7712–7727 Nucleic Acids Research, 2020, Vol. 48, No. 14
doi: 10.1093/nar/gkaa514
Published online 17 June 2020

CDK12 globally stimulates RNA polymerase II transcription elongation and carboxyl-terminal domain phosphorylation

Michael Tellier[1,*,†], Justyna Zaborowska[1,†], Livia Caizzi[2], Eusra Mohammad[2], Taras Velychko[2], Björn Schwalb[2], Ivan Ferrer-Vicens[1], Daniel Blears[3], Takayuki Nojima[1], Patrick Cramer[2] and Shona Murphy[1,*]

[1]Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK, [2]Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology, Am Fassberg 11, 37077 Göttingen, Germany and [3]Mechanisms of Transcription Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

Received April 07, 2020; Revised June 02, 2020; Editorial Decision June 03, 2020; Accepted June 04, 2020

ABSTRACT

Cyclin-dependent kinase 12 (CDK12) phosphorylates the carboxyl-terminal domain (CTD) of RNA polymerase II (pol II) but its roles in transcription beyond the expression of DNA damage response genes remain unclear. Here, we have used TT-seq and mNET-seq to monitor the direct effects of rapid CDK12 inhibition on transcription activity and CTD phosphorylation in human cells. CDK12 inhibition causes a genome-wide defect in transcription elongation and a global reduction of CTD Ser2 and Ser5 phosphorylation. The elongation defect is explained by the loss of the elongation factors LEO1 and CDC73, part of PAF1 complex, and SPT6 from the newly-elongating pol II. Our results indicate that CDK12 is a general activator of pol II transcription elongation and indicate that it targets both Ser2 and Ser5 residues of the pol II CTD.

INTRODUCTION

The phosphorylation status of the C-terminal domain (CTD) of RNA polymerase II (pol II) plays a major role in expression of non-coding and protein-coding genes (1). In humans, the CTD comprises 52 repeats of the consensus heptapeptide Tyr1/Ser2/Pro3/Thr4/Ser5/Pro6/Ser7 and is reversibly phosphorylated by several kinases during transcription (1). CDK12 is the metazoan ortholog of the *Saccharomyces cerevisiae* CTD Ser2 kinase, Ctk1 (2,3). Knockdown of CDK12 results in loss of CTD Ser2 phosphorylation in *Drosophila*, *Caenorhabditis elegans* and human cells (3–10). In addition, treatment with the CDK12/CDK13 inhibitor THZ531 in Jurkat cells affects pol II CTD Ser2 phosphorylation in a concentration-dependent manner (7). Analog-sensitive (as) kinase technology has also been used to demonstrate rapid loss of Ser2 phosphorylation after inhibiting the *Schizosaccharomyces pombe* CDK12 homologue Lsk1 (11). However, knockdown of CDK12 in HCT116 cells or inhibition of CDK12^{as} in HeLa or HCT116 cells does not always cause loss of Ser2P (12,13), and detecting loss of Ser2P after CDK12 inhibition appears to depend on the anti-Ser2P antibody used (14). In addition, CDK12 mainly phosphorylates Ser5 *in vitro* (15,16).

CDK12 forms a complex with Cyclin K and is recruited to transcription units of pol II–transcribed genes (7,17–19) and enhancers (7). Blazek *et al.* have shown that depletion of CDK12 results in reduced expression of a small subset of genes (3.9%), enriched for genes involved in the DNA damage response (DDR), such as BRCA1 and ATR (4). Knockdown of Cyclin K/CDK12 complexes also causes elevated sensitivity to DNA damage inducing agents and accumulation of the DNA damage biomarker γ-H2AX (4). In addition, CDK12(−/−) embryos have reduced expression of genes involved in the DDR (20) and mutations that impair CDK12 kinase activity are associated with misregulation of expression of DNA repair genes and ovarian carcinoma (21,22). Downregulation of key DDR genes has also been observed by RNA-seq analysis following CDK12 depletion (12) and by microarray analysis after treatment of cells with THZ531 (7). In line with these findings, CDK12 is frequently mutated in high-grade serous ovarian carcinoma (HGSOC) (23). Interestingly, small molecule inhibitors of CDK12 also affect translation (24).

Pol II ChIP-seq and transient transcriptome sequencing (TT-seq) analyses suggest that elongation of transcription of genes sensitive to THZ531 is defective (7,25). In addi-

*To whom correspondence should be addressed. Tel: +44 1865 275583; Fax: +44 1865 255515; Email: shona.murphy@path.ox.ac.uk
Correspondence may be also addressed to Michael Tellier. Tel: +44 1865 275583; Fax: +44 1865 275515; Email: michael.tellier@path.ox.ac.uk
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

tion, it has been recently shown in mES cells, human tumours and mouse and human cell lines that downregulation of DDR genes following depletion of CDK12 or inhibition of CDK12/CDK13 results from the activation of intronic poly(A) sites, leading to the production of prematurely terminated gene products (10,25). However, the primary effect of selective inhibition of CDK12 activity on transcription remained unclear.

To better understand the function of the kinase activity of CDK12 in transcription, we have used CRISPR/Cas9 gene engineering to mutate the endogenous CDK12 gene in HEK293 cells to produce an analog-sensitive CDK12 ($CDK12^{as}$) that can be selectively inhibited by bulky ATP analogs (26). We have combined the use of this $CDK12^{as}$ cell line with two complementary functional genomics methods: TT-seq (27) and native elongating transcript sequencing for mammalian cells (mNET-seq) (28,29). TT-seq combines 4sU-metabolic labeling with RNA fragmentation to monitor synthesis of newly transcribed RNAs genome-wide. mNET-seq isolates and sequences the RNA in the active site of immunoprecipitated pol II to provide single-nucleotide resolution of pol II occupancy and CTD modification status.

