2023

Uncovering the Role of SPT5 in RNA Polymerase I Transcription Through Targeted Protein Degradation Via the Auxin Inducible Degron System in Saccharomyces Cerevisiae

Nathan Bellis
University Of Alabama At Birmingham

UNCOVERING THE ROLE OF SPT5 IN RNA POLYMERASE I TRANSCRIPTION THROUGH TARGETED PROTEIN DEGRADATION VIA THE AUXIN INDUCIBLE DEGRON SYSTEM IN *SACCHAROMYCES CEREVISIAE*

by

NATHAN BELLIS

DAVID SCHNEIDER, COMMITTEE CHAIR
YVONNE EDWARDS
ROMI GUPTA

A THESIS

Submitted to the graduate faculty of the University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Master of Science

BIRMINGHAM, ALABAMA

2023

UNCOVERING THE ROLE OF SPT5 IN RNA POLYMERASE I TRANSCRIPTION THROUGH TARGETED PROTEIN DEGRADATION VIA THE AUXIN INDUCIBLE DEGRON SYSTEM IN *SACCHAROMYCES CEREVISIAE*

NATHAN BELLIS

MULTIDISCIPLINARY BIOMEDICAL SCIENCE

ABSTRACT

RNA Polymerases are the molecular machines responsible for the synthesis of RNA from the DNA template. The presence of these machines is an absolute requirement for the function and replication of all cellular organisms. In eukaryotic organisms, there are a minimum of three RNA polymerases (Pols I, II, II). Despite shared subunits and homology, these three protein complexes are functionally distinct, with many unique sub-units and additional trans-acting factors.

Pol I is responsible for the synthesis of three of the four ribosomal RNA species which is the first and rate limiting step of ribosome biogenesis. Pol I synthesizes the majority of RNA in growing cells despite only transcribing one gene. To improve our understanding of these machines we must characterize the role and function of not only the intrinsic sub-units of the polymerase but also any factors that may associate with the polymerase.

The Spt4/5 complex is a putative Pol I trans-acting factor that has been well studied for its role in Pol II. Previous research from our group has shown that Spt4/5 also interacts with Pol I in budding yeast (*Saccharomyces cerevisiae*), and that mutations in the two subunits of the heterodimer, Spt5 or Spt4, cause defects in ribosome biogenesis. Unfortunately, in all of these models it is impossible to discount that the much more established role of Spt4/5 in Pol II is responsible for the phenotypes observed. To

circumvent this issue, we used the auxin inducible degron system (AID) to deplete Spt5 to capture how its loss effected the organism in the short window of time before defects in Pol II start to impact ribosome biogenesis.

We found that depletion of Spt5 leads to severe reduction in RNA synthesis starting at thirty minutes post induction of depletion. No significant changes were observed in Pol I occupancy using Pol I native elongating transcript sequencing (NET-seq) however the vehicle control differed significantly from previously published wildtype occupancy indicating the need for better controls. Depletion of Spt5 almost certainly impacts ribosome biogenesis and nominal cell growth however as of the completion of this Master's project results remain inconclusive yet promising.

Keywords: RNA polymerase I, Spt5, Spt4, NET-seq, ribosome biogenesis, auxin inducible degron system

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Central Dogma and the Divergence of RNA Polymerases

The central dogma of molecular biology explains the genetic flow of information in biological systems. At its most simplistic, it is often summarized as DNA to RNA to protein. Throughout all domains of life, this concept and the basic molecular machinery required for its operation are preserved. RNA polymerases are the machines responsible for this first step of transcription of DNA into RNA. In prokaryotes and archaea, one RNA polymerase (RNAP) is responsible for all RNA synthesis, however in eukaryotes a minimum of three RNA polymerases (Pols I, II III) are responsible for the synthesis of the varying RNA species required for life. At their core, the action of these large multiprotein complexes is quite simple, they catalyze the addition of a single ribonucleotide to a growing chain of RNA. While simple in concept, in reality, the action is endlessly complex and requires not just the participation of the many subunits of the polymerase but the participation of many different trans-acting factors as well.

First and foremost, understanding these machines at a high level of detail is a basic science necessity. RNA polymerases play a role at every level of human health, and there are still many gaps of knowledge to fill at the transcriptional level. However, beyond a basic science interest there are many avenues for treatments and therapeutics that could work on these fundamental pathways. As we characterize and uncover more about these polymerases, understanding the similarities and differences between them,

both within species and without, expands the toolbox we have to effect specific polymerases with minimal effect on others. To do this we must not only understand the functions of each subunit both shared between, and unique to, each eukaryotic polymerase, but also any additional factors necessary for functional transcription.

RNA Polymerase I and Ribosome Biogenesis

RNA polymerase I (Pol I) is the polymerase responsible for the synthesis of three of the four ribosomal RNAs (rRNA) which form the catalytic core of ribosomes. Pol I's only function is to produce large amounts of rRNA needed for rapid cell division and nominal cell function. The three transcripts synthesized by Pol I which include the 18S, 5.8S and 25S rRNAs are synthesized as a single transcript, the 35S, with external and internal transcribed spacers that are cleaved and processed out co-transcriptionally[2]. While Pol I is responsible for the bulk of the components needed for functional ribosomes, the entire process requires the combined effort of all three ubiquitous RNA

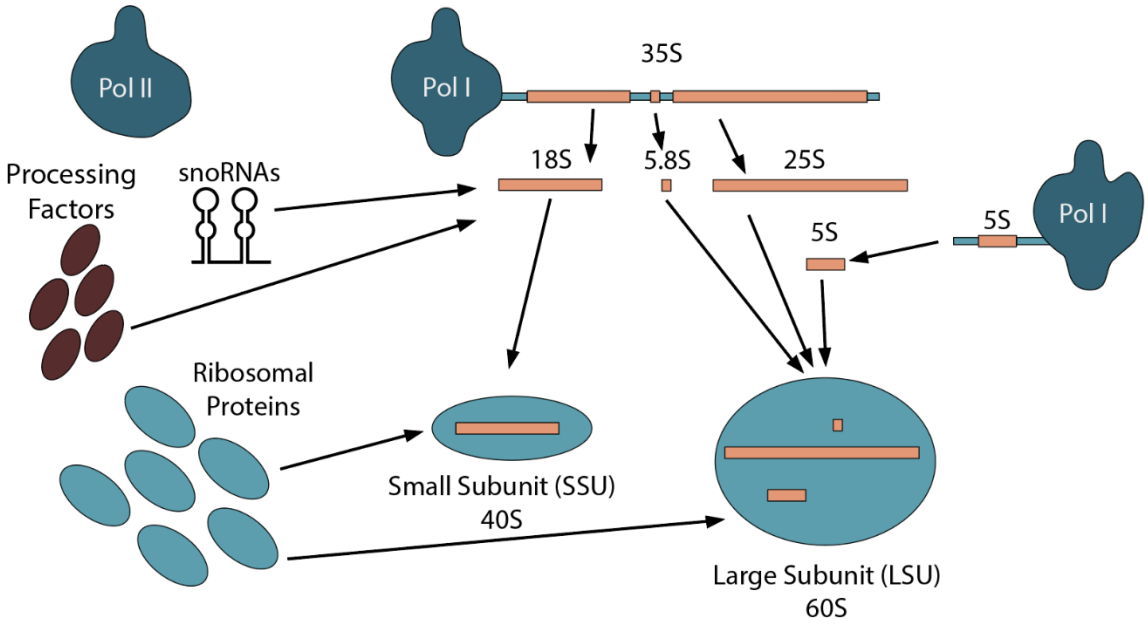

Figure 1: A simplified overview of ribosome biogenesis

polymerases visually summarized in **Figure 1**. RNA polymerase III (Pol III) is responsible for the transcription of the fourth rRNA, the 5S, and Pol II is responsible for the messenger RNA (mRNA) transcripts for ribosomal proteins or processing factors as well as non-coding RNA processing factors like snoRNAs. Targeting ribosome biogenesis is a promising target for treatment of pathogens or cancers. Some of our best antibiotics work by inhibiting the resulting ribosomes through inhibition of protein translation in the prokaryotic ribosome while allowing eukaryotic translation to continue unimpeded[3]. Targeting the machinery involve in the central dogma will always be effective, and targeting Pol I transcription is a treatment avenue that is currently being explored.

Despite the reliance on multiple polymerases for ribosome biogenesis, synthesis of rRNA by Pol I is the first and rate limiting step of ribosome biogenesis. This process is one of the most energetically costly processes in the cell and in yeast almost 80% of all newly synthesized RNA is rRNA.[4] Any eukaryotic cell, be it a cancer cell or a eukaryotic pathogen, absolutely requires highly active Pol I transcription to efficiently divide and multiply. It is this fact that has identified Pol I transcription and ribosome biogenesis as attractive targets for inhibition. In cancers, inhibition of Pol I would preferentially target the more rapidly growing cancer cells, while also minimizing DNA damage, a common side effect of many cancer treatments today[5]. As interest in targeting this pathway grows, it becomes imperative that we characterize the factors required for competent ribosome biogenesis, so that we may more selectively target Pol I and its pathway. Despite some homology between all three eukaryotic polymerases, all

three have unique machinery and factors that are required for competent transcription. Some factors are shared but most are not.

The Established Role of the Spt4/5 Complex in Pol II Transcription

Spt5, which in tandem with Pol I is the focus of this thesis project, is one protein that has been identified as a putative Pol I elongation factor. It is the only known trans-acting factor preserved over all domains of life[6]. In prokaryotes, it is known as NusG. Despite its conservation between domains of life, little is known about its conservation between eukaryotic polymerases. It was first identified in a large screen of mutations that suppressed Ty insertions (origin of spt nomenclature)[7]. The following years would establish it as an important elongation factor in Pol II. Additionally, data from our lab has shown evidence of a direct interaction between Spt5 and Pol I which has formed the underlying foundation for this project[8-10]. However before jumping into Spt5's role in Pol I transcription, it is important to look to the large body of research that has been conducted into Spt5 in Pol II. As stated previously, the eukaryotic polymerases have many conserved regions and subunits, and any data on Spt5 in Pol II, while not fully applicable, will certainly inform this investigation into Pol I.

