University of Mississippi

eGrove

2010

Replacement of the Cro Repressor in a Negative Feedback Circuit with a Slow Dimerizing Variant

Lauren Michelle Rowe

Follow this and additional works at: https://egrove.olemiss.edu/hon_thesis

Recommended Citation

Rowe, Lauren Michelle, "Replacement of the Cro Repressor in a Negative Feedback Circuit with a Slow Dimerizing Variant" (2010). *Honors Theses*. 2254.

https://egrove.olemiss.edu/hon_thesis/2254

Replacement of the Cro Repressor in a Negative Feedback Circuit with a
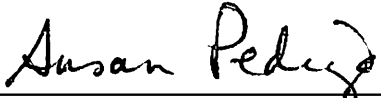Slow Dimerizing Variant

by
Lauren Michelle Rowe

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of
the requirements of the Sally McDonnell Barksdale Honors College.
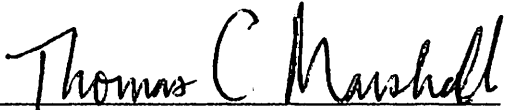

Oxford
May 2010

Approved by

Advisor: Dr. Michael C. Mossing

Reader: Dr. Susan D. Pedigo

Reader: Dr. Thomas C. Marshall

Acknowledgements

I would first like to thank Dr. Michael Mossing for allowing me to work in his lab for the last year and a half. His guidance has been invaluable, and I thank him for sharing his wealth of knowledge with me. This project may not have come to fruition without his unyielding patience and understanding.

My appreciation also extends to Jackie Harris. I thank her for showing me the ropes in the lab and helping me to accomplish my research goals. Most of all, I am thankful of the support she provided as a friend.

Words cannot begin to express the gratitude I feel toward my friends and family that are closest to me. Thank you for putting up with me and continuously providing encouragement throughout this whole process.

ABSTRACT

LAUREN M. ROWE: Replacement of the Cro Repressor in a Negative Feedback Circuit
with a Slow Dimerizing Variant

The Cro protein is a transcriptional regulator that plays an active role in the switch

between the lysogenic and lytic viral cycles of bacteriophage lambda. Functional Cro

proteins exist as dimers composed of two identical Cro subunits. It is suspected that the

dimerization rate and the affinity of Cro's subunits play a part in determining the activity

of the protein in living cells. Synthetic genetic circuits have been constructed in order to

determine the repression dynamics of Cro *in vivo*. These circuits are used to observe the

ability of wild-type Cro and a pre-dimerized variant named scCro to modulate the

expression of a reporter gene as compared to the unrepressed circuit. The unrepressed

circuit that is used as a control is named the OPO circuit, and it contains no copy of the

cro gene.

The goal of this work was to construct an additional simple circuit with another

variant *cro* gene (*croF58W*) in order to compare its *in vivo* regulatory dynamics in living

cells with the binding dynamics that have been observed in a purified system. Wild-type

Cro proteins have a phenylalanine at position fifty-eight, but CroF58W proteins have a

tryptophan at this position. *In vitro* studies done using CroF58W variants have shown that

CroF58W proteins are slower to dimerize but are thermodynamically more stable.

Steady-state assays conducted *in vivo* with the CroF58W circuit showed an increased level

of β-galactosidase compared not only to the wild-type Cro and scCro circuits but also to

our control OPO circuit.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

attP	phage attachment site
CAM	chloramphenicol
cI	gene for Lambda repressor
cro	gene for Cro repressor
croF58W	gene for the CroF58W variant
CroF58W	variant Cro protein that contains a tryptophan at position 58
DNA	deoxyribonucleic acid
IPTG	isopropyl-β-D-thio-galactosidase
lacI	genetic sequence for Lac repressor
LacR	Lac repressor protein
lacZ	gene encoding β-galactosidase
mRNA	messenger RNA
OD_{600}	optical density at a wavelength 600nm
ONPG	*o*-Nitrophenyl-β-D-Galactoside
O_R	right operator of the lambda virus
oriRγ	gamma origin of replication
PCR	polymerase chain reaction
P_R	right promoter of bacteriophage lambda
P_{RM}	the repressor maintenance promoter of lambda
RNA	ribonucleic acid
scCro	single-chain Cro repressor protein
scCro	gene encoding single-chain repressor protein

I. INTRODUCTION

When the conductor of an orchestra steps onto the stage, he or she stands solely before a score of music and anywhere between fifty and two hundred musicians. The conductor along with his musicians comprise a network. The music that the audience hears is a product of the network, and the musicians along with the conductor are responsible for making real-time adjustments to the sounds being emitted from the orchestra. The tempo may either need to be faster or slower; the trombone section may be too loud while the clarinet section is too faint. Each musician must listen to his or her neighbor and the rest of the ensemble in order to make adjustments to the pitch or loudness of the instrument. Some networks do not function correctly, and the music comes grinding to a cacophonous halt. Other networks work perfectly and provide harmonious melodies that bring audiences to their feet. Whether the orchestra is successful or not is largely dependent upon the activities of the conductor and the musicians who not only comprise the network but work to regulate it as well.

Now, think of the music not as a network of notes and sounds but as a network of proteins within a cell. Each score of music serves as the DNA of the system. The instrumentalist reads the notes off of the score and presses down the right fingers to produce sounds in the right musical key and rhythm. The resulting notes and sounds feed back into the ears of the each instrumentalist and the conductor and work to influence the succeeding sounds that are produced. RNA polymerase transcribes genes from DNA in

order to produce proteins that have the proper sequence and are functional. Each individual protein is part of a larger network of proteins in a cell that functions to give the cell a particular phenotype. The transcription of various genes is either increased or decreased according to the proteins that are present and feedback into the network at any given time. For instance, a certain protein may bind to DNA in order to increase its own production or prevent its further production. This would be an example of a negative feedback loop. The overarching goal of this research is to take a small section of a network of proteins (one instrumentalist from the orchestra) and slightly change the final product—the final protein—in order to observe the change in the dynamics of the network. By observing these changes, the knowledge of the role of that particular protein may be better developed.

A. Gene Expression and Regulation

The presence of different cell types within a multicellular organism results from one set of RNA and resulting protein molecules being accumulated in one cell and a different set accumulating in another. The human genome encodes approximately 25,000 genes with studies suggesting that a typical human cell expresses 30-60% of those genes at any one time (Alberts et al., 2008). Genes begin as DNA, a polymer of deoxy-ribonucleotides, are transcribed to RNA, a polymer of ribonucleotides, and are finally translated to proteins, a polymer of amino acids. Any aspect of this sequence of events can be acted upon in order to regulate gene expression. First, a cell can control how often a gene is transcribed from DNA. This is referred to as transcriptional control. Once a gene is transcribed, the mRNA may or may not be transported to the cytosol for

transcription. After mRNA reaches the cytosol, a cell can then determine which transcripts are translated into proteins by ribosomes. This is called translational control. After a protein is produced, it has the potential to be activated, inactivated, degraded, or specifically located in order to help with gene regulation and expression. Transcriptional control is usually the primary point for the regulation of most genes because it ensures that the cell will not make materials that it does not need.

B. Bacteriophage Lambda (λ)

Although Bacteriophage λ contains less than 50 genes, the λ lifecycle contains a simple binary developmental switch. Bacteriophage λ contains several well-studied gene-regulatory circuits that have served as an important model system in understanding simple regulatory mechanisms that are also found in more complex organisms (Schubert, 2007). Phage λ consists of one DNA molecule that contains about 50,000 base pairs. The DNA is wrapped in a protein coat that includes fifteen proteins and is divided into a head and tail section. The DNA molecule is located in the head section of the phage. Through its tail, λ phage attaches to the surface of an Escherichia coli (E. coli) cell, creates a hole in the cell wall of the bacterium, and then injects its DNA into the cell. At this point, the cell enters one of two viral cycles—the lytic cycle or the lysogenic cycle (Ptashne, 2004).

C. The Lytic and Lysogenic Cycles

If a cell enters the lytic cycle upon infection, the λ phage DNA signals a pattern of gene expression in which certain viral and host cell genes are turned off while others are turned on. The λ chromosome is replicated by the host cell, and new phage heads and tails are assembled within the host cell. Approximately forty-five minutes after a cell has been infected and entered the lytic cycle, the cell lyses and releases around one hundred new λ viruses (Ptashne, 2004).

In the lysogenic cycle, the viral DNA from the λ phage is integrated into the host chromosome and becomes a prophage. A protein called cI repressor is responsible for repressing the lytic genes of the viral DNA. This results in a stable condition where the host cell continues to grow and divide while the prophage DNA is passively replicated and passed on to the new daughter cells (Ptashne, 2004). A lysogenic cell is able to enter into the lytic cycle through a switch called prophage induction (Atsumi, 2006). This switch is described as epigenetic because the DNA of the virus is not actually altered in the process of flipping on the switch. Rather, changes in the network of proteins are initiated causing the viral DNA to be transcribed in a different way in order to produce a different set of proteins. DNA damage in the host cell induces an SOS response in which ten to twenty bacterial genes are turned on in order to help the bacterium survive the DNA damage (Ptashne, 2004). The activation of the RecA protein is a direct result of the initiation of the SOS response. This protein induces an autocatalytic response from cI repressor leaving it inactive. This allows for the viral lytic genes to be transcribed. One such lytic gene that plays a major role in this switch is called the *cro* gene which produces Cro protein (Atsumi, 2006).

4

D. Cro Repressor

Cro repressor, an early lytic protein produced by bacteriophage λ, is one of the simplest DNA binding proteins. It is part of a negative feedback loop; this means that the production of Cro protein feeds back and prevents the further transcription of its gene. The protein consists of two sixty-six amino acid residue subunits whose secondary structure consists of three α-helices and three β-sheets (Satumba, 2002; Rupert, 2000). Residues 5 and 6 make up β1 with α1 consisting of residues 7-14. α2 and α3 include residues 16-23 and 27-35 respectively. β2 includes residues 40-45, and residues 49 through 56 comprise β3. When Cro subunits dimerize, the β3 structures of each subunit interact to form an anti-parallel β-ribbon between which forms part of the dimer interface (Rupert, 2000).