We find that short-term inhibition of CDK12 causes a global decrease of pol II elongation velocity as measured by TT-seq and loss of phosphorylation of Ser2 and Ser5 of the pol II CTD as measured by mNET-seq and western blot. Elongation of transcription of the vast majority of pol II-transcribed genes is affected. CDK12 inhibition also causes the loss of key elongation factors from pol II transcription units, identifying CDK12 as a global regulator of pol II CTD Ser2 and Ser5 phosphorylation and transcription elongation in human cells.

MATERIALS AND METHODS

Cell culture

HEK293 cells were obtained from ATCC (ATCC® CRL-1573™). HEK293 parental cells and $CDK12^{as}$ HEK293 cells were grown in DMEM medium supplemented with 10% foetal calf serum, 100U/ml penicillin, 100 μg/ml streptomycin, 2 mM L-glutamine at 37°C and 5% CO_2. HEK293 cells were treated with 7.5 μM 1-NM-PP1 (Cayman Chemical Company) for 15 and 30 min. As a negative control, HEK293 cells were treated with DMSO (the resuspension vehicle for NM). Cells were routinely checked to be free of mycoplasma contamination using Plasmo Test Mycoplasma Detection Kit (InvivoGen, rep-pt1).

Analog sensitive cell line creation

Guide RNAs were computationally designed.

Guide RNA 1: GGTCCATATACTCAAATACAA

Guide RNA 2: GTATATGGACCATGACTTAAT

The modified CDK12 genomic sequence (RefSeq Accession NM_016507) (500 bp either site of the mutation) was cloned into pcDNA3 and used as the repair template for genome editing. The repair template contains a TTT to GGT mutation. The guide (g)RNAs inserts were cloned into the pX462 vector (obtained from Addgene). HEK293 cells were transfected with the gRNA vectors and correction template using Lipofectamine 2000 (Life Technologies) following the manufacturer instructions. Single clones were isolated by low density plating after Puromycin and Neomycin selection. Genomic DNA from each clone was analysed using PCR and Sanger sequencing.

gDNA preparation

Cells were incubated in 180 μl ChIP lysis buffer for 10 min and sonicated for 3 min (30 s on/30 s off) using a Q800R2 sonicator (QSONICA). 20 μl ammonium acetate was added (final conc. 400 mM) and samples were mixed with 200 μl phenol chloroform by vortexing. The samples were centrifuged (13 000 g, 5 min, 25°C) and the upper phase was transferred to a new tube. Genomic DNA was precipitated in 70% ethanol, pelleted by centrifugation (13 000 g, 5 min, 25°C) and dissolved in nuclease free water. A CDK12 fragment was PCR amplified using the following primers: Forward 5′ CCAGCCATATGGGTGG CAGAGTGAGAGTCT 3′; Reverse 5′ CCAGCGCGGC CGCGCAGAAAAGGCCAATCA 3′. DNA was purified using a QIAquick PCR purification kit (Qiagen) and sequenced by the SourceBioscience Sanger Sequencing Service, Oxford.

RNA preparation

RNA was extracted from HEK293 cells using TRIzol (Invitrogen) according to the manufacturer's instructions. Reverse-transcription (RT) was performed with 1 μg of RNA using random hexamers with the SuperScript III kit (Invitrogen) according to the manufacturer's instructions. Sequencing was performed on a cDNA PCR fragment generated with Forward primer: CGCTGTGTGGACAAGT TTGA and Reverse primer: GCTGGTGTGTAACGTT CCTC). Sequencing primer used: GCTTCCCAATCACA GCCATT.

Cell proliferation analysis

Cells were seeded at 500/well in 95 μl in 96-well microplates (Greiner, 655090) and measured every 12 or 24 h by adding alamarBlue HS (Invitrogen) (1/20) and reading in a fluorimeter after 1 hour's incubation, according to the manufacturer's instructions. NM-PP1 or DMSO was added to the cells as noted on the figure. Continuous xCelligence analysis was carried out in 96-well E-Plates (H003957) after seeding the cells at 500/well and NM or DMSO was added as noted on the figure, 36 hours after seeding.

Chromatin immunoprecipitation (ChIP)

ChIP analysis were performed as previously (30) using antibodies against control IgG (sc-2027, Santa Cruz), Pol II (MABI0601, MBL International), Ser2P (MABI0602, MBL International), Ser2P (ab5095, Abcam), Ser5P (MABI0603, MBL International), Ser5P (ab5131, Abcam), CPSF73 (A301-091A, Bethyl Laboratories), LEO1 (A300-175A, Bethyl Laboratories), CDC73 (A300-170A, Bethyl

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

Laboratories), SPT6 (15616S, Cell Signaling Technology). ChIP samples were analysed by real-time qPCR using QuantiTect SYBR Green PCR kit (Qiagen) and Rotor-Gene RG-3000 (Corbett Research). Signals are presented as percentage of Input after removing the background signal from the IP with the IgG antibody. The sequence of primers used for qRT-PCR is given in Table 1.

ChIP-seq of pol II in HEK293 (one replicate) or CDK12^{as} (biological duplicates) cells treated with DMSO or 7.5 μM 1-NM-PP1 was performed with the antibody sc-899 X (Santa Cruz Biotechnology). ChIP-seq of Ser2P (ab5095, Abcam), Ser5P (ab5131, Abcam), CPSF73 (A301–091A, Bethyl Laboratories), LEO1 (A300-175A, Bethyl Laboratories), and SPT6 (15616S, Cell Signaling Technology) were performed in biological duplicates. Preparation of ChIP-seq library and ChIP sequencing was prepared with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB), according to the manufacturer's instructions or conducted by the high throughput genomics team of the Wellcome Trust Centre for Human Genetics (WTCHG), Oxford.