Spt5 forms a heterodimer with the smaller Spt4 in eukaryotes. Spt5 consists of an acidic N-terminus, a NusG like N-terminal(NGN) region plus a varying number of KOW domains depending on species, and finally a series of C-terminal heptad repeats not dissimilar to Pol II's Rbp1 C-terminal repeats. Spt4 binds Spt5 at the NGN domain and is

thought to stabilize the NGN in its interactions[11, 12]. A simplified linear structure across all domains is shown in **Figure 2**.

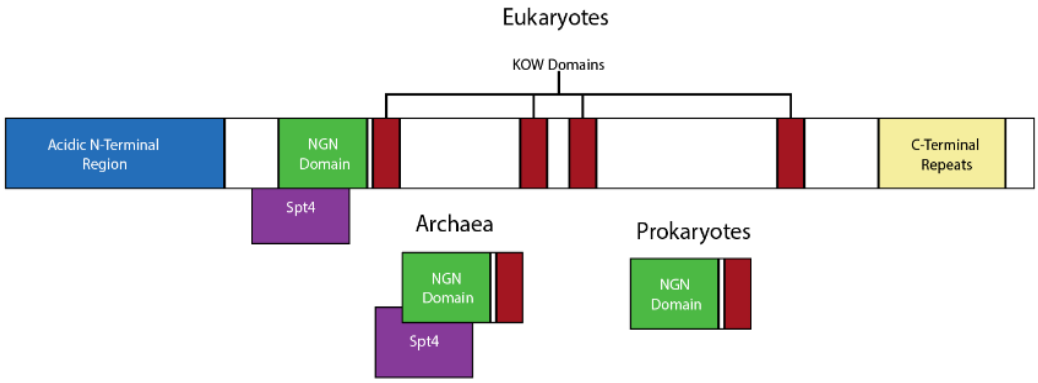

Figure 2: Simplified structures of Spt5, Spt4 and its homologs. The NGN domain and KOW region are conserved across domains, while the N-terminal Acidic region and C-terminal repeats are unique to eukaryotic organisms.

There is increasing evidence that the acidic N-terminus plays a role in nucleosome traversal[13, 14]. The NGN and KOW domains appear to sit in the cleft of Pol II and exert direct effects on the DNA and nascent RNA[15]. Recent cryo-EM data of the Spt4/5 complexed with Pol II posits that the NGN and the KOW domains may form exit channels for both the duplex DNA as well as the nascent RNA[16]. Finally, the C-terminal repeats get conditionally phosphorylated and act as a recruitment platform for other factors like Paf1c[17]; this is perhaps best characterized by its interaction with the negative elongation factor (NELF) in higher eukaryotes. Known as DSIF in metazoans, the Spt4/5 heterodimer recruits NELF leading to a long pause soon after promoter escape called promoter-proximal pausing. P-TEFb then phosphorylates Spt5 causing it to release NELF and allowing Pol II to continue to transcribe into the gene body[18-20]. Despite the ever-growing research elucidating Spt5's multiple roles in Pol II transcription, our lab is currently the only research group examining its role in Pol I transcription and the

current progress is summarized in the next section. As of this writing, no published data exists on a possible interaction of Spt5 with Pol III which is certainly an open question worth investigating.

Spt5 and RNA Polymerase I

Dr. Schneider and past students in our lab have established that Spt4/5 interacts *in vivo* and *in vitro* with Pol I subunits. Mass spectrometry of purified Pol I shows the presence of the Spt4 and Spt5 proteins with no Pol II subunits, meaning Pol II associated Spt5 contamination is unlikely.[9] Pol I and Spt5 also co-immunoprecipitate[9]. Spt5 also associates with Pol I *in vitro* with GST pull-down assays[8]; these experiments indicate that the NGN and KOW regions are the critical regions for Pol I binding These regions are also the regions necessary for Pol II binding, suggesting conserved function across the polymerases[8].

In addition to probing interactions, we have observed that genetic defects in either *SPT4* or *SPT5* also have varying effects on rRNA synthesis and coprocessing[8-10]. A mutant missing the *SPT4* gene, ***spt4Δ***, shows a modest decrease (20%) in overall growth as well as a 10-20% decrease in rRNA synthesis. ***spt4Δ*** also has defects in rRNA co-processing. ^{3}H-methyl methionine and polysome profiling indicate a reduction in mature rRNA as well as a decrease of the 60S subunit compared to the 40S. Miller EM spreads were also performed and the accumulation of uncleaved transcripts at the 3’ end of the gene was observed compared to WT[9].

Additionally, several Spt5 mutants were induced via error-prone PCR and

screened for temperature-sensitive phenotypes. One such mutant is Spt5(C292R) which grows 2.5-fold slower and has a 4-fold reduction in rRNA synthesis as determined by ^{3}H-methyl-methionine pulse chase experiments [10]. Despite its poor growth, rDNA occupancy as determined by CHIP of Spt5(C292R) was not significantly different from WT, again indicating that Spt5 may play a role in rRNA co-processing[10]. All of this evidence supports that Spt5 has many effects on Pol I transcription, and that it is likely a direct effect.

Use of Auxin Inducible Degron System for Targeted Depletion of Spt5

As mentioned previously, *SPT5* is an essential gene meaning that traditional knockout models are not an option to examine the gene's effects. Previous studies from our lab have focused on various mutations of the *SPT5* or on deletions of the associated SPT4 gene. All of these strains have provided instrumental data and evidence of Spt5's role in Pol I, however they all share one major drawback: the interaction between Pol II and Spt5. While it is most likely that the phenotypes observed were due to the disruption of Spt5/Pol I axis, it is certainly possible they were results of downstream effects of the Spt5/Pol II axis, which has been much more established in the literature. To circumvent this issue, we plan on using the Auxin Inducible Degron (AID) system to rapidly deplete Spt5 *in vivo* in yeast in hopes of capturing its effects in Pol I transcription before the defects in Pol II transcription affect the ribosome biogenesis pathway. Ribosome biogenesis requires several proteins whose expression is dependent on Pol II, for example ribosomal proteins that comprise a portion of functional ribosomes or ribosomal assembly proteins

which help fold the rRNA and incorporate the ribosomal proteins[5]. Undoubtedly defects in these protein's expression will start to affect Pol I transcription given enough time.

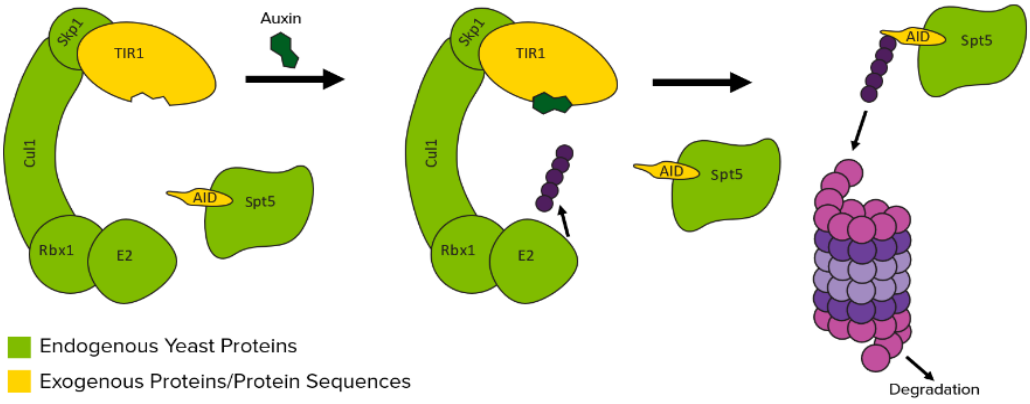

Figure 3: Auxin Inducible Degron (AID) system mechanism

The AID system uses the Tir1 protein from *Oryza sativa* and AID tag taken from the IAA17 gene from *Arabidopsis thaliana,* which is added on the protein of interest, in our case Spt5[21]. The Tir1 protein is an F-Box protein that associates with an E3 ubiquitin ligase, known as the SCF complex. The SCF complex is shared between most eukaryotes with modular F-box proteins allowing for differential degradation[22]. The Tir1 F-box is absent in yeast, yet because the other proteins of the complex are conserved, the exogenous Tir1 can incorporate in this complex and form a functional E3 ligase. In the presence of the plant hormone auxin (also referred to as indole acetic acid or IAA throughout this document), the Tir1 protein can recognize an AID tag on the N or C terminus of a protein and ubiquitinate it, sending it down the cell's endogenous protein degradation pathways. A basic schematic of the mechanism is shown in **Figure 3.** Unfortunately, while the system shows efficient degradation over a short time period, there is a major concern for the basal level of degradation that occurs even in the absence of IAA[21]. For this reason, there are several modifications that have been made to the AID system over the years since its

development [23-25]. The most promising is a recent modification to the system that use a mutated version of TIR1, OsTIR1^{F74A}, and a synthetic auxin derivative 5-Adamantyl-IAA (5-Ad-IAA), which not only reduces background degradation, but also reduces the concentration of compound needed to trigger degradation[24].

Hypothesis

The wealth of preliminary data from the previous work of our group provides strong evidence of a potential interaction between Pol I and Spt5. This project aims to confirm or refute that interaction with either being an important finding. I believe that there is strong evidence that Spt5 directly impacts Pol I transcription and therefore I hypothesize that by degrading Spt5 using the AID system, we will observe severe defects in Pol I processivity, both through RNA synthesis as well as RNA polymerase I occupancy of the rDNA.

CHAPTER 2: METHODS

AID System Deployments and Strain Constructions

For this project *Saccharomyces cerevisiae* was used as the model organism to deplete Spt5. Two differing implementations of the AID system were used during this thesis project. Both strains required the AID system in place which includes the AID tag on the C-terminus of Spt5 as well as the full F-box Tir1 protein, in addition to these AID components, the experimental strains required an HA tag on the second largest subunit of Pol I A135 which is used for immunoprecipitation in NET-Seq. We began by using the Spt5-AID strains from Maudlin and Beggs, 2021[26]. It used a β-estradiol triggered artificial promoter to induce the transcription of TIR1. This strain was attractive as it aimed to reduce the leaky degradation caused by auxin independent degradation.

Discussed further in results, we decided to move on from the first strain and onto another. We identified another strain that used the AID system that had TIR1 on a constitutive ADH1 promoter[13]. Two strains were sent to us by the Labib lab. Both strains had the TIR1 gene with a constitutive ADH1 promoter genomic inserted, but where one strain had Spt5 with an AID tag and the other had the endogenous untagged Spt5 (this strain was to be used as the control). In this case we decided to add the HA tag to A135 in both these strains rather than take its components and add them to our NET-seq wildtype strain. This required only one transformation for each strain that would transform in an HA-tag onto the C-terminus of the A135 gene.

