA sequence (in black) directly preceding PAM (green), generating two double-strand breaks (DSB) 3 bp upstream of the PAM. (c), Target fragment is cut off from the plasmid DNA, which results in two linear DNA fragments.

### 3.3.2 Materials and experimental protocol

#### Materials:

DNA templates (Plasmids, details in appendix)

sgRNAs:

sgRNA 1, agaccgctaaactgaaagtt (20 nt),

sgRNA 2, aggactcctcctgcaagac (20 nt).

Enzymes and digestion kits (New England Biolab®):

*In vitro* DNA digestion kit (item NO. M0386S), containing Cas9 protein and reaction buffer.

Restriction enzymes: EcoRI and PstI with their reaction buffer, CutSmart®.

#### Experimental protocol:

Prepare Cas9 *in vitro* reaction system mix (30  $\mu$ l in total):

1. Mix all the components, the pipetting scheme can be found in Appendix table. 2:

Plasmid DNA and sgRNA concentrations can be calculated by using the same formula

in the Appendix. For the negative control group, we add the same amount of nF H<sub>2</sub>O instead of sgRNA.

2. Incubate at 37°C for 15-20 min.
3. Denature the Cas9 protein by heating sample to 80°C, for 20 min.
4. Proceed with agarose gel electrophoresis.

**Preparation of positive control, DNA double digestion reaction system:**

1. Mix all the components, pipetting scheme can be found in appendix table 3.
2. Incubate at 37°C for 30 min.
3. Denature the Cas9 protein by heating sample to 80°C, for 20 min.
4. Proceed with agarose gel electrophoresis.

**Agarose gel electrophoresis**

1. Prepare 1% agarose by weight in TAE buffer.
2. Melt agarose mixture in microwave oven for 1-2 minutes until agarose is completely dissolved.
3. Allow the agarose solution to cool down. Add GelRed DNA dye, and mix by swirling.
4. Pour the melted agarose solution into the casting tray and put the gel combs and wait for 20-30 min.
5. After casting, remove the combs. Pour approx. 400ml TAE buffer (running buffer) in the gel tank.
6. Mix the samples (Test, positive and negative controls from above) with 6X sample loading dye (New England Biolab®) in a ratio of 5:1.
7. Carefully load the samples in each well, including the DNA ladder (Quick-Load® Purple 2-log DNA Ladder from New England Biolab®).
8. Run the gel at 140V, 400mA for 30-35 minutes. Be careful, DNA running direction should be from negative electrode to positive electrode.
9. After running, check the gel and take an image using Gel Imager (PEQLAB GmbH).
10. Analysis the gel picture by using ImageJ and calculate the intensity of each band to estimate approximate efficiency of cutting.

## 4 Quiz

1. What is the CRISPR/Cas system? What is its biological function in prokaryotic cells? What is the advantage of the CRISPR/Cas system in genome editing compared to other technologies?
2. What is the sgRNA? Which conserved sequence motifs does the sgRNA include? What should we notice when we design a sgRNA for a target on a certain genome?
3. How do you interpret the data that we obtained from the first experiment (3.1) (fluorescence signal vs time)? Could you try to calculate the repression efficiency of dCas9 from the fluoresce-time curve?
4. In a first approximation, during a typical gene expression process the protein concentration  $p$  changes over time as:
$$\dot{p} = \alpha - \delta p,$$
where  $\alpha$  is the production (expression) rate and  $\delta$  is the decay/dilution rate of the protein.
  - Which solutions does this simple differential equation have?
  - What is the steady state of the protein concentration, and what the characteristic time to reach it?
  - Which parameter is influenced by the CRISPRi process, and how does it affect steady state and rise time?
  - Now consider that we measure fluorescence and that the fluorescent protein has to mature before it starts to fluoresce. How can this be included in the kinetic model?
5. What is characteristic of the molecular structures of fluorophores with fluorescence in the visible range?
6. How do you explain the results that we obtained from experiment 3.3? If you only use a single sgRNA in the experiment, what kind of results would you expect from gel electrophoresis? Can you find a way to calculate the cutting efficiency of Cas9?
7. Please use the knowledge you learned today to design an optimal sgRNA to target yeast ATP synthase gene on the *Saccharomyces cerevisiae* genome.



## 5 References

1. Jiang, Wenyan, et al. "RNA-guided editing of bacterial genomes using CRISPR-Cas systems." *Nature biotechnology* 31.3 (2013): 233.
2. Mali, Prashant, et al. "RNA-guided human genome engineering via Cas9." *Science* 339.6121 (2013): 823-826.
3. Ran, F. Ann, et al. "Genome engineering using the CRISPR-Cas9 system." *Nature protocols* 8.11 (2013): 2281.
4. Dominguez A A, Lim W A, Qi L S. Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation[J]. *Nature reviews Molecular cell biology*, 2016, 17(1): 5.13.
5. Semenova E, Jore M M, Datsenko K A, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence[J]. *Proceedings of the National Academy of Sciences*, 2011, 108(25): 10098-10103.

