2019

DNA Template Sequence Effects on RNA Polymerase I Transcription Elongation

Andrew Martin Clarke
University of Alabama at Birmingham

DNA TEMPLATE SEQUENCE EFFECTS ON RNA POLYMERASE I TRANSCIPTION ELONGATION

by

ANDREW MARTIN CLARKE

DAVID SCHNEIDER, COMMITTEE CHAIR
CHAD PETIT
WILLIAM PLACZEK
MICHAEL CROWLEY
ELLIOT LEFKOWITZ

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, AL

2019

DNA TEMPLATE SEQUENCE EFFECTS ON RNA POLYMERASE I ELONGATION

ANDREW MARTIN CLARKE

BIOCHEMISTRY AND STRUCTURAL BIOLOGY

ABSTRACT

The production of ribosomes represents the major synthetic effort of a rapidly dividing cell, and is intimately linked to the regulation of cell growth and proliferation. Developing a greater understanding of the mechanisms that regulate ribosome biogenesis is therefore crucial to understanding cellular control of the growth cycle. Ribosome biogenesis begins with the synthesis of the 35S ribosomal RNA (rRNA) by RNA polymerase I (Pol I). This RNA is co- and post-transcriptionally processed to produce the 18S, 5.8S, and 25S rRNAs. These RNAs are critical components of ribosomes, and Pol I has been demonstrated to be a key regulation target for the ribosome biogenesis pathway. Most scholarship on the regulation of Pol I has focused on transcription initiation. More recent studies have demonstrated a link between Pol I elongation rate and efficient processing of the rRNA, indicating that Pol I transcription elongation is likely subject to regulation as well.

The effect of DNA template sequence on RNA polymerase elongation has been extensively studied in prokaryotes. By contrast, the role of DNA template sequence in regulating Pol I elongation is significantly understudied. In this thesis, we have attempted to bridge this gap by studying the relationship between Pol I elongation *in vitro* and *in vivo*. First, we demonstrated that prokaryotic *rho*-independent terminator motifs induce arrest and termination in *S. cerevisiae* (yeast) Pol I *in vitro*, indicating that Pol I is sensitive to elongation-affecting DNA template sequence motifs. We next adapted Native

Elongating Transcript Sequencing (NET-seq) to study Pol I occupancy *in vivo*. We identified reproducibly heterogeneous occupancy of Pol I throughout the 35S gene, as well as significant correlation between Pol I occupancy and the last transcribed nucleotide. Finally, we characterized occupancy in *Δrpa12* yeast, a strain with altered Pol I nucleotide addition kinetics. In this mutant we observed changes in Pol I occupancy throughout the 35S gene and identified a putative third site of Pol I transcription termination. Taken together, this work simultaneously establishes DNA template sequence as a regulator of Pol I activity, and validates a technique to better explore this relationship *in vivo*.

Keywords: Molecular Genetics, High-Throughput Sequencing, RNA Polymerase I, Transcription Elongation

ACKNOWLEDGEMENTS

I would like first to thank my family. To my parents, Elizabeth and Christopher, thank you for working so hard to make this possible. In myriad ways, I wouldn't be here if not for you. To my brother, Patrick, thank you for your support, enthusiasm, and optimism. You inspired me to keep at it, even when the going was tough. To my wonderful partner Nicki, thank you for being the wind beneath my wings, and putting up with all of the late nights and working weekends. I'd also like to thank the science faculty at the University of Puget Sound for igniting my passion in Genetics, particularly Dr. Jack Vincent for teaching an especially interesting genetics course in the spring of 2011. This is all your fault.

I'd also like to thank all of the faculty that have served on my committee, especially Dr. Placzek, the only original non-mentor member. Your insightful questions and advice helped to guide my career in more ways than you know. Dr. Crowley, your patience and input made Pol I NET-seq a reality. Dr. Lefkowitz, thank you for joining my committee later on in my training- you've been a wonderfully valuable addition.

Finally, I'd like to thank the members of both of my labs, past and present. Frank, thanks for my first tour at UAB, and for always taking the stairs. Krysta, thank you for getting me hooked on bioinformatics. Catie, thank you for always being the machete to my intellectual thicket. Saman, thank you for being an open ear and a good friend. Abby, thank you for diving in where I left off. I promise, the water's fine. Ruthie, thank you for always being on the gas and never on the brakes. To Kleinpeter and Jureka, thanks for

being truly exemplary colleagues, and for laying breadcrumbs on your way out. I'm right behind you guys. To my mentors, Drs. David Schneider and Chad Petit, you guys have helped me become a better version of myself, both professionally and personally. Because of your advice and guidance, I'm not the same person I was when I walked into this building, and I have the both of you to thank for that.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table *Page*

NETSEQ REVEALS HETEROGENEOUS NUCLEOTIDE INCORPORATION BY RNA POLYMASE I

LIST OF FIGURES

Figure *Page*

NETSEQ REVEALS HETEROGENEOUS NUCLEOTIDE INCORPORATION BY RNA POLYMASE I

LIST OF FIGURES

Figure *Page*

DEFINING THE INFLUENCE OF THE A12 SUBUNIT ON TRANSCRIPTION ELONGATION AND TERMINATION BY RNA POLYMERASE I *IN VIVO*.

INTRODUCTION

Ribosomes

Protein synthesis in all living organisms is performed by ribosomes (reviewed in [1]). In brief, these complexes bind messenger RNA molecules (mRNAs), which contain the primary sequence for proteins, encoded in nucleotide triplets called codons [2, 3]. Upon binding, the ribosome synthesizes the encoded protein one amino acid at a time by adding the amino acid corresponding to each codon to the end of the polypeptide chain [4]. Proper recruitment of the amino acid encoded by each codon is mediated by transfer RNAs (tRNAs). Specific tRNAs bind specific amino acids [5], and have regions of sequence complementarity with their corresponding codons [6].

Ribosomes are large ribonucleoprotein complexes and their general structure and function are conserved throughout all domains of life (reviewed in [7, 8]). Ribosomes consist of a large and small subunit (reviewed in [9]). The small subunit is responsible for matching each codon with the appropriate amino acid-bound tRNA [10], and the large subunit is responsible for adding that amino acid to the polypeptide chain [11]. The large and small subunits in *Saccharomyces cerevisiae* (yeast) consist of a combined four ribosomal RNAs (rRNAs) and 79 different ribosomal proteins (reviewed in [7]). The rRNAs are responsible for the catalytic activity of the ribosome (reviewed in [1]). These components need to be synchronously produced, modified, and assembled. These events are mediated by a further 76 non-coding small nucleolar RNAs (snoRNAs) and more than 200 distinct protein assembly factors (reviewed in [7]).

Ribosome Biogenesis

The production of new ribosomes is a critical activity for all rapidly dividing cells (reviewed in [7, 12]). This process (called ribosome biogenesis) is very expensive, both energetically and materially. Ribosome biogenesis accounts for more than 60% of all transcription in a eukaryotic cell (reviewed in [7]), requiring the function of all three nuclear RNA polymerases (discussed in depth below). Yeast cells must produce 200,000 ribosomes per generation at a rate of more than 2,000 ribosomes per minute [12]. Because this process is the major synthetic effort of any rapidly dividing cell, ribosome biogenesis is also a known checkpoint for cell cycle progression and a determinant of cell size [13]. In yeast, cells must meet a minimum rate of ribosome biogenesis before cell division can proceed [14]. In humans, the impairment of ribosome biogenesis prevents cell cycle progression [15] in a p53-dependent manner [16].

Though mature ribosomes are found almost exclusively in the cytoplasm, ribosome biogenesis largely occurs within the nucleolus, a sub-compartment of the cell nucleus. This process begins with transcription of the 35S gene by RNA Polymerase I (Pol I, reviewed in [17]). The 35S rRNA transcript is cleaved and extensively modified to produce three of the four rRNAs required for ribosome biogenesis: the 18S, 5.8S, and 25S rRNAs. The fourth species, the 5S rRNA, is synthesized by RNA Polymerase III (Pol III) [18]. In addition to cleavage, the rRNAs are subject to base modifications as well as chaperone-mediated folding (reviewed in [7]). Many of the cleavage and modification events required for proper rRNA maturation can occur co-transcriptionally [19]. Both co- and post-transcriptional base modifications and cleavage events are moderated by complexes of proteins and snoRNAs called snoRNPs (small nucleolar

ribonucleoproteins) (reviewed in [7]). snoRNAs contain regions of complementarity to regions of the 35S rRNA, by which snoRNPs bind directly to the nascent 35S rRNA transcript. snoRNPs either mediate binding of larger protein complexes (in the case of cleavage events), or directly catalyze specific processing events (in the case of base modifications [20]). Concurrent with transcript cleavage and base modification, ribosomal proteins bind the rRNAs, stabilizing existing transcript secondary structure or facilitating proper secondary structure formation. These binding events also confer stability on the maturing ribosomal subunits as a whole in prokaryotes and yeast [21, 22] (reviewed in [7]). Most of the rRNA processing and ribosomal protein addition steps occur in the nucleolus. The last steps of ribosome maturation begin in the nucleus, and end in the cytoplasm (reviewed in [23]).

RNA Synthesis

RNA synthesis is performed by a class of large multi-subunit complexes known as RNA polymerases, which are assisted by proteins known as transcription factors. RNA synthesis has three main steps: initiation, elongation, and termination (reviewed in [24]).

Initiation begins with recruitment of the RNA polymerase to a region upstream of the target sequence, known as the promoter. This recruitment is mediated by transcription factors bound to the template DNA. Next, the double stranded DNA is melted to create a transcription bubble, with the template strand of the DNA running through the active site of the polymerase (reviewed in [25], [26]). Beginning at the transcription start site (TSS), the RNA polymerase moves down the DNA template, performing successive nucleotide addition reactions. In brief, nucleotide addition consists of matching the DNA base in its

active center to its cognate RNA monomer (known as a nucleotide) [27, 28], catalyzing a phosphodiester bond between the RNA nucleotide and the 3' end of the growing RNA polymer (referred to above as the nascent transcript), and moving downstream to the next base in the DNA template (reviewed in [29]). The last 8 – 10 base pairs of the RNA transcript hybridize with the DNA template, creating an RNA:DNA hybrid. This hybrid stabilizes the engaged polymerase on the template [30, 31]. During transcription of the first 50 nucleotides, the polymerase is prone to abortive initiation [32, 33]. During abortive initiation, the polymerase remains engaged with the transcription factors that originally recruited it to the promoter (such as σ^{70} for prokaryotic RNA Polymerase (RNAP) and TFIIB for RNA Polymerase II (Pol II)), while "scrunching" downstream DNA to synthesize the first several bases of the transcript [34, 35]. Once the complex is able to break away from these promoter-bound transcription factors, it is considered to be an elongation complex (EC) (reviewed in [36], [37]). The elongation step consists of the EC transcribing the template DNA sequence. The final stage of transcription is termination. In this step, the elongation complex dissociates from the nascent transcript, as well as the DNA template. The transcription bubble closes, and the RNA polymerase is free to begin the transcription cycle again [38-41].

RNA Polymerase Diversity in Eukaryotes

In prokaryotes, all RNA synthesis is performed by RNAP. Eukaryotes (such as yeast) have at least 3 nuclear RNA polymerases [42], Pol I, Pol II, and Pol III. Pol I synthesizes the 35S rRNA, which is processed to produce the 18S, 5.8S, and 25S rRNAs [43]. Pol II synthesizes mRNAs [44] (reviewed in [45]), small nuclear RNAs (snRNAs)

(reviewed in [46]), and micro RNAs (miRNAs) [47]. Pol III synthesizes the 5S rRNA [18], tRNAs and snoRNAs [18] (reviewed in [24]). These three eukaryotic polymerases share ten core protein subunits that are either identical or highly conserved [48]. These core subunits are responsible for nucleotide addition and bear remarkable structural and functional homology with RNAP [49]. Unsurprisingly, the mechanism of nucleotide addition is very similar for all of these polymerases (reviewed in [24]). The source of divergence lies in the additional peripheral subunits found in the eukaryotic polymerases (reviewed in [24]). Each eukaryotic RNA polymerase also has a distinct but overlapping set of transcription factors to suit their more specialized roles (reviewed in [24]).

Interestingly, some activities that are facilitated by transiently associating transcription factors in one polymerase are instead performed by constitutively associated subunits in another. One example is how the different eukaryotic polymerases have evolved to resolve backtracking. Both prokaryotic and eukaryotic RNA polymerases backtrack on the DNA template (reviewed in [50]). This causes the active site of the RNA polymerase to lose contact with the end of the nascent transcript, preventing subsequent nucleotide addition [51, 52]. These backtracking events occur in response to RNA nucleotide mis-incorporation (which facilitates proofreading) [53, 54], or due to collision with DNA-bound proteins [55]. For short backtracking events (less than ten base pairs), polymerases generally rely on random movement along the template to realign with the end of the transcript. However, for longer backtracking events (more than 10 base pairs), the polymerase must cleave the nascent transcript in order to resume productive elongation [56]. RNA polymerases can change the conformation of their active centers to become efficient nucleases [41, 48, 57, 58], but most require the

association of a transcription factor (such as GreA and GreB in bacteria for shorter and longer backtracking events respectively [59-61] and TFIIS in yeast Pol II [58]) to do so. In contrast, this activity is mediated by the A12.2 subunit in Pol I [48] and the C11 subunit in Pol III [62]. As a result, Pols I and III have more robust cleavage activity than Pol II [56, 63].

RNA Polymerase I and the 35S Gene

Pol I is an RNA polymerase composed of 14 subunits [24]. In addition to the ten core subunits found in all eukaryotic polymerases, yeast Pol I contains an additional four subunits, all of which bear partial structural or functional homology with other eukaryotic RNA polymerase subunits or transcription factors. These subunits are present in two heterodimeric complexes: The A14/43 complex (composed of the A14 and A43 subunits) and the A49/34.5 complex (composed of the A49 and A34.5 subunits) [48]. While their roles are not fully elucidated, the A14/43 complex serves as a bridge between Pol I and the initiation factor Rrn3 via the A43 subunit. As a result this complex plays a critical role in the recruitment of Pol I to the rDNA [64]. A14 is thought to serve a structural role, serving as a binding point for both the A14/43 complex and the A49/34.5 complex. The A49/34.5 complex bears structural similarities to Pol II transcription factors TFIIF and TFIIE [65], both of which facilitate initiation of transcription by Pol II [66]. Cumulatively, the A49/34.5 complex is assumed to stimulate Pol I transcription elongation [48]. Deletion of either the A49 subunit or the A49/34.5 complex has been shown to decrease Pol I elongation rate *in vitro* and reduce yeast growth rate *in vivo* [48, 67]. While the A34.5 subunit deletion does not affect yeast growth under normal

conditions, it has been shown to interact genetically with DNA topoisomerase I and deletion of this subunit is lethal in combination with the deletion of DNA topoisomerase I [68]. The A49/34.5 complex, in concert with the A12.2 subunit, increases Pol I processivity (the number of nucleotide addition reactions performed by an elongation complex per initiation event), elongation rate, and fidelity in a synergistic manner *in vitro* [69]. Of the four Pol I-specific subunits, only A43 is essential. The A12.2 subunit is considered to be one of the ten core subunits of Pol I because it's N-terminal domain is homologous to RPB9 in Pol II. As a result, it was originally assumed to play a primarily structural role for Pol I [70]. However, the C-terminal domain of this protein is homologous to Pol II cleavage-stimulating factor TFIIS [71] and was later confirmed to be crucial for Pol I transcription termination [72] and transcript cleavage [48].

Many of the transcription factors that facilitate initiation complex formation, increase elongation rate, and allow for clearance of transcriptional blocks in the form of backtracking in Pol II are instead integrated as subunits in Pol I. As a result, Pol I transcribes faster than Pol II, pauses less often, and recovers more quickly from backtracking events [56]. These adaptations allow Pol I to better fulfill its specific transcriptional burden, which will be discussed below.

Pols III and II have hundreds and thousands of different transcriptional loci, respectively. In contrast, Pol I has a single transcriptional target- the 35S gene. Yeast has approximately 200 copies of this gene, arranged tandemly on chromosome XII (reviewed in [7]). Each 35S gene is separated from the next by an intergenic spacer that contains the 5S rRNA gene, which is transcribed by Pol III (reviewed in [7]). In rapidly dividing yeast cells, approximately 50% of these repeats are actively transcribed, and each repeat can be

host to 50 or more Pol I ECs [73]. Pol I loading at these genes is the primary determinant of rRNA synthesis volume, as opposed to the number of active repeats [73]. Pol I RNA synthesis alone accounts for more than 60% of all RNA in a yeast cell [7]. In this context, the increased processivity and ability to quickly clear backtracking observed in Pol I are likely important for maintaining steady production of rRNA.

Regulation of Ribosome Biogenesis via RNA Polymerase I

Ribosome biogenesis represents the major biosynthetic effort of rapidly dividing cells. The four rRNAs, 79 ribosomal proteins, 76 snoRNAs, and more than 200 protein assembly factors need to be synthesized, modified, and assembled in the appropriate amounts, at the appropriate times, thousands of times per minute [7, 12]. The 137 genes encoding the 79 ribosomal proteins are among the most actively transcribed by Pol II [74], though the half-lives of the corresponding mRNAs are very short [75]. Transcription of these genes is sensitive to nutrient availability [76, 77] and temperature fluctuation [78].

There is significant evidence of communication between regulatory mechanisms for the different stages of ribosome biogenesis, as failure at any one step results in down-regulation of other steps in the process. For example, decreases in ribosomal protein gene transcription or translation also decrease rRNA synthesis [79]. Additionally, defects in nuclear transport that trap nascent ribosomes inside the nucleus result in a decrease in ribosomal protein mRNA synthesis [80] and rRNA synthesis [81]. As expected, assembly of these components is hierarchical [82] and defects at early steps of processing or assembly negatively affect ribosome biogenesis as a whole [83, 84]. Because ribosome

biogenesis requires the cooperation of so many different cellular processes, efficient regulation of each step in the larger context of this pathway is necessary for efficient ribosome production.

Once generated, mature ribosomes have remarkably long half-lives (in the hundreds of hours for eukaryotes) [85], so the high rate of ribosome biogenesis in these rapidly dividing cells is primarily to ensure each daughter cell has sufficient ribosomes to function, as opposed to replenishing those lost to turnover. Ribosome biogenesis is greatly reduced in cells that do not anticipate significant further proliferation [86]. As a result, cells must also manage overall ribosome biogenesis as a factor of cell cycle progression, as cell division frequency fluctuates.

Understanding the regulation of ribosome biogenesis in yeast has yielded clinical insight, due to the conservation of this process in humans- analysis of homologous or analogous ribosomal proteins in yeast has yielded valuable insight on a host of ribosomopathies in humans, such as Diamond Blackfan Anemia [87] and Treacher Collins Syndrome [88]. Many cancers are partially products of dis-regulated ribosome biogenesis, as many known oncogenes and tumor suppressors are ribosome biogenesis regulators. For example, c-Myc is a well-known regulator of ribosome biogenesis, in addition to a number of other cellular processes [89]. It is also one of the most commonly over-expressed genes found in human cancer cells [90, 91]. Because cancer cells have a characteristic “addiction” to ribosomes (reviewed in [92]) due to unconstrained cell growth and division, ribosome biogenesis has become a promising target for cancer treatment (reviewed in [93, 94]).

Previous studies have identified Pol I transcription [95-97] and rRNA modification [98] as key targets for the regulation of ribosome biogenesis in yeast and higher eukaryotes. Historically, most studies on Pol I transcription regulation have focused on the factors that govern Pol I recruitment to the rDNA [73] as well as the activation or deactivation of 35S gene copies [99]. Recent scholarship has identified regulatory mechanisms for Pol I transcription steps beyond pre-initiation complex assembly, including promoter escape [100] and elongation [101-105]. Furthermore, Pol I elongation rate is linked to efficient rRNA processing and ribosome biogenesis in yeast [106] and *E. coli* [107].

One potential explanation for the observed relationship between Pol I elongation rate and ribosome processing is altered nascent transcript secondary structure due to fluctuations in RNA polymerase elongation rate. This phenomenon has been observed in Pol II transcription of histone genes in higher eukaryotes. The 3' end of histone transcripts contain a conserved stem loop [108], which is recognized by the Stem-Loop Binding Protein (SLBP). Transcript cleavage facilitated by SLBP binding is required for efficient translation of the histone mRNA [109]. When Pol II elongation rate is lowered by mutation or UV irradiation, the stem loop fails to form, SLBP cannot bind the transcript, the transcript is not properly cleaved, and the transcript is not translated [110]. Furthermore in *E. coli*, efficient folding of nascent transcripts can be linked to specific pauses in RNAP transcription elongation [111]. In the case of Pol I, alterations in 35S nascent transcript secondary structure due to reduced elongation rate could render snoRNP binding sites inaccessible, thus preventing crucial modification steps in rRNA

maturation. Understanding the factors that mediate Pol I elongation rate is therefore critical to fully understanding rRNA processing and ribosome biogenesis.

One factor that could affect Pol I elongation is the sequence of the template DNA. DNA template sequence is known to be a direct modulator of polymerase activity in prokaryotic systems [111-113]. Specific DNA template sequences have been shown to affect *E. coli* RNA Polymerase (RNAP) transcription in order to facilitate formation of proper nascent transcript secondary structure [111], synchronize co-transcriptional translation [114], and even terminate transcription in a largely factor-independent manner [113]. In *B. subtilis,* efficient secondary structure formation of the Flavin mononucleotide (FMN)-dependent riboswitch in the *ribDEAHT* operon is dependent on RNAP transcriptional pausing at two specific sites [115]. Proper folding of the RNAse P RNA in *B. subtilis* and *E. coli* is also dependent on RNAP transcriptional pausing at discrete sites *in vitro* [111, 116]. The relationship between Pol I elongation rate, pausing and efficient synthesis of rRNA has never been adequately explored; however, the mature rRNAs possess extensive secondary structure [7], which is frequently altered throughout the maturation process through ribosomal protein [21] and snoRNP binding [117]. We have also previously observed links between Pol I elongation rate and rRNA processing and assembly [103, 105, 106], suggesting that transcription elongation rate (and possibly transcriptional pausing) play an important role in production of functional RNAs by eukaryotic RNA Polymerase I just as they do for structurally and functionally conserved [24] prokaryotic RNA polymerases.

This dissertation will seek to address the gap in our understanding of the relationship between DNA template sequence and Pol I elongation. In the first chapter,

we demonstrate that Pol I is sensitive to specific elongation-affecting DNA template sequence motifs *in vitro*. We then determine that Pol I occupancy *in vivo* is reproducibly heterogeneous and partially dependent on DNA template sequence. In the second chapter, we show that deletion of the A12.2 subunit, which has previously been shown to affect nucleotide addition [118], alters Pol I's occupancy throughout the 35S gene in a DNA template-sequence dependent manner.

NETSEQ REVEALS HETEROGENEOUS NUCLEOTIDE INCORPORATION BY RNA POLYMERASE I

by

ANDREW M. CLARKE, KRYSTA L. ENGEL, KEITH E. GILES, CHAD M. PETIT, AND DAVID A. SCHNEIDER

Proceedings of the National Academy of Sciences. (2018). 115(50) E11633-E11641

Format Adapted for Dissertation

ABSTRACT

DNA sequence motifs that affect RNA polymerase transcription elongation are well-studied in prokaryotic organisms and contribute directly to regulation of gene expression. Despite significant work on the regulation of eukaryotic transcription, the effect of DNA template sequence on RNA Polymerase I (Pol I) transcription elongation remains unknown. In this study, we examined the effects of DNA sequence motifs on Pol I transcription elongation kinetics in vitro and in vivo. Specifically, we characterized how the spy rho-independent terminator motif from Escherichia coli directly affects Saccharomyces cerevisiae Pol I activity, demonstrating evolutionary conservation of sequence-specific effects on transcription. The insight gained from this analysis led to the identification of a homologous sequence in the ribosomal DNA of Saccharomyces cerevisiae. We then used Native Elongating Transcript Sequencing (NETSeq) to determine whether Pol I encounters pause-inducing sequences in vivo. We found hundreds of positions within the ribosomal DNA (rDNA) that reproducibly induce pausing in vivo. We also observed significantly lower Pol I occupancy at G residues in the rDNA, independent of other sequence context, indicating differential nucleotide incorporation rates for Pol I in vivo. These data demonstrate that DNA template sequence elements directly influence Pol I transcription elongation. Furthermore, we have developed the necessary experimental and analytical methods to investigate these perturbations in living cells going forward.

SIGNIFICANCE STATEMENT

It is well known that ribosomal RNA processing is directly impacted by the rate of transcription elongation by RNA polymerase I (Pol I). To understand how these

processes are orchestrated, we must carefully define transcription elongation properties in vitro and in living cells. Here, we characterize DNA sequence elements that pause and terminate Pol I transcription in vitro. We also establish methods for analyzing Pol I transcription elongation properties in vivo using Native Elongating Transcript Sequencing (NETSeq). Our NETseq data revealed frequent pausing by Pol I and decreased Pol I occupancy at G residues suggesting unequal rates of nucleotide incorporation by the enzyme. These findings redefine our understanding of Pol I transcription elongation and its heterogeneity in vivo.

INTRODUCTION

The production of a single ribosome requires the synthesis, processing, and assembly of more than 80 proteins and four non-coding RNAs known as ribosomal RNAs (rRNAs)[1]. Because this process represents the major energetic investment for a rapidly dividing eukaryotic cell, it is subject to tight regulation. In *S. cerevisiae*, the first step in ribosome biogenesis is transcription of the 35S rRNA gene (also referred to as 37S rRNA gene or *RDN37*) by RNA polymerase I (Pol I)[2]. This transcript is co- and post-transcriptionally processed to produce three of the four RNAs required for ribosome biogenesis- the 18S, 5.8S, and 25S rRNAs [3, 4]. Regulation of Pol I activity represents a robust method used by cells to control the rate of ribosome biogenesis. Previous studies concerning the regulation of transcription by Pol I have principally focused on transcription initiation. It is well-established that recruitment of Pol I to the rDNA promoter is a key regulatory target for cellular control of ribosome biosynthesis [5]. However, we and others have shown that later steps in the transcription cycle can be

influenced by transcription factors [6-9]. Furthermore, the efficiency of transcription elongation directly affects processing of the nascent rRNA [10]. Thus, trans-acting factors or template sequence features that influence transcription elongation by Pol I can have substantial consequences on cellular proliferation by affecting the synthesis or processing of rRNA.

While the effect of DNA template sequence on Pol I transcription has not been established, elongation-affecting sequence motifs have been identified as regulatory elements in prokaryotic transcription systems [11, 12]. One such motif is the rho-independent terminator motif. This motif is composed of a guanine-cytosine (G-C) rich region of dyadic symmetry, followed by a thymine-rich tract (T-tract). Transcription of the T-tract creates active site instability induced by a weak RNA-DNA hybrid [12, 13]. This weakened complex is further perturbed by the formation of an RNA hairpin in the upstream G-C rich tract of the nascent transcript. This combined effect results in efficient termination of transcription by prokaryotic RNA polymerase (RNAP) in a protein factor-independent manner [13]. Furthermore, similar motifs that induce termination of eukaryotic RNA polymerase II have also been found in viral and mammalian genes [14-16]. Because the rho-independent terminator motif affects the active site of RNAP which is conserved among all multi-subunit RNA polymerases [17], we reasoned that this motif will also affect Pol I.

In this study, we used a fully reconstituted promoter-dependent transcription assay to determine if DNA template sequence can affect Pol I transcription. The sequence motif that we used was the rho-independent terminator motif from the *spy* gene of *E. coli*. We demonstrate that this motif induces termination of Pol I transcription *in vitro*. We then

used mutational analysis of this motif to show that both the stem loop and uridine tract (U-tract) contribute to its effect on Pol I transcription. Based on these observations, we identified sequence elements present in the native 35S rRNA gene that also influence Pol I transcription elongation kinetics and demonstrated that the effects of both motifs on Pol I transcription are dependent on UTP concentration in our *in vitro* system.

To determine whether Pol I pauses *in vivo* and whether the same sequence elements contribute to this pausing, we adapted Native Elongating Transcript Sequencing (NETSeq) for use with *S. cerevisiae* Pol I. This application of NETSeq enables precise mapping of Pol I occupancy of the ribosomal DNA gene (rDNA) with single nucleotide resolution. We observed reproducibly heterogeneous occupancy of Pol I on the rDNA in rapidly growing cells. Thorough analysis of the resulting Pol I occupancy data revealed significantly lower occupancy of Pol I at G residues throughout the rDNA. Taken together, these data lead to the following conclusions: 1) The effects of strong rho-independent terminator sequences on RNA polymerase activity are conserved across domains of life and 2) Pol I elongation efficiency is sensitive to both the DNA sequence and other factors *in vivo* and *in vitro*.