PCR was used to amplify linear C-terminal fragments of the A135 gene of the strain DAS499 which already had A135 tagged with HA along with a LEU1 selection marker downstream of the gene. Using the LiAc/SS carrier DNA/PEG method which relies on homologous recombination, the linear fragment was transformed into both strains' genome[27]. Successful transformations were identified by growth on SD -Leu plates. Strains were subsequently confirmed by comparing PCR amplifications of the A135 C-terminal end of the gene on an agarose gel.

Western Blotting

Cells were inoculated and grown to OD_{600} 0.2 before 0.75 mM of IAA or the equivalent volume of the vehicle ethanol was added. Whole cell lysates were collected by mechanical lysis cells in RIPA buffer. Protein concentrations were quantified by BCA. Lysates were added to 2 X protein loading dye with BME and run on an 8% SDS-PAGE gel at 100V. Proteins were transferred to a PVDF membrane using a Semi-Dry transfer run at a constant 250 mAmps for 4.5 hours. Anti-Spt5 and anti-AID-tag antibodies were used as primaries. Secondary antibodies employed fluorescence. Blots were imaged on the ChemiDoc MP Imaging System (Bio-Rad).

Growth Curves of Spt5 Auxin Induced Degradation

50 mL of YEPD was inoculated with either the Spt5-AID strain or the endogenous untagged Spt5 control strain and allowed to grow to OD_{600} 0.2 at 30°C before the addition of 0.75 mM IAA or the equal volume of the vehicle ethanol. Starting once IAA or vehicle was added to the culture, optical density readings were taken every

30 minutes for the following four hours and then left to grow overnight with a final reading at 24 hours.

Single Plate-Serial Dilution Spotting

Both the Spt5-AID strain and the endogenous untagged Spt5 control strain were grown to saturation overnight. The saturated cultures were then diluted to an OD_{600} of 0.5 and then subsequently diluted by 10, 100, 1,000, and 10,000. 6µL of the 1X, 10X, 100X, 1,000X, and 10,000X dilutions of each strain was spotted onto either a YEPD agar plate or a YEPD + 0.5 mM IAA agar plate and allowed to grow in a 30°C incubator for 48 hours. After 48 hours, plates were photographed.

Monitoring rRNA synthesis via RT-qPCR

50 mL of YEPD was inoculated and allowed to grow to an OD_{600} of 0.2 at 30°C before the addition of 0.75 mM IAA or the equal volume of the vehicle ethanol. After desired IAA incubation time, 1 mL of culture was harvested, and cells were washed and pelleted before being frozen at -80 °C. Once ready for RNA extraction, pellets were resuspended in TES buffer. Acid phenol was added to the resuspended cells. The resuspension was incubated at 65°C for one hour with 10 seconds of vortexing every 15 minutes. After the incubation, samples were centrifuged at max for 10 minutes and the top layer was extracted. One additional phenol extraction and two chloroform extractions were performed. After extraction, the final upper layer was placed in 1 M ammonium acetate in 100% ethanol at -80°C to precipitate overnight. Precipitated RNA was centrifuged for 20 minutes at 4°C at max. The supernatant was removed, and the pellet

was washed twice with 75% ethanol. The pellet was dried, and RNA was resuspended in 100 µL of MilliQ water. RNA quality and concentration was then assessed using a NanoDrop Spectrophotometer. Samples were then reversed transcribed into cDNA using SuperScript III RTase. The subsequent cDNA was then diluted by ten by adding 10 µL of cDNA in 90 µL of MilliQ water. Probes for the ITS1 and 18S region of the rDNA were used along with Sybr Green. Reactions were run and read on the CFX Opus 96 Real-Time PCR System (Bio-Rad). The $\Delta\Delta C_t$ method was used to calculate the fold changes of the results.

Monitoring rRNA synthesis via ^{3}H-Uridine

^{3}H-uridine pulse labeling was used to measure total RNA synthesis over a 5-minute time span. Cultures were inoculated and allowed to grow in YEPD to OD_{600} of 0.2 at 30°C before the addition of 0.75 mM IAA. Time points were taken before the addition of IAA (0 minutes), then at 30 minutes, 45 minutes, and 60 minutes post IAA addition. Before beginning the time course, 10µCi of tritiated uridine was added to a plastic culture tube and the tube was placed in a 30°C incubator to warm up. At the desired time points, 1 mL of the culture was added to a culture tube containing the tritiated uridine. The cultures were allowed to grow at 30°C with shaking for 5 minutes. After five minutes, cells were lysed with 2.5 mL of 10% trichloroacetic acid (TCA) and were placed immediately on ice. After roughly 30 minutes on ice, lysates were filtered with a 0.45 µm nitrocellulose filter. The filters were washed with 10 mL of 5% TCA, and then the filters were allowed to dry for 30 minutes before being placed in a scintillation

vial with 4 mL scintillation cocktail. The relative radioactivity of each sample was then measured in counts per minute (CPM) on a liquid scintillation counter.

Examining RNA Polymerase I Occupancy through NET-Seq

A Pol I optimized version of NET-seq developed by the Schneider lab was used to probe Pol I occupancy of the rDNA[1, 28, 29]. Using 1 L of YEPD, cultures were inoculated and allowed to grow to an OD_{600} of 0.2 at 30°C before the addition of 0.75 mM IAA or an equivalent volume of ethanol. After 45 minutes, cells were filtered and then quickly frozen in liquid nitrogen. Cells were then cryo-lysed including EDTA in the lysis buffer to ensure that transcription is halted. After cryo-lysis, Pol I was immunoprecipitated with anti-HA magnetic beads. RNA was then isolated from the immunoprecipitant using two phenol extractions followed by two chloroform extractions. After precipitation in 1 M ammonium acetate overnight at -80 °C, the RNA was first subjected to linker ligation which uses the unique truncated T4 to ligate a preadenylated 5'end to a free 3' OH which allows us to ideally select only nascent transcripts with that group free. After ligation, the transcripts are then fragmented with $ZnCl_2$ and reversed transcribed before being run on a denaturing polyacrylamide/urea gel to select transcripts based on size. A region of the gel from roughly 120 nucleotides to 600 nucleotides as determined by the ladder is excised, and the reverse transcribed DNA is extracted from the excised gel. Once the reverse transcribed DNA is extracted it is circularized both for stability and to give us a known 3' end for PCR amplification. The circularized DNA is then PCR amplified and cleaned using Aline PCRCleanDX beads. Finished libraries are

then sent to be sequenced. A more in-depth protocol can be found in these references[1, 28, 29].

Data-Analysis and Snakemake Pipeline

After sequencing, data was trimmed and aligned as described in Huffines et al. [1]. The steps first include a PCR deduplication to remove any PCR duplicates using fqtrim [30]. Then cutadapt was used to trim adapter sequences[31]. First the 5' adapter is removed, followed by the 3' adapter. The sequences are then aligned to a modified version of the SacCer3 genome using STAR[32]. Modifications to chromosome XII were made to mask repeated regions of the ribosomal DNA. This removal allows for reads that map to the rDNA to be unique and thus have a stronger mapping score during alignment. To mask these regions all but one repeated rDNA sequences were replaced with N so that genome organization remained the same. Areas masked include: chromosome XII: 460,734 – 471,775, chromosome XII: 418,898 – 489,469, and chromosome XII: 489,753 – 490,545.

After alignment, the resulting bam files are sorted and indexed using SAMtools[33]. Then bedfiles are made using the genomecov tool of BEDTools which assemble the counts based on the location of the 3' end[34]. The bedfiles are trimmed to the rDNA and exported for further analysis in R. Additionally, htseq is used to create gene count files[35]. These count files will mainly be used as additional quality control to ensure that a majority of the reads map the rDNA. Once in R, the rDNA reads were normalized by dividing individual counts by the total sum of rDNA counts on the positive strand divided by 1,000,000.

Additionally, as part of this project I developed a pipeline to streamline these steps for high performance computing using snakemake[36]. Snakemake is a python-based tool for workflow management than can be used for cluster computing on a high performance computing (HPC) cluster like UAB's Cheaha. The code and a step-by-step breakdown of the rules employed by snakemake can be found here: https://github.com/nathanbellis/SchneiderNETSeqPipeline .

CHAPTER 3: RESULTS

Successful Transformations

The HA-tag transformations were successful and agarose gels of PCR confirmations are pictured below. A PCR amplified linear fragment of DAS499 which is source strain of the HA-tagged A135 C-terminus was run as the control. The same primers used to amplify the linear construct from DAS499 were also used to amplify the same region of successful transformants identified by growth on selective media. The resulting agarose gels comparing those fragments are displayed below in **Figure 4.**

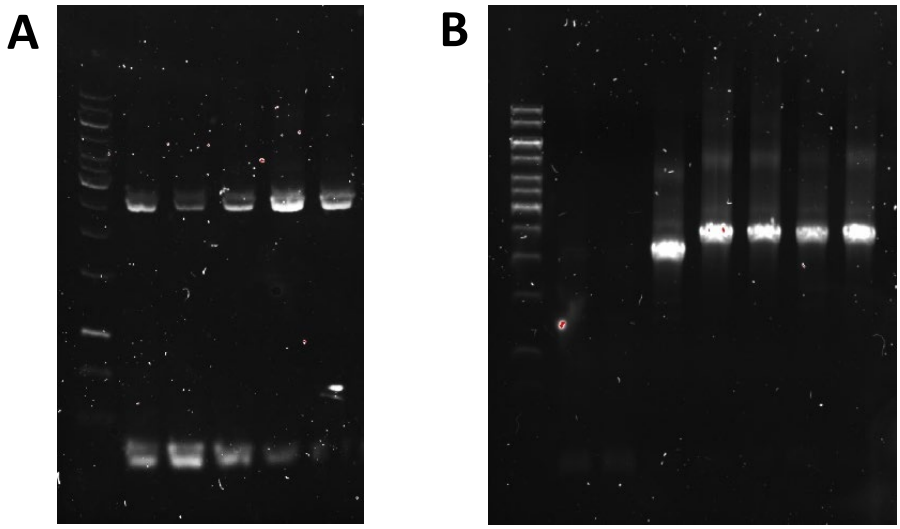

Figure 4: PCR confirmation by agarose gels. A) PCR fragments run on a 0.8% agarose gel. Lanes 1-3 are technical replicates of fragments amplified from DNA of successful transformants of the A135-HA into the strain with only ADH1::TIR1. Lanes 4-5 are fragments amplified from DAS499, source of the A135-HA tag fragment. B) PCR fragments run on a 0.8% agarose gel. Lanes 1-3 are part of another experiment and can be ignored. Lanes 4-5 are technical replicates of fragments amplified from DNA of successful transformants of the A135-HA into the strain with both ADH1::TIR1 and the AID tagged Spt5. Lanes 6-7 are fragments amplified from DAS499, source of the A135-HA tag fragment.