Current research involving Cro focuses on determining the role that Cro plays in the viral switch from the lysogenic to the lytic cycle. Research conducted by Satumba and Mossing (2002) found that the dimerization of Cro occurs at a relatively slow rate. This research also found that a weak affinity existed between the subunits of Cro dimers. Further research is being conducted to determine if Cro's slow and weak dimerization is an important aspect of prophage induction.

In order to further investigate Cro's dimerization dynamics, several different Cro variant proteins have been constructed. In addition to the wild type *cro* gene, research has been done on a mutant gene named *scCro* (for single-chain Cro). The *single-chain Cro (scCro)* gene was constructed from two copies of the *cro* gene that are linked together by twenty-four nucleotides (Jana, Hazbun, Fields, & Mossing, 1998). The scCro protein contains two direct repeats of the Cro peptide sequence linked together by eight

hydrophilic amino acids. As a result of the fact that scCro does not have to dimerize in order to bind to the lambda operator, scCro is able to bind to operator DNA more quickly and with higher affinity than wild-type Cro.

F. CroF58W

Another of variant of Cro that has been studied is known as CroF58W. In this version of the Cro protein, a tryptophan has been substituted for a phenylalanine at residue fifty-eight. The F58 phenyl side-chain serves as an anchor for the interface supplying 180 square angstroms of buried hydrophobic surface area. Despite the fact that it is a key component for the interface, its exact position is somewhat variable in different crystal structures. The substitution of a tryptophan at position 58 increases the stability of the Cro dimers (Satumba & Mossing, 2002). The added stability results from tryptophan penetrating further into the hydrophobic core of the dimer which in turn fills a cavity that is present in wild-type Cro between the first and second α-helices and the β-sheet. The F58W substitution results in the reduction of this internal cavity by approximately thirteen cubed angstroms. This reduction corresponds to a stability increase of approximately 0.4 kcal/mol per monomer (Rupert, 2000).

Research involving CroF58W mutants has thus far been done *in vitro* to determine dimerization rates and DNA binding affinities (Rupert, 2000; Satumba & Mossing, 2002). It is important to progress towards research involving CroF58W *in vivo* so that the actual dynamics may be better understood. The purpose of the present research was to construct a circuit with CroF58W that could be used to determine dimerization rates and DNA binding dynamics in living cells.

E. Circuit Constructs: Strains and Plasmids

In order to study Cro variants in living cells, simplified genetic circuits in the form of plasmids (small circular pieces of DNA) were constructed. The human genome contains around 25,000 genes. Bacteriophage λ contains fewer than fifty genes. The plasmids we use to study Cro have been simplified to contain just three genes. Plasmids were inserted into host E. coli cells or integrated into the E. coli chromosome. The E. coli host cells used in the lab were BW25113, BW25141, and BW25142. All strains lack the gene encoding β-galactosidase (the *lacZ* gene) and over-express Lac repressor encoded by the *lacI* gene. The chief difference lies in their ability to serve as hosts for replication of plasmids containing the *oriR*γ replication origin from the R6K plasmid. BW25113 cannot support plasmid replication but can host a circuit construct in a single copy at a unique chromosomal location. The BW25113 cells also contain a Lambda Int plasmid that helps a circuit construct integrate into the host cell chromosome (Haldimann & Wanner, 2001). Relevant phenotypes of each strain are shown in Table 1.

Strain	*pir* gene	Plasmid Copy #	Lac phenotype	Lambda Int
BW25113	--	0	Δ*lacZ, lacI*	(pInt-ts)
BW25141	*pir+*	15	Δ*lacZ, lacI*	--
BW25142	*pir-116*	250	Δ*lacZ, lacI*	--

Table 1: Phenotypes of E. coli Strains. The table above shows the phenotypes of the three E. coli host cell strains used in the lab. The version of the copy of the *pir* gene that is found in each strain along with the plasmid copy number of each strain are listed.

The basis for all Cro variant constructs was the pAH125 plasmid, which contains a γ replication origin (*oriR*γ) that requires the *trans*-acting Π protein for replication. The *pir* gene, which is found on the chromosome of the host E. coli cell, encodes this protein.

The BW25141 host cells contain the *pir+* gene that produces a medium plasmid copy number in the cells—about fifteen plasmids per cell. The BW25142 host cells have a high plasmid copy number as a result of containing the *pir-116* gene, which is a *pir* mutant that produces about 250 plasmids per cell. In addition to the γ replication origin and the lambda phage attachment (*attP*) site, the pAH125 plasmid also contains a selective kanamycin marker that confers resistance to the antibiotic kanamycin (Haldimann & Wanner, 2001).

Another important feature of the pAH125 plasmid is the *lacZ* gene, which - encodes β-galactosidase and serves as a reporter gene since it is absent in the chromosome of the host E. coli cells (Haldimann & Wanner, 2001). The lac operator controls the *lacZ* gene, and the product of the *lacZ* gene, β-galactosidase, is a protein that cleaves the sugar lactose into galactose and glucose. Our circuits also contain another lac gene—the *lacI* gene—which encodes Lac repressor (LacR). In the lac operon (and our circuits) the *lacI* gene is under the control of its own promoter, P_i. LacR transcripts are encoded by the *lacI* gene, and the resulting repressor proteins bind to the lac operator site and prevent RNA polymerase from binding to the lambda promoter. As a result, very low levels of β-galactosidase are present in the cell. Isopropyl-β,D-thiogalactoside (IPTG), binds to Lac repressor and reduces its affinity for the operator (Wilson, Zahn, Swint-Cruse, & Matthews, 2007).

The presence of the *lacZ* gene in our circuits is important because it is easier to assay and monitor the activity of β-galactosidase than Cro and its variants. In our circuits, the *cro* gene (wild-type or variant depending on the construct) immediately precedes the *lacZ* gene. A control unit named OPO directs transcription of both of these genes. The

OPO unit consists of the lambda operator, lambda promoter (P_R) and the lac operator. As a result of the actions of the *lacI* gene, our circuits are initially repressed—neither the *cro* gene nor the *lacZ* gene are transcribed due to the binding of LacR proteins to the lac operator. In order to induce the cells and turn "on" the transcription of *cro* and *lacZ*, we add IPTG, which removes LacR and results in P_R being open for RNA polymerase. Once *cro* is transcribed, Cro proteins are produced and can assemble and bind to the lambda operator. The binding of Cro to the lambda operator inhibits the binding of RNA polymerase and reduces the transcription of both the *cro* gene and the *lacZ* gene. By monitoring the presence of β-galactosidase after adding IPTG and inducing the cells, we can observe the activity of Cro and its variants. β-Galactosidase is detected by adding *o*-Nitrophenyl-β-D-galactoside to the sample being assayed. *o*-Nitrophenyl-β-D-galactoside is cleaved by β-galactosidase into galactose and *o*-nitrophenol. The latter of these two compounds has a yellow color. The intensity of the yellow color can be measured at a wavelength of 420 nm using a spectrometer, and this measure can then be used to calculate the amount of β-galactosidase present in the sample (Miller, 1992).

Currently assays carried out in the lab involve three genetic circuits: pAH125-lacI-OPO-lacZ (OPO), pAH125-lacI-Cro-lacZ (Cro), and pAH125-lacI-scCro-lacZ (scCro). OPO serves as the control circuit and contains no *cro* gene. After cells containing this plasmid are induced, the operator is always completely open and the rate of transcription is constant (Figure 1).

The Cro circuit contains the wild-type *cro* gene, and produces Cro proteins that must dimerize before binding to the lambda operator and repressing the circuit. In these cells, there is a burst of transcription from the circuit immediately after induction with

9

IPTG followed by a decrease in transcription as Cro proteins dimerize and bind to the

lambda operator.

As described earlier, the scCro circuit contains the variant *scCro* gene, which

encodes a pre-dimerized version of Cro. As in the other circuits, this gene directly

precedes the *lacZ* gene.

The purpose of this current work is to create a synthetic circuit identical to the

current Cro and scCro circuits but with the *croF58W* gene. This circuit can then be used

in β-galactosidase assays to determine the repression dynamics of the CroF58W proteins

in living cells. These dynamics can then be compared to the behavior of Cro and scCro

in vivo.
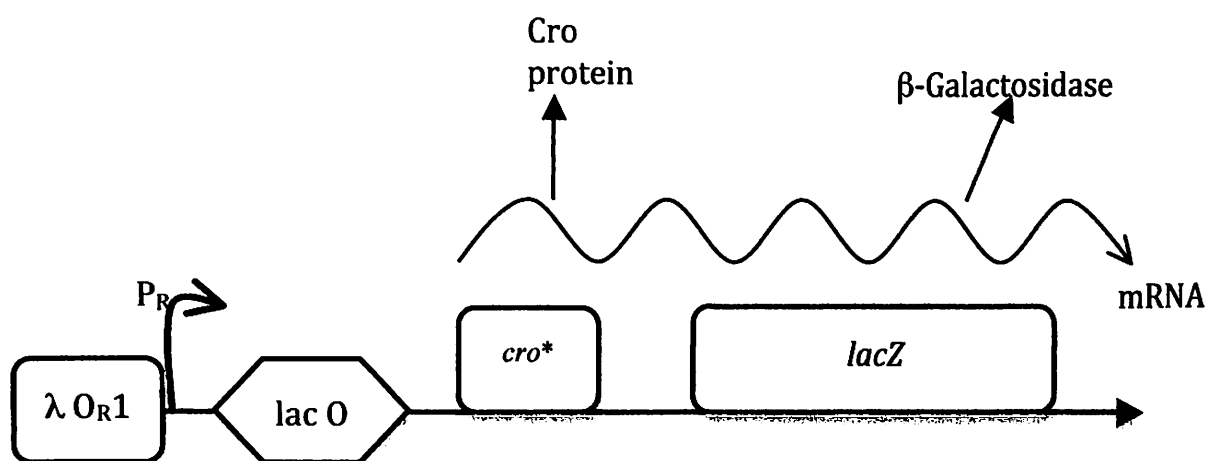
Figure 1: Basic Sequence of Cro Circuit DNA. The OPO control unit directly precedes the *cro* gene (which can be *cro*, *scCro* or *croF58W*) and the *lacZ* gene. The *cro* gene encodes Cro proteins while the *lacZ* gene encodes β-galactosidase. In the OPO circuit, no *cro* gene is present. The *lacZ* gene immediately precedes the lac operator

10

II. METHODS

A. Bacterial Cultures

When working with bacteria, it is important to begin with a single colony to ensure that all of the bacterial cells are genetically identical. In order to obtain single colonies, bacterial strains were streaked from frozen stocks stored in a -70°C freezer onto Luria broth (LB) agar plates containing 50 μg/mL of either kanamycin or ampicillin. These plates were then placed in a 30°C incubator and allowed to grow overnight. For the cloning procedures, a single colony from the LB plate was then selected using a sterile wooden stick and used to inoculate 5 mL of liquid LB media in a sterile 25 mL test tube. For the β-galactosidase assays, a single colony from the LB plate was selected and used to inoculate 5 mL of liquid M9 media in a sterile 25 mL test tube. These test tubes were then placed in a roller drum in the 30°C incubator and left to grow overnight.