Co-immunoprecipitation

For each sample: 100 μl of Dynabeads M-280 Sheep anti-mouse IgG (Thermo Fisher) were pre-blocked overnight at 4°C on a wheel in 1 ml of PBS supplemented with 0.5% BSA. The next day, the beads were washed three times in IP buffer (25 mM Tris–HCl pH 8.0, 150 mM NaCl, 0.5% NP-40, 10% Glycerol, 2.5 mM $MgCl_2$), before being incubated for 2 hours at 4°C on a wheel in 600 μl of IP buffer supplemented with 5 μg of antibody and protease inhibitor cocktail (cOmplete™, EDTA-free Protease Inhibitor Cocktail, Sigma-Aldrich). In the meantime, a 70–80% confluent 15 cm dish of CDK12^{as} cells was washed twice with ice-cold PBS before being scrapped with ice-cold PBS supplemented with protease inhibitor cocktail. The cells were pelleted at 500 g for 15 min at 4°C. The pellets were re-suspended in 800 μl of Lysis buffer (50 mM Tris–HCl pH 8.0, 150 mM NaCl, 1% NP-40, 10% glycerol, 2.5 mM $MgCl_2$, protease inhibitor cocktail, PhosSTOP (Sigma-Aldrich), 1× PMSF (Sigma-Aldrich), and 25–29 units of Benzonase (Merck Millipore)) and incubated at 4°C on a wheel for 30 min. After centrifuging for 15 min at 13 000 g at 4°C, 800 μl of Dilution buffer (150 mM NaCl, 10% glycerol, 2.5 mM $MgCl_2$, protease inhibitor cocktail, PhosSTOP, and 1× PMSF) was added to the supernatant

The beads conjugated with antibodies were washed three times with IP buffer supplemented with protease inhibitor cocktail, then the beads and 1 mg of proteins were incubated at 4°C on a wheel for 2 h. The beads were then washed three times with IP buffer containing 150 or 300 mM NaCl and supplemented with protease inhibitor cocktail and three times with IP buffer without NP-40, containing 150 or 300 mM NaCl, and supplemented with protease inhibitor cocktail. Proteins were eluted in 40 μl of 1× LDS plus 100 mM DTT for 10 min at 70°C. Western blots were performed with NuPAGE Novex 3–8% Tris-Acetate Protein Gels (Life Technologies).

The following antibody was used for immunoprecipitation: total pol II (MABI0601, MBL International).

Western blot

Western blot analysis was performed on chromatin and nucleoplasm extracts as previously described (29) until purification of the chromatin fraction. The chromatin pellet was digested in 100 μl of nuclease-free water supplemented with 1 μl of Benzonase (25–29 units, Merck Millipore) for 10 min at 37°C in a thermomixer (1400 rpm). 10 μg of proteins were boiled in 1× LDS plus 100 mM DTT. Western blots were performed with NuPAGE Novex 4–12% Bis–Tris Protein Gels (Life Technologies).

For whole cell extract, cells were washed in ice-cold PBS twice, collected in ice-cold PBS by 3000 rpm centrifuge for 5 min at 4°C. The pellets were re-suspended in RIPA buffer, kept on ice for 30 min with a vortexing step every 10 min. After centrifugation at 14 000 g for 15 min at 4°C, the supernatants were kept and quantified with the Bradford method. 20 μg of proteins were boiled in 1× LDS plus 100 mM DTT. Western blots were performed with NuPAGE Novex 4–12% Bis–Tris Protein Gels (Life Technologies).

The following antibodies were used: Pol II (MABI0601, MBL International), Ser2P (MABI0602, MBL International), Ser2P (5095, Abcam), Ser5P (MABI0603, MBL International), Ser5P (5131, Abcam), LEO1 (A300–175A, Bethyl Laboratories), CDC73 (A300–170A, Bethyl Laboratories), SPT6 (15616S, Cell Signaling Technology), CDK12 (11973S, Cell Signaling Technology), Cyclin K G-11 (sc-376371, Santa Cruz Biotechnology), β-tubulin (6046, Abcam), nucleolin (ab22758, Abcam), and histone H3 (ab1791, Abcam). Secondary antibodies were purchased from Merck Millipore (Goat Anti-Rabbit IgG Antibody, HRP-conjugate, 12-348, and Goat Anti-Mouse IgG Antibody, HRP conjugate, 12-349), the chemiluminescent substrate (SuperSignal West Pico PLUS) from Thermo Fisher, and the membranes visualized on an iBright FL1000 Imaging System (Thermo Fisher). Quantification of the western blots was performed with Image Studio Lite software.

mNET-seq and library preparation

mNET-seq was carried out in biological duplicates as previously described (29) with minor changes. In brief, the chromatin fraction was isolated from 4 × 10^7 CDK12^{as} cells treated with DMSO and 1-NM-PP1. Chromatin was digested in 100 μl of MNase (40 units/μl) reaction buffer for 2–3 min at 37°C in a thermomixer (1400 rpm). After addition of 10 μl EGTA (25mM) to inactivate MNase, soluble digested chromatin was collected by 13 000 rpm centrifuge for 5 min. The supernatant was diluted with 400 μl of NET-2 buffer and antibody-conjugated beads were added. Antibodies used: Pol II (MABI0601, MBL International), Ser2P (MABI0602, MBL International), Ser2P (5095, Abcam), and Ser5P (5131, Abcam). Immunoprecipitation was performed at 4°C for 1 h. The beads were washed with 1 ml of NET-2 buffer six times with 100 μl of 1xPNKT (1xPNK buffer and 0.05% Triton X-100) buffer once in cold room. Washed beads were incubated in 200 μl PNK reaction mix in Thermomixer (1400 rpm) at 37°C for 6 min. After the reaction, beads were washed with 1 ml of NET-2 buffer once and RNA was extracted with Trizol reagent. RNA was suspended in urea Dye and resolved on 6% TBU gel (Invitrogen) at 200 V for 5 min. In order to size select

