784–802 *Nucleic Acids Research, 2022, Vol. 50, No. 2* *Published online 30 December 2021*
https://doi.org/10.1093/nar/gkab1262

The interaction between the Spt6-tSH2 domain and Rpb1 affects multiple functions of RNA Polymerase II

Zaily Connell[1], Timothy J. Parnell[2], Laura L. McCullough[1], Christopher P. Hill[1] and Tim Formosa[1,*]

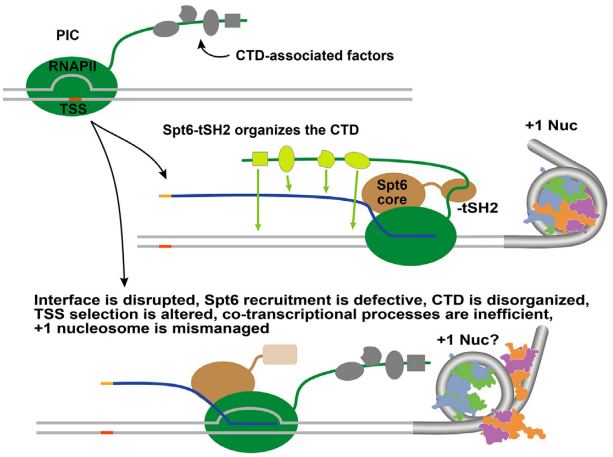
[1]Dept of Biochemistry, University of Utah School of Medicine 15 N Medical Drive, Rm 4100, Salt Lake City, UT 84112, USA and [2]Huntsman Cancer Institute, 2000 Circle of Hope, Salt Lake City, UT 84112, USA

Received May 18, 2021; Revised November 29, 2021; Editorial Decision December 06, 2021; Accepted December 09, 2021

ABSTRACT

The conserved transcription elongation factor Spt6 makes several contacts with the RNA Polymerase II (RNAPII) complex, including a high-affinity interaction between the Spt6 tandem SH2 domain (Spt6-tSH2) and phosphorylated residues of the Rpb1 subunit in the linker between the catalytic core and the C-terminal domain (CTD) heptad repeats. This interaction contributes to generic localization of Spt6, but we show here that it also has gene-specific roles. Disrupting the interface affected transcription start site selection at a subset of genes whose expression is regulated by this choice, and this was accompanied by changes in a distinct pattern of Spt6 accumulation at these sites. Splicing efficiency was also diminished, as was apparent progression through introns that encode snoRNAs. Chromatin-mediated repression was impaired, and a distinct role in maintaining +1 nucleosomes was identified, especially at ribosomal protein genes. The Spt6-tSH2:Rpb1 interface therefore has both genome-wide functions and local roles at subsets of genes where dynamic decisions regarding initiation, transcript processing, or termination are made. We propose that the interaction modulates the availability or activity of the core elongation and histone chaperone functions of Spt6, contributing to coordination between RNAPII and its accessory factors as varying local conditions call for dynamic responses.

GRAPHICAL ABSTRACT



INTRODUCTION

The central domain of Spt6 is a conserved transcription elongation factor found in eubacteria, archaebacteria and eukaryotes (1,2). In eukaryotes, this core domain is flanked by an inherently disordered region of ∼300 residues at the N-terminus and two Src-homology 2 (SH2) motifs at the C-terminus (2–6, Figure 1A). The N-terminal domain in the yeast *Saccharomyces cerevisiae* binds the essential transcription factor Spn1, and also binds nucleosomes and histones, providing at least part of Spt6's histone chaperone activity (3,7). The two C-terminal SH2 motifs form a single structural unit, called the tandem SH2 or tSH2 domain (5). The prototype SH2 motif in *Src* binds phosphorylated tyrosine residues, but the two domains in Spt6 bind a phospho-threonine and a phospho-serine located in the RNAPII catalytic subunit, Rpb1/Rpo21 (8). These targets are in a roughly 30-residue region that links the globular domains of the RNAPII catalytic core and the heptad repeats that comprise the Rpb1 C-terminal domain (CTD; Figure 1A and Supplemental Figure S1A). These repeats provide a platform for recruiting a variety of factors involved in co-transcriptional processes to the elongation complex (EC;

*To whom correspondence should be addressed. Tel: +1 801 581 5435; Fax: +1 801 581 7959; Email: Tim@biochem.utah.edu

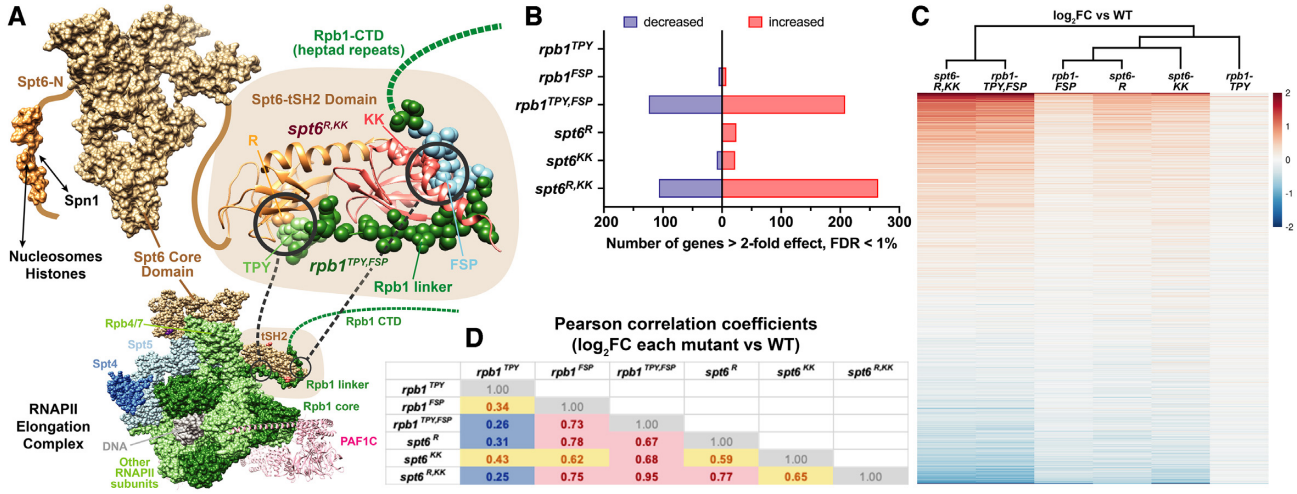


Figure 1. Context of the Spt6-tSH2:Rpb1 interaction and features of the RNA-seq screen. (**A**) Structures of Spt6 domains. A portion of the N-terminal domain as it appears when bound to Spn1 (3; 3PSI), the core domain (2; 3OAK), and the tSH2 domain bound to the Rpb1 linker (8; 5VKO) are shown connected by flexible linkers. The Rpb1 linker residues and the sites mutated in this report are shown as spheres in the tSH2:Rpb1 structure. The proximal portion of the unstructured Rpb1-CTD heptad repeats is indicated as a dotted green line. The bottom panel shows the larger context for these factors within the human RNAPII elongation complex determined by cryo-EM (13; 6GMH). (**B**) Altered transcript levels in interface mutants. The number of genes (out of 5863 annotated loci with sufficient reads to support DESeq2 analysis) whose transcript level in the RNA-seq analysis increased or decreased in the mutant strains relative to the WT strains greater than 2-fold at a false discovery rate (pAdj) less than 1% is shown for 6 genotypes (Supplementary Table S1). (**C**) Heat map comparing the $\log_2$FC values for individual genes for each genotype, sorted by the average effect and clustered by similarity of the patterns. (**D**) Pearson correlation coefficients (*r*) for all pairwise combinations of $\log_2$FC values relative to WT, with color coding indicating weak (0–0.33, blue), moderate (0.34–0.66, yellow), or strong correlations (0.67–1.0, red). Individual components of the interface had distinct functions, but were not strictly modular. For example, the effects of *rpb1*^{TPY} correlated better with loss of the C-terminal binding pocket in *spt6*^{KK} than with loss of the N-terminal SH2 domain that binds the Rpb1-TPY region.

9–11). Spt6 therefore has at least three functional domains, with the NTD interacting with chromatin proteins and at least one essential component of the transcription machinery, the conserved core promoting elongation of transcription through direct contact with components of the EC and with the emerging transcripts themselves, and the C-terminal tSH2 domain that interacts with the linker region of Rpb1 between the catalytic core and a reservoir of factors that have roles during specific phases of elongation.

The more N-terminal SH2 module recognizes a combination of phospho-threonine and unmodified tyrosine that together mimic the structure of a phospho-tyrosine (8). The C-terminal module diverges from the expected SH2 activity even further, using a non-canonical site to bind a series of residues including a phospho-serine (Figure 1A). The Bur1 kinase is responsible for phosphorylating both Rpb1 sites (12), resulting in an Spt6-tSH2:Rpb1 interaction with low nanomolar affinity (8).

A cryo-EM reconstruction of the human RNAPII EC (13) revealed multiple contacts between Spt6 and other components (Supplementary Figure S1A), including the Spt6-tSH2:Rpb1 linker interface as well as interactions with the Rpb4/Rpb7 knob, Spt5, and the nascent RNA transcript. Bur1 modifies other components of the transcription machinery during the transition from the pre-initiation complex (PIC) to the EC, including Spt5 and the Rpb1 CTD (9–11). The Spt6-tSH2:Rpb1 interaction is therefore likely to be activated by Bur1 early in transcription, possibly during the PIC to EC transition, contributing an additional contact between Spt6 and RNAPII in addition to those made by the Spt6 core domain. Consistent with this, mutations that disrupt the Spt6-tSH2:Rpb1 linker interface reduce Spt6 occupancy over transcribed regions globally, but do not eliminate the association (8). It therefore remains unclear what the role of this additional, switchable interaction between Spt6 and Rpb1 might be.

To investigate the functions of this interaction, we used genomic methods to identify genes whose transcript level, Spt6 occupancy, or chromatin architecture relied most heavily on the integrity of the Spt6-tSH2:Rpb1 interface. In addition to global changes, we identified several subsets of genes where disruption of this interaction caused more dramatic changes, including loci regulated by differential start site selection and genes encoding ribosomal proteins. Genes with increased transcript levels tended to have a characteristic accumulation of Spt6 over their 5′ ends that was diminished in the mutants, suggesting linkage between the Spt6-tSH2:Rpb1 interaction and local decisions being made during early stages of transcription. Interface mutations also caused diminished efficiency of splicing globally, and appeared to lead to delayed progression of elongating RNAPII through regions where introns are processed to release mature snoRNAs. Some of the effects were similar to those seen in strains with defects in the Spt6:Spn1 interaction or in histone chaperone activity, but others, including a role in maintaining the +1 nucleosome at RP genes, appeared to be unrelated to general chaperone functions.

We propose that this broad range of effects points to a role for the Spt6-tSH2:Rpb1 interaction in coordinating the responses of RNAPII with the distinct profile of challenges posed by each individual gene. As circumstances call for different factors to promote elongation, termination, matura-

tion of the transcript, or manipulation of nucleosomes, it seems likely that the rate of progression must be modulated to allow time for the appropriate machinery to be engaged. The mechanisms that mediate this cooperation remain unclear, but the available structures show that the Spt6 tSH2 domain is in position to coordinate the functions of its own chaperone activity, its elongation function, and the availability of factors associated with the Rpb1 CTD. We propose here that the tSH2:Rpb1 interface contributes to this process.

MATERIALS AND METHODS

RNA-seq

Three independent cultures of each genotype (Supplementary Table S1) were grown to an OD of about 0.8 in rich medium, then RNA was extracted with phenol and pelleted in 5.7 M CsCl as described (14). Samples were purified on RNeasy columns (Qiagen) following the manufacturer's instructions. Ribosomal and mitochondrial sequences were depleted using Ribo-Zero Gold kits (Illumina), and strand-specific cDNA libraries were prepared using TruSeq Stranded Total RNA Library Prep Gold kits (Illumina) with index tags. Samples were pooled and sequenced on an Illumina NovaSeq 6000 as 50 bp paired-end runs. Fastq files from separate flow cells were concatenated prior to removing sequencing adapters with Cutadapt (version 1.16), allowing a minimum overlap length of 6 (option -O) and a minimum output length of 20 (option -m). Trimmed sequences were aligned by STAR (version 2.6.1b) using the basic two-pass method against the *Saccharomyces cerevisiae* genome (version R64-1-1, ENSEMBL release 94), allowing for a maximum intron length of 1100 bp and a maximum mate gap of 2000 bp. Picard CollectRnaSeqMetrics (version 2.9.0), Samtools (version 1.5), and Fastqc (version 0.11.5) were used to judge the quality of the libraries. Alignments were counted over the middle 70% of transcripts with the application BioToolBox get_datasets (version 1.68), using the options –feature transcript –fstart 0.15 –fstop 0.85 –method ncount –strand antisense. This reduced skewing of the results by variable transcription start site usage in the mutants studied here, as well as high levels of unstable transcripts produced at some promoters. The number of genes affected by these issues was small but this class of genes was a significant focus of this study. Restricting analysis to the middle of each gene had little global effect on the results; comparing the whole gene and middle 70% data for the $\log_2$(fold change) values ($\log_2$FC) for the $spt6^{R,KK}$ and $rpb1^{TPY,FSP}$ strains gave Pearson correlation coefficients and slopes of $\geq$0.99 (not shown). Custom annotation was used that included UTRs and excluded dubious annotations (https://github.com/tjparnell/biotoolbox-nucleosome/blob/master/yeast_positioned_nucleosomes/SacCer3_R64_all_genes_NoDubious_UTR_chromo.gff3.gz). Differential analysis was performed with DESeq2 (version 1.22.2) with the aid of the hciR (https://github.com/HuntsmanCancerInstitute/hciR) package. All samples were loaded at the same time, using the genotype as the input factor. PCA plots and pairwise correlations on normalized count data were used to judge quality (Supplementary Figure S1C, D). Individual pairwise contrasts were performed against the WT sample. A heat map and hierarchical cluster dendrogram (Figure 1C) was generated with the pHeatmap R package using the DESeq2-generated $\log_2$FC values. Depth-normalized, replicate-mean, coverage tracks for visualization were generated using BioToolBox bam2wig application (version 1.68, https://github.com/tjparnell/biotoolbox) with options –smartcov –rpm –bw –mean –flip –nosecondary. An exclusion list was provided for excluding high copy number intervals (rDNA, telomeres). Figures were rendered in GraphPad Prism and assembled in Adobe Illustrator.

