2020

Investigation of the Phenotypic Effect of Mutating a Highly-Conserved Cysteine Residue in the RNA Polymerase Beta Prime Subunit of E. Coli RNA Polymerase

Meg Dillingham
Western Kentucky University, megan.e.dillingham@gmail.com

INVESTIGATION OF THE PHENOTYPIC EFFECT OF MUTATING A HIGHLY CONSERVED CYSTEINE RESIDUE IN THE BETA PRIME SUBUNIT OF *E. COLI* RNA POLYMERASE

A Capstone Experience/Thesis Project Presented in Partial Fulfillment

of the Requirements for the Degree Bachelor of Science

with Mahurin Honors College Graduate Distinction

at Western Kentucky University

By

Meg E. Dillingham

August 2020

*****

CE/T Committee:

Dr. Rodney A. King, Chair

Ms. Naomi Rowland

Dr. Claire Rinehart

ABSTRACT

All bacteria contain a multi-subunit RNA polymerase (RNAPs) that is essential for gene expression. Because of the centrality of these enzymes in cellular life, the structure and function of the various subunits is intensely studied. The primary sequence of the RNAP β' subunit contains five cysteine residues that are highly conserved. Four of the cysteines coordinate a zinc atom and form the beta prime subunit zinc binding domain (ZBD). Mutation of any one of the ZBD cysteines is lethal to the cell. However, the role of the fifth residue (C58), which is located upstream of the ZBD cysteines, has not been investigated. In previous work, we cloned a copy of the *E. coli rpoC* onto a plasmid and changed the cysteine at position 58 to an alanine (C58A). Phenotypic analysis suggested that expression of the mutant subunit from the multi-copy plasmid did not support *E. coli* growth at high temperatures. In this study, we describe the generation of the C58A mutation in single copy on the chromosome using a chromosomal engineering technique. In addition, we investigated if the mutant subunit affects RNA-mediated transcription antitermination.

ACKNOWLEDGEMENTS

I chose Dr. Rodney King to serve as my thesis advisor because I knew him to be a dedicated mentor. I have not been disappointed. Dr. King has been a constant supporter throughout my thesis work. I am so grateful not only for the time he has invested into my thesis, but also into my development and growth as a scientist.

I am also very grateful for the regular assistance of Ms. Naomi Rowland, for the thrilling discussions of molecular biology I had with Dr. Claire Rinehart, and for the dedication of all members of the WKU Biology and Chemistry departments. I would like to thank Western Kentucky University, Ogden College, and the Mahurin Honors College for their generous financial backing of my project and for their support of my development as a scientist and scholar. Lastly, I would like to thank Dr. Johnathan Whetstine, Dr. Sigrid Jacobshagen, and Ms. Jenny Wells Pyle for their incredible mentorship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

RNA Polymerase (RNAP) is an essential enzyme for all life. Its proper function is necessary for gene expression and consequently, cell function and survival. Cellular RNAPs are complex molecules that are made up of multiple subunits and each subunit has critical domains that contribute to the overall structure and function of the enzyme. If mutations occur in certain critical regions of the protein, RNA transcription will cease, and the cell will die. Alternatively, mutations may simply alter the behavior of the enzyme. For example, the RNAP found in the bacteria *Escherichia coli* has a very important region called a zinc-binding domain located in the β' subunit of the enzyme (Clerget et al., 1995). Four conserved cysteines in the β'subunit—residues C70, C72, C85, and C88—are responsible for zinc atom binding (Bergsland & Haselkorn, 1991). While the function of this highly conserved region is unknown, specific mutations in the zinc-binding domain prevent the growth of certain viruses that use *E. coli* as a host. The deletion or mutation of any of the four conserved cysteine residues in this domain is deleterious to the cell.

A fifth highly conserved cysteine, C58, is located nearby (Bergsland & Haselkorn, 1991) but the role it plays in the structure and function of the enzyme is unknown. Its conservation suggests that it too may be critical for the function of RNA polymerase. The research described here investigates the role of C58 in RNAP function.

RNA Polymerase in Prokaryotes and Eukaryotes

Though RNA Polymerase is found in organisms across all domains of life, there are important distinctions between RNAPs from the different domains. For example, prokaryotes possess a single multi-subunit RNAP which is responsible for all transcription within the cell. This polymerase contains six subunits: β, β', ω, and two α subunits form the core enzyme, and the σ initiation factor binds to the core to form the holoenzyme (Geszvain & Landick, 2005). The α dimer functions as a scaffold for the assembly of β and β', the two largest subunits of RNAP. The β and β' subunits form a cleft that allows the DNA template to enter the active site, which is composed of two double-psi beta-barrel domains, one from each of β and β'. In addition to this catalytic domain, the β' subunit of RNAP contains a trigger loop necessary for catalysis and a bridge helix required for DNA/RNA translocation as nucleotides are added (Sutherland and Murakami, 2018). The smallest subunit of the core enzyme, ω, is the only subunit that is not critical for cell survival. It is thought to play a role in maintaining RNAP catalytic activity, but given its non-crucial role, the subunit's function has not been well-characterized (Sutherland and Murakami, 2018). The σ subunit, which along with the core enzyme forms the holoenzyme, is responsible for the specific binding of the holoenzyme to promoter DNA (Geszvain & Landick, 2005). Figure 1 shows the structure of *E. coli* RNAP.

While eukaryotic RNA transcription occurs via the same fundamental mechanisms as transcription in prokaryotic cells, there are substantial differences between the process in prokaryotes and eukaryotes. Notably, eukaryotes utilize different RNAPs to transcribe distinct kinds of RNA. RNA Polymerase II is responsible for

transcription of mRNAs, while RNAP I and RNAP III are responsible for the transcription of non-protein coding RNAs such as ribosomal RNAs, transfer RNAs, snRNAs, and scRNAs (Cooper, 2000). Additionally, while prokaryotic RNAP binds directly to promoter sequences, eukaryotic RNAPs interact with additional proteins (transcription factors) to regulate and initiate transcription (Cooper, 2000).

Like prokaryotic RNAP, eukaryotic RNAPs are composed of multiple subunits, ranging from 8-14 in number. For example, human RNAP II, which is responsible for the transcription of mRNA, contains 12 subunits. Each subunit and its respective function are listed in Table 1.

Two major subunits are found in all three eukaryotic RNAPs and are related to the β and β' subunits found in prokaryotic RNAP. Given that the β and β' subunits form the active site of RNAP, it is not surprising that the structure of these regions is so well conserved. This conservation across nearly all species highlights the functional importance of this critical enzyme, demonstrating the importance of conserved regions in RNAP.

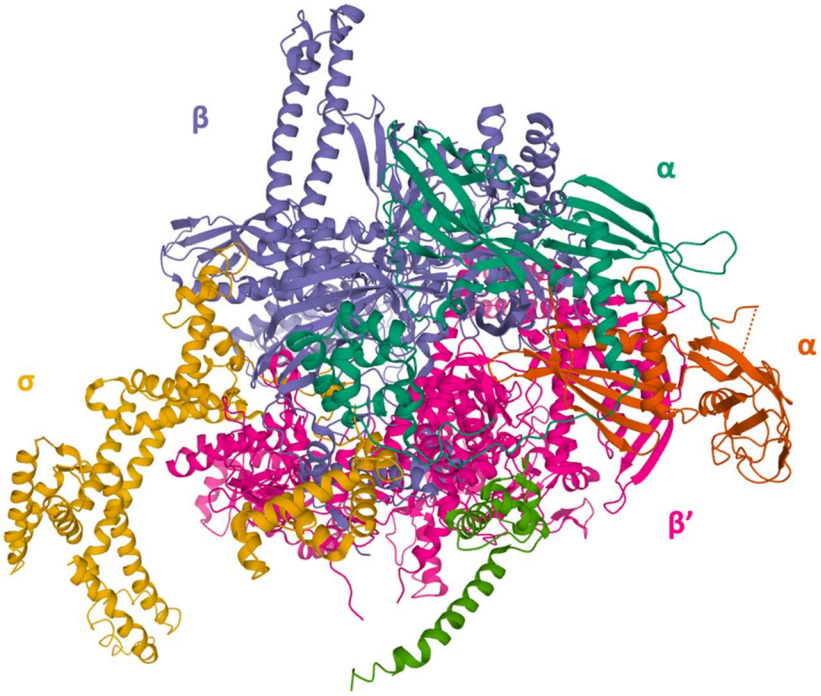


Figure 1: X-ray crystal structure of the *Escherichia coli* RNA polymerase sigma70 holoenzyme. PBD ID: 4YG2. Structure determined by Murakami, 2013. Image generated using the RCSB PBD (https://www.rcsb.org/) and Mol* (Senhal, D., D. Sehnal, Rose, A.S. , Kovca, J., Burley, S.K., and Velankar, S., 2018). Mol*: Towards a common library and tools for web molecular graphics. MolVA/EuroVis Proceedings. doi:10.2312/molva.20181103).

Subunit gene	*Function*
DNA-directed RNA polymerase II RPB1	Part of the core element of the central large cleft, which helps form the "jaw" that is thought to grab the incoming DNA template
DNA-directed RNA polymerase II RPB2	Contributes to RNAP catalytic activity and forms the active center along with RPB1
DNA-directed RNA polymerase II RPB3	Part of the core element with the central large cleft and the clamp element that moves to open and close the cleft
DNA-directed RNA polymerase II RPB4	RPB4 complexes with RPB7 to prevent double-stranded DNA from entering the active site; binds single-stranded DNA and RNA
DNA-directed RNA polymerase II RPB5	Involved in gene transcription regulation
DNA-directed RNA polymerase I, II, and III RPABC2	Part of the clamp element and along with RPB1 and RPB2 forms a pocket to which the RPB4-RPB7 subcomplex binds
DNA-directed RNA polymerase II RPB7	Complexes with RPB4 to prevent double-stranded DNA from entering the active site. Binds single-stranded DNA and RNA
DNA-directed RNA polymerases I, II, and III subunit RPABC3	Essential subunit with unclear function
DNA-directed RNA polymerase II subunit RPB9	Part of the upper jaw surrounding the large central large cleft and thought to grab the incoming DNA template
DNA-directed RNA polymerases I, II, and III RPAC5	Part of the core element with the central large cleft
DNA-directed RNA polymerase II RPB11-a	Part of the core element with the central large cleft
DNA-directed RNA polymerases I, II, and III RPABC4	Small subunit with no well-characterized function

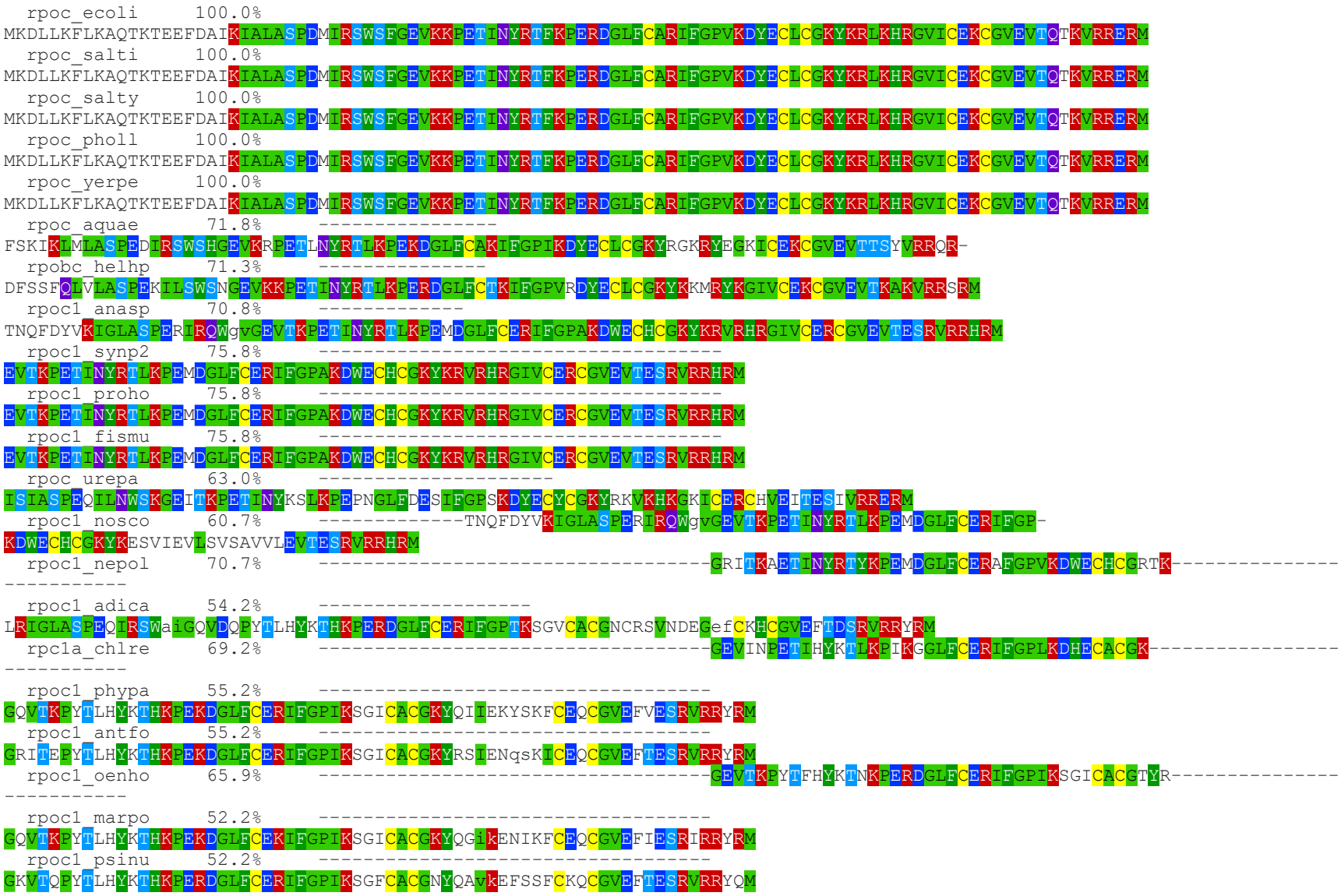
Table 1: Human RNAP II subunits and respective functions. All descriptions retrieved from the UniProt database (https://www.uniprot.org/).

