ROLE OF THE N-TERMINUS OF RPB9 IN RNA POLYMERASE II TRANSCRIPTION

An Undergraduate Research Scholars Thesis

by

KATHERINE NGUYEN

Submitted to the Undergraduate Research Scholars program
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Research Advisor: Dr. David Peterson

April 2016

Major: Biochemistry and Genetics

TABLE OF CONTENTS

ABSTRACT

Role of the N-terminus of Rpb9 in RNA Polymerase II Transcription

Katherine Nguyen
Department of Biochemistry and Biophysics
Texas A&M University

Research Advisor: Dr. David Peterson
Department of Biochemistry and Biophysics

Eukaryotic transcription is a highly regulated process that requires RNA polymerase to continuously proofread and edit the RNA stand before it continues to elongate. A subunit in RNA polymerase II (RNAPII), known as Rpb9, plays an important role in the efficiency of RNA proofreading and maintains the health of a cell as a whole. Recent studies have proven that Rpb9 serves at least two functions, one of which resides in the C-terminal half of the protein and the other in the N-terminal half. The experiments proposed here suggest a genetic approach to start to determine the function of the N-terminal domain.

ACKNOWLEDGMENTS

The process of writing my thesis was a journey through which I could have not finished without the encouragement and guidance of many individuals. First of all, I would like to thank God for the opportunity to be involved in undergraduate research. To Him be the glory.

I also wish to express my sincerest thanks to Dr. David Peterson, my research advisor, for being my biggest supporter through this entire process and for reviewing my paper on numerous occasions. I owe all of my research accomplishments to him. He took time to personally train me the first time I walked into the research lab and has preened me since then to stay eager, confident, and curious.

I would also like to thank Amber Wadle, who helped me find ways around problems I faced in various experiments. She enthusiastically explained new research techniques to me and was a major contributor to my project.

I also would like to thank the labs of Dr. Craig Kaplan and Dr. Mary Bryk for their support.

Finally, I would like to thank Texas A&M University for fostering such a great research institution, where undergraduates can excel in research. It is an opportunity that I do not take for granted.

CHAPTER I

INTRODUCTION

Transcripts produced in eukaryotic cells are constantly being proofread by RNA polymerase II during transcriptional elongation. One small subunit of RNAPII, Rpb9, a protein composed of 122 amino acids, has been proven to have a significant role in the proofreading activity of the polymerase. Recent research has shown that in the absence of Rpb9, errors incurred during transcription are not efficiently corrected. Incorrect matches to the template DNA strand are normally cleaved by an accessory factor called TFIIS, which associates with the polymerase and reprograms the active site for the removal of misincorporated nucleotides (1). Rpb9 has been proven to be crucial in slowing the elongation process after a misincorporation, providing a checkpoint during which the polymerase can backtrack and proofread the transcript (1). Therefore, the absence of Rpb9 causes errors in the transcript to be extended at a faster rate, restricting the time frame of that checkpoint (1). Not only does Rpb9 play a crucial role in TFIIS-mediated cleavage, it also affects mismatch extension (2). Random mutations caused by the decreased proofreading activity of the polymerase affects the overall health and growth of the cell (3).

Furthermore, recent studies have suggested that not only do the errors incurred during transcription impede normal cellular health but also have the potential to inflict human diseases related to abnormal protein folding, such as Alzheimer's disease (4). Proteins that are translated from erroneous mRNA strands sometimes do not have the correct three-dimensional structure

(4). In fact, some of the mutated proteins are not able to fold at all, and can cause increased levels of proteotoxic stress (4). Experimental evidence has also shown that the number of errors made during transcription is directly related to cell age (4). Therefore, it is possible that the accumulation of transcription errors made and propagated by an error-prone polymerase can be a new mechanism by which some human diseases are developed.

The structure of Rpb9 consists of two domains. The C-terminal domain of the protein has proven to be significant in reducing the number of errors present in the mRNA strand. Experiments have shown that cells containing certain mutations on the C-terminal domain of Rpb9 exhibit the same increased speed in error propagation as cells lacking the entire Rpb9 protein (5). There is convincing evidence that the N-terminal domain of Rpb9 serves a function independent from the C-terminal domain. The majority of the growth defects present in Rpb9-deficient cells can be rescued by a truncated Rpb9 containing only the first 59 amino acids of the protein. The N-terminal domain of Rpb9 has also been proposed to affect transcription start-site selection.