MNase-seq

Three replicate cultures (Supplementary Table S1) were grown in rich medium to an OD of about 0.8 and processed as described previously (15). Briefly, cultures were treated with 1% formaldehyde for 20 min, cells were spheroplasted with lyticase (Sigma-Aldrich), lysed with 1% Triton X-100 and, and the extracts were treated with MNase (Worthington). DNA was purified, libraries were constructed using NEBNext ChIP-seq kits, and fragments of the appropriate size were purified by electrophoresis through agarose and extracted (MinElute, Qiagen). Libraries were sequenced using either single-end 50 bp reads (Illumina HiSeq 2500) or paired-end 50 bp reads (Illumina NovaSeq 6000).

Reads were aligned to the genome using Novoalign (version 3.07.01), allowing for one random multi-mapping alignment. Samples had about 50–60% duplication rates. To reduce and normalize duplicate reads while maintaining biological enrichment, duplicate reads were randomly subsampled to a uniform rate of 40% using the bam_partial_dedup application (version 1.7, https://github.com/tjparnell/HCI-Scripts/blob/master/BamFile/bam_partial_dedup.pl). Normalized fragment coverage tracks were generated using BioToolBox bam2wig (v1.52), excluding intervals with high copy number such as telomeric and rDNA loci, extending the read to a fragment length of 150 bp, depth-normalizing to fragments per million, and averaging biological replicates. Delta nucleosomal coverage was generated by taking bedGraph file format versions of the coverage tracks and using the Macs2 bdgcmp function to subtract the WT coverage from mutant coverage (16, https://github.com/taoliu/MACS/). Resulting bedGraph files were converted back to bigWig for visualization and analysis.

To generate mean nucleosomal occupancy data over annotation (Figure 8E, Supplementary Figure S8A-D), the BioToolBox application get_datasets was used with the appropriate annotation and the replicate-averaged nucleosomal fragment coverage described above. The NDR and +1 nucleosome annotation was obtained from (15). To generate occupancy data relative to the TSS (Figure 8B–D, F), the BioToolBox application get_relative_data was used with windows of 10 bp and custom transcript annotation used in the RNASeq analysis above. Windows overlapping upstream transcripts were discarded with the –avoid option. To generate occupancy data over genes (Figure 8A, Supplementary Figure S8E, F), the BioToolBox application

get_binned_data was used with the custom transcript annotation used in the RNASeq analysis above. Mean occupancy was determined in intervals of 5% of the gene length, with ten additional 25 bp intervals collected upstream and downstream of the gene.

To determine nucleosomal shifts in the mutants (Supplementary Figure S8I), nucleosomal positions were first mapped in the WT strain using the package https://github.com/tjparnell/biotoolbox-nucleosome. As input data, 'skinny' nucleosome fragment coverage was generated with BioToolBox bam2wig by shifting the alignment start positions downstream by 37 bp and extending by 75 bp; this essentially generates coverage of the central 50% of the nucleosomal fragment and generally improves the reliability of nucleosomal calls over previous analyses performed with strict midpoint data by averaging nucleosomal density over the highest-confidence central portion of nucleosomes. Coverage was depth-normalized and averaged among replicates. Nucleosomes were called with map_nucleosomes.pl using a threshold of 2. Nucleosome mappings were re-verified with verify_nucleosome_mapping.pl, filtering overlapping nucleosomes with a maximum overlap of 35 bp and re-centering mapped nucleosomes as necessary (the majority of overlapping nucleosomes were in ambiguous areas with poor phasing, usually in the middle of long genes). This identified 65,082 nucleosomes in the WT genome. Nucleosomes were assigned to relative genic positions (−1, +1, +2, +3 and +4) with the script get_positioned_nucs.sh in the yeast_positioned_nucleosomes subfolder in the biotoolbox-nucleosome package. Nucleosomal shift correlations and optimal shifts were calculated with the BioToolBox application correlate_position_data.pl (v1.63) using the mapped WT nucleosomes with a radius of 50 bp from the midpoint and mutant 'skinny' nucleosome fragment coverage generated as described above. Specifically, for each queried mapped nucleosome position, the mutant skinny nucleosome coverage was incrementally shifted relative to WT skinny nucleosome coverage, and the optimal shift yielding the highest correlation was reported.

Quantitative PCR

Independent cultures of strains with the genotypes listed (Supplementary Table S1) were grown as indicated in each experiment and RNA was extracted with phenol and pelleted in 5.7 M CsCl (14). cDNA libraries were prepared with M-MLV reverse transcriptase (Promega) using the manufacturer's instructions. PCRs were performed with Taq DNA polymerase (Apex) using a Roche LC480 with SYBR Green I (Molecular Probes) and ROX (5(6)-carboxy-X-rhodamine) (Sigma-Aldrich) for detection of products. Standard curves for each primer set were determined using genomic DNA from a WT yeast strain and used to determine the concentrations of cDNAs relative to single copies of the target sequences per cell. Variance of each cDNA sample relative to other samples with the same genotype was averaged over all PCR primer sets to provide normalization values, as no suitable control gene unaffected by the mutations tested was identified.

Primer extension assay

30 μg of total RNA extracted as for qPCR assays was mixed with 300 fmol of primers (Supplementary Table S2), 100 U of reverse transcriptase (Promega), dNTPs (100 μM each) and processed essentially as described (17,18), then incubated at 37° for 30 min. After precipitation with ethanol, samples were electrophoresed through a 15% polyacrylamide gel in 1× TBE and 7 M urea. The Cy5 label on the primers was detected using a ChemiDoc scanner (BioRad) and quantitated using Image Lab 6.1 (BioRad).

RESULTS

Disrupting the Spt6-tSH2:Rpb1 interface altered the relative levels of specific transcripts

The Spt6-tSH2:Rpb1 interface has two main components, corresponding to each of the SH2 motifs of Spt6 and the two Rpb1 linker sites that they bind (Figure 1A, circles). The individual components of this interface contribute independently to the overall affinity of the interaction *in vitro* and to the functions of the interface *in vivo* (8, and Supplementary Figure S1B), but we focused here mainly on compound mutations that effectively disable the interaction completely (a detailed description of the individual mutations and the shorthand nomenclature for describing them are found in Supplementary Table S1; briefly, alleles are designated by the normal residues prior to mutation, with a comma separating the N- and C-terminal components of compound mutations, as in $spt6^{R,KK}$ to represent *spt6-R1282H,K1355A,K1435A*). Mutations were introduced into the native genomic loci (Supplementary Table S1).

To identify genes disproportionately affected by disruption of the Spt6-tSH2:Rpb1 interaction, we used RNA-seq with strand-specific cDNA libraries produced from three independent biological replicates of each genotype to screen for changes in the steady-state levels of the sense strands for all annotated genes. The effects of mutations on individual genes correlated well with one another in most cases (Figures 1C, D and Supplementary Figure S1C, D) and moderately with published values for a complete deletion of the Spt6-tSH2 domain (19, Supplementary Figure S1E). While interface mutations caused significant growth defects, deleting the tSH2 domain entirely had a larger effect (Supplementary Figure S1F), suggesting that it has functions in addition to interacting with Rpb1. Inactivating the binding pockets in Spt6-tSH2 or removing their recognition sites in Rpb1 caused similar changes in transcript levels from the same genes, supporting the interpretation that the changes were caused by loss of the Spt6-tSH2:Rpb1 interface.

This RNA-seq method would not detect uniform global changes in transcript levels per cell, but the relative transcript levels correlated strongly with WT (Supplementary Figure S1D, Pearson $r \geq 0.99$ in all cases). We anticipated that reduced association of an elongation factor with the EC would lead to a relative decrease in transcripts from some genes, and this was the case for about 100 genes in the compound mutants (Figure 1B, Supplementary Figure S1G). Scatter plots (Supplementary Figure S1H) and Gene Ontology term analysis indicated that genes with reduced tran-

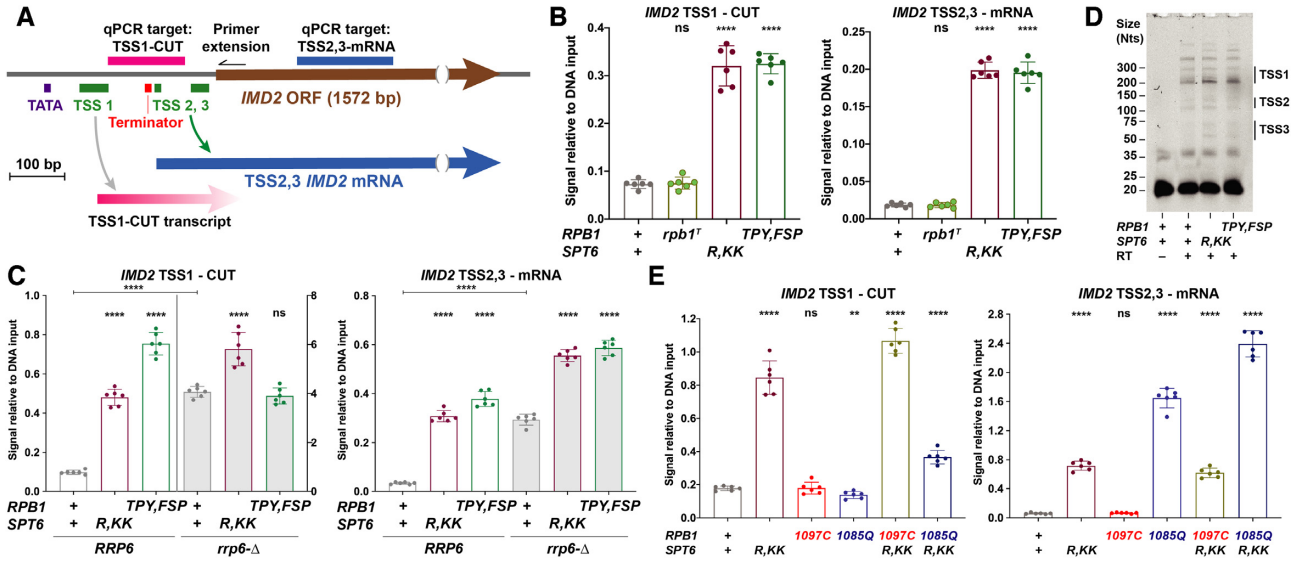


Figure 2. Interface mutations alter the profile of initiation site usage at *IMD2*. (**A**) Scale map of the *IMD2* locus. Locations of the TATA, 3 TSS clusters (as in Supplementary Figure S3A), Terminator, mRNA, and ORF are as described (22,23). PCR targets are the regions amplified by the primers listed in Supplementary Table S2. (**B**, **C**) qPCR values for 6 independent cultures of strains with the genotypes shown (Supplementary Table S1). Error bars indicate the mean and standard deviation with the results of unpaired t-tests comparing mutants to WT shown above the bars and a bar comparing *RRP6* to *rrp6-Δ* (throughout the manuscript, $P > 0.05$ is ns, $P < 0.05$ is *, $P < 0.01$ is **, $P < 0.001$ is ***, and $P < 0.0001$ is ****). Strains with and without *RRP6* in panel C are shown at different scales for TSS1-CUT. (**D**) Primer extension was performed with RNA extracted from strains with the genotypes indicated (Supplementary Table S1). Reverse transcriptase (RT) was omitted from the WT sample in the first lane. Multiple biological repeats gave similar patterns (not shown). Molecular weight standards (Fisher exACTGene) were used to determine the approximate regions for the expected products from the *IMD2* TSSs, as indicated. This region of *IMD2* is identical to the ORFs of *IMD1* and *IMD3*, so the more abundant transcripts from those loci are also detected. Only the patterns of products relative to WT can therefore be compared. (**E**) As in 2B,C but with strains carrying Rpb1 trigger-loop mutations alone or in combination with *spt6*^{R,KK} (Supplementary Table S1).

script levels tended to be highly transcribed genes, including a significant over-representation of ribosomal protein (RP) genes. Notably, more genes displayed significantly increased transcript levels (over 200 for the compound mutants, Figure 1B, Supplementary Figure S1G) and these were not associated with specific GO terms. Transcription stress can cause increased turnover of transcripts from genes associated with translation (20,21), so the decrease in RP gene transcript levels could reflect reduced ability to maintain high levels of transcription or decreased stability of these specific mRNAs. In contrast, the genes with increased levels of transcripts did not align with stress-induced genes (Supplementary Figure S1H) or other general co-regulated categories, so we examined a set of these genes individually for features that might reveal common mechanistic themes.

Effects at *IMD2* suggest a role in transcription start site selection

Transcripts from *IMD2* consistently displayed the largest increases observed in the RNA-seq data, reaching about 10-fold above WT in most interface mutants (Supplementary Figure S1D). *IMD2* is one of three genes that encode inosine monophosphate dehydrogenases (IMDHs) in *Saccharomyces cerevisiae*, all of which act at a branch-point in purine nucleotide synthesis where guanosine and adenosine production from a common intermediate are balanced with one another through competition with the activity of the product of the *ADE12* gene (Supplementary Figure S2A). *IMD2* normally contributes only about 4% of the IMDH pool, but expression of this isoform is regulated through the differential use of multiple potential transcription start sites (TSSs, Supplementary Figure 2A), allowing upregulation in response to low GTP levels (22,23). In this model, all preinitiation complexes (PICs) form at the single TATA site in the *IMD2* promoter, then scan downstream for potential initiation sites. If the first cluster of available sites is used (TSS1; Figure 2A, Supplementary Figure S3A), a terminator sequence is encountered, causing production of a short, rapidly-degraded cryptic unstable transcript (Figure 2A; TSS1-CUT). If TSS1 is not used, the PIC can continue scanning, initiate at any of several sites in the TSS2,3 cluster, and synthesize a productive mRNA. *IMD2* expression therefore depends on competition between the rate of scanning and the probability of using TSS1 for initiation, with the latter being sensitive to GTP concentration (24,25). The increase in *IMD2* transcripts in Spt6-tSH2:Rpb1 interface mutants therefore suggests that this interaction may be involved in PIC scanning or the efficiency of initiation.