With the strains PCR confirmed, freezer stocks were made and kept at -80 °C. This resulted in two strains which are identified as Spt5-AID and Control throughout the document. The genotypes listed are below in **Table 1**.

Table 1: Strain genotypes. Genotypes for the *S. cerevisiae* strains used in this thesis project.

Strain Name	Genotype
Spt5-AID	MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100 spt5-aid (HphNT) ura3-1::ADH1-OsTIR1-9Myc (URA3 & klTRP1) RPA135-(HA)3- (His)7::LEU1
Control	MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100 ura3-1::ADH1-OsTIR1-9Myc (URA3 & klTRP1) RPA135-(HA)3-(His)7::LEU1

Degradation

Initially, the first strain using β-estradiol triggered TIR1, strong degradation was never observed even after several attempts using varying incubation times and concentrations of both β-estradiol and IAA. My hypothesis with this strain is that we were never able to induce strong enough expression of TIR1 for complete degradation of a high abundance protein like Spt5. As a result, the strain was abandoned for strains that placed TIR1 on a constitutive promoter.

The main struggle with observing degration has been finding a good antibody for Spt5, as a result, I struggled to get a clean western blot with sharp depletion. **Figure 5** shows the best blot we were able to obtain, faint bands were observable, but incredibly high background kept us from longer exposures making it a sub-par blot. An anti-Spt5 antibody was used, and bands were present at the expected weight, though the signal

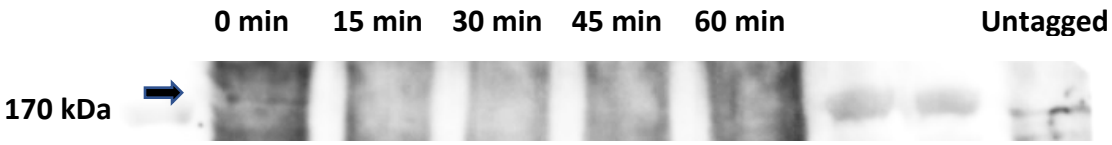

Figure 5: Spt5 degradation blot: Western Blot of Spt5 using anti-Spt5 antibody. A faint band slightly above 170 kDa is present and at the 0 time point. The band maybe slightly visible at 15 minutes and then disappears by 30 minutes and beyond. Despite apparent degradation, the background is too high to confidently quantify degradation. Additionally, a lysate with endogenous untagged Spt5 was also run in the final lane as a control. Note the slightly lower molecular weight due to lack of the AID tag.

remained low. AID tagged Spt5 is present in the 0 minute time point and disappears in the 15-60 minute time point however the background is very high and keeps us from being able to quantify the degradation. In the final lane is an untagged Spt5, its lower molecular weight is due to the lack of the AID tag. Expression of Spt5 seems low even at 0 minutes, pointing to high level of basal degradation. Additionally, there is non-specific banding with much higher signal during blotting (not shown) which could possibly be degradation intermediates. The system does appear to be working in some capacity, as evidenced by the effects we see on growth discussed below, however, there are major concerns that the basal level of degradation occurring may be too high to use the Spt5-AID + vehicle as a control.

Loss of Spt5 Severely Impacts Rapid Cell Division

Both the Spt5-AID and Control strain's growth were monitored for four hours following addition of either IAA or an equivalent volume of the vehicle ethanol; results are displayed in **Figure 6**. The Control strain's immediate growth with both IAA and vehicle was unaffected. The Spt5-AID strain with vehicle also appeared unaffected

though by the four-hour time point the difference in doubling time between the Control and Spt-AID strains due to Spt5-AID's basal level of auxin independent degradation becomes evident. The growth of the Spt5-AID strain with IAA starts to visibly diverge starting at about 1.5 hours and by 4 hours is severely impacted, indicating a detrimental effect on growth caused by the loss of Spt5. Measurements taken at 24-hours post IAA addition, not included in **Figure 6,** also suggest a failure to recover from the Spt5 degradation. At 24 hours Spt5-AID + IAA was at an average OD_{600} of 1.58 at 24 hours and the Spt5-AID + vehicle was at an average OD_{600} of 5.46.

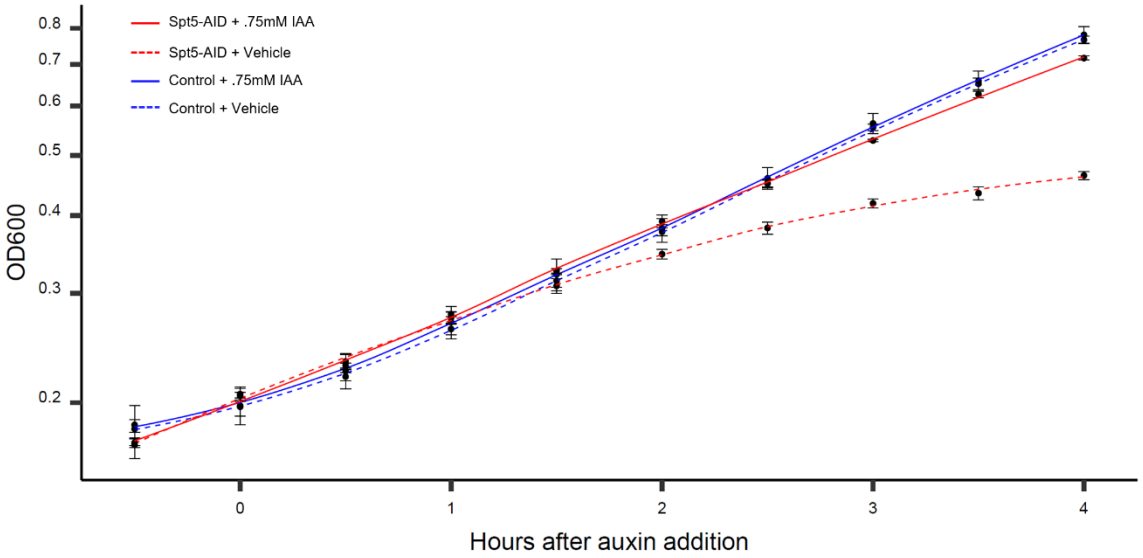

Figure 6: Growth curves of auxin induced depletion. Either 0.75 mM IAA or an equivalent volume of the vehicle ethanol was added to the Spt5-AID Strain and the Control strain with Tir1 and endogenous untagged Spt5 at time zero. Optical density was measured via spectrophotometer to monitor growth for the next four hours.

This inability to grow in the prescence of auxin is confirmed in the single-plate serial dilution spotting shown in **Figure 7**. Both the Spt5-AID and the Control strain grew well on YEPD, though again the difference in doubling time between the strains is observable as the Control strain shows stronger growth on the untreated plate. However

Figure 7: Single plate serial dilution spotting. Cultures at OD_{600} = 0.5 of both Spt5-AID Strain and the Control strain with Tir1 and endogenous untagged Spt5 were serial diluted and spotted on to either untreated YEPD plates or YEPD + 0.5mM IAA.

on the YEPD + 0.5mM IAA plate, the Spt5-AID strain failed to grow even at the highest concentration of starting cells. This supports the idea that Spt5 is absolutely essential for rapid cell division. Both the liquid and solid culture assays, strongly indicate that it is the degradation of Spt5 and not the moderate toxicity of IAA that is causing growth defects in the Spt5-AID strain.

Loss of Spt5 Severely Impacts rRNA Synthesis

Defects in growth in both liquid and solid media caused by auxin induced degradation showed evidence that the system is in place and that Spt5 is essential for cell growth. However, these results could be due to Spt5's function in Pol II transcription, so examinations of effects on rRNA synthesis were needed. I first tried qPCR to measure nascent rRNA synthesis using the ITS1 region as my reporter transcript. The ITS1 region of the 35S species of rRNA was chosen as the target to measure nascent rRNA synthesis as it is quickly processed and degraded[37]. Due to this short lifespan, it can be used as a measure of newly synthesized Pol I transcripts. The 18S species of rRNA was chosen as our normalization gene in hopes that the large population of mature and maturing 18S

rRNA would be a stable population despite perturbations at the nascent level. Traditional normalization transcripts from the mRNA population were seen as undesirable as Spt5's effect on Pol II is well established and these populations are likely altered after Spt5 degradation making them poor controls.