The β' Subunit of RNA Polymerase.

Given that the β' subunit of RNAP is one the most highly conserved subunits, it is expected to be important to the enzyme's function. Thus, the exact mechanisms of the β' function and the subunit's precise role in RNA transcription is of great interest, as understanding β' gives insight not only the function of the bacterial subunit, but also into the function of the homologous subunit found in eukaryotic RNAP (RPB1). Bacteria are excellent model organisms for the study of RNAP as they grow quickly, require simple culturing (as compared to, for example, human tissues), and are easily transformed. For these reasons, bacterial RNAPs have been heavily researched. It is important to note that the use of bacterial RNAP to study RNAP in general has some drawbacks. While the simplified bacterial RNAP contains fewer subunits and serves as a simplified model for study, the complexity of eukaryotic RNAPS, which transcribe a variety of RNA types and are temporally regulated, cannot be fully captured by studies of bacterial RNAP. Still, advancements in the understanding of prokaryotic RNAP give insight into the basic function of RNAPs across domains. The study described in this thesis investigates *E. coli* RNA polymerase, and, specifically, the effect of modifying a highly conserved domain.

Five Cysteine Residues are Conserved in the β' subunit of RNAP.

Within the β' subunit lies five highly-conserved cysteine residues. These amino acid residues are present in all prokaryotes regardless of type (*e.g.,* gram-positive, gram-negative, anaerobic, aerobic, methane-dependent etc). Additionally, the sequences are present in the mitochondria and chloroplast DNA of eukaryotes, which are derived from prokaryotic cells according to the endosymbiotic theory. The conservation of these amino acid sequences across such organisms is displayed in Figure 2.


rpoc_ecoli 100.0%
MKDLLKFLKAQTKTEEFDAIKIALASPDMIRSWSFGEVKKPETINYRTFKPERDGLFCARIFGPVKDYECLCGKYKRLKHRGVICEKCGVEVTQTKVRRERM
rpoc_salti 100.0%
MKDLLKFLKAQTKTEEFDAIKIALASPDMIRSWSFGEVKKPETINYRTFKPERDGLFCARIFGPVKDYECLCGKYKRLKHRGVICEKCGVEVTQTKVRRERM
rpoc_salty 100.0%
MKDLLKFLKAQTKTEEFDAIKIALASPDMIRSWSFGEVKKPETINYRTFKPERDGLFCARIFGPVKDYECLCGKYKRLKHRGVICEKCGVEVTQTKVRRERM
rpoc_pholl 100.0%
MKDLLKFLKAQTKTEEFDAIKIALASPDMIRSWSFGEVKKPETINYRTFKPERDGLFCARIFGPVKDYECLCGKYKRLKHRGVICEKCGVEVTQTKVRRERM
rpoc_yerpe 100.0%
MKDLLKFLKAQTKTEEFDAIKIALASPDMIRSWSFGEVKKPETINYRTFKPERDGLFCARIFGPVKDYECLCGKYKRLKHRGVICEKCGVEVTQTKVRRERM
rpoc_aquae 71.8%
FSKIKLMLASPEDIRSWSHGEVKRPETLNYRTLKPEKDGLFCAKIFGPIKDYECLCGKYRGKRYEGKICEKCGVEVTTSYVRRQR-
rpobc_helhp 71.3%
DFSSFQLVLASPEKILSWSNGEVKKPETINYRTLKPERDGLFCTKIFGPVRDYECLCGKYKKMRYKGIVCEKCGVEVTKAKVRRSRM
rpoc1_anasp 70.8%
TNQFDYVKIGLASPERIRQWgvGEVTKPETINYRTLKPEMDGLFCERIFGPAKDWECHCGKYKRVRHRGIVCERCGVEVTESRVRRHRM
rpoc1_synp2 75.8%
EVTKPETINYRTLKPEMDGLFCERIFGPAKDWECHCGKYKRVRHRGIVCERCGVEVTESRVRRHRM
rpoc1_proho 75.8%
EVTKPETINYRTLKPEMDGLFCERIFGPAKDWECHCGKYKRVRHRGIVCERCGVEVTESRVRRHRM
rpoc1_fismu 75.8%
EVTKPETINYRTLKPEMDGLFCERIFGPAKDWECHCGKYKRVRHRGIVCERCGVEVTESRVRRHRM
rpoc_urepa 63.0%
ISIASPEQILNWSKGEITKPETINYKSLKPEPNGLFDESIFGPSKDYECYCGKYRKVKHKGKICERCHVEITESIVRRERM
rpoc1_nosco 60.7%
KDWECHCGKYKESVIEVLSVSAVVLEVTESRVRRHRM
rpoc1_nepol 70.7%
rpoc1_adica 54.2%
LRIGLASPEQIRSWaiGQVDQPYTLHYKTHKPERDGLFCERIFGPTKSGVCACGNCRSVNDEGefCKHCGVEFTDSRVRRYRM
rpc1a_chlre 69.2%
rpoc1_phypa 55.2%
GQVTKPYTLHYKTHKPEKDGLFCERIFGPIKSGICACGKYQIIEKYSKFCEQCGVEFVESRVRRYRM
rpoc1_antfo 55.2%
GRITEPYTLHYKTHKPEKDGLFCERIFGPIKSGICACGKYRSIENqsKICEQCGVEFTESRVRRYRM
rpoc1_oenho 65.9%
rpoc1_marpo 52.2%
GQVTKPYTLHYKTHKPEKDGLFCEKIFGPIKSGICACGKYQGikENIKFCEQCGVEFIESRIRRYRM
rpoc1_psinu 52.2%
GKVTQPYTLHYKTHKPERDGLFCERIFGPIKSGFCACGNYQAvkEFSSFCKQCGVEFTESRVRRYQM

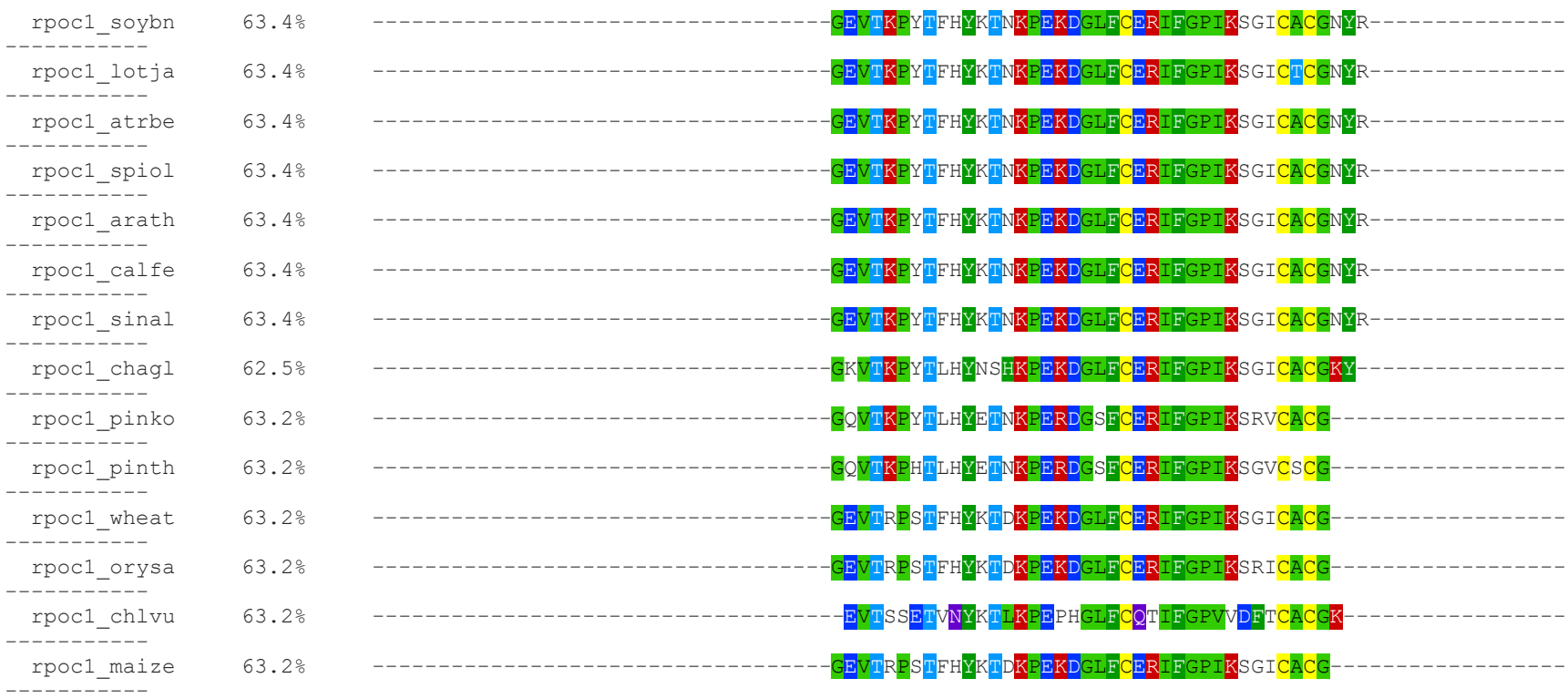


Figure 2: Conservation of cysteine residues in the *E. coli* β' subunit of RNA Polymerase (top) in various other organisms. Sample rpoc1_wheat is an example of conservation of *rpoC*-C58 in chloroplasts. Data obtained from the lab of Dr. Rodney King.

Four of the five conserved cysteines form a zinc-binding domain. Mutation of residues C70, C72, C85, and C88—which form the zinc-binding domain—is deleterious to the cell and affects response to termination and *put*-mediated antitermination (King et al., 2004). Anti-termination is a process that allows the expression of genes for which transcription is usually turned off via transcription termination. Terminators are important gene regulators and are thought to be useful in the temporal regulation of genes (Santangelo & Artsimovitch, 2011). For example, the timing of the expression of phage Lambda late genes, which are involved in cell lysis, must be carefully regulated. Transcription terminators may contribute to the temporal regulation of these genes by preventing premature expression and, consequentially, premature cell lysis (Chang, Nam, & Young, 1995). However, phage late genes must eventually be expressed, as they are required for the completion of the lytic life cycle (Wang, 2005). To accomplish expression of these genes, transcription termination is overridden by anti-termination mechanisms. In phages such as Lambda, this process is accomplished through the interaction of antitermination factors with RNAP which block termination by preventing the disassociation of the ternary complex at the termination site (Adhya & Gottesman, 1978). In an alternative mechanism for anti-termination, RNA transcripts directly modify RNAP to overcome transcription termination. The *put* (polymerase utilization) site RNA is thought to form a double stem-and-loop secondary structure that interacts with the β' subunit of bacterial RNAP to change the conformation of the elongation complex, enabling transcription of genes that lie downstream of terminators. *Put*-mediated anti-termination was discovered in the bacteriophage HK022. Mutation of the conserved residues C70, C72, C85, and C88 reduces RNAP response to intrinsic terminators and

decreases *put*-mediated antitermination (King et al., 2004.). The role of the fifth conserved cysteine residue, *rpoC*C58, in *put*-mediated anti-termination has not been investigated.

In addition to their role in anti-termination, these conserved cysteines have also been shown to play a role in the processivity of transcription elongation (Nudler et al.,1996). Nudler *et al.* substituted both C85 and C88 with serines and compared the mutant RNAP to the wild-type enzyme. To assess the effect of the mutation on RNAP function, ternary complex of both mutant and wild-type RNAP with a 20-nucleotide transcript was prepared with RNAP in excess. The mutant RNAP produced shorter RNA products of parts of the transcribed region, indicating premature disassociation of the RNAP complex. In contrast, the wild-type RNAP yielded the stochiometric amount of the transcript. Furthermore, while the wild-type RNAP complex remained intact in high-salt conditions, the mutant RNAP complex fell apart. Thus, Nudler et al. concluded that mutation of these conserved cysteines reduced RNAP processivity, increased sensitivity to salt, and prevented cell growth in the absence of the wild type *rpoC* allele (Nudler et al., 1996). However, results from King et al. (2004) contradicted these findings, indicating that the zinc binding domain is not needed for the stability or activity of the elongation complex, suggesting this proposed function is not the reason for the high conservation of the four conserved cysteine residues.

While these four cysteine residues have been shown to play an important role in *put*-mediated anti-termination, the fifth cysteine residue, *rpoC-C58,* has no clear contribution to RNAP function. The high conservation of *rpoC*-C58 and its proximity to four other critical residues suggests the cysteine residue may have an important role in

RNAP function. Previous evidence from our lab suggested a potential function for *rpoC*-C58. An Alanine substitution appeared to confer temperature sensitivity to *E. coli* cells. In previous experiments, a plasmid containing *rpoC*-C58A on a plasmid was introduced into a temperature-sensitive strain of *E. coli.* However, this plasmid failed to complement the temperature-sensitive defect, suggesting *rpoC*-C58 confers temperature sensitivity on the enzyme. It is thus necessary to study the potential temperature-related function of C58 with the *rpoC*-C58A mutation present in single copy. The research described here investigated the functional role of *rpoC*-C58A and its potential relationship with temperature sensitivity.

CHAPTER TWO: METHODS

Bacterial Strains.

The bacterial strains used in this study are detailed in Table 2.

Growth of bacterial cultures.

