Improving the efficiency of CRISPR-Cas9: The expression & purification of Cas9

By

Elaine A. Shults

Department of Biochemistry, University of Colorado Boulder

Defended April 1, 2016

Thesis Advisor:
Dr. Dylan Taatjes, Department of Biochemistry

Defense Committee:
Dr. Joseph Falke, Department of Biochemistry
Dr. Dylan Taatjes, Department of Biochemistry
Dr. Angela Thieman-Dino, Department of Anthropology

Table of Contents

Abstract

p53 is a transcription factor that plays a significant role in cell cycle regulation, senescence, and apoptosis. Through these mechanisms, p53 acts as a tumor suppressor. The p53 protein has been widely studied for its link to human cancer, and is estimated to be mutated in ~50% of all human cancers. This project focused on generating a p53-null A549 cell line using the genome editing technique CRISPR (Clustered Regularly-Interspaced Short Palindromic Repeats)-Cas9. The p53-null cell line will be used as a control for future projects that aim to identify the various roles of p53 and specific p53 isoforms, such as ΔNp53. The A549 cell line is a non-small cell human lung carcinoma line that exhibits wild-type p53 (WTp53). CRISPR-Cas9 is a revolutionary genome-editing technique that greatly minimizes the cost, time, and difficulty associated with making site-specific genetic changes. It is becoming widely adopted as the standard for genome-editing projects and already has far reaching implications in research, biotechnology, agriculture, gene therapy, and medicine. The CRISPR-Cas9 technique uses a 20 base pair-long programmable crRNA as a "guide" to direct the Cas9 enzyme to a specific location on the genome, where the Cas9 protein will create a double stranded break (DSB) in the DNA. Once this break is made, the cell can repair the damage via non-homologous end joining or homologous recombination. CRISPR-Cas9 was employed to generate a simultaneous knock-in/knockout of mCherry and WTp53, respectively. The mCherry gene was inserted at the WTp53 endogenous translational start site (TSS), thereby rendering WTp53 inactive and mCherry "active" whenever WTp53 would normally be expressed. After the initial transfection yielded low levels of mCherry fluorescence, focus shifted to next-generation CRISPR techniques that would increase the p53 knockout efficiency. One of these techniques included expressing and purifying the Cas9 protein in order to directly introduce it to target cells via electroporation. This approach has been shown to increase the success of site-specific genome editing for two main reasons: Firstly, by adding assembled Cas9 ribonucleoprotein (RNP) directly into the cell, this reduces the number of components needed to enter the cell from four to two. Secondly, it reduces off-target activity because there is a finite amount of Cas9 RNPs and they will be degraded in ~24 hours. Cas9 was overexpressed in Rosetta 2 (DE3) *E. Coli* cells and was purified through a Ni-NTA affinity column followed by a Heparin Ion-Exchange column. The presence of Cas9 throughout the purification was confirmed with Coomassie staining before and after each purification step, and as a final measure with a silver stain and western blot. The Cas9 purification process has been optimized for future uses, and the assembly and electroporation of Cas9 RNPs will likely increase efficiency for any upcoming genome-editing projects within the Taatjes lab.

Introduction

TP53/p53

p53 is a tumor suppressor protein encoded by its gene, TP53. In its full length, p53 is 393 amino acids long and binds DNA as a tetramer (**Figure 1)** (Okorokov et al., 2006). p53 is regulated at the protein level and is expressed at constant low levels until the cell is stressed, at which point kinase activity leads to the release of p53 from its repressor, Mdm2 (Zheng et al., 2015). p53 controls cell cycle arrest, leading to senescence or programmed cell death (apoptosis). p53 has been shown to regulate numerous genes involved in cell growth arrest (p21, Gadd45), DNA repair (p53R2), and apoptosis (Bax, Apaf-1, PUMA, NoxA) (Zheng et al., 2015). It is estimated that TP53 is mutated in nearly 50% of human cancers. Because of its link to preserving the DNA of healthy cells, p53 is commonly referred to as "The Guardian of the Genome," and is one of the most studied transcriptions factors of all time.

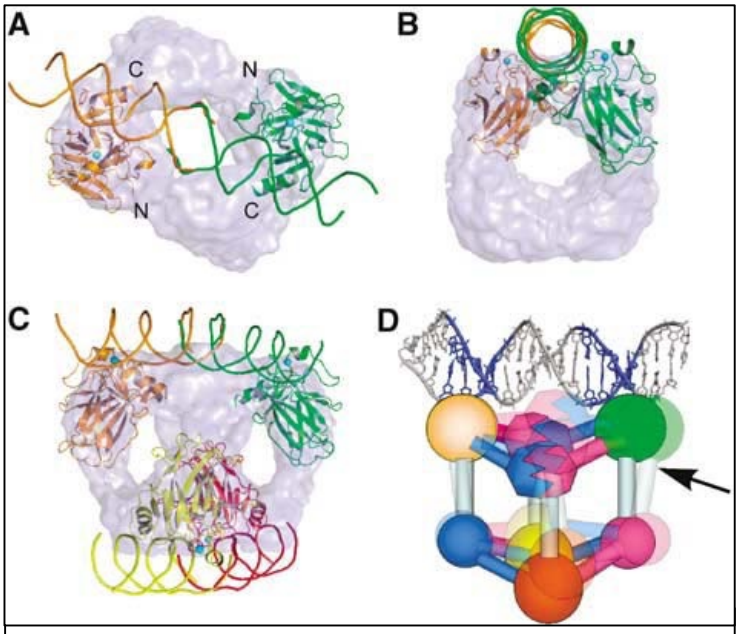


Figure 1. Model of how the p53 tetramer binds to DNA. N and C represent N- and C- termini, respectively (Okorokov et al., 2006).

The gene, TP53, is 25,772 bases pairs long and is located on the short arm of chromosome 17. It encodes for at least 15 isoforms and regulates over 100 other genes (Khoury et al., 2010). p53 is 43.6kDa in size and contains six main domains (Okorokov et al., 2006):

1. (1-67) Transcriptional activation domain
2. (67-98) Proline-rich domain
3. (98-303) Core domain
4. (303-323) Nuclear Localization containing domain

5. (323-363) Oligomerisation domain
6. (363-393) C-terminal basic domain

p53 activity has also been linked to aging (Lin & Taatjes, 2014). p53 transcriptionally regulates the Mechanistic Target of Rapamycin (mTOR) pathway, by inducing the expression of genes (Sestrin1/2 & TSC2) involved in the suppression of mTOR. mTOR has clear links to the aging process by accelerating the aging phenotype through stimulated ribosome biogenesis and mRNA translation (Goudarzi et al., 2014). Additionally, p53 has been shown to regulate ribosomal biogenesis, a key behavior in cancer proliferation. Studies that have inhibited ribosomal biogenesis lead to cell stress and activation of p53 (Goudarzi et al., 2014). Blocking the mTOR pathway with Rapamycin results in an increased lifespan for mammals (Feng et al., 2011).

p53 acts as a transcriptional activator for multiple genes (Feng et al., 2011). p53 also regulates its own levels by controlling the expression of Mdm2, which binds to unphosphorylated p53 leading to the protein's nuclear export and degradation. When p53 is phosphorylated, this restricts Mdm2's ability to bind, increasing the longevity of p53 (Feng et al., 2011). When DNA damage occurs, ATM kinases are activated, resulting in the phosphorylation of p53 and protection from degradation until the DNA is repaired (Zheng et al., 2015). This is one mechanism by which p53 remains viable in order to repair the DNA damage.

ΔNp53

ΔNp53 is a naturally occurring isoform of p53 that lacks the first 39 amino acids from its N-terminal end (**Figure 2**). ΔNp53 levels increase in response to different types of cell stress, such as: ER stress, oxidative stress, and/or

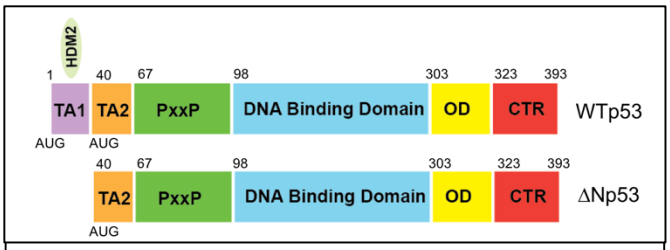


Figure 2. Schematic of WTp53 versus ΔNp53. HDM2 is a ubiquitin ligase that tags p53 for degradation. TA1: transactivation domain 1, TA2: transactivation domain 2, PxxP: Proline-rich domain, OD: Oligomerisation domain, CTR: C-terminal region (Levandowski, 2015).

serum deprivation (Lin & Taatjes, 2013). The isoform is translated via an internal ribosomal entry site (IRES), or through alternative splicing. ΔNp53 lacks the Mdm2 binding site present on WTp53, and therefore its longevity is not regulated by the mechanisms regulating WTp53 (Lin et al., 2013).

Because p53 is a tetramer, if WTp53 and ΔNp53 are co-expressed, their subunits combine in random, creating a mixture of ΔNp53:WTp53 in the ratios: 0:4, 1:3, 2:2, 3:1, and 4:0. This resulting mixture can make it challenging to identify the distinct role of WTp53 from its isoform ΔNp53. Studies have shown that ΔNp53 and WTp53 co-expression results in an accelerated aging phenotype in the mouse model (Pehar et al., 2010). Earlier death rate, increased levels of osteoporosis, and decreased learning ability were seen in a study examining ΔNp53:WTp53 genotypes in mice compared to the wild type controls (Pehar et al., 2010).

mCherry

mCherry is a monomeric red fluorescent protein with an excitation wavelength of 587 nm and an emission at 610 nm. mCherry is 27kDa in weight and exhibits a ß-barrel shape with a chromophore and its center. There is an alpha-helical domain which threads itself through the barrel's center. The mCherry gene is 1304 base pairs long and served as the gene to be inserted at the p53 endogenous TSS using CRISPR-Cas9.