Time points at 60 minutes and 90 minutes post addition of either 0.75 mM IAA or an equivalent volume of the vehicle ethanol were chosen as the starting time points. Unexpectedly, as shown in **Figure 8,** no change was observed for either time point. Some changes in rRNA synthesis were expected due to the rapid defects in growth, however qPCR suggest that there is no change in the amount of ITS1 when compared with the 18S region. However, if we examine previously published Miller EM spreads for both ***spt4Δ*** and the mutant *SPT5*C292R(not shown), both strains display processing defects on nascent rRNA with many of the RNA strands failing to cleave[8, 10]. This means that the

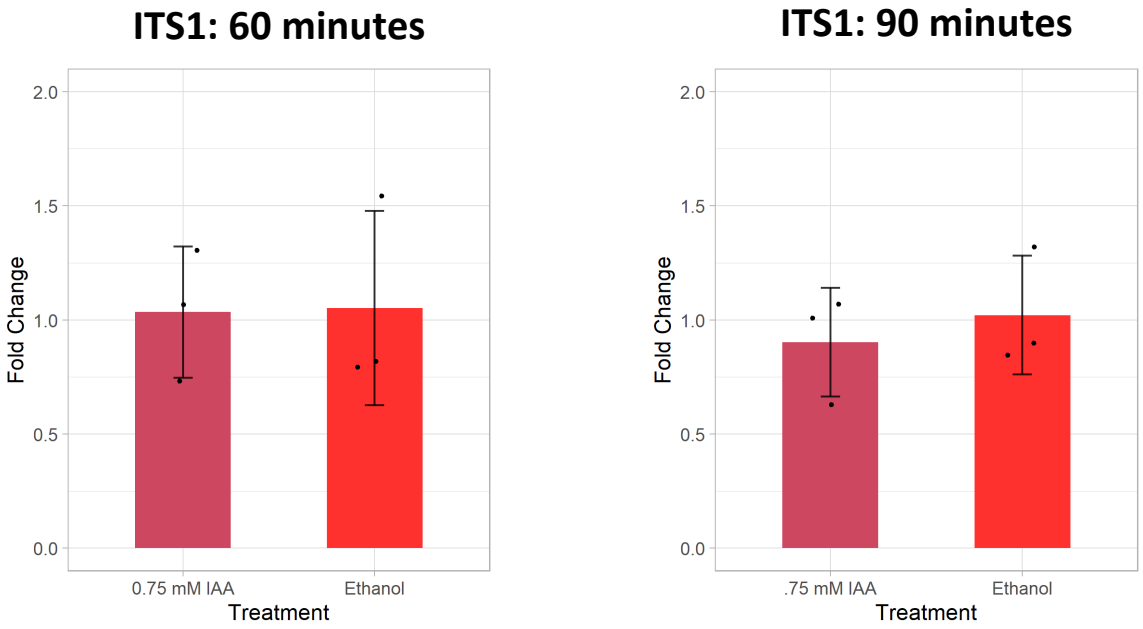

Figure 3: RT-qPCR of ITS1 region of 35S gene. RNA was extracted from Spt5-AID + IAA and Spt5-AID + vehicle at 60 minutes and 90 minutes after addition of IAA or vehicle. Primers against the ITS1 region and the 18S region were used for RT-qPCR. Using 18S for normalization the ΔΔCt method was used to calculate fold change between conditions.

ITS1 region of the rRNA could possibly persist longer than normal and maybe offset any shifts in Pol I synthesis rate in the qPCR experiments. With these results, along with the difficulty of finding solid normalization genes, we determined that ^{3}H-uridine pulse labeling would provide a better assay for monitoring nascent rRNA synthesis.

^{3}H-uridine measures the incorporation of uridine into newly synthesized RNA over a discrete window of time. The assay provides data on total RNA synthesis not just rRNA synthesis, however, in rapidly dividing yeast cells, 80% of newly synthesized RNA is rRNA and thus produced by Pol I[4]. So, if we see reduction beyond 20% percent, we can be sure that Pol I transcription is being affected. Cultures of Spt5-AID

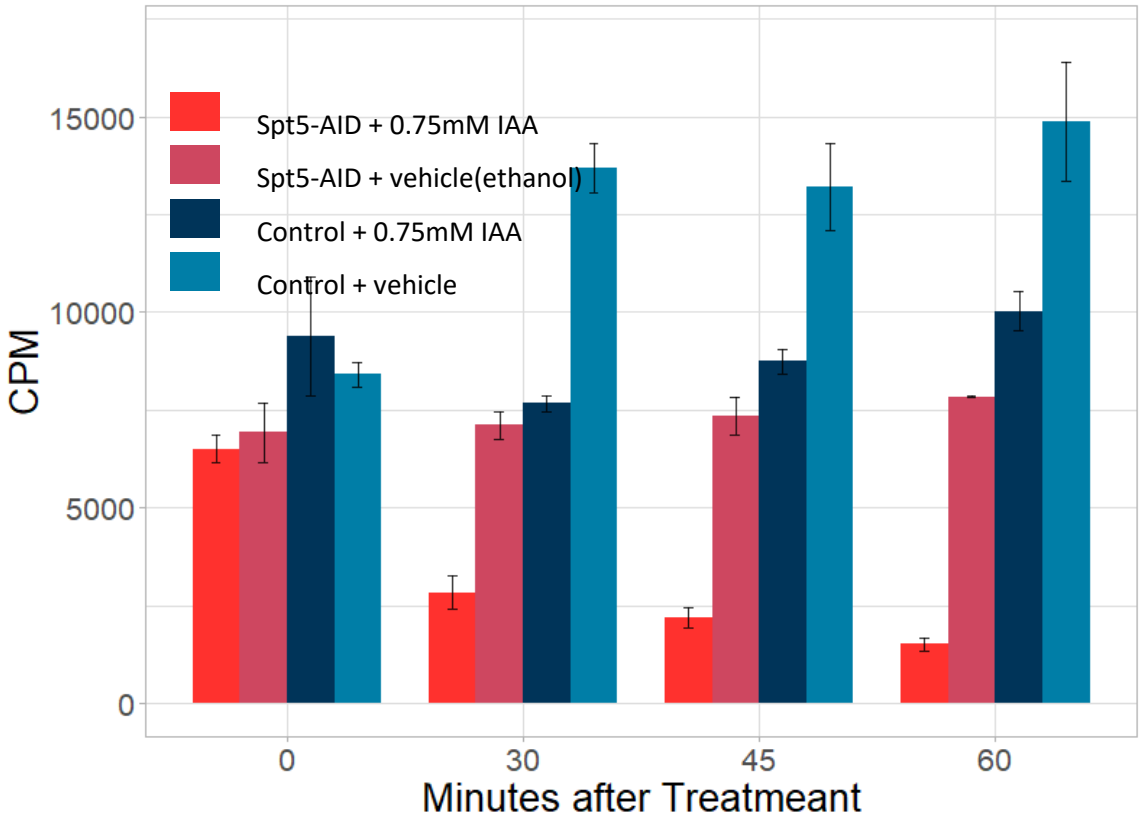

Figure 9: RNA synthesis measured via ^{3}H-uridine incorporation assay. Cultures of both the Spt5-AID and Control strain with either IAA or vehicle were pulsed for 5 minutes with tritiated uridine. Samples were collected right before the addition of either IAA or vehicle and then at 30, 45, and 60 minutes after.

and the Control strain were pulse-labeled for 5 minutes at four different time points for each strain and condition: right before addition of either IAA or vehicle (0 minutes) and then at 30 minutes, 45 minutes, and 60 minutes after.

In contrast to the qPCR results, the ^{3}H-Uridine pulse labeling **(Figure 9)** indicates that RNA synthesis and by extension Pol I transcription appear to be severely impacted by Spt5's loss. Right before the addition of either IAA or vehicle at time 0, both strains show no difference between conditions as expected. Additionally, the control strains faster growth is strongly displayed. At 30 minutes, RNA synthesis in both the Control and Spt5-AID in the presence of IAA appear strongly decreased, which is unfortunate but somewhat expected due to the toxicity of IAA at higher concentrations. The Spt5-AID + vehicle maintains its synthesis rate and the Control + vehicle increases indicating little immediate effect of ethanol on RNA synthesis rate. By 45 minutes, the Control + IAA appears to recover from the initial IAA toxicity, however the Spt5-AID + IAA fails to recover with a synthesis rate even lower than at 30 minutes. This trend continues at 60 minutes with the Spt5-AID + IAA sample synthesizing about 80% less RNA than the Spt5-AID vehicle, meaning that rRNA synthesis is certainly negatively affected. Contrarily, by 60 minutes the Control + IAA group has started to recover, and its growth rate is trending upwards compared with the 45-minute timepoint. When Control + IAA is compared with the Control + vehicle there is only 33% less RNA synthesized.

Preliminary NET-Seq Shows No Difference After Spt5 Loss

Using the ^{3}H-uridine experiments as a guide, 45 minutes was chosen as the best starting time point to capture Pol I occupancy via NET-seq. To start, samples were

harvested for the Spt5-AID strain with IAA or vehicle. There is only one replicate so far for each condition so comparisons at this point are tentative. Surprisingly, despite the difference in synthesis rate observed in the ^{3}H-uridine experiments, the samples show very little difference between each other and are very tightly correlated. Spearman rank correlation was used to compare the samples to each other. Previous Pol I NET-seq data from our lab from Huffines et al. was also included in the analysis[1]. This dataset includes three replicates of wildtype yeast as well as three replicates of a ***spt4Δ*** mutant, which is discussed briefly in the introduction. As Spt5 is part of the heterodimer Spt4/5 during transcription, the ***spt4Δ*** mutant was certainly worth comparing with the AID data.

As observed in **Figure 10**, the Spt5-AID + vehicle and the Spt5-AID + IAA are tightly correlated with each other with a Spearman's correlation coefficient of 0.984. However, neither sample correlate well with the previously published NET-seq experiments from our group. Both samples have a correlation coefficient less than 0.80 for all the previous samples. This is particularly concerning with the lack of similarity between Spt5-AID + vehicle and previously published wildtype, as the Spt5-AID +vehicle was intended as a potential control.

Contamination by mature rRNA is a known issue and can potentially drive-up correlation when looking at the 35S as a whole. To confirm the similarities of the two new samples the datasets were split by individual rDNA regions so that the transcribed spacers, which are more experimentally relevant, could be observed. The Spearman's rank tests were rerun on the split data for both the new samples as well as on averaged samples of the wildtype and ***spt4Δ*** from Huffines et al. (**Figure 11**). However even after the split, the high level of correlation between the samples is maintained. Additionally,

the Spt5-AID + vehicle is still highly divergent from wildtype, indicating the sample as a poor control for the experiment.

