Rate of transcription elongation and sequence-specific pausing by RNA polymerase I directly influence ribosomal RNA processing

Abigail K. Huffines, Krysta L. Engel, Sarah L. French, Yinfeng Zhang, Olga V. Viktorovskaya, David A. Schneider

PII: S0021-9258(22)01173-5
DOI: https://doi.org/10.1016/j.jbc.2022.102730
Reference: JBC 102730

To appear in: Journal of Biological Chemistry

Received Date: 7 July 2022
Revised Date: 18 October 2022
Accepted Date: 7 November 2022

Please cite this article as: Huffines AK, Engel KL, French SL, Zhang Y, Viktorovskaya OV, Schneider DA, Rate of transcription elongation and sequence-specific pausing by RNA polymerase I directly influence ribosomal RNA processing, Journal of Biological Chemistry (2022), doi: https://doi.org/10.1016/j.jbc.2022.102730.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 THE AUTHORS. Published by Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology.
Rate of transcription elongation and sequence-specific pausing by RNA polymerase I directly influence ribosomal RNA processing

Abigail K. Huffines1*, Krysta L. Engel1#*, Sarah L. French2, Yinfeng Zhang1, Olga V. Viktorovskaya1%, and David A. Schneider1$

1 Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, AL 35294
2 Department of Microbiology, Immunology, and Cancer Biology, University of Virginia Health System, Charlottesville, VA 22908
* These authors contributed equally to this work.
# current: Division of Hematology, University of Colorado Anschutz, Aurora, CO 80014
% current: Moderna, Inc.
$ correspondence: 720 20th Street South
Kaul Human Genetics Building, Room 442
Birmingham, AL 35294
205-934-4781
Email: dschneid@uab.edu

Running Title: Ribosomal RNA processing is linked to Pol I elongation

Keywords: ribosome, transcription, transcription regulation, rRNA processing, RNA polymerase I, NET-seq
ABSTRACT

One of the first steps in ribosome biogenesis is transcription of the ribosomal DNA (rDNA) by RNA polymerase I (Pol I). Processing of the resultant ribosomal RNA (rRNA) begins co-transcriptionally and perturbation of Pol I transcription elongation results in defective rRNA processing. Mechanistic insight regarding the link between transcription elongation and ribosome assembly is lacking due to limited in vivo methods to assay Pol I transcription. Here, we use native elongating transcript sequencing (NET-seq) with a strain of Saccharomyces cerevisiae containing a point mutation in Pol I, rpa190-F1205H, which results in impaired rRNA processing and ribosome assembly. We previously demonstrated that this mutation caused a mild reduction in the transcription elongation rate of Pol I in vitro, however, transcription elongation by the mutant has not been characterized in vivo. Here, our findings demonstrate that the mutant Pol I has an increased pause propensity during processive transcription elongation both in vitro and in vivo. NET-seq reveals that rpa190-F1205H Pol I displays alternative pause site preferences in vivo. Specifically, the mutant is sensitized to A/G residues in the RNA:DNA hybrid and at the last incorporated nucleotide position. Furthermore, both NET-seq and EM analysis of Miller chromatin spreads reveal pileups of rpa190-F1205H Pol I throughout the rDNA, particularly at the 5' end of the 35S gene. This combination of in vitro and in vivo analyses of the Pol I mutant provides novel insights into Pol I elongation properties and indicates how these properties are crucial for efficient co-transcriptional rRNA processing and ribosome assembly.

INTRODUCTION

Ribosome biogenesis requires the intricate coordination of multiple biochemical processes. The first step is transcription by RNA polymerase I (Pol I), which synthesizes ribosomal RNA (rRNA) from a ribosomal DNA (rDNA) template. In Saccharomyces cerevisiae (yeast), Pol I transcribes a single rDNA gene, the 35S, which is organized in approximately 200 tandem repeats. Processing of the rRNA transcripts is complex and requires hundreds of trans-acting factors and RNAs, ultimately giving rise to the three largest mature rRNAs that serve as the backbone of the ribosome. The earliest studies on rRNA processing suggested that it could only occur post-transcriptionally, and this was supported by the detection of full-length rRNA products in the cell (1,2). However, over the past two decades, it was shown that rRNA processing begins co-transcriptionally. The first evidence of co-transcriptional rRNA processing was from Miller chromatin spreads, which allow for the visualization of engaged Pol I transcription elongation complexes with a DNA template in vivo. Findings from that study demonstrated that nascent transcripts were co-transcriptionally cleaved while Pol I was still associated with the rDNA (3,4). Furthermore, it was discovered that when transcription elongation by Pol I is perturbed, there is a corresponding disturbance in rRNA processing and ribosome biogenesis (5). Collectively, these findings demonstrate that co-transcriptional processing is important for the efficient maturation of rRNAs, but the regulation of this process is not well-defined.
The mechanism by which transcription elongation kinetics influences the processing of nascent RNA is not entirely clear. One model proposes that the rate of RNA synthesis may impact the folding of RNA and therefore, the kinetics of co-transcriptional events that occur, such as processing (reviewed in (6-8)). Furthermore, a recent publication reported that there are sequence-specific effects on Pol I transcription rate and rRNA folding (9). Those results demonstrated that Pol I occupies G/C-rich rDNA at a higher frequency compared to A/T-rich, and that weak rRNA structures that form behind the polymerase may allow for more backtracking, which could slow transcription down. Those findings and others (10) suggest that polymerase pausing influences the kinetics of RNA folding and possibly vice versa. Additionally, the efficiency of RNA folding has been shown to have an effect on Pol II transcription rate (11), suggesting that perhaps the coupling of elongation rate, RNA processing, and RNA folding is conserved across the three Pols, or at least Pols I and II.