## 6 Appendix

### Relevant formulas for NUPACK calculation

All input sequences,  $\Phi = \{\Phi_1, \dots, \Phi_N\}$ , of one or more interacting RNA strands are specified as a list of bases with  $\Phi_i \in \{A, C, G, U\}$  for  $i = 1, \dots, N$ .

Then the secondary structures of each ordered complex,  $s$ , for one or more interacting RNA strands are defined by a set of base-pairings (Watson-Crick pairs [A-U or C-G]). A secondary structure graph is constructed by ordering the strands around a circle (hence, "circle graph"), drawing the backbones in succession from 5' to 3' around the circumference with a nick between each strand, and drawing straight lines connecting paired bases. During the computation, the partition function  $Q$ <sup>[31]</sup> (1) equilibrium base-pairing probabilities(1.2), and the minimum free energy (MFE) structure for unpseudo-knotted complexes are calculated for each ordered complex:

$$Q = \sum_{s \in \Omega} e^{-\Delta G(s)/kT} \quad (1)$$

Here,  $\Omega$  is the ensemble of unpseudo-knotted secondary structures,  $\Delta G(s)$  represents the free energy change for each ordered complex.  $k$  is Boltzmann's constant and  $T$  is temperature. This function can be used to calculate the equilibrium probability of any secondary structure  $s \in \Omega$

$$P(s) = \frac{1}{Q} e^{-\Delta G(s)/kT} \quad (1.2)$$

The free energy of an unpseudo-knotted secondary structure is calculated using

nearest-neighbor empirical parameters <sup>[32]</sup> (2) for RNA in 1M Na<sup>+</sup> solution. The MFE structure is the structure of the ordered complex which has the minimal total free energy.

$$\Delta G_T^o = \Delta H^o - T\Delta S^o \quad (2)$$

Values for free energy changes  $\Delta G_T^o$ , enthalpy  $\Delta H^o$  and entropy change  $\Delta S^o$  for certain base pair have been experimentally characterized and tabulated, T is the temperature.

### DNA and sgRNA concentration transformation formula

$$\text{Nucleic acid amount (mol)} = \frac{1 \times 10^9 \times \text{concentration}(\mu\text{g} / \mu\text{L})}{\text{molecular weight}} \quad (1)$$

Nucleic acid concentration ( $\mu\text{g}/\mu\text{L}$ ) is measured by Microvolume Spectrophotometers. Molecular weight is calculated by an online calculator (<http://biotools.nubic.northwestern.edu/OligoCalc.html>) based on their sequence.

### Experimental material

DNA templates:

dCas protein gene has been cloned into Biobrick plasmid (psB1K3,

<http://parts.igem.org/Part:psB1K3>)

RFP has also been cloned into Biobrick plasmid (psB1C3,

<http://parts.igem.org/Part:psB1C3>)

**Table.1 cell extract pipetting sheet**

Componrnnts	Amount in cell extract	Volume
dCas9 plasmid	5 nM	V1
rfp plasmid	5 nM	V2
sg RNA	5 nM	V3
T7 polymerase	6 $\mu\text{M}$	0.6 $\mu\text{L}$
Cell extract mix		11.25 $\mu\text{L}$
		<b>Add total volume to</b>
Nuclease-free H <sub>2</sub> O		<b>15 <math>\mu\text{L}</math></b>

**Table.2 Cas9 reaction system pipetting sheet**

Component	30 $\mu\text{l}$ reaction
Nuclease-free water	20 $\mu\text{l}$
300nM sgRNA	3 $\mu\text{l}$ (30nM final)

10X Cas9 Nuclease Reaction Buffer	3 $\mu$ l
1 $\mu$ M Cas9 Nuclease, <i>S. pyogenes</i> (M0386S)	1 $\mu$ l (~30 nM final)
Reaction volume	27 $\mu$ l
Pre-incubate for 10 minutes at 25 <sup>o</sup> C	
30nM plasmid DNA	3 $\mu$ l (3 nM final)
Total reaction volume	30 $\mu$ l

**Table.3 restriction enzymes reaction system pipetting sheet**

Component	25 $\mu$ l reaction
Restriction Enzymes (EcoRI and PstI)	10 units generally 1 $\mu$ l for each one
plasmid DNA	0.5 $\mu$ g
10X NEB cutsmart buffer	2.5 $\mu$ l (1X)
Nuclease-free water	Add up to total volume
Total Reaction Volume	25 $\mu$ l