B. Plasmid DNA Purification

DNA from all strains was purified using the Promega Wizard® *Plus* DNA Purification System and accompanying protocol. The bacterial cells were first pelleted and then re-suspended using the included Cell Re-suspension Solution. The cells were then lysed with the Cell Lysis Solution. Neutralization solution was then added and the clear lysate was transferred to provided Minicolumn/syringe assemblies that contained

11

resin and were attached to a vacuum pump. The stopcocks were opened as the vacuum was applied, and the solution was pulled through the column. A column wash solution containing ethanol was then added to the column and pulled through by the vacuum. The vacuum was left on for an addition 30 seconds after the liquid was gone in order to dry the resin. The Minicolumn was transferred to a 1.5 mL micro-centrifuge tube, and nuclease-free water was added to the Minicolumn. After waiting 1 minute, the tube with the Minicolumn was centrifuged. The Minicolumn was then discarded and the DNA stored in the refrigerator.

C. DNA amplification by the Polymerase Chain Reaction

During the initial cloning procedure, polymerase chain reactions (PCR) were run using 51 µL reaction volumes. Each reaction volume was composed of 45 µL of a PCR master mix which contained 40 µL of 5X buffer, 4 µL of dNTPs, 132 µL of nuclease-free water, and 1 µL of Taq buffer. The remaining 6 µL of each reaction volume included 5 µL of PCR primers and 1 µL of template DNA. The PCR parameters where as follows: a 5-minute hold at 95°C, 25 cycles of 1 minute at 95°C, 1 minute at 60°C, and 1 minute at 72°C, and finally a 5-minute hold at 72°C.

The purification of PCR products began with adding a volume of DNA binding solution that was equal to the volume of the PCR product. This mix was then transferred to a spin cap tube and centrifuged for 30 seconds. The spin cup was conserved while the supernatent was discarded. A 1X wash buffer was then added to the spin cup, and the cup was centrifuged another 30 seconds. After discarding the supernatent once again, the cup was spun for another 30 seconds. The spin cup was then transferred to a clean tube,

and the DNA was eluted using TE buffer. This mixture was allowed to sit at room temperature for five minutes and then centrifuged for 30 seconds. The spin cup was then discarded, and the purified PCR product conserved in the supernatent.

D. Restriction Enzyme Analysis and DNA Ligation

Enzyme digests were carried out using BglII, HaeIII, HindIII, and XhoI restriction enzymes from New England BioLabs. Digestion reactions were carried out according to New England BioLabs protocols. After mixing each reaction volume, all digestion reactions were allowed to incubate with gentle shaking for 1 hour in a 37°C incubator. The ligation of cut DNA was performed using New England BioLabs Quick Ligation Kit, which included 2X Ligase Buffer and T4 DNA ligase. With this kit, ligation reactions were carried out at room temperature for five minutes.

E. Bacterial Transformation with Plasmid DNA

Two different protocols were used to carry out transformations during the cloning procedure. Transformations using BW25141 competent E. coli cells were completed according to the New England BioLabs High Efficiency Transformation protocol. For this protocol, 50 μL of competent cells per reaction were thawed on ice and then combined with 2 μL of DNA by gentle pipetting. These mixtures were then incubated on ice for thirty minutes before being heat shocked for 2 minutes in a 42°C water bath. After chilling on ice for 5 minutes, the cell mixtures were combined with 950 μL of room temperature LB media. These mixtures were then incubated at 37°C with shaking at 250 rpm. After one hour had elapsed, 100 μL of each cell mixture were spread onto a LB

13

agar plate containing 50 µg/mL kanamycin. These plates were then incubated overnight at 37°C.

Transformations in which RapidTrans TAM1 competent E. coli cells were used involved the protocol provided by Active Motif that accompanied the competent cells. These competent cells came pre-packaged in 50 µL volumes and had to be thawed on ice. Once thawed, 4 µL of the ligation mixture was added to each reaction tube. These cell mixtures were then incubated on ice for thirty minutes, heat shocked at 42°C for thirty seconds, and incubated on ice for another two minutes. Following these steps, 250 µL of SOC medium was added to each cell mixture. The remaining steps of the protocol mirrored those of the other protocol with one hour of incubation at 37°C followed by the spreading of 100 µL of cells onto agar plates and incubation overnight.

F. Gel Electrophoresis

Many electrophoresis gels were run throughout the cloning procedure in order to verify the mere presence of DNA, the size of a PCR fragment, or the size of a restriction fragment. All gels that were prepared and used were 1.5% agarose. For 25 mL gels, 0.375 grams of agarose were dissolved into 25 mL of TAE electrophoresis buffer, and 1 µL of Ethidium Bromide was then added. For 200 mL gels, 3 g of agarose were dissolved into 200 mL of TAE electrophoresis buffer. For these larger gels, 5 µL of Ethidium Bromide was added. After the gels were properly mixed, the liquid was poured into a gel apparatus and combs were placed into the liquid in order to create lanes. Each gel was allowed to solidify for at least thirty minutes up to an hour. The combs were then removed and the samples loaded. Each loaded sample was 10 µL total volume and

consisted of 8 μL of DNA sample with 2 μL of GoTaq Loading Dye. The 25 mL gels were allowed to run for twenty minutes at 200V while the 200 mL gels were allowed to run for forty-five minutes at 200V.

G. β-Galactosidase Assays

After the verification of CroF58W was complete, steady-state β-galactosidase assays were carried out using the medium-copy number strains of each of our four constructs: OPO, Cro, scCro, and CroF58W. The protocol used was identical to the one laid out in Miller (1992). Overnight cultures of each strain were grown first. One culture of each strain was grown in pure liquid M9 media. A culture of each strain was also grown in liquid M9 media containing the inducer IPTG at a concentration of 1 mM. Due to the lack of inducer in the first set of cultures, those cells were said to be "un-induced." The cells that did grow in the presence of the inducer were said to be "induced." For each strain, four drops of the overnight culture were then added to 5 mL of fresh culture and allowed to shake at 37°C until a cell density of 2-5 x 10^{-8} cells/mL had been reached. This density translated to an optical density at 600 nm (OD_{600}) of 0.28-0.70. The cultures were then cooled on ice for twenty minutes in order to prevent further growth. The growth of the culture is stopped in this absorbance range because that is when the culture is said to be at steady state: the number of new cells equals the number of cells that are dying. One milliliter of each culture was then withdrawn, and the OD_{600} was measured and recorded.

Next, 100 μL of each culture were combined with 900 μL of the assay medium Z buffer. In order to open the cells, two drops of chloroform and one drop of a 0.1% sodium

15

dodecyl sulfate solution were added to each sample tube. The tubes were vortexed for

ten seconds and then placed in a 28°C water bath for five minutes. Next, the reaction was

started by adding 0.2 mL of ONPG to each tube and then shaking the tubes for a few

seconds. The reaction was allowed to proceed until sufficient yellow color had

developed. The time that had elapsed since the beginning of the reaction with ONPG was

recorded, and the reaction was stopped by adding 0.5 mL of a 1 M sodium carbonate

solution. Sodium carbonate also helps to intensify the yellow color. Finally, the OD_{420}

was measured and recorded for each sample tube. This reading reflected a combination

of the absorbance by o-nitrophenol as well as light scattering by cell debris. The

contribution of the absorbance reading from light scattering is corrected for in the results

by obtaining a reading of the absorbance at 550 nm. At this wavelength, any absorbance

recorded is the result of light scatter; o-nitrophenol does not absorb light at this

wavelength. Light scattering 420 nm is equal to the OD_{550} multiplied by a factor of 1.75.

By subtracting the light scattering correction factor from the OD_{420} reading, the

absorbance due to o-nitrophenol is obtained.

dodecyl sulfate solution were added to each sample tube. The tubes were vortexed for ten seconds and then placed in a 28°C water bath for five minutes. Next, the reaction was started by adding 0.2 mL of ONPG to each tube and then shaking the tubes for a few seconds. The reaction was allowed to proceed until sufficient yellow color had developed. The time that had elapsed since the beginning of the reaction with ONPG was recorded, and the reaction was stopped by adding 0.5 mL of a 1 M sodium carbonate solution. Sodium carbonate also helps to intensify the yellow color. Finally, the OD_{420} was measured and recorded for each sample tube. This reading reflected a combination of the absorbance by o-nitrophenol as well as light scattering by cell debris. The contribution of the absorbance reading from light scattering is corrected for in the results by obtaining a reading of the absorbance at 550 nm. At this wavelength, any absorbance recorded is the result of light scatter; o-nitrophenol does not absorb light at this wavelength. Light scattering 420 nm is equal to the OD_{550} multiplied by a factor of 1.75. By subtracting the light scattering correction factor from the OD_{420} reading, the absorbance due to o-nitrophenol is obtained.