A number of studies have described the impact of altered polymerase elongation properties on RNA processing, and this has been directly observed for Pol I. Previously, it was established that the mutation of one of the aspartate residues that coordinate Mg$^{2+}$ in the second largest subunit of Pol I, rpa135-D784G, impaired elongation rate in vitro and caused defects in rRNA processing and ribosome biogenesis (5). However, the complete characterization of Pol I transcription elongation and how the disruption of these properties can lead to defects in rRNA maturation and ribosome assembly is lacking in vivo. The co-transcriptional nature of rRNA processing, as well as the previous identification of elongation-defective Pol I mutants makes this an exceptional system to investigate the dependence of co-transcriptional events on polymerase elongation properties ((3,4,12) and reviewed in (13)).

To explore the connection between Pol I elongation kinetics and co-transcriptional events further, we characterized the effects of a Pol I mutant that contains a point mutation within the trigger loop (TL) region. The TL is a conserved domain within the active site of multi-subunit Pols across all domains of life and is known to directly affect several properties, including transcription elongation rate and pausing propensity [reviewed in (14-16)]. We have previously shown that this mutation, rpa190-F1205H, reduced elongation rate by Pol I in vitro (17). Here, we thoroughly characterize the transcription elongation efficiency of the wild-type (WT) and mutant polymerases using single-turnover in vitro transcription assays with a selection of rDNA templates. These assays revealed sequence-specific pause sites for the WT enzyme and enhanced pausing by the mutant. To characterize pausing in vivo, we used two separate methods: native elongating transcript sequencing (NET-seq) and electron microscopy of Miller chromatin spreads. These complementary methods reveal that this mutation induces significant changes to Pol I occupancy patterns throughout the 35S gene, which could be responsible for the observed dramatic changes in rRNA processing and subsequent impairments of ribosome biogenesis.

Collectively, these data suggest that increased pausing by Pol I results in aberrant co-transcriptional processing (possibly due to changes in rRNA secondary structure) and ribosome assembly events. These results support the model that transcription elongation is an important aspect of gene expression that not only influences the quantity but also
the quality of the RNA that is produced. Since both the WT and rpa190-F1205H polymerases display sensitivities to particular sequences, our findings also demonstrate that Pol I has evolved a unique set of transcription elongation properties that respond to the rDNA sequence to promote efficient ribosome biogenesis.

RESULTS

The rpa190-F1205H mutation induces sequence-specific pausing by Pol I in vitro

To investigate the effect of rDNA sequence elements on Pol I activity, we focused on a phenylalanine to histidine point mutation in a highly conserved position in the TL, rpa190-F1205H, and assayed the activity of this enzyme as compared to WT in single-turnover in vitro transcription assays. We have shown previously that the transcription elongation rate of rpa190-F1205H Pol I is ~40% of the WT rate (17), however, this enzyme has not yet been fully characterized. For templates in these reactions, we used two different regions of the S. cerevisiae rDNA locus that were cloned downstream of the Pol I core promoter (see details in the methods section and (5)). Downstream of the promoter in both templates, we included a 55-nucleotide region that does not encode any C residues in the non-template strand. Thus, we initiated transcription by incubation with 10 µM ATP, GTP, and UTP, resulting in the synchronization of elongation complexes at the first encoded C residue (at position +56, relative to the transcription start site; (17,18)). To measure transcription elongation properties of the polymerases, 10 µM CTP was added to the reactions and samples were collected as a function of time. We observed obvious sequence-specific pausing by WT Pol I at several regions within the rDNA templates (observed as distinct bands of lower molecular weight that appear and dissipate over the time course, Figure 1)). The simplest interpretation for the varied band intensities of these pauses is that the propensity for Pol I to enter the pause and escape it are based on the sequence, and some sequences confer a longer pause than others. These results demonstrate that Pol I is prone to sequence-specific pausing within the rDNA template. Since these reactions were performed with only the essential purified proteins required for in vitro transcription, these pause events are likely not dependent on trans-acting protein factors.

On each template, the rpa190-F1205H Pol I required slightly more time to produce full-length runoff product than WT Pol I, which is consistent with our previous studies demonstrating that the mutant enzyme has a reduced elongation rate (17). In addition to our previous studies showing that there was a reduction in the overall rate of transcription elongation by the mutant Pol I, here we observed altered pausing by the rpa190-F1205H enzyme. Specifically, the duration of several pause sites was lengthened when transcribed by the mutant polymerase (indicated by asterisks, Figures 1B and C). These data demonstrate that the mutation of the TL in Pol I induces pausing by the enzyme on the rDNA sequence, relative to WT Pol I, even in a highly purified system.

rpa190-F1205H affects Pol I activity and promotes pausing in vivo

The in vitro biochemical data demonstrate that the rpa190-F1205H mutation not only reduces Pol I transcription elongation rate (17), but also affects polymerase pausing on
the native rDNA sequences. To determine whether this phenomenon is also observed in vivo, we performed electron microscopy of Miller chromatin spreads and NET-seq. These techniques allow for the direct visualization of Pol I transcription on individual rDNA repeats (Miller chromatin spreads), as well as precise mapping of the positioning of individual Pol I molecules throughout the rDNA at single nucleotide resolution (NET-seq).

