*Nucleic Acids Research*, 2025, **53**, gkaf179
https://doi.org/10.1093/nar/gkaf179
**Gene regulation, Chromatin and Epigenetics**

# Overlapping and distinct functions of SPT6, PNUTS, and PCF11 in regulating transcription termination

**Fabienne Bejjani*, Emmanuel Ségéral, Kevin Mosca, Adriana Lecourieux, May Bakail, Meriem Hamoudi, Stéphane Emiliani***

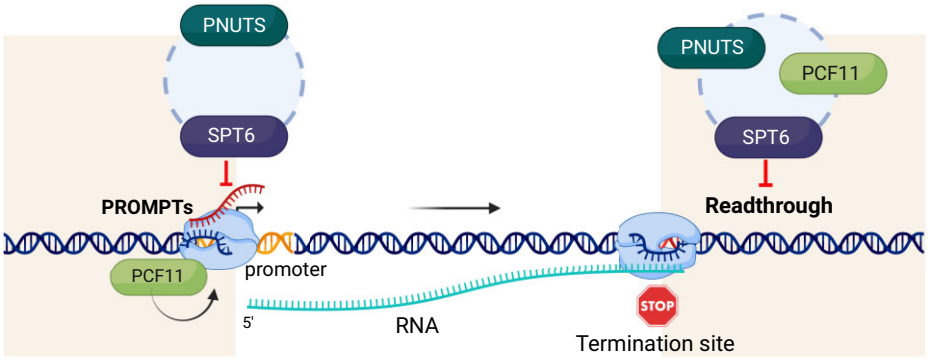
Institut Cochin, Université Paris Cité, INSERM, CNRS, Paris F-75014, France

*To whom correspondence should be addressed. Email: stephane.emiliani@inserm.fr
Correspondence may also be addressed to Fabienne Bejjani. Email: fabienne.bejjani@inserm.fr

## Abstract

The histone chaperone and transcription elongation factor SPT6 is integral to RNA polymerase II (RNAPII) activity. SPT6 also plays a crucial role in regulating transcription termination, although the mechanisms involved are largely unknown. In an attempt to identify the pathways employed by SPT6 in this regulation, we found that, while SPT6 and its partner IWS1 interact and co-localize with RNAPII, their functions diverge significantly at gene termination sites. Depletion of SPT6, but not of IWS1, results in extensive readthrough transcription, indicating that SPT6 independently regulates transcription termination. Further analysis identified that the cleavage and polyadenylation factor PCF11 and the phosphatase regulatory protein PNUTS collaborate with SPT6 in this process. These findings suggest that SPT6 may facilitate transcription termination by recruiting PNUTS and PCF11 to RNAPII. Additionally, SPT6 and PNUTS jointly restrict promoter upstream transcripts (PROMPTs), whereas PCF11 presence is essential for their accumulation in the absence of SPT6 at hundreds of genes. Thus, SPT6, PCF11, and PNUTS have both distinct and overlapping functions in transcription termination. Our data highlight the pivotal role of SPT6 in ensuring proper transcription termination at the 5′ and 3′-ends of genes.

## Graphical abstract




## Introduction

In eukaryotic organisms, the positioning and composition of nucleosomes play a crucial role in regulating gene expression. Transcription of genes by RNA polymerase II (RNAPII) collides with nucleosomes that organize chromatin within gene bodies. To overcome these nucleosomal barriers while maintaining chromatin integrity during transcription elongation, RNAPII associates with histone chaperones. Suppressor of Ty 6 (SPT6), is a histone chaperone and transcription elongation factor that is conserved throughout eukaryotes [1].

SPT6 is a large protein that contains two Src homology 2 (SH2) domains in its C-terminal region that bind directly to the phosphorylated CTD (C-terminal domain) linker of the RPB1 subunit of the RNAPII [2]. SPT6 functions as an elongation factor both *in vitro* [3] and *in vivo* [4–8]. In yeast, Spt6 interacts directly with histones via its acidic N-terminal domain (TND) [9–12]. It regulates nucleosome positioning and occupancy [13–17], as well as H3K36 methylation in gene bodies [18–23]. As a part of the RNAPII elongation complex [24], SPT6 participates in the disassembly and reassembly of nucleosomes in the wake of the RNAPII [24–26]. Depletion of Spt6 results in spurious transcription initiation and increased levels of antisense transcription [13, 16, 27, 28].

The TND of SPT6 also interacts directly with IWS1, another histone chaperone also known as Spn1 in yeast [29–31], and this interaction is essential to recruit IWS1 to the

**Received:** September 3, 2024. **Revised:** February 18, 2025. **Editorial Decision:** February 20, 2025. **Accepted:** February 24, 2025

RNAPII elongation complex [11, 21, 32]. In yeast, Spn1 binds histones H3-H4 and nucleosomes *in vitro* [33], and regulates normal transcript levels and the distribution of H3K4 and H3K36 methylation genome-wide [34]. Disrupting the Spt6–Spn1 interaction results in extensive changes in both transcription and nucleosome organization [34, 35]. However, Spn1 and Spt6 display distinct chromatin occupancy profiles near the 3′-end of coding regions, suggesting additional independent functions during transcription elongation [36]. IWS1 interacts with the methyltransferase SETD2 and was proposed to mediate H3K36 trimethylation at cellular genes [21]. IWS1 contains a TFIIS TND that binds with high affinity to a TND-interacting motif (TIM) located in the N-terminal region of SPT6 [37, 38]. IWS1 also harbors three TIMs in its disordered N-terminal region that interact with TND-containing factors including TFIIS, ELOA, PNUTS, and LEDGF/p75 [38, 39]. Thus, IWS1 could serve as a scaffold to assemble multiple factors involved in transcription processes.

Recent studies in human cells highlighted a role of SPT6 in enforcing transcription termination. Knockdown of SPT6 results in elongation defects at coding regions, while promoting transcription of long noncoding RNAs (lncRNAs) including long intergenic noncoding RNA, antisense promoter upstream transcripts (PROMPTs), and bidirectional enhancer-associated transcripts or enhancer RNA (eRNA) [18]. SPT6 is required for the association of cleavage and polyadenylation (CPA) and transcription termination factors to the RNAPII [5, 40]. Thus, in addition to the loss of processive RNAPII elongation, depletion of SPT6 also leads to increased readthrough transcription at the 3′-end of genes [5].

The CPA complex, consisting of the CPSF, CFIm, CFIIm, and CstF sub-complexes, orchestrates messenger RNA (mRNA) CPA at the 3′-ends of protein-coding (pc) genes, a process tightly coupled with transcription termination. CPSF subunits bind directly to the AAUAAA hexamer of the polyadenylation signal (PAS) and RNA cleavage is accomplished by the endonuclease CPSF3 at the cleavage site identified by the CA dinucleotide [41]. CFIIm, containing PCF11 and CLP1, interacts only transiently with the CPA complex [42]. PCF11 acts as a regulatory factor of the CPA complex activities and favors proximal PAS usage at the 3′-end of pc genes [43, 44]. In addition, PCF11 simultaneously binds RNA and phosphorylated serine2 of the RNAPII CTD via its CTD interaction domain to dismantle the elongating RNAPII complex and induce transcription termination [45–47].

Transcription termination of pc genes also relies on PNUTS, a nuclear protein phosphatase 1 (PP1) regulatory subunit that associates with the CPA complex [42]. PNUTS-PP1 dephosphorylates SPT5 to decrease RNAPII speed at the 3′-end of pc genes and promote PAS-dependent transcription termination [48]. PNUTS also cooperates with the transcription terminator complex Restrictor, composed of ZC3H4 and WDR82, to enforce transcription termination at enhancers and PROMPTs [49, 50].

Given the emerging functions of SPT6 and IWS1, we initially aimed at understanding their functional relationship during transcription. In this study, we find that SPT6 and IWS1 are similarly associated with RNAPII on gene bodies but this association diverges at the 3′-end of genes. Notably, knockdown of IWS1 results in changes in mRNA levels of very few genes when compared to SPT6. Furthermore, depletion of SPT6, but not of IWS1, induced 3′-end readthrough transcription at thousands of regions, suggesting that IWS1 is not involved in the regulation of RNAPII transcription termination by SPT6. Instead, screening for factors that could cooperate with SPT6, we identified the CPA factor PCF11 and the phosphatase regulatory protein PNUTS. We find that both PCF11 and PNUTS assist SPT6 in promoting transcription termination at the 3′-end of genes. We also show that a similar cooperation exists between SPT6 and PNUTS to restrict expression of PROMPTs. However, in the absence of SPT6, depletion of PCF11 inhibits accumulation of PROMPTs.

## Materials and methods

### Cell culture

HeLa LTR-Luc [51] and HEK293T cells (CRL-3216, obtained from the American Type Culture Collection, Rockville, MD, USA), were cultured in Dulbecco's modified Eagle's medium (DMEM) + GlutaMAX™-I (Gibco, 61965-026). J-Lat Tat-GFP Clone A1 (obtained from NIH HIV Reagent Program RRID:CVCL_1G42), were grown in Roswell Park Memorial Institute (RPMI) 1640 medium + GlutaMAX™-I (Gibco, 61870-010). Both DMEM and RPMI media were supplemented with 10% of decomplemented fetal bovine serum (Gibco, 10270-106) and 1% antibiotic–antimycotic (10 000 U/ml penicillin, 10 000 µg/ml streptomycin, and 25 µg/ml of amphotericin B; Gibco, 15240-062). All cells were grown in a humidified 5% $CO_2$ incubator at 37°C and were routinely tested for the absence of *Mycoplasma* contamination.

### siRNA transfection

A total of 100 000 HeLa cells were seeded in a 12-well plate and transfected with small interfering RNA (siRNA) at a final concentration of 20 nM (see references in Supplementary Table S1) (single transfections: 10 nM siCT + 10 nM of specific siRNA; double transfections: 10 nM of each siRNA) using Lipofectamine™ RNAiMAX Transfection Reagent (Invitrogen, 13778150) and Opti-MEM™ (Gibco, 31985070) according to the manufacturer's instructions. After 72 h, cells were harvested for RNA and protein extraction.

### Construction of shRNA plasmids and production of virus-like particles

Control short hairpin RNA (shRNA) sequence (shCT) or shRNA sequence targeting PNUTS (see references in Supplementary Table S1) were cloned into the pLKO.1-TRC cloning vector (gift from David Root, Addgene plasmid #10878; http://n2t.net/addgene:10878; RRID: Addgene_10 878) according to the manufacturer's instructions. Virus-like particles (VLPs) were produced by transient co-transfection of 3 µg of pLKO.1 vector containing either shCT or shPNUTS, 2.25 µg of the packaging plasmid psPAX2 (gift from Didier Trono; Addgene plasmid #12260; http://n2t.net/addgene:12260; RRID: Addgene_12 260), and 0.75 µg of the VSV-G expression vector (pMD.G) into 3 million HEK293T cells in a 10 cm dish using polyethylenimine (PEI) (Polysciences, 23966-1) and Opti-MEM™ (Gibco, 31985070). The following day, the medium was removed, and the cells were washed with phosphate-buffered saline (PBS). Five millilitres of fresh medium were added, and VLP stocks were collected and filtered using a 0.45 µm disposable filter (ClearLine, 146561) the next day.

### Transduction of VLPs

To transduce HEK293T cells, 300 µl of shCT or shPNUTS VLPs were mixed with 200 µl of DMEM medium, and the mixture was added to 3 million cells in a 24-well plate for 3 h at 37°C. After incubation, the cells were washed twice with PBS and transferred to six-well plates. The following day, puromycin (Gibco™, A1113803) was added at a 1 µg/ml final concentration. Puromycin selection was maintained for at least 7 days post-transduction.

### Transfection of mammalian expression vectors

The FLAG-HA_SPT6 expression vector was a kind gift from Dr Hiroshi Handa [3]. The open reading frames (ORFs) of human PNUTS or PNUTS W401A were cloned in-frame with a Flag-HA tag into a pcDNA3.1 vector (Flag-HA-PNUTS; Flag-HA-PNUTSW401A). For immunoprecipitation experiments, 5 µg of these mammalian expression vectors were transfected into 3 million HEK293T cells using PEI (Polysciences, 23966-1) and Opti-MEM™ (Gibco, 31985070). Cells were harvested 48 h post-transfection.

### Construction of IWS1 knockout cells

J-Lat A1 cells were nucleofected with LentiCRISPR v2 vector (Addgene Plasmid #5296) expressing a guide RNA targeting IWS1 (see Supplementary Table S1). The following day, puromycin (Gibco™, A1113803) was added at a 1 µg/ml final concentration. Puromycin selection was maintained for 4 days post transfection then removed. Cells were cloned by limiting dilution in 96-well flat-bottomed culture plates and expanded for 15 days. Clones were then screened by polymerase chain reaction (PCR) amplification of the targeted region of the genome.

### Immunoprecipitation

HEK293T or J-Lat A1 cells were harvested, and nuclear protein extracts were prepared following the protocol described by Matkovic *et al.* [52]. Briefly, cells were washed twice with PBS and then resuspended in Cytoplasmic Lysis Buffer (CLB) [10 mM Tris, pH = 7.4; 10 mM NaCl; 3 mM $MgCl_2$; 0.5% NP-40, 1× protease inhibitors (Roche, 05056489001)]. Nuclei were immediately pelleted by centrifugation at 300 × *g* for 4 min at 4°C. The supernatant (cytoplasmic fraction) was either discarded or kept to verify cell fractionation. Another CLB wash was performed before lysing the nuclei in a Nuclear Lysis Buffer (NLB) [20 mM Tris, pH = 8; 3 mM KCl, 150 mM NaCl, 1 mM $MgCl_2$ and 0.2% Triton™ X-100, 1× protease inhibitors (Roche, 05 056 489 001)]. After a 30 min incubation on ice, nuclear protein extracts were cleared by centrifugation at 14 000 rpm for 15 min at 4°C. The supernatant was collected and 250 µg to 1 mg of protein extracts were used for immunoprecipitation. For Fig. 1B, total protein extracts (lysed directly in NLB) were used. Primary antibodies (anti-SPT6 or anti-Flag, see references in Supplementary Table S1) were incubated with the protein extracts overnight on a rotating wheel at 4°C. The following day, Protein A (Invitrogen™, Dynabeads™ Protein A, 10002D) or Protein G (Invitrogen™, Dynabeads™ Protein G, 10004D) conjugated beads were added to the mixtures containing anti-SPT6 (rabbit) or anti-Flag (mouse), respectively. The mixtures were incubated for 30 min at room temperature on a rotating wheel. The beads were washed 3 times with NLB and then resuspended in 40 µl of 1× Laemmli buffer [diluted from NuPAGE™ LDS Sample Buffer (4×) (Invitrogen™, NP0007) and NuPAGE™ Sample Reducing Agent (10×) (Invitrogen™, NP0009)]. After heating for 10 min at 70°C, the eluted proteins were collected, and 20 µl were used for western blotting along with 10 µg of input proteins.