III. RESULTS

A. Construction of the Plasmid - pAH125-laci-OPO-CroF58W-lacZ

Two bacterial plasmids were used as the starting materials for the construction of

CroF58W—pUCroRSF58W (Figure 2) and pAH125-laci-OPO-scCro-lacZ (Figure 3).

Cells from frozen stocks of strain 1 (pUCroRSF58W) and strain 2 (pAH125-laci-OPO-

scCro-lacZ) were streaked on LB agar plus 50 µg per mL ampicillin for Strain 1 and plus

50 µg per mL kanamycin for Strain 2. Liquid cultures of these two bacterial strains were

obtained after a single colony was selected from the LB agar plates. The DNA from

these liquid cultures was then purified.

The next step of cloning involved the PCR amplification of the portion of

pUCroRSF58W DNA containing the segment of the Cro gene with the mutation at

position 58. MMUM052607c was the PCR primer that served as the top primer, and the

bottom PCR primer was MMUM0526d (Figure A-1). PCR using these two primers

produced a 308 base pair fragment beginning fifty base pairs upstream of the CroF58W

gene and ending twenty-five base pairs downstream of the beginning of the lacZ reporter

gene (Figure 4). Verification of the appropriately sized PCR product was carried out

using gel electrophoresis. The 5' tails of both PCR primers were mismatched and caused

the sequence of the plasmid to be changed to the sequence of the primers. This resulted

in the creation of an XhoI restriction site five base pairs from the end of the PCR

fragment (Figure A-2).
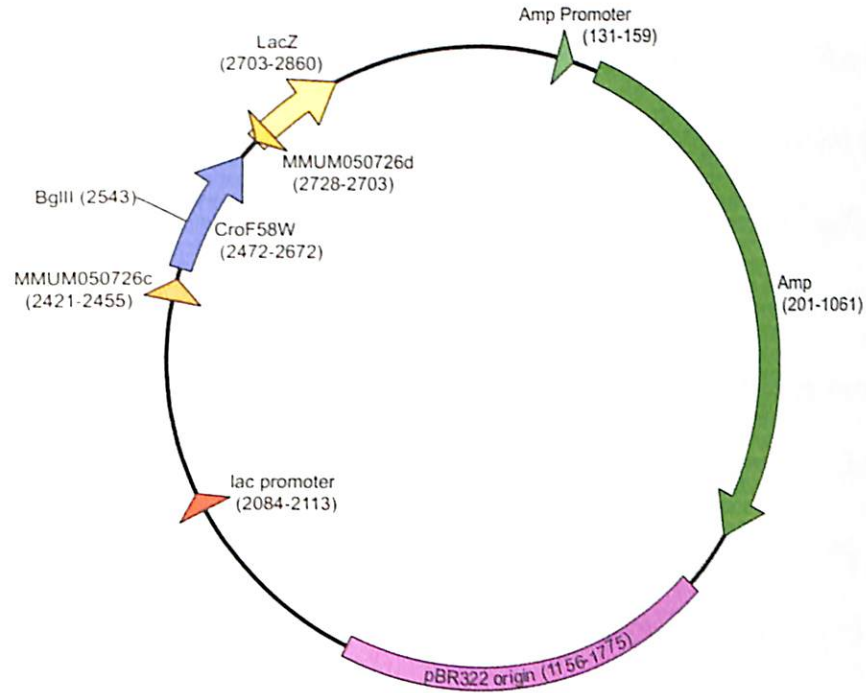
Figure 2: Strain 1—pUCroRSF58W. Depicted above is a graphic map of the puCroRSF58W plasmid. The origin of replication is represented in pink. The ampicillin resistance marker is shown in green. The *lacZ* gene is depicted in yellow. The lac promoter is represented in red. The *croF58W* is shown in purple. The two oligonucleotide primers used in the first PCR reaction are shown in orange.
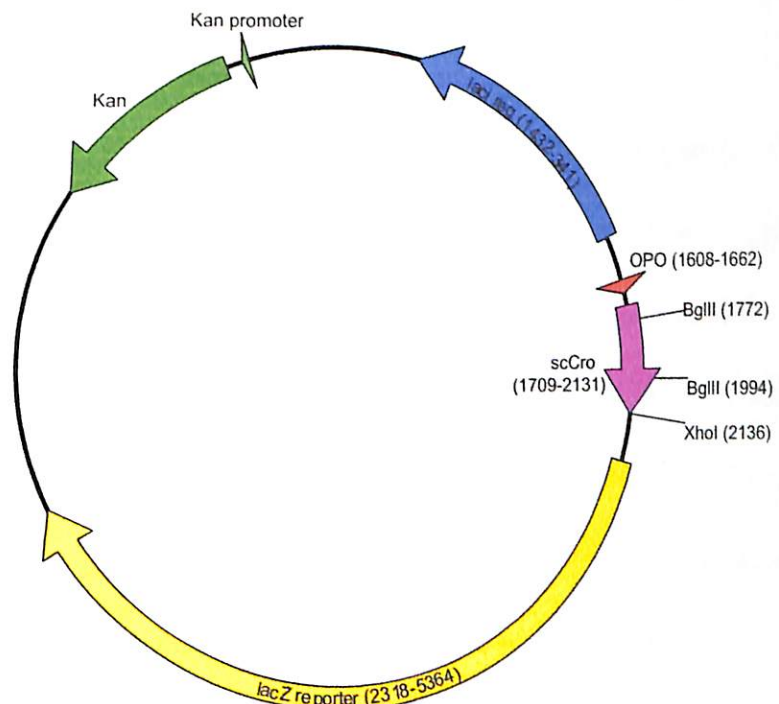
Figure 3: Strain 2—pAH125-lacI-OPO-scCro-lacZ. Depicted above is the map of the scCro circuit. The *lacI* and *lacZ* gene are shown in blue and yellow respectively. The kanamycin marker is depicted in green while the OPO control unit is shown in red. The *scCro* gene is depicted in pink.

18

Both the PCR fragment from pUCroRSF58W and the purified DNA from scCro were then subjected to a restriction enzyme digest with BglII and XhoI. The scCro plasmid has two BglII sites and one XhoI site. The first BglII site is located sixty-three base pairs downstream of the beginning of the scCro gene. The second BglII site is located thirty-seven base pairs upstream of the end of the scCro gene. The XhoI site is located five base pairs downstream of the end of the scCro gene. The segment of the scCro DNA that is relevant to the cloning of F58W is the one beginning at the first BglII site and extending to the XhoI site. This segment will ultimately serve as the vector for the cut PCR fragment during ligation. The digest of the PCR fragment cut its length down from 308 base pairs to 188 base pairs (Figure 5). The resulting restriction fragment began forty-three base pairs downstream of the CroF58W gene and ended twenty-one base pairs downstream of the beginning of the lacZ reporter. A ligation procedure was then carried out with the cut scCro DNA serving as the vector and the restriction fragment from the PCR fragment serving as the insert.

At first, transformations using the ligation product were attempted using competent BW25141 E. coli cells. However after several unsuccessful trials, highly competent RapidTrans TAMI E. coli cells were used instead. All transformations included several control agar plates including a plate streaked with unligated, cut scCro DNA and a plate containing no DNA. Successful growth on the non-control plates was obtained using the RapidTrans cells. The resulting candidate CroF58W cells should have the restriction fragment originating from pUCroRSF58W inserted into the aforementioned vector scCro DNA. The inserted restriction fragment begins sixty-three base pairs downstream of the beginning of the Cro gene and ends eighty-two base pairs

19
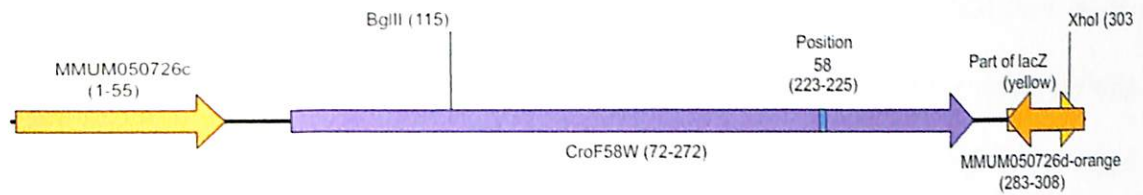
Figure 4: pUCroRSF58W PCR Fragment. The 303 base pair product of the PCR reaction done with pUCroRSF58W is shown above. The *croF58W* gene is shown in purple with its mutant position 58 shown in teal. The portion of the *lacZ* included in the PCR product is shown in yellow with the two PCR primers being shown in orange. The XhoI restriction site that is added as a result of the PCR reaction is labeled at the end of the PCR product.
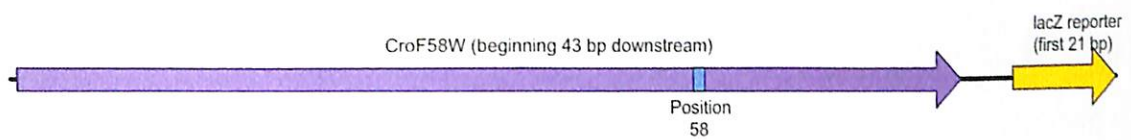
Figure 5: Product of Restriction Enzyme Digest with BglII and XhoI. Cutting the pUCroRSF58W PCR product with BglII and XhoI results in the 188 base pair fragment depicted above. This fragment is what will be inserted into the scCro vector.

20

upstream of the lacZ reporter. This newly cloned plasmid is pAH125::laci-OPO-CroF58W-lacZ and totals 7,707 base pairs in length (Figure 6).