Unless otherwise noted, all bacterial cultures were grown under the following standard conditions. Cultures were inoculated from frozen archived stocks into 5mL of LB in 15mL conical centrifuge tubes. Cultures were shaken at 200 rpm overnight at either 30°C or 37°C, with strains containing *rpoC*-C58A grown at 30°C due to suspected temperature sensitivity. Strains not suspected to be temperature sensitive were grown at 37°C. After approximately 12 hours of growth, cultures were centrifuged at 3,000 rcf (Eppendorf Centrifuge 5702 R) at 4°C for 10 minutes. The supernatant was decanted, and the pellet was suspended in 2.5mL of 10mM $MgSO_4$. Stocks were stored at 4°C up to one month.

Strain Name	*Genotype*	*Source*
RK1106	*E. coli* strain MG1655 (WT *rpoC*) transformed with an ampicillin resistant pBAD*rpoC*-C58A plasmid construct (Sen et al., 2002). Construct contains the *rpoC*-Y75N mutation and a cloned copy of the *rpoC* gene with Cysteine 58 substituted with an Alanine	King R. A., unpublished (MG1655: Guyer, M. S., 1981)
RK1008	*E. coli* strain SBS672 (*rpoC*-Y75N strain that contains HKpL-putL-Tr'T1T2-lacZ fusion) transformed with pRW4714, an ampicillin resistant plasmid that confers recombineering functions	pRW4714: Datta et al., 2006
RK1004	*E. coli* strain SBS672 (*rpoC*-Y75N strain that contains HKpL-putL-Tr'T1T2-lacZ fusion)	Sloan et al., 2007
RK1005	*E.* coli strain SBS650 (WT *rpoC* strain that contains HKpL-putL-Tr'T1T2-lacZ fusion)	Sloan et al., 2007
RK898	WT *E. coli,* MG1655	Guyer, M.S., 1981
RK1120	Strain containing plasmid PGB2ts, which confers resistance to chloramphenicol.	Clerget, 1991
RK928	Stain RW4207 (MG1655, *rpoC*397c (RK486), recA – Tn10). Contains temperature-sensitive *rpoC.*	King, R.A.
RK486	*rpoC*397c. MG1655 with temperature-sensitive *rpoC.*	Christie et al., 1996

Table 2. Bacterial strains. Abbreviations: Wild Type (WT).

Plasmid isolation.

Plasmid isolation was performed with the Qiaprep Spin Miniprep Kit (Qiagen; ID: 27104). All centrifugations were performed at 11,600 rcf (Eppendorf 5452 Minispin Centrifuge; Rotor F-45-12-11) at room temperature unless otherwise noted. A 1mL volume of cell suspension in 10mM $MgSO_4$ was added to two 1.5mL microcentrifuge tubes. Then, the tubes were spun for 3 minutes. The supernatant was decanted, and the pellets were suspended in 250µL Buffer PI. A 250µL volume of Buffer P2 was added to the suspension. Tubes were mixed by inverting until the solution turned clear. Next, 350µL of Buffer N3 were added. Tubes were mixed by inverting 8 times and centrifuged for 10 minutes. Following centrifugation, 800µL of supernatant were transferred to the spin column and centrifuged for 1 minute. The flow-through was decanted. A 500µL aliquot of Buffer PB was then added, and columns were centrifuged for 1 minute and the flow-through was decanted. Then, 750µL of Buffer PE were added to the columns and centrifuged for 1 minute. After decanting the PE buffer, the columns were centrifuged and additional 1 minute and transferred to new collection tubes. The DNA was recovered from the columns by adding 50µL of elution buffer, incubating at room temperature for 1 minute then centrifuging for 1 minute. The eluate containing the plasmid DNA was collected and stored at -20°C.

Polymerase Chain Reaction.

PCR reactions were prepared in 100µL PCR tubes. The following volumes of each reactant were added in the order given in Table 3. Note that for PCRs performed prior to sequencing, the primer concentration was 1/8X of the concentration listed in Table 3. This concentration eliminated occasional primer dimer formation observed in

earlier experiments. The PCR was run in an MJ Research PTC-200 Peltier Thermal Cycler on the cycle described in Table 4. Following completion of the program, 5µL of each PCR reaction were run on a 1.2% agarose gel at 120 volts to confirm successful amplification. Reactions were stored at -20°C until further use. Oligonucleotides used as primers are listed in Table 5.

Volume	*Component*
66µL	Deionized Water (Thermo Scientific Barnstead Nanopure)
30µL	PCR mix (Appendix A)
1µL	DNA template
1µL	Forward primer (100 µM)
1µL	Reverse primer (100 µM)
1µL	Taq DNA Polymerase (Fisher Bioreagents; 5 U/µL; FB-6000-15)
100µL	Total reaction volume

Table 3: PCR components and volumes per 100µL reaction.

Step	*Temperature (°C)*	*Duration (minutes)*
1	94	2
2	94	0.5
3	55	0.5
4	72	1
5	Repeat from step 2 for 30 repetitions	
6	4	Forever
7	END	

Table 4. PCR reaction conditions.

Name	*Sequence*	*Use*
RK121	TAGTCAACACGCTTACCGAGC	Primes within *rpoC*
RK170	ATTAAAGTTTCTGAAAGCGCAG	Primes at the beginning of *rpoC* (at approximately the fourth amino acid)
RK219	CAGAATCCTTCAACGTATTG	Primes immediately after the SphI site at the C-terminus of *rpoB*
RK229	GGGAAGCCAGTTCGATG	Primer for amplifying from *rpoC*. Located approximately 60bp upstream of the Y75N codon
RK818	AACGGTCGTACCAAGATG	Primes within *rpoB*

Table 5: Oligonucleotides used for PCR.

PCR reaction purification and concentration.

A QIAquick PCR purification Kit (Qiagen, ID: 28104) was used to concentrate PCR-amplified DNA fragments. Approximately 95µL of PCR reaction was added to a 1.5mL microcentrifuge tube. Then, 5 volumes of Buffer PB were added to the microcentrifuge tubes, mixed 8 times, and added to the provided column. The column was spun for 60 seconds at 12,000 rcf. Flow-through was discarded and the wash was repeated with 750µL of PE. The flow-through was again discarded. Columns were then placed in clean 1.5mL microcentrifuge tubes. Then, 30 µL of buffer EB were added to the columns. Columns were incubated at room temperature for 2 minutes and centrifuged as before.

Gel purification of DNA fragments.

A 10µL aliquot of 6X loading dye (Appendix A) was added to each 50µL of concentrated PCR product. A 100bp DNA ladder was prepared by mixing 1µL of DNA ladder (New England Biolabs N3231S; 500µg/mL) with 1µL of 6X loading dye (Appendix A) and 4µL of sterile Nanopure water and loaded onto the gel. The entire volume of PCR product plus loading dye was loaded into the wells of a 1.2% gel (6µL aliquots/well). The gel was run at 100V, constant voltage, for approximately 45 minutes using a Fisher Scientific FB300 Power Supply. To visualize DNA bands, gels were placed in ethidium bromide (0.5µg/mL) for 10 minutes. Gels were then placed on a transilluminator, and the desired bands were excised using a new razor blade. Gel fragments containing DNA bands were collected in 1.5mL tubes and chopped into smaller fragments using a sterile 80-gauge beveled needle. Next, 800µL of phenol were added to all tubes. The tubes were mixed well and incubated at -80°C for 5 minutes.

After freezing the samples, they were centrifuged for 10 minutes at 12,000 rcf (Eppendorf 5452 Minispin Centrifuge; Rotor F-45-12-11). Approximately 500µL of the aqueous top layer, which contained the desired *rpoC*-C58A DNA fragment, was recovered from each tube and stored at -20°C until further processing.

Processing and ethanol precipitation of gel-purified DNA fragments.

Frozen DNA samples were thawed at room temperature. Then, 500µL of chloroform were added to each tube. Tubes were vortexed for 8 seconds, then centrifuged for 1 minute at 12,000 rcf (Eppendorf 5452 Minispin Centrifuge; Rotor F-45-12-11). The upper aqueous phase was then transferred to a new tube. This process, from addition of the chloroform through removal of the aqueous phase, was repeated twice for a total of three extractions. Next, the aqueous layers from all tubes were combined, and the combined volume was equally distributed into clean microcentrifuge tubes. Then, $1/10^{th}$ of the total volume of sodium acetate (3M, pH 5.2) was added to each tube and mixed. Two volumes of 100% ethanol were then added to each tube. The tubes were mixed well and stored at -20°C until processing was continued.

After chilling at -20°C, the ethanol DNA precipitates were then pelleted by centrifugation. Samples were centrifuged at 12,000 rcf for 30 minutes at 4°C. The supernatant was decanted from all tubes, and pellets were allowed to dry lying on their sides at room temperature. Then, 100µL of 10mM Tris-Cl (pH 8) were added to the first tube and the pellet was rinsed well. This mixture was then transferred through the remaining tubes until each pellet containing the ethanol-precipitated DNA had been rinsed. To collect any remaining DNA, the tubes were rinsed a second time with 100µL of Nanopure water, which was combined with the 100µL of 10mM Tris-Cl. To

concentrate the precipitated DNA, filtration units as described in *PCR reaction purification and concentration,* were used. The final recovered product was stored at -20°C.

Measuring DNA concentration with a NanoDrop Spectrophotometer.

The DNA concentration and purity of recovered DNA samples were measured using a NanoDrop spectrophotometer (ThermoFisher Scientific; No. ND-2000). Following cleaning of the pedestal with a Kimwipe, the spectrophotometer was blanked with 2 µL of elution buffer. Then, the pedestal was wiped dry, and 2µL of sample were added. Sample concentration was measured in ng/µL, and DNA quality was assessed using the measured A260/280 and 260/230 ratios.

Gel electrophoresis.

Unless otherwise noted, a 1.2% agarose gel was prepared by adding 1.2g of DNA-grade agarose (Fisher Scientific, BP 164-500) to 100mL of 1X Tris-Acetic Acid Ethylenediaminetetraacetic acid (TAE; 40mM Tris, 20mM acetic acid, 1mM EDTA) solution in a 250mL Erlenmeyer flask. The mixture was heated until the agarose fully dissolved. If a 2% agarose gel was used, it was prepared by dissolving 2.0 g of DNA-grade agarose in 100mL of 1X TAE. Then, 30mL of the agarose mixture were transferred into a gel apparatus with a ten-well mold and allowed to solidify. The resulting gel was placed with the wells on the side of the cathode and submerged in 1X TAE.

A 100bp DNA ladder was prepared by mixing 1µL of DNA ladder (New England Biolabs N3231S; 500µg/mL) with 1µL of 6X loading dye (Appendix A) and 4µL of sterile Nanopure water. DNA ladder was loaded into the first well. A 5µL volume of each DNA sample was mixed with 1µL of 6X loading dye and added to a lane. The apparatus

was then covered with a lid and connected to a voltage supply (Fisher Scientific FB300 Power Supply) and run at 100V with constant voltage for approximately 40 minutes, or until the indicator on the 100bp ladder indicated sufficient band separation. To visualize the DNA bands, the gel was placed in an ethidium bromide solution (0.5μg/mL) for 10 minutes and imaged under UV light (FluoroChem HD2 Imager, EtBr Colorimetric).

Preparation of cells for electroporation and recombineering.

Recombineering was employed to insert the DNA containing the C58A mutation into *E. coli* chromosome. First, RK1008 cells suspended in 10mM MgSO4 were prepared for electroporation as described in *Growth of bacterial cultures*. Cultures containing 30mL LB (Appendix A) and 50μg/mL ampicillin were inoculated with 200μL of RK1008 cell suspension. Following approximately 4 hours of growth at 37°C, the optical density at 600nm of a 100μL aliquot of culture was determined (Shimadzu BioSpec-mini, CAT: 241-06250-92). Once the OD_{600} reached approximately 0.4 (indicating early logarithmic phase), cultures were placed in a 42°C water bath to induce the expression of the recombineering functions for induction; the cultures were shaken at 250 rpm for 20 minutes. After induction, cultures were transferred to an ice-water slurry and swirled for three minutes to turn off the expression of the recombineering functions. Next, 12mL of the cooled culture were transferred into two 15mL conical centrifuge tubes. Cultures were then centrifuged at 4°C and 3,000 rcf (Eppendorf Centrifuge 5702 R) for 10 minutes. The supernatant was decanted, and pellets were suspended in 5mL 10% ice-cold glycerol. The wash step was repeated for a total of 2 washes. After the washes, the cells were pelleted by centrifuging with the same settings as above and the supernatant was decanted. The pellets were suspended in approximately 1.3mL of ice-cold 10% glycerol

to achieve a total volume of 1.5mL. This suspension was transferred to microcentrifuge tubes and centrifuged at 4°C and 5,400 rcf (Eppendorf 5452 Minispin Centrifuge; Rotor F-45-12-11) for 10 minutes. Most of the supernatant was removed by pipetting until a 50µL volume remained. The pellet was suspended in the glycerol that remained in the tubes. The electrocompetent cells were used immediately or stored at -80°C until further use.

Electroporation with DNA fragment.

Electrocompetent cells were thawed on ice then electroporated with the purified DNA fragment. Typically, 441.2ng of DNA fragment were added to 50µL of electrocompetent cells in a 2mm electroporation cuvette (Bioexpress E-510-2). The mix was then electroporated with the settings given in Table 6.

Voltage	2500 V
Capacitance	25 µF
Resistance	200 Ω
Cuvette	2mm
Time Constant	5 milliseconds

Table 6: Electroporation settings.

Immediately after electroporation, 1mL of SOC media (Appendix A) was added to the cuvette and mixed. The mixture was transferred to a conical centrifuge tube and incubated at 30°C for 30 minutes to allow cells to recover.

Electroporation with plasmid.

Electroporation was performed as described in *Electroporation with DNA fragment.* Instead of electroporating with 441.2ng of DNA, cells were electroporated with a minimum of 25ng of isolated plasmid.

Screening for the desired recombinants.

