Published online 23 March 2022

Nucleic Acids Research, 2022, Vol. 50, No. 7 ***3835–3851***
https://doi.org/10.1093/nar/gkac193

ZWC complex-mediated SPT5 phosphorylation suppresses divergent antisense RNA transcription at active gene promoters

Kihyun Park[1], Jian Zhong[2], Jin Sung Jang[3,4], Jihyun Kim[1], Hye-Jung Kim[5], Jeong-Heon Lee[2,4,6,*] and Jaehoon Kim[1,*]

[1]Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea, [2]Epigenomics Program, Center for Individualized Medicine, Mayo Clinic, Rochester, MN 55905, USA, [3]Medical Genome Facility, Center for Individualized Medicine, Mayo Clinic, Rochester, MN 55905, USA, [4]Division of Experimental Pathology and Laboratory Medicine, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA, [5]New Drug Development Center, OSONG Medical Innovation Foundation, Cheongju 28160, South Korea and [6]Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN 55905, USA

Received November 01, 2021; Revised March 08, 2022; Editorial Decision March 11, 2022; Accepted March 13, 2022

ABSTRACT

The human genome encodes large numbers of non-coding RNAs, including divergent antisense transcripts at transcription start sites (TSSs). However, molecular mechanisms by which divergent antisense transcription is regulated have not been detailed. Here, we report a novel ZWC complex composed of ZC3H4, WDR82 and CK2 that suppresses divergent antisense transcription. The ZWC complex preferentially localizes at TSSs of active genes through direct interactions of ZC3H4 and WDR82 subunits with the S5p RNAPII C-terminal domain. ZC3H4 depletion leads to increased divergent antisense transcription, especially at genes that naturally produce divergent antisense transcripts. We further demonstrate that the ZWC complex phosphorylates the previously uncharacterized N-terminal acidic domain of SPT5, a subunit of the transcription-elongation factor DSIF, and that this phosphorylation is responsible for suppressing divergent antisense transcription. Our study provides evidence that the newly identified ZWC-DSIF axis regulates the direction of transcription during the transition from early to productive elongation.

INTRODUCTION

RNA polymerase II (RNAPII)-mediated transcription is a highly-coordinated and complex process regulated by a large number of transcriptional co-regulators. Transcription studies have long focused on understanding regulatory mechanisms of dominant and directional transcripts (defined as sense), especially in protein-coding genes, but recent advances in high-throughput genome-wide approaches have revealed that a large proportion of the genome also produces antisense transcripts in a range of species (1–3). Antisense transcripts are important players in gene expression and genome integrity by virtue of their ability to interfere with sense transcripts and alter chromatin structure (4,5); as such, antisense transcripts have been implicated in disease pathogenesis and are potential targets for therapeutic development (6,7). Antisense transcripts can originate from independent promoters, cryptic promoters and bidirectional promoters, which are classified according to their position relative to sense transcripts (5). The majority of mammalian promoters direct transcription initiation on both sides with opposite orientations, a phenomenon called divergent transcription (8). In a small fraction of genes, divergent transcription generates low-abundance, but lengthy, antisense RNAs (natural divergent antisense transcripts) at transcription start sites (TSSs) (9). Because the expression of antisense transcripts changes during cellular differentiation and according to environmental conditions (4,9), it is expected that divergent antisense transcription is also regulated by transcriptional co-regulators. However, how the direction of transcription is determined during the early transcription-elongation stage and what transcriptional co-regulators are involved in this process remain unclear.

WD40 repeat domain 82 (WDR82) participates in the regulation of multiple phases of the transcription process. As a subunit of SET1A/B complexes, WDR82 plays a role in histone H3K4 methylation in the promoter

*To whom correspondence should be addressed. Tel: +82 42 350 2632; Email: kimjaehoon@kaist.edu
Correspondence may also be addressed to Jeong-Heon Lee. Tel: +1 507 293 1687; Email: lee.jeongheon@mayo.edu

regions of actively transcribing genes by mediating the association of complexes with initiating and early elongating RNAPII (10,11). In addition, as a subunit of the PTW/PP1 phosphatase complex, WDR82 is involved in transcription termination and pre-mRNA 3′ end processing (12–15). The presence of a WD40 repeat domain, which often mediates protein-protein interactions (16,17), raises the possibility that WDR82 may have additional roles in the transcription process as a component of an as-yet-unidentified protein complex. In support of this, our previous immunoprecipitation-coupled mass spectrometry (IP-MS) analysis identified ZC3H4 (zinc finger CCCH-type containing 4)—a protein of unknown function—as a novel WDR82-interacting protein (15). In a similar vein, very recent studies have reported that WDR82 and ZC3H4 are involved in suppressing the transcription of long non-coding RNA (lncRNA) from promoter and enhancer regions (18,19).

Casein kinase 2 (CK2), composed of catalytic CK2α and regulatory CK2β subunits, is a constitutively active protein kinase that is ubiquitously expressed and evolutionarily conserved in eukaryotic cells (20,21). CK2 was initially reported as a serine/threonine kinase (22), but CK2-mediated tyrosine phosphorylation has also been reported (23,24). Proteomics-based studies have identified over 300 candidate proteins that can be phosphorylated by CK2 (25), suggesting versatile roles of CK2 in various cellular processes (21,26). Among several lines of evidence supporting the involvement of CK2 in the transcription process, recent studies in yeast have reported that CK2-mediated phosphorylation of the histone chaperone, Spt6, prevents aberrant antisense transcription (27,28).

DRB sensitivity-inducing factor (DSIF), composed of SPT5 and SPT4 subunits, is a transcription-elongation factor that controls RNAPII processivity during early and productive elongation stages (29). It has been well established that promoter-proximal pausing of RNAPII can be relieved by positive transcription-elongation factor b (P-TEFb)-mediated phosphorylation of DSIF, negative elongation factor (NELF), and RNAPII (29,30). P-TEFb was shown to phosphorylate multiple amino acid residues in the repetitive C-terminal region (CTR) of SPT5, thereby converting DSIF to a form that stimulates productive transcription elongation (31). Interestingly, depletion of SPT5 has been reported to cause delayed and reduced early transcription (32). Moreover, a recent study in fission yeast showed that Spt5 depletion results in altered RNAPII distribution and increased antisense transcription (33). Importantly, this study also showed that deletion of the Spt5 CTR exhibited only a small increase in antisense transcripts, suggesting that another region besides CTR is critical for regulating antisense transcription. These observations suggest an additional and uncharacterized role of DSIF in controlling antisense transcription in addition to its previously known function in promoter-proximal pausing through P-TEFb-mediated phosphorylation of the CTR of SPT5.

In this study, using biochemically defined systems and genome-wide approaches, we demonstrate that a novel ZWC complex composed of ZC3H4, WDR82 and CK2 phosphorylates the N-terminal acidic domain of SPT5, thereby suppressing divergent antisense transcription at TSSs of actively transcribed genes. We also provide mechanistic insight into ZWC complex-mediated SPT5 phosphorylation, showing that this process involves direct interactions among ZC3H4, WDR82 and the Ser5-phosphorylated RNAPII C-terminal domain (CTD), which brings CK2 to active TSSs, where it phosphorylates SPT5. In addition to the previously established function of P-TEFb-mediated phosphorylation of the SPT5 CTR in promoter-proximal pausing (P-TEFb-DSIF axis), our study reveals that ZWC complex-mediated phosphorylation of the N-terminal acidic domain of SPT5 has a role in suppressing divergent antisense transcription (ZWC-DSIF axis).

MATERIALS AND METHODS

Cell lines and RNA interference

The tetracycline-inducible FLAG-ZC3H4 stable cell line was generated by transfecting T-REx HEK293 cells (Invitrogen) with the pcDNA5/TO plasmid (Invitrogen) containing FLAG-ZC3H4 cDNA and selecting stable clones by growing in medium containing 0.2 mg/ml hygromycin B (Sigma-Aldrich). The stable FLAG-SPT5 cell line was generated by transfecting HEK293T cells with a pCAGIPuro plasmid containing FLAG-SPT5 cDNA and selecting stable clones by growing in medium containing 2 μg/ml puromycin (Sigma-Aldrich). For RNA interference, cells were transfected with small interfering RNA (siRNA) oligonucleotides using DharmaFECT1 transfection reagent (Dharmacon) according to the manufacturer's protocol. Oligonucleotide sequences for siRNA are summarized in Supplementary Table S1.

Immunoprecipitation of ZC3H4-interacting proteins and protein identification by mass spectrometry

T-REx HEK293 cells inducibly expressing FLAG-ZC3H4 or empty vector were treated with 1 μg/ml doxycycline for 1.5 days. Nuclear extracts were prepared, immunoprecipitated with anti-FLAG M2 agarose beads (Sigma-Aldrich), and extensively washed as previously described (15). Bound proteins were eluted with 10M urea. The protein sample was reduced with 10 mM DTT at 56°C for 30 min and alkylated with 20 mM iodoacetamide at 25°C for 30 min. The sample was diluted 10 times to decrease urea concentration to <1 M in solution and added 25 mM NH_4HCO_3. The protein mixture was digested by sequencing-grade modified trypsin (Promega) at 37°C for 16 h. The ratio of enzyme to protein was 1:50. Digested peptides were analyzed utilizing a Q-Exactive plus mass spectrometry with an Easy nLC-1200 pump and autosampler system (Thermo Fisher Scientific). The tryptic peptides were separated using a linear gradient of acetonitrile from 5% to 60% in water in the presence of 0.1% formic acid over a period of 60 min. The mass spectrometer was operated in data-dependent mode with a full scan (m/z 350–2000) followed by MS/MS for the top 20 precursor ions in each cycle. The acquired MS/MS spectra were subjected to searches against the Uniprot-human database (Homo sapiens, June 2021; 79 038 sequences) using SEQUEST software in Proteome Discoverer 2.4 (Thermo Fisher Scientific, version 1.4.0.288).

Two missed trypsin cleavages were allowed, and the peptide mass tolerances for MS/MS and MS were set to 10 ppm and 0.02 Da, respectively. Other parameters used for the SEQUEST searches included the fixed modification of carbamidomethylation at cysteine (+57.021 Da) the variable modification of oxidation at methionine (+15.995 Da).

cDNA, plasmids, and recombinant protein and complex purification

cDNAs for human ZC3H4 (Gene ID: 23211), WDR82 (80335), CK2α (1457), CK2β (1460), SPT5 (6829) and SPT4 (6827) were generated by reverse transcription-polymerase chain reaction (RT-PCR) from total RNA extracted from HEK293T cells. For bacterially expressed His-tagged proteins, cDNAs were subcloned into pET28 (Novagen), expressed in *Escherichia coli*, and purified on Ni-NTA beads (Qiagen) as previously described (34). For baculovirus-mediated reconstitution and purification of ZWC and DSIF complexes, cDNAs encoding complex subunits were subcloned into pFASTBAC1 (Gibco-Invitrogen) with or without a FLAG tag. Baculoviruses were generated according to the manufacturer's instructions (Gibco-Invitrogen) and coinfected into Sf9 insect cells. For the ZWC complex, cell extracts were prepared in lysis buffer (20 mM Tris–Cl [pH 7.9], 300 mM KCl, 2.5 mM $MgCl_2$, 0.25 mM EDTA, 15% glycerol, 0.1% NP-40, 1.2 mM DTT, 0.5 mM PMSF and 1× protease inhibitor cocktail [Roche]). ZWC complexes in clarified extracts were affinity purified on M2 agarose beads. After extensive washing with wash buffer (20 mM Tris–Cl [pH 7.9], 150 mM KCl, 2 mM $MgCl_2$, 0.2 mM EDTA, 15% glycerol, 0.1% NP-40, 1 mM DTT and 1 mM PMSF), complexes were eluted with wash buffer containing 0.25 mg/ml FLAG peptide. Affinity-purified complexes were loaded onto a Mono Q 5/50 GL column (Sigma-Aldrich) and then eluted with a linear KCl gradient (100–1000 mM) in buffer consisting of 20 mM Tris–Cl [pH 7.9], 0.2 mM EDTA and 5% glycerol. Peak fractions were pooled and dialyzed against dialysis buffer (20 mM Tris–Cl [pH 7.9], 100 mM KCl, 0.2 mM EDTA and 20% glycerol). For DSIF, complexes were affinity purified on M2 agarose as above except a wash buffer containing 500 mM KCl was used, and the complex solution was dialyzed against dialysis buffer.