RESULTS

A Rho-Independent Terminator Motif Affects Pol I Transcription Elongation

To test whether prokaryotic termination motifs can affect Pol I transcription elongation, we used a fully reconstituted promoter-dependent *in vitro* transcription assay [18]. Individual sequence motifs (e.g. the *spy* motif) were inserted into native rDNA sequence downstream of the transcription start site (Fig 1A). This motif was chosen

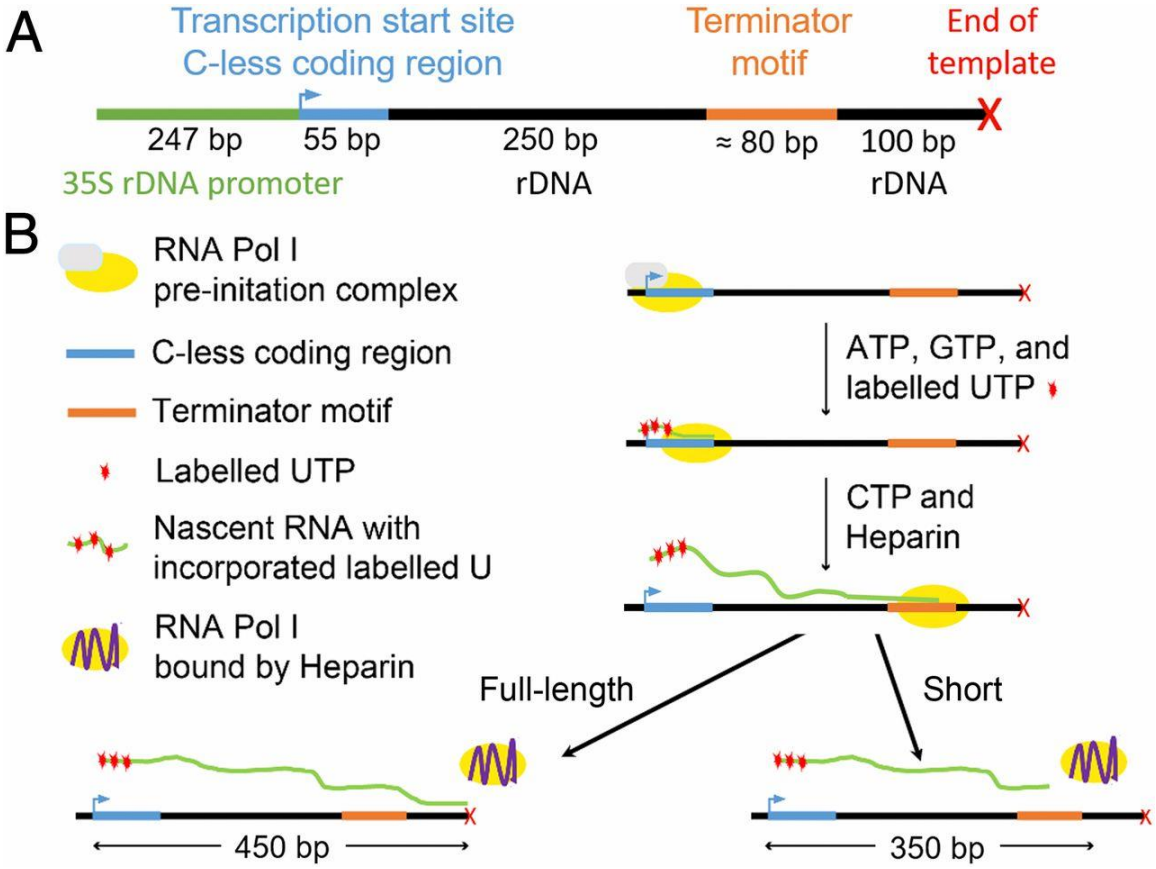

Figure 1. Promoter-dependent *in vitro* transcription assay for RNA polymerase I. 1a) linearized DNA template for promoter-dependent *in vitro* transcription by RNA polymerase I. 1b) Experimental scheme for RNA polymerase I promoter-dependent *in vitro* transcription assay.

because it was identified as the strongest terminator of *E. coli* RNA polymerase (RNAP) transcription in a recent survey [19]. To this template we added Pol I, Rrn3p, TBP, and Core Factor (Rrn6p, Rrn7p, and Rrn11p) to assemble the Pol I pre-initiation complex at the promoter. Transcription was initiated by the addition of ATP, GTP, UTP (15 μM of each), and ^{32}P-labeled UTP, resulting in synchronized transcription elongation complexes at the first encoded C (position +56). The synchronized enzymes were then released by addition of 15 μM CTP, and samples were collected as a function of time. Resultant RNA transcripts were resolved via PAGE and visualized by phosphorimage analysis. If a fraction of the Pol I population paused, arrested, or terminated at the *spy* motif, a product approximately 100 nucleotides shorter than the full length product was observed (Fig 1B). To quantify the magnitude of the effect of the motif on Pol I transcription elongation, we calculated the ratio of the shortened product to the total RNA signal in each individual lane.

Inclusion of the *spy* terminator motif in the template resulted in the appearance of a short product (Fig 2A and 2B) not seen in the negative control (Fig 2B). Since accumulation of this short product is not rescued after an extended incubation time of 20 minutes, we conclude that these complexes have either terminated transcription or are terminally arrested. These data indicate that the *spy* terminator motif directly perturbs transcription elongation by Pol I.

To define the kinetics of the *spy* effect on Pol I, we measured reaction progress during shorter time courses. The accumulation of short product reaches a maximum of 83% +/- 6.7% of total RNA product at the 30 second time point (Fig 2C). The accumulation of short product then decreases at each time point before stabilizing at 30

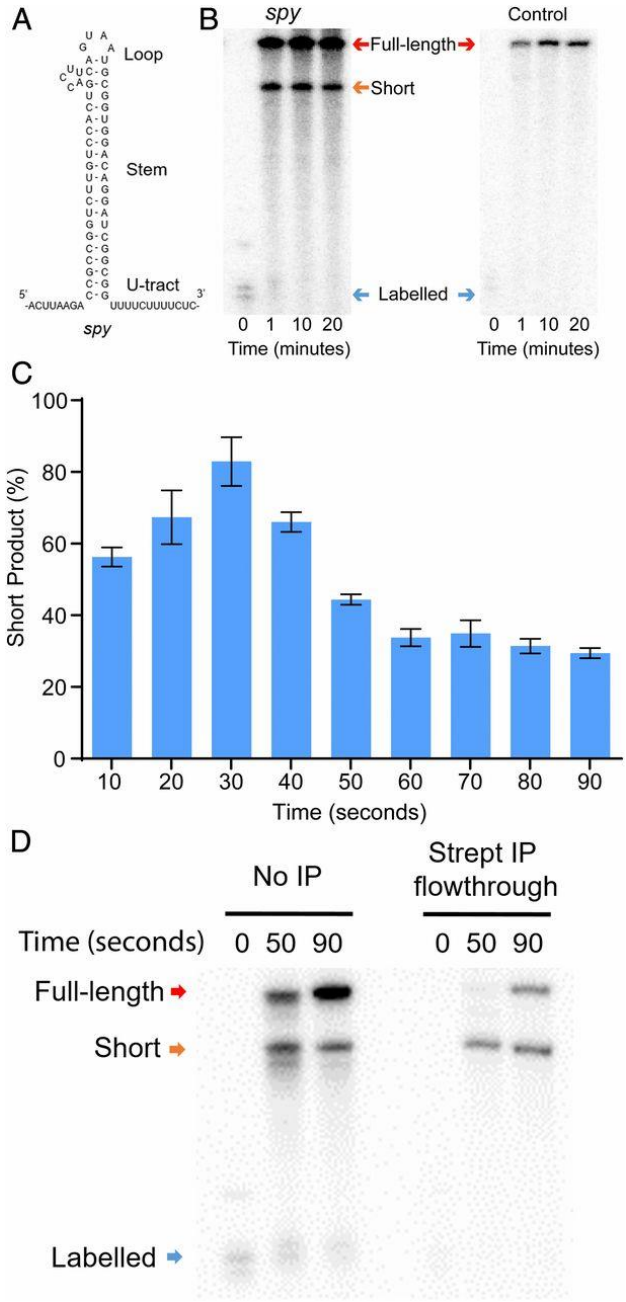

Figure 2. The *spy Rho*-Independent Terminator Motif Affects RNA Polymerase I Transcription *In Vitro*. 2a) Proposed secondary structure of transcribed *spy rho*-independent terminator motif. 2b) polyacrylamide gel of long time course *in vitro* transcription of the *spy rho*-independent terminator template by RNA Polymerase I (left panel), compared to the negative control (right panel). 2c) The short product at each time point of short time course expressed as a percentage of total signal. N=3, error bars represent standard deviation. 2d) polyacrylamide gel of time course *in vitro* transcription of the biotinylated *spy rho*-independent terminator template followed by streptavidin immunoprecipitation comparing no IP (left portion) to the IP flowthrough (right portion.)

+/- 1.6% of total RNA product after 60 seconds (Fig 2C). In order to determine if the observed effect is termination or arrest, the experiment was repeated with a biotinylated template (Fig 2D). Following immunoprecipitation of the template, the short product was observed in the flow-through while the labelled product was not. These data demonstrate that the nascent transcript is not associated with the DNA template, indicating that the motif induces termination of Pol I transcription. Based on these data, it is clear that the *spy* terminator motif induces pausing by the majority of elongation complexes. Most enzymes ultimately escape this pause and reach the end of the template, while a subset undergo termination. Thus, *rho*-independent termination motifs can both pause and terminate transcription by eukaryotic Pol I.

Multiple Regions of the spy Motif Contribute to the Effect on Pol I

To identify the features of the *spy* terminator motif that are necessary for its effect on Pol I transcription, we constructed several variants of the motif. We then repeated the transcription assay described above using these variants. In the first variant, we shortened the stem loop to ten base pairs from twenty-one (Fig 3A). Compared to the wild type *spy* terminator motif, the shortened stem loop variant accumulates substantially less short product at every time point (Fig 3A). By 90 seconds, the percentage of short product is 16 +/- 2.0% compared to 30 +/- 1.6% for wild type. For the second variant, we mutated the template strand to abolish the U-tract in the nascent transcript by exchanging each T residue with a C residue and vice versa (Fig 3B). The motivation for this mutation strategy is based on the finding that uridine enrichment in the residues directly downstream of the hairpin plays an important role in the effect of the motif on

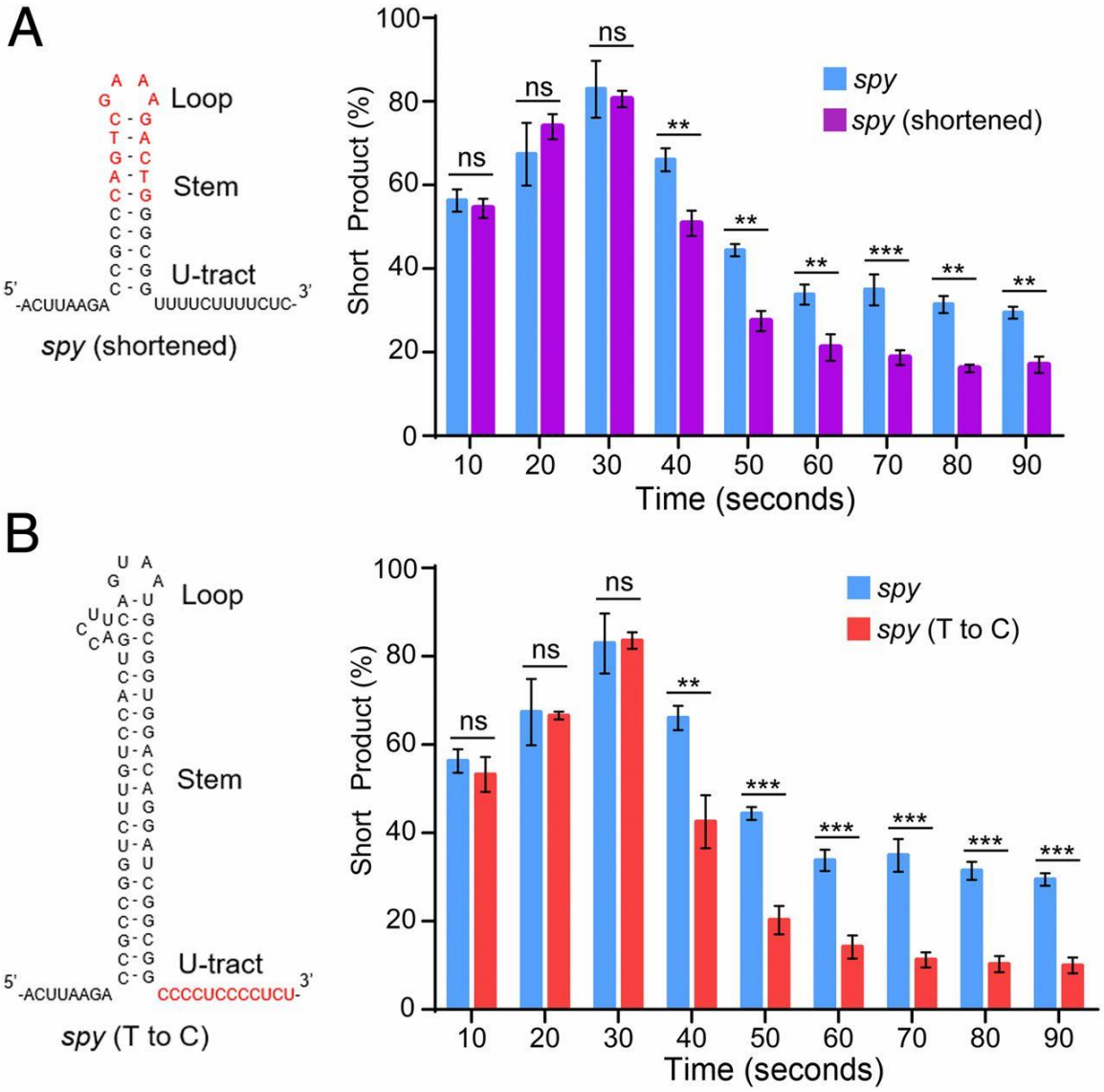

Figure 3. Both the Stem Loop and U-tract Regions of the *Spy Rho*-Independent Terminator Contribute to its Effect on RNA Polymerase I Transcription. 3a) Predicted secondary structure of the shortened stem loop *spy rho*-independent terminator motif mutant (left panel) and the mutant's short product compared to wild type (right panel). 3b) Predicted secondary structure of the T to C Switch *spy rho*-independent terminator motif mutant (left panel), and the mutant's short product compared to wild type (right panel). For 3a and 3b, N = 3, error bars represent standard deviation. Comparison by 1-tailed Student's T-test. ns = not significant, * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$

polymerase elongation [20]. These mutations stabilize the RNA-DNA complex by replacing the A-U base pairs with stronger G-C base pairs which has been previously shown to reduce pausing by prokaryotic RNAP [13, 19, 21]. By 90 seconds, the percentage of short product is just 10 +/- 1.8% compared to 30 +/- 1.6% for wild type (Fig 3B). Taken together, these data demonstrate that both the stem loop length and the U-tract are critical for the effect on Pol I transcription. However, abolishing the U-tract results in a larger decrease in termination when compared to the shortened stem loop variant. These data are consistent with prokaryotic studies showing that both of these features are critical for pausing and terminating RNA polymerases [19]. Furthermore, these findings provide guidance for predicting effects of native ribosomal DNA sequence on transcription elongation by Pol I.

Native rDNA Sequence Elements affect Pol I Transcription Elongation.

To determine if endogenous elongation-affecting motifs exist, we examined the 35S gene region of the rDNA for T-rich tracts (which encode U-tracts in the nascent transcript), which is one of the key elements of the *spy* terminator motif. We identified 54 tracts that are at least10 nucleotides long, end with three T residues, and have a T residue enrichment of 70% or more (Fig 4A). The longest element observed was in the 25S rRNA coding region of the 35S rDNA gene. We inserted this T-rich tract including 90 nucleotides upstream and 15 nucleotides downstream into our template and transcribed the template *in vitro* using the system outlined above. We observed a transient pause of a majority of the polymerases at the rDNA sequence 30 seconds after release. This effect is similar to the effect of the *spy* terminator motif. Most of these polymerases clear the

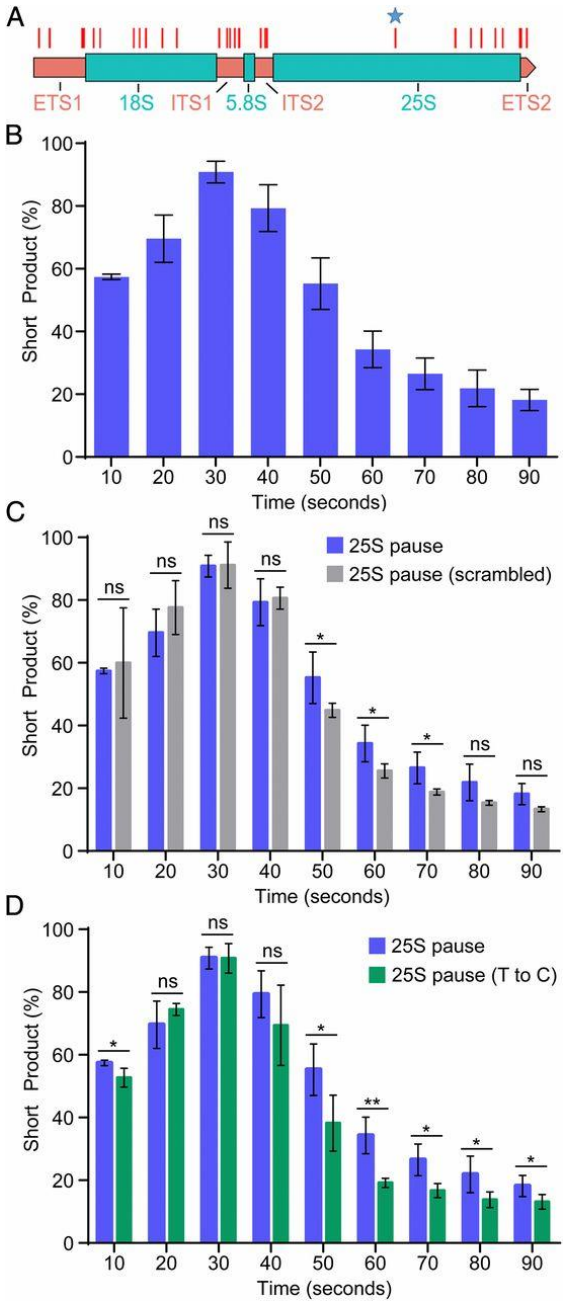

Figure 4. Endogenous rDNA Sequence Induces Site-Specific Pausing, with Contributions from the T tract and the Upstream Sequence. 4a) T-tracts (10 nucleotides or greater in length, greater than 70% T enrichment, ending with three T residue) in the 35S gene (marked with red lines.) The blue star marks the T-tract to be analyzed. 4b) rDNA Sequence Wild Type short product at each time point expressed as a percentage of total signal. 4c) rDNA Sequence Upstream Scramble product compared to wild type. 4d) rDNA Sequence T to C Switch short product compared to wild type.N = 3, error bars represent standard deviation. Comparison by 1-sided Student's T-test. ns = not significant, * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$

pause and produce full-length product with the percentage of short product stabilizing at 15 +/- 2.2% by 90 seconds post-release (Fig 4B). These data demonstrate that sequences within the rDNA inhibit transcription elongation by Pol I.

We then produced two variants to determine which parts of this sequence were necessary for its effect on Pol I transcription. In the first variant, we abolished the possibility of RNA stem loop formation by scrambling the 30 nucleotides upstream of the T-tract. This perturbation had relatively little effect on short product accumulation (Fig 4C). In the second variant, we abolished the T-tract by exchanging the T residues for C residues in the DNA template. Abolishing the T-tract decreased short product accumulation at almost every point (Fig 4D). Thus, the T-tract is required for the motif's pausing and termination effects on Pol I transcription. Taken together, these data demonstrate that T-tracts within the rDNA can directly influence Pol I activity. This observation is consistent with conclusions drawn using prokaryotic DNA elements.

Pausing by Pol I at T-tracts is sensitive to UTP concentration.

Since we have determined that the T-tract is critical for the effects on Pol I transcription, we tested whether the UTP concentration in the reaction alters the kinetics of Pol I elongation. We repeated *in vitro* transcription assays under two different UTP concentrations: 100 μM UTP and 15 μM UTP. For the *spy* terminator motif, there is a significant difference between the amount of short product in the 100 μM UTP and 15 μM UTP experiments. This difference is observed starting at 30 seconds and continues through 90 seconds after which both values stabilize (Fig 5A). At 15 μM UTP, short product accounted for 30 +/- 1.6% of all RNA by 90 seconds. At 100 μM UTP, this value

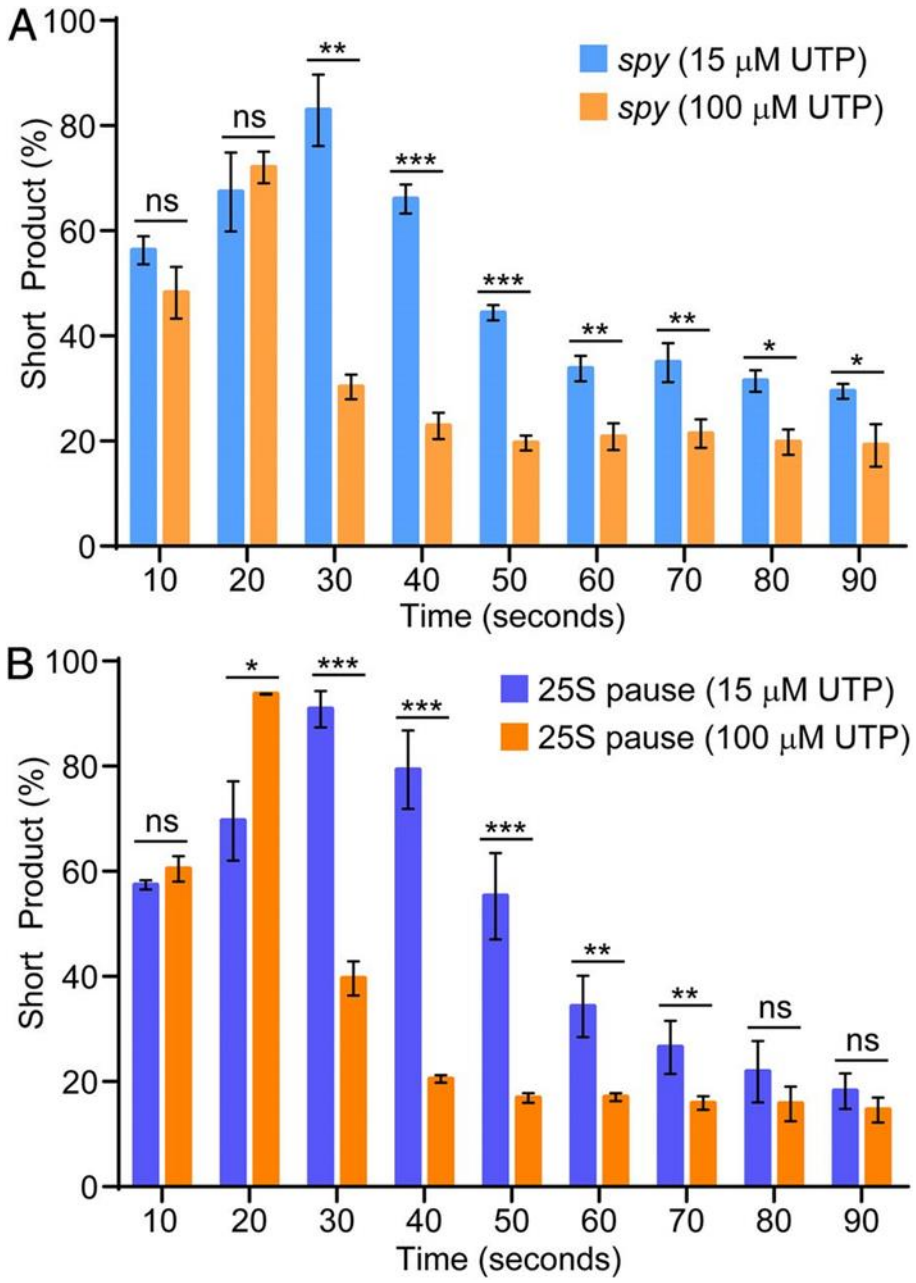

Figure 5. The effect of Both Motifs on RNA Polymerase I Transcription is UTP Concentration-Dependent. 5a) terminator wild type short product at 15 μM and 100 μM UTP. 5b) rDNA sequence wild type short product at 15 μM and 100 μM UTP. N = 3, error bars represent standard deviation. Comparison by 1-sided Student's T-test. ns = not significant, * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$

decreased to 19 +/- 4%. These data demonstrate that elevated UTP concentration suppresses, but does not eliminate, the observed pause and termination effects of the *spy* terminator motif on Pol I.

To determine if the eukaryotic sequence also responds to substrate concentration, we repeated these experiments using the T-rich 25S rDNA pause sequence detailed above. We found that, like the *spy* motif, there is a significant reduction in observed pausing at the T-tract in the presence of 100 μM UTP compared to 15 μM UTP (Fig 5B). Although the percentage of terminated polymerases does not change, the pause dwell time is reduced as UTP concentration increases. All of these *in vitro* analyses call for analysis of pausing by Pol I *in vivo*.

NETSeq reveals heterogeneous Pol I Occupancy on rDNA

To characterize Pol I pausing *in vivo*, we performed Native Elongating Transcript Sequencing (NETSeq). We adapted our methods from those used previously to characterize RNA polymerase II (see SI Appendix) [22]. Three biological replicates of NETSeq were performed with a *S. cerevisiae* strain bearing HA-tagged Pol I, and the resultant reads were mapped to the rDNA gene (Fig 6A). The 3' end of each read corresponds to the last phosphodiester bond formed by the polymerase (Figs 6A and 6B). Our data demonstrate obvious heterogeneity in Pol I occupancy throughout the gene. The amplitude of each peak reflects the frequency with which a Pol I complex is observed at that position *in vivo.* Thus, high peak height can be interpreted as a position that induces pausing by Pol I.

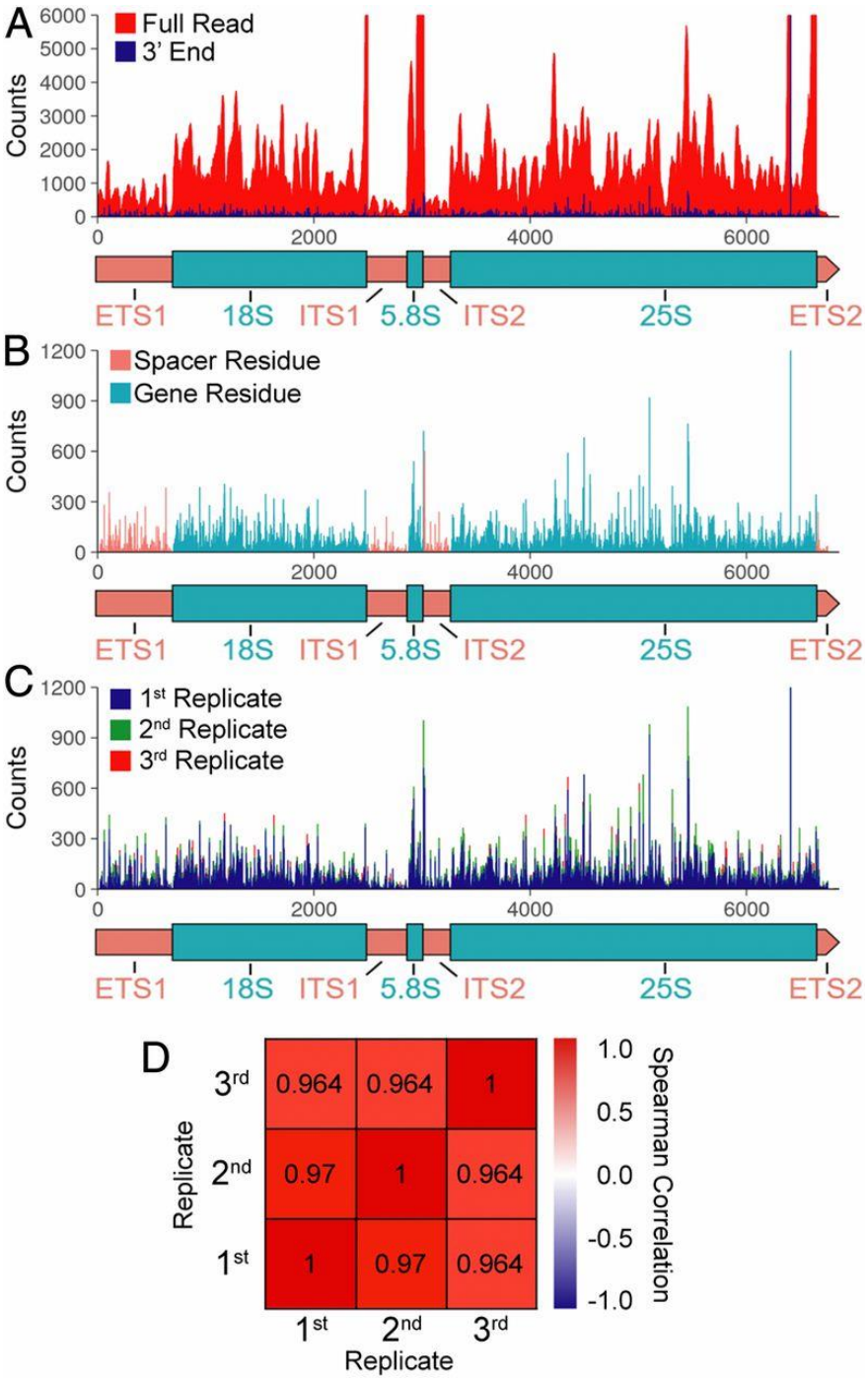

Figure 6. RNA Polymerase I NETSeq Reads Map to rDNA, and are Qualitatively Reproducible. 6a) NETSeq Pol I full read density in the 35S gene (red) compared to 3' end density (blue,) with 35S gene diagram below. 6b) NETSeq rDNA 3' end density, with 35S gene diagram below. Spacer residues highlighted in pink, gene residues highlighted in cyan. 6c) NETSeq rDNA 3' end densities from 3 biological replicates overlaid (blue, green, red,) with 35S gene diagram below. 6d) Heat map of spearman correlations between replicate position lists ranked by occupancy. All p values < $1*10^{-8}$.