To determine the proper concentration of cells for colony screening, a series of dilutions was made to determine the optimal concentration to achieve the maximum number of cells per plate without saturating plates with bacterial growth. One hundred microliters of each dilution were spread-plated onto X-gal plates (Appendix A). Plates were allowed to dry and incubated for approximately 48 hours at 30°C.

Restriction enzyme digest screen to identify rpoC-C58A mutants.

The DNA sequence that encodes cysteine at position 58 of the *RpoC* protein contains a restriction enzyme recognition site for the enzyme HhA1. When the DNA is mutated to yield an alanine at position 58 of the RPOC protein, this restriction site is destroyed. Thus, if the C58A change is successfully incorporated into the chromosome of the host, HHA1 will no longer cut. This rationale served as the foundation for a streamlined screening of potential C58A mutants that does not rely on DNA sequencing. PCR amplification was performed using primers that amplified the *rpoC*-C58 region and HhaI restriction site but no other HhaI restriction sites that are present in the *rpoC* gene sequence. Then, a Hha1 restriction digest was performed on the amplified DNA. Wild type DNA generated two DNA fragments that are smaller than the no-enzyme control DNA. By comparing DNA samples digested with HHA1 to no-enzyme controls, one can determine whether the restriction site was destroyed.

For each suspected mutant, two restriction digests were prepared according to Table 7. Reactions were prepared in 0.1mL PCR tubes. Then, reactions were incubated at 37°C for 1 hour. Following the digest, tubes were stored in -20°C until further use.

Component	*Candidate Sample – No Enzyme Control*	*Candidate Sample-Enzyme*
Enzyme (HhA1)	--	1µL (20,000 units/mL)
DNA (PCR-amplified DNA)	10µL	10µL
Buffer	2µL	2µL
Water	8µL	7µL

Table 7: Restriction enzyme digest screen setup. HhA1: New England Biolabs, R0139S.

DNA sequencing.

Samples were prepared for sequencing by combining 10µL of 1ng/µL amplified DNA with 5µL of primer at 5µM. Sequencing reactions were sequenced by a commercial vendor (Genewiz).

Quantitative β-galactosidase assays.

Overnight cultures were diluted 1/500 in 25mL of LB. Diluted cultures were grown at either 30°C or 42°C and shaken at 200 rpm. The optical density at 600nm of a 100µL aliquot of culture was monitored throughout the incubation period (Shimadzu BioSpec-mini, CAT: 241-06250-92). Once the OD_{600} reached approximately 0.4 (indicating early logarithmic phase), cultures were placed on ice. Then, β-galactosidase were performed according to the procedure given by Miller (1992). Assays were prepared by combining 0.5mL of culture with 0.5mL of Z-buffer (Appendix A) in a 10x75 mm

glass tube. Then, 50µL of 0.1% SDS (sodium dodecyl sulfate) solution and 100µL of chloroform were added to the tube. Tubes were vortexed for 8 seconds, then incubated at 28°C for a minimum of 5 minutes. To start the reaction, 200µL of ONPG (*o*-nitrophenyl-β-D-galactopyranoside; 4mg/mL; Appendix A) was added to the tube and mixed well. When a bright yellow color began to develop, 500µL of 1M sodium carbonate was added to the mixture to stop the reaction. The time it took to develop yellow color was recorded. A blank solution of Z-buffer, 0.1% SDS, chloroform, LB, and sodium carbonate was prepared in proportion to the volumes in the reaction tubes, and the absorbance of the reactions at 420nm (to measure hydrolysis of the substrate) and 550nm (to correct for light scattering by cell debris) were recorded using a Shimadzu BioSpec-mini (CAT: 241-06250-92). Then, Miller Activity Units of enzyme activity were calculated according to Equation 1.

$$Miller\ Activity\ Units = 1000 \times \frac{OD_{420} - (1.75 \times OD_{550})}{OD_{600} \times reaction\ time}$$

Equation 1: Miller Activity Units.

Archive of strains.

An 800µL aliquot of cells suspended in 10mM $MgSO_4$ was mixed with 200µL of 80% glycerol in a cryotube, mixed well, and stored at -80°C.

CHAPTER THREE: RESULTS

Generating an amino acid substitution in the E. coli β'subunit of RNA polymerase using recombineering.

Previous work in our lab suggested that an alanine substitution at position 58 of the β' subunit protein conferred a temperature sensitive phenotype: a cloned copy of a mutated *rpoC* gene failed to complement the growth defect of a temperature sensitive strain of *E. coli* (strain RK928). The cysteine at position 58 is one of the most highly conserved amino acids in the β' subunit, but its function is unknown. Thus, this study set out to determine whether the observed temperature sensitivity is a true biological phenomenon or perhaps an experimental artifact. Expressing the *rpoC*-C58A gene from a plasmid may be problematic for a number of reasons. First, many copies of the gene would be present rather than a single copy. Second, the gene would not be under the control of its natural promoter. To ensure these issues were not responsible for the previous results, the potential temperature sensitive phenotype of *rpoC*-C58A mutant must be confirmed in a strain containing the mutation in single copy and in its natural chromosomal location.

To generate the *rpoC*-C58A mutation in the chromosome, a technique called recombineering was employed. Recombineering is a genetic engineering technique based on homologous recombination that enables site-specific modification of chromosomal DNA. Like CRISPR-cas9, recombineering allows *in vivo* modification of the genome,

unlike its *in vitro* predecessor techniques of genetic engineering. A schematic of recombineering in the context of this experiment is given in Figure 3.

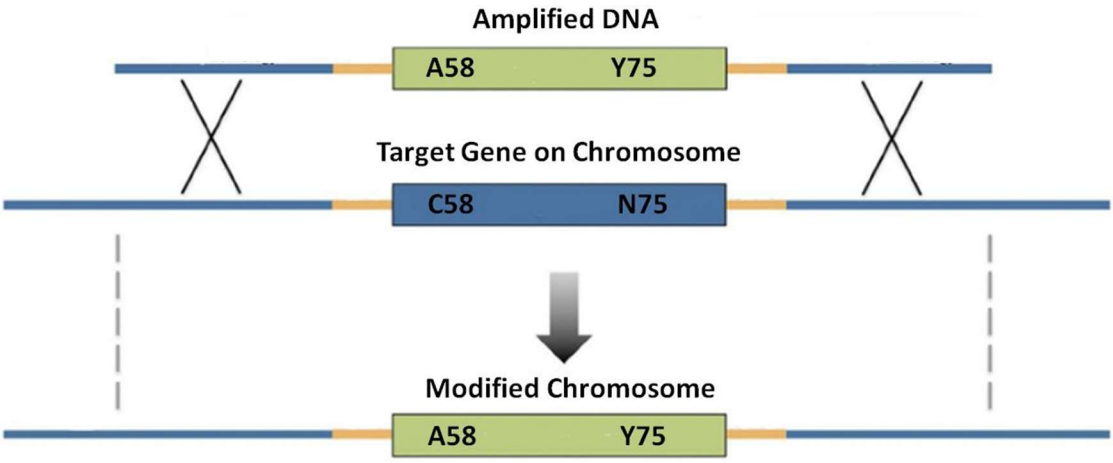


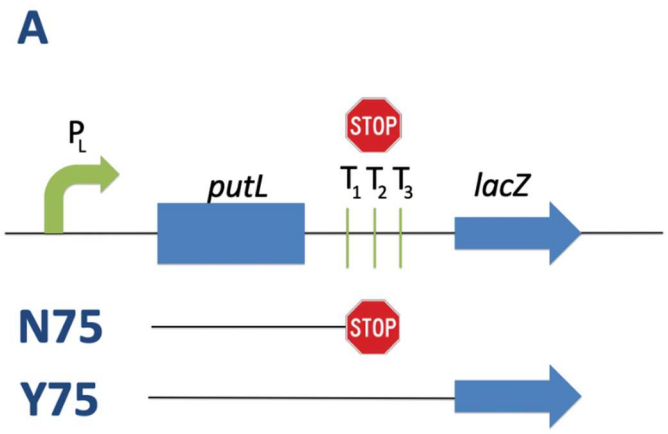
Figure 3. A diagram of recombineering. DNA containing the desired DNA sequence is added to the cells and replaces the original DNA sequence.

When recombineering is performed in *E. coli* and the recombineering functions are supplied by a defective λ prophage, recombination efficiency reaches 1% (Ellis et al., 2001). Such efficiency is possible because expression of the phage genes from their endogenous promotor enables tight regulation and coordinated expression. This prophage-based system uses a temperature-sensitive repressor to tightly repress the phage recombination functions at temperatures between 30°C and 34°C degrees but highly expresses these genes at 42°C (Thomason et al., 2005). When this prophage-based system is combined with host mismatch repair inactivation, the recombination frequency can be as high as 20% to 25% (Constantino and Court, 2003). Because of the high frequency of

recombination achieved with this method, it is possible to use a screen rather than selection to identify desired recombinants.

To facilitate the identification of recombinants, recombineering was performed on a strain that contained a *lacZ* reporter gene construct. This strain, RK1008, contains the following useful features (Figure 4): first, it contains a single copy of a *lacZ* reporter fusion in its chromosome. The *lacZ* gene in this strain can only be expressed if transcription termination at a set of terminators located between the promoter and the *lacZ* reporter are suppressed. Antitermination in these fusions is normally promoted by a cloned copy of the HK022 *putL* site located just upstream of the transcription terminators (King et al., 1996). However, RK1008 also contains the *rpoC*Y75N substitution (Clerget et al., 1995) which prevents *put*-mediated anti-termination and consequently prevents the expression of the *lacZ* reporter. By replacing the chromosomal copy of *rpoC* with DNA containing both the C58A substitution and the N75Y reversion, antitermination should be restored and the cells should express the *lacZ* reporter. Expression of *lacZ* was used to distinguish the desired recombinants from unmutated cells through a standard colorimetric assay (e.g. blue-white screening using media supplemented with the X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, a synthetic chromogenic substrate of lactose). Successfully generated mutants would produce the enzyme beta-galactosidase ("beta-gal"), the protein product of *lacZ*. Beta-gal can cleave X-gal. Cleaved X-gal spontaneously dimerizes and oxidizes to create a blue insoluble pigment, 5,5'-dibromo-4,4'-dichloro-indigo. Colonies that have been successfully recombineered should have the *rpoC*-N75Y reversion. Thus, when grown on agar plates containing X-gal, recombinants

produce a blue color, while the parental *E. coli* strain (RK1008) that contains the *rpoC*-Y75N substitution and wild-type cysteine at position form white colonies.



rpoC	Phenotype
N75	White
Y75	Blue

Figure 4: (A) Schematic of antitermination and its relation to *lacZ* reporter expression. In *rpoC*-Y75N cells, RNAP stops transcription at the terminators (indicated by the stop sign). In *rpoC*-Y75 cells, antitermination occurs. This allows read-through of *lacZ.* (B) *rpoC*-75 residues and their corresponding phenotype.

The desired DNA fragment containing the *rpoC*-C58A substitution (the recombineering substrate) and the wild type Tyrosine at position 75 was generated by PCR amplification. The 316-base pair recombineering substrate was amplified from a pBADrpoC-C58A plasmid construct (isolated from RK1106) with primers RK170 and RK229 (Figure 5). The size of the amplified fragment was verified by gel electrophoresis on a 1.2% gel (Figure 6). After purifying the amplicon, recombination was performed using appropriately prepared cells of RK1008, the recombineering strain. Electroporated cells were appropriately diluted and plated on medium containing X-gal and grown at 30°C.

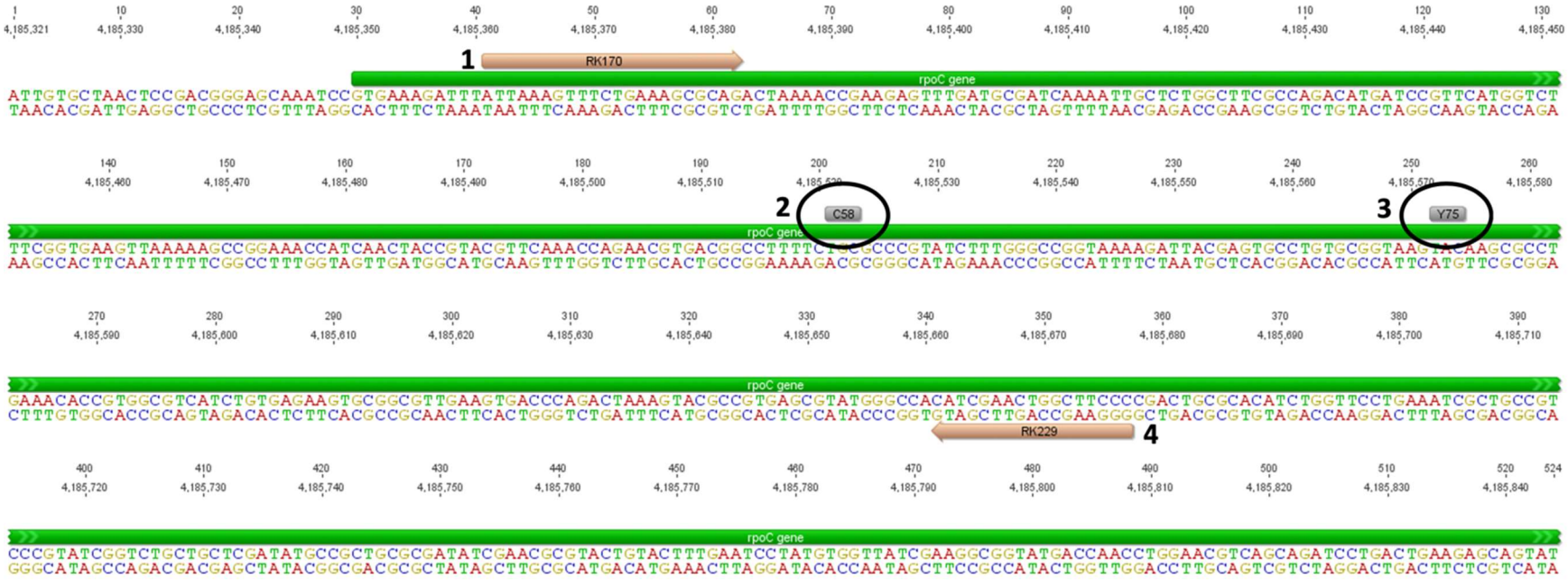


Figure 5. Location of primer RK170 (1) and RK229 (4) in relation to *rpoC.* The location of C58 (2) and Y75 (3) on *rpoC* are indicated.