ChIP-seq and data analysis

For chromatin immunoprecipitation followed by sequencing (ChIP-seq), HEK293 cells were cross-linked with 1% formaldehyde for 10 min, followed by quenching with 125 mM glycine for 5 min at room temperature. Fixed cells were washed twice with Tris-buffered saline (TBS), resuspended in cell lysis buffer (10 mM Tris–HCl [pH 7.5], 10 mM NaCl and 0.5% NP-40), and incubated on ice for 10 min. Lysates were washed with MNase digestion buffer (20 mM Tris–HCl [pH 7.5], 15 mM NaCl, 60 mM KCl and 1 mM $CaCl_2$) and incubated for 20 min at 37°C in the presence of MNase (2000 gel units/4 × 10^6 cells; New England Biolabs). After adding the same volume of sonication buffer (100 mM Tris–HCl [pH 8.1], 20 mM EDTA, 200 mM NaCl, 2% Triton X-100 and 0.2% sodium deoxycholate), the lysate was sonicated for 5 cycles (30 s on/30 s off per cycle) for chromatin-associated proteins or 15 cycles for modified histones using a Diagenode Bioruptor and centrifuged at 15 000 rpm for 10 min. The clarified supernatant, equivalent to 4 × 10^6 cells for modified histones or 10^7 cells for chromatin-associated proteins, was incubated with 2 μg of antibody at 4°C on a rocker overnight. Subsequent immunoprecipitation and library preparation were performed as previously described (35). Libraries were sequenced to 51 base pairs (bp) from both ends on an Illumina NextSeq instrument (Illumina) according to the manufacturer's instructions. One replicate of ChIP-seq data was generated for RNAPII and the histone marks H3K4me3 and H3K27me3, and compared with public datasets (GSM3272322, GSM1482819, GSM3272328, GSM327232 and GSM2643621). Two biological replicates of ChIP-seq data were generated for ZC3H4, WDR82, CK2α and SPT5 and for all targets in nonspecific control or ZC3H4 siRNA-treated cells. The ChIP-seq data are accessible through GEO accession number GSE186758.

ChIP-seq reads were mapped to the hg19 reference genome using Burrows-Wheeler Aligner (BWA, v0.7.10, http://bio-bwa.sourceforge.net) (36). Duplicates were removed using Picard (v1.67, https://broadinstitute.github.io/picard/). For visualization, read counts were generated over 200-bp sliding windows with a 20-bp step size and normalized by library size using BEDTools (v2.16.2, https://bedtools.readthedocs.io/en/latest/) (37). For H3K4me3, peaks were identified using model-based analysis of ChIP-seq data (MACS; v2.0.10, https://pypi.org/project/MACS2/) (38) at a cut-off of ≤1% and FDR ≥2-fold change relative to input. For RNAPII, S5p RNAPII, S2p RNAPII, SPT5, H3K36me3 and H3K27me3, peaks were identified using the SICER (spatial clustering approach for the identification of ChIP-enriched regions) software package (v0.1.1) (39) at an FDR ≤1%. Heatmap and aggregate per-million signal density profiles of ChIP-seq over TSSs and gene bodies were generated using the ngs.plot tool (v2.63, https://github.com/shenlab-sinai/ngsplot) (40). The peak distribution over genomic features was generated using ChIPseeker (v1.26.0, http://bioconductor.org/packages/devel/bioc/vignettes/ChIPseeker/inst/doc/ChIPseeker.html) (41) and the TxDb.Hsapiens.UCSC.hg19.knownGene annotation package (v3.2.2). Correlation plots between pairs of ChIP-seq libraries were generated by merging peaks from the two libraries and calculating raw read count per merged peak using Deeptools (v3.3.2, https://deeptools.readthedocs.io/en/latest/content/list_of_tools.html) (42). Raw read counts were normalized by library size, $\log_2$ transformed, and used as input for generating plots.

Immunoprecipitation assays

Cells were lysed in lysis buffer (20 mM Tris–Cl [pH 7.9], 150 mM KCl, 0.2 mM EDTA, 20% glycerol, 0.05% NP-40, 1 mM PMSF and 1× protease and phosphatase inhibitor cocktail [Thermo Fisher Scientific]) and cell extracts were incubated with primary antibody at 4°C overnight. Protein A/G PLUS-Agarose beads (Santa Cruz Biotechnology) were then added and extracts were further incubated

at 4°C for 1 h. Beads were extensively washed with lysis buffer, and bound proteins were monitored by immunoblot analyses. For assays with CK2 inhibition, cells were treated with 50 μM quinalizarin (a CK2 kinase-specific inhibitor; Sigma-Aldrich) or DMSO (control) for 24 h prior to immunoprecipitation.

Peptide pull-down assays

For peptide pull-down assays, 3 μg of biotin-tagged RNAPII CTD peptides (three copies of YSPTSPS heptad repeat; Anygen), immobilized on streptavidin agarose beads (Pierce), were incubated with cell extracts or 200 ng of purified proteins in binding buffer (20 mM Tris–Cl [pH 7.9], 150 mM KCl, 0.2 mM EDTA, 20% glycerol, 0.05% NP-40, 1 mM PMSF, 0.2 mg/ml bovine serum albumin [BSA] and 1× protease and phosphatase inhibitor cocktail) at 4°C for 3 h. Beads were extensively washed with binding buffer without BSA, and bound proteins were identified by immunoblot analysis.

In vitro kinase assays

In vitro kinase reactions were performed by incubating purified ZWC complex (containing 30 ng CK2α) and DSIF (containing 200 ng SPT5) in 15 μl reaction buffer (25 mM HEPES [pH 7.6], 50 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA and 10% glycerol) supplemented with 0.1 mM cold ATP (for immunoblotting) or 1 μCi [γ-^{32}P] ATP (for fluorography) at 30°C for 30 min. For assays with CK2 inhibition, 50 μM quinalizarin or DMSO (control) was added to the reaction. Proteins were resolved by SDS-PAGE and subjected to immunoblotting (cold ATP) with anti-phospho-CK2 substrate antibody or autoradiography (radiolabeled ATP).

Quantitative RT-PCR analyses

Total RNA was prepared using the TRIzol reagent (Invitrogen), and contaminating genomic DNA was removed by treatment with RNase-free DNase I (New England Biolabs). cDNA was synthesized by reverse transcription using a PrimeScript RT Master Mix (Takara) and quantitative PCR (qPCR) was performed using TOPreal qPCR PreMix (Enzynomics). Primers for qPCR are summarized in Supplementary Table S2. The data analysis was performed by calculating ΔCq normalized to β-actin and GAPDH expression by the Bio-Rad CFX Manager 3.1 based on the MIQE guidelines (43). The MIQE checklist is provided in Supplementary Table S3.

RNA-seq and data analysis

RNA-Seq libraries were generated from 200 ng of total RNA using a TruSeq Stranded Total RNA Library Prep kit (Illumina). Constructed libraries were quantified with a Bioanalyzer 2100 system using a D1000 kit (Agilent) and Qubit dsDNA BR Assay kit (Thermo Fisher Scientific). Two libraries in one flow cell were sequenced to 101 bp from both ends on an Illumina HiSeq 4000 system. FASTQ files were uploaded into Partek Flow software (Partek Inc.), used to perform primary quality control analyses. Reads were aligned to the human reference genome (hg19 and hg38) using STAR (2.6.1d) aligner. After alignment, the final BAM files were quantified with Ensembl annotations (Ensembl Transcripts release 101) using the Partek E/M algorithm (44). Strand-specific alignment and mapping were done with StringTie2 (45). The DESeq2 (46) package was used for normalizing data and determining differential expression between two groups in RNA-Seq analyses. False discovery rate (FDR), determined with the Benjamin and Hochberg method (47), was used to adjust comparisons. FDR values <0.05 and fold-changes greater than ± 2.0 were considered significant. Hierarchical clustering was performed on significantly changed genes based on the Euclidean distance and average linkage clustering algorithm using Partek Flow software (Partek). Gene Set Enrichment Analysis (GSEA; Broad Institute) was carried out in Molecular Signatures Database v5.2 (MsigDB). Per-base-pair signal track data were generated using BEDTools from STAR mapping output BAM files and normalized to the library size, separately for sense and antisense strands. Heatmap and aggregate per-million signal density profiles over the TSS and gene-body of selected genes were also generated separately for sense and antisense strands using the ngs.plot tool. The RNA-seq data are accessible through GEO accession number GSE186808.

Antibodies

The following antibodies were used for immunoblotting and immunoprecipitation: polyclonal anti-ZC3H4 antibody, developed against a purified GST-tagged human ZC3H4 fragment (amino acid residues 1114–1315) in this study; polyclonal anti-WDR82 antibody (kindly provided by the Roeder laboratory); anti-CK2α (10992-1-AP, Proteintech); anti-CK2β (ab76025, Abcam); anti-SPT5 (A300-869A for immunoblotting; A300-868A for immunoprecipitation, Bethyl Laboratories); anti-RNAPII (ab26721, Abcam); anti-S5p RNAPII (ab5131, Abcam); anti-S2p RNAPII (ab5095, Abcam); anti-β-actin (TA811000, Origene); anti-phospho-CK2 substrate [(pS/pT)DXE] (8738, Cell Signaling Technology) and anti-FLAG (A8592, Sigma-Aldrich). The following antibodies were used for ChIP-seq analyses: anti-ZC3H4 (this study) and anti-WDR82 (48), generated in house; anti-CK2α (ab70774, Abcam); anti-SPT5 (sc-133217, Santa Cruz); anti-RNAPII (A300-653A, Bethyl Laboratories); anti-S5p RNAPII (ab5131, Abcam); anti-Sp2 RNAPII (A300-654A, Bethyl Laboratories); anti-H3K4me3 (C15410003, Diagenode); anti-H3K27me3 (9733, Cell Signaling Technology) and anti-H3K36me3 (61101, Active Motif).

RESULTS

Identification of the ZWC complex, composed of ZC3H4, WDR82, CK2α and CK2β subunits

In our previous immunoprecipitation-mass spectrometric (IP-MS) analysis using a FLAG-WDR82–expressing cell line, we found that WDR82 was associated with multiple protein complexes, including histone H3K4 methyltransferase complexes, the PTW/PP1 phosphatase complex and

A

Protein name	Gene name	Fold change (FLAG-ZC3H4/ Control)	*p*-value
Zinc finger CCCH-type containing 4 (ZC3H4)	ZC3H4	1017.86	0.001
WD40 repeat domain 82 (WDR82)	WDR82	6.56	0.022
Casein kinase II subunit alpha (CK2α)	CSNK2A1	12.65	0.002
Casein kinase II subunit beta (CK2β)	CSNK2B	3.99	0.021