To test this more quantitatively, we assessed the level of transcripts initiated at the TSS clusters upstream and downstream of the terminator using quantitative PCR with additional biological replicates (Figure 2B). We observed the expected ∼10-fold increase in *IMD2* mRNA, but also about a 4-fold increase in the TSS1-CUT product. In the model proposed above, a constant pool of PICs is formed and initiation events are distributed among the available TSSs, so increased use of TSS2,3 sites should be accompanied by decreased TSS1 usage, making an increase in both products unexpected. We considered several explanations

for this outcome. First, increased promoter activity (more frequent PIC assembly) could drive an increase in products from all potential initiation sites. Second, the efficiency of termination could be decreased in the mutants, leading to production of full-length mRNAs initiated from TSS1. Third, the interface mutants could cause a defect in degrading the TSS1-CUT product. The rapid turnover of TSS1-CUT in normal cells obscures the amount produced, so we tested these models using an isogenic set of strains with and without deletions of *RRP6*, a component of the nuclear exosome responsible for degrading CUTs (26–29).

The level of TSS1-CUT signal increased ∼50-fold in an *rrp6-Δ* strain (Figure 2C), and interface mutants had smaller effects than in the *RRP6* background (no effect with $rpb1^{TPY,FSP}$, ∼40% higher with $spt6^{R,KK}$). The 4-fold increase in this transcript observed when the exosome was functional therefore does not seem to be the result of increased *IMD2* promoter activity in interface mutants, making model one proposed above unlikely.

Simple failure of termination of transcripts initiated at TSS1 would lead to synthesis of mRNAs with extended 5′ ends. We attempted to detect this by primer extension, but the site needed to perform this test is shared by the rest of the IMDH family (*IMD1*, *IMD3* and *IMD4*). We did not observe an altered pattern of primer extension products in interface mutants (Figure 2D), but given the lower level of *IMD2* mRNA expected relative to products from *IMD3* and *IMD4* this is not conclusive. A similar test of products from *ADH1* showed that the ratio of usage of two TSSs that are ∼10 bp apart and not separated by a terminator was essentially unchanged in interface mutants relative to WT (Supplementary Figure S2B). To ask more directly about termination efficiency we examined transcripts from *NRD1*. This gene encodes a component of the Nrd1–Nab3–Sen1 (NNS) termination complex whose expression is regulated by having the availability of NNS complexes determine whether initiated transcripts are completed (27,30–32). Interface mutants caused a 2–3-fold increase in both 5′ end products (initiated transcripts) and 3′ end products (completed mRNAs), with only a slight or no decrease in termination efficiency (a small increase in the completion ratio, Supplementary Figure S2C, D). Exosome mutations cause a large increase in the level of *NRD1* transcripts initiated, but a smaller increase in completion due to increased termination efficiency (33–35), but this increase was also observed in interface mutants (Supplementary Figure S2E). As a further global test of initiation and termination, we aligned the RNA-seq reads by either the annotated TSS or TTS (termination site) for all genes and observed no general changes in either profile in interface mutants (Supplementary Figure S2F). These results do not rule out defects in termination at specific loci, but they do not support this as a mechanism explaining the observed 4-fold increase in TSS1-CUT signal from *IMD2*.

In addition to the effects on TSS1-CUT, deletion of *RRP6* also resulted in a 9-fold increase in *IMD2* mRNA (Figure 2C). Interface mutants had a smaller effect in this background, causing about a further 2-fold increase. It is not clear why loss of the exosome caused this dramatic increase in *IMD2* mRNA, but overall, the results suggest that stabilization of a CUT whose transcription impinges on downstream TSSs can affect the use of those sites, that interface mutations may cause a defect in degrading products initiated at *IMD2* TSS1, and that this stabilization itself could at least partially explain the increase in mRNA from this gene.

To investigate the relationship between stabilization of CUTs produced by upstream TSSs and mRNA levels further, we examined the paralog *IMD3*. In WT cells, this gene produced about 35-fold more mRNA than *IMD2* in our hands, and while the ORFs are highly similar, the promoters are not (Supplementary Figure S2G, H). In particular, *IMD3* has a single TSS cluster and is constitutively expressed. However, we found that it also has an unannotated upstream CUT produced from an unmapped TSS, as qPCR signal using primers upstream of the annotated TSS increased 130-fold in an *rrp6-Δ* mutant (Supplementary Figure S2H). Interface mutations also stabilized this CUT 5–10-fold in *RRP6* cells but not in *rrp6-Δ* cells, and caused less than 2-fold increases in *IMD3* mRNA in cells with exosomes but no effect in cells lacking them (Supplementary Figure S2H). As at *IMD2*, interface mutants therefore appeared to have a defect in exosome-dependent degradation of a CUT that impinges on a downstream TSS at *IMD3*. In the case of *IMD2*, efficient degradation of the CUT was important for maintaining low levels of the mRNA, presumably by reducing initiation from the TSS2,3 cluster, but the effect on mRNA was smaller at *IMD3* where the choice of initiation site is not known to be involved in regulating expression.

To ask if CUT signals were globally stabilized in interface mutants, we examined the RNA-seq dataset, but were unable to obtain reliable results due to the low number of reads in the exosome-proficient cells used. We therefore tested *CUT320* as it is over 500 bp upstream of *NRD1* and transcribed in the opposite direction (Supplementary Figure S2C). It showed less than a 2-fold increase in interface mutants (compared to a 43-fold increase resulting from *rrp6-Δ*; Supplementary Figure S2E), so loss of the interaction does not appear to cause global stabilization of CUTs to the levels seen at *IMD2* and *IMD3* (Figure 2, Supplementary Figure S2H). However, *CUT320* is about 150 bp upstream of, and transcribed in the same direction as *RAD50*, and may impinge on its TSS, so the density of the yeast genome makes it difficult to fully assess the potential role of the Spt6-tSH2:Rpb1 interface in degradation of CUTs and effects on TSS usage. The interface therefore appears to contribute to exosome-dependent degradation of a subset of CUTs produced in promoter regions, and coordination between these pathways somehow affects the use of downstream TSSs.

The defects in interface mutants appear to be different from other transcription stresses

To further probe these mechanisms, we examined the effects of combining interface mutations with other factors known to alter *IMD2* expression. The trigger loop of Rpb1 participates in the polymerization reaction (18,36–38) and mutations can either increase the rate of catalysis ('fast' mutants, represented here by *rpb1-G1097C*) or decrease this rate ('slow' mutants, including *rpb1-H1085Q*). Fast mutants

accept potential initiation sites more readily than normal, and slow mutants less readily, leading to upstream and downstream shifts in the profiles of TSS usage, respectively (18,25,37). At *IMD2*, fast mutations have little effect on the mRNA level, while slow mutations increase this level by shifting the profile of initiation site usage downstream (18,25,37). Consistent with a simple model in which these mutations alter the distribution of TSS selection from a constant pool of PICs formed, a fast mutant had no effect on the level of either *IMD2* TSS1-CUT or mRNA and a slow mutant caused a small decrease in TSS1-CUT along with an increase in mRNA (Figure 2E). The effect of combining *rpb1-H1085Q* with *spt6*$^{R.KK}$ was intermediate between the single mutants for TSS1-CUT, but was roughly additive for mRNA (Figure 2E; 12-fold and 27-fold increases in mRNA for the single mutants, 41-fold for the double mutant). This could indicate that pausing caused by the slow polymerase provided additional time for coordinating with the exosome to degrade the CUT (suppression of this defect caused by interface mutations), and that the individual disturbances in *IMD2* mRNA production acted on independent mechanisms (mutual enhancement of this defect). These results support a role for the Spt6-tSH2:Rpb1 interface in selecting an initiation site at *IMD2*, but through a mechanism different from the role of the trigger loop in promoting polymerization.

Mycophenolic acid (MPA) and 6-azauracil (6AU) increase the expression of *IMD2* mRNA by reducing the pools of GTP and UTP, respectively, which causes reduced efficiency of initiation at TSS1 (22,23). MPA and 6AU each reduced the level of TSS1-CUT product and increased the amount of mRNA in both WT strains and in an interface mutant (Supplementary Figure S2I-K). This again suggests that the apparent defect in TSS1-CUT degradation in interface mutants could be partially suppressed by delayed RNAPII progression, and that the increases in mRNA production were due to different mechanisms. Overall, the Spt6-tSH2:Rpb1 interface mutations caused an increase in *IMD2* mRNA transcripts, possibly by reducing the efficiency of initiation at TSS1 through a mechanism different from catalysis of NTP incorporaton, and they also interfered with an unknown mechanism linking exosome-mediated degradation of CUTs and initiation site selection at a subset of sites.

Interface mutations affected initiation-site usage patterns at multiple genes

Interface mutations affected the profile of TSS usage at *IMD2*, but not globally or at *ADH1* (Supplementary Figure S2B,F), suggesting differential importance of the Spt6-tSH2:Rpb1 interaction at different genes. To examine this, we chose a set of genes with multiple potential initiation sites for further examination. The level of *BIO2* mRNA increased about 6-fold in our RNA-seq screen, and a published TSS-seq dataset indicated that this gene has two clusters of potential initiation sites (39; Figure 3A, Supplementary Figure S3A). Products initiated at TSS1 increased 2–3-fold in interface mutants, but the level of mRNA increased 6-fold, so the calculated ratio of TSS1/mRNA decreased, suggesting a shift towards use of TSS2 (Figure 3A, Supplementary Figure S3B). The differential use of initiation sites at this locus has not been reported to regulate the expression of this gene, but *CUT636* impinges on the promoter, so to compare the results with *IMD2* we asked if the TSS1 and mRNA signals were affected by loss of the nuclear exosome. TSS1 signal increased about 6-fold in an *rrp6-Δ* strain and the mRNA level increased 4-fold, producing a rise in the TSS1/mRNA ratio from near 1 to about 1.5 (Supplementary Figure S3F). The signal from the TSS1 target region could therefore be from *CUT636* or from unstable products initiated at TSS1 (Figure 3A). In either case, as at *IMD2*, interface mutations caused at least a partially exosome-dependent increase in signal from the TSS1 region, loss of the exosome caused an increase in mRNA by itself but interface mutations still caused a further increase, the effects of exosome and interface mutations on mRNA levels were similar to one another and not additive, and initiation shifted from primary use of TSS1 towards downstream sites (Figure 3A, Supplementary Figure S3B,F). *BIO2* and *IMD2* therefore displayed several common responses to exosome deletion and interface mutations.

Other genes encoding biosynthetic enzymes, including *ADE12* (whose product competes with the product of *IMD2* for substrate, Supplementary Figure S2A), *MET3*, and *URA8*, were also noted to have multiple potential initiation sites (Supplementary Figure S3A) with altered patterns of usage in interface mutants (Figure 3B-D, Supplementary Figure S3C–E, G–I). Like *IMD2* and *BIO2*, *ADE12* and *MET3* displayed decreased ratios of TSS1 signal to mRNA in interface mutants, indicating a shift towards use of downstream initiation sites. In the case of *ADE12*, this was due to reduced TSS1 signal without a change in mRNA, while *MET3* had a small increase in signal from the TSS1 target but a large increase in mRNA. In contrast, both TSS1 and mRNA signals from *URA8* increased, producing a slight apparent shift towards use of the minor upstream TSS1 site. The response of transcripts from the TSS1 regions of these genes to loss of nuclear exosomes also varied significantly, with *ADE12* showing little effect, *MET3* displaying a 6-fold increase from TSS1 with decreased mRNA, and *URA8* yielding large increases in both TSS1 and mRNA signals while maintaining an overall increase in the TSS1/mRNA ratio in interface mutants (Supplementary Figure S3G-I). Genes with multiple TSS options therefore used the available sites in different ways, underscoring the range of distinct situations encountered by RNAPII during transcription and the need to respond appropriately to the individual features of each gene. The Spt6-tSH2:Rpb1 interface affected each of the genes tested differently, suggesting roles for this interaction in a range of situations as RNAPII engages in dynamic responses to specific circumstances.