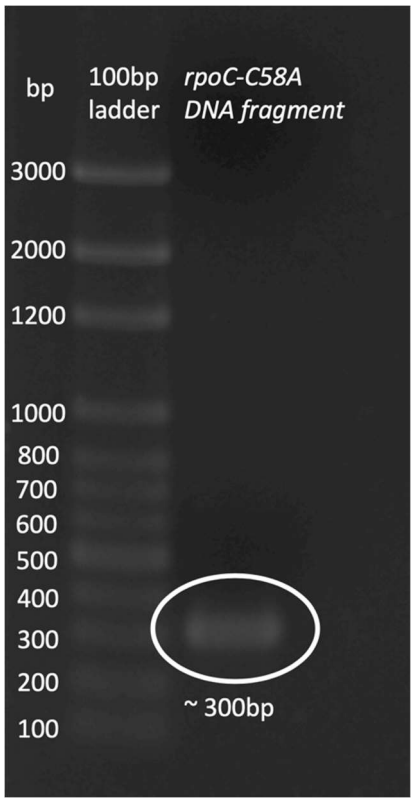


Figure 6. Agarose gel (1.2%) of the isolated *rpoC*-C58A DNA recombineering substrate. Band size of approximately 300bp in reference to a 100bp DNA ladder indicates successful amplification of the 318bp DNA fragment.

Initially, recombineering was performed according to the protocol described in Current Protocols in Molecular Biology (2007). However, very few potential recombinants were generated, and the ones that were analyzed contained the *rpoC*-N75Y reversion but not the *rpoC*-C58A substitution. Three such *rpoC*-Y75N revertants were identified in a screen of 760,000 colonies. This result led to the speculation that *rpoC*-C58A substitution might be deleterious, resulting in a RNAP protein unable to support cell growth. However, when the recombineering experiment was repeated with a concentration of the recombineering substrate double that of previous trials (441.2ng/µL vs 220.6ng/µL), we observed an increase in the frequency of probable mutants; the estimated frequency for initial trials was extremely low (~0.000394%) but increased to ~0.0145% in the second trial.

Due to the low frequency of successful mutants generated in our hands, a selection for the desired mutation would be ideal. However, a selection was not available, and it was therefore necessary to implement a screen as previously described. To identify the desired recombinants, it was necessary to plate a large quantity of recombineered cells. The dilution that led to the growth of the most colonies without reaching confluence was empirically determined and selected for spread plating. Initial plating of dilutions of recombineered RK1008 cells resulted in the formation of nine blue colonies. However, previous attempts to generate *rpoC*-C58A mutants showed that mutants containing *rpoCY75* but not *rpoC-C58A* were frequently generated. It was therefore necessary screen more cells to identify more candidate *rpoC-C58A* mutants. Spread plates were made and grown at 30°C for approximately 48 hours. In total, 99 blue colonies were ultimately identified (Figure 7) out of an estimated 681,000 screened colonies.

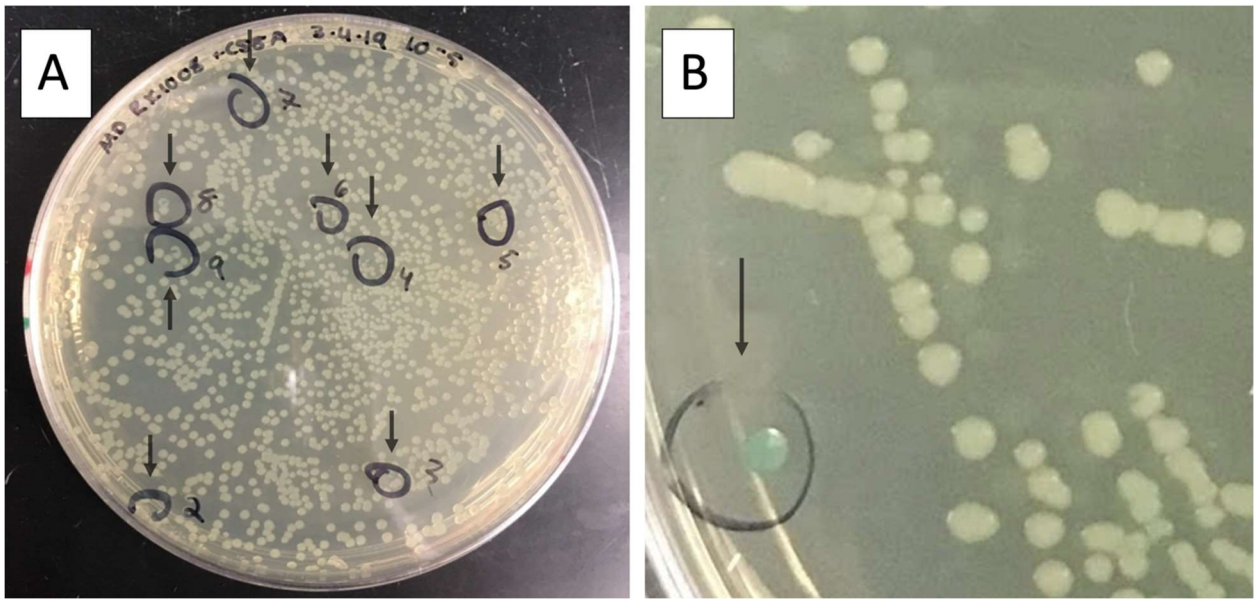


Figure 7. A: Plates containing blue colonies (circled on plates). These colonies were identified as possible *rpoC*-C58A mutants. B: Positive (blue) phenotype of recombinant cells.

Verification of the rpoC-C58A mutation.

The DNA amplicon containing the *rpoC*-C58A substitution and the N75Y reversion (the recombineering substrate) is only 318 base pairs in length and the codon for Y75 is a mere 48 base pairs away from the C58A codon. Thus, it was expected that if the Tyrosine codon at position 75 was recombineered into the chromosome of the cell, the Alanine codon at position 58 would also be incorporated because it is tightly linked. In our initial attempts to generate the C58A substitution mutants, this did not occur. DNA sequence analysis of the potential recombinants showed the sequences contained the *rpoC*-N75Y revertant, but the cysteine at position 58 of the beta subunit protein was unchanged. Thus, the presence of blue colonies, which indicates the successful insertion of the tyrosine codon at position 75, did not appear to be sufficient to verify the presence

of the C58A mutation. It was also important to distinguish potential spontaneous N75Y reversions from successfully generated recombinants. DNA sequencing of the recombineering substrate is an effective way of verifying the presence of both desired mutations, but his approach is inefficient and expensive. Thus, we searched for a different way to specifically verify the presence of the C58A mutation

The DNA sequence that encodes a cysteine at position 58 of the *RPOC* protein contains a recognition site for the enzyme Hha1. When the DNA is mutated to generate an alanine codon at this location, the restriction site is destroyed and Hha1 will no longer cut at this location. To exploit the observed restriction fragment length polymorphism (RFLP), a DNA fragment containing the single Hha1 cut site was PCR amplified with primers RK312 and RK313 (Figure 8). After digesting with HhaI, the samples were run on a 2% agarose gel. When wild type DNA was digested, two DNA fragments were generated. This cut DNA traveled further and was easily distinguished from the control uncut DNA, whereas DNA containing the C58A mutation traveled the same distance as the control, as expected. Using this method, 21 likely *rpoC*-C58A mutants were identified (Figure 9).

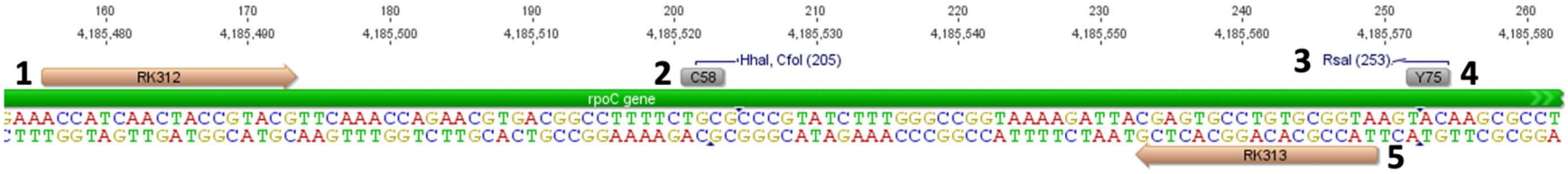


Figure 8. Geneious alignment of primer RK312 (1) and RK312 (5) in relation to *rpoC.* The location of C58 (2) and Y75 (4) as well as the Hha1 restriction site (3) on *rpoC* are indicated.

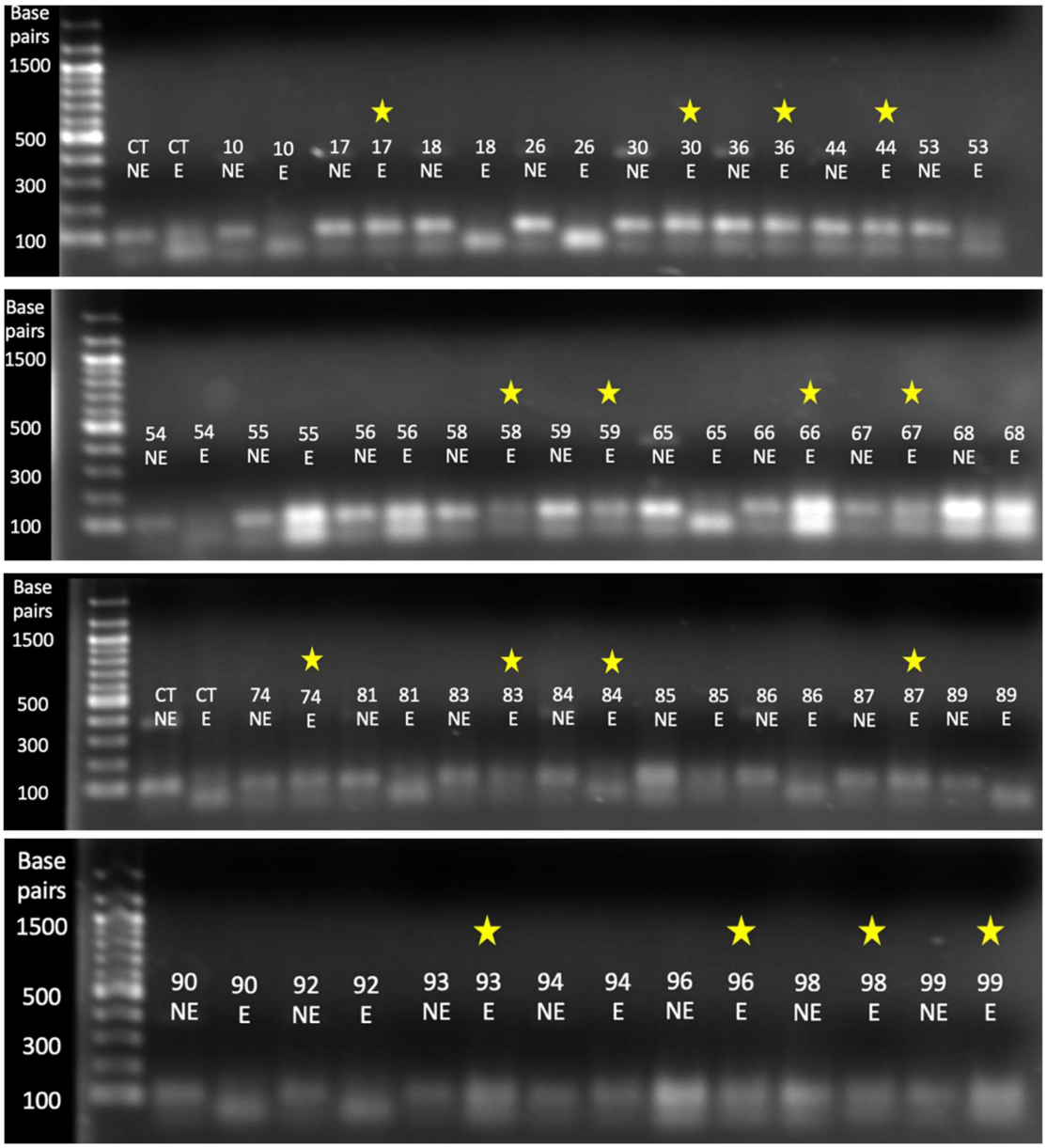


Figure 9. Gel images of Hha1 restriction digests. Stars indicate potential mutants. CT indicates controls amplified from DNA isolated from RK1004, the wild type (*rpoC*-Y75*)* strain. NE indicates no-enzyme control reactions, while E indicates reactions to which HhaI enzyme was added.