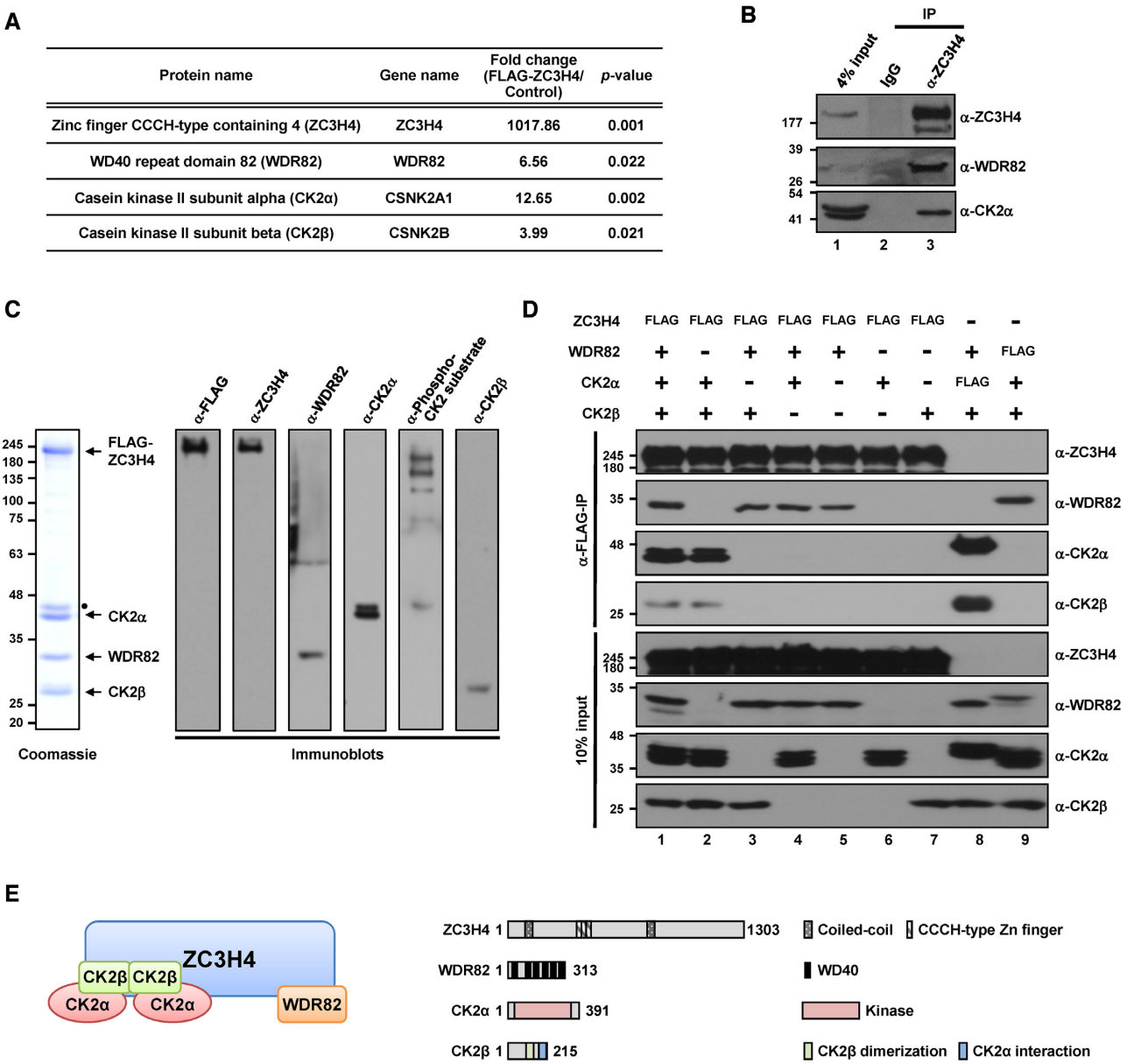

Figure 1. Identification and reconstitution of the ZWC complex. (**A**) ZC3H4-interacting proteins identified by liquid chromatography–mass spectrometry (LC–MS) analysis. T-REx HEK293 cell lines stably expressing FLAG-ZC3H4 or empty vector were induced with doxycycline and nuclear extracts were prepared and immunoprecipitated with anti-FLAG antibody. Bound proteins were identified by LC–MS analysis. All identified ZC3H4-interacting proteins are listed in Supplementary Table S4. (**B**) Intracellular interactions between ZC3H4, WDR82 and CK2α. Nuclear extracts prepared from HEK293 cells were immunoprecipitated with an anti-ZC3H4 antibody, and bound proteins were monitored by immunoblotting with the indicated antibodies. (**C**) SDS-PAGE/Coomassie Blue staining and immunoblot analyses of the reconstituted and purified ZWC complex. The dot indicates a phosphorylated form of CK2α. (**D**) Cell extracts were prepared from Sf9 insect cells coinfected with combinations of baculoviruses expressing FLAG-tagged (FLAG) or untagged proteins as indicated, and immunoprecipitated with anti-FLAG antibodies (α-FLAG-IP). Proteins were detected by immunoblotting with the indicated antibodies. (**E**) A model of subunit interactions within the ZWC complex based on the analyses shown in Figure 1D and Supplementary Figure S1C. Note that two copies of CK2α and CK2β are depicted as present within the complex based on the previously reported heterotetrameric subunit composition of the CK2 complex (20) (left). A schematic diagram of the domains in ZWC complex subunits (right).

RNAPII (15). This approach also identified ZC3H4 (also called C19orf7), but functional studies of WDR82 and ZC3H4 interactions were not carried out. To confirm WDR82 and ZC3H4 interactions, we performed FLAG-IP using nuclear extracts prepared from a doxycycline-inducible FLAG-ZC3H4 T-REx HEK293 cell line and identified interacting proteins by MS analysis (Supplementary Table S4). In addition to identifying WDR82, which validates ZC3H4–WDR82 interactions, our MS analysis also identified CK2α and CK2β that comprise CK2 with high confidence (Figure 1A) as demonstrated in a volcano plot (Supplementary Figure S1A). We have focused on CK2 because of its potential to phosphorylate transcription machinery.

In a test of intracellular interactions, we found that an anti-ZC3H4 antibody co-immunoprecipitated endogenous WDR82 and CK2α (Figure 1B), suggesting that ZC3H4, WDR82 and CK2 form a complex in cells. To examine complex formation, we reconstituted the ZC3H4–WDR82–CK2 complex (hereafter, ZWC complex) and purified it from Sf9 insect cells coinfected with baculoviruses expressing FLAG-ZC3H4 and untagged WDR82, CK2α and CK2β using anti-FLAG affinity chromatography, followed by further purification by Mono Q chromatography fractionation (Supplementary Figure S1B). Protein staining and immunoblotting clearly revealed co-purification of four protein subunits, confirming formation of a stable complex (Figure 1C). A slowly migrating band above CK2α in Coomassie Blue-stained gels was considered an autophosphorylated form of CK2α because it was detected simultaneously by an anti-CK2α antibody and an anti-phospho-CK2 substrate antibody that recognizes a CK2 phosphorylation consensus sequence. (It is also possible that CK2α can be phosphorylated by endogenous kinases other than CK2.)

Reconstituting complexes with various combinations of subunits using a baculovirus expression system (Figure 1D) led us to draw several conclusions regarding complex integrity. First, interactions of ZC3H4 with CK2 and WDR82 were maintained by omission of WDR82 (lane 2) or CK2 (lane 5), respectively, indicating that WDR82 and CK2 directly associate with ZC3H4 and demonstrating the role of ZC3H4 as a scaffold protein within the complex. Second, exclusion of ZC3H4 resulted in no interaction between WDR82 and CK2 (lanes 8 and 9), suggesting their independent association with ZC3H4. Additional complex reconstitution analyses confirmed that WDR82 and CK2 interact with C-terminal and N-terminal regions of ZC3H4, respectively (Supplementary Figure S1C). Third, individual omission of CK2α or CK2β led to depletion of both in the complex (lanes 3, 4, 6 and 7), indicating that heteromeric association of CK2α and CK2β is required for their retention within the complex. Collectively, these analyses establish subunit interactions within the ZWC complex, as depicted in Figure 1E.

The ZWC complex associates with TSSs of active genes

WDR82 participates in the regulation of various stages of transcription as a component of multi-subunit SET1A/B (10,11) and PTW/PP1 (12–15) complexes, raising the possibility that the ZWC complex is also involved in transcriptional regulation. Therefore, we first examined the genomic distribution of ZWC complex subunits on chromatin using ChIP-seq (chromatin immunoprecipitation followed by sequencing) analyses in HEK293 cells. Similar to the genomic distribution of RNAPII, more than 80% of ChIP-seq peaks for ZWC complex subunits were found to associate with the transcriptional unit. Annotation of ChIP-seq peaks revealed that about half of all ZWC complex subunits were localized in promoter regions (Supplementary Figure S2). The percentage distributions of all subunits at each genomic region were similar, suggesting formation of a stable complex of ZC3H4, WDR82 and CK2α on chromatin. Next, we examined enrichment of ZWC complex subunits, RNAPII, and histone modifications at four genes—*GAPDH*, *EIF3K*, *HDAC6* and *MYT1* (ordered from highest to lowest expression level)—that are representatives of different transcription levels (Figure 2A). As expected, RNAPII and H3K4me3 were significantly enriched around TSSs and H3K36me3 marks in gene bodies and this enrichment was positively correlated with the expression levels of the four genes. Conversely, repressive H3K27me3 marks were found to be more prevalent in genes with lower expression. More importantly, we found that ZC3H4, WDR82 and CK2α preferentially co-localized to the TSS of each gene and their degree of enrichment correlated with gene expression levels. Heat map analyses also demonstrated significant co-localization of ZC3H4 peaks with peaks from other subunits of the ZWC complex, RNAPII, and SPT5 around TSSs of transcriptionally active genes (Figure 2B). We further found that the top 10% actively expressed genes exhibited much higher levels of subunit enrichment relative to all genes (Figure 2C). Collectively, our ChIP-seq analyses indicate that the ZWC complex is preferentially localized at the TSSs of actively transcribing genes.

The ZWC complex directly binds to the Ser5-phosphorylated RNAPII CTD through ZC3H4 and WDR82 subunits

Preferential localization of the ZWC complex at TSSs of actively transcribing genes and co-enrichment with RNAPII led us to postulate that the ZWC complex interacts with the transcription machinery. In this context, our previous co-immunoprecipitation study showed that WDR82-containing SET1A complexes can be recruited to the TSS of transcribing genes through binding of WDR82 to the Ser5-phosphorylated C-terminal domain (CTD) of RNAPII (10). Accordingly, we investigated whether WDR82-containing ZWC complexes can also interact with the RNAPII CTD. Peptide pull-down assays employing HEK293T cell extracts and biotinylated RNAPII CTD heptad repeat peptides, without modification, with Ser5 phosphorylation (S5p) or Ser2 phosphorylation (S2p), showed that ZWC complex subunits are precipitated preferentially with the S5p CTD (Supplementary Figure S3A). Similar pull-down assays using purified ZWC complexes confirmed direct and selective interaction of the ZWC complex with the S5p CTD (Figure 3A). To determine which ZWC complex subunit(s) are responsible for binding to the S5p CTD, we used purified individual subunits and subunit-deficient complexes (Supplementary Figure S3B) in peptide pull-down assays (Figure 3B). Whereas complexes lacking CK2 efficiently bound to the S5p CTD (lane 5), a WDR82-deficient complex was unable to interact with it (lane 4), suggesting a critical role of WDR82 in binding the ZWC complex to S5p RNAPII. In support of this and consistent with a previous study (10), WDR82 alone exhibited an interaction, albeit weak, with the S5p CTD (lane 7). Notably, this weak interaction was significantly enhanced by inclusion of ZC3H4 (lane 5), which alone could not bind to the S5p CTD (lane 6). These results demonstrate that ZC3H4 and WDR82 are simultaneously required for efficient binding of the ZWC complex to S5p RNAPII. In support of this, reactions containing individually added ZC3H4 and WDR82 were successfully co-immunoprecipitated with an S5p CTD

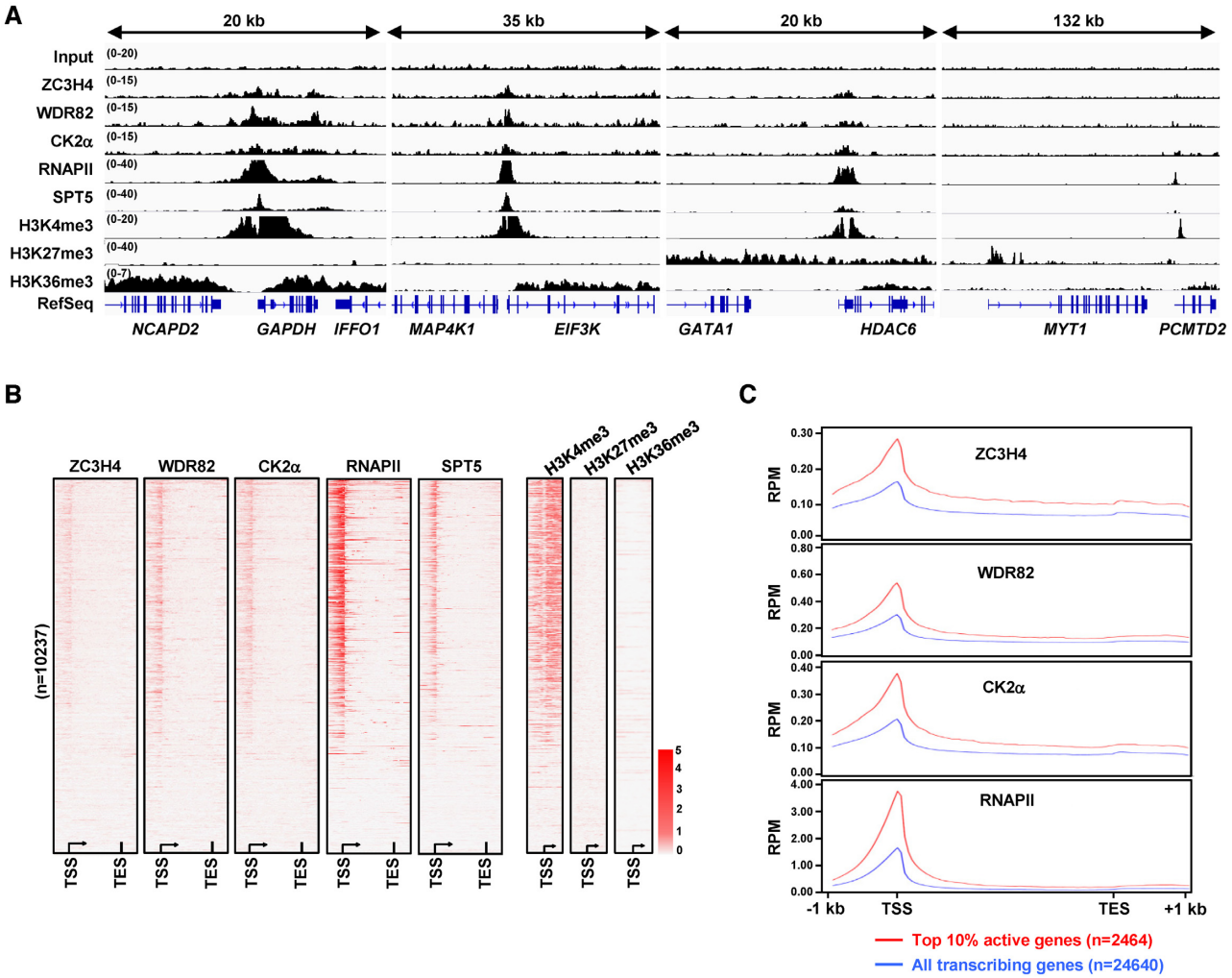

Figure 2. The ZWC complex is localized at TSSs of active genes. (**A**) Representative snapshot image of ChIP-seq results generated for ZWC complex subunits, RNAPII, SPT5 and histone modifications. Read densities were visualized in four genomic regions with different levels of transcriptional activity. (**B**) Heat map analyses of ChIP-seq peaks from ZWC complex subunits, RNAPII, SPT5 and histone modifications. ZC3H4 peaks (n = 10 237) were aligned with those of other ZWC complex subunits, RNAPII, SPT5 and modified histones. Read intensities were visualized from −1 kb relative to the TSS to +1 kb relative to the TES (transcription end site). (**C**) Transcription-dependent association of the ZWC complex with chromatin. Peak intensities are shown for the top 10% highly transcribed genes and all transcriptionally active genes (n = 24 640, RPKM >0.1). RPM, reads per million.