### Glycerol gradient sedimentation

Nuclear protein extracts from HeLa and HEK293T cells were prepared as described above. The equivalent of 900 µg of proteins were loaded onto a 5%-60% glycerol gradient prepared in NLB [20 mM Tris, pH = 8; 3 mM KCl, 150 mM NaCl, 1 mM $MgCl_2$ and 0.2% Triton™ X-100, 1× protease inhibitors (Roche, 05056489001)], and layered by 400 µl fractions into a 5 ml open-top polypropylene tube (Beckman Coulter, 326 819). Macromolecular complexes were then separated by ultracentrifugation at 45 000 rpm for 14 h at 4°C using an SW 55Ti rotor (Beckman). Subsequently, 12 equal fractions were collected in total (∼400 µl each), and 20 µl of each fraction were used for western blotting.

### Immunoblotting

Cells were incubated for 30 min on ice in an sodium dodecyl sulfate-free lysis buffer [150 mM NaCl, 10 mM Tris, pH = 8, 0.5% Triton™ X-100, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.1% sodium deoxycholate, 1× protease inhibitors (Roche, 05056489001)]. After incubation, cell lysates were clarified by centrifugation at 21 000 × *g* for 15 min at 4°C. The supernatant was collected and protein concentration was estimated using the BRADFORD assay (Bio-Rad Protein Assay Dye Reagent Concentrate, 5000006). Equal amounts of proteins were diluted in 1× Laemmli buffer [diluted from NuPAGE™ LDS Sample Buffer (4×) (Invitrogen™, NP0007) and NuPAGE™ Sample Reducing Agent (10×) (Invitrogen™, NP0009)], then heated for 10 min at 70°C. Proteins were separated in denaturing conditions using a 4%–12% gradient bis-tris polyacrylamide gel (Invitrogen™, NW04122BOX) at 35mA, then transferred onto a Polyvinylidene fluoride (PVDF) membrane (Millipore; Immobilon-P; 0, 45 µm; IPVH00010) for 2 h at 200 mA. Membranes were saturated in 5% milk in Tris-buffered saline (TBS) buffer (0.1% Tween 20) (1 h, room temperature), and incubated overnight with primary antibodies (see references in Supplementary Table S1) in 5% milk in TBS (0.1% Tween 20) at 4°C. Horseradish peroxidase-conjugated secondary antibodies (references in Supplementary Table S1) were used along with Enhanced chemiluminescence (ECL) western blotting detection reagents (Amersham™ ECL Select™ Western Blotting Detection Reagent, RPN2235) to detect proteins by chemiluminescence using Fusion FX (Vilber).

### RNA extraction and RT-qPCR

Total cellular RNAs were extracted using the MagnetaPure 32 automated magnetic bead-based RNA purification (MACHEREY-NAGEL) and the NucleoMag® RNA isolation kit (MACHEREY-NAGEL, 872168), following the manufacturer's instructions. One to two micrograms of purified RNAs were treated with TURBO™ DNase (2 U/µl) (Invitrogen™, AM2239) and reverse transcribed using either the High-Capacity complementary DNA (cDNA) Reverse Transcription (RT) Kit for random primers (Applied Biosystems™, 4368813) or the SuperScript™ II Reverse Transcriptase for

oligo(dT) primers (Invitrogen™, 18064022). After reverse transcription, cDNAs were diluted to a final volume of 200 µl (1/10 dilution) and quantitative PCR (qPCR) was performed using the SsoAdvanced Universal SYBR® Green Supermix (Bio-Rad, 1725274) on the LightCycler 480 Instrument II (Roche). RNA levels were first normalized to the control condition (using the ΔCq method) and then adjusted by a normalization factor accounting for the levels of two housekeeping genes (*KDSR* and *RN7SK*).

## Nuclear run-On

Nascent RNA transcripts were analyzed by Nuclear run-On (NRO) according to Ait Said *et al.* [51]. Briefly, HeLa cells were collected from confluent 10 cm dishes and lysed in Swelling Buffer (10 mM Tris, pH = 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$). Nuclei were then resuspended in Freezing Buffer (50 mM Tris, pH = 8.3, 40% glycerol, 5 mM $MgCl_2$, 0.1 mM EDTA). A 5-min *in-vitro* transcription was then performed at 30°C in presence of the run-on mix (10 mM Tris, pH = 8, 5 mM $MgCl_2$, 1 mM KCl, 0.5 mM of each of ATP, GTP, and Br-UTP, 2.33 mM CTP, 0.4 U/µl RNase-IN, 1% sarkosyl). Total RNAs were then extracted using TRIzol-LS reagent (Invitrogen, 10296010), and 8 µg were DNase treated as previously described and subjected to immunoprecipitation with 2 µg of anti-BrdU (see Supplementary Table S1 for reference). After two washes with PBS + 0.1% Tween-20, immunoprecipitated RNAs were extracted using TRIzol reagent. Reverse transcription and qPCR were performed as previously described. RNA levels were first normalized to the control condition (using the ΔCq method) and then adjusted by a normalization factor accounting for the levels of two housekeeping genes (KDSR and 18S) whose expression levels remain unaffected by the different protein knockdowns (Supplementary Fig. S1A) [51].

## ChIP-seq

For SPT6, IWS1, and RNAPII ChIP-seq, chromatin immunoprecipitation was performed using the iDeal ChIP-seq kit for Transcription Factors (Diagenode, C01010055) according to the manufacturer's instructions. Briefly, 2.5 million HeLa cells per IP were crosslinked with 1% formaldehyde (Thermo Scientific™, 28908) in DMEM for 10 min at room temperature. Chromatin from lysed nuclei was sheared to 200–300 bp fragments using Bioruptor® Pico sonication device (Diagenode, B01060010) and cleared by centrifugation. Sheared chromatin was incubated overnight at 4°C with 3 µg of antibody (see references in Supplementary Table S1) coupled to the Dia-Mag protein A-coated magnetic beads (from Diagenode kit). Beads and immunocomplexes were then washed and eluted per the manufacturer's instructions. For PNUTS ChIP-seq, an additional crosslink with the ChIP cross-link Gold reagent (Diagenode, C01019027) was performed before formaldehyde crosslinking. ChIP was done using the iDeal ChIP-seq kit for Transcription Factors (Diagenode, C01010055) per the manufacturer's protocol with 15 µl of anti-PNUTS (unknown concentration) (see reference in Supplementary Table S1). Sequencing libraries from immunoprecipitated DNA were prepared using the MicroPlex kit v2 (Diagenode) according to the manufacturer's protocol. In brief, immunoprecipitated DNA was quantified using Qubit fluorometer (Thermo Fisher), and 1.5–35 ng (for SPT6, IWS1, and RNAPII ChIP) or 2–10 ng (for PNUTS ChIP) were used as input material. DNA was repaired and end-blunted by enzymatic treatment. Stem-loop adaptors with blocked 5′-ends were ligated to the 5′-end of the genomic DNA, leaving a nick at the 3′-end. The 3′-ends of the genomic DNA are extended to complete library synthesis and Illumina-compatible indexes were added through amplification. Libraries were purified using AMPure XP beads protocol (Beckman Coulter, Indianapolis, IN) and quantified using Qubit fluorometer. Libraries' fragment size distribution was verified using the Bioanalyzer High Sensitivity DNA chip (Agilent Technologies, Santa Clara, CA). Libraries were pooled and spike-in PhiX Control v3 (Illumina, San Diego, CA) was added to represent 1% of the total Passing Filter (PF) reads. Clusters were generated and sequenced using a Nextseq 500 instrument (Illumina) in single read mode (75 cycles) at the GENOM'IC platform, Institut Cochin.

ChIP-seq analysis was performed using an in-house pipeline. Quality control for raw FASTQ files was conducted using FastQC (version 0.12.1). Sequencing on NextSeq500 platform (Illumina) generates one fastq file per sample and per lane (four lanes in total on the flow cell). For a given sample, the four different files—corresponding to the four lanes—were merged for easier handling. Sequencing adapters were trimmed using Trimmomatic (version 0.39). Trimmed reads were aligned to the human reference genome (GRCh38) using Bowtie2 (version 2.4.2) (bowtie2 -x index_file -U input.trim.fastq.gz -S file.sam). Aligned SAM files were converted to BAM format and sorted using SAMtools (version 1.3.1). Peaks were called using MACS2 (version 2.2.7.1) (macs2 callpeak -t file.bam -c input.bam -f BAM -B –SPMR -g hs -n sample_hg38 –outdir sample_hg38). Processed BAM files were converted to BigWig format for visualization using bamCoverage from the deepTools suite (version 3.5.0) (bamCoverage –bam input.bam -o output.bw –normalizeUsing RPKM –effectiveGenomeSize 2 747 877 777 –centerReads –ignoreDuplicates –minMappingQuality 20).

Two independent biological replicates were generated for all ChIP-seq experiments. Due to the high signal correlation between the biological replicates (Pearson correlation coefficient ranging from 0.93 and 0.99), they were merged to facilitate downstream analysis.

## RNA-seq

HeLa cells were transfected as described previously. Three independent biological replicates were prepared for each condition. Total RNAs were extracted using the ReliaPrep™ RNA Miniprep Systems (Promega, Z6012) as per the manufacturer's instructions. Sample quality was verified using the Bioanalyzer RNA 6000 Nano kit (Agilent). The libraries were prepared following the Stranded Total RNA Prep Ligation with Ribo-Zero Plus protocol from Illumina, starting from 200 ng of high-quality total RNA (RNA integrity number (RIN) > 9). Paired end (2 × 59) sequencing was performed on an Illumina NextSeq 2000 platform. Demultiplexing and quality control were conducted with the AOZAN software (ENS, Paris) [53] at the GENOM'IC platform, Institut Cochin.

RNA-seq analysis was performed using an in-house pipeline. Shortly, initial quality control of the raw FASTQ files was performed using FastQC (version 0.12.1). Raw sequencing files from lane 1 (L1) and lane 2 (L2) were merged for read 1 (R1) and read 2 (R2) using the cat command (cat L1_R1.fastq.gz L2_R1.fastq.gz > merged_R1.fastq.gz). The merged FASTQ files were aligned to the human reference genome (GRCh38) using HISAT2 (version 2.2.1) (hisat2

–dta -x genome -1 merged_R1.fastq.gz -2 merged_R2.fastq.gz -S output.sam). The SAM files were then converted to BAM format and sorted using SAMtools (version 1.3.1). Aligned reads were quantified using StringTie (version 2.1.5) (stringtie output.bam -eB -G annotation.gtf -o output.gtf -A gene_abund.tab). Differential expression analysis was performed using DESeq2 (version 1.30.1) in R (version 4.0.3) following the standard workflow. Size factor normalization was used to account for differences in sequencing depth and library size between samples. For visualization and further analysis, BAM files were processed using bamCoverage from the deepTools suite (version 3.5.0) (bamCoverage -b output.bam –scaleFactor -o output.bw). The three biological replicates were shown independently for heatmaps, and they were merged to facilitate downstream analysis.

### Analysis of readthrough transcription and PROMPTs

RNA-seq average scores at previously defined readthrough and PROMPTs (Supplementary Fig. S1E and Supplementary Fig. S6A) were calculated using "multiBigWigSummary –BED" from the deepTools suite (version 3.5.0). Regions with an average score lower than 1 reads per million (RPM) in all conditions were removed. $Log_2FC$ was calculated for each condition with respect to the siCT condition. Positive PROMPT or readthrough regions were defined as those having a $log_2FC > 1$ in the three biological replicates. For PROMPTs, an additional filter was applied to distinguish between PROMPTs and readthrough transcripts originating from a downstream gene on the opposite strand. Specifically, RNA-seq average scores were calculated within gene bodies but on the opposite strand. Genes with a $log_2FC$ exceeding 1 in both the PROMPT region and the gene body on the opposite strand were excluded from further analysis.

## Results

### SPT6 has IWS1-independent transcriptional functions

SPT6 is a large protein with multiple domains. The TND interacts with nucleosomes and harbors two conserved TIMs, capable of interacting with multiple TND-containing proteins including IWS1 [11, 32, 38] (Fig. 1A). The core region binds to the RPB4–RPB7 stalk of RNAPII [24], and the C-terminal region contains a tandem SH2 (tSH2) domain that interacts directly with the linker region of RPB1 and its phosphorylated CTD [2, 54] (Fig. 1A). Our immunoprecipitation experiments confirmed that both overexpressed (Fig. 1B) and endogenous (Supplementary Fig. S1B) SPT6 interact with IWS1 in HEK293T cells.

We first analyzed the chromatin distribution of SPT6 and IWS1 on RefSeq-annotated genes. Both proteins exhibited similar distributions, with a prominent enrichment around the transcription start site (TSS) region and downstream of genes transcription end sites (TES) (Fig. 1C). This is also reflected by a high correlation coefficient between SPT6 and IWS1 signals at the gene regions (Fig. 1D, Pearson's correlation coefficient r = 0.97). We also performed an RNAPII ChIP-seq using an antibody that recognizes the largest subunit, RPB1. The RNAPII is rather mainly enriched at gene promoters, and slightly enriched downstream of the TES (Supplementary Fig. S1C). Nevertheless, the SPT6 signal highly correlates with that of RNAPII (Fig. 1D, Pearson's correlation coefficient r = 0.94). Interestingly, the correlation between IWS1 and RNAPII signals is lower (Fig. 1D, Pearson's correlation coefficient r = 0.88). This prompted us to inspect the ChIP-seq signal ratios between SPT6, IWS1, and RNAPII. We observed that SPT6 and IWS1 remain equally associated with RNAPII on gene bodies until ∼4 kb upstream of the TES, where IWS1 dissociates more markedly than SPT6 (Fig. 1E). This suggests that, similar to what was previously observed in yeast, the transcriptional functions of SPT6 and IWS1 might diverge in the termination region [36].