The overlay of the three biological replicates revealed exceptional reproducibility between cultures (Fig 6C, 6D). In all three replicates, significant sequence coverage was observed only in rDNA genes and the mature ends of snoRNAs, short RNAs associated with nascent rRNA transcripts. In addition to the 3'-nascent RNAs detected throughout the gene, we observed large signal spikes at the position of the mature ends of all four rRNA species (25S, 5.8S, 18S, and 5S genes, SI Appendix, Fig S3A). It is well established that mature ribosomal RNA is recovered when preparing RNA sequencing libraries. Furthermore, the 5S gene is not transcribed by Pol I, yet it was observed in all of our libraries (SI Appendix, Fig S4B). Thus, we concluded that these four large peaks represented contaminating mature product, not nascent RNA, and we eliminated these peaks from subsequent analyses and plots.

Initial analysis revealed significantly lower Pol I occupancy within the spacer regions of the rDNA compared to the regions that encode mature rRNA (Fig 7). There are a number of potential explanations for this observation. The simplest interpretation is that Pol I may elongate more quickly through those regions of the template, particularly the spacer between the 18S and 5.8S regions termed "ITS1" (Fig 7A). It is known that mature rRNA folds into stable secondary and tertiary structures whereas spacer elements are less well conserved and likely less structured. If rRNA structures begin to form on the nascent RNA, these structures might slow down Pol I. This potential effect would be exacerbated in the sequences that give rise to mature product. It is also possible that there is contamination of our nascent RNA with degradation products of mature rRNA (see discussion below).

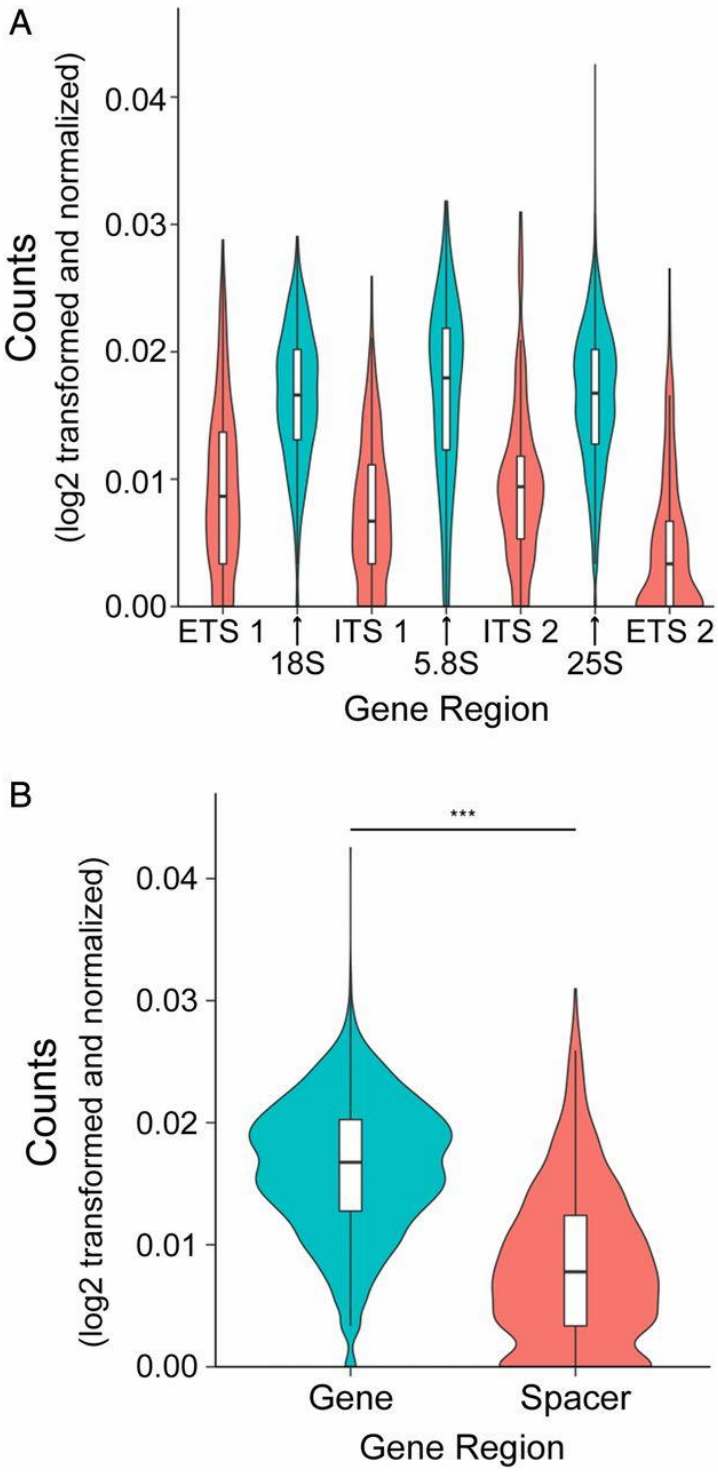

Figure 7. RNA Polymerase I Occupancy is Increased in Gene Regions. 7A) Violin plot with incorporated box plots of rDNA position occupancies sorted by region. Counts are log2 transformed. 7B) violin plot with incorporated box plots of all spacer positions and all gene positions. Counts are log2 transformed and normalized. Comparison by Mann-Whitney U test. *** = p-value < $1*10^{-8}$

To determine sites of polymerase pausing, we identified the top 2.5% of positions by Pol I occupancy (171 positions) for each replicate as top sites of polymerase enrichment/pausing. 127 positions were found in the top 2.5% by occupancy in all three sets. None of these top positions coincide with the ends of the T tracts identified by our *in vitro* studies (Figs 4A and 8A). Broadening our analysis, we examined the sequence content immediately upstream of the polymerase occupancy at each top pause site. We focused on the seven nucleotides upstream of each position, as this would reflect the minimal RNA:DNA hybrid within the polymerase. We found a significant correlation between hybrid U content and Pol I occupancy (Fig 8B). However, the correlation coefficient was very small (Spearman's correlation coefficient (ρ) = -0.043, $p < 1*10^{-8}$), indicating that that hybrid U content is not a reliable predictor of Pol I pausing *in vivo*. Generation of a sequence logo using the top sites yielded the intriguing result that only one of the 127 top sites had a G residue as the last encoded nucleotide (Fig 8C). We then looked at occupancy across the gene as a factor of the last encoded nucleotide and found a significantly lower median occupancy value for G residues (Fig 8D). Expanding this analysis, we determined the occupancy at positions proximal to G residues (Fig 8E). We observed significantly increased occupancy directly upstream of G residues. These data strongly suggest that the incorporation of G nucleotides by Pol I occurs more slowly than the other three, but addition of the next nucleotide after a G residue occurs more rapidly. The relationship between nascent pre-ribosome assembly and Pol I transcription is crucially important, and NETSeq provides a critical tool for characterization of this complex biosynthetic process.

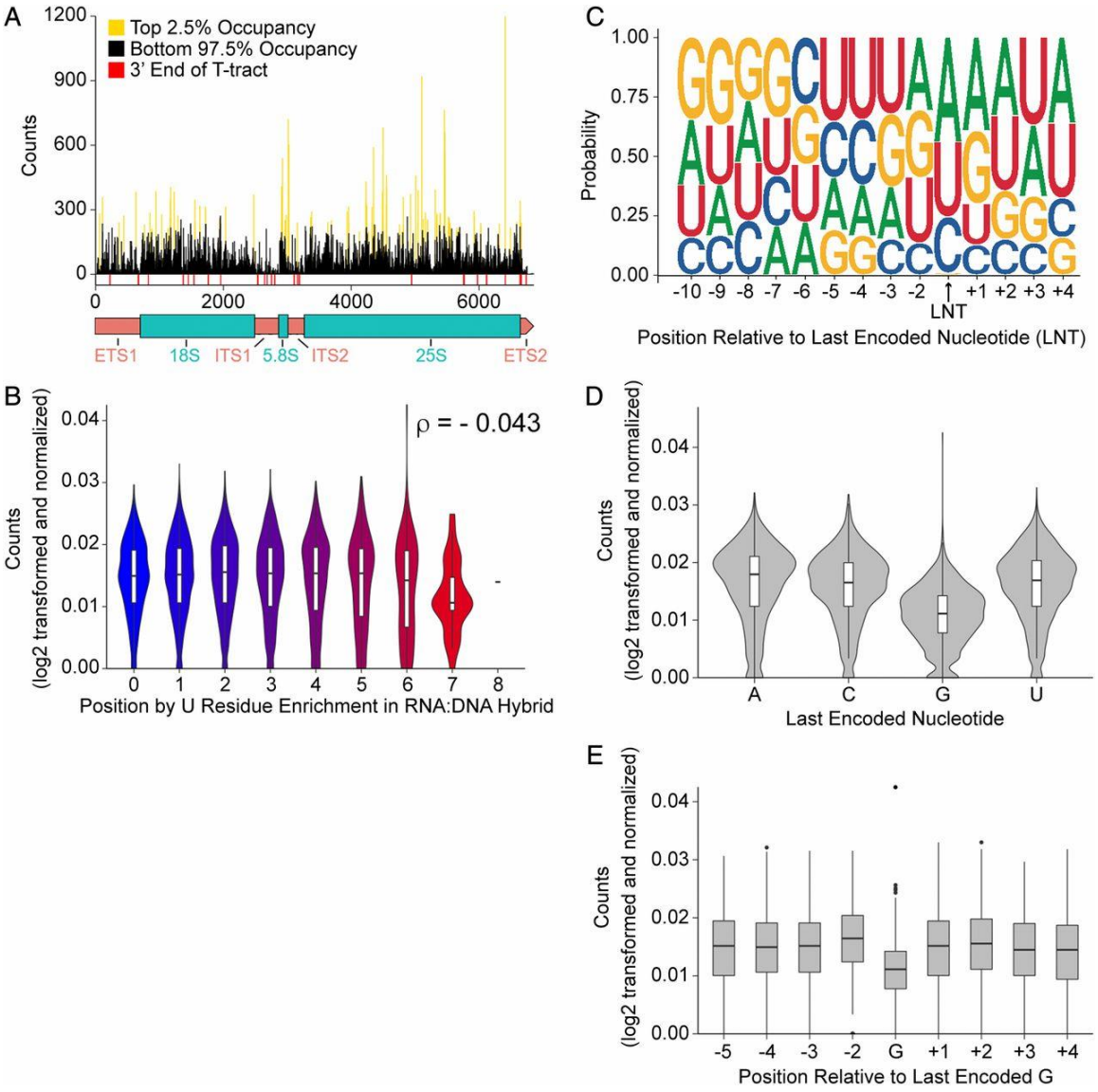

Figure 8. Low RNA Polymerase I Occupancy Correlates Strongly with Last Encoded Nucleotide. 8a) NETSeq Pol I 3' end densities (black) with significant positions indicated (wheat) and previously identified U-tracts (red,) with 35S gene diagram below. 8b) Violin plot of 3' positions sorted by RNA:DNA hybrid U content, with incorporated box plots. Counts are log2 transformed and normalized. Correlation coefficient determined by spearman correlation test, p-value < 1*10-8. 8c) sequence logo of the RNA:DNA hybrid corresponding to top 3' positions conserved in all three sets. 8d) Violin plot with incorporated box plots of rDNA position occupancies sorted by last encoded nucleotide. Comparison by Kruskal Wallis test produced a p-value < 1*10-8. 8e) box plot of position occupancies proximal to last encoded G residues. Counts are log2 transformed. Comparisons by Mann-Whitney U test. ns = not significant, * = p-value < 5*10-6, *** = p-value < 5*10-16

Certain aspects of our data suggest contamination by mature rRNA species or nascent transcripts from different polymerases. We observe no significant reads from Pol II-transcribed genes or from Pol III-transcribed tRNA genes (SI Appendix, Figs S4C and S4D). However, we observe signal at positions throughout the 5S gene, which is transcribed by Pol III (SI Appendix, Fig S4B). These data suggest that either 5S rRNA is co-precipitated with the nascent 35S rRNA or we are detecting contamination from degrading mature rRNAs and mapping them to the 35S gene. Unlike studies focused on mRNA synthesis, we cannot simply "ribo-deplete" our libraries as we are uniquely focused on the rDNA. In order to rigorously ensure that our conclusions are reflective of nascent transcription and not artifacts of bulk ribosome decay, we repeated our analysis focusing only on residues in the spacer regions of the 35S gene. These residues are not present in mature ribosomes and are rapidly degraded during pre-rRNA processing. We found that the trends observed through the entire 35S gene were recapitulated in this smaller spacer RNA set (SI Appendix, Fig S6). These data confirm that contaminating rRNA is not the driver of the observed phenomena.

Our NETSeq data are highly reproducible, revealing sites of increased and decreased Pol I occupancy *in vivo.* We've also discovered two sequence variables that correlate significantly with low Pol I occupancy. However, these data also suggest that DNA sequence elements alone do not govern Pol I occupancy. This disconnect between the biochemical studies and experiments in live cells is expected given the relative complexity of the two systems. Importantly, these data suggest that although sequence elements *can* influence Pol I directly, there may be many additional features that influence and control Pol I *in vivo*. The search for these factors (transcription elongation

factors, nucleosomes, RNA structures, as-yet undiscovered sequence elements, etc.) represents an exciting area of future investigation.

DISCUSSION

Analysis of Pol I pausing *in vitro* and *in vivo* reveals important features of the enzyme while raising key questions regarding the cellular factors and conditions that govern pausing. Our data clearly demonstrate that Pol I is prone to pausing. This observation is not surprising, since previous studies have identified or predicted pausing for many RNA polymerases [9, 23-28]. Pausing of transcription elongation can have important consequences on gene regulation. Sophisticated regulatory systems, such as attenuation mechanisms in bacteria, rely on programmed pausing of RNAP [25]. Furthermore, pausing by Pol II is a recently defined feature of eukaryotic gene regulation that has dramatic impact on cell growth and differentiation [29, 30]. In this study we begin to explore the factors that control pausing by Pol I. We not only identify key features that are conserved among RNA polymerases but also raise questions regarding the factors that govern pausing *in vivo*.

DNA sequence effects on transcription are conserved between bacteria and eukaryotes

Here, we demonstrate that a prokaryotic terminator motif can induce pausing and termination of transcription by a eukaryotic polymerase. The observed pausing and termination of transcription by Pol I can be modulated by modification of the stem loop or U-tract regions of the *spy* terminator motif. We further demonstrated that the induced pausing and termination were both dependent on UTP concentration. All of these factors

are consistent with previous observations in prokaryotic systems [12, 19, 31]. Previous studies have suggested that certain DNA sequence elements can directly affect all RNA polymerases [32]. Our findings reveal clearly conserved effects of prokaryotic terminator motifs on both Pol I and RNAP. Similar motifs also affect Pol II [15-17]. All of these data support the conclusion that critical interactions between the DNA template, the nascent RNA, and the RNA polymerase have been preserved throughout evolution.

Substrate concentration may influence rRNA synthesis directly.

The data described above provide insight into sequence motifs that affect transcription elongation by Pol I. Our search for similar T-tracts in the rDNA yielded 54 sites (Fig 4A). *In vitro* transcription of one such site from the 25S region had a similar effect on Pol I transcription as the rho-independent terminator motif. We also determined that the pause, but not the termination effect of this motif, was sensitive to UTP concentration. In the context of gene regulation, this is potentially an important finding. Given the established link between Pol I transcription rate and efficient rRNA processing [4], it is possible that these pause sites mediate RNA secondary structure formation required for rRNA processing, a phenomenon previously observed in prokaryotic expression systems [11]. Taking into account the UTP concentration-dependence of these T-tracts, these pause sites could alternatively serve as cellular sensors of nutrient concentration. Each pause site might be sensitive to UTP concentration around a specific threshold. If pausing by Pol I were responsive to substrate concentration, then rRNA processing might be linked to cellular nutrient conditions. No such factor-independent stress sensor has been identified in eukaryotic rRNA expression. However, substrate NTP

concentration plays a pivotal role in the regulation of prokaryotic rRNA synthesis [33, 34]. This previously described regulatory mechanism demonstrates that nutrient sensing via the cellular nucleoside triphosphate pool can be exploited for the control of rRNA synthesis.

In vitro pausing by Pol I is not predictive of pause sites in vivo.

To characterize Pol I pausing *in vivo*, we adapted NETSeq methods from those used previously for Pol II [22]. We observed obvious heterogeneity in Pol I occupancy. Notably, we saw no significant correlation between identified T- tracts and Pol I occupancy (Fig 8A). In addition, we found that U-residue enrichment in the rDNA hybrid had little effect on Pol I occupancy (Fig 8B). These results stand in stark contrast to what we observed *in vitro.* One potential explanation for this discrepancy is that the biochemical studies were performed under conditions of limiting substrate. Our *in vitro* experiments were performed at an NTP concentration of 15 μM for each nucleotide. Increasing the concentration to 100 μM largely abrogated the effect on Pol I elongation for both motifs (Fig 5). The $K_{1/2}$ value for ATP concentration with respect to Pol I elongation *in vitro* is 170 μM [35], and the nucleolar concentrations of NTPs in living cells is expected to be much higher. Thus, our *in vitro* experiments were performed at sub-saturating substrate concentrations. The lack of pausing at these T-tracts *in vivo* may be due to the fact that the nucleolar NTP concentrations during balanced growth were simply too high. Perhaps under conditions when substrates are limited (as discussed above), DNA sequence elements may contribute more substantially to pausing *in vivo*.

Another potential explanation for the discrepancy between pause site selection *in vitro* versus *in vivo* is the relative complexity of the two experimental systems. The biochemical assays that we deploy are fully reconstituted. The DNA template is free of histones, and all of the proteins other than Pol I are expressed and purified from *E. coli*. Thus, these assays are designed to characterize the minimal set of factors that influence RNA synthesis. Our data demonstrate that under these purified conditions DNA sequence elements can have robust effects on transcription by distantly related RNA polymerases. However, *in vivo* these potential effects may be secondary to effects by other factors (DNA binding proteins or elongation assisting factors). Based on this model, there exists a collection of factors that directly influence the rate of Pol I transcription throughout the rDNA gene.

One candidate factor is the Spt4/5 complex. *SPT5* is conserved throughout eukarya [36]. NusG, the homologue of Spt5 in bacteria, has been shown to increase the elongation rate and general processivity of the *E. coli* RNA polymerase *in vitro [37]* . The Spt4/5 complex has been shown to affect the processivity of Pol II [38]. We demonstrated previously that Spt4/5 interacts directly with Pol I [39, 40]. Furthermore, deletion of *SPT4* or mutation of *SPT5* resulted in defective rRNA synthesis and processing [40, 41]. The interaction of the Spt4/5 complex with elongating Pol I may increase its transcription elongation rate or processivity, allowing it to read through sequences which might otherwise cause Pol I to pause or arrest. Defining the factors that perturb Pol I activity *in vivo* will impact our understanding of how cells orchestrate the complex process of ribosome assembly.

Pol I occupancy is reduced at G residues.

Analysis of the sequence context of the most robust sites of pausing revealed an anti-preference for G as the last encoded nucleotide (Fig 8C). Expanded analysis revealed significantly lower occupancy at G residues throughout the rDNA, indicating that this is a gene-wide trend instead of being specific to high-occupancy sites (Fig 8D). Analysis of occupancy at positions proximal to G residues also revealed a significant increase in occupancy directly upstream. This G anti-preference has not been described in any previous NETSeq studies, suggesting that the phenomenon is not simply an artifact of library generation. While a similar phenomenon has been observed *in vivo* for RNAP [42], this is the first *in vivo* observation of differential nucleotide incorporation by Pol I in any organism. Interestingly, Pol I occupancy at A residues was not similarly reduced compared to G residues, suggesting that the difference is not due simply to purine or pyrimidine base identity. The simplest explanation for these findings is that incorporation of G nucleotides is slow, but incorporation by Pol I after addition of a G to the nascent RNA is fast. We do not see any significant effect of the identity of the next encoded nucleotide (Fig 8C). Thus, this effect is apparently due to unique features of the deoxyC:riboG hybrid in the active site of the polymerase. The structural basis for this observation is not yet clear. Perhaps the deoxyC:riboG hybrid provides a particularly efficient substrate for isomerization/translocation of the polymerase or a favorable substrate for the next nucleotide addition. Mutational analyses and biochemical characterization will reveal the nature of this unexpected effect of rDNA sequence.

METHODS

Generation of DNA Template

We synthesized linear DNA fragments containing a *NotI* restriction site, the *S. cerevisiae* RNA polymerase I distal and core promoter sequences, the first 55 transcribed residues of the 35S gene mutated such that no C residues are encoded, 250 residues of rDNA from the 25S gene, the DNA sequence to be assayed, another 100 residues of rDNA from the 25S gene, an *SfoI* restriction site, and an *XhoI* restriction site, as diagrammed in Fig 1a. These fragments served as templates for all *in vitro* assays. For the study of the effect of *rho*-independent terminator motifs on Pol I, the sequence to be assayed was the *rho*-independent terminator motif from the *E. coli* gene *spy,* using the sequence described in the supplementary information from Chen, et. al. [19] . The Terminator T to C mutant template was prepared similarly, except residues 5'-TTTTCTTTTCTCTTCT-3' in the terminator motif were replaced with 5'-CCCCTCCCCTCTCCTC-3'. For the Terminator short stem loop mutant template, the residues 5'-TCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGATC-3' in the terminator motif were replaced with 5'-CAGTCGAAAGACTG-3'. For the rDNA sequence analysis, the sequence to be assayed consisted residues 1591 to 1710 of the *S. cerevisiae* 25S gene. The rDNA sequence T to C switch mutant template was prepared similarly, except residues 5'-TTATCTTTTCTTCTT-3' from the rDNA sequence were replaced with 5'-CCACTCCCCTCCTCC-3'. In the rDNA sequence upstream scramble mutant template the residues 5'-TGGAGACGTCGGCGCGAGCCCTGGGAGGAG-3' were replaced with 5'-CCGTAGTGCTAAGGTAACCTACAACGTGCT-3'. For the negative control, the sequence to be assayed is the 52 residue scramble sequence 5'-

CGTGGCGCGGACTCGAGGAACGGACGCGAAGCTAACGGAATAGTGACTCTTC -3'. The full sequences for all linear DNA fragments are available in the SI Appendix. The fragment was ligated into the pBluescript vector plasmid using the *NotI* and *XhoI* restriction sites and linearized using the *SfoI* restriction site. For the bioitinylated template, the linear DNA fragment was amplified by PCR with a 5' biotinylated primer. DNA concentrations were determined using a Nanodrop 2000c spectrophotometer.

Purification of Proteins

S. cerevisiae Pol I, Core Factor (Rrn6p, Rrn7p, Rrn11p) (CF,) and Tata Binding Protein (TBP) were purified as previously described [43].

S. cerevisiae Rrn3p was purified from *E. coli* bearing pDAS903. This plasmid consists of the pSUMO vector plasmid with the *RRN3* gene cloned into it such that the sequence MHHHHHH followed by the 101 amino acid SUMO fusion protein was added to the N-terminus of Rrn3p. Protein expression was induced by growth in phosphate-buffered TB supplemented with 0.25% v/v glycerol, 0.025 w/v glucose, and 0.1% w/v galactose. Cells from a 1 liter culture with Abs_{600} = 0.6 were pelleted at 4,400 x g for 30 minutes. The pellet was resuspended in 50 mL breakage buffer (50 mM Tris◦Cl, 500 mM KCl, 10 mM Imidazole, 1% v/v glycerol pH, 7.8.) The cells were lysed via a french pressure cell at 15,000 PSI. The lysate was then cleared at 36,000 x g for 30 minutes. The supernatant was incubated with a 50% v/v GE Fast-Flow nickel resin slurry in breakage buffer for two hours at 4°C. The nickel resin was pelleted via centrifugation at 500 x g for 3 minutes and washed with 10 mL wash buffer (50 mM Tris◦Cl, 200 mM KCl, 10 mM Imidazole, 1% v/v Glycerol, pH 7.8.) The protein was eluted with 7.5 mL elution buffer

(50 mM Tris◦Cl, 200 mM KCl, 250 mM Imidazole, 1% v/v Glycerol, pH 7.8.) The nickel elution fraction was then supplemented with 1 mM DTT and purified SUMO protease and incubated overnight to cleave the 6xHis-SUMO tag off of Rrn3p. The nickel eluent containing now untagged Rrn3p was then loaded onto a GE Mono Q 5/50 GL anion exchange column and washed with 10 mL buffer A (50 mM Tris◦Cl, 200 mM KCl, 10% v/v Glycerol, pH 7.8.) The bound protein was eluted using a gradient of 0 – 100% buffer B (50 mM Tris◦Cl, 1 M KCl, 10% v/v Glycerol, pH 7.8.) All protein concentrations were determined via dual-beam spectrophotometry.

Promoter-Dependent In Vitro Transcription

Based on a previously described *in vitro* transcription assay [18], standard experiments were performed at 25°C with a volume of 20 µL. Experiments contained standard reaction buffer (17.5 mM tris acetate pH 7.9, 87.5 mM potassium glutamate pH 7.95, 7 mM magnesium acetate, 1.75 mM DTT, 0.035 U/µL RNAse inhibitor, 0.175 mg/mL BSA, 2.6% glycerol,) 169 nM CF, 169 nM TBP, 169 nM Rrn3p, 169 nM RNA polymerase I, 2.3 nM linearized DNA template, 15 µM ATP, CTP, GTP, UTP, ≈300 nM [α-^{32}P] NTP (UTP for the *rho*-independent terminator studies and GTP for the 25S pause region studies), and 0.025 mg/mL heparin. Rrn3p and RNA polymerase I were incubated at 25°C for one hour prior to the transcription experiment.

Linearized DNA template was added to a 2.5-fold concentrated mixture of standard reaction buffer, followed by CF, TBP, and Rrn3P/RNA polymerase I complex. ATP, GTP, UTP, and radiolabeled nucleotide were then added to the mixture to allow for transcription of the 55 nucleotide C-less coding region for three minutes. CTP and

heparin were then added to allow the polymerases to transcribe past the C-less coding region and prevent binding of un-bound polymerases to the template, respectively. After the indicated amount of time, reactions were quenched with 250 µL 1M ammonium acetate in 95% ethanol and cooled overnight at -20°C. The nascent RNA was pelleted by centrifugation at 16,800 x g for 15 minutes. The pellet was resuspended in 20 µL RNA loading dye (90% formamide, 25 mM EDTA, 0.025 mg/ml bromophenol blue, pH 8.54.) After 10 minutes of incubation, 10 µL of each sample was loaded into an 8% polyacrylamide gel, which was run for 65 minutes at 700 volts in 1x TBE. The gel was then dried for two hours, and exposed to a phosphor screen overnight.

The *in vitro* transcription experiment with the biotinylated template was performed as above- except the reactions were quenched with two volumes of 38 mM EDTA in standard reaction buffer. Half of each reaction was incubated with streptavidin-coated magnetic beads for 25 minutes at 25°C. The beads were then sequestered by magnet and the flow through was isolated. Both fractions for each time point were then combined with 250 µL 1M ammonium acetate in 95% ethanol and resolved by polyacrylamide gel as above.

Gel Imaging and Data Analysis

Gels were imaged using the GE Typhoon Scanner. All analysis was performed using the ImageQuant TL software package. The volume of the full length and truncated product bands were determined. The volume for each band was then normalized to radioactive nucleotide incorporation. To produce the fractional product value at each time point, the normalized volume of the truncated product band in that lane was divided by

the sum of the normalized truncated product volume and the normalized full length product volume, and multiplied by 100.

Native Elongating Transcript Sequencing for Pol I.