Electron micrographs of Miller chromatin spreads reveal snapshots of rDNA transcription in vivo. We and others have previously used this technique to evaluate a number of features of Pol I transcription, including polymerase pausing and nascent transcript processing (4,17,19,20). Historically, the analysis of Miller chromatin spreads has played a key role in demonstrating that rRNA processing and pre-ribosome assembly begins co-transcriptionally (3-5). Therefore, this is a valuable tool to closely investigate features of transcription elongation by Pol I. Since Pol I is densely packed on the active rDNA repeats, the effects of pausing are readily observed. If one polymerase enters a paused state while the enzymes downstream continue to elongate, an extended gap between polymerases is evident. Therefore, the presence of such gaps suggests an increase in Pol I pausing during transcription. Long gaps have been observed at a relatively low frequency in WT cells (19). We examined Miller chromatin spreads of rDNA from WT and rpa190-F1205H yeast strains to evaluate the effects of the mutation on Pol I activity in vivo (Figure 2A and Supplemental Figure 1). The analyzed Pol I distribution on the rDNA template from the rpa190-F1205H strain appeared strikingly different than those from WT yeast. First, large polymerase-free gaps were frequently observed on the 35S gene in the mutant, indicating an increased pause propensity for Pol I in the mutant vs. WT strain. Second, many analyzed Miller spreads had fewer Pol I elongation complexes in the 3’ portion of the rDNA gene relative to the 5’ portion. Finally, we observed that nascent transcripts appeared shorter in the mutant as compared to WT, particularly at the 3’ end of the rDNA template. The simplest explanation for this qualitative observation is that the rRNA may be degraded by endo- or exonucleases co-transcriptionally.

To quantify potentially long pauses, we measured the frequency of large gaps between transcription elongation complexes (>23% of the 35S gene length; as performed previously (19)). Large gaps occurred at nearly twice the frequency in the mutant as compared to WT (21% of 185 35S genes analyzed and 12% (19) of 302 35S genes analyzed, respectively, Figure 2B). We observed more gaps in the downstream half of the rDNA repeats in the mutant as compared to WT. Additionally, there was a reduction in gaps at the 5' end of the rDNA template in the rpa190-F1205H strain, suggesting that Pol I transcription elongation rate may be uniformly slower and pausing may be increased at the beginning of the 35S gene. Finally, the rpa190-F1205H Pol I appears to be more likely to accumulate in dense clusters on the rDNA template compared to WT Pol I, which is distributed more evenly across the entire 35S gene. Given that each active yeast rDNA repeat has a large number of engaged Pol I elongation complexes, these dense clusters could represent polymerase “traffic jams”, with rpa190-F1205H Pol I complexes becoming backed up on the template due to increased pausing. This in vivo evidence for pausing is consistent with in vitro observations (Figure 1).

NET-seq reveals that rpa190-F1205H Pol I is repositioned on the rDNA
To further investigate the in vivo effects on transcription conferred by the \textit{rpa190-F1205H} mutation, we used NET-seq. This technique was first developed only about a decade ago to examine transcription by Pol II at high-resolution in vivo (21). Our lab adapted these methods to investigate the properties of transcription by Pol I (22), and we have previously used it to demonstrate the role of various factors in transcription elongation (23,24) and to characterize the mechanism of action of a small-molecule Pol I inhibitor in vivo (25). Therefore, this is a powerful tool to interrogate features of transcription by Pol I, especially when characterizing a mutation such as \textit{rpa190-F1205H}. For these experiments, we generated triplicate NET-seq libraries for both the WT and \textit{rpa190-F1205H} yeast strains. We mapped the resultant reads back to the yeast genome and plotted the polymerase position on the rDNA for three replicates for each strain (Supplemental Figure 2, left and middle panels of A and B). Spearman correlation coefficient values were generated (Supplemental Figure 2, right panels of A and B) to determine the similarity between replicates (as described previously (22-25)). Consistent with previous publications from our lab (22-25), we found that within strains, Pol I occupancy was highly reproducible for both WT and \textit{rpa190-F1205H} yeast (indicated by similar overlays in the histograms and Spearman correlation coefficient values > 0.9). We performed principal component analysis (Supplemental Figure 3), which supported this conclusion, as the WT and mutant samples clustered away from each other, and the majority of the divergence (over 94%) between all samples was described by PC1, which represented the variation between WT and mutant libraries.