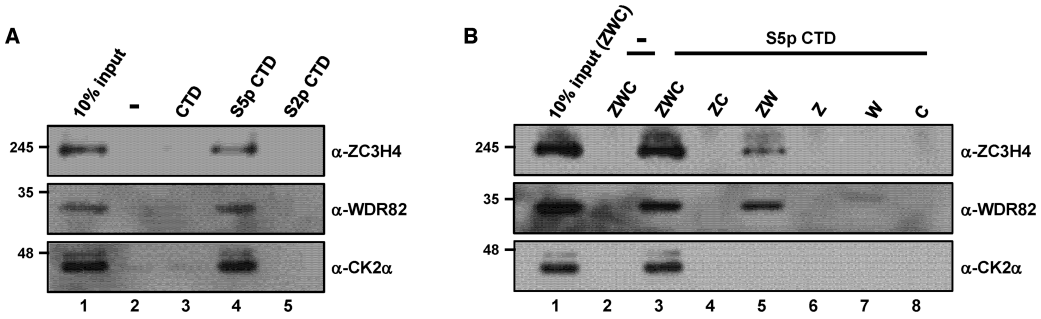

Figure 3. Direct interaction of the ZWC complex with the Ser5-phosphorylated RNAPII CTD. (**A**) Test for direct binding of the ZWC complex to unmodified, Ser5-phosphorylated (S5p) and Ser2-phosphorylated (S2p) RNAPII CTD peptides. Biotin pull-down assays employed biotinylated-RNAPII CTD peptides (three copies of YSPTSPS heptad repeats) and purified ZWC complex, and peptide-bound proteins were scored by immunoblotting with the indicated antibodies. (**B**) Requirement for binding of both ZC3H4 and WDR82 to the S5p RNAPII CTD. The biotinylated S5p RNAPII CTD peptide was tested for binding to the indicated purified complexes and proteins (Supplementary Figure S3B). Note that only the intact ZWC complex was loaded as input (lane 1) and was used for pull-down reactions without peptide (lane 2). Z, ZC3H4; W, WDR82; C, CK2.

peptide; however, all other combinations of subunits failed to exhibit comparable interactions (Supplementary Figure S3C). Collectively, these results suggest that WDR82 and the C-terminal region of ZC3H4 (Figure 1E) together create a binding platform for the interaction with S5p RNAPII, which brings CK2 to actively transcribing regions of the genome.

Depletion of ZC3H4 increases divergent antisense transcripts

To investigate the role of the ZWC complex in the transcription process, we employed siRNA-mediated depletion of ZC3H4. ZC3H4 functions as a scaffold protein in the ZWC complex (Figure 1) and is required for interaction of the complex with S5p RNAPII (Figure 3); moreover, WDR82 and CK2 are also components of other complexes. Accordingly, ZC3H4 depletion is expected to selectively abrogate function(s) of the ZWC complex in cells. Preliminary experiments confirmed selective knockdown, showing that siRNA treatment significantly reduced ZC3H4 levels without affecting those of other proteins tested (Figure 4A). In particular, WDR82 and CK2α levels were unchanged by ZC3H4 siRNA treatment, implying that ZWC complex formation is not critical for WDR82 or CK2 stability and/or that the majority of WDR82 and CK2 in cells is present as components of other complexes.

To detail genome-wide profiles of most types of transcripts in a stand-specific manner, we performed stranded total RNA sequencing (RNA-seq). First, we examined genes whose expression was changed >2-fold (FDR <0.05) in ZC3H4 siRNA-treated cells (Figure 4B, Supplementary Figure S4A). Three independent experiments showed a similar pattern of differential gene expression. A Gene Ontology analysis of up-regulated ($n = 1774$) and down-regulated ($n = 664$) genes showed no significant enrichment of particular pathways. (Supplementary Figure S4B). Instead, we found that a large percentage (∼55%) of up-regulated genes are classified as lncRNAs, suggesting that ZC3H4 (or the ZWC complex) is involved in suppression of lncRNA production (Supplementary Figure S4C).

To examine changes in the nature of transcripts caused by ZC3H4 depletion in greater detail, we investigated strand-specific RNA transcripts at a ±5-kb region around the TSS. Consistent with results shown in Figure 4B, ZC3H4 knockdown resulted in a decrease in sense transcripts for down-regulated genes ($n = 664$) (Figure 4C, top). Strikingly, we found that antisense transcripts originated at the TSS (divergent antisense transcripts) were significantly increased upon ZC3H4 depletion, whereas the levels of antisense transcripts from the gene body toward the TSS (convergent antisense transcripts) were unaltered (Figure 4C, bottom). These results suggest that, in the absence of ZC3H4, some proportion of RNAPII moves in the opposite direction to produce divergent antisense transcripts.

Next, we examined strand-specific RNA transcripts for the entire genome after treatment with ZC3H4 siRNA. To specifically focus on divergent antisense transcripts at the TSS, we chose 'non-overlapping' protein-coding genes ($n = 12\,550$) that are separated from neighboring genes by >10 kb because it is difficult to discriminate sense and antisense transcripts if genes are closely located. We analyzed transcripts with significant differential expression of genes with RPKM (reads per kilobase of transcript per million mapped reads) ≥1 at a ±2-kb region around the TSS in a strand-specific manner. We limited the analysis of transcripts to these 2-kb regions of the TSS to ensure that the transcripts likely represent transcripts of individual genes originating around the TSS (Figure 4D). A total of 587 differential transcripts were identified in sense ($n = 295$) and antisense ($n = 292$) directions after depletion of ZC3H4. Consistent with results shown in Figure 4C, 98% of antisense transcripts (285 out of 292) showed a preference for up-regulation. However, sense transcripts did not show a preference; of 295 genes, 147 were down-regulated and 148 were up-regulated. We divided the non-overlapping genes into two groups according to the production of divergent antisense transcripts under natural conditions (i.e. cells treated with non-specific control siRNA). Among genes without natural divergent antisense transcripts ($n = 10\,009$), ZC3H4 treatment led to differential expression of 285 genes, with sense transcripts ($n = 246$) predominating over antisense ($n = 39$) transcripts. Among sense transcripts, the proportions of down-regulated genes (120/246) and up-regulated genes (126/246) were similar. A small number of genes in this classification (39 of 10 009 genes) were found to produce de novo divergent antisense transcripts. Among genes that naturally produced divergent transcripts ($n = 2541$), 302 showed production of significantly differential sense ($n = 49$) and antisense ($n = 253$) transcripts upon ZC3H4 depletion. A similar number of genes exhibited down-regulated (27/49) and up-regulated (22/49) sense transcripts. Importantly, however, antisense transcripts showed a clear preference for up-regulation: ∼97% of genes (246/253) exhibited up-regulated antisense transcription. It should be noted that, among total genes with up-regulated antisense transcription ($n = 285$), ∼86% of genes (246/285) are with natural divergent antisense transcripts and ∼14% of genes (39/285) are without natural divergent antisense transcripts. This indicates that, in the absence of ZC3H4, divergent antisense transcription originating at the TSS becomes more profound in genes that generate natural divergent antisense transcripts.

Strand-specific transcription was further analyzed at the representative individual genes that we previously showed accumulated ZWC complex subunits (Figure 2A). Consistent with results presented in Figure 4D, snapshot images of strand-specific RNA-seq analyses of *GAPDH*, *EIF3K* and *HDAC6* genes, which naturally produce divergent antisense transcripts, showed increased divergent antisense transcription upon ZC3H4 depletion (Figure 4E). Notably, divergent antisense transcripts originating at TSSs of *GAPDH*, *EIF3K* and *HDAC6* genes extended into neighboring genes to become antisense (i.e. *NCAPD2* and *GATA1*) or sense (i.e. *MAP4K1*) transcripts of adjacent transcription units. To understand the mechanism underlying the increase in divergent antisense transcripts, we further investigated the distribution of RNAPII and histone modifications during early and productive elongation stages using ChIP-seq analyses (Figure 4F). We found that enrichment of total RNAPII and S5p RNAPII was not significantly changed in the absence of ZC3H4, suggesting that changes in RNAPII or the S5p form of RNAPII around TSSs are not the cause

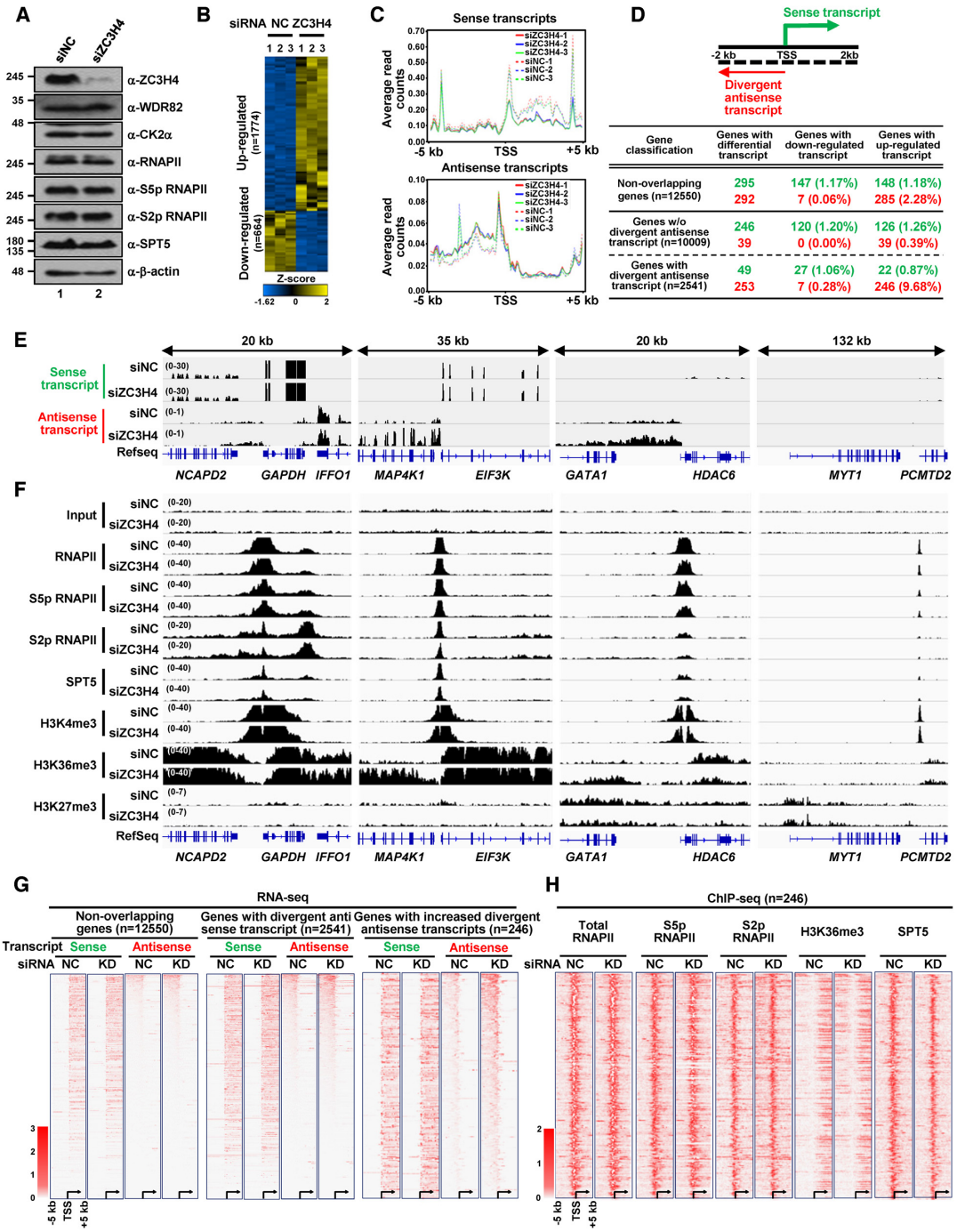

Figure 4. The ZWC complex regulates divergent antisense transcription. (**A**) Efficiency of siRNA-mediated ZC3H4 knockdown. Whole-cell lysates of HEK293T cells treated with non-specific control (NC) or ZC3H4 siRNA were immunoblotted with the indicated antibodies. (**B**) Differential gene expression analysis after ZC3H4 depletion. Heat map shows genes differentially expressed by more than two-fold (FDR <0.05) in NC versus ZC3H4 siRNA-treated cells. Three independent experiments were performed. (**C**) Strand-specific mapping analysis showing increased production of divergent antisense transcripts at the TSS after depletion of ZC3H4. Average read counts are shown for a ±5-kb region around the TSS for sense (top) and antisense (bottom) transcripts of down-regulated genes (n = 664) upon ZC3H4 depletion. (**D**) Increased production of divergent antisense transcripts at the TSS after ZC3H4 depletion. Differential transcripts with FDR <0.05 for genes with RPKM ≥1 at a ±2-kb region around the TSS were analyzed for total non-overlapping protein-coding genes separated by more than 10 kb from neighboring genes (n = 12 550) and further divided into genes without (n = 10 009) and with (n = 2541) natural divergent antisense transcripts. Numbers indicate up-regulated or down-regulated sense (green) and antisense (red) transcripts after ZC3H4 depletion in each group. Percentage is the number of genes relative to the total number of genes in each classification group. (**E** and **F**) Representative snapshot images of strand-specific RNA-seq (E) and ChIP-seq (F) at four genomic loci after depletion of ZC3H4. Note that the scale of read counts for sense transcripts is ∼30-times higher than that for antisense transcripts in (E). (**G**) Heat map analyses of strand-specific transcripts around the TSS. Read densities of sense and antisense transcripts are visualized for a ±5-kb region around the TSS in each group of genes. (**H**) Heat map analyses of ChIP-seq peaks for genes (n = 246) with up-regulated antisense transcripts. Signal intensities are shown for a ±5-kb region around TSSs of genes.

of increased divergent transcription. Similarly, H3K4me3 and H3K27me3 levels were unchanged, indicating that ZC3H4 depletion does not affect the early elongation stage of transcription. However, H3K36me3, which normally marks actively elongating transcription in the gene body, was significantly increased in the region where divergent antisense transcripts were produced, with concomitant enhancement of S2p RNAPII, a component of actively elongating transcriptional machinery, supporting the idea that the increased fraction of the productive elongation form of RNAPII moves in the opposite direction.