The experiments described here are part of a larger genetic and biochemical study to explicitly define the role of the N-terminal domain of Rpb9 on RNAPII activity and the interactions of RNAPII with other components of the eukaryotic transcription machinery. *Saccharomyces cerevisiae (S. cerevisiae),* more commonly known as Baker's yeast, was used as the model organism in these experiments, as it is a simple organism that possesses eukaryotic transcription machinery. We are interested in identifying the specific amino acids within the N-terminal domain of Rpb9 that, if changed, have visible effects on how a eukaryotic cell grows in various

environments. In order to accomplish this, an array of random mutations in the N-terminal domain of Rpb9 will be created and then screened for phenotypes associated with an *rpb9* deletion. The goal is to find a particular amino acid sequence present on the N-terminal domain of Rpb9 that, if mutated, will cause cells to mimic the growth patterns of cells missing the entire *RPB9* gene. Sequencing the mutations found in the selected alleles will identify amino acids that are critical for the function of the N-terminal domain of Rpb9 and will allow for the development of specific hypotheses related to their effect on RNAPII. The work described here is a part of efforts to develop a sensitive screen to more efficiently identify mutations in the N-terminal domain of Rpb9 that affect transcription start-site selection.

CHAPTER II

METHODS

Colony PCR

Mutants generated from random mutagenesis, Rpb9(4-6)X mutants, that showed stunted or abnormal growth were single colony purified for further testing. Colony PCR (cPCR) was performed in order to amplify the mutant DNA. During colony PCR, fresh yeast cells were placed into 10 µL of 0.02 N NaOH. A layer was paraffin oil was added to the sample to prevent evaporation in the subsequent heating process. The heat and sodium hydroxide function to lyse the yeast cells and release both RNA and DNA molecules. Water was then added to each sample of lysed cells to make a 1:5 diluted DNA sample. A master mix containing Taq polymerase, 10x PCR buffer, water, 0.2 mM dNTP's, 0.2 mM of T3 primer, and 0.2 mM of T7 primer was then added to each of the samples. Each sample contained 10 µL of the diluted sample, 31 µL of dH_2O, 5 µL of buffer, and 1 µL of each primer, dNTP's, and polymerase. The genetic sequence of each mutant was amplified using the Polymerase Chain Reaction. For colony PCR, the PCR machine ran for 1 minute at 94ºC. The PCR continued to cycle 40 times, with each cycle consisting of 30 seconds at 95ºC, then 30 seconds at 45ºC, and finally 2.5 minutes at 72ºC. At the end of the reaction, the samples were held at 10ºC. About 5 µL of the PCR products were removed from the tube. This portion of product was saved to confirm the efficiency of the PCR process. The remaining portion of cPCR product was cleaned using the Qiagen PCR Cleanup Kit.

Gel electrophoresis

Gel electrophoresis was used to determine if the mutants contained the *RPB9* construct and if purification was successful. About 5 µL of the sample not cleaned using the Qiagen PCR Cleanup Kit was mixed in with 1 µL of 6X dye. The resulting solution was loaded into a 0.8% agarose gel. One µL of 50% glycerol and 1 µL of 6X dye were added to 5 µL of the purified samples. The entire solution of pure sample, glycerol, and dye were loaded into a well. The agarose gel contained GelRed, which was used as an intercalating agent. The gel was allowed to run at 80 V in 1X TAE buffer and was analyzed under UV light. A band was present around the 1200 bp region, indicating the presence of *rpb9* on the transformed plasmid.

Amplification PCR

The purified products were amplified a second time using Amplification PCR (aPCR). To each PCR tube, 2 µL of the sample cleaned by the Qiagen PCR Cleanup Kit, 39 µL of dH_2O, 5 µL of 10X PCR buffer, 1 µL of Taq polymerase, 1 µL of 10 mM T3 primer, 1 µL of 10mM T7 primer, and 1 µL of 10 mM dNTP's were added. All tubes were placed into the thermocycler and ran under the same PCR program as cPCR. The products yielded by aPCR were cleaned a second time using the Qiagen PCR Cleanup Kit.