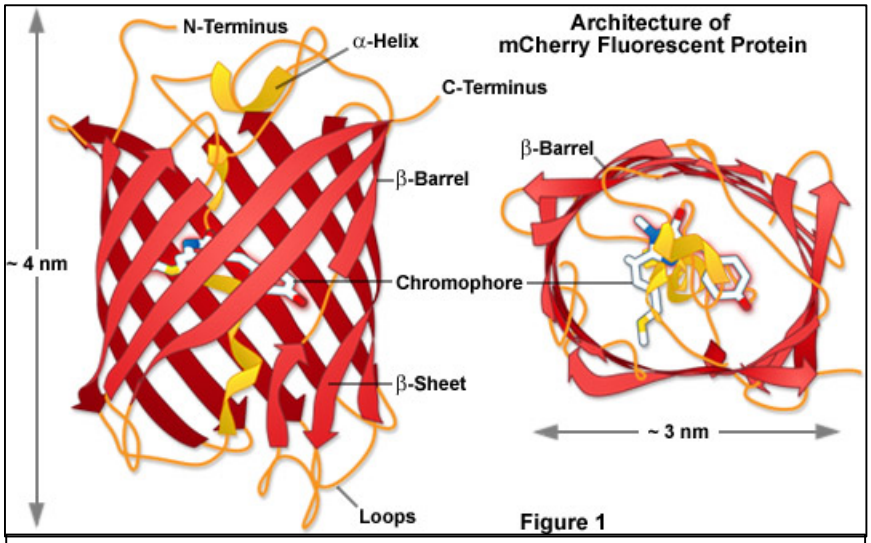


Figure 3. Structural model of mCherry florescent protein (http://zeiss-campus.magnet.fsu.edu/articles/probes/anthozoafps.html).

A549 cell line

The A549 cell line are non-small cell human lung carcinoma cells. They exhibit epithelial morphology and can be easily transfected using standard lipofectamine protocols. The population doubling time is approximately 22 hours. Most cells had two X and two Y chromosomes. These cells contain WTp53 on both alleles and are used to examine the role and function of p53 and its many isoforms within the Taatjes lab. In order to observe how ΔNp53 and WTp53 behave, it became necessary to generate a p53-null cell line within the same cell line that was being used in the ΔNp53 study. A quick, efficient and precise genetic-editing technique was required in order to generate the knockout.

Known p53-null cell lines

Although p53 plays a significant role in gene regulation, it is not an essential protein for the viability of all cells (Soussi, 2010). There are existing p53-null cell lines that have been derived from human tissue. Additionally, the term "p53-null" can prove to be ambiguous in both everyday use and in literature. In some contexts, the term "p53-null" has been used to describe cell lines that have mutated p53 genes but may exhibit normal or partial p53 responses (Soussi, 2010). In this context, a true "p53-null" cell line exhibits no p53 behavior. The following table encompasses a few of these known p53-null lines and their associated cancer type:

Cell line	Cancer
Hep3B	Hepatocellular carcinoma
KATO III	Gastric carcinoma
NCl-H1299	Non-small cell lung carcinoma
Soas-2	Osteosarcoma
LZ-Z308	Glioblastoma
HL-60	Promyelocytic leukemia
MDA - MB-157	Breast carcinoma

Table 1. Table listing multiple known p-53 null cell lines and the human cancer from which they were derived. (p53.free.fr/Database/Cancer_cell_lines/cell_lines_1.0.pdf)

To the Taatjes' lab's knowledge, a p53-null A549 cell line has not been generated before, which warrants further exploration into A549's viability and behavior as a p53-null line.

CRISPR-Cas9

CRISPR-Cas9 is a revolutionary genome-editing technology discovered by studying how bacteria defend themselves from viruses and other infectious foreign DNA through the process outlined in **Figure 4** (Mojica et al., 2005; Ebina et al., 2013). If a bacterium is infected with a virus, a CRISPR-associated (Cas) polymerase copies the viral DNA in short segments and inserts sequences of the viral DNA into a region of the bacteria's genome containing highly repetitive DNA. This region of repetitive DNA is referred to as the **C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats ("CRISPR") array. This enables the bacteria to keep a history of previous infections. The "saved" viral DNA can later be transcribed to act as "guide" RNA, which complementarily binds to invading viral DNA, directing the Cas9 nuclease to cleave it rendering the virus ineffective (Oost et al., 2014). This technique results in a robust, adaptive defense mechanism against infection.

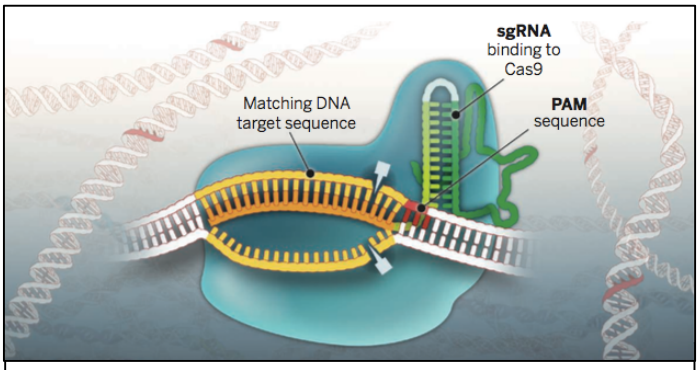


Figure 5. Cas9 RNP complex binding to DNA upstream of the PAM sequence (Doudna et al., 2014).

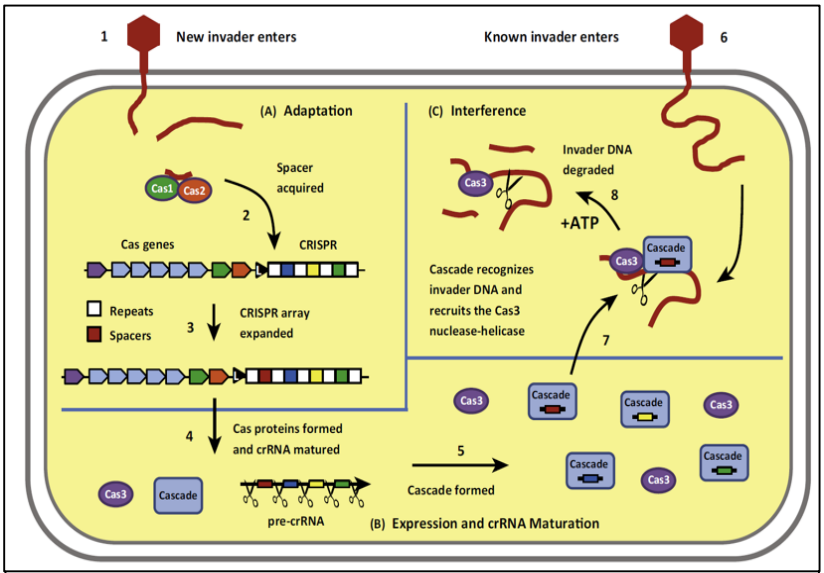


Figure 4. Overview of bacteria's adaptive immunity against foreign DNA (https://www.wageningenur.nl/en/show/CRISPRCas).

CRISPR's immense potential lies in its ability to be harnessed for precise genetic-editing purposes. There are two main components to the CRISPR-Cas9 system: The "sgRNA" and the Cas9 protein. SgRNA stands for "single guide RNA" and is a single complex engineered from two pieces of RNA: the tracRNA and crRNA. CrRNA is programmable up to 20 base pairs and is responsible for CRISPR-Cas9's location-specific activity. The crRNA Watson-Crick base-pairs with the target DNA, and the tracRNA acts as a bridge between the crRNA and the Cas9 enzyme (**Figure 5**). One requirement for CRISPR-Cas9 is the presence of a PAM sequence, which stands for "Protospacer Adjacent Motif," and is the sequence NGG, where N is any base and G is Guanine. The PAM sequence must be directly downstream of the target DNA in order for Cas9 to bind DNA (Doudna et al., 2014). Cas9 is a large (~160kDa) nuclease responsible for creating a double-stranded break (DSB) in DNA. Once the break has been made, natural DNA repair mechanisms can mend the break, often inserting a few extra nucleotides, effectively knocking-out the gene if the additional bases result in a frame-shift mutation; this is called non-homologous end joining (NHEJ) and can be a useful technique for generating a knockout (**Figure 6**). Alternatively, homologous recombination (HR) occurs when a double stranded break is repaired using homologous DNA, usually derived from the cell's homologous chromosome (McClendon et al., 2016). Using HR, foreign or "template DNA" can be inserted at the DSB, but only if the template DNA is flanked by the same DNA sequences that frame the DSB (Guilinger et al., 2014; Doudna et al., 2014). This project used a plasmid

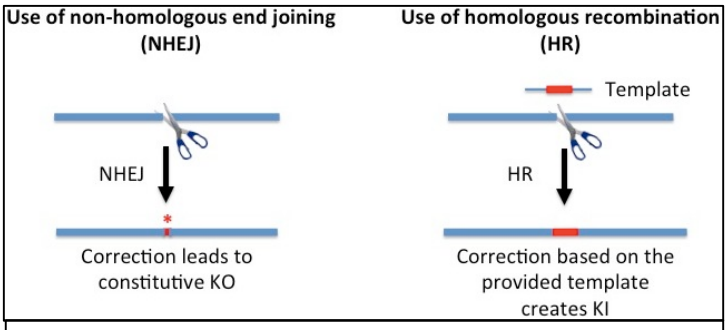


Figure 6. NHEJ vs. HR (https://www.genoway.com/crispr-cas9/nucleases.htm).

containing mCherry (1.3 kb) flanked by 1.5 kb of homologous DNA sequences directly upstream and downstream of the TSS of WTp53 in A549 cells.

The applications for CRISPR-Cas9 are seemingly endless, and are already being explored in the realms of gene therapy, agriculture, medicine, livestock, molecular biology and drug targeting (Ebnia et al., 2013; Doudna et al., 2014). The initial transfection used 1 tracRNA, 1 crRNA, 1 Cas9 plasmid, and 1 repair template DNA plasmid. The transfection of four components is not ideal for a number of reasons. Firstly, the cells must be exposed to a high level of stress in order to "take up" four separate pieces of foreign DNA and RNA. Secondly, by transfecting the Cas9 plasmid, in essence the cell is continually expressing Cas9. As the plasmid is continuously transcribed and translated, Cas9 levels stay at a relatively high level until the plasmid is degraded. One of the main problems with the CRISPR-Cas9 method is the possibility of off-target cutting, which is more likely to occur when Cas9 levels are held at a higher concentration for a relatively long period of time (Cho et al., 2014). Off-target cutting is detrimental to the health of the cell and may compromise the accuracy of the affected outcomes or goals of the project. Off-target activity is the main reason why research using CRISPR-Cas9 in human therapeutics is currently on hold, and there is a general consensus that the technology is not safe enough yet to use in viable human embryos (Cyranoski, 2015).