NETSeq was adapted for use with Pol I, based in large part on previous studies targeting Pol II [22] Detailed description of the cell growth, harvest, RNA isolation, library preparation, and data analysis are provided as in the SI Appendix.

ACKNOWLEDGEMENTS

We thank the Heflin Center for Genomic Sciences at UAB, specifically Drs. Mike Crowley and David Crossman, for their advice and technical assistance throughout this study. We also wish to thank Dr. Tom Broker providing the cell disruptor that made NETSeq possible. We would also like to thank Dr. Mikhail Kashlev for informal conversations that led us to explore the effects of prokaryotic terminators on RNA polymerase I. We thank members of the Schneider lab for critical evaluation of this project. This study was funded by NIH grant GM084946 to D.A.S.

SUPPLEMENTAL METHODS

DNA template sequences

negative control construct

GCGGCCGCctgtcactttggaaaaaaaatatacgctaagatttttggagaatagcttaaattgaagtttttctcggcgagaa
atacgtagttaaggcagagcgacagagagggcaaaagaaaataaaagtaagattttagtttgtaatgggaggggggggtttagtc
atggagtacaagtgtgaggaaaagtagttgggaggtacttcatgcgaaagcagttgaagacaagttcgaaaagagtttggaaac
gaattcgagtaggcttgtcgttcgttatgtttttgtaaatggcctcgtcaaacggtggagagagtcgctaggtgatcgtcagatctgc
ctagtctctatacagcgtgtttaattgacatgggttgatgcgtattgagagatacaatttgggaagaaattcccagagtgtgtttctttt
gcgtttaacctgaacagtctcatcgtgggcatcttgcgattccattggtgagcagcgaaggatttggtggattactagctaatagca
atctatttcaaagaattcaaacttgggggaatgccttgttgaatagccggtcgcaagactgtgattcttcaagtgtaacctcctctca
aatc**CGTGGCGCGGACTCGAGGAACGGACGCGAAGCTAACGGAATAGTGA**
CTCTTCagcgatatcaaacgtaccattccgtgaaacaccggggtatctgtttggtggaacctgattagaggaaactcaaag
agtgctatggtatggtgacggagtg***GGCGCCC*CTCGAG**

spy terminator construct

GCGGCCGCctgtcactttggaaaaaaaatatacgctaagatttttggagaatagcttaaattgaagtttttctcggcgagaaat
acgtagttaaggcagagcgacagagagggcaaaagaaaataaaagtaagattttagtttgtaatgggaggggggggtttagtcat
ggagtacaagtgtgaggaaaagtagttgggaggtacttcatgGgaaagGagttgaagaGaagttGgaaaagagtttggaaa
GgaattGgagtaggcttgtcgttcgttatgtttttgtaaatggcctcgtcaaacggtggagagagtcgctaggtgatcgtcagatct
gcctagtctctatacagcgtgtttaattgacatgggttgatgcgtattgagagatacaatttgggaagTTCAGCCAAAAA
ACTTAAGACCGCCGGTCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGA
TCGGCGGTTTTCTTTTCTCTTCTCAAcgtaccattccgtgaaacaccggggtatctgtttggtggaacctg
attagaggaaactcaaagagtgctatggtatggtgacggagtg**GGCGCCC*CTCGAG***

spy terminator construct, T to C mutant

GCGGCCGCctgtcactttggaaaaaaaatatacgctaagatttttggagaatagcttaaattgaagtttttctcggcgagaaat

acgtagttaaggcagagcgacagagagggcaaaagaaaataaaagtaagattttagtttgtaatgggaggggggggtttagtcat

ggagtacaagtgtgaggaaaagtagttgggaggtacttcatgGgaaagGagttgaagaGaagttGgaaaagagtttggaaa

GgaattGgagtaggcttgtcgttcgttatgtttttgtaaatggcctcgtcaaacggtggagagagtcgctaggtgatcgtcagatct

gcctagtctctatacagcgtgtttaattgacatgggttgatgcgtattgagagatacaatttgggaagTTCAGCCAAAAA

ACTTAAGACCGCCGGTCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGA

TCGGCGGCCCCTCCCCTCTCCTCCAAcgtaccattccgtgaaacaccggggtatctgtttggtggaacct

gattagaggaaactcaaagagtgctatggtatggtgacggagtg**GGCGCC*CTCGAG***

spy terminator construct, short stem loop mutant

GCGGCCGCctgtcactttggaaaaaaaatatacgctaagatttttggagaatagcttaaattgaagtttttctcggcgagaaat

acgtagttaaggcagagcgacagagagggcaaaagaaaataaaagtaagattttagtttgtaatgggaggggggggtttagtcat

ggagtacaagtgtgaggaaaagtagttgggaggtacttcatgGgaaagGagttgaagaGaagttGgaaaagagtttggaaa

GgaattGgagtaggcttgtcgttcgttatgtttttgtaaatggcctcgtcaaacggtggagagagtcgctaggtgatcgtcagatct

gcctagtctctatacagcgtgtttaattgacatgggttgatgcgtattgagagatacaatttgggaagTTCAGCCAAAAA

ACTTAAGACCGCCCAGTCGAAAGACTGGGCGGTTTTCTTTTCTCTTCTCAAcgta

ccattccgtgaaacaccggggtatctgtttggtggaacctgattagaggaaactcaaagagtgctatggtatggtgacggagtg

GGCGCC*CTCGAG*

rDNA sequence construct

GCGGCCGCctgtcactttggaaaaaaaatatacgctaagatttttggagaatagcttaaattgaagtttttctcggcgagaaat

acgtagttaaggcagagcgacagagagggcaaaagaaaataaaagtaagattttagtttgtaatgggaggggggggtttagtcat

ggagtacaagtgtgaggaaaagtagttgggaggtacttcatgGgaaagGagttgaagaGaagttGgaaaagagtttggaaa

GgaattGgagtaggcttgtcgttcgttatgtttttgtaaatggcctcgtcaaacggtggagagagtcgctaggtgatcgtcagatct

gcctagtctctatacagcgtgtttaattgacatgggttgatgcgtattgagagatacaatttgggaagGGAATCCGGTTA

AGATTCCGGAACCTGGATATGGATTCTTCACGGTAACGTAACTGAATGTGGA

GACGTCGGCGCGAGCCCTGGGAGGAGTTATCTTTTCTTCTTAACAGCTTATCA

CCCcgtaccattccgtgaaacaccggggtatctgtttggtggaacctgattagaggaaactcaaagagtgctatggtatggtga

cggagtg**GGCGCC*CTCGAG***

rDNA sequence construct, T to C mutant

GCGGCCGCctgtcactttggaaaaaaaatatacgctaagatttttggagaatagcttaaattgaagtttttctcggcgagaaat

acgtagttaaggcagagcgacagagagggcaaaagaaaataaaagtaagattttagtttgtaatgggaggggggggtttagtcat

ggagtacaagtgtgaggaaaagtagttgggaggtacttcatgGgaaagGagttgaagaGaagttGgaaaagagtttggaaa

GgaattGgagtaggcttgtcgttcgttatgtttttgtaaatggcctcgtcaaacggtggagagagtcgctaggtgatcgtcagatct

gcctagtctctatacagcgtgtttaattgacatgggttgatgcgtattgagagatacaatttgggaagGGAATCCGGTTA

AGATTCCGGAACCTGGATATG*G*ATTCTTCACGGTAACGTAACTGAATGTGGA

GACGTCGGCGCGAGCCCTGGGAGGAGCCACTCCCCTCCTCCAACAGCTTATC

ACCCcgtaccattccgtgaaacaccggggtatctgtttggtggaacctgattagaggaaactcaaagagtgctatggtatggt

gacggagtg**GGCGCC*CTCGAG***

rDNA sequence construct, upstream scramble mutant

GCGGCCGCctgtcactttggaaaaaaaatatacgctaagatttttggagaatagcttaaattgaagtttttctcggcgagaaat

acgtagttaaggcagagcgacagagagggcaaaagaaaataaaagtaagattttagtttgtaatgggaggggggggtttagtcat

ggagtacaagtgtgaggaaaagtagttgggaggtacttcatgGgaaagGagttgaagaGaagttGgaaaagagtttggaaa

GgaattGgagtaggcttgtcgttcgttatgtttttgtaaatggcctcgtcaaacggtggagagagtcgctaggtgatcgtcagatct

gcctagtctctatacagcgtgtttaattgacatgggttgatgcgtattgagagatacaatttgggaagGGAATCCGGTTA

AGATTCCGGAACCTGGATATG*G*ATTCTTCACGGTAACGTAACTGAATGCCGT

AGTGCTAAGGTAACCTACAACGTGCTTTATCTTTTCTTCTTAACAGCTTATCA

CCCcgtaccattccgtgaaacaccggggtatctgtttggtggaacctgattagaggaaactcaaagagtgctatggtatggtga cggagtg**GGCGCC*CTCGAG***

Cell culture and immunoprecipitation

Per replicate, three litres of *S. cerevisiae* bearing an RPA135 C-terminal (3x-HA 7x- his) tagged mutant (*MATα ade2-1 ura3-1 trp1-1 leu2-3 112 his3-11,15 can1-100 RPA135-(HA)3-(His)7:TRP1Mx6 rpa190Δ::HIS3Mx6* carrying *pRS315-RPA190*) [44] were grown in YEPD media to A_{600} = 0.3 at 30°C with nutation. The cells were collected via filtration, then frozen in liquid nitrogen. The cells were then cryogenically lysed via ten one-minute cycles of 16mm amplitude grinding in the Mikro-Dismembrator II grinding mill. The grinding cup was pre-chilled by immersion in liquid nitrogen, and the re-immersed for 1 minute in between grinding cycles.

Grindates for each liter were dissolved in five-fold weight amounts of ice-cold lysis buffer (20 mM Tris Cl pH 7.9, 0.4% Triton X-100, 0.1% NP-40, 100 mM NH_4Cl, 5 mM EDTA Na pH 8.5, 1x HALT Protease Inhibitor, 25 U/ml SUPERase-in RNAse Inhibitor). Aliquots of Pierce Anti-HA magnetic beads equal to 4% of lysis buffer volume were washed 4x with ice-cold lysis buffer. The lysates were centrifuged at 16,000 x g for 15 minutes. The lysate supernatants were combined with pre-washed aliquots of Pierce Anti-HA magnetic beads and set to nutation at 4°C for 3 hours. The beads were then isolated via magnet and washed 4x with ice-cold wash buffer (20 mM Tris Cl pH 7.9, 0.4% Triton X-100, 0.1% NP-40, 300 mM KCl, 50 mM EDTA Na pH 8.5, 25 U/ml SUPERase-in RNAse Inhibitor). Finally, the beads were resuspended in 900 µl TES (10 mM Tris Cl pH 7.5, 1% SDS) and extracted 3x with 900 µl aliquots of acidic (pH 4.3)

phenol, then 2x with 900 µl aliquots of chloroform. The remaining aqueous solutions were combined with 1.2 ml ammonium acetate precipitation solution (1M ammonium acetate, 95% ethanol) and 2 µl glycoblue, and set to -80°C for at least 2 hours. The solutions were centrifuged at 16,000 x g for 1 hour at 4°C. The pellets were then washed 2x with 750 µl 75% ethanol, and resuspended in 10 µl 10 mM tris Cl pH 7.0.

Linker ligation, zinc chloride fragmentation, and size selection

The isolated RNA samples from each liter were denatured at 80°C for 2 minutes, then put on ice. A 5' adenylated DNA linker end-blocked with a 3' di-deoxy C (5'-/5rApp/CTCCACGAGTCATCCGC/3ddc/-3', Integrated DNA Technologies) was denatured at 80°C for 3 minutes, then combined with RNA samples and buffering conditions for a labelling reaction with the following concentrations (12% PEG MW 8000, 1x T4 RNA Ligase buffer (NEB), 5 µM linker, and 10 U / µl T4 RNA Ligase 2, truncated (NEB)). The reactions were set to 25°C for three hours. In order to fragment the RNAs 2.2 µl zinc chloride fragmentation buffer (100 mM tris Cl ph 7.0, 100 mM $ZnCl_2$) was added to each ligation reaction and set to 70°C for 20 minutes. The fragmentation reactions were then quenched by that addition of 2.5 µl 200 mM EDTA Na pH 8.5, 2 µl glycoblue, and 1 ml ammonium acetate precipitation solution, and set to -80°C for at least 2 hours. The solutions were centrifuged at 16,000 x g for 1 hour at 4°C. The RNA pellets were then washed 2x with 750 µl 75% ethanol. The RNA samples were denatured at 98°C for 5 minutes, loaded on a pre-run 10% polyacrylamide gel, and run at 700 volts for 70 minutes. The gel was stained for 30 minutes at room temperature in 1X SYBR Gold stain in 1X TBE buffer. The 30 to 330 nt regions were excised for each replicate.

The gel slices were pulverized and combined with 600 µl H_2O, and set to -80°C for 14 hours. The gel slurries were then set to 70°C for 20 minutes, and the liquid was isolated via a 5 minute 16,000 x g centrifugation in Costar Spin-X Centrifuge Tube Filters (Corning). The solutions were then combined with 37.5 µl 3M ammonium acetate, 2 µl Glycoblue, and 1.125 ml isopropanol. The solutions were set to -80°C for 2 hours, then centrifuged at 16,000 x g for 1 hour at 4°C. The RNA pellets were washed twice with 750 µl 75% ethanol, and dried at room tempurature for 25 minutes. The pellets were then resuspended in 10 µl 10 mM Tris Cl pH 6.9.

Reverse Transcription

The size-selected RNA samples were combined with a 5' phosphorylated reverse transcription primer (RT primer) with two internal 18 atom hexa-ethyleneglycol spacers (5'- /5Phos/CTGTAGGCACCATCAATG ATCGTCGGA/isp18/CACTCA/isp18/CGTCTCTTCTGCGGATGACTCGTGGAG-3', Integrated DNA Technologies) with buffering conditions (1X 5X First Strand Buffer, 400 µM each dATP, dCTP, dGTP, dTTP, and 250 nM RT primer), denatured at 65°C for 5 minutes, and set on ice. The following components were then added to complete the full reverse transcription reaction (5 U/µl SUPERase-in RNAse Inhibitor, 4 mM DTT, and 8 U/ µl Superscript III (Invitrogen). The solutions were set to 50°C for 30 minutes. In order to get rid of the residual RNA 1.8 µl 1M NaOH was added to each reaction and set to 98°C for 20 minutes. The 70 to 330 nt range cDNAs for each sample were size-selected and precipitated as described above, except using 3M sodium for precipitation instead of 3M sodium acetate, and resuspended in 15 µl 10 mM Tris Cl pH 7.0.

Circularization, PCR Amplification, Size Selection, and Library Desalting

Circularization with CircLigase II ssDNA Ligase (Epicentre) was performed per manufacturer's instructions, with the addition of a second 1 µl aliquot of Circligase and 60-minute incubation step at 60°C following the first. The Circularized cDNAs were then amplified in a PCR reaction with the following concentrations: (1x Phusion Buffer, 94 nM forward primer, 94 nM reverse primer, 330 µM each dATP, dCTP, dGTP, dTTP, and 0.019 U/µl phusion polymerase (NEB)). Each sample was amplified with a unique combination of forward and reverse primers to allow for demultiplexing (table 1.) The reactions amplified by the following PCR cycle: 98°C for 30 seconds, 12x (98°C for 10 seconds, 60°C for 10 seconds, 72°C for 30 seconds.) The 160-330 nt range of library DNAs for each sample were size-selected and precipitated as described above, and resuspended in 25 µl 10 mM tris acetate pH 8.0. Libraries were desalted using illustra Microspin S-200 HR Columns (GE) according to manufacturer's instructions.

Sequencing

Libraries were sequenced on the NextSeq500 according to manufacturer's instructions, using the following index primer (5'-CCATCAATGATCGTCGGA-3', Integrated DNA Technologies) and sequencing primer (5'-CGTCTCTTCTGCGGATGACTCGTGG-3', Integrated DNA Technologies.) The image **b**ase**c**a**l**l (bcl) files from the Illumina NextSeq500 were converted to fastq file format using Illumina's bcl2fastq software (ver. 2.18.0.12.)

Analysis software

Sequence trimming, alignment, and data formatting were performed with Trim Galore (ver. 0.4.1,) Burrows-Wheeler Aligner (BWA) (ver. 0.6)[45], Samtools (ver. 1.3.1)[46], and Bedtools (ver. 2.9.1)[47]. Subsequent analysis was performed in R (ver. 3.3.1)[48] using the following packages: ggplot2 (ver. 2.2.1)[49], dplyr (ver. 0.7.4)[50], and ggseqlogo (ver 0.1)[51]. Unix and R analysis scripts available upon request.

Sequence Trimming and Alignment

All sequences were trimmed of the first two bases (AG in all sequences by library design.) Where detected, the 3' library sequence (CTGTAGGCACCAT) and all subsequent bases were also removed. The trimmed sequences were then aligned to the *S. cerevisiae* genome (SacCer3 assembly, downloaded 25 October 2017, URL: http://hgdownload.cse.ucsc.edu/goldenPath/sacCer3/bigZips/chromFa.tar.gz) via the BWA-MEM tool, with a minimum alignment score threshold of 20.

Sequence Logo Generation

127 positions were "significant" in all three replicates. To represent the RNA:DNA hybrid a 14 nt sequence (comprised of the 9 bases upstream of the "significant" position, the "significant" position itself, and the 4 nucleotides downstream) was isolated for each of these positions. These sequences were then used to produce a sequence logo via ggseqlogo using default settings.

REFERENCES

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

2. Woolford, J.L., S. J. Baserga. , *Ribosome Biogenesis in the Yeast Saccharomyces Cerevisiae.* Genetics, 2013. **195**(3): p. 643-681.

3. Hein, N., K. M. Hannan, A. J. George, E. Sanij, R. D. Hannan., *The nucleolus: an Emerging Target for Cancer Therapy.* Trends Mol. Med. , 2013. **19**(11): p. 643-654.

4. Koš, M., D. Tollervey, *Yeast Pre-rRNA Processing and Modification Occur Cottranscriptionally.* Mol. Cell, 2010. **37**(6): p. 809-820.

5. Grummt, I., *Life on a Planet of its Own: Regulation of RNA Polymerase I Transcription in the Nucleolus.* Genes Dev. , 2003. **17**: p. 1691-1702.

6. Stefanovsky, V.Y., F. Langlois, D. Bazett-Jones, G. Pelletier, T. Moss, *ERK Modulates DNA Bending and Enhancesome Structure by Phosphorylating HMG1-Boxes 1 and 2 of the RNA Polymerase I Transcription Factor UBF.* biochemistry, 2006. **45**(11): p. 3626-3634.

7. Panov, K.I., J. K. Friedrich, J. Russell, J. C. Zomerdijk, *UBF Activates RNA Polymerase I Transcription by Stimulating Promoter Escape.* Embo. J. , 2006. **25**(14): p. 3310-3322.

8. Zhang, Y., S. L. French, A. L. Beyer, D. A. Schneider, *The Transcription Factor THO Promotes Transcription Initiation and Elongation by RNA Polymerase I.* J. Biol. Chem. , 2016. **291**(6): p. 3010-3018.

9. Zhang, Y., M .L. Sikes, A. L. Beyer, D. A. Schneider, *The Paf1 Complex is Required for Efficient Transcription Elongation by RNA Polymerase.* Proc. Natl. Acad. Sci. , 2009. **106**(7): p. 2153-2158.

10. Schneider, D.A., A. Michel, M. L. Sikes, L. Vu, J. A. Dodd, S. Salgia, Y. N. Osheim, A. L. Beyer, M. Nomura, *Transcription Elongation by RNA Polymerase I is Linked to Efficient rRNA Processing and Ribosome Assembly.* Mol. Cell, 2007. **26**(2): p. 217-229.

11. Lost, I., M. Dreyfus, *The Stability of Escherichia coli lacZ mRNA Depends Upon the Simultaneity of its Synthesis and Translation.* EMBO J. , 1995. **14**(13): p. 3252-3261.

12. Wilson, K.S., P. Von Hippel, *Transcription Termination at Intrinsic Terminators: the Role of the RNA Hairpin.* Proc. Natl. Acad. Sci., 1995. **92**(19): p. 8793-8797.

13. Gusarov, I., E. Nudler, *The Mechanism of Intrinsic Transcription Termination.* Mol. Cell, 1999. **3**(4): p. 495-504.

14. Werner, F., D. Grohmann, *Evolution of Multi-Subunit RNA Polymerases in the Three Domains of Life.* Nature Reviews Microbiology, 2011. **16**(6): p. 247-250.

15. Resnekov, O., Y. Aloni, *RNA Polymerase is Capable of Pausing and Prematurely Terminating Transcription at a Precise Location in vivo and in vitro.* Proc. Natl. Acad. Sci. , 1989. **86**: p. 12-16.

16. Seiberg, M., M. Kessler, A. J. Levine, Y. Aloni, *Human RNA Polymerase II Can Prematurely Terminate Transcription of the Adenovirus Type 2 Late Transcription Unit at a Precise Site that Resembles a Prokaryotic Termination Signal.* Virus Genes, 1987. **1**(1): p. 97-116.

17. Reines, D., D. Wells, M. Chamberlin, C. M. Kane, *Identification of Intrinsic Termination Sites in vitro for RNA Polymerase II within Eukaryotic Gene Sequences.* J. Mol. Biol. , 1987. **196**: p. 299-312.

18. Schneider, D.A., *Quantitative Analysis of Transcription Elongation by RNA Polymerase I In Vitro.* Methods in Molecular Biology, 2011. **809**: p. 579-591.

19. Chen, Y.J., P. Liu, A. A. Nielsen, J. A. Brophy, K. Clancy, T. Peterson, C. A. Voigt, *Characterization of 582 Natural and Synthetic Terminators and Quantification of Their Design Constraints.* Nat. Methods, 2013. **10**(7): p. 659-664.

20. Reynolds, R., R. M. Bermudez-Cruz, M. J. Chamberlin, *Parameters Affecting Transcription Termination by Escherichia coli RNA.* J. Mol. Biol. , 1992. **224**.

21. Ray-Soni, A., M. J. Bellecourt, R. Landick, *Mechanisms of Bacterial Transcription Termination: All Good Things Must End.* Biochemistry, 2016. **85**: p. 319-347.

22. Churchman, L.S., J. S. Weissman, *Nascent Transcript Sequencing Visualizes Transcription at Nucleotide Resolution.* Nature 2011. **469**: p. 368-373.

23. Core , L.J., J. T. Lis, *Transcription Regulation Through Promoter-Proximal Pausing of RNA Polymerase II.* Science, 2008. **319**(5871): p. 1791-1792.

24. Kireeva, M.L., M. Kashlev, *Mechanism of sequence-specific pausing of bacterial RNA polymerase.* Proc. Natl. Acad. Sci., 2009. **106**(22): p. 8900-8905.

25. Yakhnin, A.V., P. Babitzke, *NusA-stimulated RNA polymerase pausing and termination participates in the Bacillus subtilis trp operon attenuation mechanism in vitro.* Proc. Natl. Acad. Sci., 2002. **99**(17): p. 11067-11072.

26. Larson, M.H., R. A. Mooney, J. M. Peters, T. Windgassen, D. Nayak, C. A. Gross, S. M. Block, W. J. Greenleaf, R. Landick, and J. S. Weissman, *A Pause Sequence*

Enriched at Translation Start Sites Drives Transcription Dynamics In Vivo. Science, 2014. **344**(6187): p. 1042 - 1047.

27. Jonkers, I., J. T. Lis, *Getting up to speed with transcription elongation by RNA polymerase II.* Nat. Rev. Mol. Cell Biol., 2015. **16**(167): p. 167-177.

28. Henkin, T.M., C. Yanofsky, *Regulation by transcription attenuation in bacteria: How RNA provides instructions for transcription termination/antitermination decisions.* BioEssays, 2002. **24**.

29. Tastemel, M., A. A. Gogate, V. S. Malladi, K. Nguyen, C. Mitchell, L. A. Banaszynski, X. Bai *Transcription pausing regulates mouse embryonic stem cell differentiation.* Stem Cell Research, 2017. **25**: p. 250-255.

30. Adelman, K., J.T. Lis, *Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans.* Nat. Rev. Gen., 2012. **13**: p. 720-731.

31. Wilson, K.S., P. Von Hippel, *Stability of Escherichia coli Transcription Complexes Near an Intrinsic Terminator.* J. Mol. Biol., 1994. **244**(1): p. 36-51.

32. Bochkareva, A., Y. Yuzenkova, V. R. Tadigotla, N. Zenkin, *Factor-independent transcription pausing caused by recognition of the RNA–DNA hybrid sequence.* EMBO J. , 2011. **31**: p. 630-639.

33. Schneider, D.A., T. Gaal, R. L. Gourse, *NTP-sensing by RNA Promoters in Escherichia coli is Direct.* Proc. Natl. Acad. Sci., 2002. **99**(13): p. 8602-8607.

34. Gaal, T., M. S. Bartlett, W. Ross, C. L. Turnbough, R. L. Gourse, *Transcription Regulation by Initiating NTP Concentration: rRNA Synthesis in Bacteria.* Science, 1997. **278**(5346): p. 2092-2097.

35. Appling, F.D., A. L. Lucius, D. A. Schneider, *Transient-State Kinetic Analysis of the RNA Polymerase I Nucleotide Incorporation Mechanism.* Biophys. J., 2015. **109**: p. 2382-2903.

36. Werner, F., *A Nexus for Gene Expression- Molecular Mechanisms of Spt5 and NusG in the Three Domains of Life.* J. Mol. Biol. , 2012. **417**(1-2): p. 13-27.

37. Herbert, K.M., J. Zhou, R. A. Mooney, A. La Porta, R. Landick, S. M. Block, *E. coli NusG Inhibits Backtracking and Accelerates Pause-Free Transcription by Promoting Forward Translocation of RNA Polymerase.* J. Mol. Biol., 2010. **399**(1): p. 17-30.

38. Hartzog, G.A., T. Wada, H. Handa, F. Winston, *Evidence that Spt4, Spt5, and Spt6 Control Transcription Elongation by RNA Polymerase II in Saccharomyces cerevisiae.* Genes Dev., 1998. **12**(3): p. 357-369.

39. Viktorovskaya, O.V., F. D. Appling, D. A. Schneider, *Yeast Transcription Elongation Factor Spt5 Associates with RNA Polymerase I and RNA Polymerase II Directly*. J. Biol. Chem., 2011. **286**(21).

40. Schneider, D.A., S. L. French, Y. N. Osheim, A. O. Bailey, L. Vu, J. Dodd, J. R. Yates, A. L. Beyer, M. Nomura, *RNA Polymerase II Elongation Factors Spt4p and Spt5p Play Roles in Transcription Elongation by RNA Polymerase I and rRNA Processing*. Proc. Natl. Acad. Sci., 2006. **103**(34): p. 12707-12712.

41. Anderson, S.J., M. L. Sikes, Y. Zhang, S. L. French, S. Salgia, A. L. Beyer, M. Nomura, D. A. Schneider, *The Transcription Elongation Factor Spt5 Influences Transcription by RNA Polymerase I Positively and Negatively*. J. Biol. Chem., 2011. **286**: p. 18816-18824.

42. Imashimizu, M., H. Takahashi, T. Oshima, C. McIntosh, M. Bubunenko, D. L. Court, M. Kashlev, *Visualizing translocation dynamics and nascent transcript errors in paused RNA polymerases in vivo* Genome Biology, 2015. **16**(98).

43. Bedwell, G., F. D. Appling, S. J. Anderson, D. A. Schneider, *Efficient Transcription by RNA Polymerase I using Recombinant Core Factor*. Genes, 2012. **492**(1): p. 94-99.

44. Viktorovskaya, O.V., K.L. Engel, S.L. French, P. Cui, P.J. Vendeventer, E.M. Pavlovic, A.L. Beyer, C.D. Kaplan, D.A. Schneider, *Divergent Contributions of Conserved Active Site Residues to Transcription by Eukaryotic RNA Polymerases I and II.* Cell Reports 2013. **4**(5): p. 974-984.

45. Li, H., R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler Transform.* Bioinformatics, 2009. **25**: p. 1754-60.

46. Li H., B.H., A. Wysoker , T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, *The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics.* Bioinformatics, 2009. **25**: p. 2078-2079.