We further investigated divergent antisense transcription and the distribution of RNAPII and histone modifications in the genes analyzed in Figure 4D. We confirmed significant changes in divergent antisense transcription in the genes that naturally produced divergent antisense transcripts ($n = 2541$) compared with all non-overlapping genes ($n = 12\,550$), a change that became more profound in the group of genes with increased divergent antisense transcripts ($n = 246$) (Figure 4G). ChIP-seq analyses around TSSs in these 246 genes showed no noticeable changes in the distribution or level of total RNAPII or S5p RNAPII, suggesting that depletion of ZC3H4 does not affect RNAPII recruitment or early transcription elongation. Importantly, this analysis showed a significant increase in H3K36me3 and a slight increase in S2p RNAPII levels in the region where divergent antisense transcription occurred (Figure 4H, Supplementary Figure S4D), suggesting that the increase in antisense transcripts is related to the changed direction of RNAPII around TSSs. Taken together, these results indicate that levels of the productive elongation form of RNAPII moving in the opposite direction from the TSS are elevated in the absence of ZC3H4, suggesting a role for ZC3H4 in suppressing divergent antisense transcription.

The ZWC complex phosphorylates eight amino acid residues in the N-terminal acidic domain of SPT5

Increased levels of divergent antisense transcription in the absence of ZC3H4 led us to speculate that ZWC complex-mediated phosphorylation of transcriptional co-regulators is required for normal movement of RNAPII during the transition from early to productive elongation around the TSS. Among co-regulators, we chose DSIF, composed of SPT5 and SPT4 subunits, as a potential target of phosphorylation by the ZWC for several reasons. First, our ChIP-seq analyses found co-enrichment of ZWC complex subunits and SPT5 at TSSs (Figure 2A and B) with decreased TSS-localization of SPT5 after depletion of ZC3H4 (Figure 4F, Supplementary Figure S4D). Second, the N-terminal region of SPT5 contains several serine and tyrosine residues with neighboring acidic amino acids, conditions that meet the consensus target sequence of CK2 (22,49) (Figure 5A). Third, a recent study reported that SPT5 depletion in *Schizosaccharomyces pombe* causes a reduced rate of RNAPII transcription and increased levels of convergent and divergent antisense transcripts (33).

In an *in vitro* kinase assay employing purified DSIF (Supplementary Figure S5A), we found that the ZWC complex can phosphorylate SPT5 (Figure 5B, lanes 1–3). SPT5 phosphorylation was completely abolished in reactions with ZWC complex containing catalytically dead K68A mutant CK2α (Supplementary Figure S5B) or those treated with the CK2 kinase-specific inhibitor quinalizarin (Figure 5B, lanes 4 and 5), indicating CK2-mediated phosphorylation of SPT5. In support of this, phosphorylated SPT5 detected by autoradiography was recognized by an anti-phospho-CK2 substrate antibody (Figure 5C).

To determine the SPT5 regions phosphorylated by the ZWC complex, we purified DSIF derivatives (Supplementary Figure S5C) that contained SPT5 fragments (Figure 5A) and tested their phosphorylation by the ZWC complex. This analysis demonstrated selective phosphorylation of the N-terminal region encompassing amino acid residues 1–108 of SPT5 (Figure 5D). Importantly, this region is rich in amino acid residues surrounded by acidic amino acids corresponding to the consensus target sequence of CK2. To identify specific phosphorylation sites on SPT5, we purified a large number of DSIF mutants in which potential serine, threonine and tyrosine residues were replaced with alanine in various combinations and subjected them to *in vitro* kinase assays (data not shown). Ultimately, these analyses revealed that the ZWC complex phosphorylates eight amino acid residues (serine 2, 4, 7, 10, 19, 36 and tyrosine 54, 86) of SPT5. Specifically, *in vitro* kinase assays using purified mutants of the SPT5 1–270 fragment (Supplementary Figure S5D) and DSIF in which these eight amino acids were replaced with alanine (S6Y2→A) (Supplementary Figure S5E) showed no phosphorylation (Figure 5E and F). It should be noted that, unlike P-TEFb-mediated phosphorylation of the CTR of SPT5, this ZWC complex-mediated phosphorylation occurred in the N-terminal acidic domain of SPT5, a finding that has not been previously reported.

To examine ZWC complex-dependent DSIF phosphorylation in cells, we treated HEK293T cells with quinalizarin, immunoprecipitated cell extracts with an anti-SPT5 antibody, and monitored phosphorylation of SPT5 by immunoblotting with an anti-phospho-CK2 substrate antibody. We found that quinalizarin treatment abolished SPT5 phosphorylation (Figure 5G), indicating CK2-dependent phosphorylation of SPT5. A similar analysis using treatment with siRNA targeting ZC3H4 resulted in a significantly decreased level of SPT5 phosphorylation (Figure 5H), suggesting that ZWC complex formation is essential for CK2-mediated phosphorylation of SPT5 in cells. However, an *in vitro* kinase assay with subunit-deficient ZWC complexes showed that CK2 alone was able to phosphorylate SPT5 (Supplementary Figure S5F), possibly reflecting the lack of constraints in a cell-free system compared with that in cells, where the accessibility of free CK2 is normally controlled. Finally, to confirm phosphorylation sites on SPT5 in cells, we generated HEK293T cell lines containing a chromosome-integrated FLAG-SPT5 expression plasmid (see below for details) and immunoprecipitated cell extracts with an anti-FLAG antibody. An anti-phospho-CK2 substrate antibody did not detect SPT5 phosphorylation in cells that expressed S6Y2→A mutant SPT5 (Figure 5I).

Co-localization of ZWC complexes and SPT5 at TSSs (Figure 2) and direct interaction of the ZWC complex with S5p RNAPII (Figure 3) suggest that the ZWC complex is re-

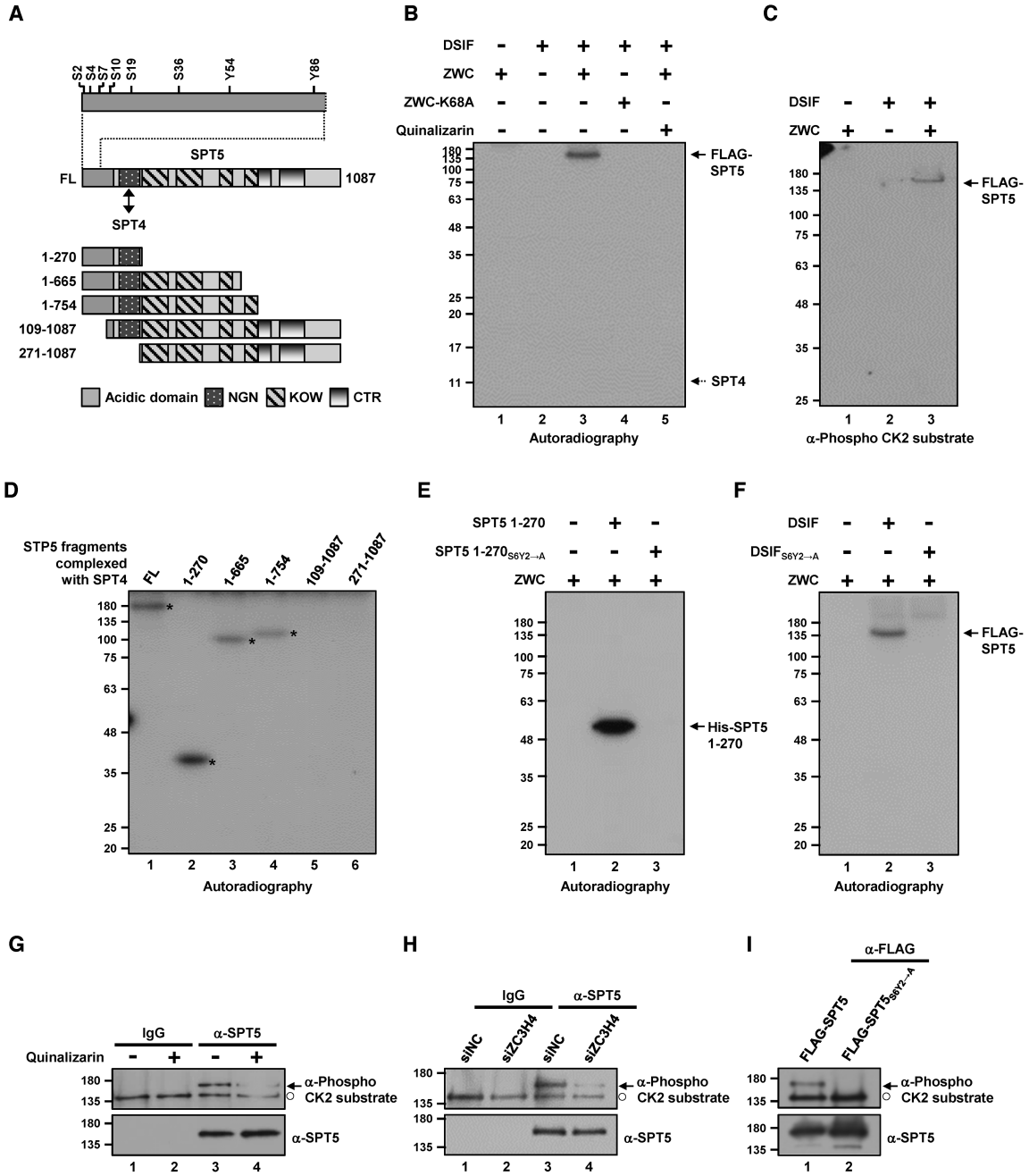

Figure 5. ZWC complex-mediated phosphorylation of eight serine and tyrosine amino acid residues in the N-terminal acidic domain of SPT5. (**A**) A schematic diagram of SPT5 and derived fragments with acidic, NGN (NusG N-terminal), KOW (Kyrpides, Ouzounis and Woese), and CTR domains. SPT4 associates with the NGN domain to form DSIF. Eight serine and tyrosine amino acids phosphorylated by the ZWC complex (identified in this study) are indicated in the enlarged N-terminal acidic domain of SPT5. Note that the N-terminal acidic domain contains many serine/threonine/tyrosine residues in addition to the eight amino acids phosphorylated by the ZWC complex. Numbers indicate amino acid residues. FL, full-length. (**B** and **C**) ZWC complex-mediated SPT5 phosphorylation *in vitro*. Purified DSIF (composed of FLAG-tagged SPT5 and untagged SPT4) was subjected to *in vitro* kinase assays with ZWC complexes containing wild-type (WT) or catalytically-dead mutant (K68A) CK2α in the presence and absence of quinalizarin, as indicated. Phosphorylation status was monitored by autoradiography (B) and immunoblotting with an anti-phospho-CK2 substrate antibody (C). Note that no autoradiographic signal was observed at the SPT4 position in (B). (**D**–**F**) ZWC complex-mediated phosphorylation of eight serine and tyrosine amino acid residues in the N-terminal acidic domain of SPT5 *in vitro*. *In vitro* kinase assays with the ZWC complex were performed on DSIFs composed of the indicated SPT5 fragments and SPT4 (D), WT and mutant (S6Y2→A, harboring alanine substitutions at eight ZWC complex-mediated phosphorylation sites) SPT5 1–270 fragments (E), and DSIFs containing WT and S6Y2→A mutant SPT5 (F). The phosphorylation status of SPT5 fragments was monitored by autoradiography. Phosphorylated SPT5 fragments are marked by asterisks in (D). (**G**–**I**) ZWC complex-mediated SPT5 phosphorylation in cells. HEK293T cells were treated with quinalizarin (G) or ZC3H4-targeted siRNA (H), as indicated, and whole-cell extracts were immunoprecipitated with anti-SPT5 antibody. Whole-cell extracts prepared from FLAG-tagged WT and S6Y2→A mutant SPT5 HEK293T cell lines were immunoprecipitated with an anti-FLAG antibody (I). SPT5 phosphorylation status was monitored by immunoblotting with an anti-phospho-CK2 substrate antibody. SPT5 is indicated by an arrow. A non-specific band recognized by the anti-phospho-CK2 substrate antibody is indicated by an empty circle.

cruited to TSSs by early elongating RNAPII, where it phosphorylates the N-terminal acidic domain of SPT5. In addition, decreased levels of SPT5 at TSSs after ZC3H4 depletion (Figure 4F, Supplementary Figure S4D) imply that the ZWC complex itself or ZWC complex-mediated phosphorylation stabilizes the chromatin association of SPT5 at TSSs.