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

Table 1. List of primers used in this study

Name	Sequence of forward primer	Sequence of reverse primer
NEG	TGGTACAACCACAGCTCAGTG	AAGCTGGACATGGTTGTGTG
KPNB1 TSS+0.2	TTACTTCCTCCCTCCAAATGGG	ACAGCCTCCCTTCCTTCTTTC
KPNB1 TSS+3.2	GCCCAGAGAACAAGAAATCG	GGAATGGACAAGCTGTGTTG
KPNB1 TSS+20.2	TGCAAGAGCCAGTGGGAACACTT	CCTCTACTCAGCAATGATACTTC
KPNB1 pA-4.8	CTGAGGAAACTGAAGAACCAAG	GAAGGCAGTGCTTGCCAGAAT
KPNB1 pA-2.9	GAGGAGTGTGCACGGATGCTGAA	CCAAGATGGCCGATGTTATGG
KPNB1 pA-0.4	TAGTTACCGTCTGCTTGGGAAGATG	CCTCTGACAGCAAGTCCAACATT
KPNB1 pA+1.4	GACTCATCACACCAAGGTCAC	GATAGTGCTGGGAAGGAAATGG
KPNB1 pA+2.6	GTACATCTCAGCTTTGGCATATG	GCCCAGAACATAGCAGGCATTGC
KPNB1 pA+4.1	GTTTCACCGTGTTAGCCAGGATGG	CCACAGCCATGTTCATTTCTGC
CCND2 TSS	AGAGCGAGACCAGTTTTAAGGG	TGGCTAAATAGGGGGTTTTCGG
H2AFZ1	ATGGGCGTTTTGTGTATGCG	CGACATCGAAACACGTGATTCC
H1F0	TGTGCAAGGACAGCAACAAC	GCCAACAAAACACACAAGCC
HIST1H1E	TAAACCAAAGACCGCCAAGC	AAGCAGTTGGCCAAAGGAAC
HIST1H2AH	AGTATGTCTGGACGTGGCAAG	TCGGCATAATTACCCTTGCG
HIST1H1C	ACTCTGGTGCAAACGAAAGG	TTAACCTTGGGCTTGGCTTC

35–100 nt RNAs, a gel fragment was cut between BPB and XC dye markers. A 0.5 mlL tube was prepared with 3–4 small holes made with 25G needle and placed in a 1.5 ml tube. Gel fragments were placed in the layered tube and broken down by centrifugation at 12 000 rpm for 1 min. The small RNAs were eluted from the gel using RNA elution buffer (1 M NaOAc and 1 mM EDTA) at 25°C for 1 h in Thermomixer (900 rpm). Eluted RNA was purified with SpinX column (Coster) with two glass filters (Millipore) and the flow-through RNA was ethanol precipitated. RNA libraries were prepared according to manual of NEBNext small RNA library prep kit (NEB). 12–14 cycles of PCR were used to amplify the library. Deep sequencing (Hiseq4000, Illumina) was conducted by the high throughput genomics team of the Wellcome Trust Centre for Human Genetics (WTCHG), Oxford.

TT-seq and library preparation

TT-seq was performed as described (27,31), with minor changes. Specifically, 6×10^7 cells, grown in absence of penicillin and streptomycin, from two biological replicates were treated for 15 and 30 min with solvent (DMSO) or 7.5μM 1-NM-PP1 (NM) at 37°C and 5% CO_2. For the 15 min treatment, cells were exposed to 10 min 500 μM of 4-thiouracil (4sU, Carbosynth, NT06186) after 5 min of DMSO or NM treatment (15 min total treatment). For the 30 min treatment, cells were exposed to 10 min 500 μM of 4-thiouracil (4sU, Carbosynth, NT06186) after 20 min of DMSO or NM treatment (30 min total treatment). Cells were lysed in 5 ml of QIAzol (Qiagen) and 300 ng of RNA spike-ins mix were added to each sample. RNA spike-ins were produced as described (27). RNAs were sonicated to obtain fragments of <10 kb using AFAmicro tubes in a S220 Focused-ultrasonicator (Covaris Inc, parameters: 10 s, peak power 100, cycles 200, duty cycle 1%). 4sU-labeled RNAs were purified from 240 μg of each of the fragmented RNAs. Biotinylation and purification of 4sU-labeled RNAs was performed as described (27,32). 100 ng of input RNA was used for strand-specific library preparation according to the Ovation Universal RNA-seq System (NuGEN). Libraries were prepared using random hexamer priming only.

QUANTIFICATION AND STATISTICAL ANALYSIS

TT-seq data preprocessing and global normalization parameters

Paired-end 75 base reads with additional 6 base reads of barcodes were obtained for each of the samples. Reads were demultiplexed and mapped using STAR 2.5.2b (33) to the hg20/hg38 (GRCh38) genome assembly (Human Genome Reference Consortium) with maximum 2% mismatches and only unique alignments were retained. SAMtools was used to quality filter SAM files, whereby alignments with MAPQ smaller than 7 (-q 7) were skipped and only proper pairs (-f2) were selected. Further data processing was carried out using the R/Bioconductor environment. We used a spike-in (RNAs) normalization strategy essentially as described (27) to allow observation of global shifts (sequencing depth) σ_j, cross-contamination rate ϵ_j (proportion of unlabeled reads purified in the TT-seq samples) and antisense bias ratio c_j (ratio of spurious reads originating from the opposite strand introduced by the reverse transcription reaction). Read counts (k_{ij}) for spike-ins were calculated using HTSeq (34). Calculations for each parameter are described in the following in more detail.