After confirming that Pol I occupancy was reproducible within strains, we further examined occupancy patterns between strains. The median Pol I occupancy for each strain was plotted along the rDNA repeat, and a t-test was performed for the median Pol I count at each individual rDNA position between strains to determine whether there was a significant difference in the \textit{rpa190-F1205H} strain as compared to WT (Figure 3A). Strikingly, there was a significant difference in occupancy across the 35S gene between strains, as indicated by the colors included below the histogram (green for increased occupancy and black for decreased occupancy in the mutant as compared to WT). From these NET-seq data, we cannot draw definitive conclusions about the kinetics of Pol I since each experiment represents a snapshot of transcription. However, we can interpret highly occupied regions as areas where Pol I is potentially paused on the template or is transcribing more slowly than positions showing lower occupancy, which could be interpreted as areas where Pol I is traversing the rDNA template more rapidly. Additionally, very low regions of occupancy could indicate that the polymerases have prematurely disengaged with the template, especially if they are widespread. Figure 3A suggests that in the \textit{rpa190-F1205H} strain, there is a significant increase in Pol I occupancy as compared to WT, and this increase seems more exaggerated at the 5' end of the 35S gene (see the external transcribed spacer 1 (ETS1)) as compared to the 3' end (see the external transcribed spacer 2 (ETS2)). This increase in occupancy is especially noticeable in the spacer regions (as evidenced by the almost solid green markers in those regions in Figure 3A vs. the mixture of black and green markers in the gene regions), though it is present across the length of the entire gene in every region. This is likely due to unavoidable mature rRNA product contamination in the NET-seq samples, a limitation of this experiment that we have discussed in a previous publication.
(25). During rRNA synthesis, the spacer regions are rapidly co- and post-transcriptionally processed and removed prior to rRNA maturation (2,4,5,12). Therefore, evaluation of Pol I occupancy in the spacer regions allows for the direct measure of nascent rRNA, without interference from contaminating mature rRNA. Therefore, we investigated the occupancy patterns in these four spacer regions (ETS1, internal transcribed spacer 1 (ITS1), internal transcribed spacer 2 (ITS2), and ETS2) more extensively to examine occupancy patterns and diminish mature product contamination as much as possible. Figure 3B indicates that there is an obvious increase in occupancy in the ETS1 and ITS1 (the most 5’ regions) in the rpa190-F1205H strain as compared to WT, with a modest occupancy increase in the ITS2 region and some occupancy differences present in the most 3’ region, the ETS2. Additionally, the Kolmogorov-Smirnov test (K-S test) was used for statistical comparison in these regions, where a p-value < 0.05 indicates that the patterns between strains are not from the same distribution. The K-S test demonstrated that the distribution patterns were significantly different between strains in all four spacer regions. The occupancy differences observed in Figure 3 were validated by generating moving average and cumulative distribution function plots (Supplemental Figures 4 and 5, respectively). Collectively, these data suggest that consistent with Figure 2, rpa190-F1205H Pol I pauses more frequently than WT Pol I in vivo, especially at the 5’ end of the 35S gene, evidenced by increased occupancy.

To gain insight into features of the rDNA that influence the strain-altered Pol I occupancy patterns shown in Figure 3, we took advantage of the high-resolution nature of NET-seq data and determined the sequence preferences for each strain by generating a DiffLogo (Figure 4). The top 2.5% occupied positions in the spacer regions were identified in both strains and the highly occupied sequences were displayed for the mutant strain on top (JS divergence < 0), compared to the WT strain below (JS divergence > 0). Just like in Figure 3B, only the spacer regions were included in this analysis to reduce sequence enrichments coming from mature rRNA product contamination. Figure 4 demonstrates that there is polymerase repositioning on the rDNA template between strains, as rpa190-F1205H Pol I readily occupies A/G-rich regions which are present in the RNA:DNA hybrid, as compared to the C/T-rich regions occupied by WT Pol I. The RNA:DNA hybrid contributes directly to the stability of transcription elongation complexes (26,27), so changes in sequence preferences in the hybrid region could directly impact pausing and nucleotide addition. Altogether, these NET-seq results suggest that rpa190-F1205H Pol I exhibits impaired transcription elongation as compared to WT Pol I in vivo, and these findings could explain some of the pausing effects seen in Figures 2 and 3.

**rRNA processing and ribosome biogenesis are defective in rpa190-F1205H yeast**

Given the previously characterized relationship between Pol I elongation defects and rRNA processing and ribosome biogenesis (5), we hypothesized that Pol I elongation rate and pausing frequency influence these co-transcriptional processes. Figures 1-4 demonstrate that rpa190-F1205H Pol I exhibits transcriptional defects and increased pausing both in vitro and in vivo. Therefore, we tested our hypothesis by determining whether ribosome biogenesis was impaired in the rpa190-F1205H yeast strain. We used Northern blot analysis to measure the level of precursor and mature rRNAs in both strains. In the rpa190-F1205H strain, we observed a reduction in the level of precursor rRNAs for
both the 18S and 25S (Figure 5A). This loss of signal was most likely due to degradation of the pre-rRNAs. However, since exosome-mediated RNA decay occurs in the 3' to 5' direction and we are confined to using a Northern probe that targets the 3' end of the species (Supplemental Figure 6), degradation of the 23S and 20S pre-rRNAs is not detectable in this experiment. Alternatively, the 5' location of the 27S probe yielded clear evidence for the degradation of this rRNA species in the mutant strain, but not in WT (Figure 5A, and Supplemental Figure 6). Quantification of this degradation was performed by measuring the ratio of the amount of full-length 27S relative to total signal within the lane. These results identified significant degradation of the 27S precursor in the mutant vs. WT (Figure 5B). To demonstrate that these processing defects are a direct result of perturbations in transcription elongation, we previously analyzed processing intermediates in exosome (rrp6) mutants (5). We found those results to be consistent with our findings in this study. From these data, we conclude that the disruption of Pol I transcription elongation results in defective processing of the nascent rRNA and induces degradation of intermediates.