Defects in ZWC complex-dependent phosphorylation of the N-terminal acidic domain of SPT5 increase divergent antisense transcription

Next, we examined whether ZWC complex-mediated phosphorylation of the N-terminal acidic domain of SPT5 regulates divergent antisense transcription. To this end, we selectively depleted endogenous SPT5 from HEK293T cell lines expressing FLAG-SPT5 from a chromosome-integrated plasmid and untagged SPT5 from endogenous genes (used for Figure 5I) by treating cells with siRNAs targeting the 3′ untranslated region (UTR) of endogenous SPT5 (thus avoiding interference from the expression of FLAG-SPT5) (Figure 6A). As expected, quantitative RT-PCR (RT-qPCR) analyses demonstrated that siRNA treatment substantially decreased endogenous SPT5 mRNA levels (Endo-SPT5; measured based on amplification of PCR products that span the region between the 5′ UTR and gene body region of the endogenous SPT5 gene) (Figure 6B, lanes 2 and 4 versus lanes 1 and 3, respectively), but slightly decreased total SPT5 mRNA levels (Total SPT5; measured based on amplification of PCR products spanning the gene body region of endogenous SPT5 and FLAG-SPT5 genes) (Figure 6B, lanes 6 and 8 versus lanes 5 and 7, respectively). Importantly, this strategy showed that a majority of SPT5 in cells was in the form of FLAG-SPT5, which was expressed at similar levels in WT and S6Y2→A mutant cell lines (Figure 6B, lane 6 versus lane 8). Selective depletion of endogenous SPT5 and similar FLAG-SPT5 levels in WT and S6Y2→A mutant FLAG-SPT5 cell lines were also confirmed by immunoblot analysis (Figure 6C).

Finally, to measure the levels of divergent antisense transcripts of *GAPDH*, *EIF3K* and *HDAC6* genes for which strand-specific RNA-seq analyses demonstrated increased divergent antisense transcription upon ZC3H4 knockdown (Figure 4E), we subjected these siRNA-treated cell lines to RT-qPCR analyses (Figure 6D). We found that, in the context of endogenous SPT5 knockdown, divergent antisense transcription of all tested genes was substantially increased in the S6Y2→A mutant FLAG-SPT5 cell line but not in the WT FLAG-SPT5 cell line (lanes 4 versus lanes 2). However, RT-qPCR measurements of sense transcript levels in gene body regions showed no significant changes (lanes 8 versus lanes 6). To rule out the off-target effect of siRNA, we repeated these experiments using another siRNA targeting the 3′ UTR of endogenous SPT5 (siSPT5 #2 in Figure 6A) and obtained consistent results (Supplementary Figure S6). Collectively, these results strongly suggest that defects in ZWC complex-mediated phosphorylation in the N-terminal acidic domain of SPT5 cause an increase in divergent antisense transcription from the TSS without large changes in the level of sense transcription.

DISCUSSION

Antisense transcription has emerged as an important process in the regulation of gene expression and chromatin structure (4,5). However, details of the molecular mechanisms underlying the biogenesis and regulation of antisense transcription remain incomplete. In this study, we report a novel ZWC complex composed of ZC3H4, WDR82 and CK2 subunits that preferentially localizes at TSSs of actively transcribing genes through a direct interaction with early elongating S5p RNAPII. We further demonstrated that the ZWC complex phosphorylates eight serine and tyrosine residues in the N-terminal acidic domain of SPT5, which is responsible for suppressing divergent antisense transcription. In accordance with a recent study reporting that Spt5 depletion causes an increase in divergent antisense transcription in fission yeast (33), we also found that ZC3H4 knockdown leads to increased divergent antisense transcription. However, unlike the study in yeast, which also reported increased convergent transcription, we found that convergent transcription was unaltered in human cells, a discrepancy that might reflect differences in organism-specific regulation of antisense transcription. Although the detailed mechanism by which the newly identified ZWC complex-mediated phosphorylation of the N-terminal acidic domain of SPT5 regulates divergent antisense transcription will require further investigation, our study provides evidence that, by being phosphorylated at different sites, SPT5 not only regulates promoter proximal pausing (P-TEFb-DSIF axis), but also regulates antisense transcription (ZWC-DSIF axis).

Simultaneous requirement for WDR82 and ZC3H4 for efficient binding to the Ser5-phosphorylated RNAPII CTD

Our peptide binding assay demonstrated that both WDR82 and ZC3H4 are required for binding of the ZWC complex to the S5p RNAPII CTD (Figure 3). In accordance with this observation, a recent study showed that association of Swd2 (a yeast homolog of WDR82) to the N-terminal region of Set1 is required for interaction of the Set1 complex with the S5p RNAPII CTD (50). Furthermore, it was reported that the PTW/PP1 complex, another WDR82-containing complex, dephosphorylates the S5p RNAPII CTD and that depletion of WDR82 weakens the dephosphorylation activity of the complex (14,15), suggesting that WDR82 may also participate in recognizing the S5p RNAPII CTD in conjunction with another subunit in the PTW/PP1 complex. These observations, taken together with our studies, suggest that WDR82 creates a new binding surface with its associating protein within the complex that enables the complex to bind to RNAPII. Additional investigations, ideally involving structural analyses, should provide important information on how WDR82 and its interacting protein collectively recognize the RNAPII CTD.

Regulation of antisense transcription by transcriptional co-regulators

In mammalian cells, the majority of promoters also direct transcription initiation in the opposite direction, and such

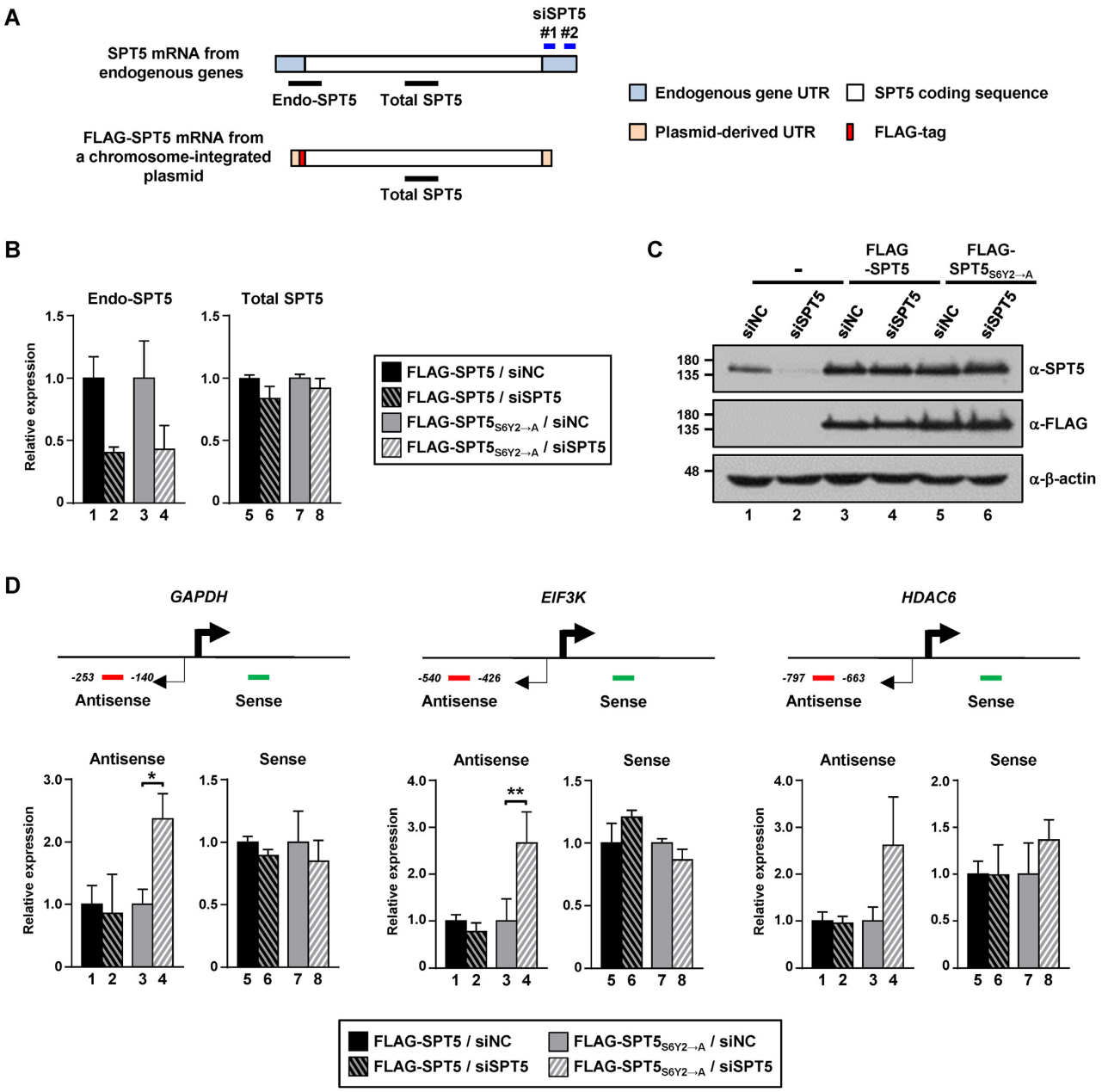

Figure 6. ZWC complex-mediated SPT5 phosphorylation regulates divergent antisense transcription. (**A**) Establishment of a cell line system expressing WT or S6Y2→A mutant FLAG-SPT5 with endogenous SPT5 depletion. Schematic representation of SPT5 mRNA from endogenous genes and FLAG-SPT5 mRNA from a chromosome-integrated plasmid indicating the SPT5 siRNAs targeting the 3′ UTR used to selectively deplete endogenous SPT5 (blue lines) and the RT-qPCR primers used to measure the levels of total and endogenous SPT5 mRNA (black lines). Note that the primers for endogenous SPT5 were designed to amplify the region between the 5′ UTR and gene body of SPT5 (Endo-SPT5), and are thus specific for endogenous SPT5 mRNA but not FLAG-SPT5 mRNA. Note also that siSPT5 #1 and #2 were used for this figure and Supplementary Figure S6, respectively. (**B** and **C**) WT and S6Y2→A mutant FLAG-SPT5 HEK293T cell lines were treated with either control or SPT5 3′ UTR-targeted (#1) siRNAs as indicated. mRNA levels of total and endogenous SPT5 were measured by RT-qPCR analyses. Error bars denote standard deviations from three independent experiments (B). Total and FLAG-tagged SPT5 protein levels were measured by immunoblot analyses with the indicated antibodies (C). (**D**) Schematic representation of *GAPDH*, *EIF3K* and *HDAC6* loci and amplicons used for qPCR measurement of the levels of sense and divergent antisense transcripts. Numbers next to the amplicons for divergent antisense transcript indicate base pairs from the TSS of the gene (top). Sense and divergent antisense transcript levels in either control or SPT5 3′ UTR-targeted (#1) siRNA-treated cell lines were measured by RT-qPCR analysis. All transcript levels were normalized to those of β-actin mRNA. Error bars denote the standard deviations from three independent experiments (*P <0.05, **P <0.01; Student's t-test) (bottom).

bidirectional transcription appears to be a hallmark of active promoters (8,51). It is plausible that antisense transcription is regulated by transcriptional co-regulators; however, no such mechanism has been fully elucidated. In this context, our study showed that the ZWC complex is preferentially enriched at TSSs of actively transcribing genes (Figure 2) and that failure of ZWC complex-mediated phosphorylation of the N-terminal acidic domain of SPT5 causes a more profound increase in divergent antisense transcripts from promoters that naturally produce divergent antisense transcripts (Figures 4 and 6), thus demonstrating a suppressive role of the co-regulator DSIF in the regulation of antisense transcription. Interestingly, ZC3H4 depletion caused increased antisense transcription of a subset of genes, but how the ZWC complex preferentially affects these genes is an unanswered question. One possible explanation is that the ZWC complex interacts with DNA or RNA in a sequence-specific manner. In support of this, the N-terminal region of ZC3H4 contains an RG (arginine and glycine)-rich sequence and a tandem repeat of a CCCH-type zinc finger domain, both of which have the potential to interact with nascent RNA. In a related observation, suppressor of sable (Su(s)), a *Drosophila* ortholog of ZC3H4, was reported to bind to RNA with high affinity and limited specificity (52).