***DNA Sequencing of probable rpoC-C58A mutants*.**

Some of the 21 *rpoC*-C58A candidates had somewhat unclear results in the restriction digest screen; most samples had bands in both positions (likely due to the formation of primer dimers), and others, like 93 (Figure 9), were somewhat smudged and difficult to interpret even after repeating the screen. Four of the 21 candidates deemed likely to contain the *rpoC*-C58A mutation were selected for sequencing verification. PCR amplification of the *rpoC*-C58A region was performed using primers RK229 and RK818 (Figure 10). The length of the amplified DNA, approximately 500 base pairs, was confirmed via gel electrophoresis. The PCR products were purified and sequenced using primers RK229 and RK818. These primers amplified the C58A and Y75 regions of *rpoC* but not the entirety of the recombineering substrate. Thus, while sequencing with these primers can be used to verify the presence of the *rpoC*-C58A mutation, it does not exclude the possibility that other, potentially compensatory, mutations were made within the recombineered region. Additional sequencing was later performed to verify that no other mutations were introduced during PCR amplification. The sequenced mutants were named RK1443A, RK1443B, RK1443C, and RK1443D. Sequencing confirmed that RK1443A, RK1443C, and RK1443D have the *rpoC*-C58A substitution. Strain RK1443A, had a secondary silent mutation. This mutation substituted a thymine for a cytosine, changing the codon for phenylalanine at position 62 from TTT to TTC. Strain RK1443C contained additional, unintended mutations within the sequenced region, and strain RK1443D contained only the desired mutations. Thus, strains RK1443A and RK1443D were selected for further study.

Additional sequencing was performed to ensure that the region in which the recombineering substrate was inserted contained only the intended genetic changes. PCR amplification was performed using primers RK123 and RK219, which amplified the entire region that was used to generate the recombineering substrate (Figure 11). Following confirmation of an amplicon of approximately 714 base pairs, DNA sequencing was performed with primers RK123 and RK219. With the exception of a silent mutation in RK1443A, no other genetic changes were present within the recombineering substrate.

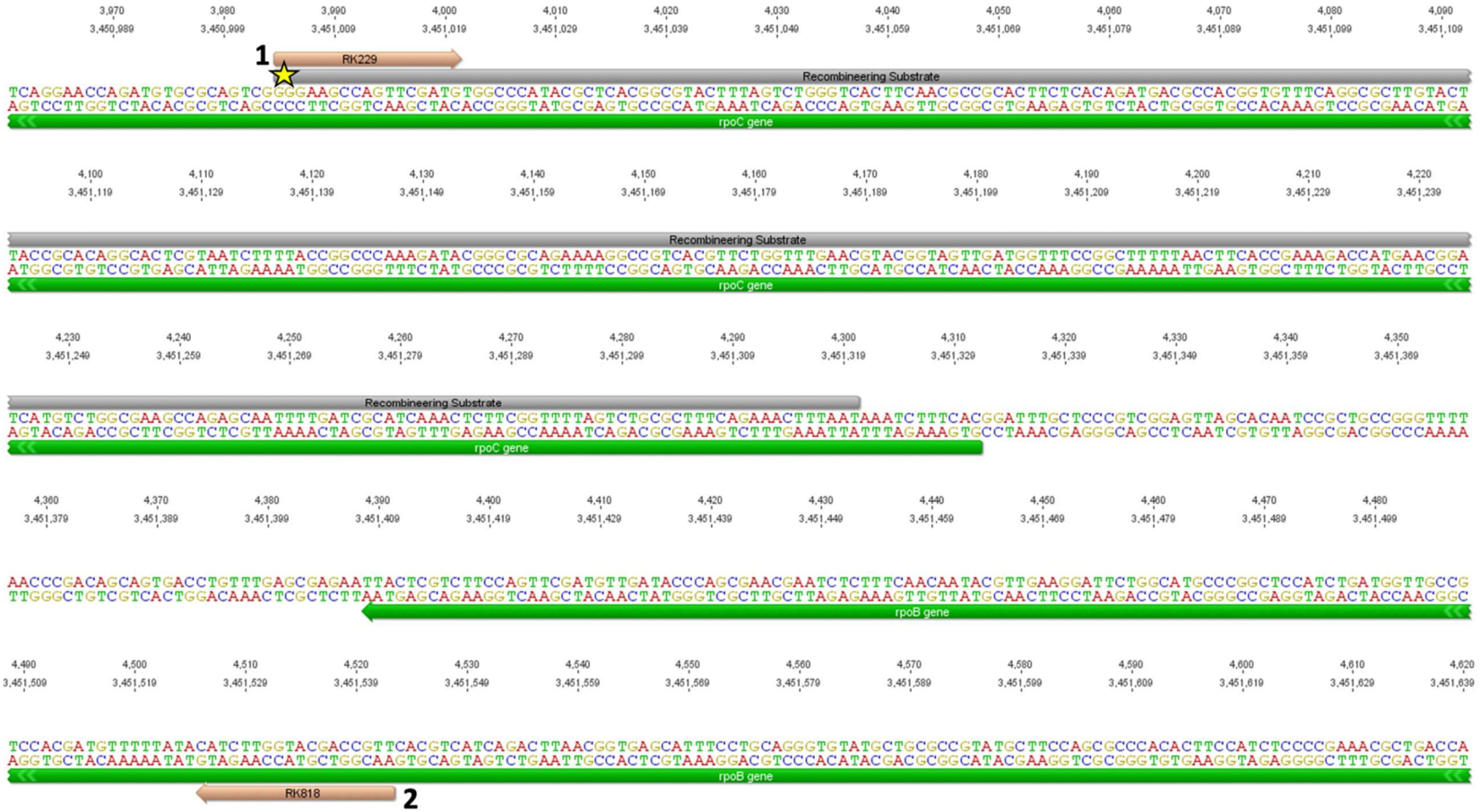


Figure 10. Location of primers RK229 (1) and RK818 (2) in relation to *rpoC* and the recombineering substrate (indicated by the star). RK818 was selected because it primes in the RPOB gene, ensuring that the insertion was indeed present within the chromosomal DNA.

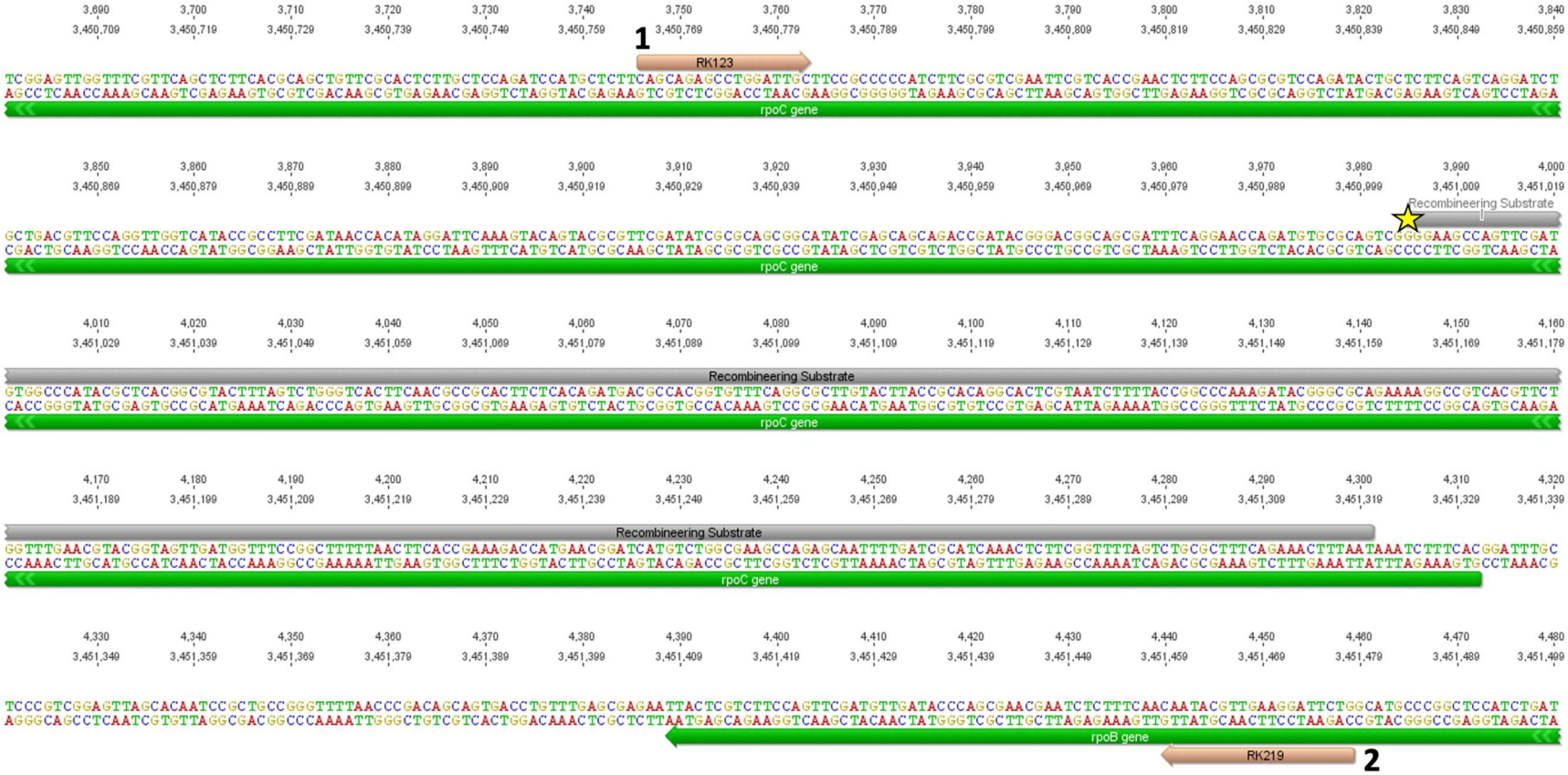


Figure 11: Location of primers RK123 (1) and RK219 (2) in relation to the recombineering substrate (indicated by the star).

Plasmid shuffling to remove the recombineering functions.

The recombineering functions in strain RK1008 are contained on a low copy plasmid that confers ampicillin-resistance. The goal of this study was to determine the effect of the *rpoC*-C58A mutation on growth at high temperatures; a temperature-sensitivity experiment would not be possible to perform if the plasmid containing recombineering functions remained in the cell because growth at high temperatures triggers overexpression of recombineering functions and, consequently, cell death. Thus, plasmid shuffling, or the replacement of one plasmid with another, was performed to remove the recombineering plasmid.

Plasmid incompatibility is a phenomenon that prevents two plasmids from being stably maintained in the same cell line. Thus, a plasmid can be removed by transforming cells with an incompatible plasmid that confers a different antibiotic resistance. This method is referred to as plasmid shuffling. To remove the recombineering plasmid pRW4714, an incompatible plasmid containing the chloramphenicol resistance gene was electroporated into strains RK1443A and RK1443D. The chloramphenicol resistance plasmid was isolated from strain RK1120 and electroporated into competent RK1443A and RK1443D cells and grown on plates containing chloramphenicol. To ensure the recombineering plasmid had been successfully kicked out, the cells were tested for growth on ampicillin (100μg/mL). Failure to grow on ampicillin plates confirmed loss of the recombineering plasmid. The newly transformed strains were named RK1445 (RK1443A) and RK1446 (RK1443D).

Growth of RK1445 and RK1446 at 42°C to test temperature sensitivity.

Previous data from our lab suggested that the *rpoC*-C58A mutation might confer a temperature sensitivity to *E. coli* cells. To test this hypothesis, strains RK1445 and RK1446 were grown at 30°C and 42°C along with a temperature-sensitive control (RK486). In two replicate experiments, there was no observable difference in size, shape, or color between *rpoC*-C58A recombinants grown at 30°C and those grown at 42°C (Figure 12). A temperature sensitive strain (RK486) was used as a control and showed an observable decrease in growth at 42°C.

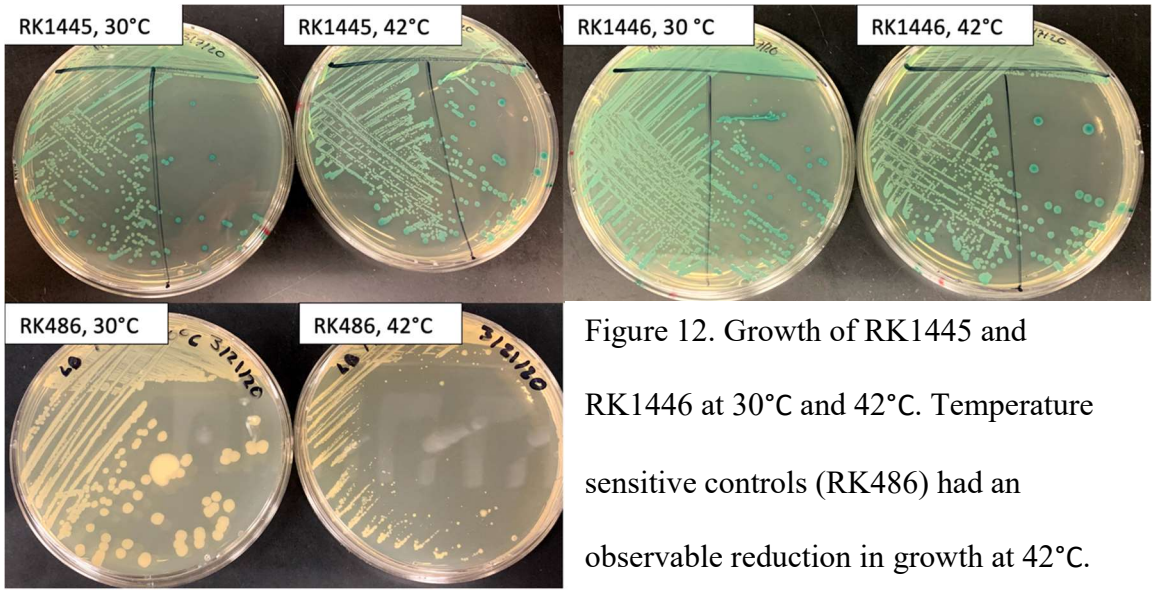


Figure 12. Growth of RK1445 and RK1446 at 30°C and 42°C. Temperature sensitive controls (RK486) had an observable reduction in growth at 42°C.

Test of rpoC-C58A complementation at 30°C and 42°C.