Binding of Spt6-tSH2 to Rpb1 might alter the conformations of one or both of these proteins, thereby directly altering their activities, or it could passively tether Spt6 to the transcription complex to indirectly enhance the local availability of the other functional domains of Spt6 (8). To test the contribution of tethering, we sought alleles of *SPT6* whose products could still be recruited through the tSH2:Rpb1 interaction but would fail to provide specific functions and might therefore phenocopy reduced localization of Spt6 to transcription complexes. However, genetic

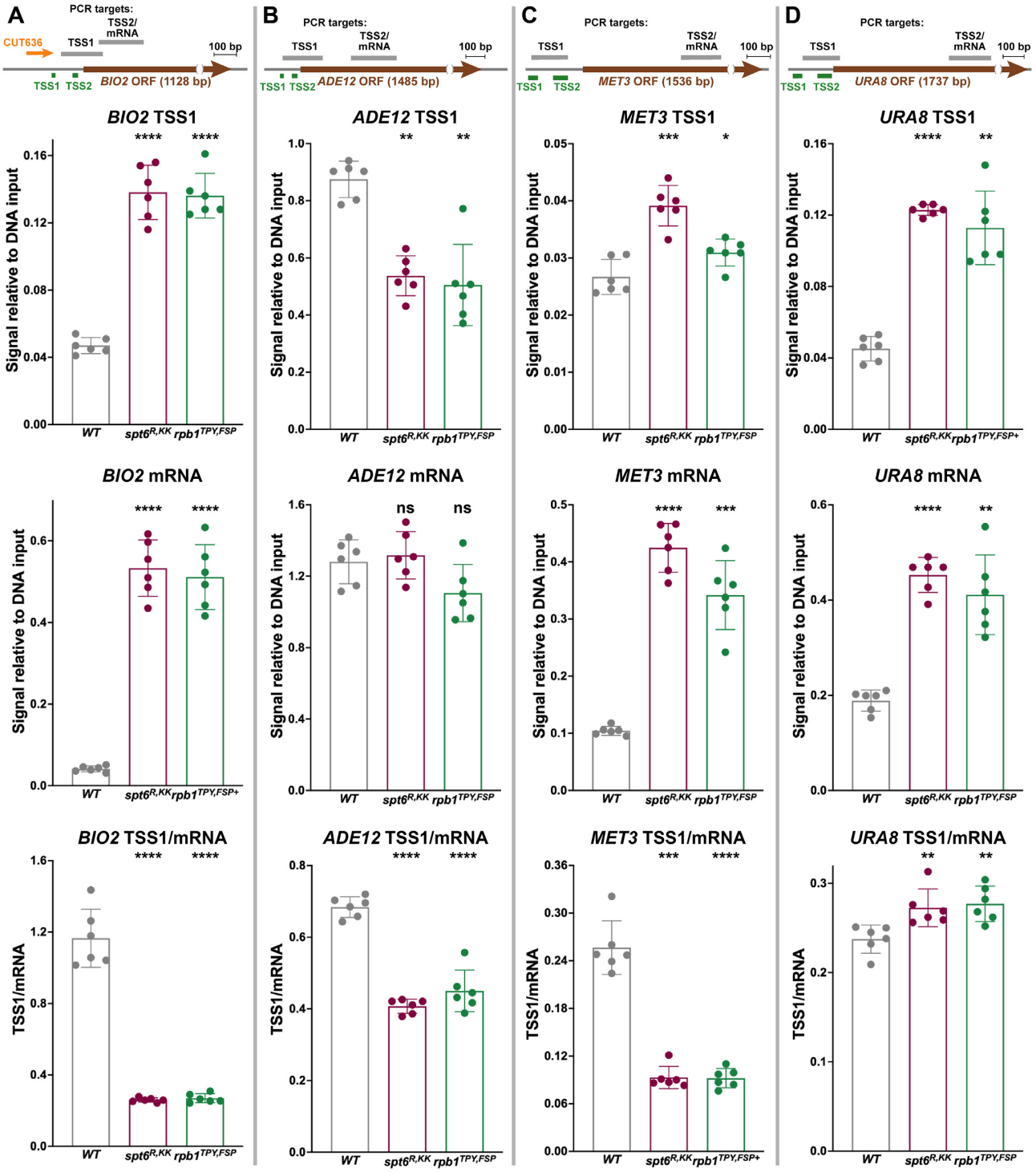


Figure 3. Interface mutations alter TSS usage profiles at additional genes. (**A–D**) The same cDNAs used in Figure 2B were tested at four additional loci that had multiple TSS clusters in published TSS-seq data (39), as diagrammed in the top panels and shown in Supplementary Figure S3A. PCR targets that amplify products initiated at the upstream site but not the downstream site were used and the ratio of TSS1 usage to total mRNAs produced from all initiation sites was calculated in the bottom panels. Statistical tests are as in Figure 2.

tools for disrupting only the individual elongation functions of the core domain or the histone chaperone activity of the NTD are not currently available as residues supporting these activities have not been precisely mapped or disrupted. The *spt6*^{F249K} mutation disturbs the interaction of the Spt6 NTD with the multi-functional factor Spn1, reducing the ability of Spn1 to compete with Spt6 for binding to nucleosomes (3; see Figure 1A). This allele does not appear to directly impact histone or nucleosome binding, but it impairs recruitment of Spn1 to transcription complexes through Spt6. We therefore compared its effects with those of interface mutants as a partial probe of how reduced Spt6 recruitment alone might affect transcript levels. The functions of Spn1 and its interaction with Spt6 display complex genetic interactions with the histone chaperone FACT (40), so to attempt to determine the contribution of a generic reduction in global chaperone activity to the phenotypes of interface mutants we also tested the effects of two FACT alleles that cause distinct but overlapping defects in chromatin architecture (*pob3*^{Q308K} and *spt16*$^{-11}$; 15,41).

All of the tested mutations caused apparent shifts toward TSS2 usage at *BIO2*, but *spt6*^{R,KK} had the largest effect (Supplementary Figure S3B, F). *spt6*^{R,KK} also increased the *MET3* mRNA level the most, with only small effects of any of the mutants on TSS1 signal (Supplementary Figure S3D). Both *spt6* alleles caused a reduction in the TSS1/mRNA ratio at *ADE12*, with smaller effects of FACT mutations (Supplementary Figure S3C), while only the interface mutant caused a shift to apparent upstream TSS1 usage at *URA8* (Supplementary Figure S3E). This analysis again revealed the distinct nature of regulation for each of these genes, with overlap among the activities disturbed by this set of mutations but differential profiles of contributions at each locus. The effect of the Spt6-tSH2:Rpb1 interface at *URA8* was unlike the other factors tested here and may reveal a function that is distinct from recruitment of Spn1 or histone chaperone activities.

The *spt6*^{R,KK} mutation also caused the largest apparent stabilization of the *IMD2* and *IMD3* CUTs, even though all of the tested mutants displayed increases in mRNA similar to or greater than those caused by the interface mutations (Supplementary Figure S3J). The potential defect in coordinating CUT degradation with initiation site usage therefore appeared to track mainly with loss of the Spt6-tSH2:Rpb1 interaction.

The Spt6-tSH2:Rpb1 interaction contributed to co-transcriptional processing of products

The association of reduced transcript levels with ribosomal proteins (RPs) in interface mutants could be the result of the active degradation of these mRNAs in response to transcription stress (20,21). Alternatively, this class of genes is also enriched for introns in *S. cerevisiae,* so we asked if interface mutations caused a defect in splicing. The RNA-seq data were consistent with a reduction in splicing efficiency globally (Supplementary Figure S4A–D), but due to the high efficiency of splicing, this analysis relied on small numbers of reads from only three biological replicates. Both statistical methods and direct measurements of intron signals have been used to overcome this limitation (42), and we chose the latter approach to allow additional genotypes and culture conditions to be tested. We selected genes from distinct classes to give broad representation. This included three RP genes with introns (*RPL17A*, *RPL18A* and *RPL23A*) based on their apparent low, moderate, or high intron retention in the RNA-seq data (Supplementary Figure S4A). *TEF4* was also selected as a gene involved in translation but not in ribosome biogenesis, and as the gene with the highest apparent intron retention in the RNA-seq data (Supplementary Figure S4A,B). *IMD4* was added as an intron-containing paralog of the intron-less *IMD2* and *IMD3* genes studied extensively above, *YPR063C* as a well-studied target of splicing efficiency (43), and *URA2* as a gene with an intron in the 5′ UTR instead of in the coding sequence. These targets span a range of functions, intron configurations, apparent effects on intron retention in the RNA-seq data, and changes in mRNA accumulation in interface mutants.

The level of intron signal relative to exons increased significantly for each of the RP genes tested (Figure 4B–E), as well as for all other genes tested (Supplementary Figure S4E–I). The interface mutations therefore caused a more uniform global splicing defect than was suggested by the RNA-seq data. Enhanced mRNA turnover could artificially elevate the apparent retention of intron sequences at steady-state, as mature mRNA would be preferentially degraded, passively increasing the fraction of nascent unprocessed transcripts. However, the level of intron retention did not correlate with the changes in mRNA in our tests, suggesting this is unlikely to be a major contributor to the apparent splicing defect (Figure 4, Supplementary Figure S4; similar increases in intron retention were observed for genes with increased, unaffected, or decreased mRNA levels).

A simple lack of intron removal should produce similar amounts of signal for the 5′ exon-intron junction, the intron itself, and the 3′ intron-exon junction. Tests of all three regions at the *IMD4* locus instead revealed a gradient of signals with 5′ junction signals higher than 3′ junctions (Figure 5A, C, Supplementary Figure S5A). Excess signal from the 5′ junction of *IMD4* was also visible in RNA-seq browser tracks (Supplementary Figures S3A, S5E), indicating this was not an artefact of the PCR assay. We normalized the signals two ways, first to the exon level for the same sample to remove the effects of variation in total transcripts and reveal absolute retention of each region, and second relative to the value for the same region in WT cells to isolate the effects of mutations (Figure 5C). Excess 5′ junction signal could indicate incomplete transcripts, suggesting stalling of polymerase in intron regions, potentially revealing a role for the Spt6-tSH2:Rpb1 interface in coordinating polymerase progression with transcript processing.

The intron of *IMD4* encodes a snoRNA (*SNR54*) that is processed from the primary transcript in parallel with splicing of the intron (Figure 5A; reference 44). *TEF4* also encodes a snoRNA (*SNR38*; Figure 5B) and it produced an even more prominent gradient of intron retention (Figure 5B,D, Supplementary Figures S3A, S5B, S5E). Neither the defect in splicing nor the apparent delay in progression were observed as prominently in FACT mutants, so these features

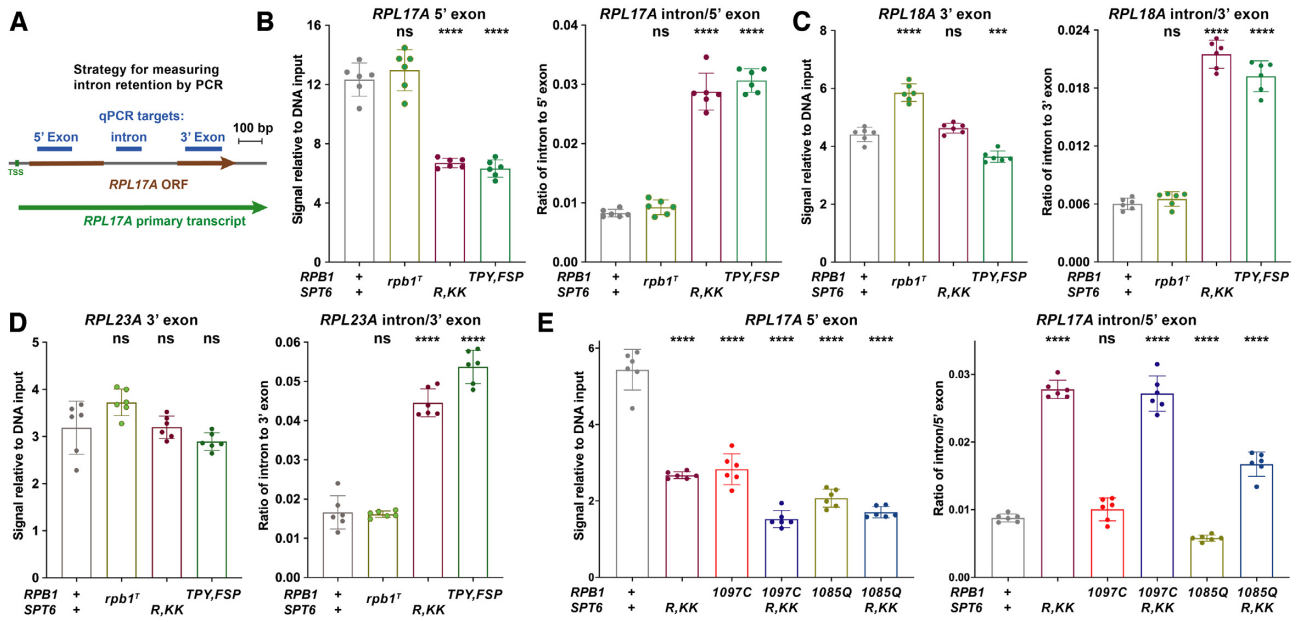


Figure 4. Interface mutations caused increased intron retention at three RP genes. (**A**) The *RPB17A* locus is drawn to scale, illustrating the strategy used for primer placement for measuring intron retention for all genes. The 5′ exon of *RPL18A* is too small to provide a PCR target, so only the 3′ exon was used. (**B–D**) Signal from the intron and exon regions of three genes encoding RPs was measured using the same samples and statistical analysis as in Figure 2 (Supplementary Table S1). The ratio of intron to exon signal was calculated for each sample and in this case paired t-tests were calculated for the comparison with WT. (**E**) As above, but with the same samples used in Figure 2E. A 'fast' Rpb1 mutation caused a reduction in *RPL17A* exon signal similar in magnitude to the $spt6^{R,KK}$ mutation, but did not affect intron retention, suggesting that the apparent splicing defect caused by the interface disruption is not caused by all forms of transcription stress. Similarly, a 'slow' Rpb1 mutation caused decreases in both mRNA level and intron retention.

are unlikely to be generic signatures of histone chaperone defects (Figure 5E, F, Supplementary Figure S5F, G). The $spt6^{F249K}$ mutant displayed elevations of both 5′ and 3′ junction signals relative to the intron itself, a mixture of the patterns observed with interface mutations and $spt16^{-11}$, consistent with potential decreases in both Spn1 recruitment and histone chaperone activity. These results suggest that interface mutations may cause RNAPII to be delayed as it traverses regions where complex co-transcriptional maturation of the product is required, and that other domains of Spt6 may contribute to avoiding these delays through their elongation functions or ability to recruit Spn1.

To test association of this pattern with snoRNAs within introns, we tested four additional genes that do not have this feature (Supplementary Figure S5C–E, H–K). *RPL17A*, *RPL18A* and *RPL23A* also produced differential signals across introns, but in this case with elevated 3′ junction levels, consistent with a delay after the 5′ donor site has been disrupted but before completion of splicing. FACT mutations and $spt6^{F249K}$ caused either uniform retention across the introns or increased 3′ junction signal at these loci (Supplementary Figure S5H-J). The analysis of *URA2* revealed very high levels of 3′ junction signal, but this appeared to be the result of initiation of transcription from an unannotated TSS within the intron, as discussed below (see Supplementary Figures S3A, S5E, H). The apparent delay of polymerase progression through intron regions therefore correlated with a requirement for additional processing of the intron after splicing, and may reveal a role for the Spt6-tSH2:Rpb1 interaction in coordinating progression of transcription with processing of challenging transcripts.