Antisense transcription in the majority of mammalian promoters is less frequent relative to sense transcription because of inefficient elongation and early termination upstream of the TSS. (53). On a similar note, very recent papers have reported that ZC3H4 plays a suppressive role in antisense transcription by causing early termination of antisense transcripts (18,19). Our study further revealed that, rather than acting alone, ZC3H4 forms a complex with WDR82 and CK2 to phosphorylate SPT5, which is responsible for suppression of divergent antisense transcription. In addition to SPT5, as demonstrated in this study, CK2 has been reported to phosphorylate additional transcription-elongation factors, including SPT6 and the PAF complex (27,28,54). Importantly, CK2-dependent SPT6 phosphorylation was reported to prevent antisense transcription (27,28). These results suggest that the ZWC complex may regulate RNAPII processivity and transcription elongation during the transition from early to productive stages by phosphorylating multiple transcription-elongation factors. In relation to the direct demonstration of the involvement of CK2 in the suppression of divergent antisense transcription, we found that inactivation of CK2 by treating cells with siRNA targeting CK2α or with quinalizarin showed no significant changes in the levels of divergent antisense transcripts from the tested genes (data not shown). However, it is noteworthy that CK2 was reported to have the potential to phosphorylate over 300 proteins besides transcription-elongation factors in mammalian cells (25). Inactivation of CK2 results in the failure of phosphorylation of a large number of proteins, which may directly or indirectly affect the final readout of antisense transcription.

The Integrator complex has also been implicated in early termination events. In comparing the effects of ZC3H4 and INTS1 (a subunit of the Integrator complex) depletions on non-coding RNA expression, a recent paper showed that each ZC3H4 and INTS1 knockdown increased antisense transcription from several hundred genes, but there was little overlap between these genes (19). This result suggests that the ZWC and Integrator complexes function in separate pathways during the early termination of antisense transcripts.

Our study revealed a regulatory function of the newly identified ZWC complex-mediated phosphorylation of the N-terminal acidic domain of SPT5 (ZWC-DSIF axis) in divergent antisense transcription. It has also been proposed that P-TEFb-mediated phosphorylation of the repetitive CTR of SPT5 (P-TEFb-DSIF axis) is important for regulation of processive elongation (31). These observations suggest that the activity of SPT5 can vary depending on which of its sites are phosphorylated. In relation to the role of ZWC complex-mediated phosphorylation of SPT5, it is worth noting that the N-terminal acidic domain of SPT5 is essentially an intrinsically disordered region (IDR). Acidic IDRs are often found in N-terminal or C-terminal regions of histone chaperones (55). Interestingly, the N-terminal acidic domain of SPT5 is conserved only in eukaryotes in which a nucleosome structure exists. A recent structural study showed that the N-terminal acidic domain of SPT5 is positioned close to H2A–H2B in a neighboring nucleosome (56); thus, ZWC complex-mediated SPT5 phosphorylation may affect interactions between DSIF and nucleosomes to regulate RNAPII processivity. This scenario is supported by CK2-mediated phosphorylation of Spt6. Studies in yeast have reported that CK2 phosphorylates the N-terminal IDR of Spt6, which is required for nucleosome occupancy at the TSS, and thereby prevents aberrant antisense transcription and enforces transcriptional directionality (27,28). The nucleosomal barrier and RNAPII processivity are closely related to regulation of antisense transcription. A recent study showed that slow transcription leads to an increase in transcription-elongation markers, including S2p RNAPII CTD and H3K36me3, at the 5′ ends of the genes, a finding that is correlated with increased divergent transcription (57). Consistent with this, we found that ZC3H4 depletion resulted in higher levels of S2p RNAPII CTD and H3K36me3 upstream of the TSS (Figure 4). Similarly, it was recently reported that depletion of SPT5 causes delayed and reduced early transcription (32). These reports raise the possibility that ZWC complex-mediated SPT5 phosphorylation regulates antisense transcription by controlling the speed of early transcription.

Effects of divergent antisense transcripts

Divergent antisense transcripts regulate gene expression by binding to proteins and nearby RNA in a sequence-specific manner (4,5). For example, interactions of divergent antisense transcripts with polycomb repressive complex 2 (PRC2) (58) and Mediator (59) regulate sense transcription by altering histone modifications and the chromatin structure of neighboring promoters. Also, altered histone methylation patterns generated by overlapping antisense transcripts can affect the expression of adjacent genes by recruiting histone deacetylases like Rpd3S and Set3C (60,61). In addition, divergent antisense transcripts may directly impact neighboring genes by transcribing into their transcription unit. A closer investigation of our RNA-seq data showed that the read density of divergent antisense

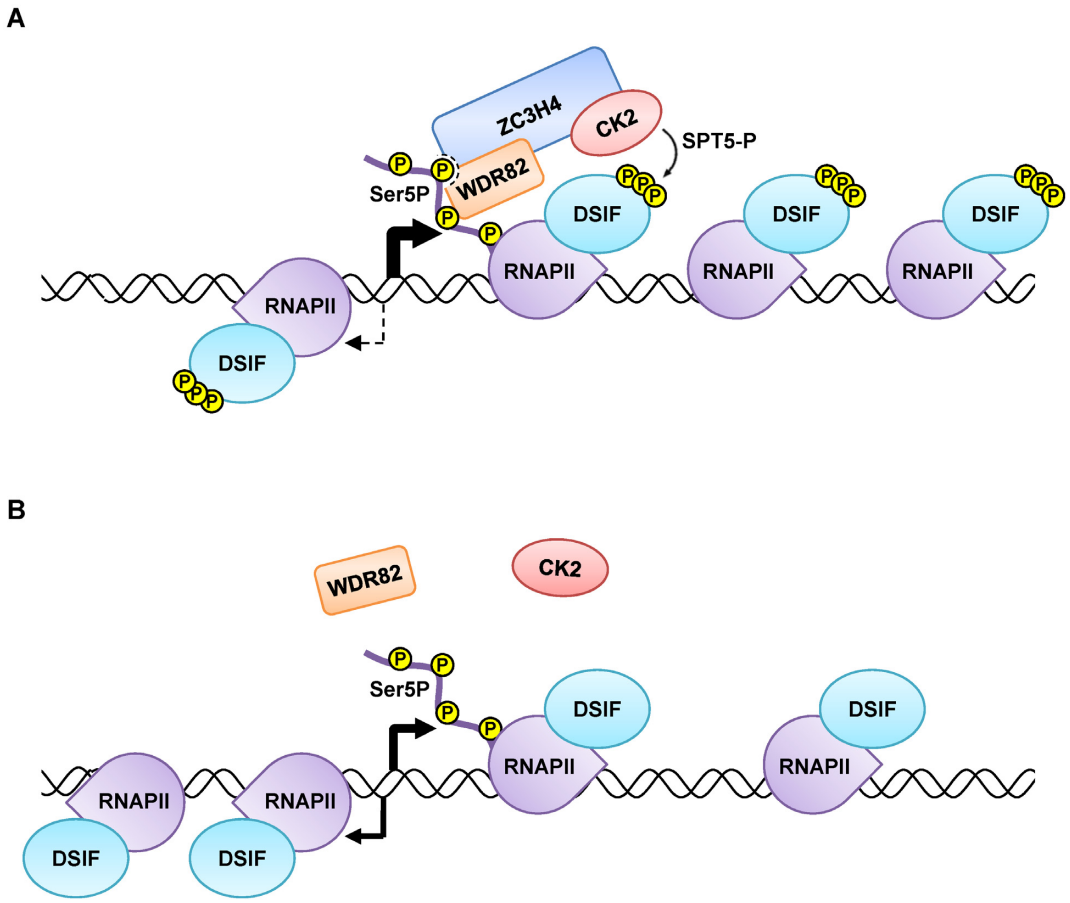

Figure 7. A mechanistic model of the suppression of divergent antisense transcription by ZWC complex-mediated SPT5 phosphorylation. (**A**) The ZWC complex is recruited to the TSSs of active genes through direct, simultaneous binding of ZC3H4 and WDR82 to the S5p RNAPII CTD. This allows CK2 to phosphorylate eight newly identified serine and tyrosine residues in the N-terminal acidic domain of SPT5, a subunit of DSIF, which may facilitate the transcriptional transition from early to productive elongation through nucleosomes (not depicted) in the sense transcription direction. How SPT5 N-terminal phosphorylation supports this transition needs to be further investigated. Although not shown, the ZWC complex may phosphorylate additional transcriptional co-regulators that also participate in the transition from early to productive transcription elongation. (**B**) In the absence of ZC3H4, the ZWC complex is destabilized, and CK2 is unable to phosphorylate SPT5. Failure to phosphorylate the N-terminal acidic domain of SPT5 prevents SPT5 from fully supporting the transition from early to productive elongation in the sense transcription direction; this increases the RNAPII population that moves in the opposite direction and results in increased divergent antisense transcription and decreased sense transcription. This phenomenon is more profound at promoters that naturally produce divergent antisense transcripts. Failed phosphorylation also increases S2p RNAPII CTD and H3K36me3 marks upstream of the TSS (not depicted). Note that the total number of RNAPII molecules around the TSS is unchanged.

transcripts gradually decreased at a distance from the TSS, but transcripts are lengthy (>10 kb) to invade into neighboring genes. (Figure 4E, see divergent antisense transcripts originating from *EIF3K* and *HDAC6* TSSs). Given the typical splicing pattern of mRNAs, antisense transcripts originating from the *EIF3K* TSS become sense transcripts of the *MAP4K1* gene. Also, antisense transcripts originating from the *HDAC6* TSS become antisense transcripts of the *GATA1* gene together with aberrant histone modifications and a decrease in *HDAC6* sense transcripts. Thus, production of divergent antisense transcripts could change levels of the transcripts of neighboring genes. Misregulation of antisense transcription has been reported to cause various diseases, including cancers (4). ZC3H4 also has been implicated in silica-induced epithelial-mesenchymal transition (62) and early embryogenesis (63). We also observed retarded cell proliferation in ZC3H4-depleted cells (data not shown). If these defects are related to aberrant divergent antisense transcript production, ZC3H4 could be a potential biomarker for disease pathogenesis and a target for therapeutic development.

A model of ZWC complex-mediated SPT5 phosphorylation and suppression of divergent antisense transcription

By newly identifying ZWC complex-mediated phosphorylation of the N-terminal acidic domain of SPT5, our biochemical and genome-wide approaches provide valuable insight into the molecular mechanism underlying SPT5 regulation of divergent antisense transcription (Figure 7). ZC3H4, WDR82 and CK2 form a stable ZWC complex that can be recruited to the TSS of active genes during the early elongation stage through direct binding of both ZC3H4 and WDR82 to the S5p RNAPII CTD. In turn, CK2-mediated phosphorylation of eight amino acids in the N-terminal acidic domain of SPT5 suppresses the di-

vergent antisense transcription that originates around the TSS in an opposite direction to sense transcription during the transition from early to productive elongation. How ZWC complex-mediated SPT5 phosphorylation regulates divergent antisense transcription will require further investigation. One possible scenario based on the function of DSIF as a transcription-elongation factor is that phosphorylated SPT5 in the N-terminal acidic domain accelerates the transition of RNAPII from early to productive elongation and facilitates movement of RNAPII in the sense direction by increasing transcription-elongation efficiency through nucleosomes. In addition to SPT5, other transcription-elongation factors and histone chaperones can also be phosphorylated by the ZWC complex and could act together to control transcription elongation. In the absence of ZC3H4, the ZWC complex is destabilized and CK2 cannot be recruited to the TSS. Unphosphorylated SPT5 in the N-terminal acidic domain is no longer able to fully support sense transcription, resulting in increased divergent antisense transcription in the opposite direction. Additional biochemical and structural analyses, ideally involving forms of SPT5 phosphorylated in the N-terminal acidic domain and a biochemically defined eukaryotic transcription system, should provide detailed information relevant to this model. The mechanistic insights into ZWC complex-mediated SPT5 phosphorylation provided here set the stage for a detailed investigation of the regulation of antisense transcription during the transition from early to productive elongation in mammalian cells.

DATA AVAILABILITY

All ChIP-seq and RNA-seq data in this publication have been deposited in the Gene Expression Omnibus with the accession number GSE186809.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Robert G. Roeder (Rockefeller University) for anti-WDR82 antibody. We also thank Dr Huihuang Yan (Mayo Clinic) for initial data analysis.

FUNDING

National Research Foundation of Korea [NRF-2019R1A2C2090830 and NRF-2018R1A5A1024261 to J.K.]; National Institutes of Health [P01DK068055-Core B to J.H.L.]; Mayo Clinic Center for Individualized Medicine (to J.H.L.). Funding for open access charge: National Research Foundation of Korea [NRF-2019R1A2C2090830].
Conflict of interest statement. None declared.