To evaluate consequences of the rpa190-F1205H mutation on the populations of ribosomal subunits, we performed sucrose gradient centrifugation. Sucrose gradient centrifugation provides an overall view of the abundance and activity of cytoplasmic ribosomes involved in translation. This technique is used to separate ribosomal subunits, fully intact ribosomes, and polyribosomes, and the resultant data can be used to make qualitative predictions about the efficiency of ribosome assembly and function in cells. Using this technique, we observed a substantial reduction in the relative abundance of 60S ribosomal subunits (vs. either 40S subunits or 80S monosomes) in the rpa190-F1205H cells compared to WT (Figure 5C and Supplemental Figure 7). Consistent with the imbalanced proportion of the large subunit population relative to that of the small subunit, we also observed half-mers, which indicate that a small subunit has engaged the mRNA template but has yet to associate with a large subunit. These half-mers appear as shoulder peaks detectable on the monosome and disome peaks (labelled in Figure 5C, model of half-mer formation in Figure 5D). Thus, the appearance of half-mers is consistent with the observation of reduced 60S subunit abundance relative to 40S subunits in the mutant strain.

Sucrose gradient analyses were performed in triplicate, and representative traces are shown in Figure 5C. Quantification of sucrose gradient traces is complicated because gradients are loaded for equal amounts of material that absorb UV light at 260 nm. Since many molecules absorb UV, the baseline absorption does not uniquely reflect ribosome abundance. Thus, the absolute amplitude of any position in the trace does not necessarily reflect a biologically significant difference (e.g., the apparent increase in polysomes in the mutant trace). When the baseline is calculated by extrapolating between “valleys” in the trace, we determined that there was approximately a 4-fold decrease in the average 60S abundance relative to 40S in the mutant strain, as compared to an almost 3-fold increase in WT yeast (Supplemental Figure 7). These observations demonstrate that in rpa190-F1205H yeast, the production of the 60S subunit is preferentially impaired compared to the 40S, which leads to inefficient translation initiation. Altogether, these findings
demonstrate that transcription elongation is directly coupled to rRNA processing and ribosome assembly.

**DISCUSSION**

**Pol I transcription elongation properties directly affect ribosome biogenesis**

Co-transcriptional processing of nascent RNA has been appreciated for many years (for example, the 5' capping of mRNA). Interestingly, the elongation phase of transcription has also been implicated in events such as alternative splicing (8,28,29). Pre-rRNA processing serves as an excellent model system to study the connection between transcription elongation properties and RNA processing efficiency due to its cellular abundance, intimate connection to cell viability, and clearly defined co-transcriptional nature. In this study, we used a Pol I TL mutant, *rpa190-F1205H*, which exhibits a reduction in transcription elongation rate (17). Characterization of the transcription elongation properties of *rpa190-F1205H* Pol I revealed an increased tendency to pause both *in vitro* and *in vivo*. Consistent with previous studies investigating elongation-defective Pol I mutants (5), rRNA processing and ribosome biogenesis were impaired by this mutation. These results suggest that the transcription elongation properties of Pol I have co-evolved with early rRNA processing steps to promote efficient co-transcriptional processing of rRNA and ribosome assembly.

Previous results suggested that Pol I transcription elongation directly influences the efficiency of ribosome biogenesis (30-32). A critical finding in support of this idea was the characterization of an elongation-impaired Pol I mutant, *rpa135-D784G* (5). This mutant exhibited a ~10-fold-defect in transcription elongation rate *in vitro* compared to WT Pol I. Furthermore, in the *rpa135-D784G* yeast strain, there was aberrant rRNA processing and defects in ribosome biogenesis (evidenced by half-mers and a reduced 60S/40S ratio). A separate study demonstrated that deletion of the gene that encodes for the Rpa49 subunit of Pol I also resulted in accumulation of half-mers and a reduction in 60S compared to 40S subunit abundance (20). Those findings showed that *rpa49Δ* Pol I had a reduced elongation rate as compared to WT, which resulted in increased spacing between polymerases and suggested additional torsional stress in the rDNA. In WT yeast cells, Pol I occupancy on the rDNA is high and the polymerases are densely packed, with about 50 elongation complexes per 35S gene (33). The close proximity between Pol I elongation complexes prevents accumulation of positive supercoils in front and negative supercoils behind each polymerase. Positive supercoiling in front of polymerases could present an obstacle during transcription elongation, and severe supercoiling could even inhibit transcription altogether. These data highlight an important question: is the elongation rate, pausing propensity of Pol I, or spacing between polymerases important for efficient rRNA processing and ribosome biogenesis? One study was performed to test this via mutation of an important preinitiation complex factor, Rrn3. The findings from that study demonstrated that increased spacing between polymerases (due to impaired transcription initiation), does not lead to ribosome assembly defects (5). Therefore, we propose that transcription elongation properties rather than Pol I spacing are critical for co-transcriptional processes. Along these lines, the effect of a Pol I mutant which exhibits
a gain-of-function phenotype (such as increased elongation rate) on ribosome biogenesis would be a further test of this hypothesis. Indeed, one would hypothesize that an alteration of elongation rate through either a decrease (as shown here) or increase would negatively impact rRNA processing and ribosome biogenesis. Interestingly, a hyperactive Pol I mutant has been described in recent literature (34). Those findings demonstrated that the Pol I mutant, rpa135-F301S, produced more rRNA overall compared to WT Pol I both in vitro and in vivo, but the effects of this mutation on rRNA processing have not yet been characterized.