Then, to determine whether the divergence in transcriptional function of SPT6 and IWS1 could be attributed to their presence within separate macromolecular complexes, we inspected their sedimentation profiles. The mixture of multiprotein complexes from HeLa cells (Fig. 1F) and HEK293T cells (Supplementary Fig. S1D) was separated using a 5%–60% glycerol gradient sedimentation and their distribution was examined in different fractions by western blotting. The results indicated that, although SPT6 and IWS1 were found within complexes of similar molecular weights (fractions 6–8 in Fig. 1F, fractions 6–12 in Supplementary Fig. S1D), IWS1 was also enriched in independent complexes (fractions 4 and 5, Fig. 1F and Supplementary Fig. S1D), suggesting potential independent roles for this protein in the nucleus.

To assess the impact of these factors on gene expression, we performed total RNA-seq experiments in HeLa cells upon depletion of SPT6 and IWS1 (Fig. 1G). Target genes were identified using a 1.5 FC threshold and a padj < 0.05. In order to eliminate false-positive differentially expressed genes, we kept only those with an RNAPII ChIP-seq peak in their promoter region (TSS ± 1.5 kb). A total of 3405 and 119 genes were differentially expressed upon SPT6 and IWS1 depletion, respectively (Fig. 1H), with similar numbers of genes up- and downregulated (1844 versus 1561 upon SPT6 depletion and 64 versus 55 upon IWS1 depletion) (Supplementary Fig. S1E). Surprisingly, only 47 genes were commonly regulated by both factors, suggesting that the transcriptional roles of SPT6 and its closest partner, IWS1, rarely overlap.

Recent studies have highlighted an important role for SPT6 in restricting lncRNAs and readthrough transcripts [5, 18]. To determine whether IWS1 similarly affects these transcripts, we first defined readthrough regions for RefSeq-annotated genes as 5 kb-regions downstream of genes TES. We excluded regions overlapping with annotated genes on the same strand to avoid contaminating signal from intragenic transcription, and retained regions with an average RNA-seq signal higher than 1 RPM in at least one condition and corresponding to genes having an RNAPII ChIP-seq peak within their promoter region (TSS ± 1.5 kb) (Supplementary Fig. S1F). We obtained a total of 8725 readthrough regions. In line with the previous studies, SPT6 significantly affected readthrough transcripts, while IWS1 had minimal impact (Fig. 1I and J).

We further examined whether SPT6 and IWS1 collaborate to regulate transcription termination by qPCR analysis of transcripts produced at one of SPT6 targets, the *SEC22B* gene, upon single or double knockdown of SPT6 and IWS1 (Fig. 1K). We designed amplicons to detect readthrough and mRNA transcripts as a control. While SEC22B mRNA levels remained largely unchanged in all conditions, our results showed increased SEC22B readthrough transcripts upon SPT6 depletion, and no significant change upon IWS1 depletion (Fig. 1L). Double knockdown of SPT6 and IWS1

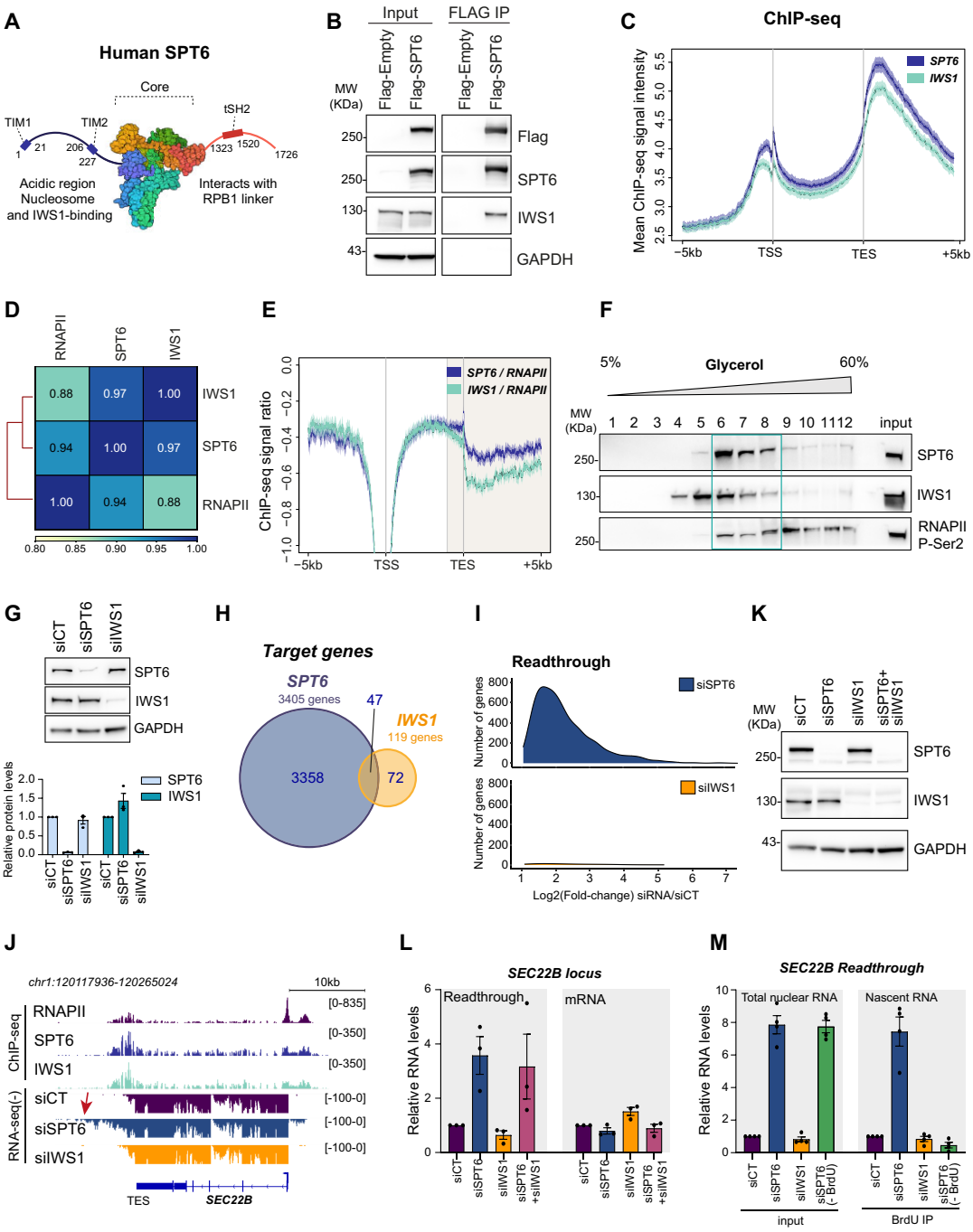



**Figure 1.** SPT6 And IWS1 chromatin binding and transcriptional roles. (**A**) A schematic representation of SPT6 N-terminal (1–283 amino acids), core (284–1287 amino acids), and C-terminal (1288–1726 amino acids) regions. The illustration denotes N to C terminal protein domains. (**B**) Overexpressed Flag-tagged SPT6 from HEK293T total protein extracts was immunoprecipitated using an anti-Flag antibody and analyzed by western blotting using indicated antibodies. (**C**) SPT6 and IWS1 mean ChIP-seq signal intensities plotted at RefSeq-annotated genes. The solid line represents the mean. The dark area represents the standard error and light area represents the 95% confidence interval. (**D**) Heatmap representing the Pearson correlation coefficients calculated for IWS1, SPT6, and RNAPII merged ChIP-seq signals, on RefSeq-annotated genes (±1 kb). (**E**) SPT6/RNAPII and IWS1/RNAPII mean ChIP-seq signal ratios calculated on RefSeq-annotated genes in HeLa cells. The solid line represents the mean. The dark area represents the standard error and light area represents the 95% confidence interval. (**F**) Total protein extracts from HeLa cells were separated using a 5%–60% glycerol gradient. A total of 12 fractions were recovered after ultracentrifugation and analyzed by western blotting. (**G**) Western blot showing SPT6 and IWS1 depletions upon siRNA transfection in RNA-seq experiments. Lower panel shows protein quantification relative to GAPDH and to the siCT condition ($n = 3$). (**H**) SPT6 and IWS1 target genes in HeLa cells. Positive targets correspond to genes having a $|FC| > |\pm 1.5|$ and a padj $< 0.05$. (**I**) Density plot showing the gene count of SPT6 and IWS1 readthrough targets, highlighting the distribution of their $\log_2$ fold-change (FC) values. (**J**) *SEC22B* gene locus featuring RNAPII, SPT6, and IWS1 ChIP-seq profiles, as well as RNA-seq profiles upon depletion of SPT6 and IWS1 (negative strand). The arrow highlights readthrough transcription (**K**) Western blot showing SPT6 and IWS1 depletion upon siRNA transfection (see the "Materials and methods" section for details). (**L**) SEC22B readthrough and mRNA levels were assayed by RT-qPCR in three independent experiments. Values were normalized to the siCT condition arbitrarily set to 1. (**M**) SEC22B readthrough levels were assayed by nuclear run-on experiments. The input represents total nuclear RNAs. The siSPT6 (-BrdU) condition is used to control the specificity of the anti-BrdU immunoprecipitation. Values were normalized to the siCT condition arbitrarily set to 1, and to the KDSR and 18S housekeeping genes.

affected readthrough transcripts similarly to SPT6 knockdown alone (Fig. 1L). This pattern was also observed for another gene, CHST3 (Supplementary Fig. S1G). Notably, readthrough transcripts obtained upon SPT6 depletion were also detected using oligo(dT) primers during reverse transcription (Supplementary Fig. S1H), suggesting that these transcripts are polyadenylated.

Given that we are analyzing total cellular RNAs, we wondered whether the effect of IWS1 could be masked by its effect on unstable RNAs that could be degraded at the time of detection. To investigate this, we examined nascent transcripts at the SEC22B and CHST3 loci. We performed nuclear run-on experiments after depleting SPT6 and IWS1, using a 5-min *in-vitro* transcription in presence of BrdU. We also included a condition without BrdU to verify the specificity of our anti-BrdU immunoprecipitation. SPT6 had a similar effect on nascent and total transcripts for both genes (Fig. 1M and Supplementary Fig. S1I). Consistent with its effect on total transcripts, IWS1 did not affect the levels of nascent readthrough transcripts at both the SEC22B and CHST3 loci (Fig. 1M and Supplementary Fig. S1I). Nascent full-length mRNA levels of both genes were unaffected by the different knockdowns (Supplementary Fig. S1J), suggesting a specific effect of SPT6 on transcription termination. Altogether, these findings suggest an IWS1-independent role for SPT6 in transcription termination.

## Investigating the multifaceted pathways of transcription termination involving SPT6

In order to gain a comprehensive view of the pathways employed by SPT6 to terminate transcription, we investigated its functional cooperation with the main protein complexes described to have roles in mammalian transcription termination: (i) Canonical CPA factors, represented by the endonuclease of the CPSF complex (CPSF3/CPSF73) and the CFIIm factor PCF11 [44, 55], (ii) Torpedo factors PNUTS, SPT5, and XRN2 [48, 56], (iii) Restrictor factors WDR82 and ZC3H4 [50, 57, 58] and (iv) Integrator, represented by its catalytic component INTS11 [59].

RNA levels produced at four gene loci (SEC22B, ZNF180, CHST3, and FLNB) were assessed using primers that overlap (i) readthrough transcripts, and (ii) the last exon, to account for possible roles on gene full-length transcript levels (Fig. 2A). The nine factors were depleted independently using siRNAs in HeLa cells, or co-depleted with SPT6 to assess potential functional cooperation (Supplementary Fig. S2).

For the four genes, the individual depletions of SPT6, PNUTS, and PCF11 had the most consistent effect on increasing readthrough transcripts (Fig. 2B, upper panel). The depletion of CPSF3, SPT5, WDR82 also showed mild effects depending the gene. Depletion of the other factors (XRN2, ZC3H4, and INTS11) had little to no impact on readthrough transcription at these loci (Fig. 2B, upper panel). Importantly, a synergistic increase in readthrough transcripts was observed when SPT6 was co-depleted with either PCF11 or PNUTS, indicating cooperation between the three factors (Fig. 2B, lower panel). Of note, no major effect was detected on the last exon, suggesting that the observed increase in readthrough transcripts is due to a termination defect rather than intragenic transcription regulation.

Overall, these findings underscore a diversity in the mechanisms of transcription termination that involve SPT6, with a potential collaboration with PNUTS and PCF11 at several genes.

## Interaction dynamics between PNUTS, PCF11, and SPT6

PNUTS is an RNA-binding protein that harbors a TND in its N-terminal region (Fig. 3A, upper panel). This domain was described by Cermakova and colleagues to interact with an IWS1-TIM domain [38]. PCF11 is a component of the CFIIm complex that interacts with the CTD of RNAPII via its N-terminal region and with the transcribed RNA via its C-terminal region, therefore coupling transcription to 3′-end processing [60, 61] (Fig. 3A, lower panel). To investigate whether PNUTS and PCF11 interact with SPT6 and whether this potential interaction is mediated by IWS1, we immunoprecipitated SPT6 in both wild-type (WT) and IWS1-knockout J-Lat A1 cells. Our results show that both PNUTS and PCF11 co-immunoprecipitate with SPT6 in WT cells (Fig. 3B and Supplementary Fig. S3A) and that these interactions are unaffected by IWS1 knockout (Fig. 3B and Supplementary Fig. S3B).

Then, to determine whether the interaction between SPT6 and PNUTS might depend on the phosphatase subunit of the PTW:PP1 complex, we tested the interaction between SPT6 and either WT PNUTS or a mutant, PNUTS–W401A [62], that does not interact with PP1. As shown in Fig. 3C and Supplementary Fig. S3C, the interaction between PNUTS and SPT6 was independent of PP1 phosphatase. These results highlight an interaction between SPT6 and PNUTS that appears to be independent of IWS1 and PP1-dephosphorylation.