Concentration measurement

The concentrations of the resulting samples were measured using BioTek's Epoch Micro-Volume Spectrophotometer System. Only samples containing more than 100 ng/µL of DNA qualified for sequencing.

Sequencing PCR

Two µL of each of the doubly purified samples, 1 µL of water, 1 µL of Big Dye, and 1 µL of T3 primer were added into clean test tubes. Sequencing PCR was then used to amplify the purified sample solutions. At the beginning of this process, the PCR machine ran for 4 minutes at 96ºC. This allows for the initial denaturation of the DNA strands. The PCR continued to cycle 30 times, with each cycle consisting of 30 seconds at 96ºC, then 15 seconds at 50ºC, and finally 4 minutes at 60ºC. At the end of the reaction, the samples are held at 10ºC.

DNA sequencing

In preparation for DNA sequencing, the samples were placed into size-exclusion gel columns. Before the addition of the sample, however, the gel columns were centrifuged at 1000 g for 2 minutes. The buffer supernatant was then disposed and the column was placed into another tube. Water was added to each sample to make a total volume of 50 µL, and then all of the diluted sample was dispensed through the gel column. After centrifugation at 1000 g for 4 minutes, the tubes were dried for an hour using a RoboVap spinning apparatus. The dehydrated products and sequences were then sent to the Texas A&M Gene Technology Lab in the Biology Department for analysis. The gene technology lab ran an automated sequencing procedure that generated the sequence of DNA. By sequencing the Rpb9 mutants, we sought to identify the amino acid sequence located in the N-terminal domain that is crucial in maintaining the function of the Rpb9 protein.

Transformation using yeast construct

Yeast cells with a *rpb9*Δ*::NATMX4 imd2*Δ*::HIS3* genome were transformed with trialanine mutant plasmids. One thousand µL of the transformed cultures were centrifuged and resuspended in 800 µL of transformation solution. The transformation mix consisted of 700 µL of Magic Mix (0.2 M lithium acetate, 40% polyethylene glycol 3350), 72 µL of dH_2O, 8 µL of 1 M DTT, and 20 µL of 2 mg/mL single-strand salmon sperm DNA. All transformations were grown on -leu plates.

Replica plating

Transformed cells were replica plated onto -leu, -leu -his, -leu with 0.04 µg/mL of MPA, -leu -his with 0.04 µg/mL of MPA, -leu -his with 0.04 µg/mL of MPA and 5 mM of 3AT (3-amino-triazole), and -leu -his with 0.04 µg/mL of MPA and 10 mM of 3AT. All plates were grown at 30ºC for a total of 4 days.

CHAPTER III

RESULTS

Site-directed trialanine mutagenesis

A library of mutations that introduced amino acid substitutions into the N-terminal domain of Rpb9 has already been made using site-directed mutagenesis. These mutations were generated on a plasmid of *Escherichia coli* (*E.coli*) using primers to introduce changes in the DNA that caused three consecutive amino acids in Rpb9 to be replaced with alanine. The growth patterns of the trialanine mutants were then assessed, and colonies that exhibited similar growth to Rpb9-deficient cells were selected for further study. Yeast cells that contained the trialanine mutation at the 4-6 amino acid position yielded a promising phenotype. The mutations were made according to Figure 1.