Antisense bias ratio c_j

Antisense bias ratios were calculated for each sample j according to

$$c_j = \underset{i}{\text{median}} \left(\frac{k_{ij}^{antisense}}{k_{ij}^{sense}} \right)$$

for all available spike-ins i.

Cross-contamination rate ϵ_j

The cross-contamination rate ϵ_j was calculated for each sample j as

$$\epsilon_j = \underset{i}{\text{median}} \left(\frac{k_{ij}}{l_i} \right) / \sigma_j$$

using the unlabeled spike-ins i for TT-seq samples.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

Definition of transcription units based on the UCSC RefSeq genome assembly GRCh38 (RefSeq-TUs)

For each annotated gene, transcription units were defined as the union of all existing inherent transcript isoforms (UCSC RefSeq GRCh38).

TT-seq processing with global normalization parameters

The number of transcribed bases (tb_j) or read counts (k_j) for RefSeq-TUs were normalized and corrected for antisense bias c_j, sequencing depth σ_j as follows using the parameters calculated as described above. The cross-contamination rate estimates for TT-seq in HEK293 were very low and were thus not corrected for.

Antisense bias correction

The real number of read counts or coverage s_{ij} for transcribed unit i in sample j was calculated as

$$s_{ij} = \frac{S_{ij} - c_j A_{ij}}{1 - c_j^2}$$

where S_{ij} and A_{ij} are the observed numbers of read counts or coverage on the sense and antisense strand.

Read counts per kilobase (RPK)

RPKs were calculated upon antisense bias corrected read counts (k_j) falling into the region of a RefSeq-TU divided by its length in kilobases.

TT-seq normalization

For a robust normalization, we first identified a subgroup of RefSeq-TU's as non-differentially expressed (non-DE) for both time points over the response of NM-PP1(CDK12^{as} inhibitor) using DESeq2 package. Based on the antisense bias corrected RPKs, a group of expressed TUs [n = 5631] was defined to comprise all TUs with an RPK of 100 or higher in two summarized replicates of TT-seq 15 min samples without inhibitor treatment. An RPK of 100 corresponds to approximately a coverage of 10 per sample due to an average fragment size of 200. Within this group, a subset of TUs having a length >50 kb [$n = 1867$] was used to identify the non-DE TU's.

The antisense bias corrected coverage s_{ij} of transcribed unit i in sample j was normalized for sequencing depth as

$$t_{ij} = \frac{s_{ij}^{\mathrm{TT-seq}}}{\sigma_j^{\mathrm{TT-seq}}}$$

where σ_j was balanced between replicates via classical size factor normalization to gain statistical power in the differential expression analysis.

On the resulting 140 non-DE RefSeq-TUs, size factors for each sample were then estimated as described by Equation (5) in DESeq package (34) to correct for library size and sequencing depth variations.

Visualization

For visualization, read coverages were summed up over replicates, size factors for each condition were calculated on the 140 non-DE RefSeq-TUs by Equation (5) in DESeq package (34) and used to correct for library size and sequencing depth variation.

For plotting, the expressed gene set is defined as follows: Based on the antisense bias corrected RPKs, a group of expressed TUs [$n = 11\,282$] was defined to comprise all TUs with an RPK of 20 or higher in two summarized replicates of TT-seq 15 min samples without inhibitor treatment. An RPK of 20 corresponds to approximately a coverage of 2 per sample due to an average fragment size of 200.

Simulation of TT-seq data based on elongation velocity profiles

Based on the following model we simulated TT-seq coverage values by providing elongation velocity profiles $v(t)$, a labeling duration t^{lab} and a uracil content dependent labeling bias

$$l_f = 1 - \left(1 - p^{lab}\right)^{\#u_f}$$

p^{lab} denotes the labeling probability (set to 0.01) and $\#u_f$ the number of uracil residues of a given fragment f (set to 0.28 times fragment length). Elongating polymerases are propagated at nucleotide resolution along the template in a time-dependent manner. Each position of elongating polymerases τ^* can be mapped to its respective time of transcription $t^*$with an elongation velocity profile $v(t)$ as

$$\tau^* = \int_0^{t^*} v\,(t)\,dt$$

Consequently, the progression of a polymerase at any given position τ^* with velocity $v(t^*)$for time $t_p$can be calculated as

$$\tau_{t_p}\,(t) = \int_0^{t_p+t*} v\,(t)\,dt - \int_0^{t^*} v\,(t)\,dt = \int_{t*}^{t_p+t*} v\,(t)\,dt$$

while $v(t)$ is the current velocity function at absolute time.