PNUTS was also described to affect RNAPII CTD phosphorylation in *Drosophila melanogaster* [63]. Given that SPT6 interacts with a phosphorylated form of RNAPII [64], we wondered whether PNUTS could regulate the interaction between SPT6 and the RNAPII. We immunoprecipitated SPT6 in HeLa cells transduced with either a control shRNA or a PNUTS-targeting shRNA. SPT6 levels were slightly increased upon PNUTS down-regulation (Fig. 3D), but its interaction with the RNAPII remained unaffected (see RNAPII:SPT6 ratio in Supplementary Fig. S3D).

## PNUTS, PCF11, and SPT6 chromatin occupancy in HeLa cells

Both PNUTS and PCF11 play significant roles in mammalian transcription termination. PCF11 is known to be enriched at both the 5′- and 3′-ends of genes in HeLa cells and depletion of PCF11 has been reported to shift and reduce RNAPII pausing downstream of the PAS, indicating impaired transcription termination [44].

PNUTS has been shown to bind gene promoters and transcription end sites [65]. Additionally, it co-purifies with the human pre-mRNA 3′-processing complex [42], and regulates RNAPII speed at the 3′-end of genes [48]. Despite these findings, there is a scarcity of published ChIP-seq data for endogenous PNUTS. Therefore, we set out to evaluate its enrichment in the promoter and termination regions of genes, as well as other regulatory regions, such as enhancers and repressed regions. We performed PNUTS ChIP-seq in HeLa cells and defined gene promoter and termination zones as the 3kb regions surrounding the RefSeq-annotated gene TSS and TES, respectively. Enhancers and repressed regions were de-

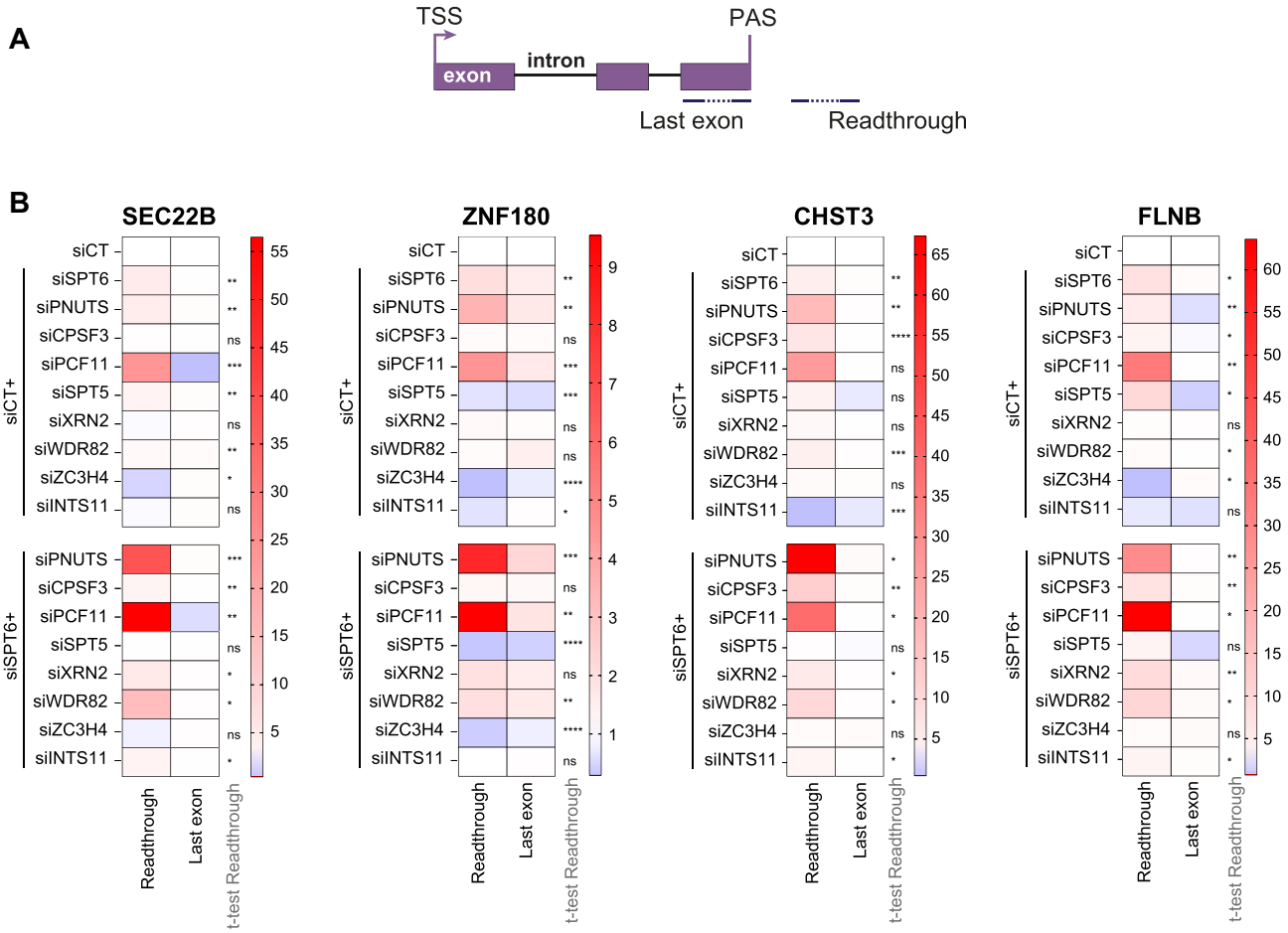



**Figure 2.** Evaluating functional collaboration between SPT6 and multiple termination pathways. (**A**) Amplicons designed to evaluate readthrough transcription at various genes. The last exon amplicon measures full-length gene transcription. (**B**) Heatmap displaying the relative RNA levels assayed by RT-qPCR in five independent experiments using the amplicons designed in panel (A) at the SEC22B, *ZNF180*, *CHST3*, and *FLNB* genes. Values were normalized to the siCT condition arbitrarily set to 1. Statistical significance of readthrough transcripts was assessed using an unpaired *t*-test between each condition and the siCT, with *P*-values indicated as follows: .1234 (ns, nonsignificant), .0332 (*), .0021 (**), .0002 (***), <.0001 (****).

fined based on their epigenetic state in HeLa cells as described in Supplementary Fig. S4A and excluded promoter regions.

Our analysis revealed that PNUTS peaks were predominantly enriched in gene regions (~78% in gene body, promoter and termination regions), with only a few peaks identified in distal regulatory regions (~4% in repressed regions, active and inactive enhancers) (Fig. 4A and Supplementary Fig. S4B). Similar to PCF11, we noted a slight discrepancy in PNUTS enrichment between promoter regions (33%) and termination regions (18%), which led us to further analyze their binding pattern to gene regions.

We divided RefSeq-annotated genes into three categories based on their RNAPII binding profile: Category A, where RNAPII peaks are enriched both in promoter and termination regions ($n = 8180$), Category B where RNAPII is mainly enriched in the promoter region ($n = 4484$) and Category C where RNAPII is predominantly enriched in the termination region ($n = 1238$) (Fig. 4B). We also calculated the pausing index, expression level, and gene length for each category (Fig. 4C–F). Consistent with existing literature [66], these categories reflect three gene and transcriptional states: Category A includes genes that are paused and highly expressed, with a median length of 17 kb. Category B consists of genes with a higher pausing index and lower expression, and Category C comprises genes with a low pausing index and low expression level. Categories B and C have similar median gene lengths of 32 and 33 kb, respectively (Figs 4C–F). Interestingly, the three categories were also reflected in the binding patterns of PNUTS, PCF11, and SPT6 (Fig. 4G), indicating that the binding of these factors depends on the gene transcriptional state. Examples are shown in Supplementary Fig. S4C.

Furthermore, to determine whether PNUTS, PCF11, and SPT6 affect transcription of a specific gene category, we examined the RNAPII ChIP-seq profiles following the depletion of PNUTS (our data), PCF11 [44], and SPT6 [5]. Our analysis revealed that the most significant decrease in RNAPII pausing downstream of the PAS occurred in Category A for all three factors. Depletion of PNUTS and PCF11 also affected Category C, but to a lesser extent (Fig. 4H).

Overall, our data show that while PNUTS, PCF11, and SPT6 display distinct chromatin binding profiles compared to RNAPII, they are notably enriched at RNAPII pause sites, whether in the gene promoter or termination regions. Their effect on RNAPII enrichment at the 3′-end of genes is primarily observed in medium-length, highly expressed genes. As pile-up of RNAPII at the 3′-end of genes is indicative of a slowdown in elongation [44], our results suggest that SPT6, PNUTS, and PCF11 could collaborate to affect transcription termination.

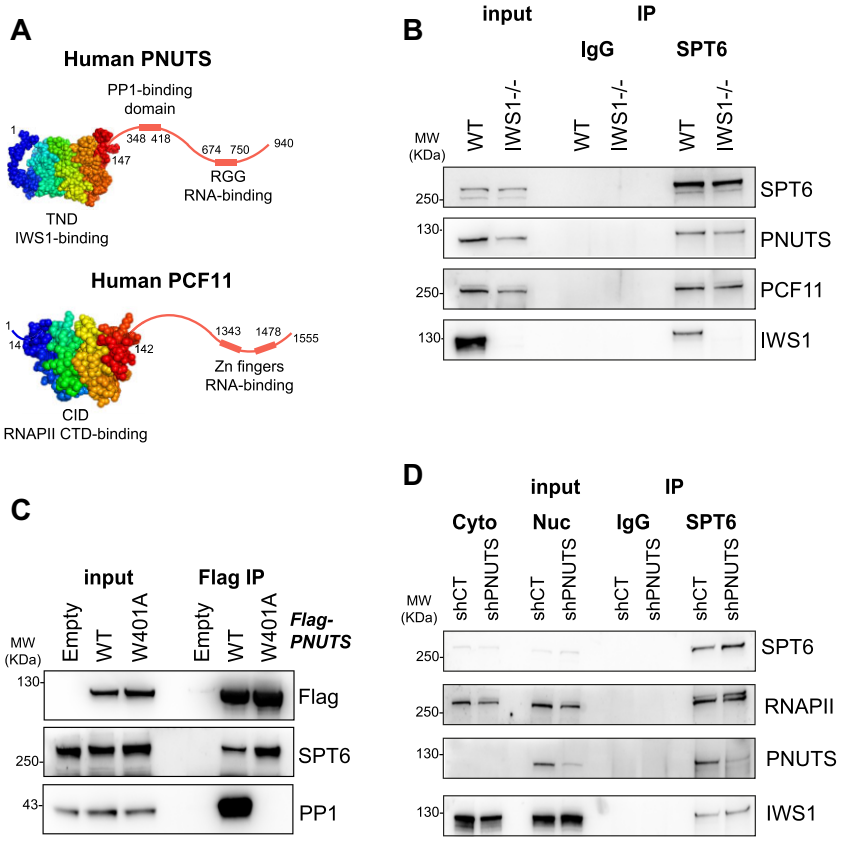



**Figure 3.** SPT6 interacts with IWS1, PNUTS, and PCF11. (**A**) Upper panel: Diagram depicting PNUTS domains: the TND (1–147 amino acids), the PP1-binding domain (348–418 amino acids) and the RNA-binding domain (674–750 amino acids). Lower panel: Diagram depicting PCF11 domains: the CTD-interacting domain (CID) (14–142 amino acids), and the RNA-binding zinc fingers (between 1343 and 1478 amino acids). The illustration denotes N-to-C terminal protein domains. (**B**) Nuclear protein extracts from WT or IWS1-knockout (IWS1-/-) J-Lat A1 cells were used to immunoprecipitate endogenous SPT6. Co-immunoprecipitation of IWS1, PCF11 and PNUTS was assessed ($n = 3$). (**C**) Overexpressed Flag-tagged WT PNUTS (WT) or Flag-tagged PNUTS–W401A (W401A), a mutant for PP1 interaction, were immunoprecipitated using an anti-Flag antibody from HEK293T total protein extracts. SPT6 co-immunoprecipitation was examined for both WT and mutant PNUTS ($n = 3$). (**D**) HeLa cells infected with lentiviruses containing either a control shRNA (shCT) or an shRNA targeting PNUTS (shPNUTS). Cells nuclear fractions were used to immunoprecipitate endogenous SPT6 ($n = 2$). "Cyto" stands for Cytoplasmic fraction, and "Nuc" stands for Nuclear fraction used as input.

## Functional cooperation of SPT6, PNUTS, and PCF11 in transcription termination

Considering the interaction of SPT6 with PNUTS and PCF11, as well as their shared impact on transcription termination, we decided to assess their functional cooperation. PNUTS and PCF11 were depleted either independently or in combination with SPT6 (Supplementary Fig. S5A).

We then analyzed readthrough regions defined previously (Supplementary Fig. S1E). Regions showing a FC >2 ($\log_2$FC > 1) upon protein depletion compared to the control condition were identified as positive readthrough regions. In addition to SPT6, PNUTS, and PCF11 depletion also increased readthrough levels (Supplementary Fig. S5B). However, given the less efficient depletion of PNUTS compared to SPT6 and PCF11, as shown in Supplementary Fig. S5A, our ability to detect readthrough transcripts upon PNUTS depletion is reduced compared to SPT6 or PCF11 depletion (Supplementary Fig. S5B).

We then focused on SPT6 readthrough targets ($n = 1252$). Overall, the combined depletion of SPT6 and PNUTS resulted in increased readthrough levels compared to SPT6 depletion alone (Fig. 5A). To determine if this pattern is consistent across all the studied regions, we clustered them based on their readthrough patterns.