Cas9

The Cas9 enzyme is approximately 160kDa and is a part of the Cas family of proteins originating from bacteria. The Cas9 enzyme has four main domains shown in **Figure 7**. RuvC (grey) cleaves the non-target strand. HNH (cyan) cleaves the target strand. Guide RNA Domain (blue) is responsible for binding to the guide RNA, specifically the tracRNA component, which leads to Cas9's site-specific binding on DNA. PI (PAM Interacting) Domain (orange) binds to the downstream PAM sequence (NGG) and is the most critical component for Cas9-DNA binding (Nishimasu et al., 2014). The double stranded break is generated 3-4 nucleotides upstream from the PAM sequence. As indicated, Cas9 has two catalytic sites, each one is responsible for the cleavage of one strand of DNA. Mutating aspartate to alanine in the RuvC domain successfully converted the enzyme to a nickcase that only cleaves one strand of DNA. This engineered nickcase is called SpCas9n and is currently being studied for its potential applications (Cong et al., 2013).

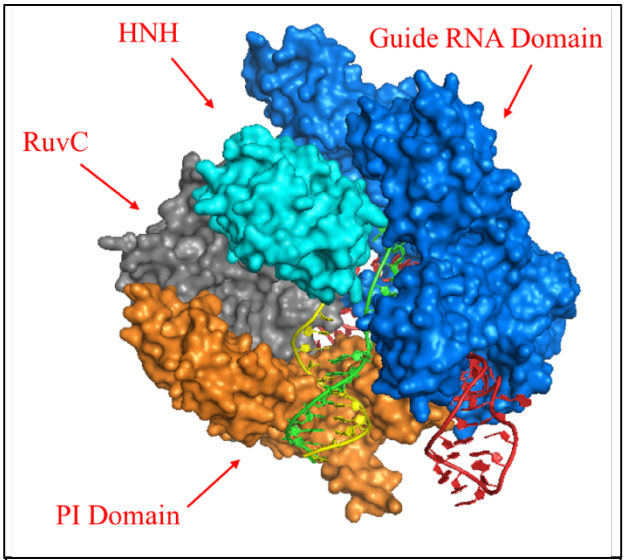


Figure 7. Cas9 crystallography structure. Adapted from Anders *et al* in the 2014 *Nature* paper.

pMJ915

pMJ915 is the recombinant plasmid designed by the Doudna Lab and supplied by Addgene that contains the Cas9 gene under control of the T7 promoter (Appendix A). The Cas9 contains 6xHis-MBP fusion tag on its N-terminal end. The polymerase T7 is specific to the T7 promoter and will only transcribe DNA downstream of its promoter. The total vector size is 10,124 bp in length. The plasmid confers ampicillin resistance and is designed for expression in bacteria.

Techniques to improve efficiency

There are a number of proposed and implemented techniques used to increase CRISPR Cas9's efficiency in cells and to minimize off-target activity. Synchronizing the cell's cycle increases the efficiency of CRISPR activity because the cells can be programmed to all enter metaphase simultaneously so the DNA will be bound less tightly to histones and the chromatid in anticipation of replication (Lin et al., 2014). This makes the DNA more accessible to the CRISPR-Cas9 complex and leads to increased efficiency of targeted cleaving.

Other techniques to reduce off-target cutting include truncating the length of the crRNA (<20 bp), which has been shown to decrease off-target mutagenesis by 5,000 fold. Additionally, inhibiting one of the catalytic sites on Cas9 so it cleaves just 1 strand (SpCas9n) has been shown to increase efficiency (Fu et al., 2014; Shen et al., 2014). Instead of generating a double stranded break, the mutated Cas9 cleaves a single strand of DNA but is used in tandem with another Cas9 nickcase acting on the opposite end of the target DNA. The target DNA is then "flanked" by two Cas9 enzymes, each mutated in one catalytic site. This increases the site-specific accuracy up to 1,000-fold, in turn reducing cell death, as it is much easier for the cell to repair a single strand DNA break than a double stranded break in the event of off target effects (Ran et al., 2013).

The next generation efficiency-improving CRISPR technique pursued by this project was to directly add the Cas9 ribonucleoprotein (RNP) complex into the cells as opposed to transfecting the Cas9 plasmid. This will limit the presence of Cas9 in the cell to approximately 24 hours; it has been shown that Cas9 will be degraded by that time (Doudna et al., 2014). In order to directly insert the Cas9 protein it must be pure and catalytically active. Purified Cas9 can be purchased but it is costly and a short-term solution, so the decision was made to express and purify Cas9 so it can be used in future genome-editing experiments via RNP assembly and electroporation.

Results

Project Overview

The project was divided into two parts. Part I attempted to generate a p53 knockout using CRISPR-Cas9 in A549 cells, and Part II expressed and purified the Cas9 protein. Part II came about in the wake of low efficiency after the initial transfection, so the focus of the project shifted to next generation CRISPR-Cas9 techniques to increase efficiency.

Part I: Generate p53 knockout via Homologous Recombination of mCherry

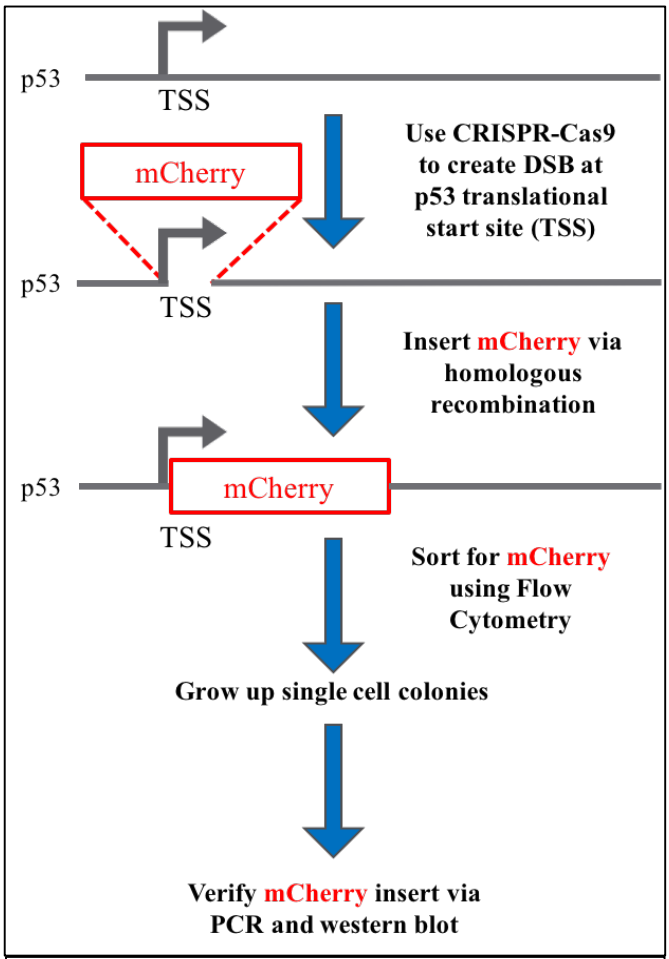


Figure 8. Part 1 Flowchart. p53 knockout/mCherry knock-in using CRISPR-Cas9 at the endogenous p53 translational start site (TSS) in A549 cells.

Day 1 of the transfection began when A549 non-small cell human lung carcinoma cells were grown up to confluence. To transfect, appropriate transfection solutions (serum free media, crRNA, tracRNA, Cas9 plasmid, mCherry plasmid, & DuoFECT) were added at appropriate concentrations. The reagents incubated for 20 minutes while the cells were prepared at their desired concentration (100,000 cells per well on 24-well plate). 52uL of transfection reagent (containing all CRISPR-Cas9 components described above) were added to 448uL of cell preparation (~200,000 cells/mL), and allowed to incubate for 48 hours. On Day 3,

the transfected cells were selected by adding 1.5 ug/mL puromycin. On Day 5, the selection media was removed and replaced with regular media.

The media was replaced every two days, and the surviving cells were grown until confluent and passaged to 60mm plates. Next, Flow Cytometry was used to select for cells exhibiting mCherry fluorescence with a negative control reading used to create a "gated" positive region (**Figure 9**). The negative control sample contained all the reagents minus the mCherry plasmid. Only 0.012% of cells fluoresced out of the thousands that were imaged. The cells that were selected for positive fluorescence (0.012%) were grown up as single-cell colonies in 96-well plates. All single-cell colonies exhibiting growth were imaged using 16 photos/well in the Spencer Lab, and all wells with signs of fluorescence were expanded. The final percentage of cells exhibiting fluorescence are below:

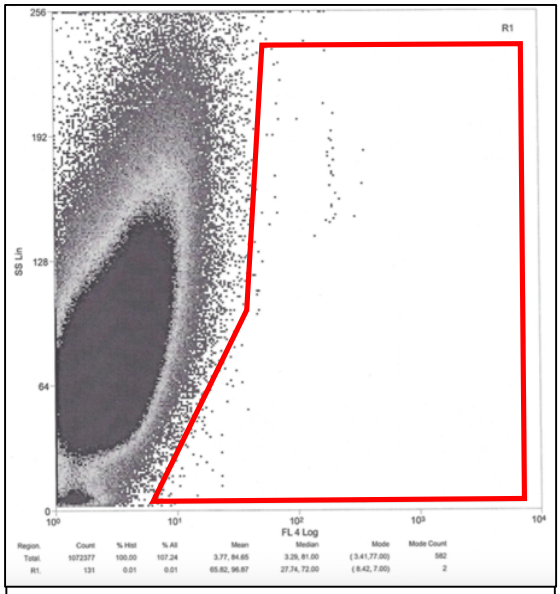


Figure 9. Flow Cytometry Data. "Gated" positive region delineated in red. Each dot represents the reading of one cell. X-axis: Fluorescence detected (log scale). Y-axis: Side Scatter - angle of orthogonal light scatter (linear scale).