With this equation, the movement of single polymerases over the course of labeling can be tracked. Start and end positions of fragments can be derived given the positions of any given polymerase at $\tau(t, t - t^{lab})$ for any given transcription time. We used the number of uracil residues present in the RNA fragment $\#u_f$ to weight the amount of coverage contributed by this fragment as l_f. Additionally, we applied a size selection similar to that in the original protocol for fragments below 80 bp in length with a sigmoidal curve that mimics a typical size selection spread. This was done for templates resembling genes of sizes ~0,1-2000 kbp. The resulting gene-wise RNA synthesis profiles were subsequently accumulated to yield meta-gene profiles.

mNET-seq data processing

mNET-seq data were processed as follows: adapters were trimmed with Cutadapt in paired-end mode with

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

the following parameters: -q 15, 10 –minimum-length 10 –A GATCGTCGGACTGTAGAACTCTGAAC –a AGATCGGAAGAGCACACGTCTGAACTCCAGTC AC. Trimmed reads were mapped to the human hg38 reference sequence with Tophat2 and the parameters –g 1 –r 3000 –no-coverage-search. SAMtools was used to retain only properly paired and mapped reads (-f 3). A custom python script (28) was used to obtain the 3′ nucleotide of the second read and the strandedness of the first read. Strand-specific bam files were generated with SAMtools. FPKM normalized bigwig files were generated for each bam files with Deeptools2 (35) bamCoverage tool (-bs 1 –p max –normalizeUsing RPKM). For plotting, the expressed gene set is defined as follows: Based on the antisense bias corrected RPKs, a group of expressed TUs [n = 11282].

mNET-seq and ChIP-seq normalization

The total pol II mNET-seq treated with DMSO and NM 15 min were re-normalized to a set of 140 genes found to be non-affected in the spiked-in TT-seq analysis after 15 and 30 min inhibition. The re-normalization factor was calculated from the average fold change on total pol II signal between DMSO and the NM conditions across the gene bodies of this set of non-affected genes found in TT-seq. For the total pol II mNET-seq metaprofiles and total pol II quantification, the re-normalization factors for NM 15 min R1 and R2 are 0.90 and 1.11, respectively. The Ser2P and Ser5P 15 min were normalized to the CTD phosphorylation signals in histone genes (see Results and Supplementary Figure S4) with the following normalization factors for the merged biological replicates: Ser2P MABI0602: 0.626, Ser2P ab5095: 0.544 and Ser5P ab5131: 0.791.

ChIP-seq data processing

Adapters were trimmed with Cutadapt in paired-end mode with the same parameters as mNET-seq. Obtained sequences were mapped to the human hg38 reference genome with Bowtie2. Properly paired and mapped reads were filtered with SAMtools. PCR duplicates were removed with Picard MarkDuplicates tool. FPKM normalized bigwig files were generated for each bam files with Deeptools2 bamCoverage tool (-bs 10 –p max –normalizeUsing RPKM –e). For the HEK293 ChIP-seq presented in Supplementary Figure S2B, the pol II signal was normalized to the *KPNB1* gene body (TSS + 500 bp to poly(A) site, NM normalization factor: 0.494) as this gene was found to be non-affected by pol II ChIP-qPCR. CDK12as pol II ChIP-seq NM normalization factors are 1.345 and 0.840 for replicates 1 and 2, respectively. Ser2P NM normalization factors are 1 and 1.429 for replicates 1 and 2, respectively. Ser5P NM normalization factors are 1.275 and 1.269 for replicates 1 and 2, respectively.

Metagene profiles

Metagene profiles of genes scaled to the same length were then generated with Deeptools2 computeMatrix tool with a bin size of 10 bp and the plotting data obtained with plotProfile –outFileNameData tool. Graphs representing the (IP – Input) signal (ChIP-seq) or the mNET-seq signal were then created with GraphPad Prism 8.4.0. Metagene profiles are shown as the average of two biological replicates.

P-values and significance tests

P-values were computed with an unpaired two-tailed Student's *t* test. Statistical tests were performed in GraphPad Prism 8.4.0.

RESULTS

Rapid and specific CDK12^{as} inhibition in human cells

Several studies using RNAi-mediated knockdown of CDK12 have highlighted roles for this kinase in CTD phosphorylation and expression of DDR genes (3,4,17). To assess whether these effects are mediated by CDK12 kinase activity, the loss of CDK12 itself, or indirect effects due to the long time-frame of knockdown, we produced a CDK12^{as} HEK293 cell line, using CRISPR/Cas9, where the gatekeeper phenylalanine residue (codon TTT) is mutated to glycine (codon GGT) (Supplementary Figure S1A). This allows rapid and selective inhibition of CDK12 kinase activity with a bulky ATP-analog 1-NM-PP1 (NM) (14,16,36), without affecting levels of the protein itself, which has been shown to interact with other components of the transcription/RNA processing machinery (12,19,37). Importantly, the growth rate of the CDK12^{as} cells is not lower than that of the parental HEK293 cell line (Supplementary Figure S1B). Also, CDK12 and Cyclin K protein levels are unchanged in CDK12^{as} cells compared to the parental cell line, as measured by western blot analyses (Supplementary Figure S1C and D).

To determine which concentration of NM to use, we followed the growth of CDK12^{as} and the parental HEK293 cells after the addition of 5, 7.5 and 10 μM of NM to the medium using either an alamarBlue HS cell viability assay or continuous label-free monitoring by xCelligence (Supplementary Figure S1E and F). In both cases, 7.5 μM NM affects the viability and growth of CDK12^{as} cells with less effect on HEK293, whereas 10 μM NM affects both cell lines. This is in agreement with previously published data on CDK12^{as} inhibition in HeLa cells (14). Importantly, short-term treatment of HEK293 and CDK12^{as} cells with 7.5 μM NM does not affect the protein levels of CDK12 and Cyclin K (Supplementary Figure S1G and H) and takes many hours to affect cell viability (Supplementary Figure S1E). We have therefore treated cells with this inhibitor concentration for 15 or 30 min, in order to determine the immediate effect of CDK12 inhibition. Importantly, pol II distribution on expressed transcription units in the parental cells is not affected by short-term NM treatment as measured by pol II chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR) of *KPNB1* and ChIP-sequencing (ChIP-seq) (Supplementary Figure S2A and B).