For SPT6 and PNUTS, Cluster I includes readthrough transcripts upregulated upon the combined knockdown of SPT6 and PNUTS compared to SPT6 knockdown alone. Cluster II shows readthrough transcripts downregulated upon the combined knockdown of SPT6 and PNUTS compared to SPT6 knockdown alone, while Cluster III display no substantial change in readthrough transcripts upon individual or combined depletion of SPT6 (Fig. 5B and C). Notably, 61% of SPT6 readthrough transcripts were found in Cluster I (769 out of 1252) (Fig. 5C). Moreover, in 82% of the readthrough regions in Cluster I (634 out of 769), the observed increased in the combined knockdown condition is nonsynergistic compared to the individual depletions (Bliss independence test, Supplementary Fig. S5C, left panel), suggesting that, in half of the studied regions, PNUTS and SPT6 operate within the same termination pathway.

PCF11 depletion had a more pronounced effect on increasing readthrough transcripts (Supplementary Fig. S5B). Similar to the combined depletion of SPT6 and PNUTS, co-depletion of PCF11 and SPT6 further increased readthrough transcript levels compared to their individual depletions (Fig. 5D). The same 1252 readthrough regions were then similarly clustered based on transcript levels upon SPT6 and PCF11 co-depletion. Approximately 77% of the readthrough transcripts

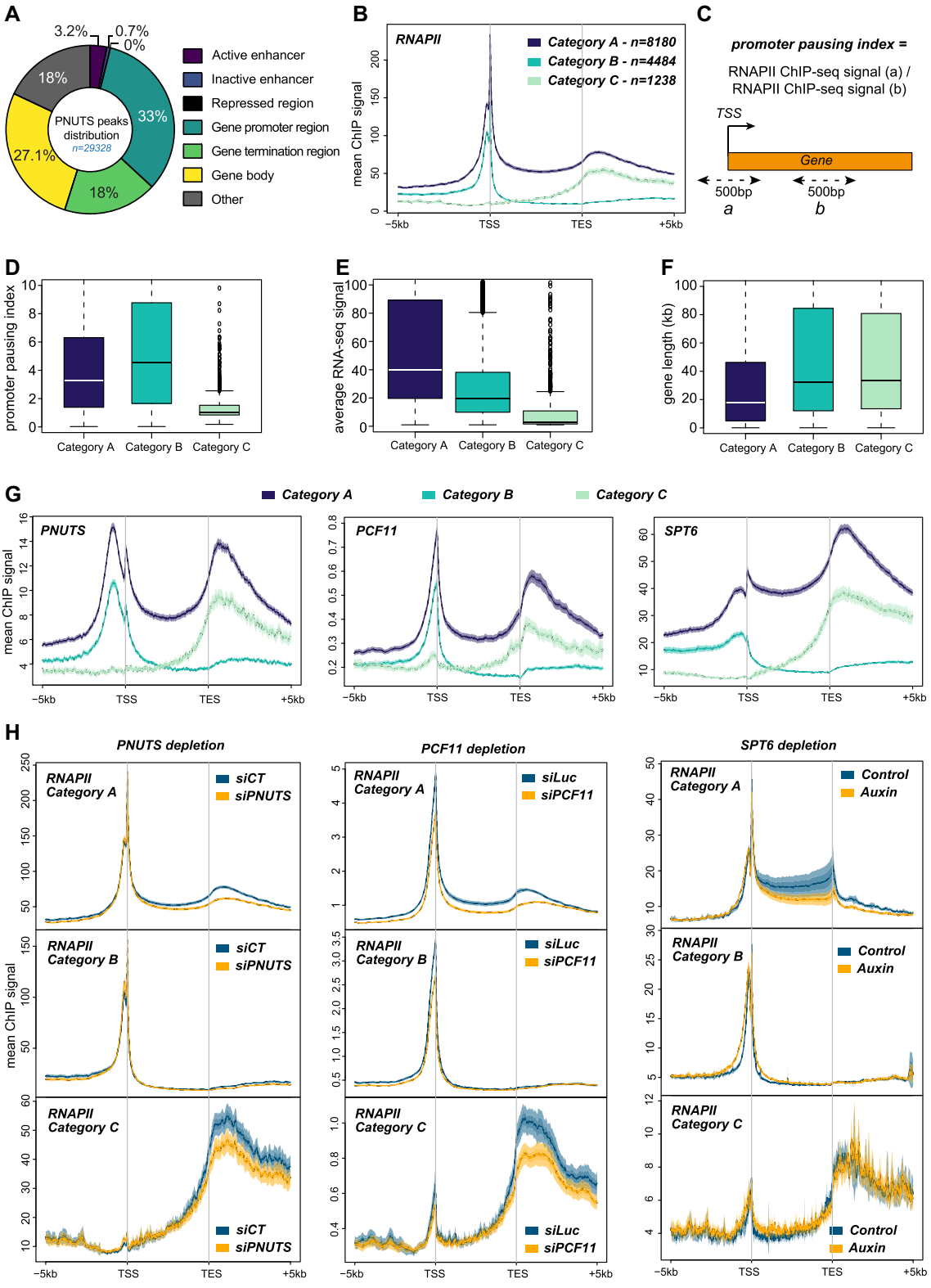



**Figure 4.** PNUTS, PCF11, and SPT6 chromatin distribution and transcriptional effect according to gene categories. (**A**) PNUTS ChIP-seq peaks ($n = 29\,328$) distribution was assessed at active and inactive enhancers, repressed regions, and RefSeq-gene promoter (TSS ± 1.5 kb), body and termination (TES ± 1.5 kb) regions. (**B**) RNAPII displaying promoter and termination region enrichment (Category A), promoter-only enrichment (Category B) or termination-region-only enrichment (Category C). The plots represent the mean RNAPII ChIP-seq signals. (**C**) Promoter-pausing index defined by the ratio between RNAPII signal in a 500 bp window surrounding the TSS and RNAPII signal in a 500 bp window starting 750 bp downstream of the TSS (regions a and b in the scheme, respectively). (**D**) Promoter-pausing index for the three gene categories defined in panel (B). (**E**) Average RNA-seq signal calculated in the promoter region (TSS ± 1.5 kb) of genes belonging to Categories A, B, and C. (**F**) Length (in kilobases) of genes belonging to Categories A, B, and C. (**G**) PNUTS, PCF11, and SPT6 binding profiles depending on gene category. Mean ChIP-seq signals are plotted. (**H**) RNAPII binding profile upon PNUTS depletion (left panel, our data), PCF11 depletion [44] and SPT6 depletion [5], depending on the gene category. Figure 4B, G, and H: The solid line represents the mean. The dark area represents the standard error and light area represents the 95% confidence interval.

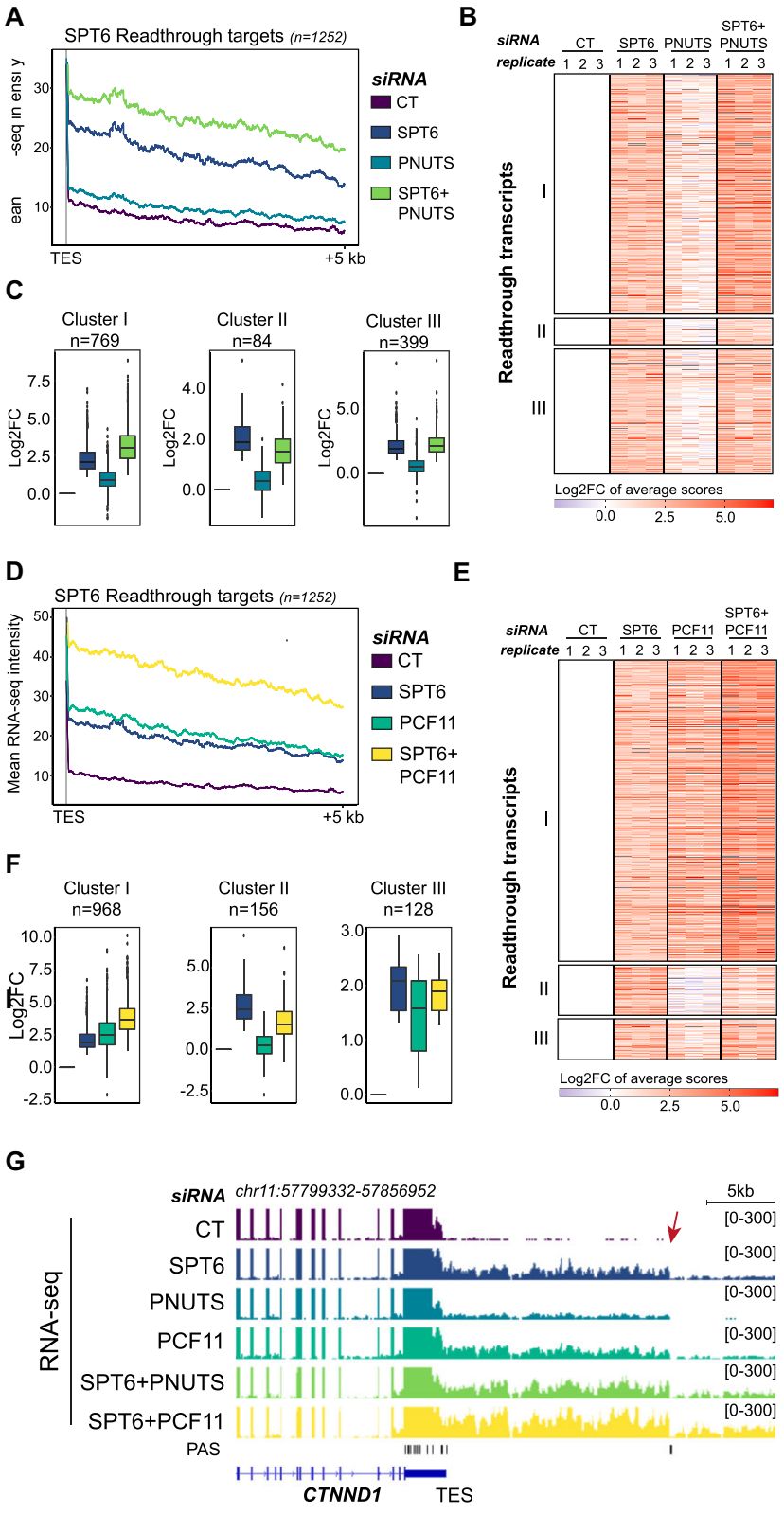



**Figure 5.** Differential readthrough transcripts upon depletion of SPT6, PNUTS, and PCF11. (**A**) Mean RNA-seq signal intensity shown at SPT6 readthrough targets (TES + 5 kb) in the different conditions. (**B**) Heatmap displaying differentially expressed readthrough transcripts in HeLa cells following SPT6 and/or PNUTS depletion. Readthrough transcripts with a $\log_2$FC >1 in the siSPT6 condition compared to the control condition were considered positive. The readthrough regions were identified as described in Supplementary Fig. S1F. The heatmap plots the average log2FC scores (siRNA/siCT) of these regions. Three distinct clusters were defined to characterize the varying responses to SPT6 and/or PNUTS depletion. The analysis was conducted using three independent biological replicates. (**C**) Box-plots showing the median $\log_2$FC of readthrough transcripts relative to the control condition for the three clusters shown in panel (B). (**D–F**) Same as in panels (A)–(C) but for SPT6 and/or PCF11 depletion. (**G**) *CTNND1* gene locus featuring RNA-seq data (positive strand) following the use of the indicated siRNAs. The lower panel (PAS) displays the locations of human PASs as defined by Zhang *et al.* [82]. The arrow highlights readthrough transcription. The scale is shown at the top right corner of the figure.

(968 out of 1252) were found in Cluster I (Fig. 5E and F). Although the simultaneous depletion of SPT6 and PCF11 had a greater impact than their individual depletions in this cluster, it produced a synergistic effect at only 5% of the regions (Supplementary Fig. S5C, right panel), indicating that SPT6 and PCF11 also function within the same termination pathway. Clusters II and III exhibit either a downregulation or no significant change, respectively, in readthrough transcripts following the combined knockdown of SPT6 and PCF11 compared to the knockdown of SPT6 alone.

Notably, 65% of the genes undergoing increased readthrough transcripts upon SPT6, PNUTS or PCF11 depletion belong to Category A genes, i.e. highly expressed genes with RNAPII ChIP-seq peaks on both gene ends (Supplementary Fig. S5D). Category C genes were excluded from our study, as they present low RNAPII at their promoter region.

The increase in readthrough transcripts observed in total RNA-seq upon SPT6, PNUTS, or PCF11 depletion was further corroborated by analyzing nascent RNA levels using nuclear run-on experiments, which showed similar fold changes in readthrough levels, but not full-length mRNA levels, between total nuclear RNAs and nascent RNAs (Supplementary Fig. S5E). This suggests that the impact of these factors on readthrough levels occurs primarily at the transcriptional level.

PCF11 is known to favor the use of proximal polyadenylation sites [44]. Remarkably, inspection of individual genes showed that readthrough transcripts usually end with a PAS (Fig. 5G and Supplementary Fig. S5F). Furthermore, analysis of published 3′-RNAseq data [18] showed a significant increase in the polyadenylated 3′-end signal upon SPT6 depletion at the SPT6 readthrough regions defined in our study (Supplementary Fig. S5G). These findings suggest an alternative polyadenylation mechanism leading to a downstream-shifted termination in the absence of either SPT6, PNUTS, or PCF11.

## PNUTS assists SPT6 in PROMPTs restriction, while PCF11 is essential for PROMPTs activation in the absence of SPT6

In addition to regulating transcription termination at gene 3′-ends, SPT6 has also been implicated in regulating antisense transcription at gene promoters [18]. In order to evaluate whether SPT6, PCF11, and PNUTS regulate PROMPTs and readthrough transcripts in the same manner, we first defined PROMPT regions for RefSeq-annotated genes as 5 kb-regions upstream of the gene TSS. We selected regions that (i) do not overlap with annotated genes on the same strand, (ii) correspond to genes with an RNAPII ChIP-seq peak within their promoter region (TSS ± 1.5 kb) (genes of Category A and B), and (iii) have an average RNA-seq signal higher than 1 RPM in at least one condition (Supplementary Fig. S6A). This process resulted in the identification of 4131 PROMPT regions.

For each factor, we then excluded the PROMPT regions corresponding to genes that were positively regulated by the same factor on the opposite strand, to eliminate PROMPTs arising from readthrough transcripts of downstream genes (see Supplementary Fig. S6B for example).