Detecting progression delays by reducing NTP pools

To further test the model that interface mutations cause delayed progression of RNAPII in the introns of *IMD4* and *TEF4*, we asked whether other forms of transcription stress cause similar effects. 6-Azauracil (6AU) causes global pausing of RNAPII progression by reducing the availability of UTP (45). Unlike the strong induction of *IMD2* observed with 6AU treatment (Supplementary Figure S2J), *IMD3* (not shown) and *IMD4* mRNA levels did not increase (Figure 6A, Supplementary Figure S6), consistent with their constitutive expression. Greater increases in 5′ junction signal were observed for *IMD4* and *TEF4*, although the gradient was less pronounced for *TEF4* due to increased retention across the intron region (Figure 6A,B). *TEF4* mRNA dropped ∼3-fold during 6AU treatment, possibly leading to passive increases in intron retention due to active degradation of the mature transcript (20,21). Consistent with this, *RPL17A* and *RPL18A* mRNA levels also dropped sharply and relative intron signals increased, without (*RPL17A*) or with a small amount of skewing towards 5′ junctions (*RPL18A*; Figure 6C, D). Signals were also uniform across the intron regions of *YPR063C* and *URA2* after treatment with 6AU, with no change in the mRNA level or a slight increase, respectively (not shown). Pausing of RNAPII progression caused by 6AU therefore also revealed an apparent delay of polymerase progression in snoRNA-containing intron regions, with smaller effects at some other loci, complicated by a global response to the transcription stress caused by this drug.

The preferential increase in 5′ junction signals observed in interface mutants and in cells treated with 6AU

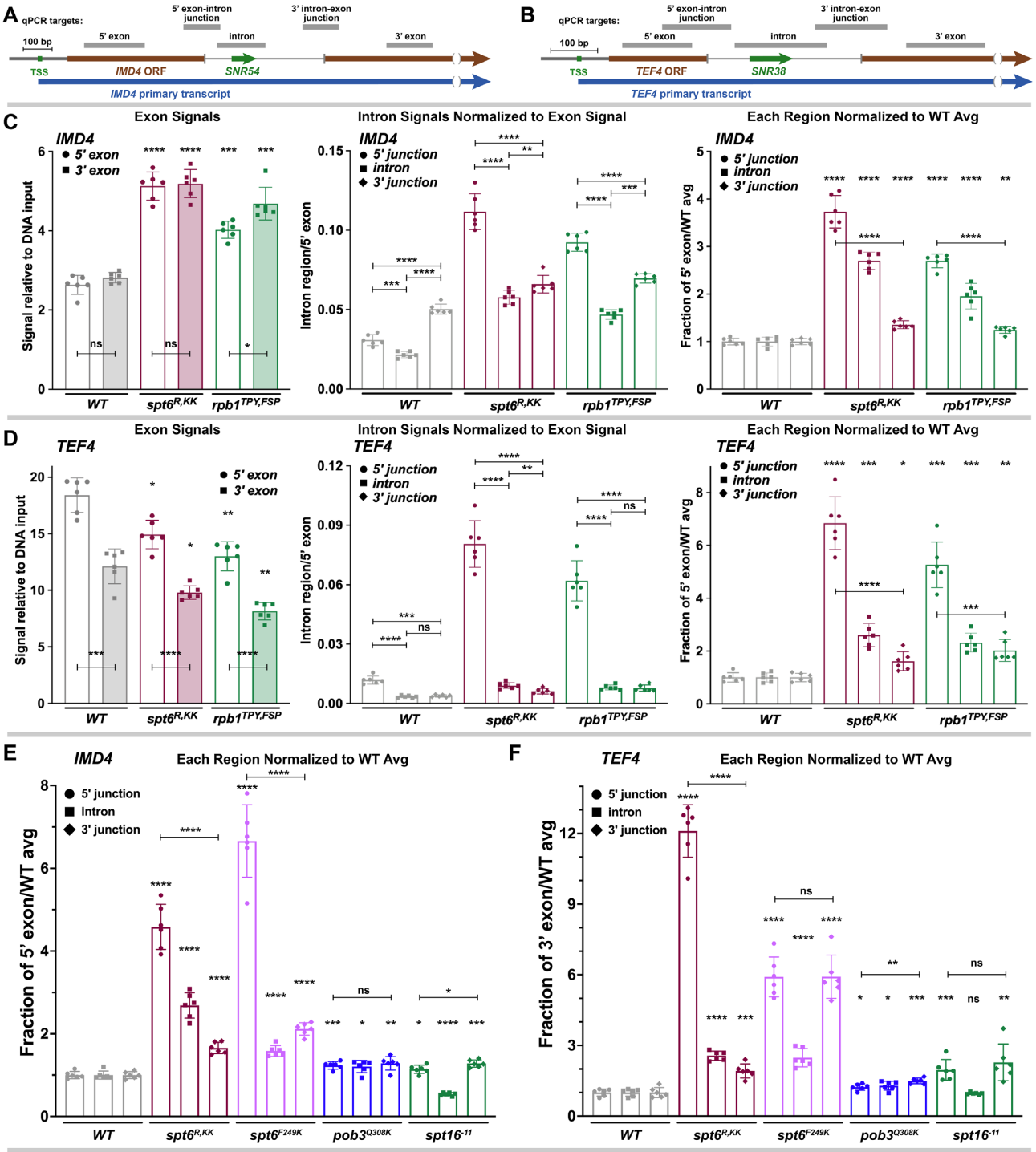


Figure 5. Interface mutations caused gradients of intron retention at two genes encoding snoRNAs. (**A**, **B**) Maps of the *IMD4* and *TEF4* loci are shown drawn to scale. (**C**, **D**) The left panels show qPCR measurements of the signal from 5′ and 3′ exons and the middle panel shows the signals from regions across the intron normalized to the 5′ exon as in Figure 4. The right panel shows the same ratios further normalized to the average value for the WT strain for the same region. The results from unpaired *t*-tests (comparing mutants to WT) or paired tests (comparing regions within the same genotype) are shown with the same symbols as in Figure 2. (**E**, **F**) As in panels (D) and (E) except using strains with a different allele of *SPT6* and two FACT mutations (Supplementary Table S1), and showing only the final results normalized to the average WT values as in the right panels of (D) and (E).

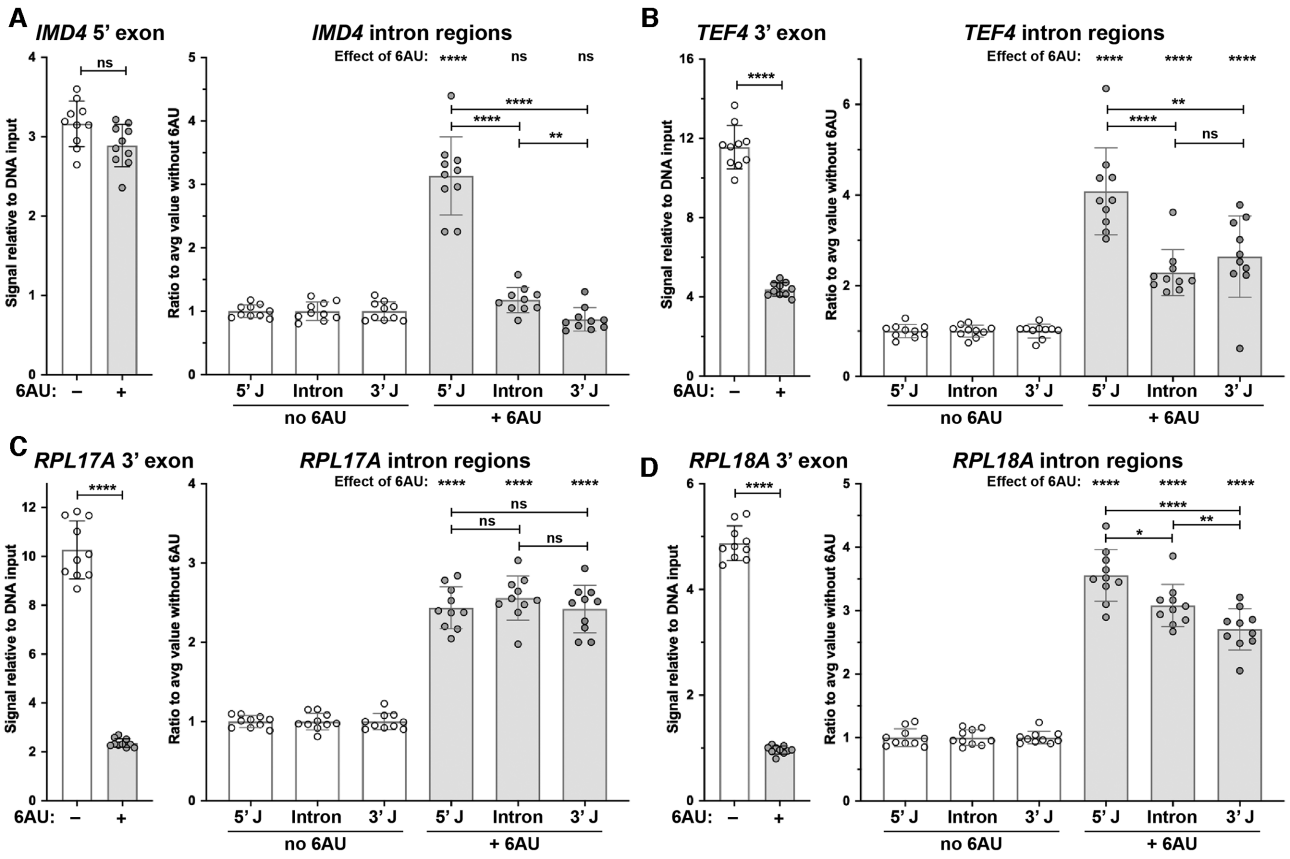


Figure 6. Treatment with 6AU reveals intron-retention gradients at some genes in WT cells. (**A–D**) 10 biological replicates of a WT strain (Supplementary Table S1) were grown in synthetic medium lacking uracil or in the same medium with 75 μg/ml 6AU for 2 h. qPCR signals across the introns of 4 genes (maps shown in Figures 4, 5) were measured and normalized as in Figure 4 except using the average value for each region in untreated cells. Paired t-test values for the differences between the distributions at the 5′ and 3′ junction regions are given using the same symbols as in Figure 2.

were additive (Supplementary Figure S6), again suggesting that these perturbations affected different mechanisms. An Rpb1 trigger loop mutation that increased the rate of catalysis caused a small increase in 5′ junction signal at *TEF4*, but neither this allele nor a slow mutant altered the effect of *spt6*^{R,KK} significantly (not shown). Apparent delays to progression were therefore induced independently by decreased UTP availability and loss of the Spt6-tSH2:Rpb1 interface, but the rate of the polymerization reaction itself had less effect on the degree of challenge posed by these difficult template regions.

Interface mutations affected the profile of Spt6 occupancy across specific gene bodies

Interface mutations cause a global decrease in Spt6 occupancy across transcription units (8), indicating that this interaction is at least partly responsible for maintaining the association of Spt6 with the elongation complex. However, the remaining occupancy was still proportional to the level of transcripts produced (Figure 7A) and to the occupancy in normal cells (Figure 7B, C). The Spt6-tSH2:Rpb1 interaction therefore contributes to the stability of the association of Spt6 with transcription complexes as detected by ChIP, but recruitment still occurs, possibly through the other contact sites noted above (Supplementary Figure S1A).

The decrease in Spt6 occupancy was relatively uniform, although the differential increased somewhat with increasing transcript levels, especially for genes with the highest transcript abundance (Figure 7A–C; the gap between WT and mutant distributions increased at higher transcript or occupancy levels). Decreased association of an essential elongation factor with transcription units would be expected to cause a decrease in transcripts, and this was the case for the RP genes (Figure 7D; large decreases in Spt6 occupancy at the RP genes were generally associated with large decreases in transcript levels in the lower left quadrant). However, while most genes experienced decreased Spt6 occupancy (most points were below zero on the y axis in Figure 7D), most did not display corresponding changes in transcript levels (most points clustered near zero on the *x*-axis in 7D and were not significantly affected in Figure 1B, C, Supplementary Figure S1D, G), producing a moderate overall correlation (Pearson $r = 0.41$). The quantitative relationship between Spt occupancy and transcript levels therefore appears to vary with specific genes or classes of genes rather than being a uniform feature of all transcription units.

The profile of Spt6 occupancy across gene bodies was not uniform, with lower levels over promoters and a peak near the 3′ end (Figure 7E, gray symbols). The apparent accumulation near the transcription termination site (TTS) was