CDK12^{as} inhibition globally decreases RNA synthesis

To monitor changes in RNA synthesis upon CDK12^{as} inhibition, we performed TT-seq with RNA spike-ins after 15 and 30 min of NM treatment (Figure 1A). Also using

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

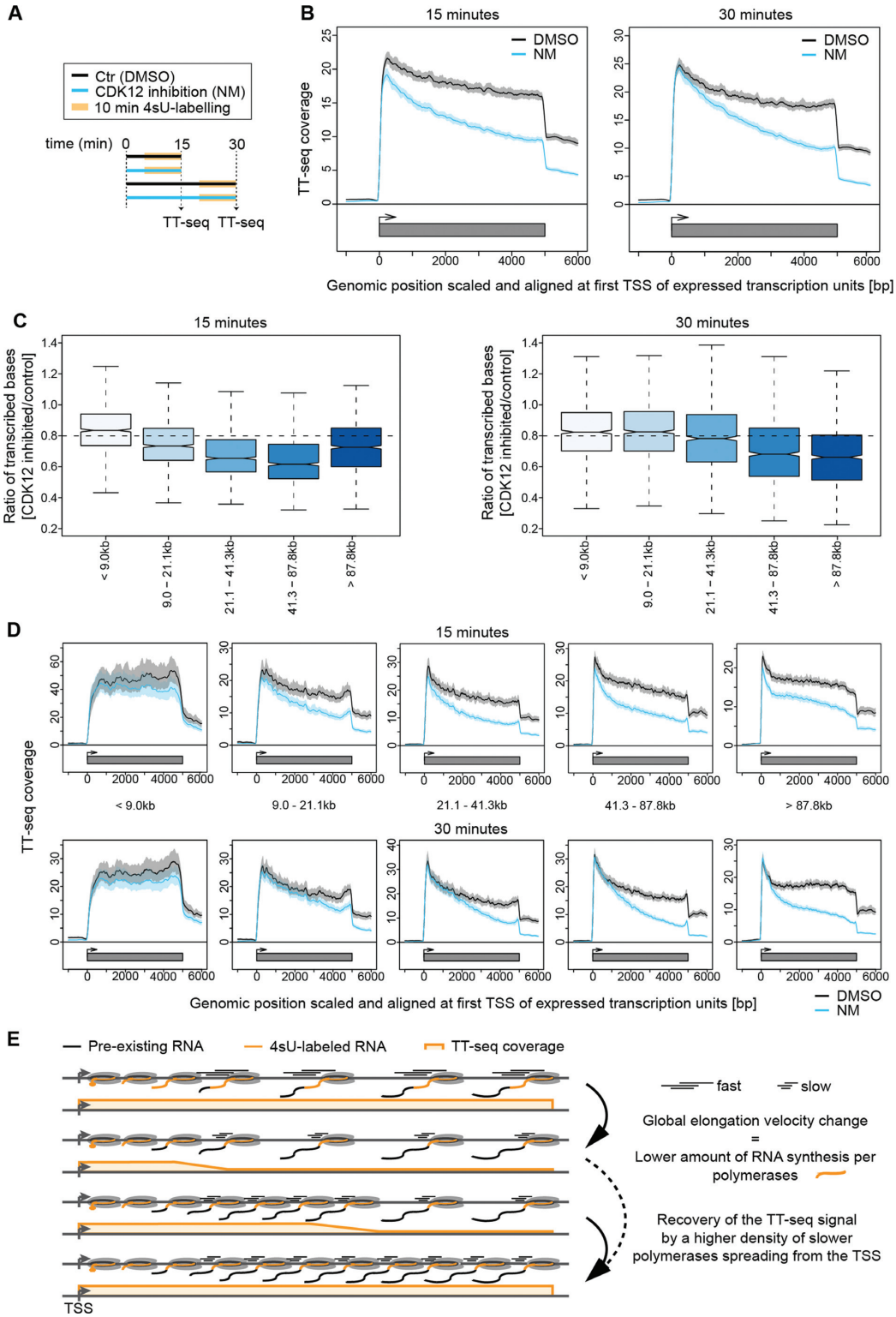

Figure 1. CDK12^{as} inhibition globally decreases RNA synthesis. (**A**) Experimental design. (**B**) Metagene analysis of TT-seq signal for expressed genes after treatment of cells with DMSO (black) versus 7.5 μM NM treatment (blue) for 15 min (left) or 30 min (right). The TT-seq coverage was averaged and aligned at their transcription start sites (TSSs) and polyadenylation (pA)-sites. Shaded areas around the average signal (solid lines) indicate confidential intervals. (**C**) Box plots of different length classes show the ratio of transcribed bases after 15 min (left) and 30 min (right) inhibition of CDK12^{as} compared to control. (**D**) Metagene analysis for different length classes comparing the average TT-seq signal before (DMSO treatment, black) and after CDK12^{as} inhibition (7.5 μM NM treatment, blue) for 15 min (upper panel) and 30 min (lower panel). The TT-seq coverage was averaged and aligned at their transcription start sites (TSSs) and polyadenylation (pA)-sites. Shaded areas around the average signal (solid lines) indicate confidential intervals. (**E**) Schematic representation of TT-seq signal changes along the gene body upon elongation velocity change. Upper panel: steady state transcription. Lower panels show TT-seq signal recovery spreading from the TSS.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

this analysis, treatment with NM has little effect on transcription in the parental cells (Supplementary Figure S2C and D). However, metagene profiles of TT-seq signals averaged over expressed genes (see Materials and Methods) in CDK12^{as} cells show a decreased TT-seq signal across gene bodies after 15 min of inhibition, indicating reduced RNA synthesis, either as the result of reduced elongation or loss of polymerase, in 11 182 transcription units, including 10,393 protein-coding genes and 789 unclassified transcription units (Figure 1B–D). CDK12^{as} inhibition for 30 min instead leads to a recovery of RNA synthesis activity at the beginning of genes (Figure 1C, right, D lower panel).