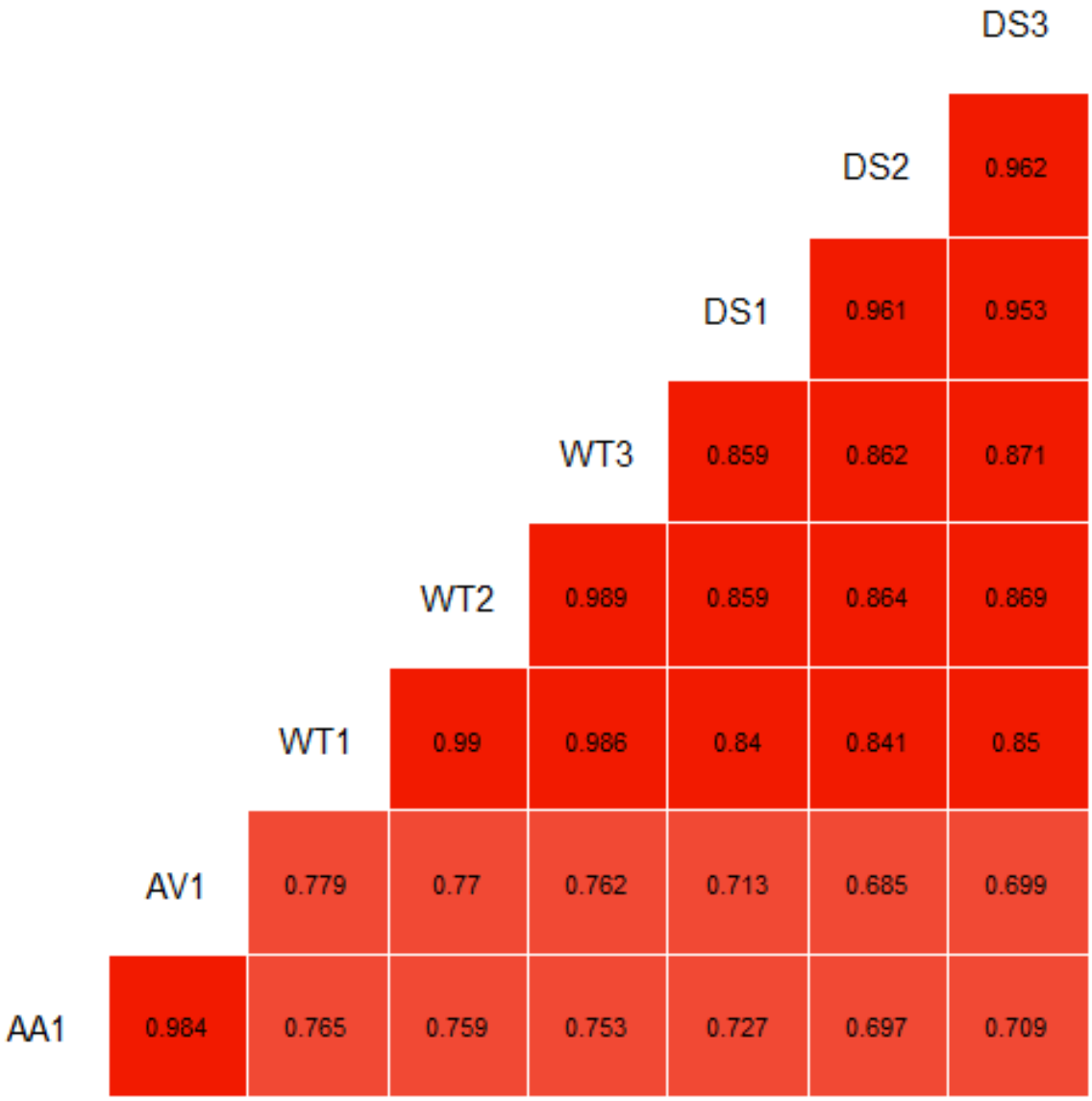

Figure 10: Spearman's correlation of full 35S gene. Spearman's rank correlation of the 35 S gene. Samples compared are one replicate of Spt5-AID + IAA(AA1), one replicate of Spt5-AID + vehicle (AV1) and three replicates of previously published wildtype (WT1-3) and three replicates of previously published ***spt4Δ*** mutant (DS1-3) from Huffines et al[1].

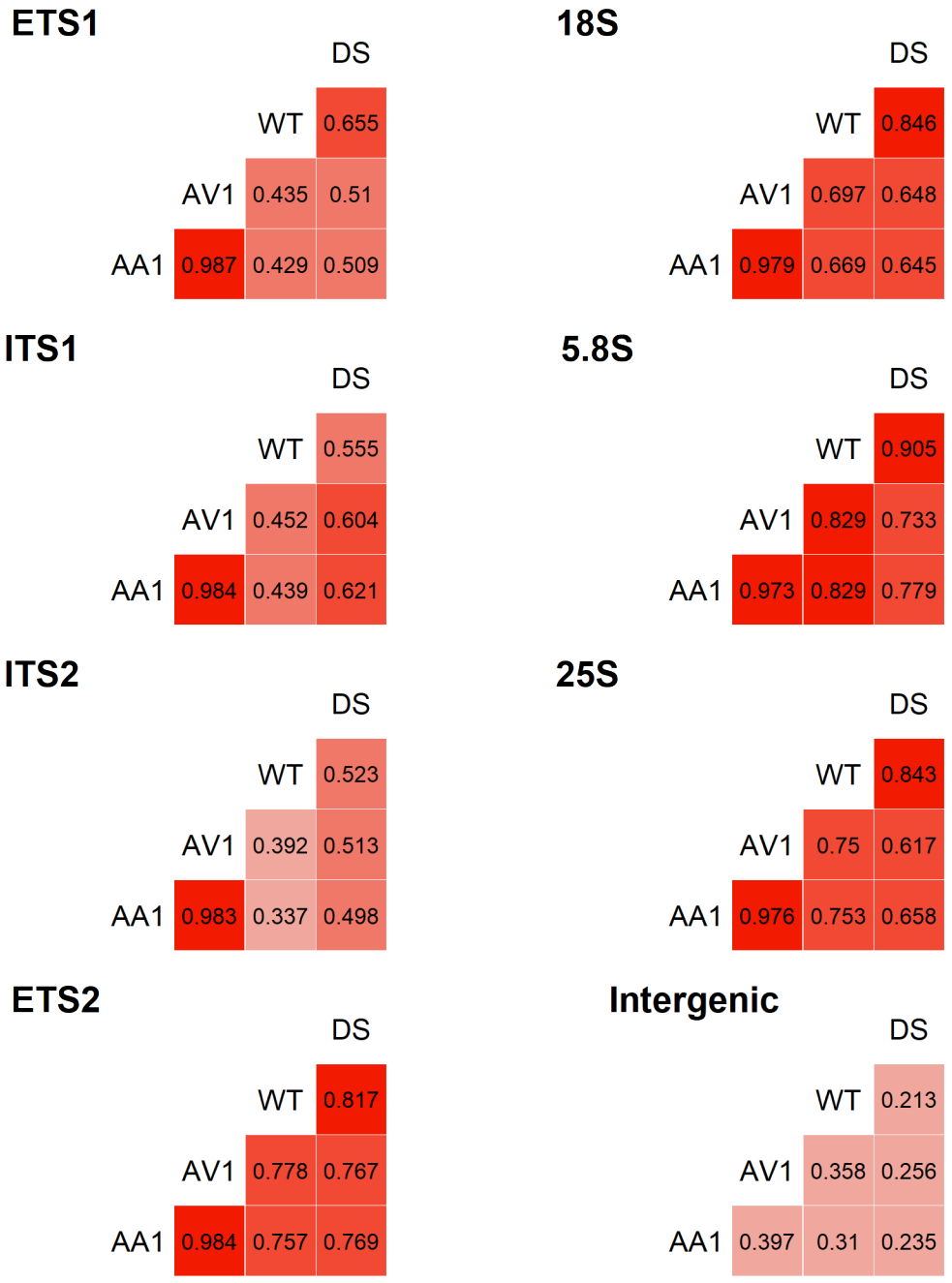

Figure 4: Spearman's correlation matrix of 35S gene split by region. Spearman's rank correlation of the 35 S gene split by region (ETS1, 18S, ITS1, 5.8S, ITS2, 25S, ETS2, and the intergenic region. Samples compared are one replicate of Spt5-AID + IAA (AA1), one replicate of Spt5-AID + vehicle (AV1), the average of three replicates of previously published wildtype (WT) and the average of three replicates of previously published ***spt4Δ*** mutant(DS) from Huffines et al[1].

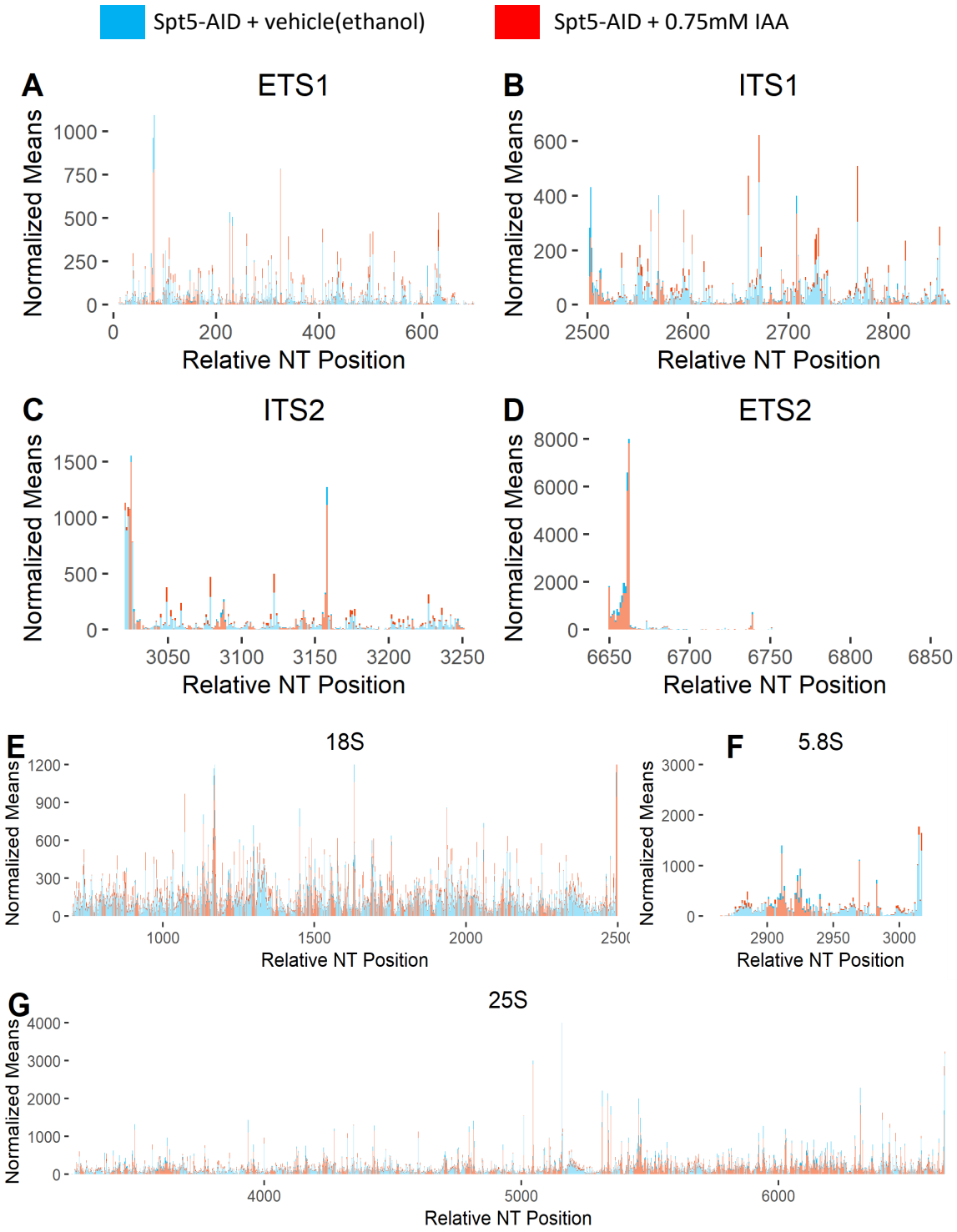

Figure 12: NET-seq of the 35S gene split by region. Histograms of NET-seq counts of the 35S gene of the rDNA. The graphs plot the normalized counts for each individual nucleotide coordinate of the rDNA and is split by region. Histograms are plotted with the lower value layered in front.