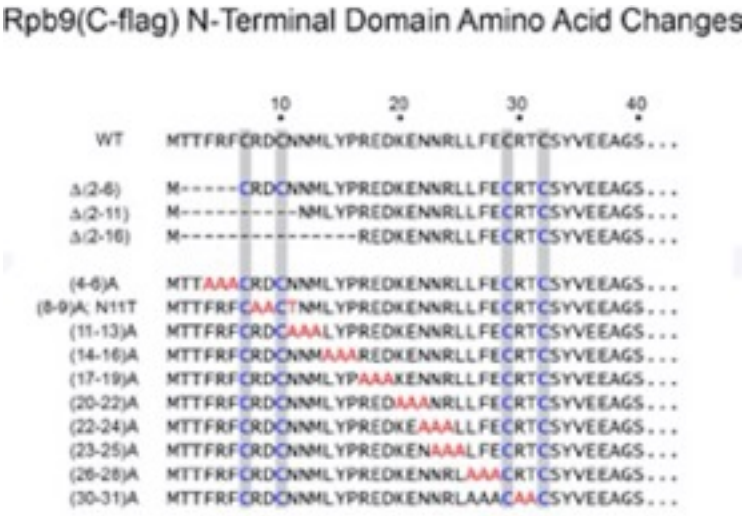

Figure 1. Rpb9(C-flag) N-terminal Domain Amino Acid Changes: Mutations generated through site-directed trialanine mutagenesis

Random mutagenesis

Targeted yet random mutations were then constructed in wild-type yeast cells at the 4-6 residue position of *RPB9*. Each construct contained three different mutant amino acid sequences at the 4-6 position and were made using PCR-mediated mutagenesis. Multiple assays were carried out in order to phenotype the newly created Rpb9 mutant yeast colonies, denoted as Rpb9(4-6)X mutants. In one experiment, the Rpb9(4-6)X cells were plated onto mycophenolic acid (MPA) and their growth patterns were accessed as the drug concentration increased. MPA inhibits the production of GTP, which results in stunted growth of hypersensitive yeast cells. Some cells, however, have the ability to induce the production of inosine monophosphate dehydrogenase 2 (*IMD2*), an enzyme used to overcome the deficiency of GTP caused by the presence of MPA. Interestingly, the induced transcription of *IMD2* is correlated with the speed at which RNA polymerase initiates and elongates transcription. Therefore, growth patterns on MPA plates provides information about the speed of the polymerases present in the generated mutants. Preliminary results suggest that some *rpb9* mutants had higher levels of sensitivity for the drug and were selected for further study. In another experiment, the mutant yeast cells were grown in an environment of 37°C. Some candidates were more sensitive to higher temperatures and did not grow as well as the wild-type yeast cells. These results indicated that certain amino acids present in the 4-6 residue region of the *RPB9* drastically affected the growth and viability of the cell.

DNA sequencing

There is a single amino acid change in the N-terminus of Rpb9, specifically in the 4-6 nucleotide region, that can significantly affect the growth of yeast cells in various stress conditions. Yeast cells that contained the T3S variant of the Rpb9 protein were more sensitive to mycophenolic acid than their wild-type counterparts. The significant amino acid change was specifically caused by two nucleotide changes. Nucleotide 8, usually containing C, was changed to a G. Secondly, nucleotide 9 was changed from a G to a C. By obtaining the specific amino acid changes that are critical for the function of the N-terminal domain of Rpb9, we can start to develop specific hypotheses related to their effect on RNAPII and eukaryotic transcription.

Replica plating

After 72 hours of growth, the -leu -his plate containing 0.04 μg/mL of MPA and 5 mM of 3AT showed the best gradation in cellular growth. The *IMD2* gene contains two transcription sites, but only the downstream transcription site produces a functional protein. The *IMD2* open reading frame was replaced with the *HIS3* open reading frame to provide a much more sensitive assay for the induction of transcription from the *IMD2* promoter. MPA is a drug that induces the selection of the downstream transcription start-site in the presence of Rpb9. Without the proper function of Rpb9, MPA fails to induce the selection of the downstream transcription start-site. This should result in the stunted growth of Rpb9 mutants in the presence of MPA. Addition of 3AT, an inhibitor of the *HIS3* gene product, can be used to adjust the sensitivity of the assay.