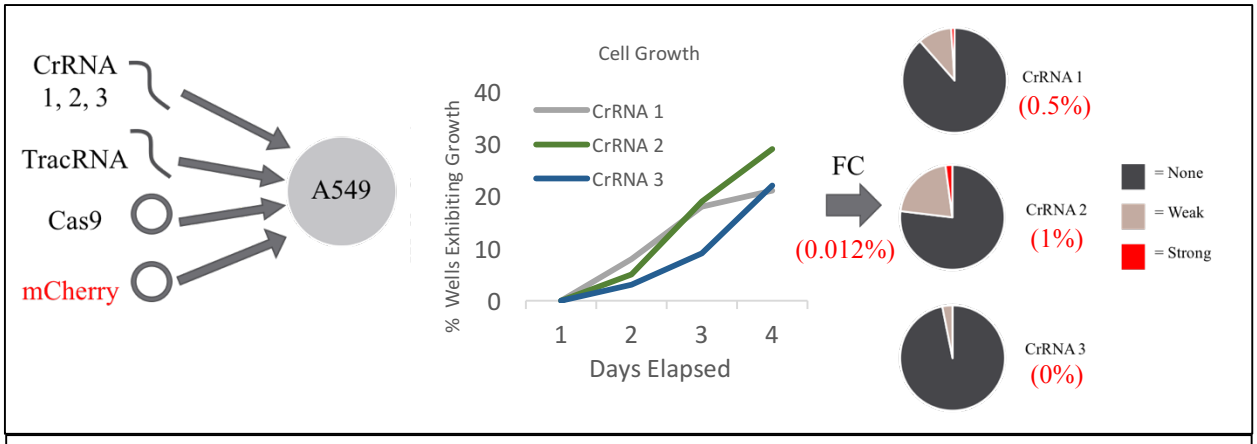


Figure 10. Part 1 Overview of Transfection Efficiency. Four components were transfected into A549 cells. The resulting cell growth data was obtained. Approximately 30% of each crRNA group grew to be passaged to undergo Flow Cytometry (FC). ~ 0.012% of cells exhibited fluorescence and were grown up as single cell colonies on 96-well plates for further imaging. Further imaging revealed ≤ 1% of cells "strongly" fluoresced.

Based on the low levels of fluorescence (≤1% of 0.012%) and the weakness of the fluorescence it was decided that the efficiency of the transfection was too low to continue. The project then shifted to expressing and purifying the Cas9 protein to increase genome editing efficiency.

Part II: Expression & Purification of Cas9 protein

Rosetta 2 (DE3) *E. Coli* cells were grown up in order to overexpress Cas9. A mini-prep purified the plasmid obtained from Addgene and the Cas9 plasmid was transformed into Rosetta 2 (DE3) cell lines. 4 L of Terrific Broth were used throughout the transformation to maximize Cas9 expression because its expression is typically low. The OD600 (optical density) is an indication of bacterial growth, and once it reaches 0.8-1.0, the plasmid can be induced through IPTG addition, as the Cas9 plasmid is under control of the lac operon. LacL inhibits the promoter, but IPTG binds to LacL, releasing it from the promoter and allowing for the transcription of downstream (Cas9) genes (Marbach et al., 2011). The OD600 measurements and the moment of IPTG addition are shown in Appendix B & C. In **Figure 12**, uninduced (U) and induced (I) samples were taken at increasing concentrations and run on a polyacrylamide gel stained with Coomassie to show the appearance of an induction band, which is Cas9.

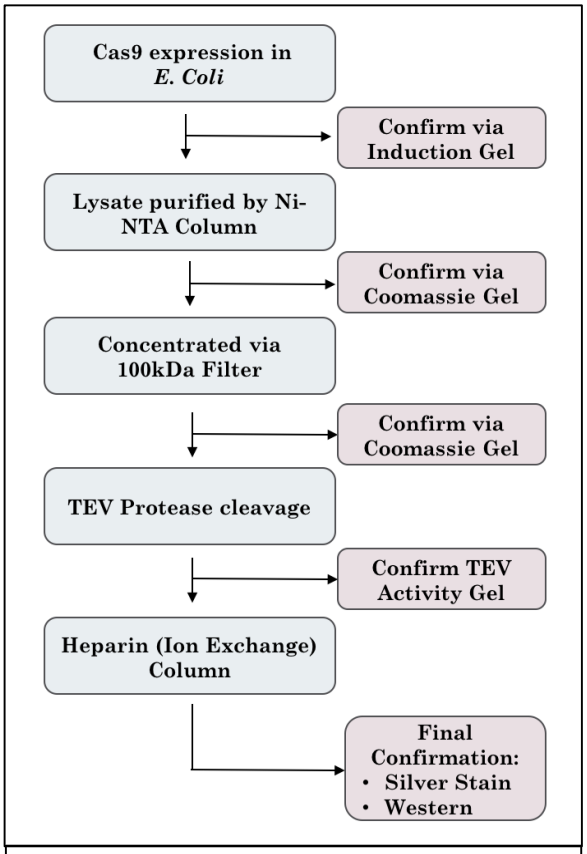


Figure 11. Part 2 Flowchart. Cas9 Purification.

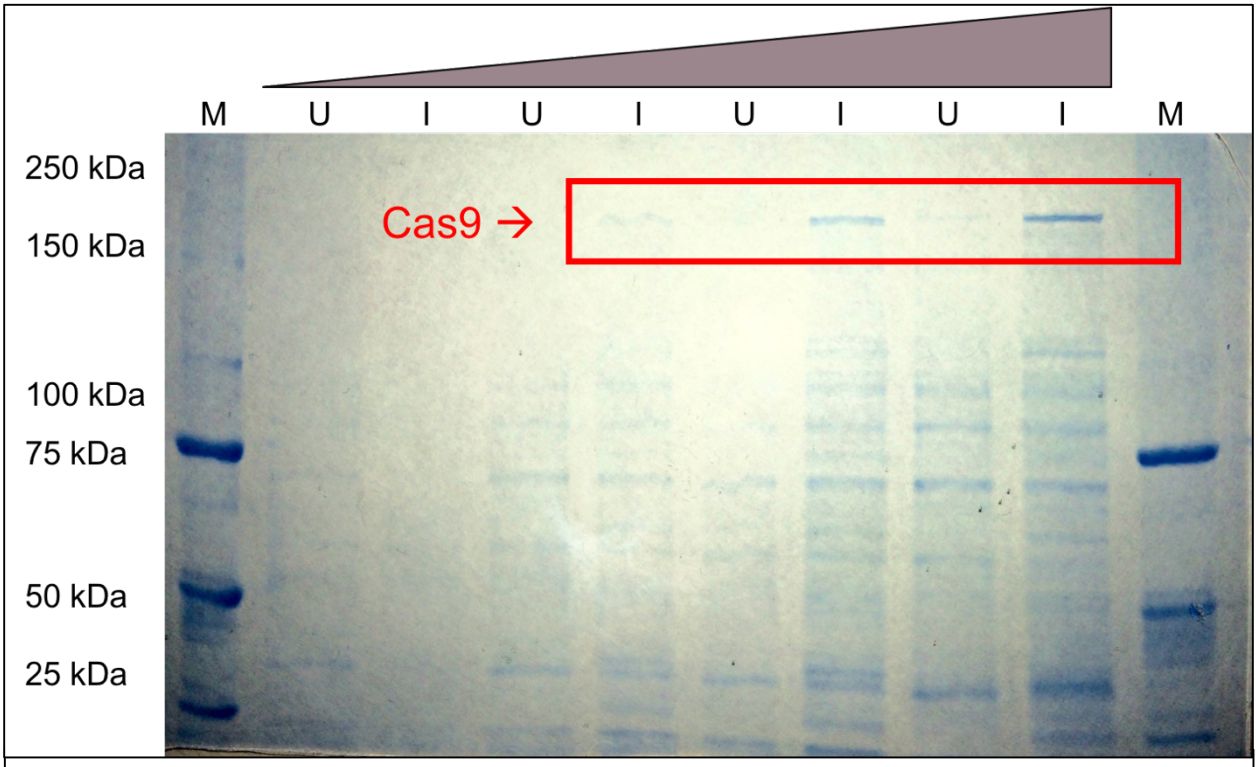


Figure 12. Induction Gel on SDS-PAGE 7.5% polyacrylamide gel in Coomassie stain. M: "Marker;" U: "Uninduced;" I: "Induced." U and I pairs were loaded at increasing volumes from 3 uL, 5uL, 10uL to 15 uL, left to right. The appearance of the Cas9 band indicates a successful induction.

Ni-NTA column

The Ni-NTA purification method is effective in purifying proteins that contain a series of Histidine residues. The Cas9 in question has a 6xHis-MPB fusion tag on its N-terminal end, which qualifies it to be used in a Ni-NTA purification. The Ni-NTA column is composed of beads attached to the molecule NTA, nitrotriaceticacid, which is a chelating ligand in the presence of an immobilized Ni^{2+} ion. The negative neighboring histidine residues interact with the column's complex through coordination

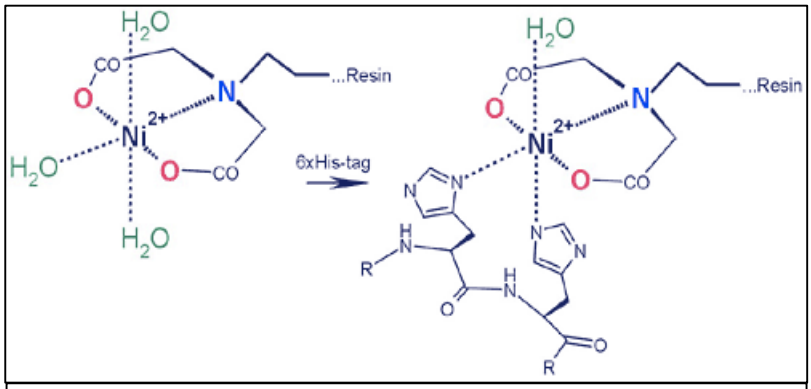


Figure 13. Ni-NTA coordination bonds interacting with water and histidines (http://www.gentoprice.com/ni-ida-agarose.htm).

bonds. The coordination complex is shown in **Figure 13**. A coordination bond is similar to a covalent bond with the exception that both shared electrons come from the same atom. The Ni-NTA column was performed in a batch purification using Ni-NTA beads, but was rocked for a period of 10 minutes per elution as opposed to using a vertical flow-through column. Increasing concentrations of imidazole (which structurally resembles histidine) were used to compete with histidine's affinity to the NTA complex, eventually washing out the desired Cas9, shown in the Coomassie stained gel under "Elution" (**Figure 14)**.