To see if there was any difference at the single nucleotide level, the normalized reads were again split by region and plotted separately as shown in **Figure 12**. ITS and ETS reads will be the primary regions we will examine as we expect them to be the most reflective of nascent transcription due to the lack of contaminating mature and intermediate rRNA products.

These plots visually validate the similarities we saw in the Spearman rank correlation tests between the samples, as there are no obvious differences between the two samples. It is clear that more replicates as well as additional controls will be needed to draw any definitive conclusions. The Control + IAA would be the first choice followed by the Control + vehicle. Despite similarities between our IAA and vehicle samples, some preliminary comparisons can be made to previously collected wildtype and ***spt4Δ*** NET-seq libraries as there are some noticeable differences between them.

One of the strengths of NET-seq is the ability to examine polymerase occupancy at single nucleotide resolution. This means single-nucleotide locations with significantly high reads mapping can be identified as "pausing sites" and the sequence surrounding that location can be used to examine if particular motifs or residues impact pausing. To examine the regions surrounding these "pause sites", Difflogos were generated using the DiffLogo package in R to compare the sequences surrounding the positions with the highest counts. Again, to control for the contaminating mature rRNA, only the spacer regions were looked at. Out of the 1515 nucleotide positions between the spacers, the 75 highest positions (5%) were selected and compared. The Spt5-AID + IAA and the Spt5-AID + vehicle DiffLogo (**Figure 13**) show very little difference, which by this point after seeing the Spearman's correlation as well as the occupancy histogram is unsurprising.

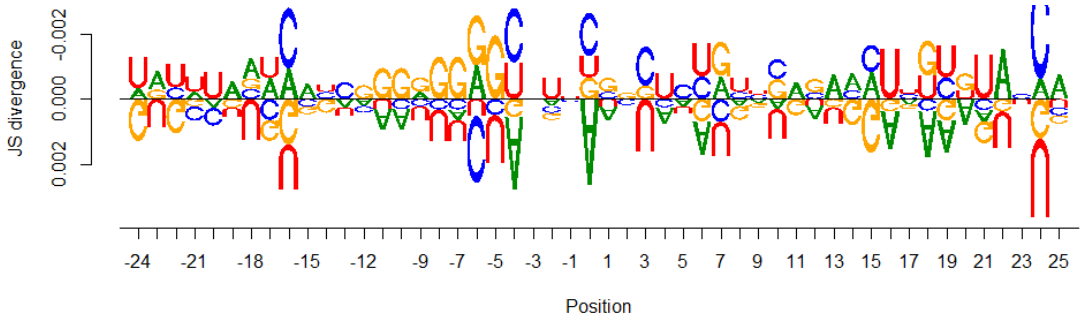

Figure 5: DiffLogo of Spt5-AID + IAA compared with Spt5-AID + vehicle. The top 5% positions with the highest counts of each sample were compared and a DiffLogo was generated. No significant difference was observed. Note the incredibly small y-axis.

However, both the IAA and Vehicle samples of the Spt5-AID have some striking differences in surrounding sequences when compared to previously published wild type and ***spt4Δ*** (**Figure 14**). When compared to both the previously published samples, the new Spt5-AID samples display a significant preference for the pause sites to be located in sequences where there is a U as last incorporated nucleotide (position 0) as well as at the four upstream nucleotides (positions -4 through -1). It is important to note that while these differences are interesting, our intended control Spt5-AID + vehicle did not correlate with wildtype meaning this difference could simply be caused by technical differences during library preparation. Reproduction of these data along with reproduction of a more "wildtype" NET-seq phenotype are needed to validate these results as biologically. Either way the vast divergence of the samples to previously published data is worth exploring.

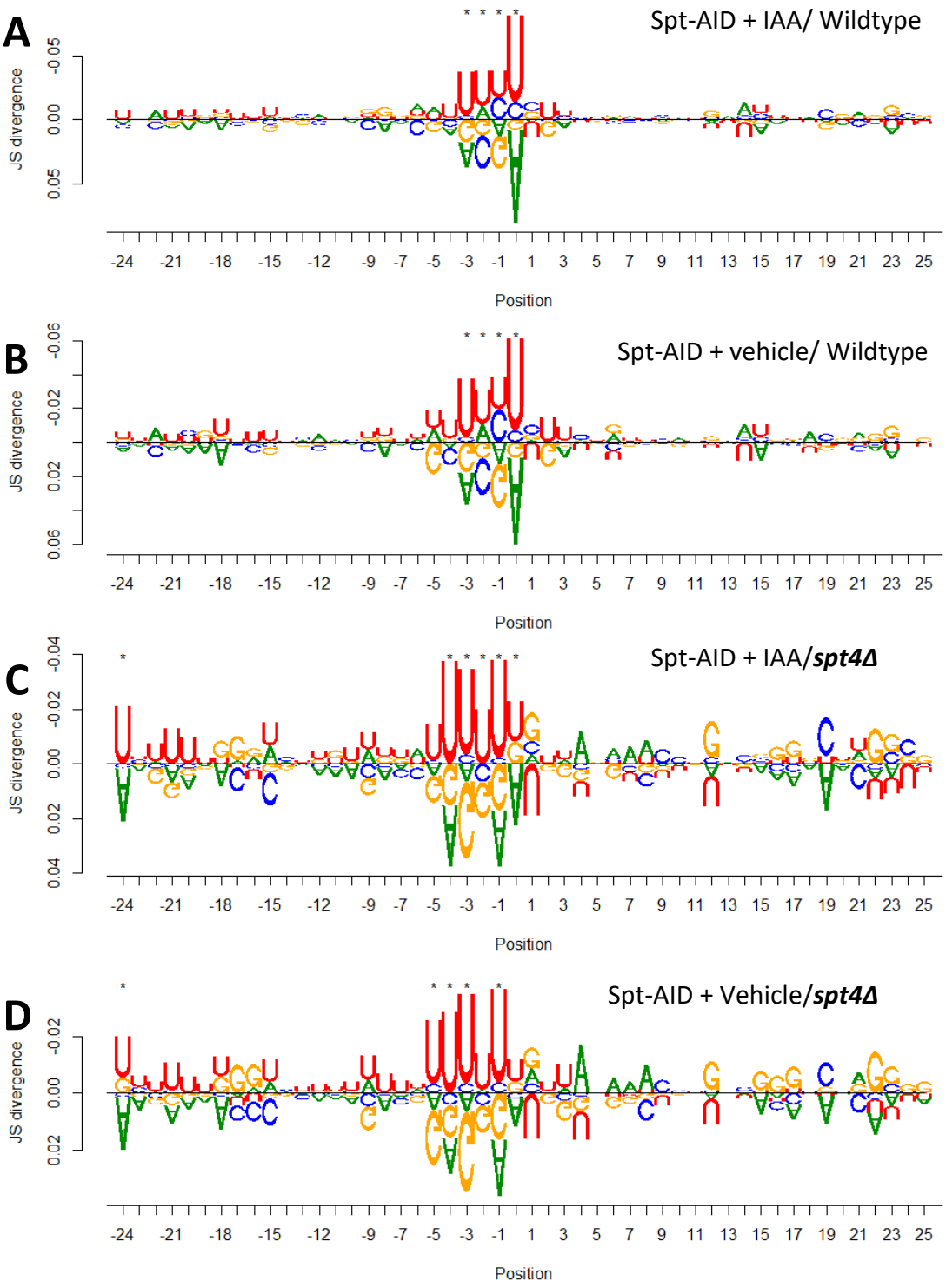

Figure 6: DiffLogos of Spt5-AID and previously published NET-seq data. DiffLogos were generated by comparing Spt5-AID + IAA and Spt5-AID + vehicle with wildtype and ***spt4Δ*** from Huffines et al. In all figures, preferential nucleotides from Spt5-AID are displayed on top of the logo.

CHAPTER 4: DISCUSSION

Interpretation of results: Overall the results seem in conflict. While the ^{3}H-uridine experiments strongly suggest that loss of Spt5 causes a decrease in rRNA synthesis, both the qPCR and the NET-Seq data show very few differences between the auxin and vehicle at the similar time points. However, with the NET-seq data, we are limited to one replicate for each condition and our intended control strain is not similar to wildtype as we had hoped, so any conclusions from NET-seq are still preliminary. Perhaps most concerning is the vast divergence that the vehicle control has from previously published wildtype data. There are a few explanations for these conflicting results varying from technical to biological. The first explanation for the striking differences could be due to the Spt5-AID and previous wildtypes differences in parental strains, however the genotypes between the strains are very similar both originating from the commonly used lab strain w303 with the only genetic differences being genes for the AID system. An alternative explanation could be that the decrease in rRNA synthesis is a result of dysregulation of Pol II transcription and what we observe with the ^{3}H-uridine experiments is simply the result of a decrease in rRNA synthesis due to a lack of proteins required for Pol I transcription. Additionally, the AID's location on Spt5 could also somehow hinder its function. And finally, ethanol is known to effect growing yeast and could potentially be causing the differences from previously published WT. Either way

the conclusion remains the same more controls are necessary if this strain is to be a viable experimental strain.