47. Quinlan A. R., I.M.H., *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-842.

48. Team, R.C., *R: A language and environment for statistical computing.* 2016.

49. Whickham, H., *ggplot2: Elegant Graphics for Data Analysis*. 2009, New York: Springer-Verlag.

50. Hadley W., F.R., L. Henry, K. Müller, *dplyr: A Grammar of Data Manipulation.* 2017.

51. Wagih, O., *ggseqlogo: A 'ggplot2' Extension for Drawing Publication-Ready Sequence Logos*. 2017.

Supplementary Table 1. Library Amplification Primers for NETSeq

Replicate	reverse primer sequence	forward primer sequence
1st	AATGATACGGCGACCACCGAGATCTACACtagatcgc	CAAGCAGAAGACGGCATACGAGATtcgcctta
	CGTCTCTTCTGCGGATGACTCG	TCCGACGATCATTGATGGTGCC
2nd	AATGATACGGCGACCACCGAGATCTACACtagatcgc	CAAGCAGAAGACGGCATACGAGATctagtacg
	CGTCTCTTCTGCGGATGACTCG	TCCGACGATCATTGATGGTGCC
3rd	AATGATACGGCGACCACCGAGATCTACACtagatcgc	CAAGCAGAAGACGGCATACGAGATttctgcct
	CGTCTCTTCTGCGGATGACTCG	TCCGACGATCATTGATGGTGCC

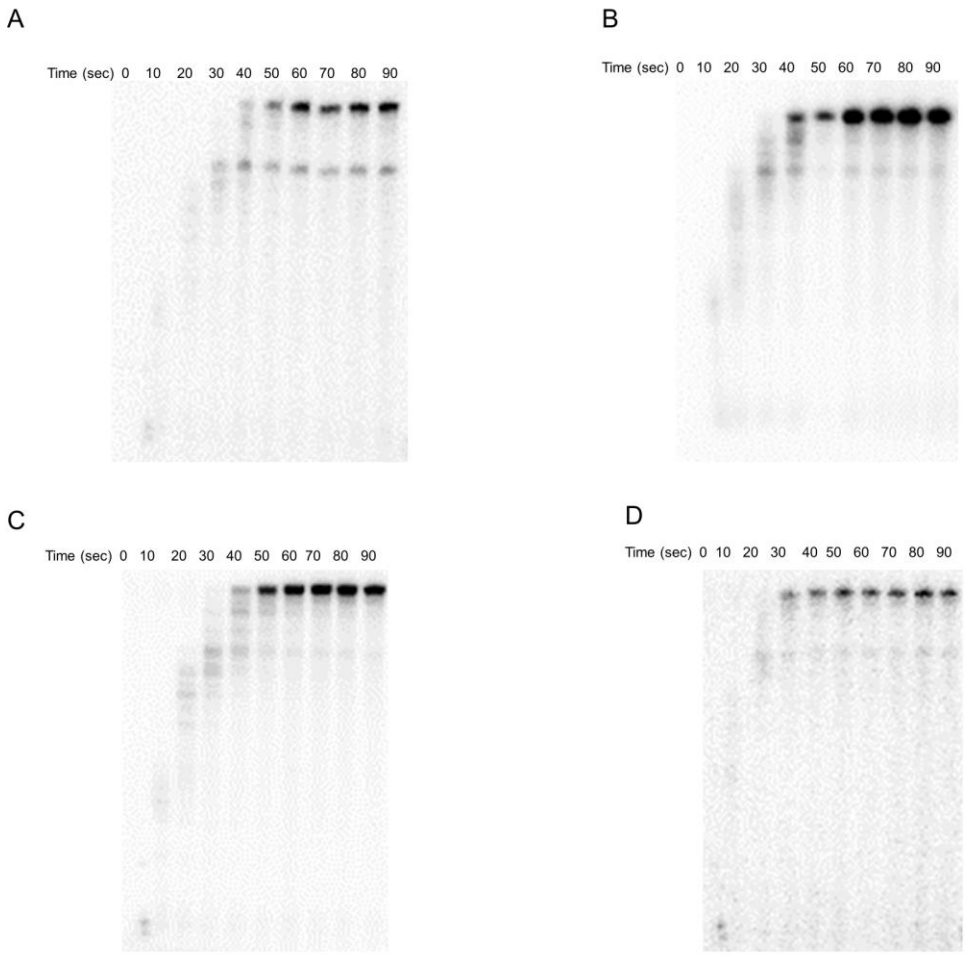

Figure S1. Representative Gels for Figures 2c, 3b, 3c, and 5a. S1a) Representative polyacrylamide gel for Fig 2c. S1b) representative polyacrylamide gel for Fig 3b. S1c) Representative polyacrylamide gel for Fig 3c. S1d) Representative polyacrylamide gel for Fig 5a, 100 µM UTP.

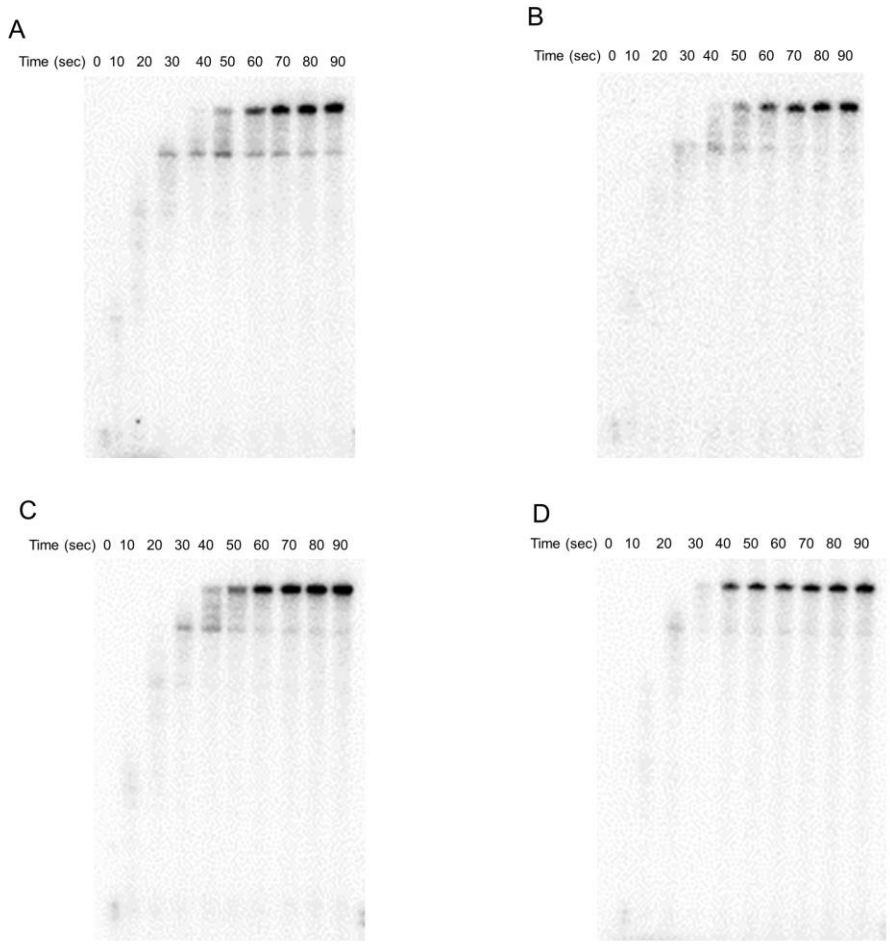

Figure S2. Representative Gels for Figures 4a, 4b, 4c, and 5c. S2a) Representative polyacrylamide gel for Fig 4a. S2b) representative gel for Fig 4b. S2c) representative polyacrylamide gel for 4c. S2d) representative polyacrylamide gel for Fig 5b, 100 µM UTP.

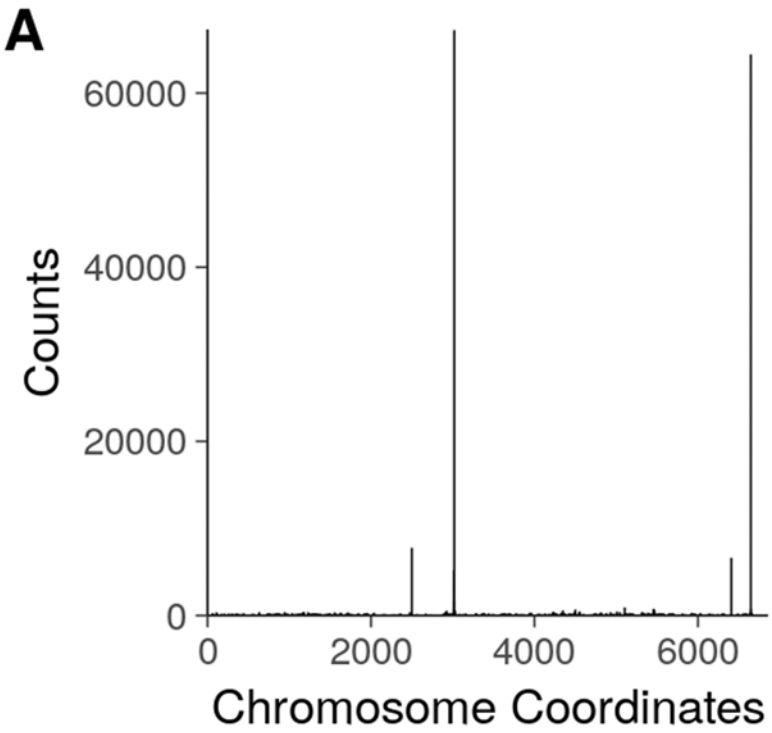

Figure S3. NETSeq Pol I 3' End Densities in the 35S Gene Including the Mature Ends of rRNA Species. S3a) NETseq Pol I 3' end densities in the 37S gene including the ends of mature rRNA species.

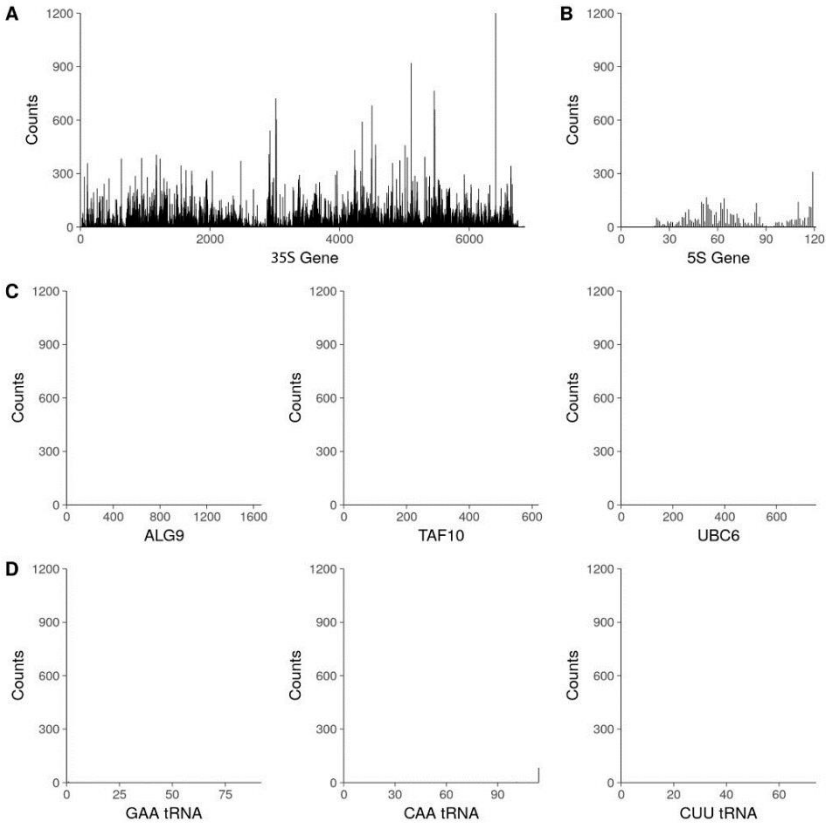

Figure S4. NETSeq 3' End Densities in Pol I, Pol II, and Pol III Genes. S4a) NETSeq 3' end densities in the 35S gene. S4b) NETSeq 3' end densities in the 5S gene. S4c) NETSeq 3' end densities in three Pol II transcribed genes. D) NETSeq 3' end densities in three Pol III transcribed genes.

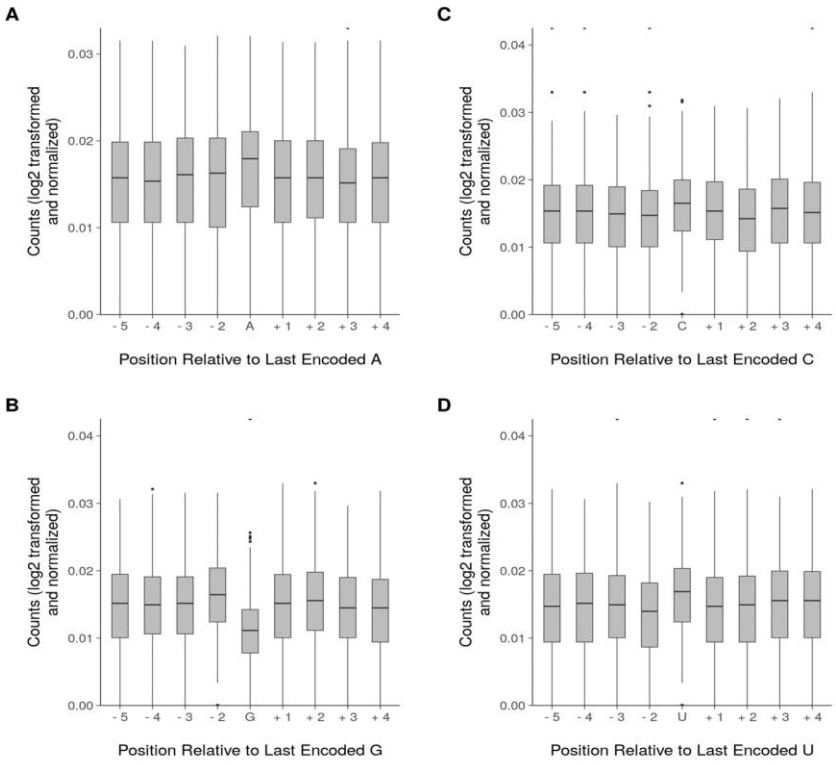

Figure S5. Position Occupancies Proximal to All Last Encoded Nucleotides. S5a-d) box plots of position occupancies proximal to last encoded A, C,G, and U residues, respectively. Counts are log2 transformed and normalized.

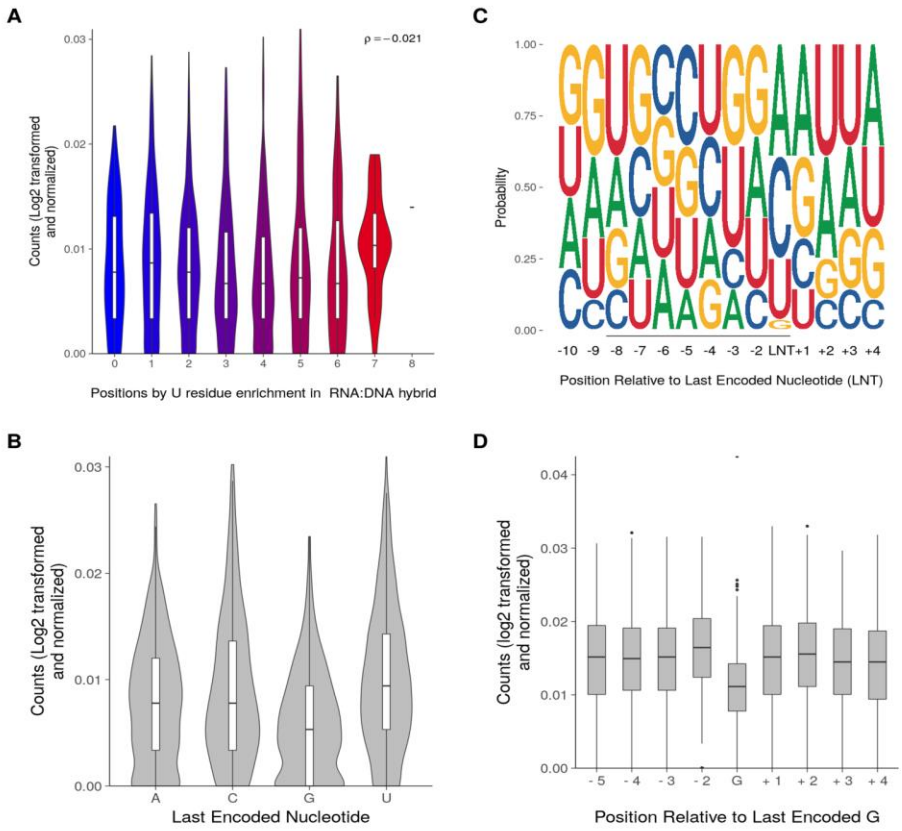

Figure S6. Repetition of Figure 8 Analysis Excluding Gene Residues. S6a) Violin plot of 3' spacer positions sorted by RNA:DNA hybrid U content, with incorporated box plots. Counts are log2 transformed and normalized. Correlation coefficient determined by spearman correlation test, p-value < 1*10-8. S6b) Violin plot with incorporated box plots of rDNA spacer position occupancies sorted by last encoded nucleotide. Comparison by Kruskal Wallis test produced a p-value < 1*10-8. S6c) Sequence logo of the RNA:DNA hybrid corresponding to top 3' spacer positions conserved in all three sets. S6d) box plot of spacer position occupancies proximal to last encoded G residues. Counts are log2 transformed and normalized. Comparisons by Mann-Whitney U test. ns = not significant, * = p-value < 5*10-6, *** = p-value < 5*10-16

DEFINING THE INFLUENCE OF THE A12 SUBUNIT ON TRANSCRIPTION ELONGATION AND TERMINATION BY RNA POLYMERASE I *IN VIVO*.

by

ANDREW M. CLARKE, CHAD M. PETIT, AND DAVID A. SCHNEIDER

In Preparation for the *Journal of Biological Chemistry*

Format adapted for dissertation

ABSTRACT

S. cerevisiae has approximately 200 copies of the 35S rRNA gene, arranged tandemly on chromosome XII. This gene is transcribed by RNA Polymerase I (Pol I) and the 35S rRNA transcript is processed to produce three of the four rRNAs required for ribosome biogenesis. An intergenic spacer (IGS) separates each copy of the 35S gene and contains the 5S gene (transcribed by RNA Polymerase III), an origin of DNA replication, and the promoter for the adjacent 35S gene. Efficient transcription termination by Pol I is required to prevent the collision of actively transcribing Pol I elongation complexes and the various DNA binding proteins found in the IGS. The A12.2 subunit of Pol I plays an important role in cleavage, termination, and nucleotide addition. Deletion of this subunit causes alteration of nucleotide addition kinetics, as well as read-through of transcription termination sites. To interrogate both of these phenomena we performed Native Elongating Transcript Sequencing (NET-seq) on a *rpa12Δ* strain of *S. cerevisiae* and evaluated the resultant change in Pol I occupancy throughout the 35S gene and the IGS. Compared to WT, we observed template sequence-specific changes in Pol I occupancy throughout the 35S gene. We also observed *rpa12Δ* Pol I occupancy downstream of both termination sites and throughout most of the IGS, including the 5S gene. Occupancy of *rpa12Δ* Pol I increased just upstream of the promoter proximal Reb1 binding site and dropped significantly after, implicating this site as a third terminator for Pol I transcription.

INTRODUCTION

Ribosome biogenesis in *S. cerevisiae* (yeast) begins with transcription of the 35S gene by RNA polymerase I (Pol I) to produce the 35S ribosomal RNA (rRNA), which is co- and post-transcriptionally cleaved to produce the 18S, 5.8S, and 25S rRNAs (Fig 1). These rRNAs, along with the 5S rRNA, which is synthesized by RNA Polymerase III, form the RNA backbone of ribosomes. Rapidly dividing yeast cells have a tremendous demand for ribosomes, which is reflected by the fact that more than 60% of the cell's RNA is ribosomal RNA[1]. Yeast has between 150 and 200 copies of the 35S gene to accommodate this demand, arranged tandemly on chromosome XII [1]. Half of these copies are transcriptionally active, and each actively transcribed copy is host to approximately 50 Pol I transcription elongation complexes (ECs) [2]. In between each copy is an intergenic spacer (IGS). The IGS consists of the 5S gene flanked by intergenic spacers 1 and 2 (IGS 1 and IGS 2) (Fig 1). IGS 1 contains the second of two termination sites of Pol I transcription as well as a replication fork blocking (RFB) region composed of three Fob1 binding sites [3] (Fig 1, red brackets). It also contains a bi-directional Pol II promoter (E-pro) [4]. IGS 2 contains an autonomously replicating sequence (rARS) as well as the upstream and core promoter elements for the adjacent 35S gene [5] (Fig 1, black brackets). RNA polymerase II (Pol II) transcription from E-pro (normally suppressed by silencing protein Sir2) is implicated in rDNA repeat number expansion by preventing cohesin association with the rDNA [4]. The RFB region in IGS 1 prevents DNA replication in the opposite direction of Pol I transcription to prevent collision of transcription and replication machineries [6]. These collisions result in DNA damage as evidenced by Pol I transcription-dependent increased copy number variation

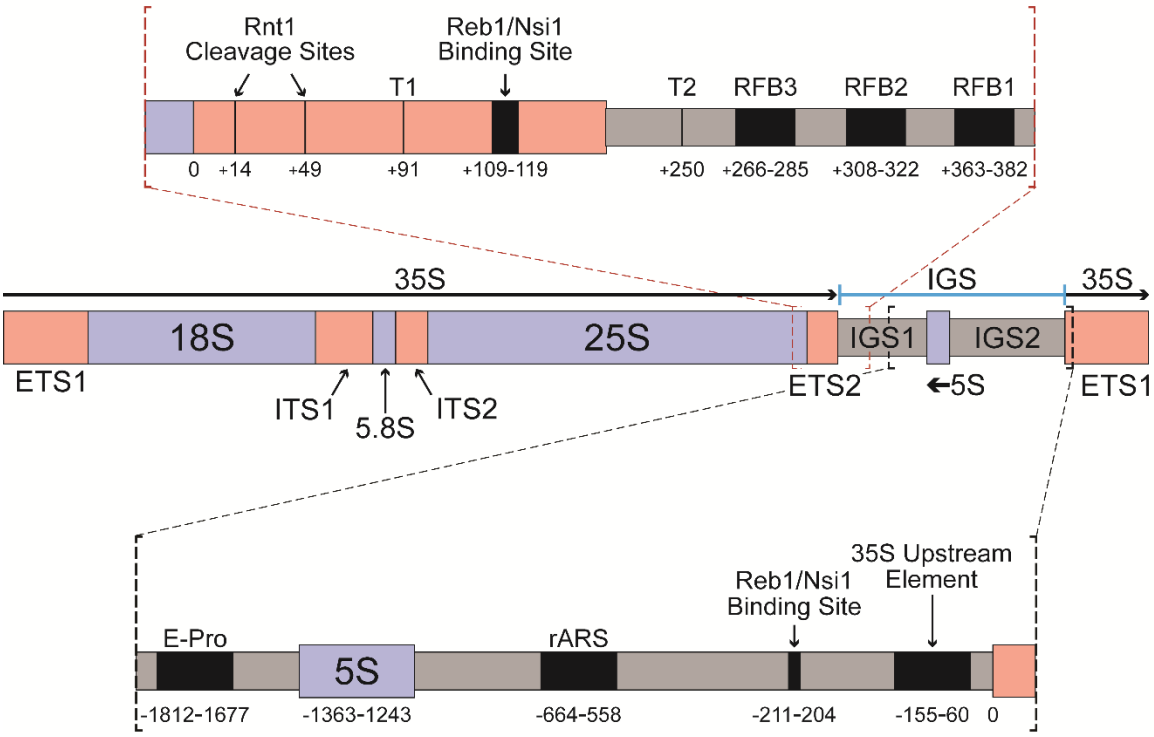

Figure 1. The *S. cerevisiae* rDNA repeat. Gene diagram of *S. cerevisiae* rDNA repeat. Black arrows indicate direction of transcription. red brackets) Portion of the rDNA containing the two sites of Pol I termination as well as the RFB region, feature coordinates are relative to the mature end of the 25S rRNA. black brackets) Portion of the rDNA containing the E-pro, 5S gene and IGS 2, feature coordinates relative to the transcription start site of the downstream 35S repeat.

and generation of extra-ribosomal circles (ERCs) in low RFB activity yeast strains [7]. Given that the RFB region lies just downstream of the two terminator regions for Pol I, efficient Pol I transcription termination is therefore also crucial to preventing collisions between transcription and replication machineries.

Early studies identified two sites of Pol I transcription termination [8]. The first site (referred to as T1) is 91 base pairs downstream of the mature end of the 25S rRNA, and represents the main site of Pol I transcription termination [9]. T1 lies at the end of a 20 bp tract of A/T-rich template rDNA. The binding site for the DNA binding protein Reb1 is located downstream of this A/T-rich DNA sequence. In combination with Reb1 these two sequence elements were sufficient to terminate Pol I transcription *in vitro* [10]. The proposed mechanism for termination by Pol I consisted of a Reb1-mediated "pause" followed by "release" of the nascent transcript due to the A-rich template DNA tract [11]. Previous studies identified that murine transcription termination factor I (mTTF-I) could also terminate yeast Pol I transcription *in vitro* [12]. Extensive sequence similarity between mTTF-I and Reb1 [13] supported this mechanism and established the possibility of a universally conserved mechanism of Pol I transcription termination in all eukaryotes. However, subsequent analyses indicated that the mechanism of Pol I termination in yeast is more complicated. *In vivo* studies implicated a host of new factors in Pol I termination, including RNase III Rnt1 [9] and 5' exonuclease Rat1 [14]. Other studies have cast doubt on Reb1's role in yeast Pol I termination [15, 16]. Reb1 has an additional binding site downstream of T1, 210 bp upstream of the transcription start site for the adjacent 35S gene (Fig 1, black brackets). It has long been known that Reb1's binding affinity for this promoter-proximal site is greater than for the terminator-proximal site near T1 [17]. More

recent Chromatin Endogenous Cleavage studies have demonstrated that Nsi1 (a Reb1 homologue) has a much stronger affinity for the terminator-proximal Reb1 binding site, and is required for efficient Pol I termination *in vivo* [16], unlike Reb1 [15]. Interestingly, Reb1 is also implicated in the initiation and termination of Pol II transcription in yeast. Reb1 binding at sites within a subset of Pol II promoters is required for the formation of Nucleosome Free Regions (NFRs), allowing transcription of those genes [18, 19]. Reb1 binding to sites downstream of Pol II genes has also been shown to pause Pol II transcription *in vivo* in a polar manner, prompting ubiquitination and subsequent Pol II transcription termination [20], similar to its originally suggested role in Pol I transcription.

The current model for Pol I termination at T1 is similar to the "torpedo" model for Pol II termination. Pol I transcribes to position +91 downstream of the mature end of the 25S rRNA and halts upstream of the binding site for Nsi1. Concurrently, Rnt1 recognizes a stem loop formed in the nascent RNA transcript just downstream of the 25S end, which cleaves between positions +14/+15 and positions +49/+50 [21]. This cleavage produces a 5' monophosphate terminus on the nascent transcript which is the appropriate substrate for Rat1 [22]. As in Pol II termination, Rat1 is proposed to be the "release factor" for Pol I; when Rat1 reaches the stalled Pol I EC, termination occurs [14]. However, no mechanism has been put forth describing how this release occurs. Termination at the "failsafe" site downstream of T1, known as T2, is less well-studied, though it has been shown to involve RFB region binding protein Fob1 and Rat1 [14]. Previous work also suggests an additional termination site adjacent to the promoter-proximal binding site for Reb1 [23].

One of the known Pol I subunits involved in both intrinsic cleavage [24] and transcription termination [25] is Rpa12 (A12.2). This subunit is homologous to Rpb9 in Pol II, and Rpc11 in Pol III. The C-terminal domain of A12.2 also bears homology to Pol II cleavage factor TFIIS [26]. *rpa12Δ* strains grow very slowly compared to wild type at 30°C [27] and are Pol I transcription termination deficient [25]. Structural studies of the interaction between Reb1 and Pol I ECs in *Schizosaccharomyces pombe* revealed that protein-protein interactions between Reb1 and the A12.2 subunit are specifically required for transcriptional arrest and termination of Pol I ECs [28], possibly explaining the previously identified orientation dependence in Reb1 transcriptional arrest activity with respect to *S. pombe* Pol I [29] and *S. cerevisiae* Pol II [20]. In addition to its roles in cleavage and termination, recent work from our lab demonstrated that A12.2 also affects stability of the Pol I EC [30] and is directly involved in nucleotide addition by Pol I [31, 32].

Given A12.2's involvement in nucleotide addition, cleavage, and termination, we set out to investigate how deletion of this subunit perturbs Pol I occupancy *in vivo* using Native Elongating Transcript Sequencing (NET-seq.) This technique allows us to observe reproducible Pol I occupancy with single-nucleotide resolution [33]. Upon deletion of the this gene, we observe re-ordering of median polymerase occupancy throughout the 35S gene, favoring the 5' end of the transcript. Furthermore, we identified key changes in Pol I EC occupancy in response to upstream nucleotide enrichment between the WT and *rpa12Δ* strains, confirming that A12.2 affects Pol I nucleotide addition *in vivo.* Our experiments also confirmed that the *rpa12Δ* strain is termination-deficient *in vivo,* with significant Pol I occupancy observed downstream of T1, T2, and the 5S gene. Strikingly,

median occupancy increases dramatically just upstream of the promoter-proximal Reb1 binding site, indicating that the *rpa12Δ* Pol I EC is paused or halted stably at this site. These findings provide insight into A12.2's role in growing cells and pose intriguing new questions about Pol I termination and the conserved genomic architecture of the IGS.