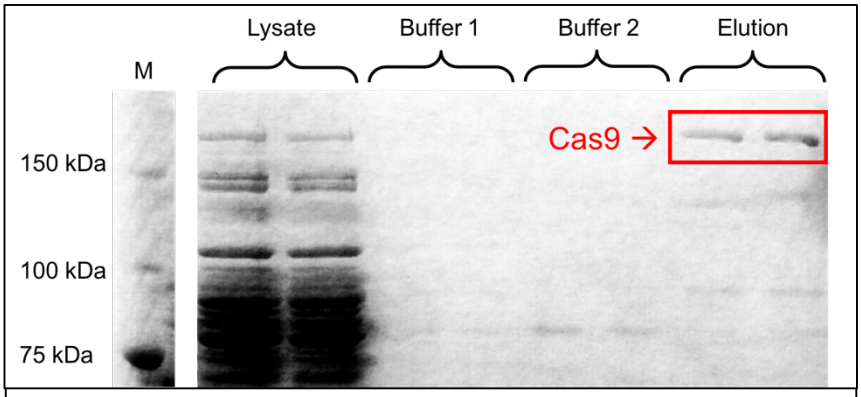


Figure 14. Ni-NTA batch duplicate elutions on SDS-PAGE 7.5% polyacrylamide gel in Coomassie stain. M: "Marker;" Buffer 1: 0 mM Imidazole; Buffer 2: 10 mM Imidazole, Elution: 300 mM Imidazole.

100kD Amicon Filter

Once the elution containing Cas9 was collected from the Ni-NTA batch, it was condensed via a 100kDa Amicon filter. The elution was placed in the top portion of the condenser and centrifuged until the sample was concentrated from ~10mL to 2x 500 uL samples. All protein less than 100kDa in size is free to pass through the membrane, whereas larger proteins, such as Cas9, will not flow through the membrane, allowing it to remain at the top of the tube in the solution at a higher concentration. During this concentration process, a wispy white precipitant began to form in the solution. This was both a positive and negative

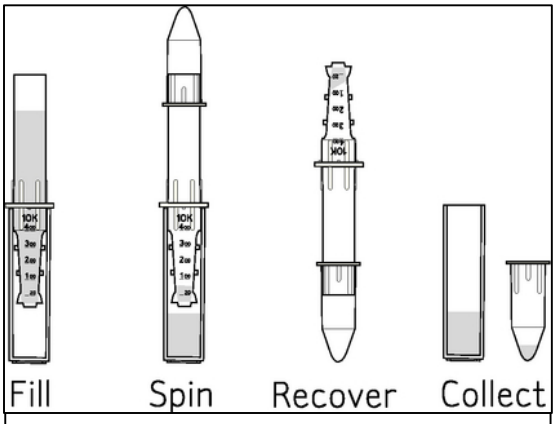


Figure 15. Schematic of Amicon 100kDa protocol (https://www.merckmillipore.com).

sign. It was positive because it indicated protein was present in the sample. It was negative because it meant protein was "crashing" out of the solution, which means the protein's conformation changed so its hydrophobic residues, which usually reside on the inside of its 3D shape, began flipping outward. When hydrophobic resides come into contact with polar solution, they react insolubly, lending the protein to "crash out." This can occur when the ion concentration is high as this leads to an "ion shield" around the protein, which prevents the protein's hydrophilic residues from interacting with water molecules. Instead, the "ion shield" permits the protein to interact with other proteins, which eventually aggregate together and precipitate. Once a protein crashes out of a solution, it is lost to the rest of the purification process as it is energetically unfavorable for it to go back into the solution.

TEV Protease

Once the elution was concentrated, TEV protease was added to the sample to cleave off the 6xHis-MPG fusion tag to continue the purification. TEV protease was isolated from the Tobacco Etch Virus and is a Cysteine protease that specifically recognizes the sequence ENLYFQ/S: Glu-Asn-Leu-Tyr-Phe-Gln-Ser and cleaves between Gln and Ser. The TEV Protease was added overnight at a concentration of 1ug per ~50 ug of substrate. A polyacrylamide gel was stained with Coomassie gel (**Figure 17**) was run as a confirmation for its effectiveness.

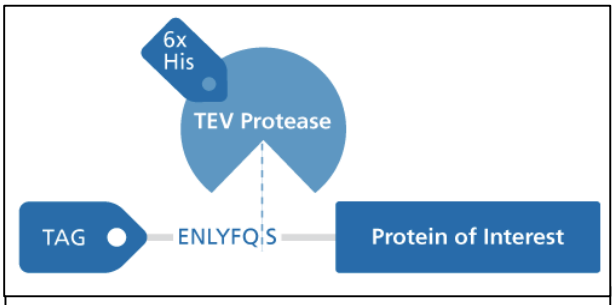


Figure 16. TEV Protease Graphic (https://www.biomol.com/tev-protease.html?id=1211).

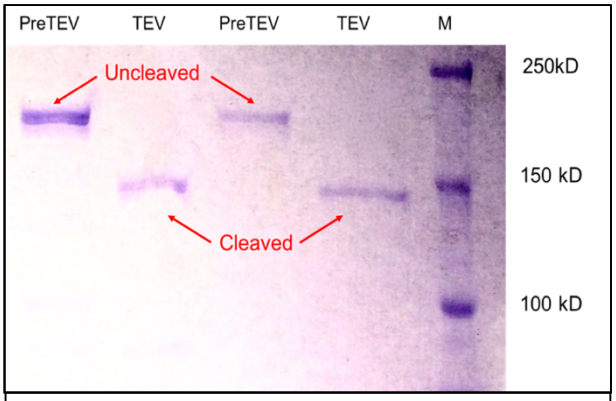


Figure 17. TEV Protease activity on SDS-PAGE 7.5% polyacrylamide gel in Coomassie stain.

Heparin Ion Exchange Column

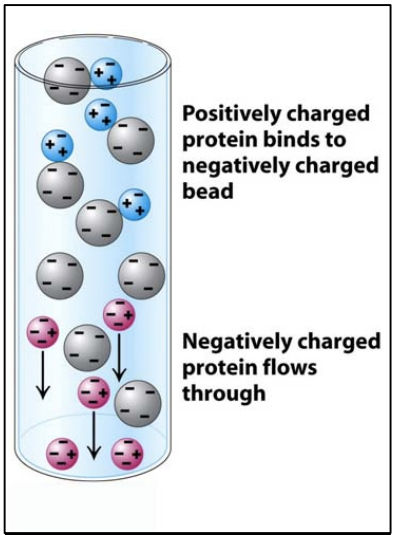


Figure 18. Cation exchange column (Freeman et al., 2012).

Once the sample had been cleaved of its 6xHis-MBP tag, it was run through a HiTrap Heparin Ion Exchange Column. This ion exchange column is a cation-exchange column, meaning it contains negatively charged resin that ionically binds with molecules that are positively charged (**Figure 18**). A key importance is that the protein of interest must carry a net positive charge at the working pH, and the column must carry a net negative charge at the same pH. The difference of charges between the protein of interest and the resin at the working pH determines the strength of the binding. The pI (Isoelectric point) of Cas9 is ~pH 8. This is the pH at which its net overall charge is neutral. When the working pH < pI, the overall charge of the protein will be positive, as there are more H+ ions in the solution interacting with the protein's residues. In this case the working solution is pH=7.5, so the overall charge of Cas9 is positive, hence, the column's resin should be negative. Increasing salt concentrations begin to compete with the resin-protein interactions and eventually the protein is eluted as the cations in the salt solution form ionic bonds with the resin and the charged residues on the protein.

As the salt concentration gradient increased, any remaining uncleaved Cas9 sample eluted first and the cleaved Cas9 sample eluted second, as indicated by the peaks on **Figure 19**. The peaks show absorbance at wavelength = 280 nm, which is where Tryptophan and Tyrosine absorb. Cas9 contains 7 Tryptophan's and 54 Tyrosine's, so it is a good candidate for this method of detection (Nishimasu, 2014). The fractions with the highest peaks were combined and sampled on a silver stain gel, shown in **Figure 20**. Silver stain can be 200x more sensitive than Coomassie stain, so it

was a strong indication that the final Cas9 sample was remarkably pure (Bio-Rad). A final affirmation of Cas9's presence was confirmed via a western blot, using antibodies specific to Cas9, shown in **Figure 21**.

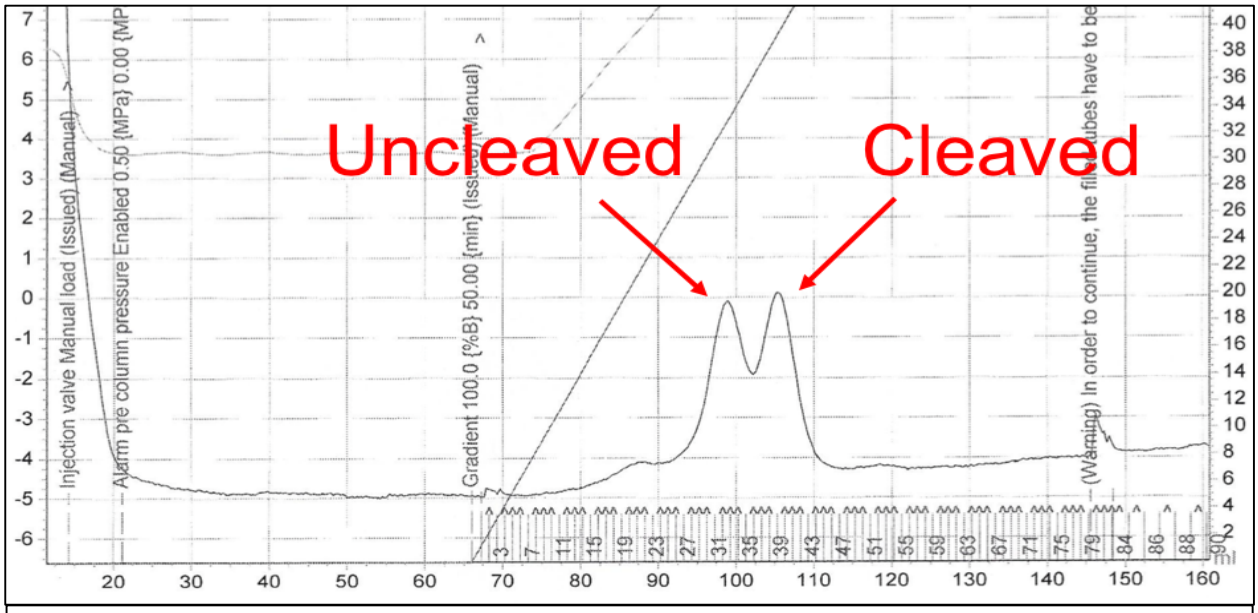


Figure 19. Elution peaks from HiTrap Heparin Ion-Exchange column with the uncleaved Cas9 eluting first and the cleaved Cas9 eluting second. Left Y-axis: UV=280 nm absorbance (mAU) reading. X-axis: Fractions collected. Right Y-axis: % of total solution.