**Which features of transcription elongation influence rRNA processing?**

The goal of this study was to investigate the importance of elongation rate versus pausing with respect to co-transcriptional events; however, the two features are difficult to disentangle. To influence transcription elongation rate directly, the Pol I catalytic functions must be impaired. As a result, we expect most elongation-defective enzymes to exhibit altered pausing properties as well. Additionally, any increase in pausing will inherently decrease overall transcription elongation rate. Interestingly, the rpa190-F1205H mutation did not simply increase the dwell time of pause sites already exhibited by the WT enzyme in vivo, rather, the mutation resulted in a redistribution of pausing sites with preference for the 5' end of the rDNA template (Figure 3). Based on these data, we conclude that the introduction of abnormal pause sites for Pol I, together with a modest defect in transcription elongation rate, cause severe defects in rRNA processing.

**Does Pol I pausing facilitate co-transcriptional ribosome assembly?**

This study focused on the characteristics of the rpa190-F1205H Pol I and how defective transcription elongation properties can perturb rRNA processing events. Since NET-seq data can be viewed as a snapshot of Pol I occupancy in vivo, one may interpret large peaks as sites of polymerase pausing. Therefore, Figure 3 demonstrates the heterogeneity of Pol I occupancy in vivo, consistent with previous findings (9,22-25,37),
and indicates that Pol I faces barriers to transcription elongation, even though the rDNA is generally thought to be mostly free of phased nucleosomes (38). Whereas these data highlight an apparent nucleotide incorporation preference and multiple sites of increased Pol I occupancy throughout the 35S gene, there are many open questions: do these pause sites play a role in co- or post-transcriptional processing? If so, are they conserved among eukaryotic species? We posited previously that because sequence similarity and secondary structure proximal to the active site are highly conserved throughout both eukaryotic and prokaryotic DNA-dependent RNA polymerases, sequence-dependent pausing effects observed in yeast Pol I might be conserved in other polymerases as well. On the other hand, the primary sequence for the rDNA in various species can differ a great deal (though core mature secondary structure is largely conserved). Are sequence effects consistent for Pol I throughout Eukarya? All of these questions are raised by the novel capabilities provided by experimental approaches like NET-seq. As these questions are answered over time, we expect to determine the significant evolutionary constraints applied to the rDNA to optimize ribosome function as well as assembly and potentially regulation of ribosome biosynthesis.

**Are trans-acting factors involved in coordinating Pol I elongation and co-transcriptional ribosome assembly?**

These data suggest that Pol I elongation kinetics are important for co-transcriptional processes, therefore, it is conceivable that the actions of transcription factors, which modulate Pol I elongation, may also affect ribosome assembly. Indeed, one such factor might be Paf1. It was previously established that the deletion of genes that encode Paf1 subunits leads to reduced Pol I elongation rate and aberrant co-transcriptional rRNA processing (19). Additionally, Spt4 and 5 have been shown to influence Pol I elongation *in vivo* (23), and deletion of Spt4 results in a reduction in Pol I elongation rate and rRNA processing defects (39). These studies suggest that various transcription factors may facilitate proper elongation and possibly co-transcriptional ribosome assembly.

The effects of transcription factors on pausing have previously been shown to alter the folding of RNA co-transcriptionally. In *E. coli*, the folding of RNase P domains was influenced by the transcription elongation factor NusA (10). These data suggest that transcription factors can directly affect co-transcriptional RNA processing and maturation. It is reasonable to suggest that in cells, trans-acting factors might manipulate pause probability or rates of nucleotide addition to tune the assembly or activity of processing factors on the nascent rRNA. Currently, we have only begun to reveal the complex relationship between transcription elongation and the myriad of steps required to build functional eukaryotic ribosomes.

**EXPERIMENTAL PROCEDURES**

**Yeast strains, plasmids, and growth conditions**

Yeast cultures were grown at 30°C in standard growth conditions and yeast strain manipulation was performed using standard methods (40). For isolation of *rpa190-F1205H*, the mutation was introduced into a pBlueScript derivative carrying a region of
**In vitro transcription**

Pol I was purified and in vitro experiments were performed as previously described (17,18). Briefly, the template sequence included the native promoter and rDNA sequence, except for a stretch of six Gs in the template strand. This enables synchronization of the polymerases at that position with the omission of CTP and is key for elongation rate assays. The sequences used in vitro were cloned from positions in the rDNA 5’ ETS (+405 to +604) and 18S (+2246 to +2445) and placed immediately downstream of the native rDNA promoter and the initially transcribed 55-nucleotide C-less stretch. The rDNA position is numbered with respect to the transcription start site. In vitro reactions (20 µL each) were performed in 1X transcription buffer (20 mM Tris-OAc (pH 7.9), 100 mM potassium glutamate, 8 mM magnesium acetate, 2 mM DTT, RNasin Plus (Promega), 0.2 mg/mL acetylated bovine serum albumin), 10 µCi of [α\(^{32}\)P] UTP (Perkin Elmer), 10 µM ultrapure NTPs (GE Life Sciences), and quenched with 1 mL of 1.25 M ammonium acetate in 95% ethanol.