B. Verification of CroF58W Clones

To confirm the identity of candidate colonies as CroF58W, colony PCR was conducted on a number of candidate colonies and compared to colony PCR from scCro. Twenty colonies were selected from the successful transformation plates, and each was subjected to PCR using MMUM050726b as the top primer and MMUM051214a as the bottom primer. The PCR product for the candidate CroF58W was expected to be 351 base pairs in length with the fragment beginning fifty base pairs upstream of the CroF58W gene and ending eighty-five base pairs downstream of the CroF58W gene (Figure 7; Figure A-3). As each colony was selected from the transformation plates, it was also re-streaked onto LB agar plates and allowed to grow overnight in order to obtain more single colonies of each candidate. DNA from scCro was also put through PCR with the same two primers. The PCR fragment from scCro began fifty base pairs upstream of the scCro gene and ended sixty base pairs downstream of the scCro gene. The total length of the scCro PCR product was expected to be 527 base pairs (Figure A-4). After all PCR products were purified, a 200 mL electrophoresis gel was run with both the scCro PCR product and each candidate's PCR product.

Most of the candidates were found to have identical DNA bands as those of scCro. However, one candidate was found to have a DNA band in the 300 base pair range as would be expected from CroF58W. Several enzymatic digestions were then conducted to further verify the candidate's identity as CroF58W. First, purified DNA
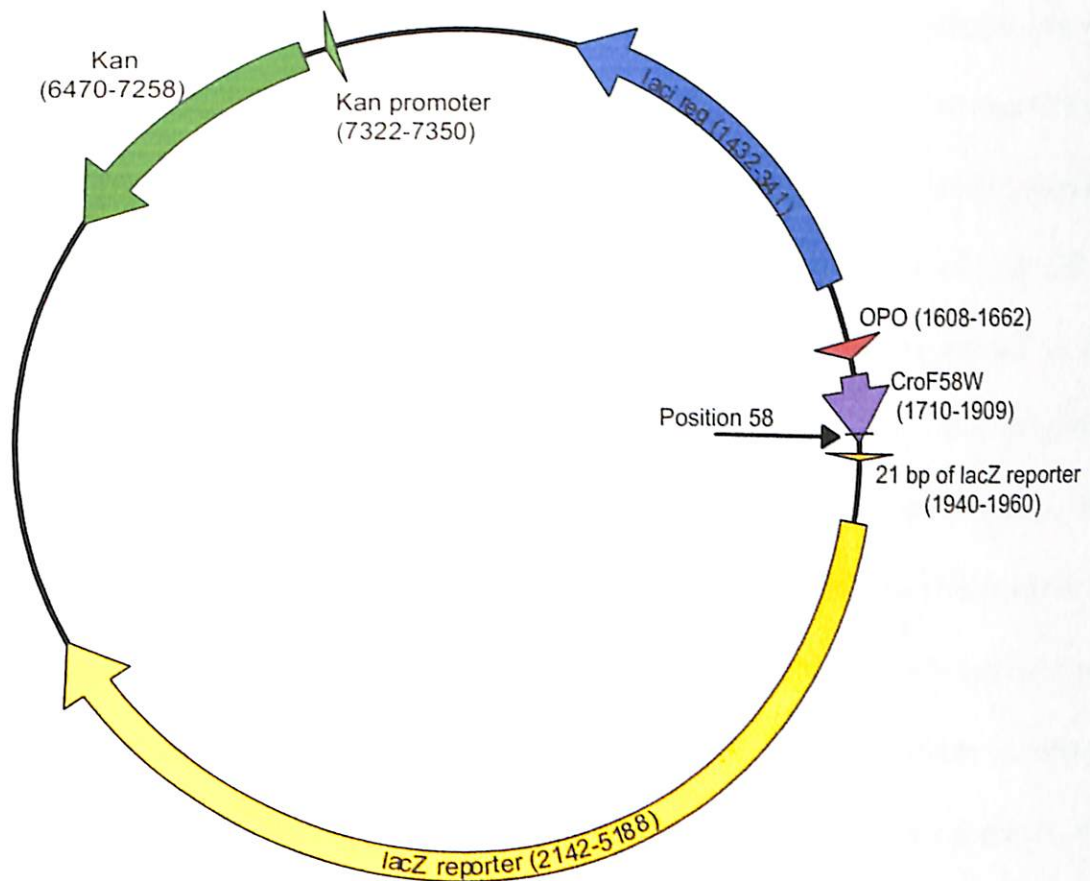
Figure 6: Newly Cloned pAH125-lacI-OPO-CroF58W-lacZ. After carrying out a transformation using the restriction enzyme fragment and scCro as the vector, CroF58W is produced. As seen in previous constructs, the *lacI*, *lacZ*, kanamycin, and OPO components of the circuit are depicted. The CroF58W gene is shown in purple.
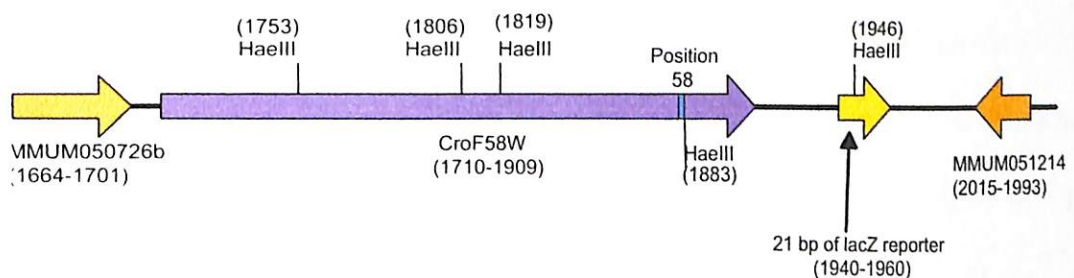
Figure 7: Candidate Colony PCR Product Used for the Verification of CroF58W. The graphic above shows the placement of PCR primers that were used to produce a 351 base pair PCR product. This fragment was then used to help verify that the candidate colonies were in fact CroF58W. This fragment comes from the plasmid shown in Figure 5. The PCR primers are depicted in orange with the *croF58W* gene being depicted in purple. HaeIII restriction sites used for verification procedures are labeled above.

from Cro, scCro and the candidate CroF58W colony was digested with the HindIII restriction enzyme. Through the cloning procedure an extra HindIII restriction site was added twenty base pairs downstream of the end of the CroF58W gene. Neither Cro nor scCro has this extra HindIII site. As a result, CroF58W contains three HindIII sites as opposed to the two that both Cro and scCro have. Enzymatic digestion with HindIII should result in three DNA fragments in CroF58W and only two in Cro and scCro. After this restriction digest was completed, a 25 mL electrophoresis gel was run with lanes containing cut and uncut Cro, scCro, and candidate CroF58W. As expected, three DNA bands were present for cut CroF58W while cut Cro and scCro only had two present.

Another enzymatic digest was then done with the restriction enzyme HaeIII and the PCR products from Cro and the candidate CroF58W. First, PCR had to be conducted using purified DNA from Cro. MMUM050726b was again used as the top primer, and MMUM051214a served as the bottom primer (Figure 8). The PCR product for Cro contains three HaeIII sites (Figure 9; Figure A-5) whereas the PCR product for CroF58W (Figure 7) contains the three from Cro plus two additional sites. One of these additional sites is the result of the amino acid change from phenylalanine to tryptophan at position 58, and the other additional site is the result of an extra twenty-one base pairs that are present at the end of the CroF58W gene as a result of cloning. Due to these extra sites, enzymatic digest with HaeIII should result in six DNA fragments in CroF58W and only four in Cro. Once the purified PCR products from Cro and the candidate CroF58W had been digested, a 25 mL electrophoresis gel was run to confirm the presence of the restriction bands. The results were as expected with the CroF58W candidate having more DNA bands than Cro. By taking these results and combining them with those of the
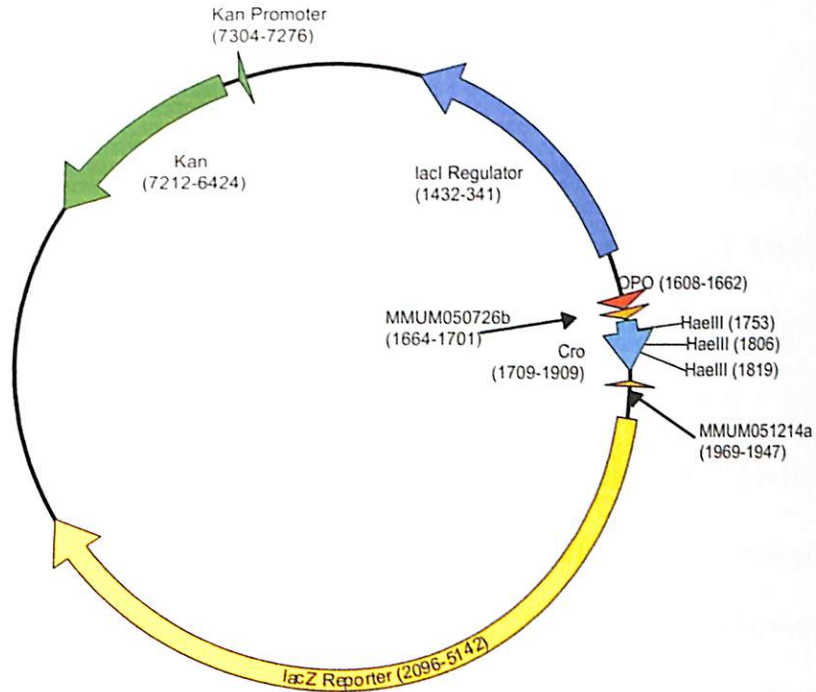
Figure 8: The Cro Plasmid with PCR Primers Used in the Verification of CroF58W. A PCR reaction was carried out using Cro as the template. In the figure above, the Cro circuit is shown along with the PCR primers used in the reaction (represented in orange). The HaeIII restriction sites that were involved in the subsequent enzyme digest of the PCR product are labeled above.
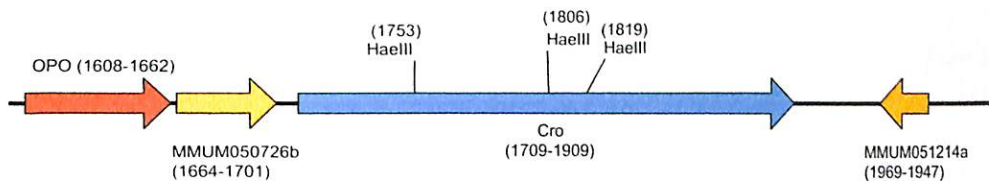
Figure 9: Cro PCR Product Used in the Restriction Enzyme Digest to Verify CroF58W. After carrying out a PCR reaction with the two primers shown above in orange, a 305 base pair product is obtained. This PCR fragment contains three HaeIII sites (labeled above). This product was digested with HaeIII and used as a comparison in procedures done to verify the candidate colony as CroF58W.