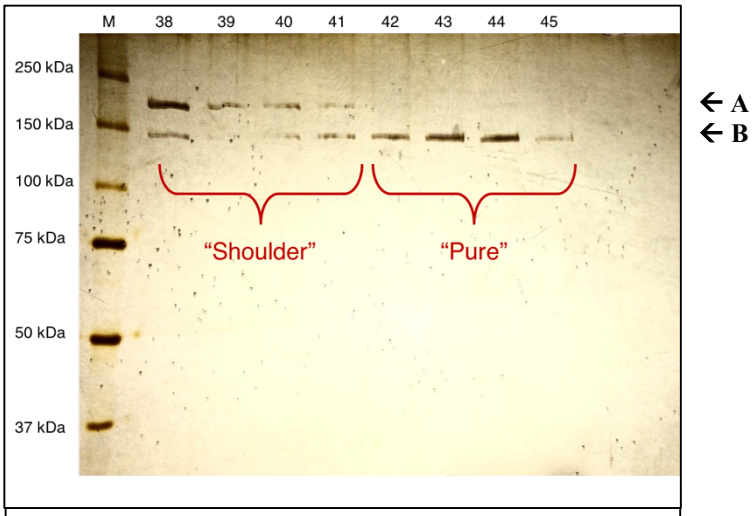


Figure 20. Silver stain of peak fractions. **A**: uncleaved Cas9. **B**: cleaved Cas9. Fractions 38-41 were combined as "Shoulder" sample and fractions 42-45 were combined as "Pure." Both were run on western blot shown in Figure 21.

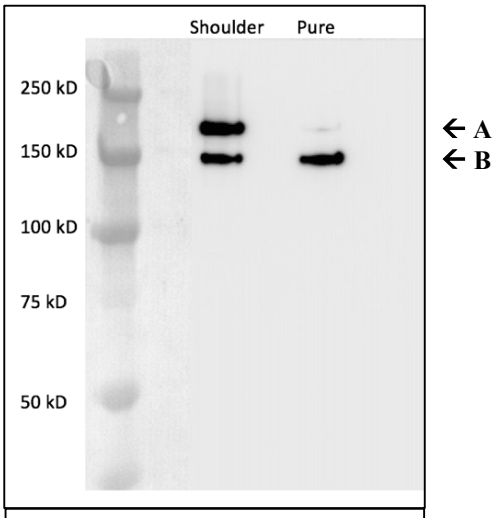


Figure 21. Western blot of combined fractions selected from the silver stain. **A**: uncleaved Cas9. **B**: cleaved Cas9.

Discussion

The original transfection, which relied on transfecting four separate components: 1 crRNA 1 tracRNA, 1 Cas9 plasmid and 1 mCherry plasmid, was low in efficiency (<1% of 0.012% of cells fluoresced). As an alternative practice, the direct electroporation of Cas9 was discussed as a possible way to increase efficiency. The second half of the project focused on the expression and purification of Cas9. Throughout the purification, Cas9 proved to be a robust enzyme, and although every step was preformed at 4Cº, it remained resilient with the exception of minor "crashing out" during its concentration. Samples were taken for SDS-PAGE analysis before and after every purification step to ensure it hadn't been lost throughout the process.

The direct electroporation of Cas9 RNP will result in finite levels of Cas9, resulting in less off-target cutting. A study showed that by direct electroporation, Cas9 was degraded in 24 hours: contrast this with transfecting its plasmid, which expresses Cas9 for several days (Kim et al., 2014). Additionally, by reducing the number of components needed to be transfected from 4 to 2, this increases the likelihood of a successful overall transfection. By comparing the two transfections side-by-side and checking for fluorescence via Flow Cytometry and visual imaging, one would anticipate an increase in efficiency. A recent study that delivered the Cas9 RNP complex to human cells saw on-target mutations as high as 79% of the time with a reduction of off-target activity as compared with transfecting plasmids (Kim et al., 2014). The expression and purification of Cas9 was successful and has been optimized for future uses. The direct assembly of the Cas9 RNP complex is a next-generation CRISPR-Cas9 technique that yields higher efficiencies, and will likely be invaluable for future genetic-engineering projects in the Taatjes lab. Purifying Cas9 was the first step to assembling the complex, and its use will hopefully provide key insights into the importance of p53 and help demystify the roles of its many isoforms.

Future Directions

Now that a purification protocol has been established, optimized, and proven effective, the next steps are to first validate the purified Cas9 activity *in vitro*, then to repeat the purification process with a higher initial volume of cells to end with a greater Cas9 yield. A single transfection typically calls for ~30ug of Cas9. Once an adequate amount of Cas9 has been purified and validated, the ribonucleoprotein complex (RNP), will be assembled and added into the target cells directly via electroporation, along with the repair template, mCherry. This protocol would ideally be done side-by-side with the original transfection, which was reliant on transfection of 4 separate components. Ideally, by direct electroporation of the Cas9 RNA complex and the mCherry plasmid, efficiency of the mCherry insert and resulting p53 knockout will increase significantly.

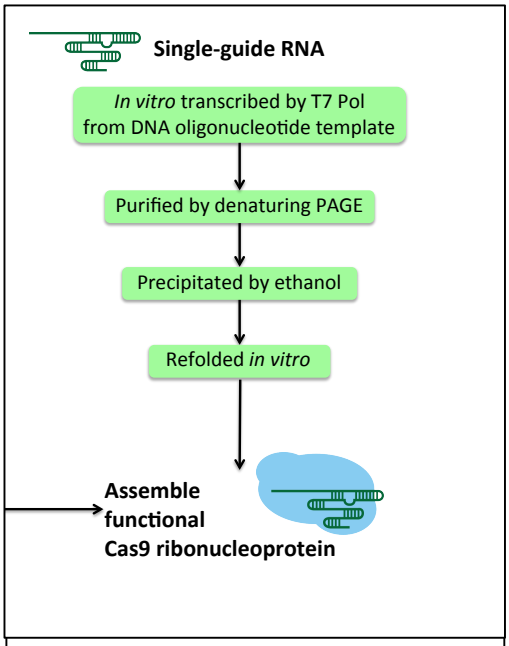


Figure 22. Flowchart of RNP assembly protocol (Doudna et al., 2014).

Methods

A549 cell culture and transfection

Day 1: Transfect Cells with crRNA, tracRNA, template plasmid, and Cas9 plasmid. Day 3: Select with Puromycin. Day 5: Remove selection media. Day 12-19: Sort cells for mCherry marker into single cell colonies in 96 well plates with conditioned media. Day 19-26: Look at single cell colonies to ensure one colony is in each well. Grow cells up enough to extract DNA and freeze down tubes, which takes at least 2-3 more weeks. Spencer lab imaged cells 16 photos/well to see which single cell colonies express mCherry.

Flow Cytometry

Cells were washed twice with PBS and harvested with trypsin. Trypsin was quenched and the solution was centrifuged at 1300rpm for 5 minutes. The cells were then resuspended in 1 mL of media and filtered through a CellTric 50 uM filter. The cells were transferred into 5 mL Falcon round-bottom polypropylene tubes. The Falcon tubes were inserted into a MoFlo Legacy Flow Cytometry machine and sorted at 561 nm.

pMJ915 Transformation into Rosetta 2 (DE3) cells

Added 30 uL of Rosetta cells per reaction. Pre-chilled BD Falcon polypropylene round-bottom tubes on ice. Thawed cells on ice, and added 50 ng of experimental DNA (pMJ915). Swirled the tube gently and incubated on ice for 30 minutes. Heat-pulsed the tubes in a 42℃ water bath for 30 seconds. Incubated on ice for 2 minutes. Added 970 uL of LB to tube and incubated at 37℃ on shaker for 1 hour. Plated 100 uL of 1 mL solution on pre-warmed plate. Spun down the remaining solution after transferring to 1.5 mL Eppendorf tube at less than 3000 rpm. Resuspended the pellet in 100uL of LB and plated on pre-warmed amp plates. Incubated overnight at 37℃ and selected colonies for induction the next day.

E. coli cell culture and Cas9 protein expression

Overnight culture was made by selecting colonies containing the Cas9 plasmid pMJ915 (containing antibiotic resistance and an IPTG inducible promoter) in a small sample of Terrific Broth (800 mL of distilled water, 12 g Tryptone, 24 g Yeast extract, 4 mL Glycerol, adjust to 900 mL with distilled H2O). Autoclaved and let cool to room temp, and adjusted to 1000 mL with filtered sterilized 100 mL solution of 0.17 M KH2PO4 and 0.72 M K2HPO4) (TB) overnight at 37℃. Ampicillin and chloramphenicol were added in final concentrations of 0.1 mg/mL to select for antibiotic resistant strains. The overnight culture was then transferred to a larger flask containing Terrific Broth and antibiotics and incubated at 37℃ for approximately 20 hours. The cells were subcultured at 1:100 ratios into 1 L of TB containing 100ug/ml of amp and 30 ug/ml of chloramphenicol. The cells were incubated for approximately 4 hours until their ocular density (OD600) reached 0.8-1.0. The culture was then chilled in an ice bath and IPTG was added to 0.5 mM to induce protein expression. The culture was grown overnight in an 18℃ shaker. The culture was transferred to 1-L centrifugal bottles and spun down at 3500g for 10 minutes at 4℃. The cells were resuspended in 50 mL or Buffer 1 (20 mM HEPES, pH 7.5, 1 M KCl, 10 % glycerol, 1 mM

TCEP and 0.1 % Triton X-100) per liter of culture. 1mM of PMSF was added. The suspension was stored at -80℃ until the purification process began.