REFERENCES

1. Katayama,S., Tomaru,Y., Kasukawa,T., Waki,K., Nakanishi,M., Nakamura,M., Nishida,H., Yap,C.C., Suzuki,M., Kawai,J. *et al.* (2005) Antisense transcription in the mammalian transcriptome. *Science*, **309**, 1564–1566.
2. He,Y., Vogelstein,B., Velculescu,V.E., Papadopoulos,N. and Kinzler,K.W. (2008) The antisense transcriptomes of human cells. *Science*, **322**, 1855–1857.
3. David,L., Huber,W., Granovskaia,M., Toedling,J., Palm,C.J., Bofkin,L., Jones,T., Davis,R.W. and Steinmetz,L.M. (2006) A high-resolution map of transcription in the yeast genome. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 5320–5325.
4. Barman,P., Reddy,D. and Bhaumik,S.R. (2019) Mechanisms of antisense transcription initiation with implications in gene expression, genomic integrity and disease pathogenesis. *Noncoding RNA*, **5**, 11.
5. Pelechano,V. and Steinmetz,L.M. (2013) Gene regulation by antisense transcription. *Nat. Rev. Genet.*, **14**, 880–893.
6. Esteller,M. (2011) Non-coding RNAs in human disease. *Nat. Rev. Genet.*, **12**, 861–874.
7. Wahlestedt,C. (2013) Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nat. Rev. Drug Discov.*, **12**, 433–446.
8. Seila,A.C., Core,L.J., Lis,J.T. and Sharp,P.A. (2009) Divergent transcription: a new feature of active promoters. *Cell Cycle*, **8**, 2557–2564.
9. Latge,G., Poulet,C., Bours,V., Josse,C. and Jerusalem,G. (2018) Natural antisense transcripts: molecular mechanisms and implications in breast cancers. *Int. J. Mol. Sci.*, **19**, 123.
10. Lee,J.H. and Skalnik,D.G. (2008) Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol. Cell. Biol.*, **28**, 609–618.
11. Wu,M., Wang,P.F., Lee,J.S., Martin-Brown,S., Florens,L., Washburn,M. and Shilatifard,A. (2008) Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.*, **28**, 7337–7344.
12. Austenaa,L.M., Barozzi,I., Simonatto,M., Masella,S., Della Chiara,G., Ghisletti,S., Curina,A., de Wit,E., Bouwman,B.A., de Pretis,S. *et al.* (2015) Transcription of mammalian cis-regulatory elements is restrained by actively enforced early termination. *Mol. Cell*, **60**, 460–474.
13. Cortazar,M.A., Sheridan,R.M., Erickson,B., Fong,N., Glover-Cutter,K., Brannan,K. and Bentley,D.L. (2019) Control of RNA Pol II speed by PNUTS-PP1 and Spt5 dephosphorylation facilitates termination by a "sitting duck torpedo" mechanism. *Mol. Cell*, **76**, 896–908.
14. Landsverk,H.B., Sandquist,L.E., Bay,L.T.E., Steurer,B., Campsteijn,C., Landsverk,O.J.B., Marteijn,J.A., Petermann,E., Trinkle-Mulcahy,L. and Syljuasen,R.G. (2020) WDR82/PNUTS-PP1 prevents transcription-replication conflicts by promoting RNA polymerase II degradation on chromatin. *Cell Rep.*, **33**, 108469.
15. Lee,J.H., You,J., Dobrota,E. and Skalnik,D.G. (2010) Identification and characterization of a novel human PP1 phosphatase complex. *J. Biol. Chem.*, **285**, 24466–24476.
16. Schapira,M., Tyers,M., Torrent,M. and Arrowsmith,C.H. (2017) WD40 repeat domain proteins: a novel target class?*Nat. Rev. Drug Discov.*, **16**, 773–786.
17. Jain,B.P. and Pandey,S. (2018) WD40 repeat proteins: signalling scaffold with diverse functions. *Protein J.*, **37**, 391–406.
18. Austenaa,L.M.I., Piccolo,V., Russo,M., Prosperini,E., Polletti,S., Polizzese,D., Ghisletti,S., Barozzi,I., Diaferia,G.R. and Natoli,G. (2021) A first exon termination checkpoint preferentially suppresses extragenic transcription. *Nat. Struct. Mol. Biol.*, **28**, 337–346.
19. Estell,C., Davidson,L., Steketee,P.C., Monier,A. and West,S. (2021) ZC3H4 restricts non-coding transcription in human cells. *Elife*, **10**, e67305.
20. Litchfield,D.W. (2003) Protein kinase CK2: structure, regulation and role in cellular decisions of life and death. *Biochem. J.*, **369**, 1–15.
21. St-Denis,N.A. and Litchfield,D.W. (2009) Protein kinase CK2 in health and disease: from birth to death: the role of protein kinase CK2 in the regulation of cell proliferation and survival. *Cell. Mol. Life Sci.*, **66**, 1817–1829.
22. Meggio,F., Marin,O. and Pinna,L.A. (1994) Substrate specificity of protein kinase CK2. *Cell Mol. Biol. Res.*, **40**, 401–409.
23. Vilk,G., Weber,J.E., Turowec,J.P., Duncan,J.S., Wu,C., Derksen,D.R., Zien,P., Sarno,S., Donella-Deana,A., Lajoie,G. *et al.* (2008) Protein kinase CK2 catalyzes tyrosine phosphorylation in mammalian cells. *Cell. Signal.*, **20**, 1942–1951.

24. Basnet,H., Su,X.B., Tan,Y., Meisenhelder,J., Merkurjev,D., Ohgi,K.A., Hunter,T., Pillus,L. and Rosenfeld,M.G. (2014) Tyrosine phosphorylation of histone H2A by CK2 regulates transcriptional elongation. *Nature*, **516**, 267–271.
25. Bian,Y., Ye,M., Wang,C., Cheng,K., Song,C., Dong,M., Pan,Y., Qin,H. and Zou,H. (2013) Global screening of CK2 kinase substrates by an integrated phosphoproteomics workflow. *Sci. Rep.*, **3**, 3460.
26. Filhol,O. and Cochet,C. (2009) Protein kinase CK2 in health and disease: cellular functions of protein kinase CK2: a dynamic affair. *Cell. Mol. Life Sci.*, **66**, 1830–1839.
27. Dronamraju,R., Kerschner,J.L., Peck,S.A., Hepperla,A.J., Adams,A.T., Hughes,K.D., Aslam,S., Yoblinski,A.R., Davis,I.J., Mosley,A.L. *et al.* (2018) Casein kinase II phosphorylation of Spt6 enforces transcriptional fidelity by maintaining Spn1-Spt6 interaction. *Cell Rep.*, **25**, 3476–3489.
28. Gouot,E., Bhat,W., Rufiange,A., Fournier,E., Paquet,E. and Nourani,A. (2018) Casein kinase 2 mediated phosphorylation of Spt6 modulates histone dynamics and regulates spurious transcription. *Nucleic Acids Res.*, **46**, 7612–7630.
29. Decker,T.M. (2021) Mechanisms of transcription elongation factor DSIF (Spt4-Spt5). *J. Mol. Biol.*, **433**, 166657.
30. Dollinger,R. and Gilmour,D.S. (2021) Regulation of promoter proximal pausing of RNA polymerase II in metazoans. *J. Mol. Biol.*, **433**, 166897.
31. Yamada,T., Yamaguchi,Y., Inukai,N., Okamoto,S., Mura,T. and Handa,H. (2006) P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol. Cell*, **21**, 227–237.
32. Diamant,G., Bahat,A. and Dikstein,R. (2016) The elongation factor Spt5 facilitates transcription initiation for rapid induction of inflammatory-response genes. *Nat. Commun.*, **7**, 11547.
33. Shetty,A., Kallgren,S.P., Demel,C., Maier,K.C., Spatt,D., Alver,B.H., Cramer,P., Park,P.J. and Winston,F. (2017) Spt5 plays vital roles in the control of sense and antisense transcription elongation. *Mol. Cell*, **66**, 77–88.
34. Kim,J. and Roeder,R.G. (2011) Nucleosomal H2B ubiquitylation with purified factors. *Methods*, **54**, 331–338.
35. Zhong,J., Ye,Z., Lenz,S.W., Clark,C.R., Bharucha,A., Farrugia,G., Robertson,K.D., Zhang,Z., Ordog,T. and Lee,J.H. (2017) Purification of nanogram-range immunoprecipitated DNA in ChIP-seq application. *BMC Genomics*, **18**, 985.
36. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
37. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
38. 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.
39. Zang,C., Schones,D.E., Zeng,C., Cui,K., Zhao,K. and Peng,W. (2009) A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics*, **25**, 1952–1958.
40. Shen,L., Shao,N., Liu,X. and Nestler,E. (2014) ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*, **15**, 284.
41. Yu,G., Wang,L.G. and He,Q.Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383.
42. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.
43. Bustin,S.A., Benes,V., Garson,J.A., Hellemans,J., Huggett,J., Kubista,M., Mueller,R., Nolan,T., Pfaffl,M.W., Shipley,G.L. *et al.* (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.*, **55**, 611–622.
44. Xing,Y., Yu,T., Wu,Y.N., Roy,M., Kim,J. and Lee,C. (2006) An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs. *Nucleic Acids Res.*, **34**, 3150–3160.
45. Kovaka,S., Zimin,A.V., Pertea,G.M., Razaghi,R., Salzberg,S.L. and Pertea,M. (2019) Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.*, **20**, 278.
46. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
47. Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.*, **57**, 289–300.
48. Lee,J.H., Tate,C.M., You,J.S. and Skalnik,D.G. (2007) Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J. Biol. Chem.*, **282**, 13419–13428.
49. Meggio,F. and Pinna,L.A. (2003) One-thousand-and-one substrates of protein kinase CK2? *FASEB J.*, **17**, 349–368.
50. Bae,H.J., Dubarry,M., Jeon,J., Soares,L.M., Dargemont,C., Kim,J., Geli,V. and Buratowski,S. (2020) The Set1 N-terminal domain and Swd2 interact with RNA polymerase II CTD to recruit COMPASS. *Nat. Commun.*, **11**, 2181.
51. Core,L.J., Waterfall,J.J. and Lis,J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, **322**, 1845–1848.
52. Murray,M.V., Turnage,M.A., Williamson,K.J., Steinhauer,W.R. and Searles,L.L. (1997) The Drosophila suppressor of sable protein binds to RNA and associates with a subset of polytene chromosome bands. *Mol. Cell. Biol.*, **17**, 2291–2300.
53. Wu,X. and Sharp,P.A. (2013) Divergent transcription: a driving force for new gene origination? *Cell*, **155**, 990–996.
54. Bedard,L.G., Dronamraju,R., Kerschner,J.L., Hunter,G.O., Axley,E.D., Boyd,A.K., Strahl,B.D. and Mosley,A.L. (2016) Quantitative analysis of dynamic protein interactions during transcription reveals a role for casein kinase II in polymerase-associated factor (PAF) complex phosphorylation and regulation of histone H2B monoubiquitylation. *J. Biol. Chem.*, **291**, 13410–13420.
55. Warren,C. and Shechter,D. (2017) Fly fishing for histones: catch and release by histone chaperone intrinsically disordered regions and acidic stretches. *J. Mol. Biol.*, **429**, 2401–2426.
56. Ehara,H., Yokoyama,T., Shigematsu,H., Yokoyama,S., Shirouzu,M. and Sekine,S.I. (2017) Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science*, **357**, 921–924.
57. Fong,N., Saldi,T., Sheridan,R.M., Cortazar,M.A. and Bentley,D.L. (2017) RNA Pol II dynamics modulate co-transcriptional chromatin modification, CTD phosphorylation, and transcriptional direction. *Mol. Cell*, **66**, 546–557.
58. Pontier,D.B. and Gribnau,J. (2011) Xist regulation and function explored. *Hum. Genet.*, **130**, 223–236.
59. Luo,S., Lu,J.Y., Liu,L., Yin,Y., Chen,C., Han,X., Wu,B., Xu,R., Liu,W., Yan,P. *et al.* (2016) Divergent lncRNAs regulate gene expression and lineage differentiation in pluripotent cells. *Cell Stem Cell*, **18**, 637–652.
60. Kim,J.H., Lee,B.B., Oh,Y.M., Zhu,C., Steinmetz,L.M., Lee,Y., Kim,W.K., Lee,S.B., Buratowski,S. and Kim,T. (2016) Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. *Nat. Commun.*, **7**, 13534.
61. Kim,T., Xu,Z., Clauder-Munster,S., Steinmetz,L.M. and Buratowski,S. (2012) Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell*, **150**, 1158–1169.
62. Jiang,R., Zhou,Z., Liao,Y., Yang,F., Cheng,Y., Huang,J., Wang,J., Chen,H., Zhu,T. and Chao,J. (2019) The emerging roles of a novel CCCH-type zinc finger protein, ZC3H4, in silica-induced epithelial to mesenchymal transition. *Toxicol. Lett.*, **307**, 26–40.
63. Su,J., Miao,X., Archambault,D., Mager,J. and Cui,W. (2020) ZC3H4, a novel CCCH-type zinc finger protein, is essential for early embryogenesis in mice. *Biol. Reprod.*, **104**, 325–335.