RESULTS

RPA12 deletion alters Pol I occupancy throughout the 35S gene

We performed NET-seq in triplicate on WT and *rpa12Δ S. cerevisiae* strains bearing HA-tagged Pol I as previously described [33]. The resultant reads were mapped to the 35S gene(Fig 2A.) The amplitude at each position corresponds to the number of polymerases which mapped to that position, representing a snap shot of Pol I occupancy. The counts were normalized to the sum of reads mapping to the 5S gene. These rRNAs are synthesized by Pol III and represent a previously-identified contaminant in Pol I NET-seq library generation. These counts were assessed for statistically significant increase (blue) or decrease (red) in the *rpa12Δ* strain relative to WT. Qualitative analysis of these data reveals a 5' shift of polymerase occupancy, with increased occupancy in ETS 1, 18S, and both ITS regions in the mutant strain. Analysis of median occupancy confirms these findings, revealing statistically significant increases in occupancy in the above regions, along with ETS 2 (Fig 2B). To characterize the influence of DNA sequence on Pol I occupancy, we created sequence logos of the positions with significantly increased or decreased occupancy to assess whether there are conserved sequence elements that correlate with these changes

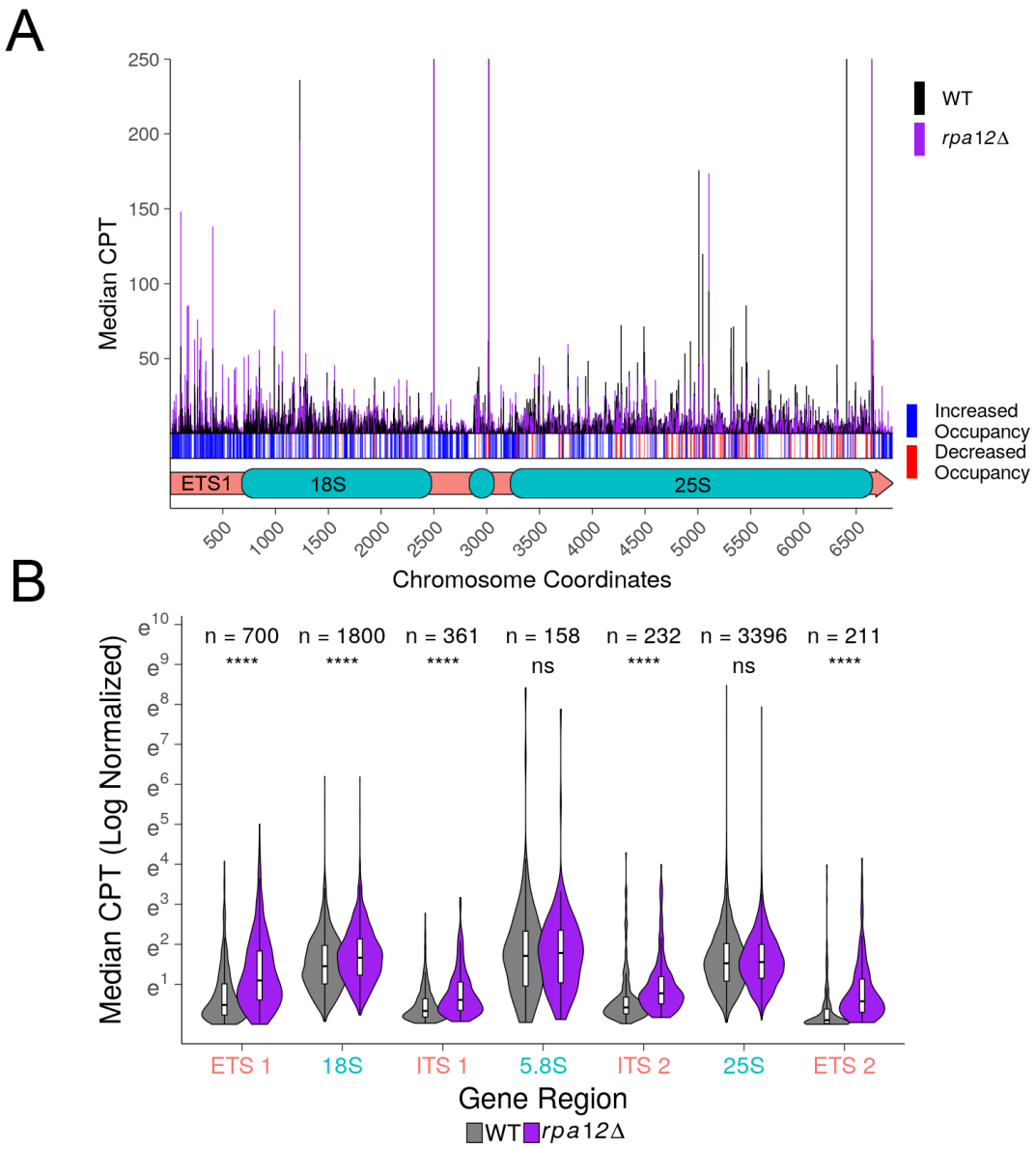

Figure 2. *RPA12* deletion shifts Pol I occupancy towards the 5' end of the 35S gene. 2A) Pol I NET-seq median 3' end densities for WT (black) and *rpa12Δ* (Purple) strains. Red underlines indicate statistically significant decreases in occupancy in *rpa12Δ* strain compared to wild type, Blue underlines indicate statistically significant increases in occupancy (Student's T test, n = 3, P value < 0.01). 35S gene diagram is color coded for gene (cyan) and spacer (salmon) regions. Counts reported in counts per thousand, normalized to the sum of 5S signal. 2B) Median densities for each strain in each region of the 35S gene. Comparison by Wilcoxon test (n noted above each comparison, **** = P value < $5*10^{-5}$). Counts reported in counts per thousand, normalized to the sum of 5S signal and natural log transformed.

(Fig 3). Increased U and A residue enrichment in the RNA:DNA hybrid correlates with significantly increased Pol I occupancy in the *rpa12Δ strain* (Fig 3A). In examining sequence context from decreased occupancy positions (Fig 3B) we observed increased U and A residue enrichment directly downstream of the last nucleotide added by the EC (LNT). These finding suggest that *rpa12Δ* alters Pol I's response to A and U residues both in the RNA-DNA hybrid, and downstream DNA.

These sequence logos are aggregates of sequence context for a small subset of positions within the 35S gene. In order to examine possible relationships between Pol I occupancy and sequence context gene-wide, we assessed polymerase occupancy as a factor of specific nucleotide enrichment in the RNA:DNA hybrid (Fig 4) and the eight base pairs immediately downstream (Fig 5) of each position in the 35S gene. Deletion of *RPA12* alters Pol I's response to increasing A/T and C/G enrichment in the RNA:DNA hybrid differently. For A and U residues, the occupancy difference peaks in the middle range of enrichment (between 2 and 4 residues in the RNA:DNA hybrid) (Figs 4A and 4D). For G and C residues, the occupancy difference is greatest at low levels of enrichment (between 0 and 2 residues in the RNA:DNA hybrid) (Figs 4B and 4C), with little effect at higher G and C enrichment. Pol I occupancy changes in response to downstream nucleotide enrichment is more consistent for all four nucleotides, *RPA12* deletion increases Pol I occupancy in response to modest nucleotide enrichment (fewer than 4 nucleotides in the 8 bp window downstream of the last encoded nucleotide) for all four residues (Fig 5). These data show cumulatively that *rpa12Δ* Pol I is differentially affected by increasing nucleotide enrichment both up- and downstream, indicating that

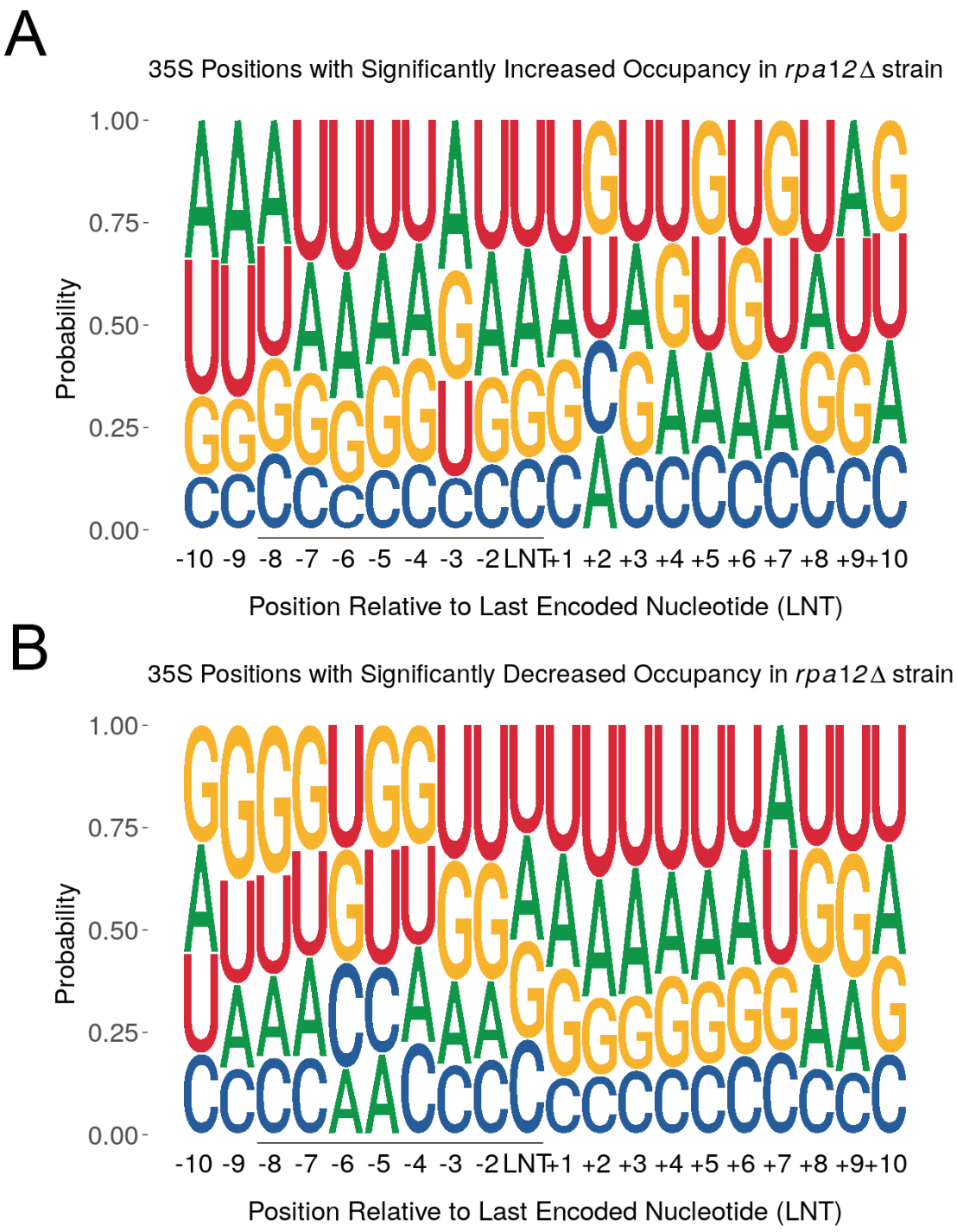

Figure 3. Increased and decreased occupancy positions have distinct sequence features. 3A) Sequence logo for all positions with significantly increased intra-region normalized mean occupancy in *rpa12Δ* strain (student's T test, P value < $5*10^{-3}$). 3B) Sequence logo for all positions with significantly decreased intra-region normalized mean occupancy in *rpa12Δ* strain (student's T test, P value < $5*10^{-3}$).

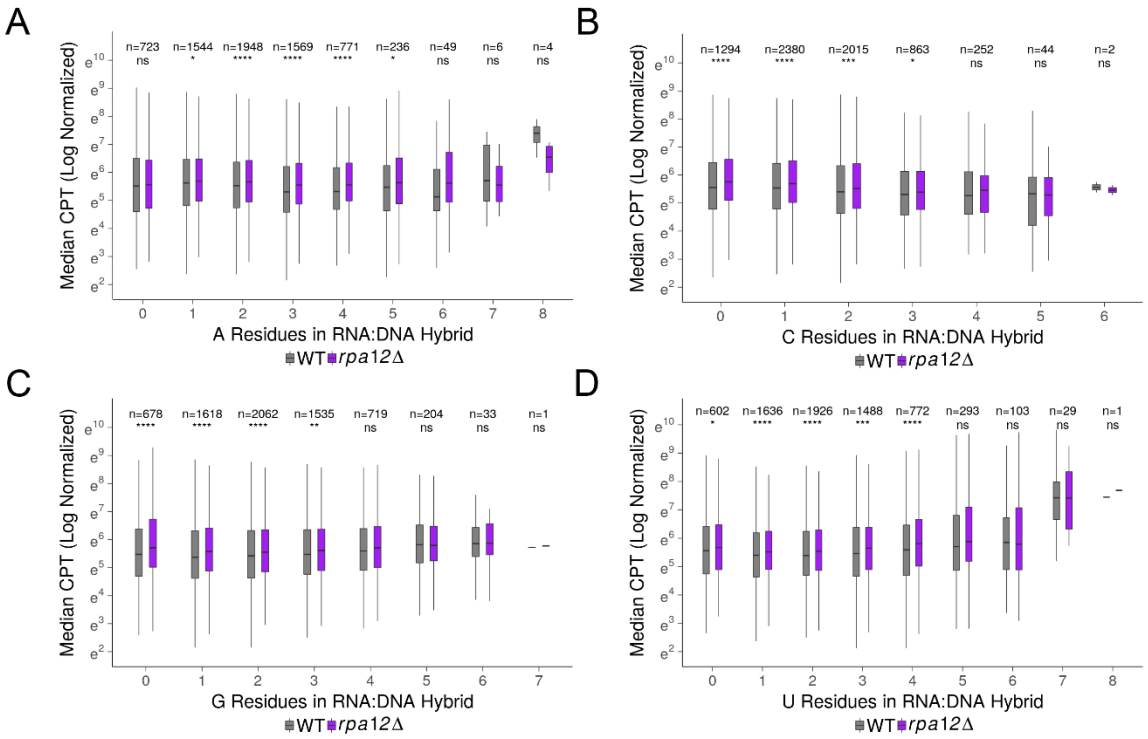

Figure 4. *RPA12* deletion alters relationship between Pol I occupancy and nucleotide enrichment in the RNA:DNA hybrid. Violin plots comparing median occupancy of WT (grey) and *rpa12Δ* (purple) Pol I strains by enrichment for 4A) A, 4B) C, 4C) G, 4D) U residues in the RNA:DNA hybrid (eight base pair region upstream of the 3' transcript end) for each position. Comparison by Wilcoxon test (n noted above each comparison; * = P value < $5*10^{-2}$, ** = P value < $5*10^{-3}$, *** = P value < $5*10^{-4}$, **** = P value < $5*10^{-5}$). Counts reported in counts per thousand, intra-region normalized and natural log transformed.

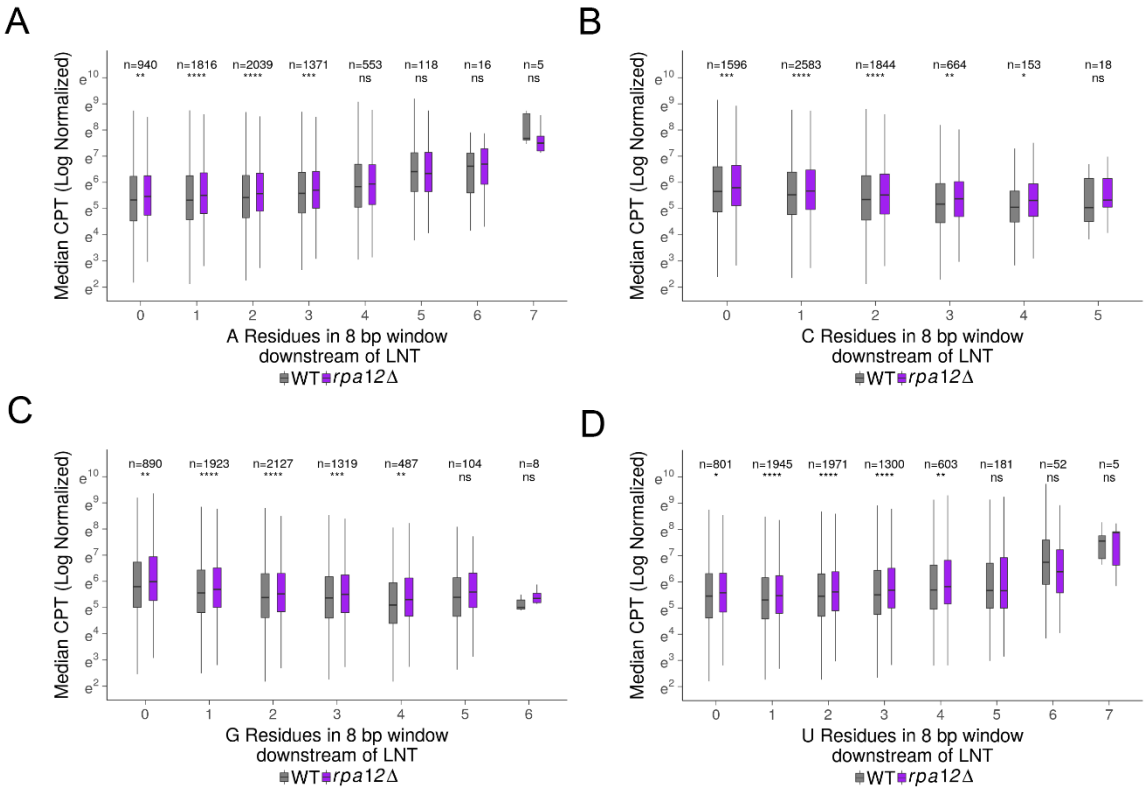

Figure 5. *RPA12* deletion alters relationship between Pol I occupancy and nucleotide enrichment downstream of the last encoded nucleotide. Violin plots comparing median occupancy of WT (grey) and *rpa12Δ* (purple) Pol I strains by enrichment for 5A) A, 5B) C, 5) G, 5D) U residues in the 8 bp region directly downstream of each position. Comparison by Wilcoxon test (n noted above each comparison; * = P value < $5*10^{-2}$, ** = P value < $5*10^{-3}$, *** = P value < $5*10^{-4}$, **** = P value < $5*10^{-5}$). Counts reported in counts per thousand, intra-region normalized and natural log transformed.

deletion of *RPA12* perturbs nucleotide addition by Pol I *in vivo*, as suggested previously based on *in vitro* studies [31].

Deletion of RPA12 Results in Significant Occupancy Downstream of T1 and T2

Previous work has demonstrated significant read through of transcriptional terminators by *rpa12Δ* Pol I [25]. In order to characterize this finding in more detail we examined median polymerase occupancy both up- and downstream of T1 and T2 (Fig 6A, green arrows). As expected, we observed increased WT Pol I occupancy just upstream of each terminator site, followed by significant decrease directly downstream (Figs 6B (top panel), 6D (top panel)). This pattern is consistent with the initial transcriptional pause required for termination of Pol I transcription. In contrast, we do not see any such buildup and subsequent decrease for *rpa12Δ* Pol I occupancy with respect to either terminator site (Figs 6B (bottom panel), 6D (bottom panel)), indicating that *rpa12Δ* Pol I is not pausing at these sites like WT Pol I.

We also observed persistent *rpa12Δ* Pol I occupancy throughout IGS 1, the 5S gene, and IGS 2 (Fig 6C). The detection of these ECs via NET-seq requires that they be associated with a nascent transcript, indicating that these polymerases are actively transcribing. These data further suggest that the elongation rate of Pol I is greater than the progressive exonucleotide excision rate of Rat1. Alternatively, Rat1 binding to the nascent transcript may rely on transcriptional pausing at T1 or assembly of precise RNA structures that may be depleted in the mutant strain. Interestingly, WT Pol I signal was

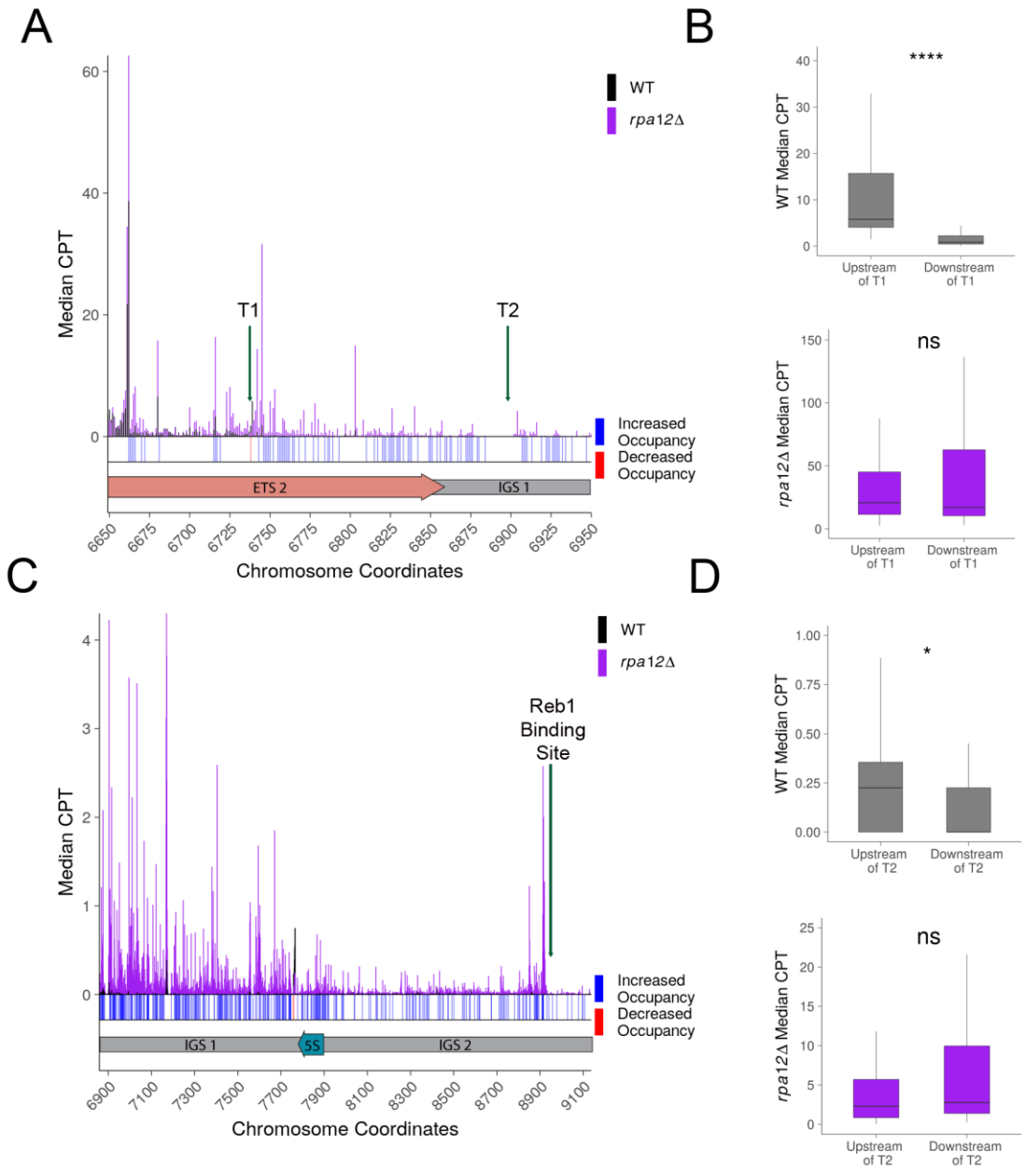
A
B
C
D
Median CPT
Chromosome Coordinates
WT
rpa12Δ
T1
T2
ETS 2
IGS 1
Increased Occupancy
Decreased Occupancy
Reb1 Binding Site
5S
IGS 2
WT Median CPT
rpa12Δ Median CPT
****
ns
*
Upstream of T1
Downstream of T1
Upstream of T2
Downstream of T2

Figure 6. *RPA12* deletion reveals a putative third Pol I transcription termination site in the IGS. 6A) Pol I NET-seq median occupancy for WT (black) and *rpa12Δ* (purple) strains in ETS 2 and IGS 1. Termination sites are highlighted with green arrows. Red underlines indicate statistically significant decreases in occupancy in *rpa12Δ* strain compared to wild type, Blue underlines indicate statistically significant increases in occupancy (Student's T test, n = 3, P value < 0.01). Gene diagram is color coded for ETS 2 (salmon) and IGS 1 (grey). Counts reported in counts per thousand, normalized to the sum of 5S signal. 6B) Comparison of local median occupancy 50 bp upstream and downstream of T1 for WT (top panel) and *rpa12Δ* (bottom panel) strains. Comparison by Mann-Whitney U test; ns = not significant, **** =P value < $1*10^{-16}$. 6C) Pol I NET-seq median occupancy for WT and *rpa12Δ* strains in IGS. Promoter Proximal Reb1 binding site is highlighted with a green arrow. Red underlines indicate statistically significant decreases in occupancy, Blue underlines indicate statistically significant increases in occupancy (Student's T test, n = 3, P value < 0.01). Gene diagram is color coded for IGSs 1 and 2 (grey) and the 5S gene (cyan). Counts reported in counts per thousand, normalized to the sum of 5S signal. 6D) Comparison of local median occupancy 50 bp upstream and downstream of T2 for WT (top panel) and *rpa12Δ* (bottom panel) strains. Comparison by Mann-Whitney U test; ns = not significant, * =P value < 0.05).

observed downstream of T2, directly upstream (relative to Pol I transcription) of the 5S gene. Sporadic, low levels of WT Pol I signal were also observed throughout IGS 2.

These data indicate that WT Pol I ECs infrequently read through both the primary termination site (T1) and the secondary fail-safe site (T2), ultimately halting upon encountering the 3' end of the 5S gene (6D). The sporadic signal observed downstream of this site shows that some Pol I ECs transcribe through this region, though these polymerases appear to represent an extremely small fraction of all transcribing Pol I. By contrast, *rpa12Δ* results in significantly increased Pol I EC read-through of T1 and T2, as well as the 5S gene. Our lab has identified that the *rpa12Δ* Pol I EC is much more stable than WT *in vitro* [30] and it is likely this additional stability allows *rpa12Δ* Pol I ECs to read through the Pol III EC-occupied 5S gene. Compared to WT we see much more consistent *rpa12Δ* Pol I occupancy throughout IGS 2, culminating in a cluster of positions with sharply increased occupancy approximately 215 base pairs upstream of the transcription start site for the adjacent 35S gene (Fig 6D). Interestingly, this site is also just upstream of the promoter-proximal binding site for Reb1 (Fig 6D, green arrow). Furthermore, this site lies at the end of an A/T rich template DNA tract, similar to T1. These data suggest that this position represents a third site for Pol I EC pausing/arrest, and possibly termination.

DISCUSSION

RPA12 deletion alters Pol I occupancy within the 35S gene

Initial characterization of the mechanistic contribution of A12.2 to Pol I function focused on transcript cleavage and termination [24-26]. However, recent studies of Pol I

nucleotide addition *in vitro* showed that removal of the A12.2 subunit from the Pol I EC changes the kinetics of nucleotide incorporation [30, 31], indicating that A12.2 plays a role in transcription elongation as well as termination and transcript cleavage. Analysis of our NET-seq data corroborate these findings. NET-seq revealed substantial shifts in *rpa12Δ* Pol I occupancy throughout the 35S gene, including increased occupancy in the 5' end of the gene relative to WT (Figs 2A, 2B). By preparing these libraries in triplicate we determined that many of these changes in occupancy were statistically significant (Fig 2B). These data are consistent with transcription run-on data published by the Proudfoot lab [25], which also revealed a 5' bias in *rpa12Δ* Pol I EC distribution. The lack of cleavage activity may play a role in this observed 5' occupancy shift. Reductions in TFIIS-mediated transcript cleavage by Pol II have been shown to impair yeast viability and Pol II transcription elongation *in vivo* [34]. The permanent integration of the TFIIS paralogue A12.2 into Pol I suggests that cleavage activity is important for the proper function of Pol I. Deletion of *RPA12* induces a two-fold decrease in the growth rate of yeast. Pol I ECs are also very densely packed on each active rDNA repeat, often exceeding 50 ECs per 35S gene [1]. Thus, stable *rpa12Δ* polymerase stalling on the rDNA template (coupled with the increased stability of the *rpa12Δ* Pol I EC [30, 35]) may represent a substantial barrier for Pol I ECs in the 5' end of the gene. This accumulation of "road blocks" might explain the observed accumulation of *rpa12Δ* Pol I ECs in the 5' end of the 35S gene. Previous studies on head-to-tail collision events between Pol II ECs have indicated that collision with the leading polymerase causes the trailing polymerase to backtrack *in vitro*, relying on TFIIS-mediated cleavage to continue productive elongation [36]. This finding for Pol II is consistent with our observations,

and similar analyses of both WT and *rpa12Δ* Pol I *in vitro* should further elucidate this phenomenon.