In line with the previous study by Nojima *et al.* [18], the effect of SPT6 depletion on PROMPTs was noticeable (Supplementary Fig. S7A). Similar to what we observed for readthrough transcripts, IWS1 depletion had minimal impact on PROMPT levels. However, unlike readthrough transcripts, the effect of PCF11 or PNUTS on PROMPTs was mild (Supplementary Fig. S7A). Additionally, 77% of the PROMPTs affected by SPT6, PNUTS, or PCF11 belonged to gene Category A, defined earlier (Supplementary Fig. S7B).

PROMPTs having a FC >2 ($\log_2$FC > 1) upon SPT6 depletion compared to the control condition ($n = 1111$) were studied. Overall, SPT6 and PNUTS co-depletion had a greater impact on PROMPTs than SPT6 depletion alone (Fig. 6A). These PROMPT regions were then classified into three clusters, similar to readthrough transcripts in Fig. 5. For most genes (Cluster I, 680 out of 1111), the combined depletion of SPT6 and PNUTS had a stronger effect on PROMPTs accumulation than SPT6 depletion alone (Fig. 6B and C). Notably, around 80% of Cluster I genes (538 out of 680) showed no synergistic effect upon SPT6 and PNUTS double knockdown (Bliss independence test, Supplementary Fig. S7C), suggesting that, in half of the studied regions, PNUTS reinforces SPT6 in restricting PROMPTs within the same pathway. An example is shown for the PXDN gene (Fig. 6D), with amplification of this gene's first intron used as a control (Supplementary Fig. S7D). Nevertheless, depleting both SPT6 and PNUTS can also have the opposite effect and reduce PROMPT expression (Cluster II, Fig. 6B and C).

In the case of PCF11, we observed an overall robust decrease in the effect of SPT6 on PROMPT accumulation upon combined depletion of PCF11 and SPT6 (Fig. 6E). This pattern was noted in 58% of all SPT6 PROMPTs (Cluster II, 648 out of 1111) (Fig. 6F and G). An example is shown for the GGCT gene (Cluster II) (Fig. 6H and Supplementary Fig. S7E). This suggests that PCF11 is necessary for PROMPT accumulation in the absence of an SPT6-dependent repression. Yet, for a smaller number of PROMPTs (Cluster I), an increase in their accumulation was observed upon the combined depletion of PCF11 and SPT6 (Fig. 6F and G). Of note, the observed effects are specific to PROMPTs, as we did not detect similar changes in intragenic transcription [Supplementary Fig. S7D and E (first intron) and Supplementary Fig. S8A and B].

Similar to readthrough transcripts, the effect of SPT6 and PNUTS on PROMPT transcripts, but not intragenic transcripts, is observed at the transcriptional level, as demonstrated by nuclear run-on experiments (Supplementary Fig. S8C). In contrast, PCF11 has little to no effect on PROMPT transcripts, even at the transcriptional level.

Additionally, similar to readthrough transcripts, we observed that PROMPTs end at a polyadenylation site for several genes (see examples in Fig. 6I and Supplementary Fig. S7F), and that the polyadenylated 3′-RNAseq signal significantly increases upon SPT6 depletion at SPT6 PROMPTs (Supplementary Fig. S8D) indicating a downstream-shifted termination as previously proposed [18].

Overall, our results highlight a complex network involving SPT6, PNUTS, and PCF11 in regulating transcription termination. At the 3′-end of genes, SPT6 strongly restricts readthrough accumulation in collaboration with PCF11 and PNUTS for most of the studied genes. Meanwhile, SPT6 and PNUTS jointly inhibit hundreds of PROMPTs. However, in this case, PCF11 appears to be necessary for the expression of PROMPTs in the absence of an SPT6:PNUTS repression (Fig. 7).

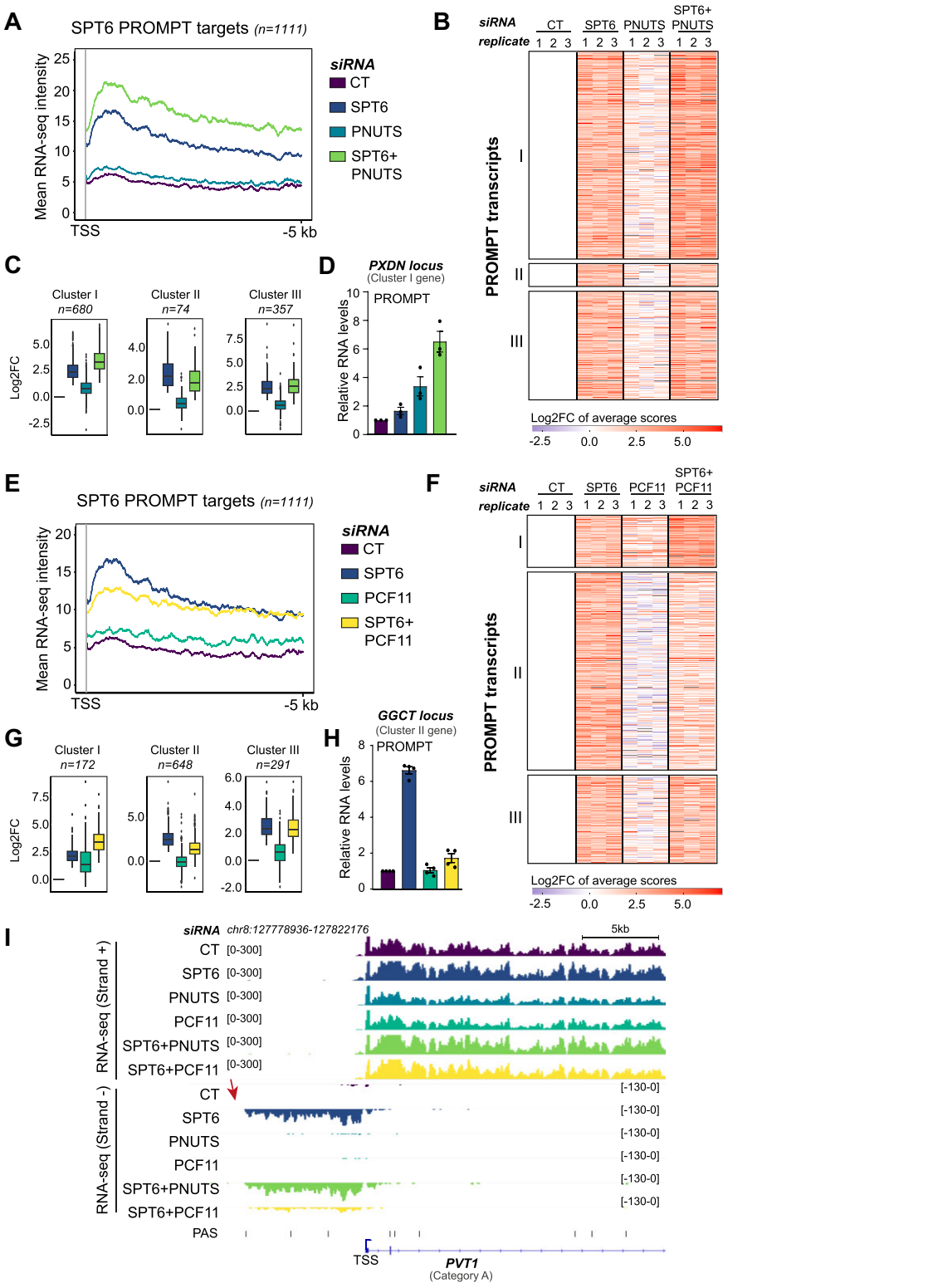



**Figure 6.** Differential PROMPTs regulation upon depletion of SPT6, PNUTS, and PCF11. (**A**) Mean RNA-seq signal intensity shown at SPT6 PROMPT targets (TSS, -5 kb) in the different conditions. (**B**) Heatmap displaying differentially expressed PROMPTs in HeLa cells following SPT6 and/or PNUTS depletion. PROMPTs with a $\log_2$FC >1 in the siSPT6 condition compared to the control condition were considered positive. PROMPT regions were identified as described in Supplementary Fig. S6A. The heatmap plots the average $\log_2$FC scores (siRNA/siCT) of these regions. Three distinct clusters were defined to characterize the varying responses to SPT6 and/or PNUTS depletion. The analysis was conducted using three independent biological replicates. (**C**) Box-plots showing the median $\log_2$FC of PROMPTs relative to the control condition for the three clusters shown in panel (B). (**D**) PXDN PROMPT levels were assayed by RT-qPCR in three independent experiments. Values were normalized to the siCT condition arbitrarily set to 1. (**E**–**G**) Same as in panels (A)–(C) but for SPT6 and/or PCF11 depletion. (**H**) Same as in panel (D) but for the GGCT gene. (**I**) *PVT1* gene locus featuring RNA-seq data (positive strand is in positive values and negative strand in negative values) following the use of the indicated siRNAs. The lower panel (PAS) displays the locations of human PASs as defined by Zhang *et al.* [82]. The arrow highlights PROMPTs. The scale is shown at the top right corner of the figure.

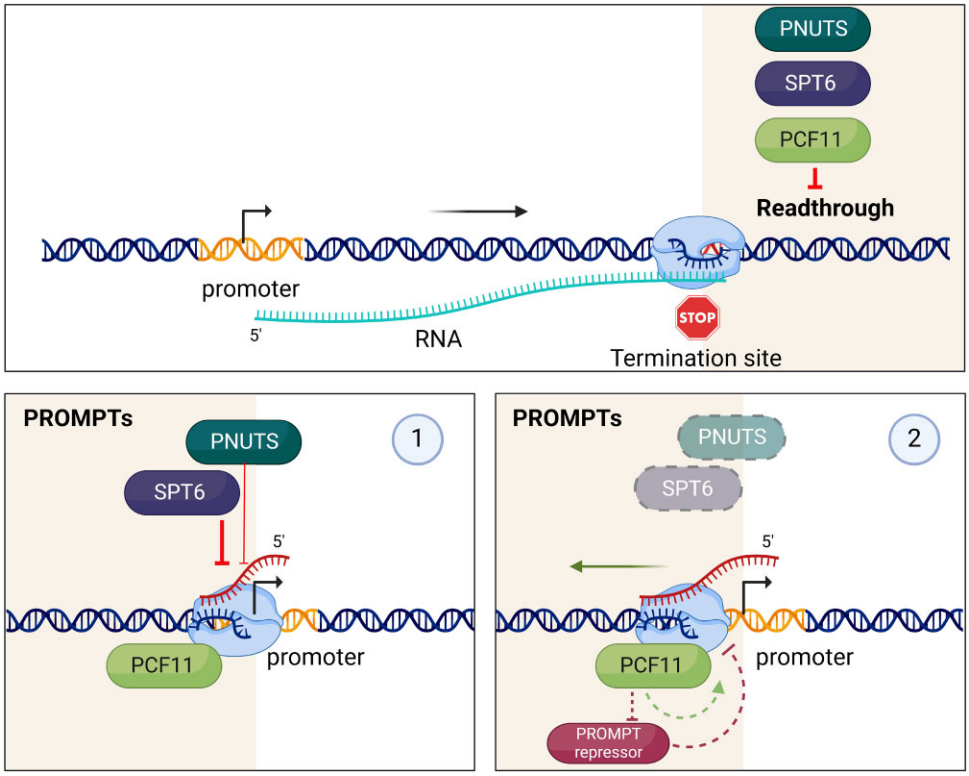



**Figure 7.** Model for the regulation of transcription termination by SPT6, PNUTS, and PCF11 at promoters and gene ends. At transcription termination sites, SPT6, PNUTS, and PCF11 prevent readthrough transcription. At gene promoters, SPT6 and PNUTS establish an initial block to limit PROMPTs. When this primary block is removed, PCF11 promotes PROMPTs, either directly or by inhibiting the recruitment of another PROMPT repressor. Created in BioRender. Bejjani, F. (2025) https://BioRender.com/f07e489.

## Discussion

In this study, we investigated the roles of SPT6 and its interaction partners in transcription termination. We initially assessed the cooperation between SPT6 and its closest partner, IWS1. While our results show a very strong correlation between the chromatin occupancy of IWS1, SPT6, and RNAPII, we observed that the IWS1/RNAPII ratio decreases at the 3′-end of genes when compared to that of SPT6/RNAPII. Similarly, distinct occupancy profiles at the 3′-end of genes are also observed in yeast for Spt6 and Spn1 [36]. However, opposite effects were observed, as Spn1 occupancy increases much more than Spt6 towards the 3′-end of the gene and peaks just upstream and downstream of the poly(A) site [36]. Nevertheless, taken together these studies suggest that SPT6 and IWS1/Spn1 could have separate functions in regulating gene expression.

While SPT6 knockdown regulated the expression of thousands of genes, very few genes were affected upon IWS1 knockdown in our experiments. Interestingly, even if depletion of Spn1 in yeast resulted in a more widespread decrease of transcripts, affecting more than 1500 genes, the effects were still mild compared to SPT6 depletion [16, 34]. While in our experiments we cannot exclude the possibility of a difference in depletion efficacy between IWS1 and SPT6, this discrepancy suggests that another cellular factor may functionally compensate for IWS1 function in human cells. In yeast, an SPT6 mutation impaired for Spn1 interaction leads to an increase in the recruitment of FACT to chromatin [35]. FACT is a histone chaperone that associates with elongating RNAPII and facilitates transcription through nucleosomes [26]. As FACT was shown to be highly enriched at active genes in cancer cells [67], it is plausible that the requirement for IWS1 is compensated by elevated levels of FACT in our IWS1-knockdown experiments using cancer cell lines. Another possibility is that IWS1 is not required for gene expression under control conditions in human cells but is rather involved in the transcriptional response to external signals. Indeed, IWS1 is the substrate of several kinases including CDK9 and AKTs [68, 69], and it was proposed that IWS1 phosphorylation by AKT kinases is required for its interaction with SETD2 and favors the recruitment of SETD2 to the RNAPII elongation complex [68]. Further experiments will be needed to determine whether IWS1 is required to regulate gene expression in response to external stimuli in human cells.