24

HindIII restriction digest and the colony PCR, it was concluded that the candidate colony was in fact CroF58W.

C. β-Galactosidase Assays

After confirming that our cloning procedures had resulted in the production of CroF58W, the CroF58W plasmid was used to transform competent BW25141 E. coli cells. After this was complete, β-galactosidase steady-state assays were conducted using the medium-copy number strains of our four Cro constructs: BW25141::pAH125-lacI-OPO-lacZ, BW25141::pAH125-lacI-OPO-Cro-lacZ, BW25141::pAH125-lacI-OPO-scCro-lacZ, and BW25141::pAH125-lacI-OPO-CroF58W-lacZ. Cells were grown both with and without inducer (IPTG) present. After the reaction with ONPG was complete, the recorded OD_{420} reading for each sample (after correcting for light scattering) was multiplied by 1000 and then divided by the time the reaction with ONPG progressed in minutes, the volume of culture used in the assay in milliliters, and the OD_{600} which indicated the cell density just before the assay. The resulting number was equal to units of β-galactosidase, which are proportional to the increase of o-nitrophenol per minute per bacterium (Miller, 1992). Table 1 lists the units of β-galactosidase for each strain both induced and un-induced as well as the induction level for each strain. Figure 10 shows the percent repression of each circuit under steady state conditions.

	OPO	Cro	scCro	CroF58W
Un-Induced (U)	169 units	135 units	100 units	1170 units
Induced (I)	17700 units	1400 units	407 units	13000 units
Ratio I/U	105	10.4	4.1	11.1

Table 2: CroF58W shows a higher level of β-galactosidase units than expected. In the un-induced strain data shown above, the CroF58W circuit contained a β-galactosidase level that was almost ten-fold higher than our control circuit OPO. The β-galactosidase level of CroF58W was also heightened in the induced cells as compared to the Cro and scCro circuits. By dividing the level of β-galactosidase in the induced cells by that found in the un-induced cells, the induction ratio for each strain is found.
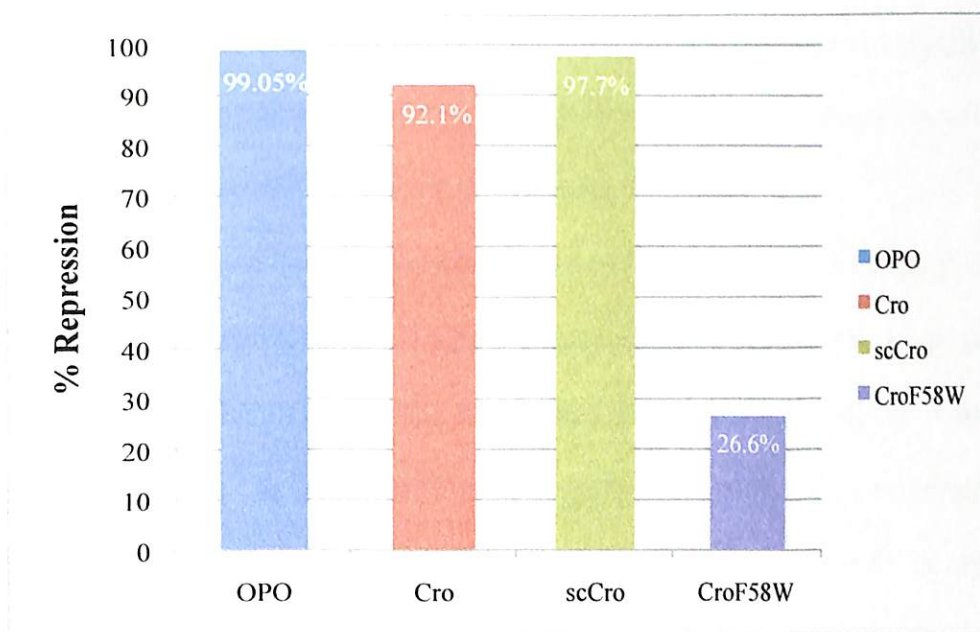
Figure 10: The Observed Percent Repression in the Induced Cells of each Circuit. The OPO repression level is indicative of the repressive capabilities of LacR. The other three circuits' repression levels indicate the ability of each Cro protein (wild-type, scCro, or CroF58W) to repress the circuit through a negative feedback loop.

IV. DISCUSSION

CroF58W is more thermodynamically stable but has slower dimer assembly kinetics as compared to wild-type Cro (Satumba & Mossing, 2002). The enhanced thermodynamic stability of CroF58W was attributed to the amino acid substitution at position 58. The substituted tryptophan residue penetrates deeper into the hydrophobic core of CroF58W dimers. Therefore, CroF58W dimers are more stable and less likely to dissociate as compared to Cro dimers. Since Cro dimers are more likely to separate and lose their functionality than are CroF58W dimers, CroF58W was expected to perhaps be better at repressing the circuit than Cro. CroF58W dimers can still separate, however, so scCro was expected to still be the best at repressing the circuit.

Two versions of the steady-state β-galactosidase assay were conducted; one involved assaying cells that had been grown under normal media conditions, and the other involved assaying cells that had been grown under inducing conditions. When the cells are not induced, LacR produced from the *lacI* gene occupies the promoter most of the time. Under inducing conditions, IPTG works to remove LacR from the lac operator, which results in RNA polymerase being able to access the lambda promoter and then transcribe both the *cro* gene and the *lacZ* gene. The differences in the steady-state β-galactosidase levels that are seen between our various constructs are the result of Cro proteins (wild-type Cro, scCro, or CroF58W) binding to the lambda operator and blocking the access of RNA polymerase.

The β-galactosidase activity of the CroF58W circuit, depicted in Table 1, is noticeably larger than those of the other Cro circuits. When comparing the induction ratios of each strain, however, CroF58W is found to have about the same ratio as wild-type Cro. The wild-type Cro circuit shows an approximate ten-fold increase in β-galactosidase between the un-induced and induced cells while the CroF58W circuit shows an approximate eleven fold increase. These results are relatively close to what is expected.

The results in Figure 10 show the percent repression of each circuit. This percentage is an indication of how effective each Cro variant is at repressing the circuit. In order to obtain this number, it is assumed that the induced OPO circuit gives a β-galactosidase level that is indicative of our promoter's maximal transcriptional capacity. This assumption is made because LacR is the only repressor present in the OPO construct, and this repressor is not occupying the operator under inducing conditions. This means that the promoter is always free and open for RNA polymerase binding. By dividing the un-induced β-galactosidase level of OPO by its induced level, the amount of activity in the form of transcription that LacR allows from the circuit can be seen. By subtracting this value from 1, the amount of repression that results from LacR can be seen. LacR was found to repress transcription from the circuit 99.05% of the time meaning that LacR was bound to the lac operator 99.05% of the time. Since the binding of LacR to the lac operator is a non-covalent interaction, this percentage can never reach 100% there will always be a time when that non-covalent interaction is not taking place.

By dividing the induced β-galactosidase levels of each circuit by the induced β-galactosidase level of the OPO circuit, we are able to obtain the amount of activity that

each Cro protein allows from the circuit. By comparing the percent of repression, it is seen that scCro is almost as effective as LacR in shutting down transcription from the circuit with a percent repression of 97.7%. The wild-type Cro protein does a relatively good job at repressing the circuit, but the CroF58W protein does a fairly poor job with a percent repression of 26.6%.

The simplest and most un-interesting explanation of the circuit's lack of repression is that there is something defective about both the *lacI* gene and the lac operator that are present in the CroF58W construct. A detrimental mutation could have occurred to the gene or the operator during any one of the cloning steps. The now imperfect gene is producing imperfect LacR, and the LacR proteins are not able to bind with the mutated lac operator. However, if the problem did lie with the LacR proteins and the lac operator, transcription from the promoter would still result in CroF58W proteins which could then negatively feedback into the circuit, bind to the lambda promoter, and shut down transcription. This should result in a β-galactosidase level that is around the level observed in induced wild-type Cro cells. Since the CroF58W β-galactosidase levels are significantly higher than the induced Cro levels, the CroF58W protein would also have to have a much weaker binding affinity that originally expected.

The increased β-galactosidase level of CroF58W could also be due to an extra fragment of DNA that is present at the end of the *croF58W* gene but not at the end of any of the other *cro* genes. As an artifact of the cloning process, the CroF58W construct has twenty-five extra base pairs in between the end of the *croF58W* gene and the beginning of the *lacZ* gene. Twenty-one of those twenty-five base pairs are the beginnings of the *lacZ* gene meaning that the *lacZ* gene begins ten base pairs downstream of *croF58W*,

29

runs for twenty-one base pairs, ends, and begins again at the actual whole *lacZ* gene 182 base pairs later.

The mRNA transcripts produced from the CroF58W circuit could also have an effect on the β-galactosidase levels observed in the cells. These mRNA molecules could be more stable and have a longer half-life than other Cro transcripts. This would result in the transcripts being able to be translated more often than other transcripts that deteriorate at a faster rate. There could also be something about the mRNA transcripts that causes them to be recognized and translated more quickly by ribosomes, which would also result in an increased amount of translation for each transcript. If each CroF58W transcript is being translated at a higher level and resulting in ten times more β-galactosidase per mRNA molecule, the observed increase in β-galactosidase levels of the circuit would be explained.