The positions with statistically significant changes in occupancy displayed unique sequence trends both within the RNA:DNA hybrid, and directly downstream (Figs 3A, 3B). Some of these trends were also observed gene-wide (Figs 4, 5). The large-scale change in Pol I occupancy observed above suggests that nucleotide addition is perturbed by *RPA12* deletion *in vivo* as was previously observed *in vitro* [31]. Furthermore, the data indicate that this effect on nucleotide addition is at least partially template sequence specific, as demonstrated by the distinct sequence elements proximal to the increased and decreased occupancy positions, as well as the differing responses to proximal template nucleotide enrichment upon *RPA12* deletion. Additional mutational analysis should further elucidate the role that the A12.2 subunit plays in nucleotide addition. For example, the C-terminal domain (CTD) of A12.2 is responsible for conferring intrinsic cleavage activity to Pol I [24], whereas the N-terminal domain is responsible for anchoring A12.2 within Pol I [26]. By analyzing pol I occupancy in A12.2 CTD deletion yeast strains, we could determine the specific contributions of the transcript cleavage and anchoring domains to the perturbations in Pol I occupancy observed above.

S. cerevisiae contains a putative third Pol I Terminator region

Our NET-seq data confirmed that deletion of *RPA12* rendered Pol I termination deficient *in vivo*, as we observed significant *rpa12Δ* Pol I EC readthrough of both previously defined sites of Pol I termination (Figs 6A, 6B, 6D). Furthermore, we observed dramatically increased *rpa12Δ* Pol I occupancy throughout the IGS compared to

WT (Fig 6C). Interestingly, *rpa12Δ* Pol I occupancy increased greatly at a site just upstream of the promoter-proximal Reb1 binding site, indicating *rpa12Δ* Pol I is pausing or arresting at this point. Reb1 has a higher affinity for this site than its terminator-proximal binding site, and appears capable of halting *rpa12Δ* Pol I, while Nsi1 binding to the T1-proximal site is not (Fig 6B) [16]. As NET-seq is not sensitive to termination events, we cannot conclude that this site induces termination of Pol I transcription; however, Pol I transcription termination involves pausing of the EC as the first step [6]. These data raise the question of whether this site represents an even stronger Pol I terminator than T1. This third site could function as a final failsafe, preventing the collision of Pol I ECs and the transcription initiation machinery bound to the promoter region of the adjacent 35S repeat. This arrangement has been observed in higher eukaryotes such as *Xenopus laevis* [37] and mice [38]. However, our data do not suggest that WT Pol I reaches this site *in vivo* under normal conditions. Why then has this site been preserved, if it is unnecessary under normal growth conditions? In the WT strain, a population of polymerases read through both T1 and T2, resulting in Pol I occupancy near the 3' end of the 5S gene. Under our growth conditions we did not observe consistent WT signal beyond the 5S gene, suggesting that the WT Pol I EC cannot transcribe past actively transcribed Pol III repeats. It is reasonable to suggest that Pol III ECs on the 5S gene interfere with Pol I transcription, resulting in Pol I transcriptional arrest. Perhaps in slower growth conditions or under stress when ribosome synthesis is reduced, Pol III loading on the 5S genes would be decreased, increasing the probability of Pol I EC read-through into IGS 2. Under such conditions a promoter proximal terminator

site would prevent Pol I EC collisions with transcription initiation factors at the next 35S gene.

Analysis of RNA Polymerase I occupancy in the ribosomal DNA via NET-seq refines our understanding of the role for A12.2 in Pol I function *in vivo*, as well as the architecture of the *S. cerevisiae* IGS. At the same time, these data suggest new lines of inquiry. How does A12.2 moderate Pol I occupancy in a sequence-dependent manner? Which domain of A12.2 is responsible for this behavior? Does the promoter-proximal binding site represent a third site of Pol I termination, and if so, what purpose does it serve in *S. cerevisiae*? How well conserved are these termination sites in other eukaryotic species? Answering these questions will bring us closer to a fundamental understanding of Pol I, and the genomic region it transcribes.

METHODS AND MATERIALS

Native Elongating Transcript Sequencing (NET-seq) Library Preparation and Sequencing

NET-seq library generation was performed as previously described using the original HA-tagged Rpa135 *S. cerevisiae* strain [33] and a derivative of this strain carrying a full deletion of *RPA12*. In brief, early log-phase cultures were collected via filtration, then flash-frozen in liquid nitrogen and cryogenically lysed by mixer mill. Pol I elongation complexes were isolated via immunoprecipitation with anti-HA magnetic beads. Associated RNA transcripts were then isolated using an acidic phenol-chloroform extraction. The RNA transcripts were then 3’ appended with a DNA linker and reverse transcribed to produce complementary DNAs (cDNAs). The cDNAs were circularized

and amplified by PCR to produce high-throughput sequencing libraries. These libraries were then sequenced as previously described [33].

NET-seq data formatting

Sequencing trimming, alignment, and data formatting were performed as previously described [33]. Unix scripts available upon request.

NET-seq data analysis

The three replicates for the WT and *rpa12Δ* yeast strains were aggregated. The counts for each position in the rDNA repeat represents the median signal on the negative strand for the three replicates of each strain. For the data in figures 2 and 6, the counts were normalized to the sum of all signal on the positive strand of the 5S gene for each replicate. For the data in figures 3-5, the counts in each region were normalized to the sum of all signal on the negative strand in that region. Analyses were performed in R (ver. 3.4.2) [39], using the following packages: dplyr (ver. 0.7.5) [40], plyr (ver. 1.8.4) [41], ggplot2 (ver. 2.2.1) [42], ggseqlogo (ver. 0.1) [43], ggpubr (ver. 0.1.7) [44], cowplot (ver. 0.9.2) [45], devtools (ver. 2.0.1) [46], matrixStats (ver. 0.53.1) [47], metaMA (ver. 3.1.2) [48], and scales (ver. 0.5.0) [49]. R scripts available upon request**.**

ACKNOWLEDGEMENTS

We thank the members of the Heflin Center for Genomic Sciences for assistance with sequencing. We also thank the current and past members of the Schneider lab for

thoughtful critique and discussion of this work. This study was funded by NIH Grant #GM084946 to D.A.S.

REFERENCES

1. Woolford, J.L., and S.J. Braserga, *Ribosome Biogenesis in the Yeast Saccharomyces cerevisiae.* Genetics, 2013. **195**(3): p. 643 - 681.

2. French, S.L., Y.N. Osheim, F. Cioci, M. Nomura, and A.L. Beyer *In Exponentially Growing Saccharomyces cerevisiae Cells, rRNA Synthesis is Determined by the Summed RNA Polymerase I Loading Rate Rather than by the Number of Active Genes.* Mol Cell Biol, 2003. **23**(5): p. 1558-1568.

3. Kobayashi, T., *The Replication Fork Barrier Site Forms a Unique Structure with Fob1p and Inhibits the Replication Fork.* Mol Cell Biol, 2003. **23**(24): p. 9178-9188.

4. Kobayashi, T., and A. R. D.Ganley, *Recombination Regulation by Transcription-Induced Cohesin Dissociation in rDNA Repeats.* Science, 2005. **309**(5740): p. 1581 - 1584.

5. Nomura, M., *Transcription factors used by Saccharomyces cerevisiae RNA Polymerase I and the mechanism of initiation*, in *Transcription of ribosomal RNA genes by eukaryotic RNA polymerase I*, M.R. Paule, Editor. 1998, Springer-Verlag: Berlin, Germany.

6. Nemeth, A., J. Perez-Fernandez, P Merkl,S. Hamperl, J.Gerber, J. Griesenbeck, and H. Tschochner, *RNA Polymerase I Termination: Where is the End?* . Biochimica et Biophysica Acta- Gene Regulator Mechanisms, 2013. **1829**(3 - 4): p. 306 - 317.

7. Takeuchi, Y., T.Horiuchi, and T. Kobayashi, *Transcription-Dependent Recombination and the Role of Fork Collision in Yeast rDNA.* Genes Dev. , 2003. **17**(12): p. 1497 - 1506.

8. Lang, W.H., and R.H. Reeder, *The REB1 Site is an Essential Component of a Terminator for RNA Polymerase I in Saccharomyces cerevisiae.* Molecular and Cell Biology, 1993. **13**(1): p. 649 - 658.

9. Reeder, R.H., P. Guevara, and J.G. Roan, *Saccharomyces cerevisiae RNA Polymerase I Terminates Transcription at the Reb1 Terminator In Vivo.* Mol Cell Biol. , 1999. **19**(11): p. 7369 - 7376.

10. Lang, W.H., and R.H. Reeder, *Transcription Termination of RNAPolymerase I Due to a T-rich Element Interacting with Reb1p.* Proc. Natl. Acad. Sci. USA, 1995. **92**(92): p. 9781- 9785.

11. Lang, W.H., B.E. Morrow, Q. Ju, J.R. Warner, and R.H. Reeder, *A Model for Transcription Termination by RNA Polymerase I.* Cell, 1994. **79**: p. 527-534.

12. Kuhn, A.E., I. Bartsch, and I. Grummt, *Specific Interaction of the Murine Transcription Termination Factor TTF I With Class-I RNA Polymerases.* Nature, 1990. **344**: p. 559 - 562.

13. Evers, R., A. Smid, U. Rudloff, F. Lottspeich, and I. Grummt, *Different Domains of the Murine RNA Polymerase I-Specific Termination Factor mTTF-I Serve Distinct Functions in Transcription Termination* Embo J. , 1995. **14**(6): p. 1248 - 1256.

14. El Hage, A., M. Koper, J. Kufel, and D. Tollervey, *Efficient Termination of Transcription by RNA Polymerase I Requires the 5' Exonuclease Rat1 in Yeast.* Genes Dev., 2008. **22**(8): p. 1069 - 1081.

15. Kawauchi, J., H. Mischo, P. Braglia, A. Rondon, and N.J. Proudfoot, *Budding Yeast RNA Polymerases I and II Employ Parallel Mechanisms of Transcriptional Termination.* Genes Dev., 2008. **22**(8): p. 1082 - 1092.

16. Reiter, A., S. Hamperl, H. Seitz, P. Merkl, J. Perez-Fernandez, L. Williams, J. Gerber, A. Nemeth, I. Leger, O. Gadal, P. Milkereit, J. Griesenbeck, and H. Tschochner, *The Reb1-homologue Ydr026c/Nsi1 is Required for Efficient RNA Polymerase I Termination in Yeast.* Embo J. , 2012. **31**(16): p. 3480 - 3493.

17. Morrow, B.E., Q. Ju, and J.R. Warner, *Purification and Characterization of the Yeast rDNA Binding Protein REB1.* Journal of Biological Chemistry, 1990. **265**: p. 20778 - 20783.

18. Raisner, R.M., P. D. Hartley, M. D. Meneghini, M. Z. Bao, C. L. Liu, S.L. Schreiber, O. J. Rando, and H. D. Madhan, *Histone Variant H2A.Z Marks the 5' Ends of Both Active and Inactive Genes in Euchromatin.* Cell, 2005. **123**(2): p. 233 - 248.

19. P. D. Hartley, a.H.D.M., *Mechanisms that Specify Promoter NucleosomeLocationand Identity.* Cell, 2009. **137**(3): p. 445 -458

20. Colin, J., T. Candelli, O. Porrua, J. Boulay, C. Zhu, F. Lacroute, L. M. Steinmetz, and D. Libri *Roadblock Termination by Reb1p Restricts Cryptic and Readthrough Transcription.* Mol Cell, 2014. **56**(5): p. 667 - 680.

21. Kufel, J., B. Dichtl, and D. Tollervey, *Yeast Rnt1p is Required for Cleavage of the Pre-Ribosomal RNA in the 3' ETS but not in the 5' ETS.* . RNA, 1999. **5**(7): p. 909-917.

22. Stevens, A., and T.L. Poole, *5'Exonuclease-2 of Saccharomyces cerevisiae Purification and Features of Ribonuclease Activity with Comparison to 5'-Exonuclease-1.* Journal of Biological Chemistry, 1995. **270**: p. 16063 - 16069.

23. van der Sande, C.A.F.M., T. Kulkens, A.B. Kramer, I.J. de Wijs, H. van Heerikhuizen, J. Klootwijk, and R.J. Planta, *Termination of Transcription by Yeast RNA Polymerase I.* Nucleic Acids Research, 1989. **17**(22): p. 9127 - 9146.

24. Kuhn, C.-D., S.R. Heiger, S. Baumli, M. Gartmann, J. Gerber, S. Jennebach, T. Mielke, H. Tschochner, R. Beckmann, and P. Cramer, *Functional Architecture of RNA Polymerase I.* Cell, 2007. **131**(7): p. 1260 - 1272.

25. Prescott, E.M., Y.N. Osheim, H.S. Jones, C.M. Alen, J.G. Roan, R.H. Reeder, A.L. Beyer, and N.J. Proudfoot *Transcriptional Termination by RNA Polymerase I Requires the Small Subunit Rpa12P.* Proc. Natl. Acad. Sci. USA, 2004. **101**(16): p. 6068 - 6073.

26. Van Mullem, V., E. Landrieux, J. Vandenhaute, and P. Thuriaux, *Rpa12p, a ConservedRNA Polymerase I Subunit with Two Functional Domains.* Molecular Microbiology, 2002. **43**(5): p. 1105 - 1113.

27. Nogi, Y., R. Yano, J. Dodd, C. Carles, and M. Nomura, *Gene RRN4 in Saccharomyces cerevisiae Encodes the A12.2 Subunit of RNA Polymerase I and Is Essential Only at High Tempuratures.* Mol Cell Biol, 1993. **13**(1): p. 114 - 122.

28. Jaiswal, R., M. Choudhury, S. Zaman, S. Singh, V. Santosh, D. Bastia, and C. R. Escalante, *Functional Architecture of the Reb1-Ter Complex of Schizosaccharomyces pombe.* Proc. Natl. Acad. Sci. USA, 2016. **113**(16): p. E2267 - E2276.

29. Zhao, A., A. Guo, Z. Liu, and L. Pape, *Molecular Cloning and Analysis of Schizosaccharomyces pombe Reb1p: Sequence-specific Recognition of Two Sites in the Far Upstream rDNA Intergenic Spacer.* Nucleic Acids Research, 1997. **25**(4): p. 904 - 910.

30. Appling, F.D., C.E. Scull, A.L. Lucius, and D.A. Schneider, *The A12.2 Subunit is an Intrinsic Destabilizer of the RNA Polymerase I Elongation Complex.* Biophys. J., 2018. **114**(11): p. 2507 - 2515.

31. Appling, F.D., D.A. Schneider, and A.L. Lucius, *Multisubunit RNA Polymerase CLeavage Factors Modulate the Kinetics and Energetics of Nucleotide Incorporation: An RNA Polymerase I Case Study.* Biochemistry, 2017. **56**: p. 5654 - 5662.

32. Darriere, T., M. Pilsl, A. Chauvier, T. Genty, S. Audibert, C. Dez, I. Leger-Silvestre, C. Normand, O. Calvo, C. Fernandez-Tornero, H. Tschochner, O. Gadal, *Genetic Analysis of RNA Polymerase I Led to theDiscovery of a Super-*

Active Mutant of the RNA Polymerase I PLOS Genetics, 2019. **15**(5): p. e1008157.

33. Clarke, A.M., K.L. Engel, K.E. Giles, C.M. Petit, and D.A. Schneider, *NET-seq Reveals Heterogeneous Nucleotide Incorporation by RNA Polymerase I.* PNAS, 2018. **115**(50): p. E11633 - E11641.

34. Sigurdsson, S., A. B. Dirac-Svejstrup, J. Q. Svejstrup, *Evidence that Transcript Cleavage Is Essential for RNA Polymerase II Transcription and Cell Viability.* Mol Cell, 2010. **38**(2): p. 202-210.

35. Scull, C.E., Z. M. Ingram, A. L. Lucius, and D. A. Schneider, *A Novel Assay for RNA Polymerase I Transcription Elongation Sheds Light on Evolutionary Divergence of Eukaryotic RNA Polymerases.* Biochemistry, 2019. **58**(16): p. 2116-2124.

36. Saeki, H., and J. Q. Svejstrup, *Stability, Flexibility, and Dynamic Interactions of Colliding RNA Polymerase II Elongation Complexes.* Mol Cell, 2009. **35**(2-3): p. 191 - 205.

37. McStay, B., and R.H. Reeder, *A Termination Site for Xenous RNA Polymerase I Also Acts as an Element of an Adjacent Promoter.* Cell, 1986. **47**: p. 913 - 920.

38. Grummt, I., H. Rosenbauer, I. Niedermeyer, U. Maier, and A. Ohrlein, *A Repeated 18 bp Sequence Motif in the Mouse rDNA Spacer Mediates Binding of a Nuclear Factor and Transcription Termination.* Cell, 1986. **45**(6): p. 837 - 846.

39. Team, R.C., *R: A Language and Environment for Statistical Computing*. 2017, R Foundation for Statistical Computing: Vienna, Austria.

40. Wickham, H., R. Francois, L. Henry, and K. Muller, *dplyr: A Grammar of Data Manipulation*. 2018.

41. Wickham, H., *The Split-Apply-Combine Strategy for Data Analysis.* Journal of Statistical Software, 2011. **40**(1): p. 1 - 29.

42. Wickham, H., *ggplot2: Elegant Graphics for Data Analysis*. 2009, New York, NY: Springer-Verlag.

43. Wagih, O., *ggseqlogo: A 'ggplot2' Extension for Drawing Publication-Ready Sequence Logos*. 2017.

44. Kassambara, A., *ggpubr: 'ggplot2' Based Publication Ready Plots*. 2018.

45. Wilke, C.O., *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'*. 2017.

46. Wickham, H., J. Hester, and W. Chang, *devtools: Tools to Make Developing R Packages Easier*. 2018.

47. Bengtsson, H., *matrixStats: Functions that Apply to Rows and Columns of Matrices (and to Vectors)*. 2018.

48. Marot, G., J.-L. Fouley, C.-D. Mayer, and F. Jaffrezic, *Moderated Effect Size and P-value Combinations for Microarray Meta-Analyses.* Bioinformatics, 2009. **25**(20): p. 2692 - 2699.

49. Wickham, H., *scales: Scale Functions for Visualization*. 2017.

SUMMARY OF FINDINGS AND FINAL CONCLUSIONS

Summary of Findings

In this dissertation, we demonstrated that Pol I occupancy is DNA template sequence-dependent *in vitro* and *in vivo*, and that mutations that alter nucleotide addition by Pol I also affect the relationship between Pol I occupancy and DNA template sequence.

In the first chapter, we characterized the effect of a prokaryotic *rho* independent terminator motif from the *E. coli Spy* gene on yeast Pol I *in vitro*. This terminator motif was recently identified as the most efficient terminator of RNAP transcription in the *E. coli* genome [119]. Similar motifs have previously been shown to terminate transcription by other eukaryotic RNA polymerases *in vitro* and *in vivo* [120-122]. We determined that this motif is capable of terminating Pol I transcription *in vitro*, though at reduced efficiency compared to RNAP. We also demonstrate that both components of the *rho*-independent terminator motif are required for termination and identified a sequence motif from the 35S gene that has a similar effect on Pol I transcription. Finally, we observed that the effects on Pol I transcription of both the *rho*-independent terminator motif and the endogenous 35S sequence were UTP concentration dependent. We next examined Pol I occupancy *in vivo* by adapting NET-seq [123] for use with yeast Pol I. We identified reproducible heterogeneity of occupancy throughout the 35S gene. Within the highest occupancy sites, we observed an overwhelming anti-preference for G residues in the last encoded nucleotide position. Expanding our analysis, we determined that this anti-

preference exists gene-wide, indicating that Pol I occupancy is at least partially dependent on sequence context *in vivo*.

In the second chapter, we used NET-seq to observe Pol I occupancy in a Pol I subunit deletion (*Δrpa12*) strain of yeast. The A12.2 subunit has domains that are homologous to Pol II subunit Rpb9 and Pol II transcription factor TFIIS [71] and is responsible for stimulating transcript cleavage by Pol I [48] and mediating Pol I transcription termination [72]. Previous work from the Schneider lab identified that deletion of the gene that encodes the A12.2 subunit also affects nucleotide addition by Pol I [118] and the overall stability of the Pol I EC [124]. Deletion of the A12.2 subunit resulted in a 5' shift of Pol I occupancy within the 35S gene. We also observed alterations in the relationship between Pol I occupancy and increasing nucleotide enrichment in both up- and downstream DNA, confirming previous studies that posited a role for the A12.2 subunit beyond transcript cleavage and termination [118, 124]. Additionally, we identified distinct changes in occupancy within the Intergenic Spacer (IGS) region, downstream of the 35S gene. As previously reported, we found that *Δrpa12* Pol I is termination-deficient, with consistent occupancy observed downstream of both known sites of Pol I transcription termination [72]. We also observed a significant increase in occupancy at a site in the 3' end of the IGS, upstream of the promoter region for the adjacent 35S gene. Interestingly, this site lies just upstream from a promoter proximal binding site for DNA-binding protein Reb1. Reb1 binding to the rDNA mediates termination of Pol I transcription in *Schizosaccharomyces pombe* [125] and was long assumed to play the same role in yeast [126-128], though recent scholarship has identified a Reb1 homologue that is actually responsible for mediating yeast Pol I

termination [129]. These findings suggest that yeast possesses a third site of Pol I termination proximal to the promoter for downstream 35S genes, which is present in other eukaryotes [130] but not previously described in yeast.

Together, these data demonstrate that DNA template sequence affects Pol I elongation *in vitro* and *in vivo*, and that this relationship is perturbed by mutations in Pol I. The implications of these discoveries on our understanding of efficient rRNA synthesis by Pol I are discussed below.

Implications

Formation of functionally viable ribosomes has been shown to be RNA polymerase elongation rate-dependent in both prokaryotic and eukaryotic organisms. In *E. coli*, transcription of the ribosomal RNA by bacteriophage T7 RNA polymerase (which has a 5-fold faster elongation rate compared to RNAP) results in completely inactive ribosomes [107]. In yeast, the Schneider lab and others have demonstrated that perturbations in Pol I elongation rate also affect rRNA processing and efficient ribosome biogenesis. Deletion of Paf1 complex subunits reduces Pol I elongation rate *in vivo* and also results in reduced co-transcriptional rRNA cleavage [102]. Deletion of the conserved transcription factor *SPT4* negatively affects efficient synthesis of the large ribosomal subunit, and this defect was determined to occur during transcription elongation [103]. Finally, mutations of Pol I subunits that impair Pol I elongation rate also result in abnormal rRNA processing and ultimately dysfunctional ribosome biogenesis [106]. All of these data indicate a link between Pol I elongation rate and efficient ribosome biogenesis.

In other transcription systems, transcript processing steps are linked to specific transcriptional pauses that facilitate appropriate secondary structure formation. One example is proper folding of the RNAse P RNA found in *E. coli* and *B. subtilis*. As with ribosomes, the RNA in RNAse P is responsible for the catalytic activity of its parent complex [131], and contains extensive secondary structure, including helices composed of distal elements [132]. In *B. subtilis,* proper secondary structure formation within the RNAse P RNA was linked first to RNA polymerase transcription rate in *B. subtilis,* and then to transcriptional pausing [116]. Analysis of RNAse P RNA folding in *E. coli* revealed that transcriptional pausing at discrete sites allowed for the efficient formation of helices with distal elements. Pausing after the transcription of the 5' portion of the helix allowed for sequestration of that portion until the 3' portion could be transcribed and quickly folded into the mature structure [111].

The ribosomal RNA in yeast contains significant secondary structure, much of it requiring the efficient folding of distal elements (reviewed in [7]). In addition, the bulk of RNA processing steps require the accessibility of primary 35S transcript sequence for snoRNP binding (including the co-transcriptional cleavage events perturbed in the yeast mutational studies described above [102, 103, 106]) (reviewed in [7]). Our work has shown that Pol I occupancy is highly heterogeneous *in vivo*, suggesting that previously observed bulk elongation rates represent a composite of fast and slow nucleotide addition events, including numerous instances of polymerase pausing. It is therefore possible that the perturbations of rRNA processing observed in yeast strains with reduced Pol I elongation rate are due to altered polymerase pausing at specific sites, as opposed to changes in overall transcription rate.

We also observed that elongation affecting DNA sequence motifs from other transcription systems, as well as a similar sequence from the yeast 35S gene affect Pol I elongation *in vitro*. Crucially, we also demonstrated that the observed effect for both motifs is nucleotide concentration-dependent. Nutrient concentration-dependent regulation of transcription via differential secondary structure formation in nascent transcripts has long been observed in prokaryotes (reviewed in [133]) and more recently in eukaryotes as well [134]. It is therefore possible that the observed polymerase pausing *in vivo* is dictated by free nucleotide availability, and that reduction in local nucleotide concentration may perturb RNA polymerase pausing events important for transcript processing similarly to mutations that negatively affect Pol I elongation rate. Future NET-seq studies on mutations that reduce Pol I elongation rate as well as comparisons of Pol I occupancy before and after nucleotide depletion, coupled with development of experimental and computational techniques for analyzing co-transcriptional formation of nascent transcript secondary structure will allow us to explore this development further.

Final Conclusions

The data presented above demonstrate that Pol I occupancy is reproducibly heterogeneous, and sensitive to DNA template sequence *in vitro* and *in vivo*. They also show that mutations that affect nucleotide addition by Pol I also affect Pol I occupancy. These findings suggest the existence of another layer of information embedded in the rDNA, which dictates changes in elongation rate to Pol I. Given previously observed relationships between Pol I elongation rate and rRNA processing, it is also likely that these encoded pauses are integral to successful rRNA maturation, and ribosome

biogenesis as a whole. Taken together, the contents of this dissertation establish DNA template sequence as a potential regulator of rRNA synthesis, and raise interesting questions about the co-evolution of the ribosomal DNA, and the specialized RNA polymerase that transcribes it.

GENERAL LIST OF REFERENCES

1. Moore, P.B., and T. A. Steitz, *The Involvement of RNA in Ribosome FUnction.* Nature, 2002. **418**: p. 229 - 235.

2. Crick, F.H.C., L. Barnett, S. Brenner, and R. J. Watts-Tobin, *General Nature of the Genetic Code for Proteins.* Nature, 1961. **192**(4809): p. 1227 -1232.

3. Brenner, S., *RNA, Ribosomes, and Protein Synthesis.* Cold Spring Harb. Symp. Quant. Biol., 1961. **XXVI**: p. 101 - 109.

4. Barta, A., et al., *Identification of a site on 23S ribosomal RNA located at the peptidyl transferase center.* Proceedings of the National Academy of Sciences of the United States of America, 1984. **81**(12): p. 3607-3611.

5. Stephenson, B., M. L. Scott, J. F. Hecht, P. C. Zamecnik, *A Soluble RIbonucleic Acid Intermediate in Protein Synthesis.* J Biol Chem, 1958. **231**: p. 243 - 257.

6. Crick, F.H.C., *Codon-Anticodon Pairing: The Wobble Hypothesis.* J Mol Biol, 1966. **19**: p. 548 - 555.

7. Woolford, J.L., and S.J. Braserga, *Ribosome Biogenesis in the Yeast Saccharomyces cerevisiae.* Genetics, 2013. **195**(3): p. 643 - 681.

8. Lafontaine, D.L.J. and D. Tollervey, *The function and synthesis of ribosomes.* Nature Reviews Molecular Cell Biology, 2001. **2**(7): p. 514-520.

9. Steitz, T.A., *A structural understanding of the dynamic ribosome machine.* Nature Reviews Molecular Cell Biology, 2008. **9**: p. 242.

10. Ogle, J.M., et al., *Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit.* Science, 2001. **292**(5518): p. 897.

11. Nissen, P., et al., *The Structural Basis of Ribosome Activity in Peptide Bond Synthesis.* Science, 2000. **289**(5481): p. 920.

12. Warner, J.R., *The Economics of Ribosome Biosynthesis in Yeast.* Trends in Biochemical Sciences, 1999. **24**(11): p. 437 - 440.

13. Lempiäinen, H. and D. Shore, *Growth control and ribosome biogenesis.* Current Opinion in Cell Biology, 2009. **21**(6): p. 855-863.

14. Bernstein, K.A., F. Bleichert, J. M. Bean, F. R. Cross, and S. J. Braserga, *Ribosome Biogenesis is Sensed at the Start Cell Cycle Checkpoint.* Mol Biol Cell, 2007. **18**(3): p. 953 - 964.

15. Volarevic, S., M. J. Stewart, B. Ledermann, F. Zilberman, L. Terracciano, E. Montini, M. Grompe, S. C. Kosma, and G. Thomas, *Proliferation, But Not*

Growth, Blocked by Conditional Deletion of 40S Ribosomal Protein S6. Science, 2000. **288**(5473): p. 2045 - 2047.