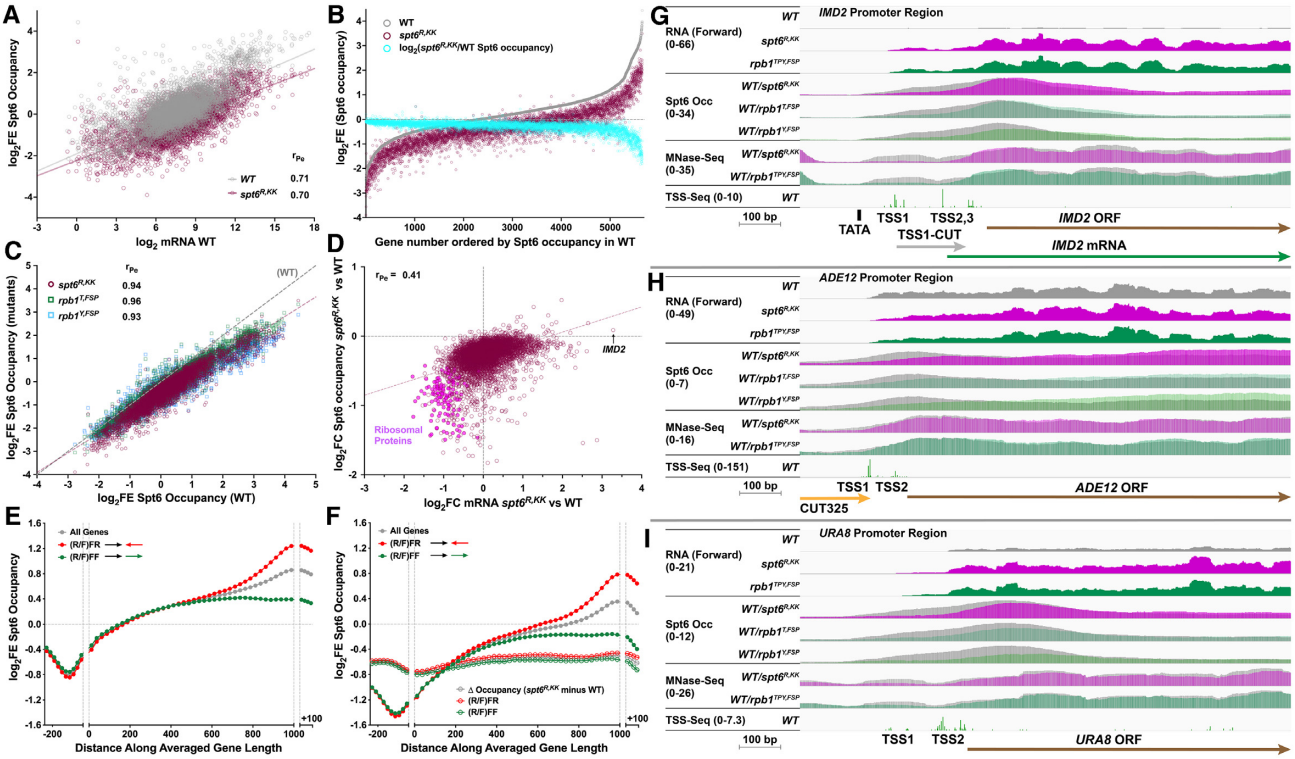


Figure 7. Interface mutations caused both generic and localized effects on Spt6 occupancy. (**A**) ChIP-seq was performed and $\log_2$FE values relative to pre-immune serum were calculated for each gene as described previously (8; Supplementary Table S1, 3 replicates for each genotype). These values are shown plotted against the total sense RNA signal for each gene (rlog$_2$ values) from the RNA-seq screen for WT and *spt6*^{R,KK} strains, with the Pearson correlation coefficients shown. The least-squares fit line for the mutant had a lower slope, revealing greater divergence at higher transcript levels. (**B**) The same values as in panel A were sorted by the transcript level in WT cells, and the difference between the $\log_2$FE values in mutant and WT cells was calculated (blue circles, this is the same as the $\log_2$FC for Spt6 occupancy in *spt6*^{R,KK} relative to WT). This value was consistently below zero indicating a global reduction in occupancy, but was lowest for the genes producing the most transcripts. (**C**) The $\log_2$FE values for Spt6 occupancy for mutants were plotted against the WT values for an *spt6*^{R,KK} mutant and two compound *rpb1* mutants that disrupt individual components of the N-terminal SH2 binding pocket (Supplementary Table S1). Overall correlations were high (Pearson r given), with greater divergence at higher occupancy levels, consistent with a global loss of Spt6 roughly proportional to the level observed in normal cells. (**D**) The $\log_2$FC values for Spt6 occupancy were plotted against the $\log_2$FC values in sense RNA for each gene with the Pearson correlation shown. *IMD2* and the RPs are indicated for reference. (**E, F**) The $\log_2$FE values for Spt6 occupancy across all genes normalized to the same averaged length are shown for WT and *spt6*^{R,KK} strains (closed circles). Genes were separated into classes by the orientation of the next downstream gene, and the difference between the WT and mutant values was calculated and plotted in panel (F) (open circles). (**G–I**) Browser tracks from RNA-seq, ChIP-seq (8), MNase-seq, and TSS-seq (39) were rendered in IGV (56) for 3 genes that illustrate a distinct pattern of Spt6 accumulation over the 5′ end in WT and associated changes in the interface mutants. Scales were adjusted in groups for each locus, as indicated. WT (gray) and mutant (lavender, green) are shown separately for RNA signals and overlaid for Spt6 occupancy and MNase-seq tracks. The *rpb1*^{T,FSP} and *rpb1*^{Y,FSP} strains gave similar overall patterns of Spt6 occupancy changes as *spt6*^{R,KK}, consistent with similar effects of mutations on phenotypes and Spt6 binding (Supplementary Figure S1B, reference 8).

due to genes whose downstream neighbor was oriented in opposition to the target gene, suggesting that this feature of the profile is caused by spillover from transcription of neighboring genes or some process associated with colliding transcription complexes (Figure 7E, green *versus* red symbols). The gradual ramping of occupancy near TSSs suggests that recruitment of Spt6 to transcription complexes might be gradual, or might be less stable during early stages of transcription and therefore not detected as efficiently by ChIP. These occupancy profiles did not depend on the Spt6-tSH2:Rpb1 interaction as the same shapes were observed in mutants, just at lower absolute levels (Figure 7F).

In contrast to the aggregated data, specific genes like *IMD2* had both an unusual distribution in WT cells and a characteristic set of changes in the interface mutants. Spt6 occupancy was higher over the 5′ end of *IMD2* than over the gene body, possibly due to association of Spt6 with the high level of constitutive transcription of TSS1-CUT, although the peak was downstream of the position of the CUT (Figure 7G, Supplementary Figure S3A). The accumulation was reduced over TSS1-CUT in interface mutants, even though transcripts from this site were maintained at their normal level or higher in these cells (Figure 2B, C, E, Supplementary Figure S2J, K), but occupancy eventually reached the WT level or higher in the gene body. Similar patterns were observed at *ADE12*, which had no evidence of CUT production from TSS1 (Figure 7H, Supplementary Figure S3G), *MET3*, which had some CUT formation (Supplementary Figures S7C, S3H), and *URA8* and *IMD3*, which each had strong evidence for CUT formation (Figure 7I, Supplementary Figure S2H, S3I, S7A). Spt6 occupancy was also reduced over the TSSs of genes with introns, although some of the decrease could be attributed to reduced transcript levels (Supplementary Figure S3A). Taken together, these results

show that a subset of genes have a peak of Spt6 occupancy over their 5′ ends, that interface mutations caused delayed, reduced, or unstable recruitment of Spt6 at many of these sites, and that loss of this accumulation was associated with a range of defects in downstream processes including a shift in the profile of TSS usage.

This pattern of Spt6 occupancy over 5′ ends of genes was not observed in averaged profiles so it is not a general feature of transcription units (Figure 7E, F). We therefore used an unbiased informatic screen to identify genes with peaks of Spt6 occupancy near the TSS, and found that genes with this feature also displayed strong depletion of occupancy over this region in interface mutants (Supplementary Figure S7D, E), indicating that this pattern of 5′ end accumulation of Spt6 depends on a functional Spt6-tSH2:Rpb1 interface. Notably, all of the examples described above that were initially identified by manual inspection ranked in the 97th percentile or higher by this metric (Supplementary Figure S7F). High scores were also associated with increased mRNA levels, suggesting that this pattern is usually associated with repression (Supplementary Figure S7F). We propose that Spt6 is generally recruited to transcription complexes as they transition from the PIC to the EC forms, the Spt6-tSH2:Rpb1 interface contributes to the efficiency or stability of this association, and also helps to establish the appropriate composition or configuration of the EC that is optimal for the circumstances encountered at each individual gene. The interaction appears to be especially important at loci where high levels of aborted transcripts are produced in the promoter region, where TSS selection is relevant for regulation of expression, or where chromatin structure creates strong links between initiation, elongation, and termination.

Interface mutations affected nucleosome occupancy and positioning

The NTD of Spt6 binds to nucleosomes and also to Spn1, these interactions are competitive with one another *in vitro*, and the function of the Spt6:Spn1 interface is important for balancing the activities of Spn1 and FACT *in vivo* (3,42,40; see Figure 1A). The Spt6-tSH2:Rpb1 interface could affect chromatin management by altering the localization or orientation of the Spt6-NTD:Spn1 module, or it could affect nucleosome integrity before, during, or after transcription and therefore dictate the need for cooperation among the histone chaperone activities of Spt6, Spn1, and FACT. To investigate these possibilities, we used MNase-seq to analyze nucleosome occupancy and positioning in interface mutants.

In general, interface mutations caused increased nucleosome occupancy over the 5′ ends of genes and decreased occupancy over about the last half of the average transcription unit (Figure 8A). *spt6*^{R,KK} and *rpb1*^{TPY,FSP} caused similar changes in nucleosome occupancy over different regions of the same genes (Supplementary Figure S8A), and similar profiles of effects across averaged transcription units, although *rpb1*^{TPY,FSP} had larger effects over 5′ ends (Figure 8A, B, Supplementary Figure S8A–D). The magnitude of the change in nucleosome occupancy was proportional to the level of transcripts, suggesting that these alterations in chromatin were likely to be associated with a defect in chromatin management during active transcription (compare the blue and red lines in Figure 8D). A FACT mutation also caused increased nucleosome occupancy over 5′ ends and a decrease over downstream regions (Supplementary Figure S8F), but this was associated with increased nucleosome occupancy over highly transcribed genes relative to WT (15), which was not observed in interface mutants (Supplementary Figure S8E, compare the red lines for each genotype). While the average nucleosome occupancy increased over 5′ ends in interface mutants, the decile of genes producing the most transcripts lost nucleosomes from this region. Importantly, this effect was largely limited to the +1 nucleosome at the 132 RP genes (Figure 8D, Supplementary Figure S8G; lavender lines). Occupancy returned to normal levels at the +2 site, then increased above the WT level as observed with average genes. This focused reduction of the +1 nucleosome of RP genes was not observed in a FACT mutant (Figure 8D), was also visible in browser tracks (Supplementary Figure S3A, see *RPL17A* or *RPL18A*), was much less pronounced or not seen in other highly expressed genes (Figure 8D, dotted red lines), and was the opposite of the general effect of interface mutations which was a specific increase in +1 nucleosome occupancy (Supplementary Figure S8B–D). The Spt6-tSH2:Rpb1 interaction therefore has a specific role in managing the +1 nucleosome, and this role is especially prominent for RP genes.

This analysis also revealed a distinction between disabling the interface from the Rpb1 side or the Spt6 side, as *spt6*^{R,KK} strains displayed minimal changes in nucleosome occupancy in regions upstream of genes/within NDRs, while these levels were increased globally in *rpb1*^{TPY,FSP} strains (Figure 8D, compare the regions upstream of the TSS). The integrity of the Rpb1 linker region was therefore important for keeping nucleosome occupancy low in NDRs, but this function appears to be independent of the interaction with Spt6-tSH2. This suggests that Spt6-tSH2:Rpb1 binding may be in competition with other factors that contribute to chromatin structure in promoters.

Nucleosome occupancy and transcript changes

If nucleosomes are simple barriers to transcription, then decreased occupancy should be associated with increased transcript levels, either because loss of the barrier allows more transcription or because more transcription erodes the barrier. RP genes displayed the opposite effect at the +1 nucleosome in interface mutations, with sharp reductions in both nucleosome occupancy and transcript levels. However, interpreting this subset of genes is complicated by the active turnover of some of these transcripts in response to transcription stress (20,21). To examine the relationship between nucleosome occupancy and transcript levels more globally, we compared the changes in nucleosome occupancy for NDRs, +1 nucleosomes, and gene bodies in interface mutants with the changes in transcript levels for all genes.

The $\log_2$FC values for transcript levels in interface mutants gave negative correlations with $\log_2$FC values for nucleosome occupancy upstream of the annotated TSS and also with average occupancy in whole gene bodies, as ex-

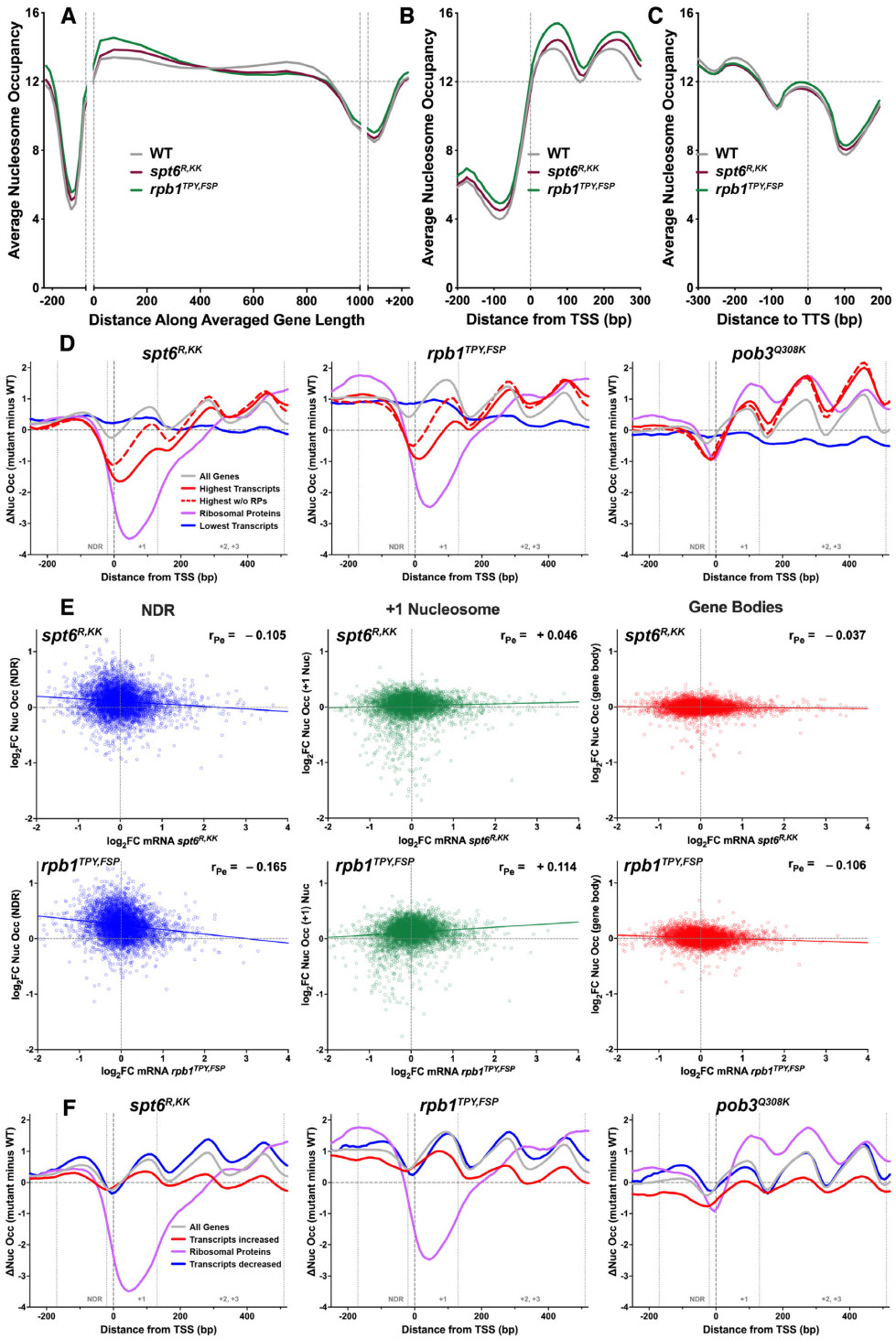


Figure 8. Interface mutations caused gene-specific changes in nucleosome occupancy. (**A–C**) MNase-seq was performed as described in Materials and Methods and the average nucleosome occupancy for each genomic sequence was determined. Values for WT and interface mutant strains were aligned for averaged gene locations (arbitrarily numbered 1–1000) and the regions 250 bp upstream and downstream for all genes in panel A, and for the regions near the annotated TSS and TTS sites in (B) and (C). (**D**) The differences in nucleosome occupancy near the TSS were determined for interface mutants for parallel analysis of results from a FACT mutant, *pob3*^{Q308K}, reported previously (15). Genes were sorted into deciles by transcript levels (DESeq2 rlog$_2$ values), with the highest and lowest deciles plotted. RP genes were removed from the highest group, replaced with the next-highest genes, and these two groups are shown separately to illustrate the distinct patterns observed with the RP genes and highly transcribed genes. (**E**) The log$_2$FC values for nucleosome occupancy in mutant cells relative to WT were calculated for each gene in the compound interface mutants using the NDR (the 130 bp upstream of the TSS), the +1 nucleosome (20 bp upstream to 130 bp downstream of the TSS) and averaged gene bodies, and plotted against the log$_2$FC in sense RNA for each gene in the same strain. Pearson correlation coefficients were calculated for each comparison, as indicated. (**F**) The change in nucleosome occupancy for each region was calculated as in panel D, but genes were sorted by the log$_2$FC in sense transcripts, and RPs were removed from the 'transcripts decreased' decile.