An additional explanation is that the basal level of degradation of Spt5 is too high. NET-seq is sensitive to pausing of RNA polymerases, which means that even if a small percentage of polymerases manage to transcribe without Spt5 in the vehicle control, they could shift the profile to a more Spt5-less like pausing pattern. It helps to think of transcription like cars on a single lane highway, only a few cars are really needed to cause a traffic jam even if a majority of the cars are working. As this is the case, only a few polymerases without Spt5 compared to those with Spt5 would be needed to significantly alter the pausing landscape. In support of this explanation, even without IAA or ethanol, these cells already grow slower than wildtype yeast, indicating that perhaps a larger portion of polymerases are transcribing without Spt5. It is possible that these cells are right on the threshold of amount of Spt5 needed to function, and the induced degradation just barely moves them past that threshold.

If the lack of change is due to biological reasons and not technical error, it seems in conflict with the ^{3}H-uridine results which show a severe decrease in newly synthesized RNA. One possible explanation for these seemingly conflicting results is that Spt5's main role in Pol I transcription could be in RNA processing. If rRNA processing starts to be severely affected, misprocessed transcripts could be quickly recognized and degraded showing as less RNA synthesis. In fact, there is a certain logic to this from many different angles. First, Spt5's C-terminal region is known to recruit certain transcription factors to the elongation complex[17, 18, 38-41]. Perhaps there are important rRNA processing factors that are recruited by Spt5, and as rRNA get synthesized, the nascent transcripts

miss important processing checkpoints. Additionally, Cryo-EM data positions Spt5's NGN and KOW4-5 regions near the RNA and DNA and have putative roles as exit channels for the nascent RNA and exiting nt-DNA. Spt5's role on processing could be due to important direct interactions on either the DNA or RNA. Poor processing could occur either by hindering important checkpoints that are policed by transcription rate or perhaps by causing or preventing secondary structures immediately after RNA synthesis. Additionally Spt5 could prevent the formation of DNA:RNA hybrid R-loops by keeping RNA and DNA separate. The bacterial homolog of Spt5, NusG, has been implicated in the prevention of R-loops. NusG-deficient strains show a prevalence of R-loops compared to wildtype[42]. Either way, improper folding or processing leading to degradation of defective rRNA could explain both the defects in rRNA synthesis rate as well as the lack of difference in the qPCR and the NET-seq.

It is also important to note that occupancy is only indirect data when it comes to inferring about transcription elongation rate. The polymerase could be moving slowly, while maintaining the same pausing profile, and there would be no way to identify that decrease in rate from NET-seq. However, if that were the case we would expect to see differences in the qPCR results.

Hypothesis: Overall the hypothesis is still inconclusive. Our best suggestion that Spt5 loss affects Pol I processivity are the tritium uridine experiments, however this data is indirect and a much wider picture of RNA synthesis as a whole. Our attempts to directly examine Pol I either by directly probing transcripts produced or by examining occupancy show little variation. More research is needed to make further conclusions.

Future directions: There are many avenues to take with this project. First would be to examine the AID system further and perhaps attempt another deployment of the system. Both systems using AtAFB or OsTIR1^{F74A} as an alternative to OsTIR1 have shown promise at eradicating basal levels of degradation as well as requiring less of the triggering compound[24, 25].

Whether or not the best implementation of the AID system is in place, it becomes necessary to look at rRNA processing more closely. Assays such as ^{3}H methyl methionine pulse-chase or polysome profiling could be used to observe the landscape of both intermediate and mature rRNA products. Using these techniques, we could track the post-processing species of the rRNA present in the degron strain. We would look for less mature rRNA product and an accumulation of intermediate species. If we decide to continue with the same strains used for this thesis project, most likely both the Control + IAA and the Control + vehicle will be needed to draw any stronger conclusions from NET-seq. Despite the uncertainty surrounding the conclusions, the results so far suggest that Spt5 is affecting Pol I transcription in addition to Pol II. This could be one in a many series of steps, that identifies Spt5 as a shared transcription factor across eukaryotic polymerases.

LIST OF REFERENCES

1. Huffines, A.K., Y.J.K. Edwards, and D.A. Schneider, *Spt4 Promotes Pol I Processivity and Transcription Elongation.* Genes (Basel), 2021. **12**(3).

2. Schneider, D.A., et al., *Transcription elongation by RNA polymerase I is linked to efficient rRNA processing and ribosome assembly.* Mol Cell, 2007. **26**(2): p. 217-29.

3. McCoy, L.S., Y. Xie, and Y. Tor, *Antibiotics that target protein synthesis.* Wiley Interdiscip Rev RNA, 2011. **2**(2): p. 209-32.

4. Warner, J.R., *The economics of ribosome biosynthesis in yeast.* Trends Biochem Sci, 1999. **24**(11): p. 437-40.

5. Pelletier, J., G. Thomas, and S. Volarevic, *Ribosome biogenesis in cancer: new players and therapeutic avenues.* Nat Rev Cancer, 2018. **18**(1): p. 51-63.

6. Guo, M., et al., *Core structure of the yeast spt4-spt5 complex: a conserved module for regulation of transcription elongation.* Structure, 2008. **16**(11): p. 1649-58.

7. Swanson, M.S., E.A. Malone, and F. Winston, *SPT5, an essential gene important for normal transcription in Saccharomyces cerevisiae, encodes an acidic nuclear protein with a carboxy-terminal repeat.* Mol Cell Biol, 1991. **11**(8): p. 4286.

8. Viktorovskaya, O.V., F.D. Appling, and D.A. Schneider, *Yeast transcription elongation factor Spt5 associates with RNA polymerase I and RNA polymerase II directly.* J Biol Chem, 2011. **286**(21): p. 18825-33.

9. Schneider, D.A., et al., *RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing.* Proc Natl Acad Sci U S A, 2006. **103**(34): p. 12707-12.

10. Anderson, S.J., et al., *The transcription elongation factor Spt5 influences transcription by RNA polymerase I positively and negatively.* J Biol Chem, 2011. **286**(21): p. 18816-24.

11. Yamaguchi, Y., et al., *Structure and function of the human transcription elongation factor DSIF.* J Biol Chem, 1999. **274**(12): p. 8085-92.

12. Ivanov, D., et al., *Domains in the SPT5 protein that modulate its transcriptional regulatory properties.* Mol Cell Biol, 2000. **20**(9): p. 2970-83.

13. Evrin, C., et al., *Spt5 histone binding activity preserves chromatin during transcription by RNA polymerase II.* EMBO J, 2022. **41**(5): p. e109783.

14. Farnung, L., et al., *Structure of a backtracked hexasomal intermediate of nucleosome transcription.* Mol Cell, 2022. **82**(17): p. 3126-3134 e7.

15. Guo, G., et al., *Structural and biochemical insights into the DNA-binding mode of MjSpt4p:Spt5 complex at the exit tunnel of RNAPII.* J Struct Biol, 2015. **192**(3): p. 418-425.

16. Ehara, H., et al., *Structure of the complete elongation complex of RNA polymerase II with basal factors.* Science, 2017. **357**(6354): p. 921-924.

17. Squazzo, S.L., et al., *The Paf1 complex physically and functionally associates with transcription elongation factors in vivo.* EMBO J, 2002. **21**(7): p. 1764-74.

18. Yamaguchi, Y., et al., *NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation.* Cell, 1999. **97**(1): p. 41-51.

19. Wada, T., et al., *Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro.* EMBO J, 1998. **17**(24): p. 7395-403.

20. Ni, Z., et al., *P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo.* Mol Cell Biol, 2008. **28**(3): p. 1161-70.

21. Nishimura, K., et al., *An auxin-based degron system for the rapid depletion of proteins in nonplant cells.* Nat Methods, 2009. **6**(12): p. 917-22.

22. Xie, J., Y. Jin, and G. Wang, *The role of SCF ubiquitin-ligase complex at the beginning of life.* Reprod Biol Endocrinol, 2019. **17**(1): p. 101.

23. Mendoza-Ochoa, G.I., et al., *A fast and tuneable auxin-inducible degron for depletion of target proteins in budding yeast.* Yeast, 2019. **36**(1): p. 75-81.

24. Nishimura, K., et al., *A super-sensitive auxin-inducible degron system with an engineered auxin-TIR1 pair.* Nucleic Acids Res, 2020. **48**(18): p. e108.

25. Li, S., et al., *An efficient auxin-inducible degron system with low basal degradation in human cells.* Nat Methods, 2019. **16**(9): p. 866-869.

26. Maudlin, I.E. and J.D. Beggs, *Spt5 modulates cotranscriptional spliceosome assembly in Saccharomyces cerevisiae.* RNA, 2019. **25**(10): p. 1298-1310.

27. Gietz, R.D. and R.H. Schiestl, *Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method.* Nat Protoc, 2007. **2**(1): p. 35-7.

28. Clarke, A.M., et al., *NETSeq reveals heterogeneous nucleotide incorporation by RNA polymerase I.* Proc Natl Acad Sci U S A, 2018. **115**(50): p. E11633-E11641.

29. Churchman, L.S. and J.S. Weissman, *Nascent transcript sequencing visualizes transcription at nucleotide resolution.* Nature, 2011. **469**(7330): p. 368-73.

30. Pertea, G., *fqtrim*. 2015.

31. Martin, M., *Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads.* EMBnet Journal, 2011. **17**: p. 10-12.

32. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

33. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.

34. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-2.

35. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.

36. Molder, F., et al., *Sustainable data analysis with Snakemake.* F1000Res, 2021. **10**: p. 33.

37. Henras, A.K., et al., *An overview of pre-ribosomal RNA processing in eukaryotes.* Wiley Interdiscip Rev RNA, 2015. **6**(2): p. 225-42.

38. Aoi, Y., et al., *SPT5 stabilization of promoter-proximal RNA polymerase II.* Mol Cell, 2021. **81**(21): p. 4413-4424 e5.

39. Wen, X., et al., *The transcription factor Spt4-Spt5 complex regulates the expression of ATG8 and ATG41.* Autophagy, 2020. **16**(7): p. 1172-1185.

40. Mayer, A., et al., *The spt5 C-terminal region recruits yeast 3' RNA cleavage factor I.* Mol Cell Biol, 2012. **32**(7): p. 1321-31.

41. Wier, A.D., et al., *Structural basis for Spt5-mediated recruitment of the Paf1 complex to chromatin.* Proc Natl Acad Sci U S A, 2013. **110**(43): p. 17290-5.

42. Leela, J.K., et al., *Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in Escherichia coli.* Proc Natl Acad Sci U S A, 2013. **110**(1): p. 258-63.