16. Pestov, D.G., Z. Strezoska, and L. F. Lau, *Evidence of p53-Dependent Cross-Talk between RIbosome Biogenesis and the Cell Cycle: Effects of Nucleolar Protein Bop1 on G_1/S Transition.* Mol Cell Biol, 2001. **21**(13): p. 4246 - 4255.

17. Thomson, E., S. Ferreira-Cerca, and E. Hurt, *Eukaryotic Ribosome Biogenesis at a Glance.* Journal of Cell Science, 2013. **126**: p. 4815 - 4821.

18. Weinmann, R. and R.G. Roeder, *Role of DNA-Dependent RNA Polymerase III in the Transcription of the tRNA and 5S RNA Genes.* Proceedings of the National Academy of Sciences, 1974. **71**(5): p. 1790-1794.

19. Kos, M., and D. Tollervy, *Yeast Pre-rRNA Processing and Modification Occur Cotranscriptionally`.* Mol Cell, 2010. **37**(6): p. 809 - 820.

20. Henras, A.K., C. Plisson-Chastang, M. F. O'Donohue, A. Chakriborty, and P. E. Gleizes *An Overviewof pre-ribosomal RNA processing in Eukaryotes.* WIREs RNA, 2014. **6**(2): p. 225 - 242.

21. Ramaswamy, P., and S. A. Woodson, *Global Stabilization of rRNA Structure by Ribosomal Proteins S4, S17, and S20.* J Mol Biol, 2010. **392**(3): p. 666 - 677.

22. Deshmukh, M., et al., *Yeast ribosomal protein L1 is required for the stability of newly synthesized 5S rRNA and the assembly of 60S ribosomal subunits.* Molecular and cellular biology, 1993. **13**(5): p. 2835-2845.

23. Fatica, A. and D. Tollervey, *Making ribosomes.* Current Opinion in Cell Biology, 2002. **14**(3): p. 313-318.

24. Werner, F., and D. Grohmann, *Evolution of Multisubunit RNA Polymerases in the Three Domains of Life.* Nat Rev Microbiol, 2011. **9**: p. 85 - 98.

25. deHaseth, P.L., M.L. Zupancic, and M.T. Record, *RNA Polymerase-Promoter Interactions: the Comings and Goings of RNA Polymerase.* Journal of Bacteriology, 1998. **180**(12): p. 3019-3025.

26. Roeder, R.G., *The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly.* Trends in Biochemical Sciences, 1991. **16**: p. 402-408.

27. Kettenberger, H., K. J. Armache, and P. Cramer, *Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS.* Mol Cell, 2004. **16**(6).

28. Vassylyev, D.G., et al., *Structural basis for substrate loading in bacterial RNA polymerase.* Nature, 2007. **448**: p. 163.

29. Brueckner, F., J. Ortiz, and P. Cramer, *A movie of the RNA polymerase nucleotide addition cycle.* Current Opinion in Structural Biology, 2009. **19**(3): p. 294-299.

30. Kireeva, M.L., N. Kommisarova, D. S. Waugh, and M. Kashlev *The 8-Nucleotide-long RNA:DNA Hybrid is a Primary Stability Determinant of the RNA Polymerase II Elongation Complex.* J Biol Chem, 2000. **275**: p. 6530-6536.

31. Sidorenkov, I., N. Komissarova, and M. Kashlev, *Crucial Role of the RNA:DNA Hybrid in the Processivity of Transcription.* Mol Cell, 1998. **2**(1): p. 55 - 64.

32. Straney, D.C. and D.M. Crothers, *Intermediates in transcription initiation from the E. coli lac UV5 promoter.* Cell, 1985. **43**(2, Part 1): p. 449-459.

33. Luse, D.S., and G. A. Jacob, *Abortive initiation by RNA polymerase II in vitro at the adenovirus 2 major late promoter.* J Biol Chem, 1987. **262**: p. 14990-14997.

34. Kapanidis, A.N., E. Margeat, S. O. Ho, E. Kortkhonjia, S. Weiss, and R. H. Ebright, *Initial Transcription by RNA Polymerase Proceeds Through a DNA-Scrunching Mechanism.* Science, 2009. **314**(5802): p. 1144- 1147.

35. Liu, X., D. A. Bushnell, D. Wang, G. Calero, and R. D. Kornberg, *Structure of the RNA Polymerase II-TFIIB Complex and the Transcription Initiation Mechanism.* Science, 2010. **327**(5962): p. 206 - 209.

36. McClure, W.R., *Mechanism and Control of Transcription Initiation in Prokaryotes.* Ann Rev Biochem, 1985. **54**: p. 171 - 204.

37. Shandilya, J. and S.G.E. Roberts, *The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling.* Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2012. **1819**(5): p. 391-400.

38. Henkin, T.M., *Transcription termination control in bacteria.* Current Opinion in Microbiology, 2000. **3**(2): p. 149-153.

39. El Hage, A., M. Koper, J. Kufel, and D. Tollervey, *Efficient Termination of Transcription by RNA Polymerase I Requires the 5' Exonuclease Rat1 in Yeast.* Genes Dev., 2008. **22**(8): p. 1069 - 1081.

40. Kuehner, J.N., E.L. Pearson, and C. Moore, *Unravelling the means to an end: RNA polymerase II transcription termination.* Nature Reviews Molecular Cell Biology, 2011. **12**: p. 283.

41. Nielsen, S., Y. Yuzenkova, and N. Zenkin, *Mechanism of Eukaryotic RNA Polymerase III Transcription Termination.* Science, 2013. **340**(6140): p. 1577 - 1580.

42. Roeder, R.G. and W.J. Rutter, *Multiple Forms of DNA-dependent RNA Polymerase in Eukaryotic Organisms.* Nature, 1969. **224**(5216): p. 234-237.

43. Zylber, E.A., and S. Penman, *Products of RNA Polymerases in HeLa Cell Nuclei.* PNAS, 1971. **68**(11): p. 2861 - 2865.

44. Kedinger, C. and R. Simard, *THE ACTION OF α-AMANITIN ON RNA SYNTHESIS IN CHINESE HAMSTER OVARY CELLS.* Ultrastructural and Biochemical Studies, 1974. **63**(3): p. 831-842.

45. Young, R.A., *RNA POLYMERASE II.* Annual Review of Biochemistry, 1991. **60**(1): p. 689-715.

46. Busch, H., R. Reddy, L. Rothblum, and Y. C. Choi, *snRNAs, snRNPs, and RNA Processing.* Ann Rev Biochem, 1982. **51**(617 - 654).

47. Lee, Y., et al., *MicroRNA genes are transcribed by RNA polymerase II.* The EMBO Journal, 2004. **23**(20): p. 4051-4060.

48. Kuhn, C.-D., S.R. Heiger, S. Baumli, M. Gartmann, J. Gerber, S. Jennebach, T. Mielke, H. Tschochner, R. Beckmann, and P. Cramer, *Functional Architecture of RNA Polymerase I.* Cell, 2007. **131**(7): p. 1260 - 1272.

49. Lane, W.J., and S. A. Darst, *A Molecular Evolution of Multisubunit RNA Polymerases: Sequence Analysis.* Journal of Molecular Biology, 2010. **395**(4): p. 671 - 685.

50. Nudler, E., *RNA Poymerase Backtracking in Gene Regulation and Genome Instability.* Cell, 2012. **149**(7): p. 1438 - 1445.

51. Nudler, E., A. Mustaev, A. Goldfarb, and E. Lukhtanov, *The RNA-DNA Hybrid Maintains the Register of Transcription by Preventing Backtracking of RNA Polymerase.* Cell, 1997. **89**(1): p. 33 - 41.

52. Wang, D., D. A. Bushnell, X. Huang, K. D. Westover, M. Levitt, and R. D. Kornberg, *Structural Basis of Transcription: Backtracked RNA Polymerase II at 3.4 Angstrom Resolution.* Science, 2009. **324**(5831): p. 1203 - 1206.

53. Thomas, M.J., A. A. Platas, and D. K. Hawley, *Transcriptional Fidelity and Proofreading by RNA Polymerase II.* Cell, 1998. **93**(4).

54. Erie, D.A., O. Hajiseyedjavadi, M. C. Young, and P. H. von Hippel, *Multiple RNA Polymerase Conformations and GreA: Control of the Fidelity of Transcription.* Science, 1993. **262**(5135).

55. Epshtein, V., F. Toulme, A. R. Rahmouni, S. Borukhov, and E. Nudler, *Transcription Through the Roadblocks: the Role of RNA Polymerase Cooperation.* Embo J, 2003. **22**(18): p. 4719 - 4727.

56. Lisica, A., C. Engel, M. Jahnel, E. Roldan, E. A. Galburt, P. Cramer,and S. W. Grill, *Mechanisms of Bactrack Recoveryby RNA Polymerase I and II.* PNAS, 2016. **113**(11): p. 2946 - 2951.

57. Orlova, M., J. Newlands, A. Das, A. Goldfarb, and S. Borukhov, *Intrinsic Transcript Cleavage Activity of RNA Polymerase.* PNAS, 1995. **92**: p. 4596 - 4600.

58. Reines, D., *Elongation Factor-dependent Transcript Shortening by Template-engaged RNA Polymerase II.* J Biol Chem, 1992. **267**(6): p. 3795 - 3800.

59. Borukhov, S., A. Polyakov, V. Nikiforov, and A. Goldfarb, *GreA Protein: A Transcription Factor form Escherichia coli.* PNAS, 1992. **89**(19): p. 8899 - 8902.

60. Borukhov, S., V. Sagitov, and A. Goldfarb, *Transcript Cleavage Factors from E. coli.* Cell, 1993. **72**(3): p. 459 - 466.

61. Toulme, F., C. Mosrin-Huaman, J. Sparkowski, A. Das, M. Leng, and A. R. Rahmouni, *GreA and GreB Proteins Revive Backtracked RNA Polymerase in vivo by Promoting Transcript Trimming.* Embo J, 2000. **19**(24): p. 6853 - 6859.

62. Chedin, S., M. Rivera, P. Schultz, A. Sentenc, and C. carles, *The RNA cleavage activity of RNA Polymerase III is Mediated by an Essential TFIIS-like Subunit and is Important for Transcription Termination.* Genes Dev, 1998. **12**: p. 3857 - 3871.

63. Alic, N., N. Ayoub, E. Landrieux, E. Favry, P. Baudouin-Cornu, M. Riva, and C. Carles, *Selectivity and Proofreading Both Contribute Significantly to the Fidelity of RNA Polymerase III Transcription.* PNAS, 2007. **104**(25): p. 10400 - 10405.

64. Peyroche, G., P. Milkereit, N. Bischler, H. Tschochner, P. Schultz, A. Sentenac, C. Carles, and M. Riva, *The Recruitment of RNA Polymerase I on rDNA is Mediated by the Interaction of the A43 Subunit with Rrn3.* Embo J, 2000. **19**(20): p. 5473 - 5482.

65. Geiger, S.R., K. Lorenzen, A. Schreieck, P. Hanecker, D. Kostrewa, A. J. R. Heck, and P. Cramer, *RNA Polymerase I Contains a TFIIF-Related DNA-Binding Subcomplex.* Mol Cell, 2010. **39**(4): p. 583 - 594.

66. Ohkuma, Y., S. Hashimoto, C. K. Wang, M. Horikoshi, and R. G. Roeder, *Analysis of the Role of TFIIE in Basal Transcription and TFIIH-Mediated Carboxy-Terminal Domain Phosphorylation through Structure-Function Studies of TFIIE-a* Mol Cell Biol, 1995. **15**(9): p. 4854 - 4866.

67. Liljelund, P., S. Mariotte, J. M. Buhler, and A. Sentenac, *Characterization and Mutagenesis of the Gene Encoding the A49 Subunit of RNA Polymerase A in Saccharomyces cerevisiae.* PNAS, 1992. **89**: p. 9302 - 9305.

68. Gadal, O., S. Mariotte-Labarre, S. Chedin, E. Quemeneur, C. Carles, A. Sentenac, and P. Thuriaux, *A34.5, a Nonessential Component of Yeast RNA Polymerase I, Cooperates with Subunit A14 and DNA Topoisomerase I To Produce a Functional rRNA Synthesis Machine.* Mol Cell Biol, 1997. **17**(4): p. 1787 - 1795.

69. Merkl, P.E., M. Pilsl, T. Fremter, G. Langst, P. Milkereit, J. Griesenbeck, and H. Tschochner, *RNA Polymerase I Transcription Fidelity, Speed, and Processivity Depend on the Interplay of its Lobe Binding Subunits.* BioRxiv, 2018.

70. Nogi, Y., R. Yano, J. Dodd, C. Carles, and M. Nomura, *Gene RRN4 in Saccharomyces cerevisiae Encodes the A12.2 Subunit of RNA Polymerase I and Is Essential Only at High Tempuratures.* Mol Cell Biol, 1993. **13**(1): p. 114 - 122.

71. Van Mullem, V., E. Landrieux, J. Vandenhaute, and P. Thuriaux, *Rpa12p, a ConservedRNA Polymerase I Subunit with Two Functional Domains.* Molecular Microbiology, 2002. **43**(5): p. 1105 - 1113.

72. Prescott, E.M., Y.N. Osheim, H.S. Jones, C.M. Alen, J.G. Roan, R.H. Reeder, A.L. Beyer, and N.J. Proudfoot *Transcriptional Termination by RNA Polymerase I Requires the Small Subunit Rpa12P.* Proc. Natl. Acad. Sci. USA, 2004. **101**(16): p. 6068 - 6073.

73. French, S.L., Y.N. Osheim, F. Cioci, M. Nomura, and A.L. Beyer *In Exponentially Growing Saccharomyces cerevisiae Cells, rRNA Synthesis is Determined by the Summed RNA Polymerase I Loading Rate Rather than by the Number of Active Genes.* Mol Cell Biol, 2003. **23**(5): p. 1558-1568.

74. Velculescu, V.E., L. Zhang, W. Zhou, J. Vogelstein, M. A. Basrai, D. E. Basset jr., P. Heiter, B. Vogelstein, and K. W. Kinzler, *Characterization of the Yeast Transcriptome.* Cell, 1997. **88**(2): p. 243 - 251.

75. Li, B., C.R. Nierras, and J.R. Warner, *Transcriptional Elements Involved in the Repression of Ribosomal Protein Synthesis.* Molecular and Cellular Biology, 1999. **19**(8): p. 5393.

76. Donovan, D.M., and N. J. Pearson, *Transcriptional Regulation of RIbosomal Proteins During a Nutritional Upshift in Saccharomyces cerevisiae.* Mol Cell Biol, 1986. **6**(7): p. 2429 - 2435.

77. Herruer, M.H., W. H. Mager, L. P. Woudt, R. T. Nieuwint, G. M. Wassenaar, P. Groenveld, and R. J. Planta, *Transcriptional Control of Yeast Ribosomal Protein Synthesis During Carbon-Source Upshit.* Nucleic Acids Res, 1987. **15**(24): p. 10133 - 10144.

78. Kim, C.H., and J. R. Warner, *Mild Temperature Shock Alters the Transcription of a Discrete Class of Saccharomyces cerevisiae Genes.* Mol Cell Biol, 1983. **3**(3): p. 457 - 465.

79. Reiter, A., et al., *Reduction in ribosomal protein synthesis is sufficient to explain major effects on ribosome production after short-term TOR inactivation in Saccharomyces cerevisiae.* Molecular and cellular biology, 2011. **31**(4): p. 803-817.

80. Mizuta, K., and J. R. Warner, *Continued Functioning of the Secretory Pathway is Essential for Ribosome Synthesis.* Mol Cell Biol, 1994. **14**(4): p. 2493 - 2502.

81. Liang, S., F. Lacroute, and F. Kepes, *Multicopy STS1 Restores Both Protein Transport and Ribosomal RNA Stability in a New Yeast sec23 Mutant Allele.* European Journal of Cell Biology, 1993. **62**(2): p. 270 - 281.

82. Pérez-Fernández, J., et al., *The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism.* Molecular and cellular biology, 2007. **27**(15): p. 5414-5429.

83. Gallagher, J.E.G., D. A. Dunbar, S. Granneman, B. M. Mitchell, Y. Osheim, A. L. Beyer, and S. J. Baserga, *RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components.* Genes Dev, 2004. **18**: p. 2506 - 2517.

84. Dragon, F., et al., *A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis.* Nature, 2002. **417**(6892): p. 967-970.

85. Nikolov, E.N., M. D. Dabeva, and T. K. Nikolov, *Turnover of Ribosomes in Regenerating Rat Liver.* International Journal of Biocehmistry, 1983. **15**(1): p. 1255- 1260.

86. Ju, Q. and J.R. Warner, *Ribosome synthesis during the growth cycle of Saccharomyces cerevisiae.* Yeast, 1994. **10**(2): p. 151-157.

87. Leger-Silvestre, I., J. M. Caffrey, R. Dawaliby, D. A. Alvarez-Arias, N. Gras, S. J. Bertolone, P. E. Gleizes, and S. R. Ellis, *Specific Role for Yeast Homologs of the Diamond Blackfan Anemia-associated Rps19 Protein in Ribosome Synthesis.* J Biol Chem, 2005. **280**: p. 38177-38185.

88. Hayano, T., M. Yanagida, Y. Yamauchi, T. Shinkawa T. Isobe, and N.Takahashi *Proteomic Analysis of Human Nop56p-associated Complexes.* J Biol Chem, 2003. **278**: p. 34309 - 34319.

89. Miller, D.M., et al., *c-Myc and cancer metabolism.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2012. **18**(20): p. 5546-5553.

90. Beroukhim, R., et al., *The landscape of somatic copy-number alteration across human cancers.* Nature, 2010. **463**(7283): p. 899-905.

91. Lin, C.Y., et al., *Transcriptional amplification in tumor cells with elevated c-Myc.* Cell, 2012. **151**(1): p. 56-67.

92. Ruggero, D., *Revisiting the Nucleolus: From Marker to Dynamic Integrator of Cancer Signaling.* Science Signaling, 2012. **5**(241): p. pe38-pe38.

93. van Sluis, M. and B. McStay, *Ribosome biogenesis: Achilles heel of cancer?* Genes & cancer, 2014. **5**(5-6): p. 152-153.

94. Pelletier, J., G. Thomas, and S. Volarević, *Ribosome biogenesis in cancer: new players and therapeutic avenues.* Nature Reviews Cancer, 2017. **18**: p. 51.

95. Grummt, I., *Life on a Planet of Its Own: Regulation of RNA Polymerase I Transcription in the Nucleolus* Genes Dev, 2003. **17**: p. 1691 - 1702.

96. Laferte, A., A. Favry, A. Santenac, M. Riva, C. Carles, and S. Chedin, *The Transcriptional Activity of RNA Polymerase I is a Key Determinant for the Level of all Ribosome Components.* Genes Dev, 2006. **20**(15): p. 2030 - 2040.

97. Bywater, M.J., G. Poortinga, E. Sanij, N. Hein, A. Peck, C. Cullinane, M. Wall, L. Cluse, D. Drygin, K. Anderes, N. Huser, C. Proffitt, J. Bliesath, M. Haddach, M. K. Schwaebe, D. M. Ryckman, W. G. Rice, C. Schmitt, S. W. Lowe, R. W. Johnstone, R. B. Pearson, G. A. McArthur, and R. D. Hannan, *Inhibition of RNA Polymerase I as a Therapeutic Strategy to Promote Cancer-Specific Activation of p53.* Cancer Cell, 2012. **22**(1): p. 51 - 65.

98. Kos-Braun, I.C., and M. Kos, *Post-Transcriptional Regulation of RIbosome Biogenesis in Yeast.* Microbial Cell, 2017. **4**(5): p. 179 - 181.

99. Sandmeier, J.J., S. French, Y.Osheim, W. L. Cheung, C. M. Gallo, A. L. Beyer, and J. S. Smith, *RPD3 is Required for the Inactivation of Yeast Ribosomal DNA Genes in Stationary Phase.* Embo J. , 2002. **21**(18): p. 4959 - 4968.

100. Panov, K.I., J. K. Friedrich, J. Russell, and J. C. B. M. Zomerdijk, *UBF activates RNA Polymerase I Transcription by Stimulating Promoter Escape.* EMBO J, 2006. **25**(14): p. 3310 - 3322.

101. Zhang, Y., S. L. French, A. L. Beyer, and D. A. Schneider, *The Transcription Factor THO Promotes Transcription Initiation and Elongation by RNA Polymerase I.* J Biol Chem, 2016. **291**(6): p. 3010 - 3018.

102. Zhang, Y., M. L. Sikes, A. L. Beyer, and D. A. Schneider, *The Paf1 Complex is Required for Efficient Transcription Elongation by RNA Polymerase I.* PNAS, 2009. **106**(7): p. 2153 - 2158.

103. Schneider, D.A., S. L. French, Y. N. Osheim, A. O. Bailey, L. Vu, J. Dodd, J. R. Yates, A. L. Beyer, and M. Nomura, *RNA Polymerase II Elongation Factors*

Spt4p and Spt5p Play Roles in Transcription Elongation by RNA Polymerase I and rRNA Processing. PNAS, 2006. **103**(34): p. 12707 - 12712.

104. Zhang, Y., S. J. Anderson, S. L. French, M. L. Sikes, O. V. Viktorovskaya, J. Huband, K. Holcomb, J. L. Hartman 4th, A. L. Beyer, and D. A. Schneider *The SWI/SNF Chromatin Remodeling Complex Influences Transcription by RNA Polymerase I in Saccharomyces cerevisiae.* PLoS One, 2013. **8**(2): p. e56793.

105. Zhang, Y., A. D. Smith IV, M. B. Renfrow, and D. A. Schneider, *The RNA Polymerase-Associated Factor 1 Complex (Paf1C) Directly Increases the Elongation Rate of RNA Polymerase I and is Required for Efficient Regulation of rRNA Synthesis.* J Biol Chem, 2010. **285**(19): p. 14152 - 14159.

106. Schneider, D.A., A. Michel, M. L. Sikes, L. Vu, J. A. Dodd, S. Salgia, Y. N. Osheim, A. L. Beyer, and M. Nomura, *Transcription Elongation by RNA Polymerase I is Linked to Efficient rRNA Processing and Ribosome Assembly.* Mol Cell, 2007. **26**(2): p. 217 - 229.

107. Lewicki, B.T.U., T. Margus, J. Remme, and K. H. Nierhaus, *Coupling of rRNA Transcription and Ribosomal Assembly in vivo: Formation of Active Ribosomal Subunits in Escherichia coli Requires Transcription of rRNA Genes by Host RNA Polymerase which Cannot be Replaced by Bacteriophage T7 RNA Polymerase.* JMB, 1993. **231**(5): p. 581 - 593.

108. Birchmeier, C., R. Grosschedl, and M.L. Birnstiel, *Generation of authentic 3' termini of an H2A mRNA in vivo is dependent on a short inverted DNA repeat and on spacer sequences.* Cell, 1982. **28**(4): p. 739-745.

109. Sànchez, R. and W.F. Marzluff, *The Stem-Loop Binding Protein Is Required for Efficient Translation of Histone mRNA In Vivo and In Vitro.* Molecular and Cellular Biology, 2002. **22**(20): p. 7093-7104.

110. Saldi, T., N. Fong, and D. L. Bentley, *Transcription Elongation Rate Affects Nascent Histone Pre-mRNA Folding and 3' End Processing.* Genes Dev, 2018. **32**(3-4): p. 297-308.

111. Wong, T.N., T. R. Sosnick, and T. Pan, *Foldng of Noncoding RNAs During Transcription Facilitated by Pause-Induced Nonnative Structures.* PNAS, 2007. **104**(46): p. 17995 - 18000.

112. Lost, I., and M. Dreyfus, *The Stability of Escherichia coli lacZ mRNA Depends Upon the Simultaneity of its Synthesis and Translation.* Embo J., 1995. **14**(13): p. 3252 - 3261.

113. Wilson, K.S., and P. H. von Hippel, *Transcription Termination at Intrinsic Terminators: the Role of the RNA Hairpin.* PNAS, 1995. **92**(19): p. 8793 - 8797.

114. Landick, R., J. Carey, and C. Yanofsky, *Translation Activates the Paused Transcription Complex and Restores Transcription of the trp Operon Leader Region.* PNAS, 1985. **82**(14): p. 4663 - 4667.

115. WIckiser, J.K., W. C. Winkler, R. R. Breaker, and D. M. Crothers, *The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch.* Mol Cell, 2005. **18**(1): p. 49 - 60.

116. Pan, T., I. Artsimovitch, X. w. Fang, R. Landick, and T. R. Sosnick, *Folding of the Large Ribozyme During Transcription andthe Effect of the Elongation Factor NusA.* PNAS, 1999. **96**(17): p. 9545 - 9550.

117. Dutca, L.M., J. E. G. Gallagher, and S. J. Braserga, *The Initial U3 snoRNA:pre-rRNA Base Pairing Interaction Required for Pre-18S rRNA Folding Revealed by in vivo Chemical Probing.* Nucleic Acids Res, 2011. **39**(12): p. 5164 - 5180.

118. Appling, F.D., D.A. Schneider, and A.L. Lucius, *Multisubunit RNA Polymerase CLeavage Factors Modulate the Kinetics and Energetics of Nucleotide Incorporation: An RNA Polymerase I Case Study.* Biochemistry, 2017. **56**: p. 5654 - 5662.

119. Chen, Y.-J., et al., *Characterization of 582 natural and synthetic terminators and quantification of their design constraints.* Nature Methods, 2013. **10**: p. 659.

120. Reines, D., et al., *Identification of intrinsic termination sites in vitro for RNA polymerase II within eukaryotic gene sequences.* Journal of Molecular Biology, 1987. **196**(2): p. 299-312.

121. Seiberg, M., M. Kessler, A. J. Levine, and Y. Aloni, *Human RNA polymerase II can prematurely terminate transcription of the adenovirus type 2 late transcription unit at a precise site that resembles a prokaryotic termination signal.* Virus Genes, 1987. **1**(1): p. 97 - 116.

122. Resnekov, O. and Y. Aloni, *RNA polymerase II is capable of pausing and prematurely terminating transcription at a precise location in vivo and in vitro.* Proceedings of the National Academy of Sciences, 1989. **86**(1): p. 12-16.

123. Churchman, L.S., and J. S. Weissman, *Nascent Transcript Sequencing Visualizes Transcription at Nucleotide Resolution.* Nature, 2011. **469**: p. 368 - 373.

124. Appling, F.D., C.E. Scull, A.L. Lucius, and D.A. Schneider, *The A12.2 Subunit is an Intrinsic Destabilizer of the RNA Polymerase I Elongation Complex.* Biophys. J., 2018. **114**(11): p. 2507 - 2515.

125. Jaiswal, R., M. Choudhury, S. Zaman, S. Singh, V. Santosh, D. Bastia, and C. R. Escalante, *Functional Architecture of the Reb1-Ter Complex of Schizosaccharomyces pombe.* Proc. Natl. Acad. Sci. USA, 2016. **113**(16): p. E2267 - E2276.

126. Lang, W.H., and R.H. Reeder, *Transcription Termination of RNAPolymerase I Due to a T-rich Element Interacting with Reb1p.* Proc. Natl. Acad. Sci. USA, 1995. **92**(92): p. 9781- 9785.

127. Lang, W.H., and R.H. Reeder, *The REB1 Site is an Essential Component of a Terminator for RNA Polymerase I in Saccharomyces cerevisiae.* Molecular and Cell Biology, 1993. **13**(1): p. 649 - 658.

128. Reeder, R.H., P. Guevara, and J.G. Roan, *Saccharomyces cerevisiae RNA Polymerase I Terminates Transcription at the Reb1 Terminator In Vivo.* Mol Cell Biol. , 1999. **19**(11): p. 7369 - 7376.

129. Reiter, A., S. Hamperl, H. Seitz, P. Merkl, J. Perez-Fernandez, L. Williams, J. Gerber, A. Nemeth, I. Leger, O. Gadal, P. Milkereit, J. Griesenbeck, and H. Tschochner, *The Reb1-homologue Ydr026c/Nsi1 is Required for Efficient RNA Polymerase I Termination in Yeast.* Embo J. , 2012. **31**(16): p. 3480 - 3493.

130. McStay, B., and R.H. Reeder, *A Termination Site for Xenopus RNA Polymerase I Also Acts as an Element of an Adjacent Promoter.* Cell, 1986. **47**: p. 913 - 920.

131. Guerrier-Takada, C., et al., *The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme.* Cell, 1983. **35**(3, Part 2): p. 849-857.

132. James, B.D., et al., *The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme.* Cell, 1988. **52**(1): p. 19-26.

133. Winkler, W.C. and R.R. Breaker, *REGULATION OF BACTERIAL GENE EXPRESSION BY RIBOSWITCHES.* Annual Review of Microbiology, 2005. **59**(1): p. 487-517.

134. McCown, P.J., K. A. Corbino, S. Stav, M. E. Sherlock, and R. R. Breaker, *Riboswitch Diversity and Distribution.* RNA, 2017. **23**(7): p. 995 - 1011.