Isolation of cell lysate

All purification steps were performed at 4℃. Cells were thawed under cold tap water and transferred to 100mL beaker containing a stir bar. The beaker was placed on ice. One tablet of Roche EDTA-free protease inhibitor was added per 50 mL of suspension. The cells were lysed via sonication (power output 5, pulse-on 1 seconds, pulse-off 1 seconds, for a total of 2 minutes) under constant stirring. The lysate was centrifuged at 15,000 g for 45 minutes.

Ni-NTA batch purification

1 mL of Qiagen Ni-NTA beads were equilibrated in a 50 mL conical tube by washing with water. The beads were centrifuged at 3000 g for 5 minutes to collect. A 10 uL lysate sample was taken for SDS-PAGE analysis. The remainder of the lysate incubated with the beads on an orbital shaker on low speed for 30 minutes. The tube was centrifuged at 3000 g for 5 minutes to collect the beads. The beads were washed with 20 column volumes of Buffer 1, with low-speed rocking for 5 minutes before centrifugation to collect the beads between each wash. The beads were then washed with 20 column volumes of Buffer 2 (20 mM HEPES, pH 7.5, 1 M KCl, 10 % glycerol, 1 mM TCEP and 10 mM imidazole, 0.2 um filtered and degassed). The protein was then eluted in Buffer 3 (20 mM HEPES, pH 7.5, 0.5 M KCl, 10 % glycerol, 1 mM TCEP and 300 mM imidazole, 0.2 um filtered and degassed). The elution was rocked in Buffer 3 for 15 minutes in 10 column volumes, equaling a total of 10 mL. Elution fractions were sampled and run on SDS-PAGE in Figure 14. The 10 mL of elution was condensed to 2 samples of 500uL using 100kDa Amicon filter cuvettes. TEV protease was added to cleave the 6xHis tag and MBP fusion at 4℃ overnight.

Ion Exchange Chromatography

A 5-mL HiTrap Heparin HP column was connected to a GE FPLC machine and the flow rate was set to 2 mL/min for the duration of the purification. The column was equilibrated with 10 column volumes of Buffer 4 (20 mM HEPES, pH 7.5, 300 mM KCl, 10 % glycerol and 1 mM TCEP). 1 mL of Ni-NTA Cas9 elution was diluted in 9 mL of Buffer 4 for a total of 10 mL to bring the KCl concentration to approximately 300 mM. The solution was then filtered with a 0.45 um syringe filter and injected into the column. The column was washed with 10 column volumes of Buffer 4 until the OD280nm reached baseline. The FPLC was programmed to elute Cas9 by a 20 column volume linear gradient to 100% of Buffer 5 (20 mM HEPES, pH 7.5, 1 M KCl, 10 % glycerol and 1 mM TCEP) at 2ml/min. Every third fraction throughout the entire elution and every fraction where peaks occurred were taken for SDS-PAGE analysis. The fractions were combined and concentrated by Amicon 100kDa filter.

Amicon 100kDa Concentration

Added up to 4.5 mL of elution to the top portion of the Amicon tube. Centrifuged at 3000 rpm for 5 minute pulses until 500 uL elution remained. Saved and sampled flow-through to run on SDS-

PAGE for analysis. Solutions with higher glycerol content will flow much slower through the membrane and require longer periods of centrifugation.

Coomassie stain

A 7.5% polyacrylamide SDS-PAGE gel was run with the samples. The gel was then placed in blue BioSafe Coomassie stain overnight. The gel was removed from the stain, washed in deionized water, and imaged.

Silver stain

A silver stain was used to visualize the protein content of the fractions collected after identifying major peaks of the Heparin Ion-Exchange column. Samples from each fraction were run in a 7.5% polyacrylamide SDS-PAGE gel. The gel was then transferred to a solution of 50% methanol for 10 minutes, then a 5% methanol solution for 10 minutes, a solution of 35 mM DTT for five minutes, and then a solution of .006M silver nitrate for 10 minutes. The gel was developed using a solution of 280 mM sodium carbonate anhydrous and quenched using powdered citric acid monohydrate.

Western blot

A western blot was used to confirm the presence of the final sample of Cas9. SDS-PAGE was run using a 7.5% polyacrylamide gel. The gel was transferred onto a nitrocellulose membrane at 0.25 mAmps and 100 V for 1 hour. The nitrocellulose membrane was blocked in 5% milk in TBST. The primary antibodies were then diluted in 5% milk in TBST 1:500 and the nitrocellulose membrane was left to incubate overnight. The primary antibody was an Anti-Cas9 purified mouse monoclonal IgG1κ in buffer containing 0.1 M Tris-Glycine (pH 7.4), 150 mM NaCl with 0.05% sodium azide from Millipore (MAC 133). After incubating in the primary overnight, the membrane was washed using a TBST wash buffer for three sessions ten minutes long each. Secondary mouse antibody bound with horseradish peroxidase was diluted 1:5000 in 5% milk in TBST. The nitrocellulose membrane incubated in the secondary antibody for 45 minutes before being washed for 10 minutes three times. The nitrocellulose membrane was then developed using peroxidase substrate for enhanced chemiluminescence (ECL).

References

1. A. Ran et al. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154, 1380–1389.
2. B. Shen et al. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat. Methods 11, 399–402.
3. C. Pourcel, G. Salvignol, G. Vergnaud. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology 151, 653–663.
4. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., … Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science (New York, N.Y.), 339(6121), 819–823. http://doi.org/10.1126/science.1231143
5. Cyranoski, D. (2015). A call by scientists to halt precision gene editing of DNA in human embryos would allow time to work out safety and ethical issues. *Nature*.
6. Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science.* Vol. 346. Is. 6213, 1077-1258096-9.
7. F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, E. Soria. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60, 174–182.
8. Feng Z, Lin M and Wu R. (2011). The Regulation of Aging and Longevity: A New and Complex Role of p53. Genes Cancer. 2:443-452.
9. Goudarzi, K. M., Nistér, M., & Lindström, M. S. (2014). mTOR inhibitors blunt the p53 response to nucleolar stress by regulating RPL11 and MDM2 levels.*Cancer Biology & Therapy*, *15*(11), 1499–1514. http://doi.org/10.4161/15384047.2014.955743
10. H. Ebina, N. Misawa, Y. Kanemura, Y. Koyanagi. (2013). Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. Sci. Rep. 3, 2510.
11. H.Yinetal. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat. Biotechnol. 32, 551–553.
12. J. P. Guilinger, D. B. Thompson, D. R. Liu. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat. Biotechnol. 32, 577–582.
13. J. van der Oost, E. R. Westra, R. N. Jackson, B. Wiedenheft. (2014). Unravelling the

structural and mechanistic basis of CRISPR-Cas systems. Nat. Rev. Microbiol. 12, 479–492.

14. Khoury, M. P., & Bourdon, J.-C. (2010). The Isoforms of the p53 Protein. *Cold Spring Harbor Perspectives in Biology*, *2*(3), a000927.
15. Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, *24*(6), 1012–1019.
16. Lin, S.-C., & Taatjes, D. J. (2013). ΔNp53 and aging. Aging, 5(10), 717-718.
17. Lin, S.-C., Karoly, E. D., & Taatjes, D. J. (2013). The human ΔNp53 isoform triggers metabolic and gene expression changes that activate mTOR and alter mitochondrial function. Aging Cell, 12(5), 863-872.
18. Lin, S., Staahl, B. T., Alla, R. K., & Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife*, *3*, e04766.
19. Marbach, A., & Bettenbrock, K. (2011). Lac operon induction in Escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. 157(1):82-8.
20. McClendon, T.B., Sullivan, M., Bernstein, K., & Yanowitz, J. (2016). Promotion of Homologous Recombination by SWS-1 in Complex with RAD-51 Paralogs in Caenorhabditis elegans. Genetics, 115(2).
21. Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S., Dohmae, N., … Nureki, O. (2014). Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*, *156*(5), 935–949.
22. Okorokov, A. L et al. (2006). The structure of p53 tumour suppressor protein reveals the basis for its functional plasticity. The EMBO Journal, 25(21), 5191-5200.
23. Pehar M, O'Riordan KJ, Burns-Cusato M, Andrzejewski ME, del Alcazar CG, Burger C, Scrable H, Puglielli L (2010). Altered longevity-assurance activity of p53:p44 in the mouse causes memory loss, neurodegeneration and premature death. Aging Cell, 9:174–190
24. S. A. Shah, S. Erdmann, F. J. M. Mojica, R. A. Garrett. (2013). Protospacer recognition motifs: Mixed identities and functional diversity. RNA Biol. 10, 891–899.
25. S. W. Cho et al. (2014). Analysis of off-target effects of CRISPR/Cas- derived RNA-

guided endonucleases and nickases. Genome Res. 24, 132–141.

26. Schumann, K., Lin, S., Boyer, E., Simeonov, D. R., Subramaniam, M., Gate, R. E., et al. (2015). Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(33), 10437–10442.