The simplest way to start resolving some of these questions and hypotheses would be to create another synthetic CroF58W circuit that is the same as the wild-type Cro circuit except for the one codon change at position 58 in the *cro* gene. In order to compare the assay results of each circuit to each other with a certainty that the comparisons are valid, everything in each circuit must be identical and controlled except for the variant *cro* genes. Unfortunately, the extra fragment at the end of the *croF58W* gene in our current circuit makes this circuit different from our other circuits in more ways than just the variant gene. Ideally a process called cassette mutagenesis could be conducted to obtain an identical circuit (save the *croF58W* gene). This process is used to produce targeted mutations in DNA. The functioning wild-type Cro circuit that we have could be used as the starting material, and cassette mutagenesis would make mutations

30

necessary for changing the codon at position 58. By doing this procedure, we can be assured that the only difference that exists between the Cro, scCro, and CroF58W circuits are the variants of the *cro* genes. Further assays can then be carried out and their results analyzed.

Once enough assay data has been accumulated for CroF58W, computer-modeling programs can be used to try and fit the data in order to determine the dynamics of protein folding, dimerization and DNA binding. Computer modeling is currently being used to fit the data for wild type Cro and scCro. Once all of these models are finished, the dynamics of each circuit can be compared to those of the others in order to determine the role that folding, dimerization, and DNA binding play in each of the circuits.

LIST OF REFERENCES

Alberts, B. et al. (2008). *Molecular Biology of the Cell, fifth edition.* New York: Garland Science.

Atsumi, S. and J.W. Little. (2006). Coupled energetics of lambda cro repressor in prophage induction of phage λ as analyzed by a module-replacement approach. *Proceedings of the National Academy of Sciences USA, 103,* 4558-63.

Haldimann, A. and B.L. Wanner. (2001). Conditional-Replication, Integration, Excision, and Retrieval Plasmid-Host Systems for Gene-Structure-Function Studies of Bacteria. *Journal of Bacteriology, 183* (21), 6384-6393.

Jana, R., Hazbun, T., Fields, J., and Mossing, M. (1998). Single-Chain Lambda Cro Repressors Confirm High Intrinsic Dimer-DNA Affinity. *Biochemistry, 37,* 6446-6455.

Miller, J.H. (1992). Experiments in Molecular Genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Ptashne, M. (2004). A Genetic Switch 3rd Edition: Phage Lambda Revisited. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Rupert, P., Mollah, A.K.M.M., and Mossing, M. (2000). The Structural Basis for Enhanced Stability and Reduced DNA Binding Seen in Engineered and Second-generation Cro Monomers and Dimers. *Journal of Molecular Biology, 296,* 1075-1090.

Satumba, J.W. and M. Mossing. (2002) Folding and Assembly of Lambda Cro Repressor Dimers Are Kinetically Limited by Proline Isomerization. *Biochemistry, 41,* 14216-224.

Schubert, R., Dodd, I., Egan, J.B., and Shearwin, K. (2007). Cro's role in the CI-Cro bistable switch is critical for λ's transition from lysogeny to lytic development. *Genes and Development, 21,* 2461-2472.

Wilson, C.J., Zahn, H., Swint-Kruse, L., and Matthews, K.S. (2007). The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cellular and Molecular Life Sciences, 64,* 3-16.

APPENDIX

```
                                5'  CACTCGACGAATTCCAGGAAACAGAATGGTTGCAT 3' MMUM050726c
                                                                                      M
ctcgTATAATgtgtggAATTGTgAgCgGaTaACAATTcacacaggaaacagaatggttgcatgtactaaggaggttccA
       ^2410        ^2420        ^2430        ^2440        ^2450        ^2460        ^2470
       ....         .....  ..... .......tcctttgtcttaccaacgtacatgattcctccaaggT
--------------------------------------------------------------------------------


     E   Q   R   I   T   L   K   D   Y   A   M   R   F   G   Q   T   K   T   A   K   D   L   G   V   Y   Q
tgGaaCaaCgcAtaActTtaAaaGatTatGcaAtgCgcTttGgcCaaAccAagAcaGctAaaGatCtcGgcGtgTatCaa
       ^2490        ^2500        ^2510        ^2520        ^2530        ^2540        ^2550
       .....  .....  ..... .....AaaCgGttTggTtcTgtCgaTttCtaGagCcgCacAtaGtt
--------------------------------------------------------------------------------


 S   A   I   N   K   A   I   H   A   G   R   K   I   F   L   T   I   N   A   D   G   S   V   Y   A   E   E
AgcGcgAtcAacAagGccAtcCatGccGgcCgaAaaAtaTttTtgActAtaAacGctGatGgaTccGtaTacGcgGaaGa
       ^2570        ^2580        ^2590        ^2600        ^2610        ^2620        ^2630
       ....         ...... AaaAacTgaTatTtgCgaCtaCctAggCatAtgCgcCttCt
--------------------------------------------------------------------------------


                                                                        3' GTGACCGGCA
 V   K   P       P   S   N   K   K   T   T   A   Z
gGtaAagCccTggCcgAgtAacAaaAaaAcaAcaGcataaataaacccgacCTGCAGgcatgcAAGCTTggcactggccgt
       ^2650        ^2660        ^2670        ^2680        ^2690        ^2700        ^2710
       .....        ..... ..... .....ttattgggctgGACGTCcgtacgTTCGAAccgtgaccggca
--------------------------------------------------------------------------------


GCAAAATGTGAGCTCC 5' MMUM050726d
cgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagct
       ^2730        ^2740        ^2750        ^2760        ^2770        ^2780        ^2790
       .....        ..... ..... ....gaaattagcggaacgtcgtgtaggggggaaagcggtcga
--------------------------------------------------------------------------------
```

Figure A-1: Segment of pUCroRSF58W with Aligned PCR Primers. Above shows a DNA segment from the pUCroF58W plasmid that ranges from base pair 2401 to base pair 2800. The top strand PCR primer is shown in green while the bottom strand PCR primer is shown in red. Each primer is aligned above its corresponding DNA sequence. The amino acid sequence of the CroF58W protein that is encoded by the *croF58W* gene is shown in blue. The tryptophan amino acid that replaces a phenylalanine at position 58 is shown in orange.

35

CAGTCGACGAATTCCAGGAAACAGAATGGTTGCAT 3' MMUM050726c

```
                                                   M  E  Q  R  I  T  L  K  D  Y
CAGTCGACGAATTCCAGGAAACAGAATGGTTGCATgtactaaggaggttccAtgGaaCaaCgcAtaActTtaAaaGatTa
        ^10          ^20          ^30         ^40         ^50        ^60        ^70
                                        attcctccaaggTacCttGttGcgTatTgaAatTttCtaAt
-----------------------------------------------------------------------------------
```

```
                                   BglII
 A   M   R   F   G   Q   T   K   T   A   K   D   L   G   V   Y   Q   S   A   I   N   K   A   I   H   A   G
tGcaAtgCgcTttGgcCaaAccAagAcaGctAaaGatCtcGgcGtgTatCaaAgcGcgAtcAacAagGccAtcCatGccG
        ^90          ^100         ^110        ^120        ^130       ^140       ^150
aCg                                   adagCcgCacAtaGttTcgCgcTagTtgTtcCggTagGtaCggC
-----------------------------------------------------------------------------------
```

```
 R   K   I   F   L   T   I   N   A   D   G   S   V   Y   A   E   E   V   K   P   W   P   S   N   K   K
gcCgaAaaAtaTttTtgActAtaAacGctGatGgaTccGtaTacGcgGaaGagGtaAagCccTggCcgAgtAacAaaAaa
        ^170         ^180         ^190        ^200        ^210       ^220       ^230
cggc                             act AagCatAtgCgcCttCtcCatTtcGggAccGgcTcaTtgTttTtt
-----------------------------------------------------------------------------------
```

```
                                        3'  GTGACCGGCAGCAAAATGTGAGCTCC 5' MMUM050726d
                                                                       XhoI
 T   T   A   Z
AcaAcaGcataaataacccgacCTGCAGgcatgcAAGCTTggCACTGGCCGTCGTTTTACACTCGAGG
        ^250         ^260         ^270        ^280        ^290       ^300
Tgt                             acaTTCGAAccGTGACCGGCAGCAAAATGTGAGCTCC
-----------------------------------------------------------------------------------
```

Figure A-2: PCR Product from pUCroRSF58W. The DNA sequence of the 308 base pair PCR product from pUCroF58W is shown above. The top strand primer is shown in green, and the bottom strand primer is shown in red. The amino acid sequence of the CroF58W gene is shown in blue with the tryptophan amino acid shown in orange. The XhoI restriction site that is added as a result of the bottom strand primer is shown in purple. The BglII restriction site that is also utilized in the subsequent restriction enzyme digest is also shown in purple.