pected for changes in the level of a barrier to transcription (Figure 8E). The correlation values were weak for all regions, but it is not clear what the quantitative relationship should be between reduced nucleosome occupancy and transcript levels. Notably, the correlation was positive for +1 nucleosome occupancy globally, just as it was for RP genes (Figure 8E, middle panel), especially for *rpb1*^{TPY,FSP}, which also had the greater effect on occupancy at this site (Figure 8A,B, Supplementary Figure S8A–D). This suggests that higher +1 nucleosome occupancy is generally associated with increased transcript levels, the opposite of the expected outcome for a barrier to transcript production. A FACT mutant produced the expected negative correlations for all regions in a similar analysis (15). The Spt6-tSH2:Rpb1 interaction therefore has a distinct role in managing +1 nucleosomes that is not shared by the general histone chaperone FACT, and maintaining this nucleosome may have a positive contribution to transcript levels.

To further examine this relationship, we sorted the nucleosome occupancy profiles into deciles by the $\log_2$FC values for mRNA levels. Genes with the greatest reductions in transcript levels (considering RP genes independently, Figure 8F, Supplementary Figure S8H; blue and lavender lines) had more than the average increase in nucleosome occupancy, and genes with the largest increase in transcripts had less than the average increase (Figure 8F, Supplementary Figure S8H; red lines). The differences were small, especially at the +1 nucleosome for genes with decreased transcripts, where mutants and WT were essentially equal, and nucleosome occupancy was generally higher than WT across the region shown for all genes except at the +1 site of RP genes. A FACT mutant again had a different pattern of effects, with very little difference in the nucleosome occupancy profiles among classes except for a reduction at the subset of genes with the greatest increase in transcripts, and increased +1 and +2 occupancy for RP genes (Figure 8F, Supplementary Figure S8H). Changes in absolute nucleosome occupancy near the TSS were therefore small and the patterns of changes caused by interface mutations were different from those caused by a FACT mutation.

Nucleosome positioning near TSSs was also affected by interface mutations, shifting downstream 2–3 bp per nucleosome, similar to the effect of a FACT mutation (15; Figure 8B, Supplementary Figure S8I). The magnitude of the shift was proportional to transcript levels in all of the mutants (Supplementary Figure S8I), consistent with our previous speculation that it is caused by delayed restoration of nucleosomes after passage of RNAPII (15). Positioning near the TTS was unaffected (Figure 8C), so the shift appears to reveal a generic function of chaperones near initiation sites. The role of the Spt6-tSH2:Rpb1 interface may be even more focused on the TSS region than it is for FACT, as interface mutants appeared to alter primarily the +1 and +2 nucleosomes, while the effects of a FACT mutant extended downstream further (Supplementary Figure S8I).

In addition to these global comparisons, we also examined how interface mutations affected the expression of specific genes that are particularly sensitive to chromatin-mediated repression. Mutating any component of the interface except *rpb1*^{TPY} caused the Spt- phenotype (8 and Supplementary Figure S1B), and the *SER3* mRNA level, which is tightly linked to restoration of chromatin integrity after transcription of the upstream ncRNA *SRG1* (15,46–50), was consistently increased in the RNA-seq data. qPCR revealed a small increase in *SRG1* transcripts in interface mutants, and a larger increase in *SER3* mRNA (Supplementary Figure S8J-M). As noted above for several other genes with ncRNA transcripts produced in promoter regions, deletion of *RRP6* caused increases in both *SRG1* and *SER3* transcript levels, but loss of the exosome had no further effect on the elevated level of *SER3* mRNA produced in interface mutants. FACT and *spt6*^{F249K} also caused increased *SER3* mRNA, with smaller, mixed effects on *SRG1* levels, supporting a role for chromatin integrity in controlling *SER3* transcription, but also raising further questions about the precise mechanistic links between ncRNA transcription, termination and degradation of the ncRNA, chromatin restoration, and *SER3* promoter activity. The Spt6-tSH2:Rpb1 interface therefore appears to contribute to maintaining chromatin after transcription, but its precise role in the interaction between ncRNA transcription and mRNA production remains to be further investigated.

DISCUSSION

Spt6 makes multiple contacts with RNAPII subunits and accessory factors in the elongation complex, including a high-affinity interaction between the Spt6-tSH2 domain and phosphorylated residues in the linker region of Rpb1 (8,12,13, Figure 1A, Supplementary Figure S1A). We disrupted this interface by mutating either the binding pockets in Spt6-tSH2 or the bound sites on Rpb1 to determine the function of this interaction. Some defects were global, including reduced growth rates, reduced Spt6 occupancy (8), and altered chromatin architecture, but others were specific to subsets of genes. The Spt6-tSH2:Rpb1 interface therefore appears to serve a general function by contributing to the tethering of an essential elongation factor to the RNAPII complex, but it also has roles whose importance varies with local conditions at specific genes. This interaction therefore enhances the availability of the histone chaperone and core elongation functions of Spt6 to RNAPII, and may also provide a mechanism to coordinate these functions as changing local template conditions call for adjustments in the activities required. We do not know how information would be communicated through this interaction, but it is notable that the interface sits at the junction between the catalytic core of Rpb1 and the reservoir of accessory factors bound to its CTD. Simple mechanical tension on the tSH2 domain might then alter the proximity of those factors to their substrates, or change the position of the core elongation-factor domain of Spt6 relative to RNAPII. Other possibilities include inducing conformational changes in Spt6 or Rpb1 that affect their activities and therefore their responses to dynamic challenges encountered during transcription, or the availability of individual domains of Spt6 due to altered local recruitment. This work does not answer these mechanistic questions but identifies multiple situations where this interface might play a role.

We propose four sets of circumstances where Spt6-tSH2:Rpb1 interaction could function based on our observations (Figure 9). In Figure 9A, a PIC scans downstream

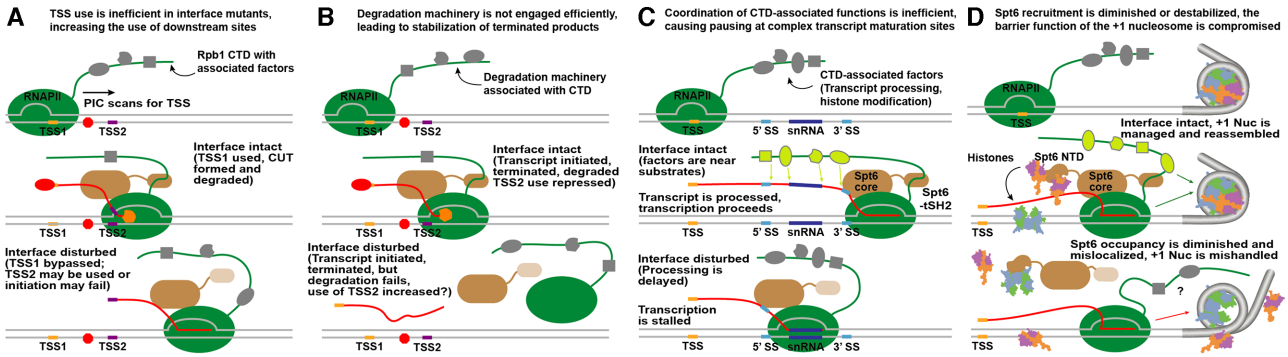


Figure 9. Proposed models for the functional roles of the Spt6-tSH2:Rpb1 interface in coordinating events during transcription elongation by RNAPII. RNAPII is represented as a green oval, with a green line indicating the Rpb1 CTD. Gray shapes indicate factors that can associate with the CTD that are inactive or unable to interact efficiently with their substrates, with green shading indicating their activation. Tan ovals represent the Spt6 core and tSH2 domains, with lighter tan in the bottom panels indicating interface mutations. (**A**) The interface integrity contributes to initiation-site selection. (**B**) Loss of the interface causes decreased efficiency of TSS-CUT degradation, which increases the use of a downstream TSS. Exosome deletion has the same effect but the mechanism coupling CUT degradation with TSS selection is unknown. (**C**) Loss of coordination with the reservoir of factors bound to the CTD causes a reduction in splicing globally and delayed progression through challenging template regions. (**D**) Coordination of chaperone activities with RNAPII progression is important for maintaining nucleosome integrity globally, with a special function at +1 nucleosomes, especially those at RP genes.

for an initiation site (25,51), and this process is influenced by interactions with Spt6 that are partially dependent on the Spt6-tSH2:Rpb1 interaction. TSS selection might be impacted globally by interface mutations, but this would only affect the relative level of expression of the subset of genes where TSS selection is a major component of regulation, such as *IMD2*. Use of the TSS1-CUT initiation site at this locus normally exceeds the use of downstream sites by over 100-fold (Figure 2C), so the increase in the number of PICs that bypass this site could significantly increase the opportunity to use the TSS2,3 cluster and therefore the level of mRNA produced even if the efficiency of using the TSS2,3 sites is reduced somewhat. TSS selection could therefore be impacted globally but have gene-specific effects, as also seen with the genes examined in Figure 3 and Supplementary Figure S3B-I.

We observed that unstable transcripts that overlap sites of productive initiation, such as the *IMD2* TSS1-CUT, the CUT initiated upstream of *IMD3*, the TSS1 products from *MET3* and *URA8,* and the ncRNA from *SRG1* all accumulated at abnormally high levels in interface mutants. This did not appear to be the result of increased promoter activity, as double mutants with both interface and exosome mutations generally displayed either small increases in these transcripts or decreases relative to exosome mutants alone (Figure 2C, Supplementary Figures S2E, J, S3F–I, S8L). Termination appeared to be normal globally (Supplementary Figures S2D–F), so we propose that these results indicate a defect in degradation of this class of CUTs, due to a lack of coordination between termination, exosome recruitment, and the use of downstream initiation sites (Figure 9B). RNAPII may need to delay progression to allow termination and exosome engagement to be efficient, and the interface may be a conduit for processing this information and responding. This again raises questions that are not answered by this work, suggesting a complex relationship among transcription that impinges on potential downstream initiation sites, the efficiency of degradation of those transcripts, and initiation from these sites, with this relationship somehow being affected by the integrity of the Spt6-tSH2:Rpb1 interaction.

We also observed a general defect in splicing efficiency, which could be caused by inefficient positioning of factors associated with the CTD (Figure 9C). In this model, factors might be recruited normally to the CTD, but disruption of the Spt6-tSH2:Rpb1 interface alters the orientation of the CTD relative to the emerging transcript, diminishing the proximity of the factors to their substrates. Alternatively (or in addition to this model), RNAPII progression might be delayed in regions where transcript maturation occurs, and the interface could serve as a way for the splicing factors to communicate the need for a delay to the catalytic core by altering the elongation factor activity of Spt6. Failure of this mechanism could lead to diminished maturation of mRNAs, or a lack of coordination between the processes could impair progression through particularly complex regions such as the snoRNA-containing introns of IMD4 and TEF4 as multiple minor delays added in series cause a major delay.

The NTD of Spt6 binds nucleosomes and Spn1, supporting a complex relationship among the histone chaperone activities of Spt6, Spn1 and FACT (3,40), so inefficient recruitment of Spt6 to transcription complexes could lead to changes in chromatin due to a simple reduction in available chaperone capacity (Figure 9D). Comparison of the changes in nucleosome occupancy and positioning in interface mutants with those caused by a FACT mutant supported this interpretation for some genes but also revealed distinct properties for the Spt6-tSH2:Rpb1 interaction. For example, FACT and interface mutations caused similar global downstream shifts in nucleosome occupancy near the 5′ ends of genes, especially at highly transcribed loci, and increased transcripts from *SER3*, whose regulation requires efficient restoration of nucleosomes after passage of the transcription complex (15,47–50,52). In contrast, interface mutants had focused effects on maintaining +1 nucleosomes in general, with a particularly strong role at RP genes, that were not shared by a FACT mutant. This nucle-

osome has an unusual composition, containing much of the variant H2A.Z found in yeast cells, and is expected to play a distinct role in regulating initiation of transcription (53,54). The role of the Spt6-tSH2:Rpb1 interface in managing this nucleosome therefore underscores not only a specific role for the Spt6 chaperone function at subsets of genes, but also the potentially positive role for this nucleosome in initiation at some genes or perhaps more globally.