As initial attempts to generate a *rpoC*-C58A mutants were unsuccessful, we performed additional experiments with *rpoC*-C58A on a plasmid in case investigations with the mutation in single copy would not be possible. These experiments yielded results worth mentioning. Two plasmids, pBAD*rpoC*-C58A and pBAD*rpoC*, were

electroporated into strains RK1004 and RK1005 (see Table 2, Figure 13). The names and descriptions of these new strains are given in Table 8. Ampicillin resistant transformants were selected and single colonies were isolated and grown according to *Growth of bacterial cultures.*

Strain Name	*Genotype*
RK1447	*E. coli* strain RK1004 transformed with a pBAD*rpoC* (WT *rpoC*) plasmid that confers ampicillin resistance.
RK1448	*E. coli* strain RK1004 transformed with pBAD*rpoC*-C58A plasmid that confers ampicillin resistance.
RK1449	*E. coli* strain RK1005 transformed with a pBAD*rpoC* (WT *rpoC*) plasmid that confers ampicillin resistance.
RK1450	*E. coli* strain RK1005 transformed with pBAD*rpoC*-C58A plasmid that confers ampicillin resistance.

Table 8. Generated RK1004 and RK1005 transformants.

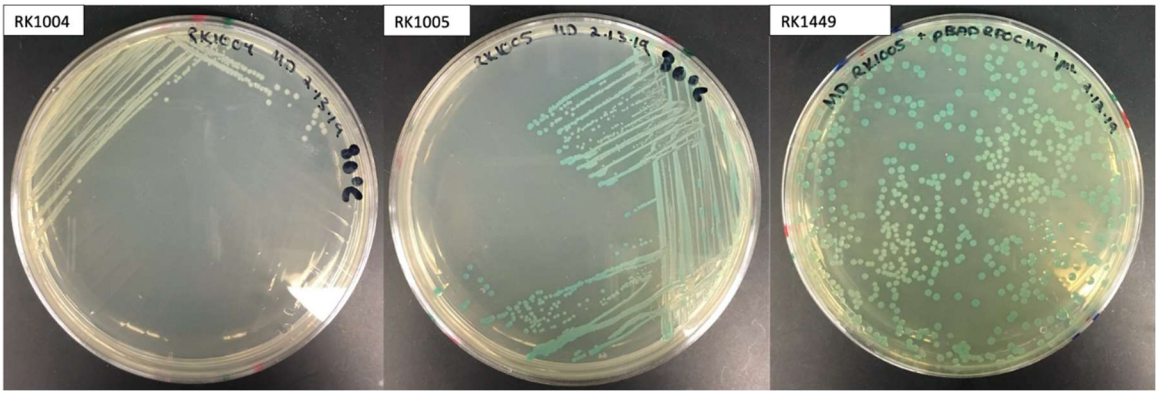


Figure 13. Streak plates of RK1004, RK1005, and RK1449.

If *rpoC*-C58A hinders RNAP from functioning at 42°C, the plasmid copy of RNAP, which enables anti-termination, would not be functional at 42°C. To determine whether complementation in RK1448 was affected by growth at high temperatures, strains RK1447 and RK1448 were plated and grown at 30°C and 42°C (see Figure 14). Anti-termination, which could only occur if complementation occurred, was not affected. This result suggests complementation was not hindered by growth at 42°C.

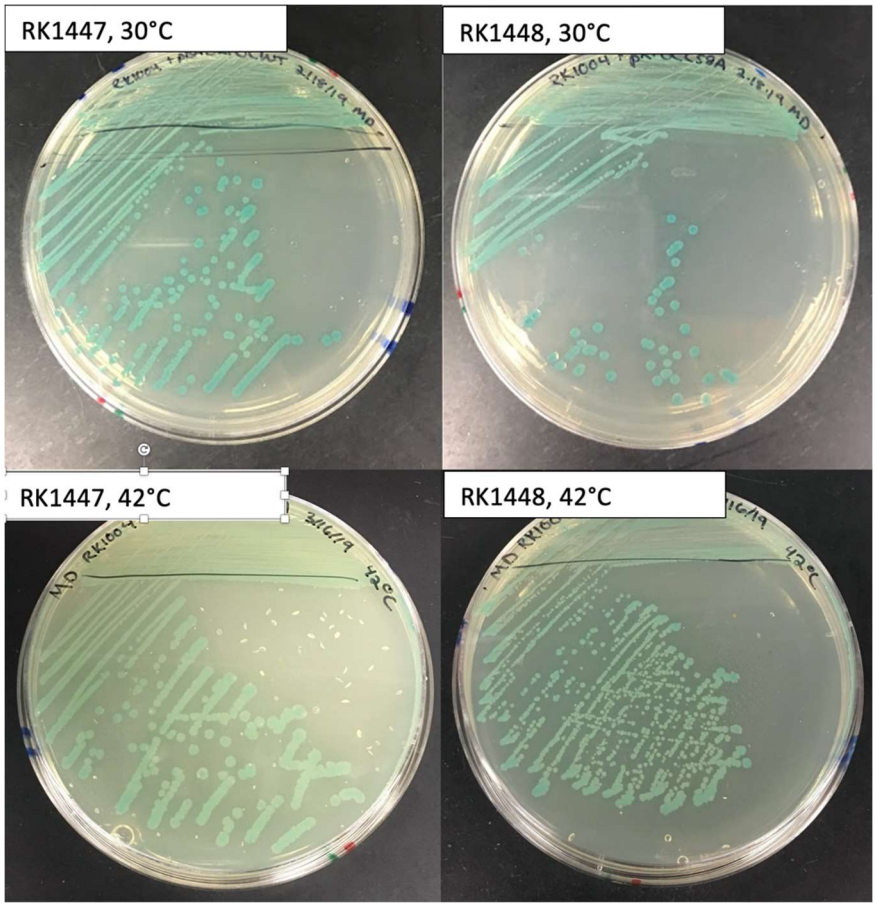


Figure 14. Growth of RK1447 and RK1448 at 30°C and 42°C. No distinct phenotypic differences were observable.

Quantitative β-galactosidase Assays.

Quantitative analysis of anti-termination was performed using β-galactosidase assays. Reactions were run in duplicate on RK1004 (*rpoC*-Y75N), RK1005 (WT *rpoC*), RK1445 (*rpoC*-C58A), and RK1446 (*rpoC*-C58A). In this assay, cells were supplied with ONPG, which acts as an artificial substrate for β-galactosidase. If β-galactosidase cleaves ONPG, o-nitrophenol is produced and gives the solution a bright yellow color. The absorbance of the solution can then be measured by spectrometry and, along with other measured values, used to calculate Miller activity units (see Methods), which indicate a standardized amount of β-galactosidase activity. These values can then be used to compare differences in expression of the reporter gene, *lacZ*, and therefore differences in anti-termination. Tables 9 and 10 contain the Miller activity units for the tested strains at 30°C and 42°C respectively. Results are averages of two internal trials and are given in Table 9 and Table 10.

30°C	Avg Miller Units	Standard Error
RK1445 (*rpoC*- C58A)	178	0.3
RK1446 (*rpoC*- C58A)	97.1	0.8
RK1004 (*rpoC*-Y75N)	0	0
RK1005 (WT *rpoC*)	64.4	2.3

Table 9. Average Miller units and standard error of the mean for β-galactosidase assays of *rpoC*-C58A mutants and controls at 30°C. RK1004 represents minimum *lacZ* expression, and RK1005 represents constitutive *lacZ* expression. Note that no yellow color developed in RK1004 samples after 1 week; this result has been equated to zero β-galactosidase activity.

42°C	Avg Miller Units	Standard Error
RK1445 (*rpoC*- C58A)	25.6	3.9
RK1446 (*rpoC*- C58A)	28.4	0.9
RK1004 (*rpoC*-Y75N)	0	0
RK1005 (WT *rpoC*)	17.3	0.8

Table 10. Average Miller units and standard error of the mean for β-galactosidase assays of *rpoC*-C58A mutants and controls at 42°C. RK1004 represents minimum *lacZ* expression, and RK1005 represents constitutive *lacZ* expression. Note that no yellow color developed in RK1004 samples after 1 week; this result has been equated to zero β-galactosidase activity.

Quantitative analysis of anti-termination in RK1147, RK1448, RK1449, and RK1450 was also performed using β-galactosidase assays. The assays were performed according to *Quantitative β-galactosidase assays.* Results are given in Table 10 and Table 11.

30°C	Avg Miller Units	Standard Error
RK1447 (rpoC-Y75N + WT)	31.7	0.8
RK1448 (rpoC-Y75N + *rpoC*- C58A)	2.7	0.3
RK1449 (WT *rpoC* + WT *rpoC*)	51.4	8.4
RK1450 (WT *rpoC* + *rpoC*-C58A)	84.8	3.8
RK1004 (*rpoC*-Y75N)	0	0
RK1005 (WT *rpoC*)	64.4	2.3

Table 11. Average Miller units and standard error of the mean for β-galactosidase assays of RK1147, RK1448, RK1449, and RK1450 and controls at 30°C. RK1004 represents minimum *lacZ* expression, and RK1005 represents constitutive *lacZ* expression. Note that no yellow color developed in RK1004 samples after 1 week; this result has been equated to zero β-galactosidase activity.

42°C	Avg Miller Units	Standard Error
RK1447 (rpoC-Y75N + WT)	5.1	0.4
RK1448 (rpoC-Y75N + *rpoC*- C58A)	1.5	0
RK1449 (WT *rpoC* + WT *rpoC*)	17.1	0.2
RK1450 (WT *rpoC* + *rpoC*-C58A)	23.9	0.1
RK1004 (*rpoC*-Y75N)	0	0
RK1005 (WT *rpoC*)	17.3	0.8

Table 12. Average Miller units and standard error of the mean for β-galactosidase assays of RK1147, RK1448, RK1449, and RK1450 mutants and controls at 42°C. RK1004 represents minimum *lacZ* expression, and RK1005 represents constitutive *lacZ* expression. Note that no yellow color developed in RK1004 samples after 1 week; this result has been equated to zero β-galactosidase activity.

CHAPTER FOUR: DISCUSSION

The high conservation of *rpoC*-C58 indicates that the cysteine has an important function and raises two primary questions. First, can a functional *rpoC*-C58 amino acid substitution be generated, or is it deleterious? If such a mutant can be generated, how does mutation of the conserved cysteine residue affect RNAP function?

The first of these questions has been answered by the work shown in this thesis. Despite the conservation of *rpoC*-C58, an alanine substitution of the cysteine residue was successfully generated, and the mutant is viable. Note, however, that the frequency for the generation of these mutants was relatively low. The reported frequency of recombinants when recombineering functions are supplied from a defective λ prophage is approximately 1%. However, initial recombineering with the *rpoC*-C58A DNA fragment (the recombineering substrate) led to the generation of *rpoC*-N75Y revertants with a frequency of .000394%. A frequency so low compared to the expected frequency suggests that these are likely to be revertants and not recombinants. Given that *E. coli* cells have a generation time of a mere 30 minutes, there is ample opportunity for spontaneous mutations to occur. The reversion only requires a single nucleotide mutation (TAC to AAC), increasing the likelihood we could identify these rare mutations with our screen. The DNA sequence results that showed the suspected recombinants generated in early experiments contained the *rpoC*-N75Y reversion but not the *rpoC*-C58A mutation. This supports the possibility these were true revertants. An experiment could be designed

to determine the frequency at which spontaneous reversion occurs; performing the recombineering procedure without addition of the recombineering substrate would be one such experiment. Though the experiments described in this thesis did not explicitly measure the frequency of *rpoC*-N75Y reversions, the arguments given above suggest reversion did occur. Sequencing of the potential recombinants generated in these experiments did not contain the *rpoC*-C58A mutations, thus our initial attempts to recombineer the desired mutation were unsuccessful.

To generate the *rpoC*-C58A mutation, alteration of the recombineering protocol was necessary. A 2X increase in the concentration of recombineering substrate led to the successful generation of multiple *rpoC*-C58A mutants which were eventually verified by DNA sequencing. Still, the observed frequency of recombination, 0.0145%, is much lower than the expected frequency of 1%. It is possible that the low frequency of recombinants was indicative of a deleterious effect of the *rpoC*-C58A mutation. It is possible, for example, that a cell can only survive the *rpoC*-C58A mutation if a simultaneous compensatory mutation is made. The requirement of generating two mutations—one intended and one spontaneous—could result in a low frequency of recombinants being generated.

To ensure that no potentially compensatory mutations were present within the immediate region of the *rpoC*-C58A mutation, a region containing the entire recombineering substrate was sequenced for all suspected *rpoC*-C58A mutants. Mutants containing additional alterations within the recombineering substrate region were eliminated from further study. While this measure decreases the likelihood that a compensatory mutation was present in these mutants, it did not exclude the possibility

that a compensatory mutation had occurred in other regions of *rpoC* or elsewhere in the *E. coli* genome. To reach such certainty, it would be necessary to sequence the entire genome of *rpoC*-C58A mutants. Such an experiment is beyond the scope of this study but might be considered for future work. However, evidence of complementation by plasmids containing *rpoC*-C58A in strain RK1448 does suggest that a compensatory mutation is not needed to produce a functional RNAP.

Now that it is known that an amino acid substitution of *rpoC*-C58 can be generated, the function of the amino acid can be investigated. Preliminary results from our lab showed that plasmids containing *rpoC*-C58A could not complement a temperature sensitive copy of *rpoC* (RK928). This result led to the question of whether the *rpoC*-C58A mutation might confer a temperature sensitivity to RNAP. The *rpoC*-C58A chromosomal mutants were therefore tested for a temperature sensitivity. Despite the observations of the earlier data, mutants RK1445 and RK1446 did not appear to demonstrate any change in colony size, shape, or color. Note, however, that there are a few potential problems with the experimental design. First, while the incubators were set to 42°C, the actual temperature continuously cycled between 40.5°C and 43.5°C. This variation may have affected the outcome of this experiment. Second, this temperature growth experiment was designed to be preliminary and to be followed by a quantitative growth experiment but could not be completed due to COVID-19. Since a standard number of cells was not added to each plate, we cannot draw conclusions about the effect of temperature on colony number. In the future, such an experiment might reveal an effect on mutant growth that may have been overlooked here.