**Miller chromatin spreads**

Electron microscopy of Miller chromatin spreads was performed as previously described (33) and quantification of gaps was performed as described in (19). Briefly, a ‘large gap’ was defined as a stretch of rDNA that could accommodate 40 polymerases (~25% of the total rDNA length). The position of the start (promoter-proximal edge) of the gap was determined. The data was binned in intervals of 10% of the gene length (~670 bp/bin) and the frequency of observing the start of a gap within each bin was plotted at the position of the midpoint of the bin. The rDNA repeats shown are representative of the populations observed from mutant or WT cells.

**Native elongating transcript sequencing (NET-seq)**

NET-seq was performed exactly as previously described (25) using the rpa190-F1205H strain with an additional 3x-HA tag on the C-terminus of RPA135 and the appropriate parent strain. Briefly, one liter cultures (a total of three liters per strain) were harvested and lysed under cryogenic conditions. Pol I elongation complexes were immunoprecipitated and the RNA was extracted. A DNA linker containing a unique molecular identifier (UMI) was ligated onto the 3’ end of the isolated nascent transcripts (23) to preserve the last incorporated nucleotide position and then reverse transcribed. Finally, the complementary DNA (cDNA) was circularized and amplified via PCR to produce high-throughput sequencing libraries using a NextSeq 500 as previously described (22,23). Amplification primers for each sample are included in Supplemental Table 1.
NET-seq data analysis was performed exactly as published, using the same options and function settings as previously described (23-25). In summary, reads were deduplicated based on the UMI sequence as previously described using fqtrim (version 0.9.7, (41)). Next, the 5’ and 3’ adaptors were trimmed off of resultant reads using cutadapt (version 3.4 (42)). Reads were aligned to the yeast genome (*Saccharomyces cerevisiae* genome assembly 64-1-1) using the STAR aligner (version 2.7.1a (43)). Finally, resultant BAM files were sorted and indexed using SAMTools (version 1.6 (44)), converted to BED files, and genome coverage files were generated using BEDTools (version 2.28.0 (45)).

Following genome coverage file generation, all data analysis and visualization was performed using R (version 4.1.3) and RStudio (version 2022.02.1). An organized data frame containing coordinate number, region identity, nucleotide, and normalized counts was created exactly as previously described (23). Histograms were generated either by using the built-in plotting function of R or with ggplot2 (version 3.3.5), and Spearman correlation coefficients were calculated using R. The DiffLogo package (version 2.18.0) was used to create the DiffLogo plot. A complete list of software packages and versions used are available in Supplemental Table 2. Raw data can be accessed through NCBI's Gene Expression Omnibus via series accession number GSE196146. The data processing code and R scripts available upon request.

**Northern blot analysis**

RNA was isolated from cells harvested during exponential phase by phenol-chloroform extraction followed by an ethanol precipitation. The RNA pellets were suspended in water and the concentration of RNA was normalized between samples. Equal amounts of RNA were run on 0.9% formaldehyde:agarose gels and then transferred to a nylon membrane. The Northern procedure was performed using an HB-500 Minidizer hybridization oven (UVP). The membrane was pre-hybridized at 65ºC for 4 hours in 20 mL hybridization buffer (0.5 M sodium phosphate (pH 7.2), 7% SDS (w/v)). Oligo probes (0.5 µM) were labeled using [γ-32-P] ATP and polynucleotide kinase (NEB) for 1-2 hours. The entire oligo labelling reaction (20 µL) was added to the hybridization buffer and allowed to hybridize at 65ºC overnight. Prior to washing, the blot was incubated at room temperature for 2 hours and then washed in 20 mL wash buffer (40 mM sodium phosphate (pH 7.2), 1% SDS (w/v)) for 15 minutes at room temperature, 10 minutes at 42ºC, and then 15 minutes at room temperature. The labelled blots were exposed to a phosphor screen and developed using a phosphorimager (Typhoon Trio, GE). Bands were quantified using ImageQuant. Including the sequence of the probes used, this technique was performed as previously described (17). Hybridization probe sequences are included in Supplemental Table 3 and a probe map can be found in Supplemental Figure 6.

**Sucrose gradient sedimentation**

The separation of ribosomal subunits and polysomes was performed using sucrose gradient sedimentation. Immediately prior to harvesting, cells were treated with cycloheximide (100 µg/mL, final concentration). Cells were lysed in breakage buffer (20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 30 mM MgCl2·6H2O, 0.1 mg/mL cycloheximide, 0.2 mg/mL heparin) using a FastPrep with four cycles at 4.5 m/s. Sucrose gradients were
poured in a solution base of 50 mM Tris-HCl (pH 7.5), 50 mM ammonium chloride, 12 mM MgCl₂·6H₂O, and 1 mM DTT, with 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, and 5% sucrose (from bottom to top, respectively). Lysates were normalized to RNA concentration (A₂₆₀), loaded onto the top of the sucrose gradient, and spun at 30,000 rpm for 5 hours at 4°C in a Beckman SW41 rotor. Finally, 60% sucrose (pumped through a pierced bottom of the column) was used to push the gradient through an FPLC and absorbance of the elution was recorded at 254 nm.