IWS1 connects several factors involved in transcription through specific TND–TIM interactions [38]. IWS1 TND interacts with TIMs in the N-terminal region of SPT6, while each of its three TIM domains can selectively bind to TND-containing factors, including PNUTS [37–39]. It was proposed that, via its TND–TIM interfaces, IWS1 could bridge SPT6 and PNUTS [38]. However, our co-immunoprecipitation results show that SPT6 still interacts with PNUTS in IWS1 knockout cells, suggesting that IWS1 is not absolutely required for this interaction. It remains unclear whether SPT6 TIMs could directly bind to PNUTS TND *in vivo* or whether their interaction involves another factor.

The limited ChIP-seq data available for PNUTS prompted us to characterize the chromatin binding profile of the endogenous protein. Our ChIP-seq analysis revealed that, in our model, PNUTS peaks were predominantly located in gene re-

gions. Despite previous reports suggesting PNUTS involvement in eRNA transcription [70], we observed very few, low-abundance peaks in enhancer regions. This suggests that PNUTS binding to enhancers might be highly dynamic and technically challenging to detect by ChIP, or that its effect on enhancers is indirect. In line with studies on both epitope-tagged PNUTS [48] and endogenous PNUTS in U2OS cells [65], we detected significant enrichment of PNUTS at TSS and TES. However, our data showed similar abundance of PNUTS peaks at TES and TSS on average. Furthermore, the binding profile of PNUTS varied depending on gene characteristics such as expression level, pausing index, and gene length. Notably, we observed the highest PNUTS enrichment in highly transcribed genes, as reported in earlier studies [58, 65].

SPT6 depletion resulted in increased readthrough transcripts at ~1250 RefSeq-annotated genes. While our results stand in contrast with studies in yeast which suggest that the loss of SPT6 leads to earlier transcription termination [71, 72], they confirm a recent study in mammalian cells showing that acute SPT6 depletion increases readthrough transcripts [5]. We found that PCF11 and PNUTS, two factors involved in transcription termination, strongly cooperate with SPT6 to regulate transcription readthrough. In agreement with previous analysis showing that depletion of SPT6 reduced the association of transcription termination factors, including PNUTS and PCF11, with RNAPII [5, 40], we show that SPT6 interacts with both factors. PNUTS targets the PP1 protein phosphatase to SPT5 and promotes its dephosphorylation to slow down transcription at the 3′-end of genes [48]. We speculate that by recruiting PNUTS to the RNAPII complex, SPT6 could regulate the dephosphorylation of SPT5 and facilitate the deceleration of RNAPII at poly(A) sites of termination.

We noticed that similarly to what we observed for PCF11, the depletion of SPT6 and/or PNUTS also led to a shift towards the use of the most distal PAS at several annotated genes. Recent studies show that PCF11 is a regulatory factor of the CPA complex which was shown to affect alternative CPA (APA) by favoring the use of proximal PAS at a subset of pc genes [43, 44, 73]. Thus, our results suggest that SPT6 cooperates with PCF11 and PNUTS to enforce transcription termination at proximal PAS. It also indicates that in the absence of SPT6, distal PAS-dependent transcription termination can still occur. Thus, rather than having a global effect on transcription termination, we propose that SPT6 could regulate APA at a subset of cellular genes by promoting the use of the proximal PAS through its interaction with PCF11 and PNUTS.

Human gene promoters are generally considered to be inherently bidirectional, though not all RNAPII promoters exhibit such bidirectionality. In fact, various transcription elongation factors, nucleosome remodelers and RNA surveillance factors have been identified to limit promoter-antisense transcription [18, 70, 74–76]. Here, we report a significant effect of SPT6 on PROMPT restriction. This observation aligns with DeGennaro *et al.*'s study [13] in *Schizosaccharomyces pombe*, where the loss of Spt6 induced significant chromatin structure alterations and increased antisense transcription in over 70% of genes. Additionally, in humans, SPT6 depletion was found to increase lncRNA transcription, including PROMPTs, which was accompanied by elevated H3K36me3 histone mark in the antisense direction [18].

PROMPTs were initially characterized in 2008 by Preker *et al.* [74] as a class of short RNAs, ranging from 0.5 to 2.5 kb, transcribed by the RNAPII upstream of annotated gene promoters. These RNAs are polyadenylated but highly unstable, as indicated by their high enrichment upon depletion of exosome subunits. Indeed, SPT6 has been shown to co-localize with exosome subunit Rrp6 in *Drosophila* [77], without direct binding in pulldown assays [37], implying an indirect association between SPT6 and Rrp6. Our findings also reveal that PNUTS depletion upregulates PROMPT levels, albeit less effectively than SPT6, likely due to less efficient siRNA-mediated knockdown of PNUTS expression in our assays. Interestingly, PNUTS was recently described to associate with the restrictor complex (WDR82-ZC3H4) and ARS2, which in turn recruits the nuclear exosome to terminate unproductive transcription [50, 78, 79]. Thus, we propose that SPT6, by interacting with PNUTS, recruits the RNA exosome to terminate antisense transcription.

In contrast, we show that PCF11 depletion alone does not greatly impact PROMPT levels. However, for hundreds of genes, in the absence of SPT6, PROMPT accumulation following PCF11 depletion is restricted. This implies that PCF11 is either necessary for PROMPT transcription/stabilization, or competes with another PROMPT-restricting factor for RNAPII binding. In yeast, the CID-containing Nrd1 protein was shown to regulate cryptic unstable transcripts through the yeast nuclear exosome and TRAMP components [80]. The termination potential of the Nrd1 protein was shown to be affected by PCF11 binding to RNAPII [81].

In this study, we used size factor normalization for RNA-seq data analysis, a widely used and robust method. While this approach has limitations, particularly in datasets with a large proportion of differentially expressed genes, we assessed its applicability by recalculating size factors with and without differentially expressed genes in our datasets. The results showed that the size factors remained highly consistent across conditions, confirming that our normalization approach is stable and primarily influenced by the majority of unaffected genes. Additionally, the size factor was consistently higher in the siSPT6 condition than in the control condition, indicating that any potential bias introduced by normalization would lead to an underestimation of effects rather than overestimation.

Similarly, the use of total RNA-seq rather than nascent RNA-seq may also contribute to an underestimation of transcriptional effects. For highly expressed genes, such as SEC22B (Supplementary Fig. S9), readthrough transcripts can be difficult to visualize on a linear scale. However, these readthrough transcripts are easily detectable when zooming into the region or when using RT-qPCR and NRO experiments. In contrast, for moderately expressed genes (e.g. CTNND1 and DHX34; Supplementary Fig. S9), the effects are more readily detectable. Notably, since PROMPTs are transcribed from the opposite strand relative to mRNAs, they are less masked by the high abundance of mRNA transcripts in our datasets (e.g. DOCK5 and PVT1, Supplementary Fig. S9). Furthermore, the effects observed across all conditions have been validated using complementary methods, including nuclear run-on experiments, which specifically measure nascent transcription and provide strong evidence that the changes we report reflect genuine transcriptional regulation by SPT6, PNUTS, and PCF11.

Overall, our data highlight a crucial role for SPT6 in regulating transcription termination at the 3′-end of genes, in collaboration with PNUTS and PCF11. At the 5′-end of genes, we propose a model where SPT6 and PNUTS establish a primary

checkpoint for PROMPT regulation. Following the release of this initial checkpoint, PROMPT regulation is further managed by PCF11 and other potential termination factors.

## Acknowledgements

We thank all members of the Host-Virus Interactions Lab at Institut Cochin for their helpful discussions, Melissa Ait Said for setting up the tools used for PCF11 studies, and Monica Naughtin for cloning the shPNUTS vector. Special thanks go to Rosemary Kiernan and Isabelle Jariel-Encontre for their critical reading of the manuscript, and to Roy Matkovic for sharing the cell fractionation protocol and for scientific discussions. We are grateful to Eric Verdin for the gift of the J-Lat A1 cell line, Dr David Skalnik for the gift of the PNUTS plasmid, and Dr Hiroshi Handa for the gift of the SPT6 plasmid. We thank Juliette Hamroune, Lucie Adoux, and Benjamin SaintPierre at the Institut Cochin Genom'IC core facility, and Olivier Cuvier at CBI Toulouse for their expertise. The following reagent was obtained through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH: J-Lat Tat-GFP Cells (J-Lat A1). The SE team is a member of the Groupement de recherche DYNAVIR (GDR2194): "Dynamics of interactions between viral and cellular chromatins", supported by the CNRS.

*Author contributions*: Fabienne Bejjani (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Validation, Visualization, Writing—original draft), Emmanuel Ségéral (Investigation), Kevin Mosca (Data curation, Formal analysis), Adriana Lecourieux (Data curation, Formal analysis), Meriem Hamoudi (Investigation), May Bakail (Investigation), Stéphane Emiliani (Conceptualization, Formal analysis, Funding acquisition, Project administration, Methodology, Supervision, Writing—original draft)

## Supplementary data

Supplementary data is available at NAR online.

## Conflict of interest

None declared.

## Funding

This work was supported by the "Agence nationale de recherches sur le SIDA et les hépatites virales - maladies infectieuses émergentes" (ANRS-MIE) and Sidaction to S.E. F.B. was supported by post-doctoral fellowships from ANRS-MIE and Sidaction. M.H. and M.B. were supported by a post-doctoral fellowship from ANRS-MIE. Funding to pay the Open Access publication charges for this article was provided by laboratory funds.

## Data availability

ChIP-seq and RNA-seq files were deposited in the Gene Expression Omnibus database under accession GSE272847 and GSE272848, respectively.

## References

1. Miller CLW, Warner JL, Winston F. Insights into Spt6: a histone chaperone that functions in transcription, DNA replication, and genome stability. *Trends Genet* 2023;**39**:858–72. https://doi.org/10.1016/j.tig.2023.06.008
2. Sdano MA, Fulcher JM, Palani S *et al.* A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *eLife* 2017;**6**:e28723. https://doi.org/10.7554/eLife.28723
3. Endoh M, Zhu W, Hasegawa J *et al.* Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro. *Mol Cell Biol* 2004;**24**:3324–36. https://doi.org/10.1128/MCB.24.8.3324-3336.2004
4. Ardehali MB, Yao J, Adelman K *et al.* Spt6 enhances the elongation rate of RNA polymerase II in vivo. *EMBO J* 2009;**28**:1067–77. https://doi.org/10.1038/emboj.2009.56
5. Narain A, Bhandare P, Adhikari B *et al.* Targeted protein degradation reveals a direct role of SPT6 in RNAPII elongation and termination. *Mol Cell* 2021;**81**:3110–27. https://doi.org/10.1016/j.molcel.2021.06.016
6. Kaplan CD, Morris JR, Wu C *et al.* Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in D. melanogaster. *Genes Dev* 2000;**14**:2623–34. https://doi.org/10.1101/gad.831900
7. Žumer K, Maier KC, Farnung L *et al.* Two distinct mechanisms of RNA polymerase II elongation stimulation *in vivo*. *Mol Cell* 2021;**81**:3096–109. https://doi.org/10.1016/j.molcel.2021.05.028
8. Andrulis ED, Guzman E, Doring P *et al.* High-resolution localization of drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. *Genes Dev* 2000;**14**:2635–49. https://doi.org/10.1101/gad.844200
9. Bortvin A, Winston F. Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* 1996;**272**:1473–6. https://doi.org/10.1126/science.272.5267.1473
10. McCullough L, Connell Z, Petersen C *et al.* The abundant histone chaperones Spt6 and FACT collaborate to assemble, inspect, and maintain chromatin structure in *Saccharomyces cerevisiae*. *Genetics* 2015;**201**:1031–45. https://doi.org/10.1534/genetics.115.180794
11. McDonald SM, Close D, Xin H *et al.* Structure and biological importance of the Spn1–Spt6 interaction, and its regulatory role in nucleosome binding. *Mol Cell* 2010;**40**:725–35. https://doi.org/10.1016/j.molcel.2010.11.014
12. Kasiliauskaite A, Kubicek K, Klumpler T *et al.* Cooperation between intrinsically disordered and ordered regions of Spt6 regulates nucleosome and pol II CTD binding, and nucleosome assembly. *Nucleic Acids Res* 2022;**50**:5961–73. https://doi.org/10.1093/nar/gkac451
13. DeGennaro CM, Alver BH, Marguerat S *et al.* Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in fission yeast. *Mol Cell Biol* 2013;**33**:4779–92. https://doi.org/10.1128/MCB.01068-13
14. Ivanovska I, Jacques PE, Rando OJ *et al.* Control of chromatin structure by spt6: different consequences in coding and regulatory regions. *Mol Cell Biol* 2011;**31**:531–41. https://doi.org/10.1128/MCB.01068-10
15. Jeronimo C, Watanabe S, Kaplan CD *et al.* The histone chaperones FACT and Spt6 restrict H2A.Z from intragenic locations. *Mol Cell* 2015;**58**:1113–23. https://doi.org/10.1016/j.molcel.2015.03.030
16. Doris SM, Chuang J, Viktorovskaya O *et al.* Spt6 Is required for the fidelity of promoter selection. *Mol Cell* 2018;**72**:687–99. https://doi.org/10.1016/j.molcel.2018.09.005
17. Bakel H, Tsui K, Gebbia M *et al.* A compendium of nucleosome and transcript profiles reveals determinants of chromatin architecture and transcription. *PLoS Genet* 2013;**9**:e1003479. https://doi.org/10.1371/journal.pgen.1003479
18. Nojima T, Tellier M, Foxwell J *et al.* Deregulated expression of mammalian lncRNA through loss of SPT6 induces R-loop