27. Soussi T: Handbook of p53 mutation in cell lines. 2010,http://p53.free.fr/.

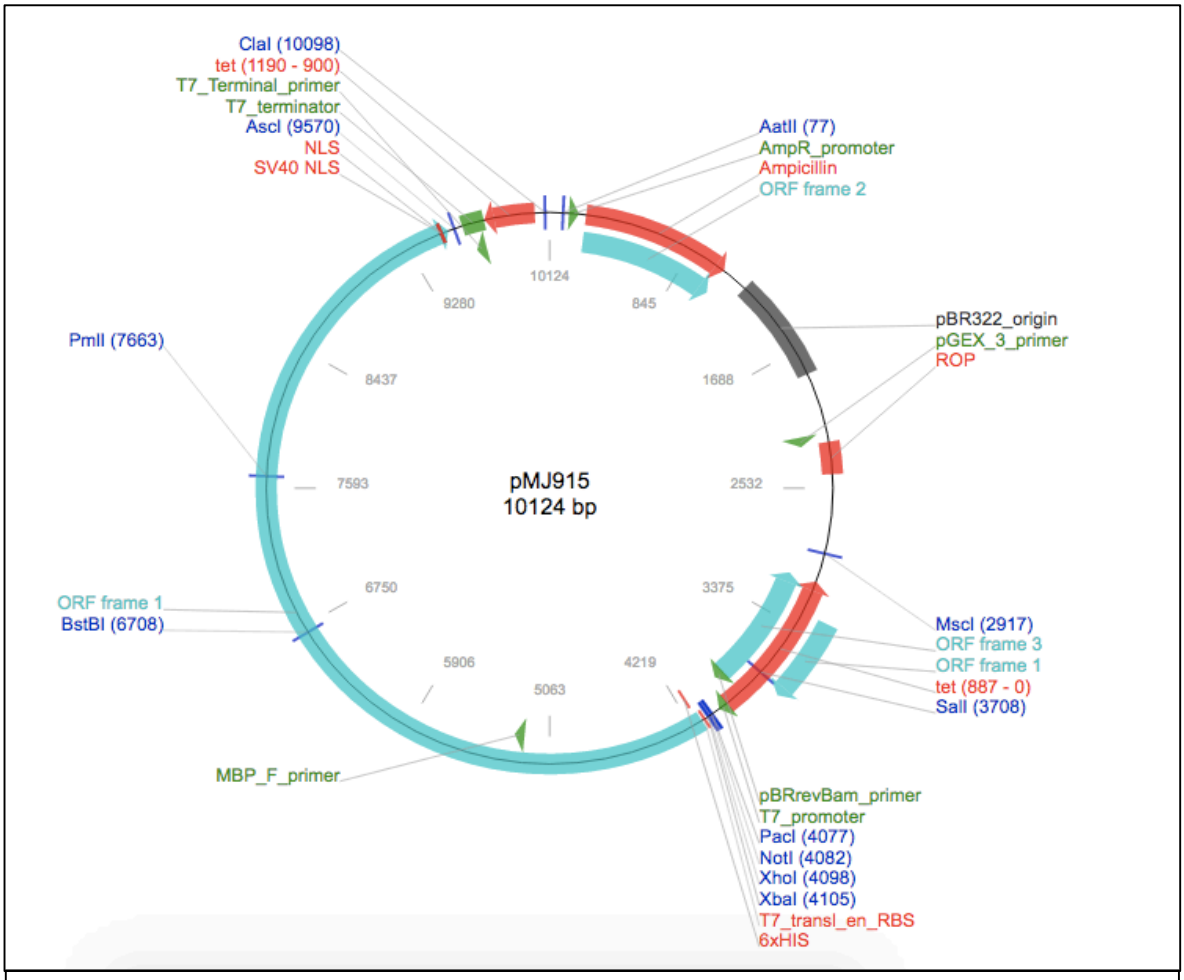
28. Y. Fu et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. 31, 822–826.

29. Y. Fu, J. D. Sander, D. Reyon, V. M. Cascio, J. K. Joung. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. 32, 279–284.

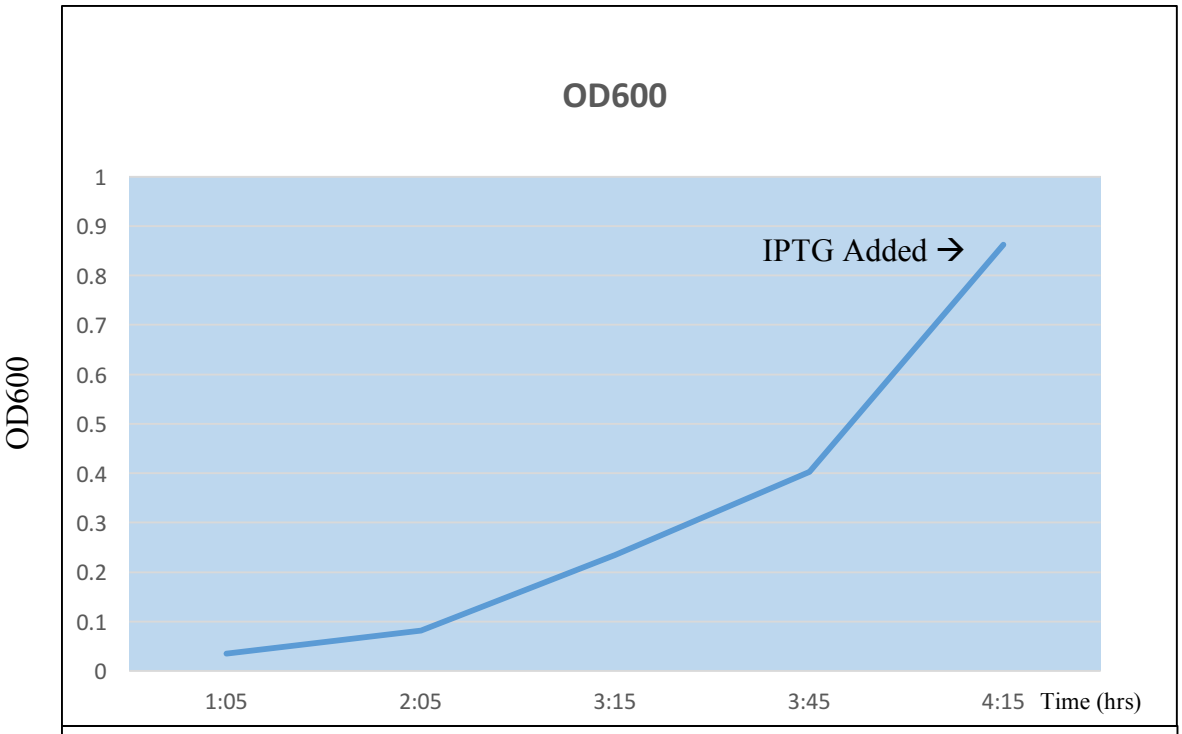
30. Y. Zhao et al. (2014). Sequence-specific inhibition of microRNA via CRISPR/CRISPRi system. Sci. Rep. 4, 3943.

31. Zheng J, Lang Y, Zhang Q, Cui D, Sun H, Jiang L, Chen Z, Zhang R, Gao Y, Tian W, et al. (2015). Structure of human MDM2 complexed with RPL11 reveals the molecular basis of p53 activation. Genes Dev; 29:1524–1534.

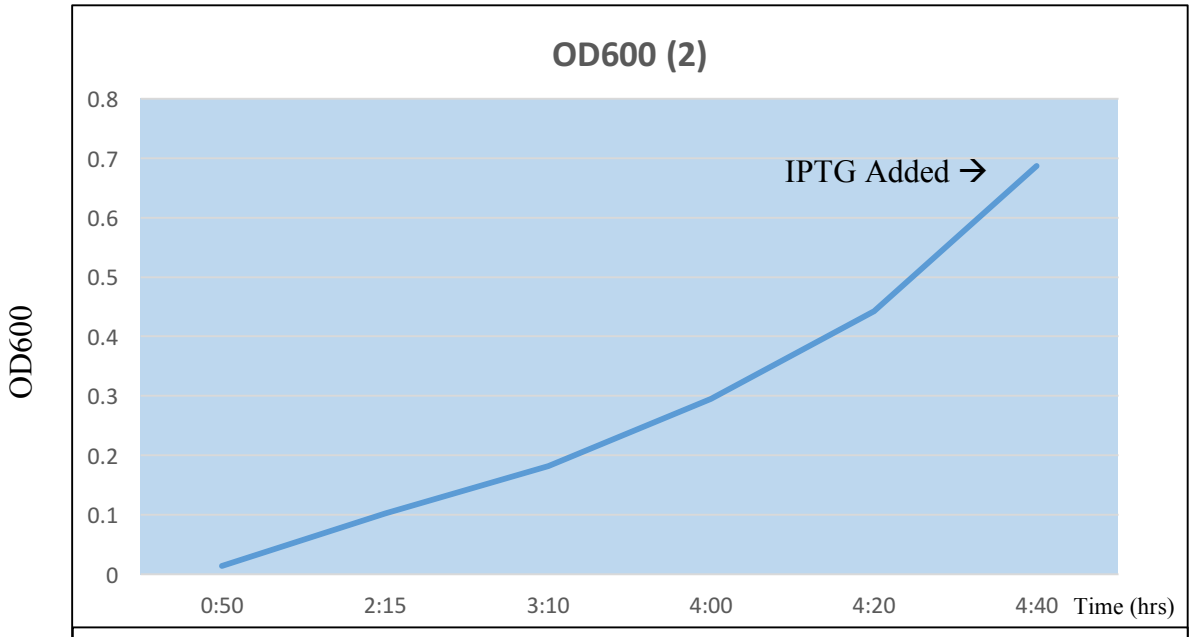
Appendix



Appendix A. Map of pMJ915 plasmid.



Appendix B. First growth curve of Rosetta 2 (DE3) cells and induction time.



Appendix C. Second growth curve of Rosetta 2 (DE3) cells and induction time.

Acknowledgments

I would like to thank Dylan Taatjes for his invaluable guidance and for providing me an opportunity to work in his lab through the duration of this project. Thank you also to Cecie Levandowski for her continuous mentorship and support. Thank you to the entirety of the Taatjes Lab: Ben Allen, Charli Fant, Jonathan Rubin, Radhika Rawat, and Meagan Paul for their help and advice. Thank you especially to Abigail Horn and the Goodrich Lab for sharing supplies and information, the Ahn and Spencer Lab for lending their equipment, Theresa Nahreini from cell culture, and the Doudna Lab for their insights. Finally, thanks to the Howard Hughes Medical Institute and CU Boulder's Biological Science Initiative for funding and various opportunities. This experience has deeply furthered my understanding about fundamental biochemistry lab techniques, and it couldn't have been possible without each of these groups and individuals.