GcgcgtgTTGACTattttacctctggcggtGATAATggttgaattgtgagcggataacaattcccctCTAGAAATAATTT
          ^1610   ^1620   ^1630   ^1640   ^1650   ^1660   ^1670

                                       ...ATTAccaacttaacactcgcctattgttaaggggaGATCTTTATTAAA

TGTTTAACTTTAAGAAGGAGA 3' MMUM050726b

                                                                              HaeIII

                    M   E   Q   R   I   T   L   K   D   Y   A   M   R   F   G   Q   T   K
TGTTTAACTTTAAGAAGGAGATATAcatATGGAACAACGCATAACTTTAAAAGATTATGCAATGCGCTTTGGCCAAACCA
          ^1690   ^1700   ^1710   ^1720   ^1730   ^1740   ^1750

                                    ...TTGTTGCGTATTGAAATTTTCTAATACGTTACGCGAAACCGGTTTGGT

                                                              HaeIII
                                  HaeIII
    T   A   K   D   L   G   V   Y   Q   S   A   I   N   K   A   I   H   A   G   R   K   I   F   L   T   I
AGACAGCTAAAGatCtcGgcGtgTatCaaAgcGcgAtcAacAagGccAtcCatGccGgcCgaAaaAtaTttTtgActAta
          ^1770   ^1780   ^1790   ^1800   ^1810   ^1820   ^1830

                          ...TcgCgcTagTtgTtcCggTagGtaCggCcgGctTttTatAaaAacTgaTat

---

                                      HaeIII
    N   A   D   G   S   V   Y   A   E   E   V   K   P   X   P   S   N   K   K   T   T   A   Z
AacGctGatGgaTccGtaTacGcgGaaGagGtaAagCccTggCcgAgtAacAaaAaaAcaAcaGcataaataacccgacC
          ^1850   ^1860   ^1870   ^1880   ^1890   ^1900   ^1910

                      ...TtcGagAccGgcTcaTtgTttTttTgtTgtCgtatttattgggctgG

                                                              3' GACGATTG

                        HaeIII
TGCAGgcatgcAAGCTTggCACTGGCCGTCGTTTTACACTCGAGcaccaccacCACCACCACTGAGATCCGGCTGCTAAC
          ^1930   ^1940   ^1950   ^1960   ^1970   ^1980   ^1990

          ...AGCAAAATGTGAGCTCgtggtggtgGTGGTGGTGACTCTAGGCCGACGATTG

TTTCGGGCTTAAGA 5' MMUM051214a
AAAGCCCGaattcccggggatccgacaaccgatgaaagcggcgacgcgcagttaatcccacagccgccagttccgctggc
          ^2010   ^2020   ^2030   ^2040   ^2050   ^2060   ^2070

          ...actttcgccgctgcgcgtcaattagggtgtcggcggtcaaggcgaccg

---

Figure A-3: CroF58W Plasmid with Aligned PCR Primers. The PCR primers that were used to obtain the 351 base pair PCR product from CroF58W are shown aligned with their complementary DNA sequences above. The top strand primer is shown in green, and the bottom strand primer is shown in red. The HaeIII restriction sites that were utilized in the restriction enzyme digest are shown in purple. The amino acid sequence of the CroF58W gene is shown in blue with the tryptophan amino acid at position 58 is shown in orange.

```
                                                      5' CAGTCGACGAATTCTTT
GcgcgtgTTGACTattttacctctggcggtGATAATggttgaattgtgagcggataacaattcccctCTAGAAATAATTT
        ^1610      ^1620      ^1630      ^1640      ^1650      ^1660      ^1670
                                                               aaggggaGATCTTTATTAAA
--------------------------------------------------------------------------------


TGTTTAACTTTAAGAAGGAGA 3' MMUM050726b
                           M  E  Q  R  I  T  L  K  D  Y  A  M  R  F  G  Q  T  K
TGTTTAACTTTAAGAAGGAGATATAcatATGGAACAACGCATAACTTTAAAAGATTATGCAATGCGCTTTGGCCAAACCA
        ^1690      ^1700      ^1710      ^1720      ^1730      ^1740      ^1750
                          TTTGTTGCGTATTGAAATTTTCTAATACGTTACGCGAAACCGGTTTGGT
--------------------------------------------------------------------------------


   T  A  K  D  L  G  V  Y  Q  S  A  I  N  K  A  I  H  A  G  R  K  I  F  L  T  I
AGACAGCTAAAGATCTCGGCGTGTATCAAAGCGCGATCAACAAGGCCATCCATGCCGGCCGAAAAATATTTTTGACTATA
        ^1770      ^1780      ^1790      ^1800      ^1810      ^1820      ^1830
                          AGTTGTTCCGGTAGGTACGGCCGGCTTTTTATAAAAACTGATAT
--------------------------------------------------------------------------------


N  A  D  G  S  V  Y  A  E  E  V  K  P  F  P  S  N  K  K  T  T  A  A  G  T  G  G
AACGCTGATGGATCCGTATACGCGGAAGAGGTAAAGCCCTTCCCGAGTAACAAAAAAACAACAGCGGCCGGTACCGGTGG
        ^1850      ^1860      ^1870      ^1880      ^1890      ^1900      ^1910
                          ATTTCGGGAAGGGCTCATTGTTTTTTTGTTGTCGCCGGCCATGGCCACC
--------------------------------------------------------------------------------


S  G  G  M  E  Q  R  I  T  L  K  D  Y  A  M  R  F  G  Q  T  K  T  A  K  D  L  G
CTCTGGCGGCATGGAACAACGCATAACTTTAAAAGATTATGCAATGCGCTTTGGCCAAACCAAGACAGCTAAAGATCTCG
        ^1930      ^1940      ^1950      ^1960      ^1970      ^1980      ^1990
                          ATTGAAATTTTCTAATACGTTACGCGAAACCGGTTTGGTTCTGTCGATTTCTAGAGC
--------------------------------------------------------------------------------


   V  Y  Q  S  A  I  N  K  A  I  H  A  G  R  K  I  F  L  T  I  N  A  D  G  S  V
GCGTGTATCAAAGCGCGATCAACAAGGCCATCCATGCCGGCCGAAAAATATTTTTGACTATAAACGCTGATGGATCCGTA
        ^2010      ^2020      ^2030      ^2040      ^2050      ^2060      ^2070
                          TTTATTCCGGTAGGTACGGCCGGCTTTTTATAAAAACTGATATTTGCGACTACCTAGGCAT
--------------------------------------------------------------------------------


Y  A  E  E  V  K  P  F  P  S  N  K  K  T  T  A  Z
TACGCGGAAGAGGTAAAGCCCTTCCCGAGTAACAAAAAAACAACAGCATAAataCTCGAGcaccaccacCACCACCACTG
        ^2090      ^2100      ^2110      ^2120      ^2130      ^2140      ^2150
ATGCGCTTCTCCATTTCGGGAAGGGCTCATTGTTTTTTTGTTGTCGTATTtatGAGCTCgtggtggtgGTGGTGGTGAC
--------------------------------------------------------------------------------


   3' GACGATTGTTTCGGGCTTAAGA 5' MMUM051214a
AGATCCGGCTGCTAACAAAGCCCGaattcccggggatccgacaaccgatgaaagcggcgacgcgcagttaatcccacagc
        ^2170      ^2180      ^2190      ^2200      ^2210      ^2220      ^2230
TCTAGGCCGACGATTGTTTCGGG cttaagggcccctaggctgttggctactttcgccgctgcgcgtcaattagggtgtcg
--------------------------------------------------------------------------------
```

Figure A-4: DNA Sequence from the scCro Plasmid with Aligned PCR Primers. The DNA sequence shown above is from the scCro plasmid. The two PCR primers that were used to obtain a 527 base pair PCR product are aligned with their complementary DNA sequences. The top strand primer is shown in green, and the bottom strand primer is shown in red. The amino acid sequence of the scCro protein is shown in blue.

38

```
                                                           5' CAGTCGACGAATTCTTT
GcgcgtgTTGACTattttacctctggcggtGATAATggttgaattgtgagcggataacaattcccctCTAGAAATAATTT
         ^1610       ^1620       ^1630       ^1640       ^1650       ^1660       ^1670
                              ...ATTAccaacttaacactcgcctattgttaaggggaGATCTTTATTAAA
----------------------------------------------------------------------------------


TGTTTAACTTTAAGAAGGAGA  3'  MMUM050726b
                                                                           HaeIII
                    M   E   Q   R   I   T   L   K   D   Y   A   M   R   F   G   Q   T   K
TGTTTAACTTTAAGAAGGAGATATAcatAtgGaaCaaCgcAtaActTtaAaaGatTatGcaAtgCgcTttGgcCaaAccA
         ^1690       ^1700       ^1710       ^1720       ^1730       ^1740       ^1750
                          ...GcgTatTgaAatTttCtaAtaCgtTacGcgAaaCcgGttTggT
----------------------------------------------------------------------------------


                                                            HaeIII
                                                HaeIII
     T   A   K   D   L   G   V   Y   Q   S   A   I   N   K   A   I   H   A   G   R   K   I   F   L   T   I
agAcaGctAaaGatCtcGgcGtgTatCaaAgcGcgAtcAacAagGccAtcCatGccGgcCgaAaaAtaTttTtgActAta
         ^1770       ^1780       ^1790       ^1800       ^1810       ^1820       ^1830
                          ...cTagTtgTtcCggTagGtaCggCcgGctTttTatAaaAacTgaTat
----------------------------------------------------------------------------------


N   A   D   G   S   V   Y   A   E   E   V   K   P   F   P   S   N   K   K   T   T   A   Z
AacGctGatGgaTccGtaTacGcgGaaGagGtaAagCccTtcCcgAgtAacAaaAaaAcaAcaGcataaataCTCGAGca
         ^1850       ^1860       ^1870       ^1880       ^1890       ^1900       ^1910
                          ...at TtcGggAagGgcTcaTtgTttTttTgtTgtCgtatttatGAGCTCgt
----------------------------------------------------------------------------------


                        3' GACGATTGTTTCGGGCTTAAGA 5'  MMUM051214a
ccaccacCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGaattcccgggatccgacaaccgatgaaagcggcgacg
         ^1930       ^1940       ^1950       ^1960       ^1970       ^1980       ^1990
                      ...GACGATTGTTTCGGGCttaagggcccctaggctgttggctactttcgccgctgc
----------------------------------------------------------------------------------
```

Figure A-5: Cro with Aligned PCR Primers Used for Verification. The above DNA sequence is from the wild-type Cro plasmid. The top strand and bottom strand primers are shown in green and red respectively and are aligned with their corresponding DNA sequences. These two primers resulted in a 305 base pair PCR product that was used in a restriction enzyme digest used to verify CroF58W. The HaeIII restriction sites utilized in the digest are shown in purple. The amino acid sequence of the wild-type Cro protein is shown in blue.