formation, replication stress, and cellular senescence. *Mol Cell* 2018;**72**:970–84. https://doi.org/10.1016/j.molcel.2018.10.011
19. Oqani RK, Lin T, Lee JE *et al.* Iws1 and Spt6 regulate trimethylation of histone H3 on lysine 36 through akt signaling and are essential for mouse embryonic genome activation. *Sci Rep* 2019;**9**:3831. https://doi.org/10.1038/s41598-019-40358-3
20. Youdell ML, Kizer KO, Kisseleva-Romanova E *et al.* Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. *Mol Cell Biol* 2008;**28**:4915–26. https://doi.org/10.1128/MCB.00001-08
21. Yoh SM, Lucas JS, Jones KA. The Iws1:spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev* 2008;**22**:3422–34. https://doi.org/10.1101/gad.1720008
22. Gopalakrishnan R, Marr SK, Kingston RE *et al.* A conserved genetic interaction between Spt6 and Set2 regulates H3K36 methylation. *Nucleic Acids Res* 2019;**47**:3888–903. https://doi.org/10.1093/nar/gkz119
23. Carrozza MJ, Li B, Florens L *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 2005;**123**:581–92. https://doi.org/10.1016/j.cell.2005.10.023
24. Vos SM, Farnung L, Boehning M *et al.* Structure of activated transcription complex pol II-DSIF-PAF-SPT6. *Nature* 2018;**560**:607–12. https://doi.org/10.1038/s41586-018-0440-4
25. Farnung L, Ochmann M, Garg G *et al.* Structure of a backtracked hexasomal intermediate of nucleosome transcription. *Mol Cell* 2022;**82**:3126–34. https://doi.org/10.1016/j.molcel.2022.06.027
26. Ehara H, Kujirai T, Shirouzu M *et al.* Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT. *Science* 2022;**377**:eabp9466. https://doi.org/10.1126/science.abp9466
27. Cheung V, Chua G, Batada NN *et al.* Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. *PLoS Biol* 2008;**6**:e277. https://doi.org/10.1371/journal.pbio.0060277
28. Kaplan CD, Laprade L, Winston F. Transcription elongation factors repress transcription initiation from cryptic sites. *Science* 2003;**301**:1096–9. https://doi.org/10.1126/science.1087374
29. Lindstrom DL, Squazzo SL, Muster N *et al.* Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Mol Cell Biol* 2003;**23**:1368–78. https://doi.org/10.1128/MCB.23.4.1368-1378.2003
30. Krogan NJ, Kim M, Ahn SH *et al.* RNA polymerase II elongation factors of *Saccharomyces cerevisiae*: a targeted proteomics approach. *Mol Cell Biol* 2002;**22**:6979–92. https://doi.org/10.1128/MCB.22.20.6979-6992.2002
31. Fischbeck JA, Kraemer SM, Stargell LA. SPN1, a conserved gene identified by suppression of a postrecruitment-defective yeast TATA-binding protein mutant. *Genetics* 2002;**162**:1605–16. https://doi.org/10.1093/genetics/162.4.1605
32. Diebold M-L, Koch M, Loeliger E *et al.* The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26. *EMBO J* 2010;**29**:3979–91. https://doi.org/10.1038/emboj.2010.272
33. Li S, Edwards G, Radebaugh CA *et al.* Spn1 and its dynamic interactions with Spt6, histones and nucleosomes. *J Mol Biol* 2022;**434**:167630. https://doi.org/10.1016/j.jmb.2022.167630
34. Reim NI, Chuang J, Jain D *et al.* The conserved elongation factor Spn1 is required for normal transcription, histone modifications, and splicing in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 2020;**48**:10241–58. https://doi.org/10.1093/nar/gkaa745
35. Viktorovskaya O, Chuang J, Jain D *et al.* Essential histone chaperones collaborate to regulate transcription and chromatin integrity. *Genes Dev* 2021;**35**:698–712. https://doi.org/10.1101/gad.348431.121
36. Mayer A, Lidschreiber M, Siebert M *et al.* Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* 2010;**17**:1272–8. https://doi.org/10.1038/nsmb.1903
37. Yoh SM, Cho H, Pickle L *et al.* The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev* 2007;**21**:160–74. https://doi.org/10.1101/gad.1503107
38. Cermakova K, Demeulemeester J, Lux V *et al.* A ubiquitous disordered protein interaction module orchestrates transcription elongation. *Science* 2021;**374**:1113–21. https://doi.org/10.1126/science.abe2913
39. Gérard A, Ségéral E, Naughtin M *et al.* The integrase cofactor LEDGF/p75 associates with Iws1 and Spt6 for postintegration silencing of HIV-1 gene expression in latently infected cells. *Cell Host Microbe* 2015;**17**:107–17. https://doi.org/10.1016/j.chom.2014.12.002
40. Aoi Y, Shah AP, Ganesan S *et al.* SPT6 functions in transcriptional pause/release via PAF1C recruitment. *Mol Cell* 2022;**82**:3412-3423. https://doi.org/10.1016/j.molcel.2022.06.037
41. Shi Y, Manley JL. The end of the message: multiple protein–RNA interactions define the mRNA polyadenylation site. *Genes Dev* 2015;**29**:889–97. https://doi.org/10.1101/gad.261974.115
42. Shi Y, Di Giammartino DC, Taylor D *et al.* Molecular architecture of the human pre-mRNA 3′ processing complex. *Mol Cell* 2009;**33**:365–76. https://doi.org/10.1016/j.molcel.2008.12.028
43. Li W, You B, Hoque M *et al.* Systematic profiling of poly(A)+ transcripts modulated by core 3′ end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet* 2015;**11**:e1005166. https://doi.org/10.1371/journal.pgen.1005166
44. Kamieniarz-Gdula K, Gdula MR, Panser K *et al.* Selective roles of vertebrate PCF11 in premature and full-length transcript termination. *Mol Cell* 2019;**74**:158–72. https://doi.org/10.1016/j.molcel.2019.01.027
45. Zhang Z, Fu J, Gilmour DS. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. *Genes Dev* 2005;**19**:1572–80. https://doi.org/10.1101/gad.1296305
46. Zhang Z, Gilmour DS. Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol Cell* 2006;**21**:65–74. https://doi.org/10.1016/j.molcel.2005.11.002
47. West S, Proudfoot NJ. Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. *Nucleic Acids Res* 2008;**36**:905–14. https://doi.org/10.1093/nar/gkm1112
48. Cortazar MA, Sheridan RM, Erickson B *et al.* Control of RNA pol II speed by PNUTS-PP1 and Spt5 dephosphorylation facilitates termination by a "sitting duck torpedo" mechanism. *Mol Cell* 2019;**76**:896–908. https://doi.org/10.1016/j.molcel.2019.09.031
49. Russo M, Piccolo V, Polizzese D *et al.* Restrictor synergizes with Symplekin and PNUTS to terminate extragenic transcription. *Genes Dev* 2023;**37**:1017–40. https://doi.org/10.1101/gad.351057.123
50. Estell C, Davidson L, Eaton JD *et al.* A restrictor complex of ZC3H4, WDR82, and ARS2 integrates with PNUTS to control unproductive transcription. *Mol Cell* 2023;**83**:2222–39 https://doi.org/10.1016/j.molcel.2023.05.029
51. Ait Said M, Bejjani F, Abdouni A *et al.* Premature transcription termination complex proteins PCF11 and WDR82 silence HIV-1 expression in latently infected cells. *Proc Natl Acad Sci USA* 2023;**120**:2222–39. https://doi.org/10.1073/pnas.2313356120
52. Matkovic R, Morel M, Lanciano S *et al.* TASOR epigenetic repressor cooperates with a CNOT1 RNA degradation pathway to repress HIV. *Nat Commun* 2022;**13**:66. https://doi.org/10.1038/s41467-021-27650-5
53. Perrin S, Firmo C, Lemoine S *et al.* Aozan: an automated post-sequencing data-processing pipeline. *Bioinformatics* 2017;**33**:2212–3. https://doi.org/10.1093/bioinformatics/btx154

54. Connell Z, Parnell TJ, McCullough LL *et al.* The interaction between the Spt6-tSH2 domain and Rpb1 affects multiple functions of RNA Polymerase II. *Nucleic Acids Res* 2022;**50**:784–802. https://doi.org/10.1093/nar/gkab1262
55. Nojima T, Gomes T, Grosso AR *et al.* Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* 2015;**161**:526–40. https://doi.org/10.1016/j.cell.2015.03.027
56. Fong N, Brannan K, Erickson B *et al.* Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. *Mol Cell* 2015;**60**:256–67. https://doi.org/10.1016/j.molcel.2015.09.026
57. Estell C, Davidson L, Steketee PC *et al.* ZC3H4 restricts non-coding transcription in human cells. *eLife* 2021;**10**:e67305. https://doi.org/10.7554/eLife.67305
58. Austenaa LMI, Piccolo V, Russo M *et al.* A first exon termination checkpoint preferentially suppresses extragenic transcription. *Nat Struct Mol Biol* 2021;**28**:337–46. https://doi.org/10.1038/s41594-021-00572-y
59. Dasilva LF, Blumenthal E, Beckedorff F *et al.* Integrator enforces the fidelity of transcriptional termination at protein-coding genes. *Sci Adv* 2021;**7**:eabe3393. https://doi.org/10.1126/sciadv.abe3393
60. Noble CG, Hollingworth D, Martin SR *et al.* Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nat Struct Mol Biol* 2005;**12**:144–51. https://doi.org/10.1038/nsmb887
61. Meinhart A, Cramer P. Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors. *Nature* 2004;**430**:223–6. https://doi.org/10.1038/nature02679
62. Kim YM, Watanabe T, Allen PB *et al.* PNUTS, a protein phosphatase 1 (PP1) nuclear targeting subunit. Characterization of its PP1- and RNA-binding domains and regulation by phosphorylation. *J Biol Chem* 2003;**278**:13819–28. https://doi.org/10.1074/jbc.M209621200
63. Ciurciu A, Duncalf L, Jonchere V *et al.* PNUTS/PP1 regulates RNAPII-mediated gene expression and is necessary for developmental growth. *PLoS Genet* 2013;**9**:e1003885. https://doi.org/10.1371/journal.pgen.1003885
64. Brázda P, Krejčíková M, Kasiliauskaite A *et al.* Yeast Spt6 reads multiple phosphorylation patterns of RNA polymerase II C-terminal domain *In vitro*. *J Mol Biol* 2020;**432**:4092–107. https://doi.org/10.1016/j.jmb.2020.05.007
65. Cossa G, Roeschert I, Prinz F *et al.* Localized inhibition of protein phosphatase 1 by NUAK1 promotes spliceosome activity and reveals a MYC-sensitive feedback control of transcription. *Mol Cell* 2020;**77**:1322–39. https://doi.org/10.1016/j.molcel.2020.01.008
66. Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* 2012;**13**:720–31. https://doi.org/10.1038/nrg3293
67. Jeronimo C, Robert F. The histone chaperone FACT: a guardian of chromatin structure integrity. *Transcription* 2022;**13**:16–38. https://doi.org/10.1080/21541264.2022.2069995
68. Sanidas I, Polytarchou C, Hatziapostolou M *et al.* Phosphoproteomics screen reveals akt isoform-specific signals linking RNA processing to lung cancer. *Mol Cell* 2014;**53**:577–90. https://doi.org/10.1016/j.molcel.2013.12.018
69. Sansó M, Levin RS, Lipp JJ *et al.* P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. *Genes Dev* 2016;**30**:117–31. https://doi.org/10.1101/gad.269589.115
70. Austenaa LMI, Barozzi I, Simonatto M *et al.* Transcription of mammalian *cis*-regulatory elements is restrained by actively enforced early termination. *Mol Cell* 2015;**60**:460–74. https://doi.org/10.1016/j.molcel.2015.09.018
71. Kaplan CD, Holland MJ, Winston F. Interaction between transcription elongation factors and mRNA 3′;-end formation at the *Saccharomyces cerevisiae* GAL10-GAL7 locus. *J Biol Chem* 2005;**280**:913–22. https://doi.org/10.1074/jbc.M411108200
72. Geisberg JV, Moqtaderi Z, Struhl K. Chromatin regulates alternative polyadenylation via the RNA polymerase II elongation rate. *Proc Natl Acad Sci USA* 2024;**121**:e2405827121. https://doi.org/10.1073/pnas.2405827121
73. Wang R, Zheng D, Wei L *et al.* Regulation of intronic polyadenylation by PCF11 impacts mRNA expression of long genes. *Cell Rep* 2019;**26**:2766–78. https://doi.org/10.1016/j.celrep.2019.02.049
74. Preker P, Nielsen J, Kammler S *et al.* RNA exosome depletion reveals transcription upstream of active human promoters. *Science* 2008;**322**:1851–4. https://doi.org/10.1126/science.1164096
75. Eaton JD, Board J, Davidson L *et al.* Human promoter directionality is determined by transcriptional initiation and the opposing activities of INTS11 and CDK9. *eLife* 2024;**13**:RP92764. https://doi.org/10.7554/eLife.92764
76. Andersen PR, Domanski M, Kristiansen MS *et al.* The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat Struct Mol Biol* 2013;**20**:1367–76. https://doi.org/10.1038/nsmb.2703
77. Andrulis ED, Werner J, Nazarian A *et al.* The RNA processing exosome is linked to elongating RNA polymerase II in *Drosophila*. *Nature* 2002;**420**:837–41. https://doi.org/10.1038/nature01181
78. Rouvière JO, Salerno-Kochan A, Lykke-Andersen S *et al.* ARS2 instructs early transcription termination-coupled RNA decay by recruiting ZC3H4 to nascent transcripts. *Mol Cell* 2023;**83**:2240–57. https://doi.org/10.1016/j.molcel.2023.05.028
79. Bentley D, Treisman R, Erickson B *et al.* PP1 PNUTS binds the restrictor and dephosphorylates RNA pol II CTD Ser5 to stimulate transcription termination. bioRxiv, https://doi.org/10.1101/2024.07.12.603302, 13 July 2024, preprint: not peer reviewed.
80. Arigo JT, Eyler DE, Carroll KL *et al.* Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol Cell* 2006;**23**:841–51. https://doi.org/10.1016/j.molcel.2006.07.024
81. Grzechnik P, Gdula MR, Proudfoot NJ. Pcf11 orchestrates transcription termination pathways in yeast. *Genes Dev* 2015;**29**:849–61. https://doi.org/10.1101/gad.251470.114
82. Zhang H, Hu J, Recce M *et al.* PolyA_DB: a database for mammalian mRNA polyadenylation. *Nucleic Acids Res* 2004;**33**:D116–20. https://doi.org/10.1093/nar/gki055


**Received:** September 3, 2024. **Revised:** February 18, 2025. **Editorial Decision:** February 20, 2025. **Accepted:** February 24, 2025