As other conserved cysteines have a demonstrated role in antitermination (King et al., 2004), β-gal assays were performed to determine whether the *rpoC*-C58A mutation affects anti-termination. Surprisingly, RK1445 and RK1446, which contain the *rpoC*-C58A mutation, produced substantially higher Miller units than the RK1005 control (which represents constitutive expression). At 30°C, mutants RK1445 and RK1446 had Miller unit values of 178.0 and 97.1, respectively, compared to 64.4 units in RK1005. Note that no external replicates were performed, and it is possible this result is not replicable. However, if the result is reproducible, such an increase in antitermination is notable. Future work should determine whether this difference is biologically significant, and if so, it might be interesting to further evaluate the relationship between *rpoC*-C58A and antitermination.

Given that *rpoC*-C58A mutants were originally hypothesized to be temperature-sensitive, β-galactosidase assays of RK1445 and RK1446 grown at 42°C were also performed. For all samples, the Miller units were quite low. This finding may be the result of experimental error; during early growth, the incubator reached a temperature of 44.5°C; such a temperature is barely tolerable for *E. coli* cells and could have affected the results.

Regardless of the overall low levels of β-galactosidase activity, a pattern of relative activity at 42°C exists. Both mutants, RK1445 and RK1446, show higher levels of β-galactosidase activity (25.6 and 28.4 units respectively) than RK1005 (17.3 units). This result is consistent with the earlier observations at 30°C. Again, these assays were replicated internally, and the experiment should be repeated to ensure that the result is true. Still, the observation of higher β-galactosidase activity in mutant samples compared

to controls in two separate experiments suggests that *rpoC*-C58A mutants may antiterminate more efficiently than wild-type *rpoC E. coli* strains.

To determine whether *rpoC*-C58A mutants truly do antiterminate more efficiently than *rpoC*-C58 strains, a few potential future experiments might be performed. First, antitermination in other confirmed *rpoC*-C58A mutants should be tested and compared to these mutants. The finding that other mutants display similar phenotype would warrant further investigation into how *rpoC*-C58A affects antitermination. First, it would be important to ensure that the mutation itself is responsible for the increase in antitermination. As previously discussed, it is possible that a compensatory mutation might occur for *rpoC*-C58A mutants to survive. If the same compensatory mutation occurred in each of the *rpoC*-C58A mutants, it is possible that the compensatory mutation is responsible for the increase in antitermination rather than *rpoC*-C58A itself. To determine whether the *rpoC*-C58A mutation is responsible for the phenotype, the mutants could be recombineered to change the mutant *rpoC-A58* back to wild type *rpoC* (C58). If the effect on antitermination is lost, one could conclude that *rpoC*-C58A increases antitermination.

β-galactosidase assays were also performed on strains RK1447 (*rpoC*-Y75N transformed with pBAD*rpoC)*, RK1448 (*rpoC*-Y75N transformed with pBAD*rpoC*-C58A), RK1449 (*rpoC*-Y75 transformed with pBAD*rpoC*), and RK1450 (*rpoC*-Y75 transformed with pBAD*rpoC*-C58A). A few interesting trends are apparent in these data. At 30°C, complementation by the wild type copy of *rpoC* occurs in RK1447. However, the plasmid copy of *rpoC*-C58A complements very poorly in RK1448. Interestingly, RK1450, which contains a copy of *rpoC*-C58A on a plasmid, had a higher value of Miller

units than RK1449 (which contained wild type *rpoC*). This result is consistent with the observation that *rpoC*-C58A mutants appeared to have increased antitermination, substantiating the suggestion that there may be a relationship between *rpoC*-C58A and antitermination. These assays were also repeated at 42°C. The high-temperature assays were subject to the same experimental limitations described previously, and thus the observed Miller units are much lower in all assays performed at 42°C. Still, the same pattern of complementation seen at 30°C was observed.

This work has demonstrated that the *rpoC*-C58A mutation can be generated in the chromosome of *E. coli*. However, the reason for the high conservation of *rpoC*-C58 remains unclear. Our results do not indicate a temperature-sensitive phenotype in *rpoC*-C58A mutant. but additional experiments are necessary to support this conclusion. A screen for an effect of the mutation on other phenotypes, such as RNAP response to different growth conditions, might also reveal the reason for *rpoC*-C58A conservation. Experiments such as these could determine the basis for *rpoC*-C58 conservation and will lead to a better understanding of this critical RNA polymerase domain.

REFERENCES

Adhya, S., & Gottesman, M. (1978). Control of Transcription Termination. *Annual Review of Biochemistry, 47*(1), 967-996. doi:10.1146/annurev.bi.47.070178.004535

Bergsland, K. J., & Haselkorn, R. (1991). Evolutionary relationships among eubacteria, cyanobacteria, and chloroplasts: Evidence from the rpoC1 gene of Anabaena sp. strain PCC 7120. *Journal of Bacteriology, 173*(11), 3446-3455. doi:10.1128/jb.173.11.3446-3455.1991

Chang, C. Y., Nam, K., & Young, R. (1995). S gene expression and the timing of lysis by bacteriophage lambda. *Journal of Bacteriology, 177*(11), 3283-3294. doi:10.1128/jb.177.11.3283-3294.1995

Christie, G. E., Cale, S. B., Isaksson, L. A., Jin, D. J., Xu, M., Sauer, B., & Calendar, R. (1996). Escherichia coli rpoC397 encodes a temperature-sensitive C-terminal frameshift in the beta' subunit of RNA polymerase that blocks growth of bacteriophage P2. *Journal of Bacteriology, 178*(23), 6991-6993. doi:10.1128/jb.178.23.6991-6993.1996

Clerget M. (1991). Site-specific recombination promoted by a short DNA segment of plasmid R1 and by a homologous segment in the terminus region of the Escherichia coli chromosome. The New biologist, 3(8), 780–788.

Clerget, M., Jin, D. J., & Weisberg, R. A. (1995). A Zinc-binding Region in the β′ Subunit of RNA Polymerase is Involved in Antitermination of Early Transcription of Phage HK022. *Journal of Molecular Biology, 248*(4), 768-780. doi:10.1006/jmbi.1995.0259

Cooper, G. M. (2000). Eukaryotic RNA Polymerases and General Transcription Factors. In The Cell: A Molecular Approach. 2nd edition. Sunderland, MA: Sinauer Associates. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9935/

Costantino, N., & Court, D. L. (2003). Enhanced levels of Red-mediated recombinants in mismatch repair mutants. *Proceedings of the National Academy of Sciences, 100*(26), 15748-15753. doi:10.1073/pnas.2434959100

Datta, S., Costantino, N., & Court, D. L. (2006). A set of recombineering plasmids for gram-negative bacteria. *Gene, 379*, 109-115. doi:10.1016/j.gene.2006.04.018

Ellis, H. M., Yu, D., Ditizio, T., & Court, D. L. (2001). High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proceedings of the National Academy of Sciences, 98*(12), 6742-6746. doi:10.1073/pnas.121164898

Geszvain, K., & Landick, R. (2005). The Structure of Bacterial RNA Polymerase.

Guyer, M. S., Reed, R. R., Steitz, J. A., & Low, K. B. (1981). Identification of a Sex-factor-affinity Site in E. coli as. Cold Spring Harbor Symposia on Quantitative Biology, 45(0), 135–140. https://doi.org/10.1101/sqb.1981.045.01.022

King, R. A., Banik-Maiti, S., Jin, D. J., & Weisberg, R. A. (1996). Transcripts That Increase the Processivity and Elongation Rate of RNA Polymerase. *Cell, 87*(5), 893-903. doi:10.1016/s0092-8674(00)81996-0

King, R. A., Markov, D., Sen, R., Severinov, K., & Weisberg, R. A. (2004). A Conserved Zinc Binding Domain in the Largest Subunit of DNA-dependent RNA Polymerase Modulates Intrinsic Transcription Termination and Antitermination but does not Stabilize the Elongation Complex. *Journal of Molecular Biology, 342*(4), 1143-1154. doi:10.1016/j.jmb.2004.07.072

Miller, J. H. (1992). Unit 3: The lac System. In *A short course in bacterial genetics. a laboratory manual and handbook for Escherichia coli and related bacteria* (pp. 72-74). Plainview, NY: Cold Spring Harbor Laboratory Press.

Murakami K. S. (2013). X-ray crystal structure of Escherichia coli RNA polymerase σ70 holoenzyme. The Journal of biological chemistry, 288(13), 9126–9134. https://doi.org/10.1074/jbc.M112.430900

Nudler, E., Avetissova, E., Markovtsov, V., & Goldfarb, A. (1996). Transcription Processivity: Protein-DNA Interactions Holding Together the Elongation Complex. *Science, 273*(5272), 211-217. doi:10.1126/science.273.5272.211

Santangelo, T. J., & Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase runs a stop sign. *Nature Reviews Microbiology, 9*(5), 319-329. doi:10.1038/nrmicro2560

Sehnal, D., Rose, A. S., Koča, J, Burley, S. K., & Velankar , S. (2018). Mol*: towards a common library and tools for web molecular graphics. *Proceedings of the Workshop on Molecular Graphics and Visual Analysis of Molecular Data (MolVA '18).* Eurographics Association, Goslar, DEU, 29–33.

Sloan, S., Rutkai, E., King, R. A., Velikodvorskaya, T., & Weisberg, R. A. (2007). Protection of antiterminator RNA by the transcript elongation complex. *Molecular Microbiology, 63*(4), 1197-1208. doi:10.1111/j.1365-2958.2006.05579.x

Sutherland, C., & Murakami, K. S. (2018). An Introduction to the Structure and Function of the Catalytic Core Enzyme of Escherichia coli RNA Polymerase. *EcoSal Plus, 8*(1). doi:10.1128/ecosalplus.esp-0004-2018

Thomason, L., Court, D. L., Bubunenko, M., Costantino, N., Wilson, H., Datta, S., & Oppenheim, A. (2005). Escherichia coli, Plasmids, and Bacteriophages. In *Current protocols in molecular biology* (pp. 1.16.1-1.16.21). S.l.: John Wiley & Sons.

Wang, I. (2005). Lysis Timing and Bacteriophage Fitness. *Genetics, 172*(1), 17-26. doi:10.1534/genetics.105.045922

APPENDIX A

Luria Broth (LB), Lennox Recipe

The following were combined:
-5 g NaCl (Fisher BioReagents; BP358-212)
-5 g Yeast Extract (BD; 212750)
-10 g Tryptone (Fisher BioReagents; BP9726-500)
-1 L DI water
Contents were heated until solution was dissolved. Solution was autoclaved and stored at room temperature.

To make LB plates, 15g of agar (Difco; 281810) were added and the solution was heated to dissolve prior to autoclaving.

Luria Broth (LB) with Ampicillin, Lennox Recipe

LB was made as described. Then, 100µg/mL Ampicillin (Fisher BioReagents; BP-1760-25) was added. Mixture was stored at 4°C for up to 14 days.

To make LB plates with ampicillin, 15g of agar (Difco; 281810) were added and the solution was heated to dissolve prior to autoclaving.

Polymerase Chain Reaction Mix

The following were combined:
-6 µL 10 mM dATP (Fisher Scientific Kit; FB-6000-10)
-6 µL 10 mM dCTP (Fisher Scientific Kit; FB-6000-10)
-6 µL 10 mM dGTP (Fisher Scientific Kit; FB-6000-10)
-6 µL 10 mM dTTP (Fisher Scientific Kit; FB-6000-10)
-330 µL 10X Buffer A (Fisher Scientific Kit; FB-6000-10)
Mix was stored at - 20°C and kept on ice during use.

Super Optimal Broth (SOC Broth)

The following were combined:
-20 g Tryptone (Fisher BioReagents; BP9726-500)
-5 g Yeast Extract (BD; 212750)
-0.5 g NaCl (Fisher BioReagents; BP358-212)
-1.25 mL 2 M KCl (Acros; CAS: 744-40-7)
-10 mL 1 M MgCl2 (Fisher Chemical; M13448)
-1 L DI water

The solution was heated to dissolve, autoclaved, and cooled to room temperature. Then, the following were added:
-1 mL 1 M $MgSO_4$ (Sigma; M2773)
-2 mL 10M glucose (Fisher Scientific; D16500)
The mixture was stored at room temperature.

X-gal plates

LB was made as described. Following autoclaving, 2mL of 20mg/mL X-gal were added prior to pouring plates.

Z-buffer

The following were combined:
-16.1 g $Na_2HPO_4 \bullet 7H_2O$ (Sigma-Aldrich; 7782-85-6)
-5.5 g $NaH_2PO_4 \bullet H_2O$ (Sigma-Aldrich; 10049-21-5)
-0.75 g KCl (Acros; CAS: 744-40-7)
-0.246 g $MgSO_4 \bullet H_2O$ (Sigma-Aldrich; 10034-99-8)
-57 2.7 mL β-mercaptoethanol (SigmaAldrich; M-3148)
-1 L water
After the salts were dissolved, the pH was adjusted to 7.0, and the solution was stored at 4°C.

6X Loading Dye

The following were combined:
-0.25% Bromophenol Blue (Sigma-Aldrich; CAS: 115-39-9)
-0.25% Xylene Cyanol FF (Research Organics; 7113X)
-30% Glycerol (Fisher Chemical; CAS: 56-81-5)
Dye was stored at room temperature.

50X Tris-Acetic Acid Ethylenediaminetetraacetic acid (TAE)

The following were combined:
-242 g Tris-base (Fisher BioReagents; M-11645)
-57.1 mL Glacial Acetic Acid (Amresc; 0714-4L)
-100 mL 0.5 M EDTA (Sigma; E-5134)
The volume was brought to 1L with DI water, and the solution was stored at room temperature.