### Plasmid maps

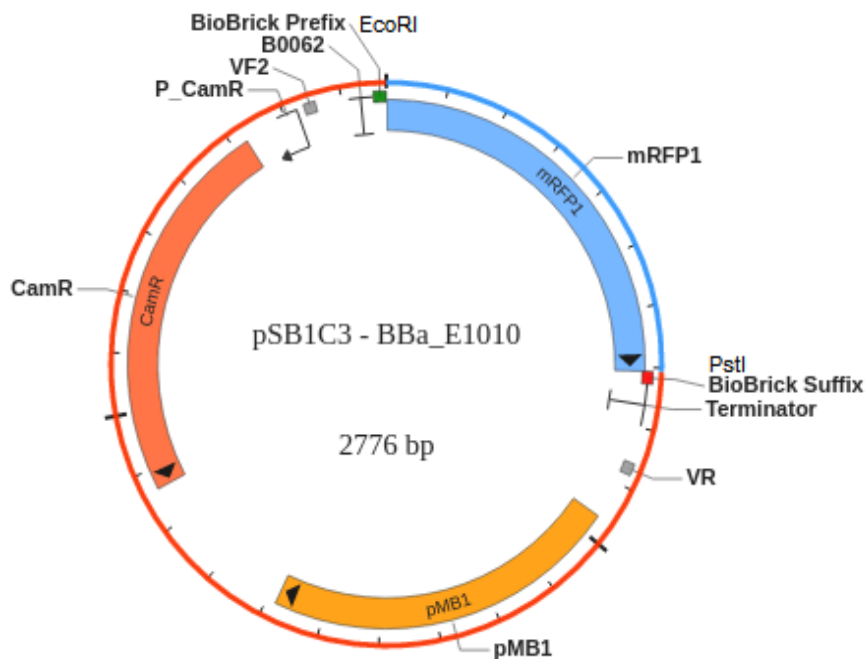


Figure.1 Plasmid BBa\_E1010 (mRFP) in pSB1C3 vector. Biobrick prefix which EcoRI is on the 5' upstream of the *mrfp* gene and the suffix which includes PstI sites is on 3' downstream of the *mrfp* gene.

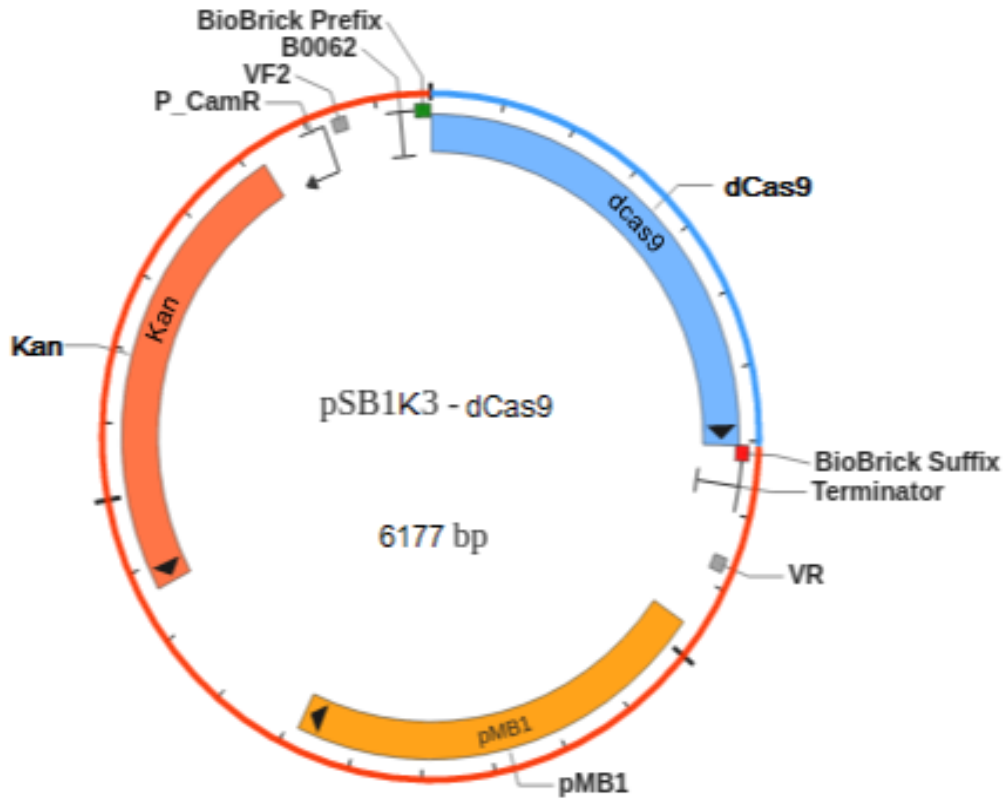


Figure.2 Plasmid dCas9 in pSB1k3 vector. Biobrick prefix which EcoRI is on the 5' upstream of the *dcas9* gene and the suffix which includes PstI sites is on 3' downstream of the *dcas9* gene.