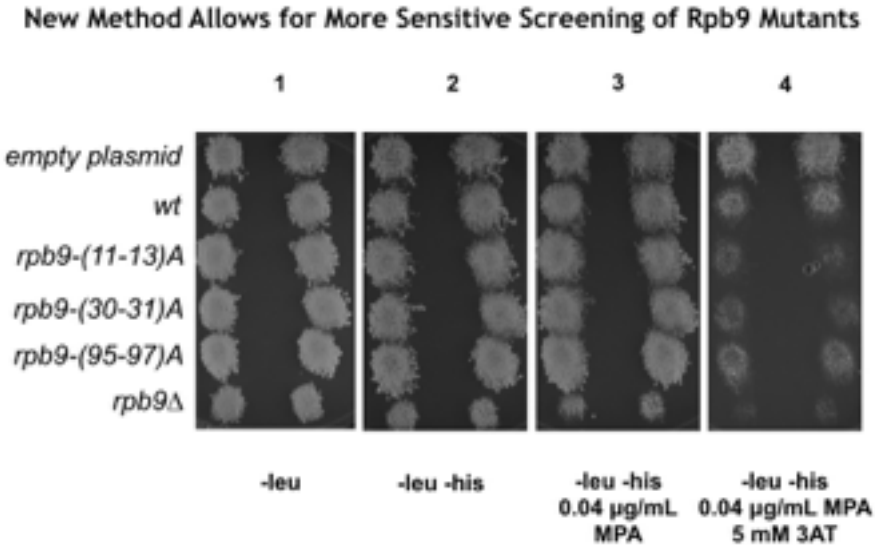

Figure 2. New Method Allows for More Sensitive Screening of Rpb9 mutants: All plates pictured are after 72 hours of growth. Plates **1-3** show normal growth patterns for all yeast strains. Plate **4** demonstrates stunted growth patterns for N-terminal *rpb9* mutants and essentially no growth for *rpb9*Δ mutants. MPA or mycophenolic acid induces *IMD2* expression, and 3-amino triazole inhibits the *HIS3* gene product.

CHAPTER IV

CONCLUSION

Using the *rpb9*Δ*::NATMX4 imd2*Δ*::HIS3* yeast strain, a more sensitive method was generated for screening Rpb9 N-terminal mutants. By replacing the *IMD2* open reading frame with the *HIS3* open reading frame, the transcription start-site defects could be phenotypically observed. The replacement of the genomic, wild-type *RPB9* gene with *NATMX4*, a gene that encodes an enzyme that inactivates the antibiotic nourseothricin, allowed for the knockout of genomic *RPB9* to be phenotypically verified. Plasmids containing mutant *rpb9* genes were used to introduce different alleles of *rpb9* into the yeast cells. These plasmids enabled the determination of whether the different alleles affect the ability of the yeast cells to induce transcription from the *IMD2* promoter, which, in this strain, results in expression of *HIS3*. The addition of MPA induces transcription start-site selection. The addition of 3AT, an inhibitor of the *HIS3* gene product, can be used to fine tune how much *HIS3* needs to be expressed to allow the cells to grow in the absence of added histidine. This inhibitory protein can be used to adjust the sensitivity of the assay.

This new method of mutant screening will allow for the identification of the most important residues on the N-terminus of Rpb9. More site-specific mutants can be generated to identify specific alleles that significantly affect Rpb9 function. This study can also be expanded to investigate associations between start-site selection and other functions of Rpb9.

REFERENCES

1. Sigurdsson, S., Dirac-Svejstrup, A. B., and Svejstrup, J. Q. (2010) Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. Mol Cell 38, 202-210

2. Knippa, K., and Peterson, D. O. (2013) Fidelity of RNA Polymerase II Transcription: Role of Rbp9 in Error Detection and Proofreading. Biochemistry 52, 7807-7817

3. Nesser, N. K., Peterson, D. O., and Hawley, D. K. (2006) RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. Proc Natl Acad Sci U S A 103, 3268-3273

4. Vermulst, M., Denney, A. S., Lang, M. J., Hung, C. W., Moore, S., Mosely, A. M., Thompson, W. J., Madden, V., Gauer, J., Wolfe, K. J., Summers, D. W., Schleit, J., Sutphin, G. L., Haroon, S., Holczbauer, A., Caine, J., Jorgenson, J., Cyr, D., Kaeberlein, M., Strathern, J. N., Duncan, M. C., and Erie, D. A. (2015) Transcription errors induce proteotoxic stress and shorten cellular lifespan. Nat Commun 6, 8065

5. Kaster, B., Texas A&M University