Spt6 occupancy was not uniform across gene bodies or among all genes. Levels were notably high at genes like *IMD2* over the region where the decision to either terminate and degrade the TSS1-CUT product or initiate from the TSS2,3 cluster are made. This accumulation was decreased in interface mutants, even though transcripts from these regions remained high or increased. To the extent that the level of transcripts observed reflects the frequency of transcription, this could suggest that high Spt6 occupancy was not a passive consequence of high transcription. A similar pattern of high Spt6 accumulation over regions where choices among optional outcomes were being made, and reduced occupancy in interface mutants, was observed at several genes. We propose that lowered or weakened recruitment of Spt6 diminished the ability to form effective transcription complexes, leading to inefficient or uncoordinated functioning of the mechanisms outlined in Figure 9. In these models, the Spt6-tSH2:Rpb1 interaction promotes proper orientation of the Rpb1 CTD and possibly serves as a conduit for communication between the catalytic core of RNAPII and accessory factors needed only in specific circumstances. This includes interacting appropriately with nucleosomes as they are encountered, especially the unusual +1 nucleosome, and promoting efficient restoration of chromatin after transcription. The different profiles of effects at different genes underscores the individual nature of each gene, and suggests that the composition and properties of the RNAPII EC may need to be tailored specifically for each locus to produce optimal regulation. Our results therefore support a role for the Spt6-tSH2 in tethering the elongation factor and histone chaperone domains of Spt6 to the RNAPII elongation complex, but that role appears to involve more than mechanical attachment. Instead, this interaction appears to participate in coordinating functional transitions that are important for making appropriate decisions at multiple stages of transcription from initiation through termination, affecting the efficiency of multiple pathways.

DATA AVAILABILITY

The RNA-seq and MNase-seq data have been deposited in NCBI's Gene Expression Omnibus (GEO) database (55) and are accessible through the superfamily accession number GSE184955 or its components GSE174427 (RNA-seq), GSE184954 (single-end MNase-seq), and GSE189831 (paired-end MNase-seq).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank members of David Stillman's lab, especially Emily Parnell, Robert Yarrington, and Yaxin Yu, for assistance with quantitative PCR and RNA extraction methods, Jeff Morgan and Jared Rutter for assistance with the primer extension assay, Fred Winston for providing materials for strain construction, and Mahesh Chandrasekharan for assistance with genomic methods.

FUNDING

NIH [GM116560 to T.F., C.P.H., AI150464 to C.P.H., GM064649 to T.F.]. Funding for open access charge: NIH.
Conflict of interest statement. None declared.

REFERENCES

1. Johnson,S.J., Close,D., Robinson,H., Vallet-Gely,I., Dove,S.L. and Hill,C.P. (2008) Crystal structure and RNA binding of the Tex protein from *Pseudomonas aeruginosa*. *J Mol Biol*, **377**, 1460–1473.
2. Close,D., Johnson,S.J., Sdano,M.A., McDonald,S.M., Robinson,H., Formosa,T. and Hill,C.P. (2011) Crystal structures of the *S. cerevisiae* Spt6 core and C-terminal tandem SH2 domain. *J Mol Biol*, **408**, 697–713.
3. McDonald,S.M., Close,D., Xin,H., Formosa,T. and Hill,C.P. (2010) Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding. *Mol Cell*, **40**, 725–735.
4. Diebold,M.L., Koch,M., Loeliger,E., Cura,V., Winston,F., Cavarelli,J. and Romier,C. (2010) The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26. *EMBO J.*, **29** , 3979–3991.
5. Diebold,M.L., Loeliger,E., Koch,M., Winston,F., Cavarelli,J. and Romier,C. (2010) A non-canonical tandem SH2 enables interaction of elongation factor SPT6 with RNA polymerase II. *J. Biol. Chem.*, **285**, 38389–38398.
6. Sun,M., Lariviere,L., Dengl,S., Mayer,A. and Cramer,P. (2010) A tandem SH2 domain in transcription elongation factor Spt6 binds the phosphorylated RNA polymerase II CTD. *J. Biol. Chem.*,**285**, 41597–41603.
7. Bortvin,A. and Winston,F. (1996) Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science*, **272**, 1473–1476.
8. Sdano,M.A., Fulcher,J.M., Palani,S., Chandrasekharan,M.B., Parnell,T.J., Whitby,F.G., Formosa,T. and Hill,C.P. (2017) A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *Elife*, **6**, e28723.
9. Buratowski,S. (2009) Progression through the RNA polymerase II CTD cycle. *Mol Cell*, **36**, 541–546.
10. Corden,J.L. (2013) RNA polymerase II C-terminal domain: tethering transcription to transcript and template. *Chem Rev*, **113**, 8423–8455.
11. Eick,D. and Geyer,M. (2013) The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem Rev*, **113**, 8456–8490.
12. Chun,Y., Joo,Y.J., Suh,H., Batot,G., Hill,C.P., Formosa,T. and Buratowski,S. (2019) Selective kinase inhibition shows that Bur1 (Cdk9) phosphorylates the Rpb1 linker in vivo. *Mol. Cell. Biol.*, **39**, e00602-18.
13. Vos,S.M., Farnung,L., Boehning,M., Wigge,C., Linden,A., Urlaub,H. and Cramer,P. (2018) Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature*, **560**, 607–612.
14. Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.E., Seidman,J.G., Smith,J.A. and Struhl,K. (1988) In: Ausubel,M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds). *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., Media, PA, Vol. **1 and 2**.
15. McCullough,L.L., Pham,T.H., Parnell,T.J., Connell,Z., Chandrasekharan,M.B., Stillman,D.J. and Formosa,T. (2019) Establishment and maintenance of chromatin architecture are promoted independently of transcription by the histone chaperone FACT and H3-K56 acetylation in *Saccharomyces cerevisiae*. *Genetics*, **211**, 877–892.

16. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, **9**, R137.
17. Hahn,S.L., Hahn,M. and Hayward,W.S. (1989) Structural organization of upstream exons and distribution of transcription start sites in the chicken c-myb gene. *Mol. Cell. Biol.*, **9**, 837–843.
18. Kaplan,C.D., Jin,H., Zhang,I.L. and Belyanin,A. (2012) Dissection of Pol II trigger loop function and Pol II activity-dependent control of start site selection in vivo. *PLoS Genet*, **8**, e1002627.
19. Dronamraju,R., Hepperla,A.J., Shibata,Y., Adams,A.T., Magnuson,T., Davis,I.J. and Strahl,B.D. (2018) Spt6 association with RNA polymerase II directs mRNA turnover During Transcription. *Mol Cell*, **70**, 1054–1066.
20. Sun,M., Schwalb,B., Pirkl,N., Maier,K.C., Schenk,A., Failmezger,H., Tresch,A. and Cramer,P. (2013) Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. *Mol Cell*, **52**, 52–62.
21. O'Duibhir,E., Lijnzaad,P., Benschop,J.J., Lenstra,T.L., van Leenen,D., Groot Koerkamp,M.J., Margaritis,T., Brok,M.O., Kemmeren,P. and Holstege,F.C. (2014) Cell cycle population effects in perturbation studies. *Mol Syst Biol*, **10**, 732.
22. Kuehner,J.N. and Brow,D.A. (2008) Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol Cell*, **31**, 201–211.
23. Jenks,M.H., O'Rourke,T.W. and Reines,D. (2008) Properties of an intergenic terminator and start site switch that regulate IMD2 transcription in yeast. *Mol Cell Biol*, **28**, 3883–3893.
24. Malik,I., Qiu,C., Snavely,T. and Kaplan,C.D. (2017) Wide-ranging and unexpected consequences of altered Pol II catalytic activity in vivo. *Nucleic Acids Res*, **45**, 4431–4451.
25. Qiu,C., Jin,H., Vvedenskaya,I., Llenas,J.A., Zhao,T., Malik,I., Visbisky,A.M., Schwartz,S.L., Cui,P., Cabart,P. *et al.* (2020) Universal promoter scanning by Pol II during transcription initiation in *Saccharomyces cerevisiae*. *Genome Biol*, **21**, 132.
26. Wyers,F., Rougemaille,M., Badis,G., Rousselle,J.-C., Dufour,M.-E., Boulay,J., Régnault,B., Devaux,F., Namane,A., Séraphin,B. *et al.* (2005) Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell*, **121**, 725–737.
27. Arigo,J.T., Eyler,D.E., Carroll,K.L. and Corden,J.L. (2006) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell*, **23**, 841–851.
28. Neil,H., Malabat,C., Aubenton-Carafa,Y., Xu,Z., Steinmetz,L.M. and Jacquier,A. (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature*, **457**, 1038–1042.
29. Xu,Z., Wei,W., Gagneur,J., Perocchi,F., Clauder-Munster,S., Camblong,J., Guffanti,E., Stutz,F., Huber,W. and Steinmetz,L.M. (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature*, **457**, 1033–1037.
30. Steinmetz,E.J., Conrad,N.K., Brow,D.A. and Corden,J.L. (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. *Nature*, **413**, 327–331.
31. Arndt,K.M. and Reines,D. (2015) Termination of transcription of short noncoding RNAs by RNA polymerase II. *Annu. Rev. Biochem.*, **84**, 381–404.
32. Porrua,O. and Libri,D. (2015) Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell. Biol.*, **16**, 190–202.
33. Fox,M.J., Gao,H., Smith-Kinnaman,W.R., Liu,Y. and Mosley,A.L. (2015) The exosome component Rrp6 is required for RNA polymerase II termination at specific targets of the Nrd1-Nab3 pathway. *PLoS Genet*, **11**, e1004999.
34. Moreau,K., Le Dantec,A., Mosrin-Huaman,C., Bigot,Y., Piegu,B. and Rahmouni,A.R. (2019) Perturbation of mRNP biogenesis reveals a dynamic landscape of the Rrp6-dependent surveillance machinery trafficking along the yeast genome. *RNA Biol*, **16**, 879–889.
35. Villa,T., Barucco,M., Martin-Niclos,M.J., Jacquier,A. and Libri,D. (2020) Degradation of non-coding RNAs promotes recycling of termination factors at sites of transcription. *Cell Rep.*, **32**, 107942.
36. Wang,D., Bushnell,D.A., Westover,K.D., Kaplan,C.D. and Kornberg,R.D. (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell*, **127**, 941–954.
37. Kaplan,C.D., Larsson,K.M. and Kornberg,R.D. (2008) The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol Cell*, **30**, 547–556.
38. Larson,M.H., Zhou,J., Kaplan,C.D., Palangat,M., Kornberg,R.D., Landick,R. and Block,S.M. (2012) Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 6555–6560.
39. Doris,S.M., Chuang,J., Viktorovskaya,O., Murawska,M., Spatt,D., Churchman,L.S. and Winston,F. (2018) Spt6 is required for the fidelity of promoter selection. *Mol. Cell*, **72**, 687–699.
40. Viktorovskaya,O., Chuang,J., Jain,D., Reim,N.I., Lopez-Rivera,F., Murawska,M., Spatt,D., Churchman,L.S., Park,P.J. and Winston,F. (2021) Essential histone chaperones collaborate to regulate transcription and chromatin integrity. *Genes. Dev.*, **35**, 698–712.
41. Formosa,T. and Winston,F. (2020) The role of FACT in managing chromatin: disruption, assembly, or repair? *Nucleic Acids Res*, **48**, 11929–11941.
42. Reim,N.I., Chuang,J., Jain,D., Alver,B.H., Park,P.J. and Winston,F. (2020) The conserved elongation factor Spn1 is required for normal transcription, histone modifications, and splicing in Saccharomyces cerevisiae. *Nucleic Acids Res*, **48**, 10241–10258.
43. Neves,L.T., Douglass,S., Spreafico,R., Venkataramanan,S., Kress,T.L. and Johnson,T.L. (2017) The histone variant H2A.Z promotes efficient cotranscriptional splicing in S. cerevisiae. *Genes Dev*, **31**, 702–717.
44. Peck,S.A., Hughes,K.D., Victorino,J.F. and Mosley,A.L. (2019) Writing a wrong: Coupled RNA polymerase II transcription and RNA quality control. *Wiley Interdiscip. Rev. RNA*, **10**, e1529.
45. Powell,W. and Reines,D. (1996) Mutations in the second largest subunit of RNA polymerase II cause 6- azauracil sensitivity in yeast and increased transcriptional arrest in vitro. *J Biol Chem*, **271**, 6866–6873.
46. Kaplan,C.D., Laprade,L. and Winston,F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science*, **301**, 1096–1099.
47. Hainer,S.J. and Martens,J.A. (2011) Identification of histone mutants that are defective for transcription-coupled nucleosome occupancy. *Mol. Cell Biol.*, **31**, 3557–3568.
48. Hainer,S.J., Pruneski,J.A., Mitchell,R.D., Monteverde,R.M. and Martens,J.A. (2011) Intergenic transcription causes repression by directing nucleosome assembly. *Genes Dev.*, **25**, 29–40.
49. Martens,J.A., Laprade,L. and Winston,F. (2004) Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature*, **429**, 571–574.
50. Thebault,P., Boutin,G., Bhat,W., Rufiange,A., Martens,J. and Nourani,A. (2011) Transcription regulation by the noncoding RNA SRG1 requires Spt2-dependent chromatin deposition in the wake of RNA polymerase II. *Mol Cell Biol*, **31**, 1288–1300.
51. Kuehner,J.N. and Brow,D.A. (2006) Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. *J Biol Chem*, **281**, 14119–14128.
52. Martens,J.A., Wu,P.Y. and Winston,F. (2005) Regulation of an intergenic transcript controls adjacent gene transcription in *Saccharomyces cerevisiae*. *Genes Dev*, **19**, 2695–2704.
53. Clapier,C.R., Iwasa,J., Cairns,B.R. and Peterson,C.L. (2017) Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol*, **18**, 407–422.
54. Luk,E., Ranjan,A., Fitzgerald,P.C., Mizuguchi,G., Huang,Y., Wei,D. and Wu,C. (2010) Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. *Cell*, **143**, 725–736.
55. Edgar,R., Domrachev,M. and Lash,A.E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*, **30**, 207–210.
56. Thorvaldsdottir,H., Robinson,J.T. and Mesirov,J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform*, **14**, 178–192.