DATA AVAILABILITY

All data described within the manuscript are included in this article and supporting information. NET-seq raw data has been deposited into NCBI’s Gene Expression Omnibus, and can be accessed via series accession number GSE196146.

SUPPORTING INFORMATION

This article contains supporting information.

ACKNOWLEDGEMENTS

We thank Ann L. Beyer for her assistance with the analysis of the Miller chromatin spreads in this study. We thank the current and former members of the Schneider lab for their input to these studies. We thank the UAB fermentation facility for growth of yeast cells used for purification of polymerases. We are also grateful to the Heflin Genomic Center and the resources provided by the University of Alabama at Birmingham IT-Research Computing group for high performance computing (HPC) support and CPU time on the Cheaha computer cluster. We thank Dr. Yvonne Edwards at UAB for assistance with bioinformatic analysis.

AUTHOR CONTRIBUTIONS

This study was conceived by DAS, KLE, and OVV. The experiments were conducted by AKH, KLE, YZ, SLF, and OVV. Data analysis was performed by AKH, KLE, SLF, and DAS. The manuscript was written by AKH, KLE, and DAS.

FUNDING AND ADDITIONAL INFORMATION

This study was supported by the following grants from the National Institutes of Health: GM084946 and GM140710 to D.A.S. and GM008111 to K.L.E. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.
REFERENCES


FIGURE LEGENDS

**Figure 1.** *rpa190-F1205H* Pol I has an increased pause propensity *in vitro.* Purified *rpa190-F1205H* and WT Pol I were used to determine the pause propensity *in vitro*. The reactions were performed using 10 μM NTPs and the RNA transcripts were labeled using [α-32P] GTP. **A.** A schematic (drawn to scale) indicating the positions of individual templates used for transcription in B and C. **B and C.** Asterisks denote sites of polymerase pausing that are exaggerated by the *rpa190-F1205H* mutation. Assays on each template were performed at least four times, with representative gels shown for each.

**Figure 2.** Miller chromatin spreads identify defects in Pol I distribution on the rDNA in *rpa190-F1205* yeast. **A.** Miller chromatin spreads of WT and *rpa190-F1205H* yeast strains. The representative WT rDNA gene shows the typical features of Pol I transcription, including high polymerase density, relatively uniform distribution of transcripts across the gene, and nascent transcript cleavage (cleaved transcripts indicated by bracket). The representative *rpa190-F1205H* genes show large polymerase-free gaps, a higher polymerase density at the 5' end of the genes, and some evidence for nascent transcript degradation (in regions denoted by arrows). See Supplemental Figure 1 for comparison of these three rDNA repeats with additional repeats from the *rpa190-F1205H* strain. **B.** Positions of large polymerase-free gaps observed in chromatin spreads from WT and *rpa190-F1205H* yeast. The frequency of start sites (promoter-proximal ends) of gaps large enough to encompass at least 40 polymerases are plotted vs. position along a normalized rDNA gene. The absence of start sites for gaps at the 3' end of the gene results from the fact that only large gaps representing a quarter of the gene were scored. Data for the WT strain has been reported previously (19).

**Figure 3.** NET-seq indicates an altered pause profile in *rpa190-F1205H* yeast. **A.** The median Pol I occupancy in the WT (blue) and *rpa190-F1205H* (red) strains were mapped to the rDNA gene. At each position, a t-test was executed to determine whether there was a significant difference in occupancy (*p* < 0.05) between strains. These results (green (increased), black (decreased), and white (no change)) are indicated below the histogram for the *rpa190-F1205H* strain with respect to WT. **B.** Histograms depicting median Pol I occupancy differences in the spacer regions only (ETS1, ITS1, ITS2, and ETS2). The Kolmogorov-Smirnov (K-S) test was used to determine whether the patterns in each strain were from the same distribution (see *p*-values located above each graph).

**Figure 4.** The *rpa190-F1205H* mutation alters the nucleotide preference at pause sites *in vivo.* A DiffLogo plot is shown, which demonstrates the sequence preference differences for the top 2.5% occupied positions in the spacer regions only for the *rpa190-F1205H* strain (top, JS divergence < 0) and the WT strain (bottom, JS divergence > 0). The black arrow indicates the position of the last incorporated nucleotide into the transcript. Red asterisks indicate positions of significance (*p*-value < 0.05), indicating that the sequences at that position are not from the same distribution.
Figure 5. The *rpa190-F1205H* mutation induces rRNA processing and ribosome biogenesis defects.

A. Northern blot analysis of precursor and mature rRNAs. The *rpa190-F1205H* mutation leads to a loss of the precursor rRNAs. B. Degradation of the 27S species is observed as the smear of signal below the full-length 27S. The degradation of 27S was quantified by measuring the ratio of full-length 27S divided by the total lane signal (n=4). C. Polysome profile analysis (sucrose gradient centrifugation) was performed on WT and mutant strains. Free ribosomal subunits (40S and 60S) were separated from monosomes (80S) and polysomes. Half-mers (H) were detected in the *rpa190-F1205H* strain, along with an altered relative abundance of 60S. The traces shown here are representative of three biological replicates (n=3). D. Schematic indicating how a half-mer is formed.
CONFLICT OF INTEREST
Authors declare no competing interests.
AUTHOR CONTRIBUTIONS