Recovery of the TT-seq signal after 30 min could be explained by a higher density of slower polymerases spreading from the transcription start site and gradually populating the gene body (Figure 1E). This effect would not require any changes in initiation frequencies, but could stem only from an increase in polymerase density over genes due to slower elongation. After CDK12^{as} inhibition, polymerases dramatically slow down, resulting in a lower amount of RNA synthesis, which may be successively restored by a higher number of polymerases occupying the gene body provided the same number of initiation events. This model predicts that recovery of RNA synthesis takes longer for long genes and shorter for short genes (Figure 1C–E). As human genes take on average longer than 30 min to be transcribed, recovery of RNA synthesis and TT-seq signal would be limited to the 5′-region of genes and not observed for many 3′-regions, leading to the observed slope in the TT-seq metagene profile (Figure 1C–E).

To test this model further, we simulated TT-seq metagene profiles based on a kinetic model that computes RNA synthesis at each gene position based on initiation rates and elongation velocities. The model readily recapitulates the observed TT-seq profiles if we assume that CDK12 inhibition induces rapid and global downregulation of RNA elongation without affecting transcription initiation frequency (Supplementary Figure S2E and F). Thus, TT-seq analysis and kinetic modeling supports the notion that CDK12 is required for normal transcription elongation, and the elongation defect resulting from its inactivation is readily observed after rapid inhibition followed by immediate monitoring of RNA synthesis.

Inhibition of CDK12^{as} affects transcription elongation

To confirm that the observed decrease in RNA synthesis activity results from reduced elongation, we performed mNET-seq of total pol II in CDK12^{as} cells after 15 min of NM treatment (Figure 2A). mNET-seq identifies the last nucleotide transcribed by sequencing the RNA present in the active site of immunoprecipitated pol II (29). The mNET-seq data were normalized to a set of genes found to be unaffected in TT-seq (see Materials and Methods). Inhibition of CDK12^{as} results in an increase of pol II signal in the gene body of the vast majority of expressed pol II-transcribed genes within 15 min of inhibition (Figure 2B and C), consistent with higher densities of more slowly-elongating pol II. ChIP-seq of pol II gives a similar picture after 15 min inhibition of CDK12^{as} (Supplementary Figure S3A). We have confirmed this increase of pol II occupancy in the gene body of a model gene, *KPNB1*, by performing pol II ChIP-qPCR after 15 and 30 min treatment of CDK12^{as} cells with NM (Supplementary Figure S3B). The higher pol II signal results from an increased pol II residence time in gene bodies due to slower elongation, and not from an increased amount of pol II entering elongation, as shown by the decrease in RNA synthesis in TT-seq (Figure 1).

Further analysis shows that transcription of pol II-transcribed genes of all lengths is affected by CDK12^{as} inhibition, with a wave of elongation-compromised pol II reaching the ends of the shortest genes first (Figure 2D). Genes <9 kb have a less pronounced elongation defect after CDK12^{as} inhibition than genes >9 kb as measured by both TT-seq and mNET-seq (Figures 1D and 2D). In addition, the elongation defect increases with increasing distance from the TSS for all gene length classes (Figure 1D). This is more apparent after analysis of TT-seq and mNET-seq data for the first few kilobases of genes >6.5 kb (Supplementary Figure S3C and D). Thus, the effect of CDK12^{as} inhibition starts early in the transcription cycle and increases with increasing distance from the TSS (Supplementary Figure S3C and D). Many of the genes <9 kb will therefore be too small to exhibit a major effect of CDK12^{as} inhibition on elongation as pol II will reach the end of genes soon after the elongation defect starts to occur.

Interestingly, on genes longer than 21.1 kb, pol II increases in the gene body but is specifically reduced downstream of the poly(A) site (Figure 2D), before the wave of slower polymerases reaches this region. Single gene examples of this phenomenon are shown in Supplementary Figure S3E and this can also be observed by pol II ChIP-qPCR of the *KPNB1* gene (Supplementary Figure S3B).

Since transcription of intron-containing protein-coding genes is affected by short-term inhibition of CDK12^{as}, we investigated whether transcription of intronless ($n = 5396$), histone ($n = 118$), and snRNAs ($n = 36$) genes is also altered (Supplementary Figure S4A and B). TT-seq signals for intronless and pol II-transcribed snRNA genes decrease after 15 or 30 min inhibition, whereas mNET-seq signals on intronless genes are not markedly affected and slightly reduced on snRNA genes. We interpret this to mean that pol II is slower on these genes but pol II is not building up as they are relatively short (the intronless and snRNAs have an average size of 948 and 150 bp, respectively). TT-seq signals for histone genes instead are either unaffected or slightly increased and mNET-seq signals unaffected by CDK12^{as} inhibition. We therefore conclude that transcription of histone genes does not require CDK12 activity. We have confirmed this by pol II ChIP-qPCR for five different histone genes (Supplementary Figure S4C).

The primary effects of CDK12^{as} inhibition on transcription of intron-containing protein-coding genes are therefore an elongation defect in gene bodies and loss of pol II downstream of the poly(A) site. Taken together, the results indicate that CDK12 plays a role in maintaining efficient transcription elongation velocity and processivity on these human genes. Efficient transcription of intronless genes and snRNA genes also appears to require CDK12, although the defects are less apparent. Histone genes, on the other hand, are transcribed efficiently in the absence of CDK12.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

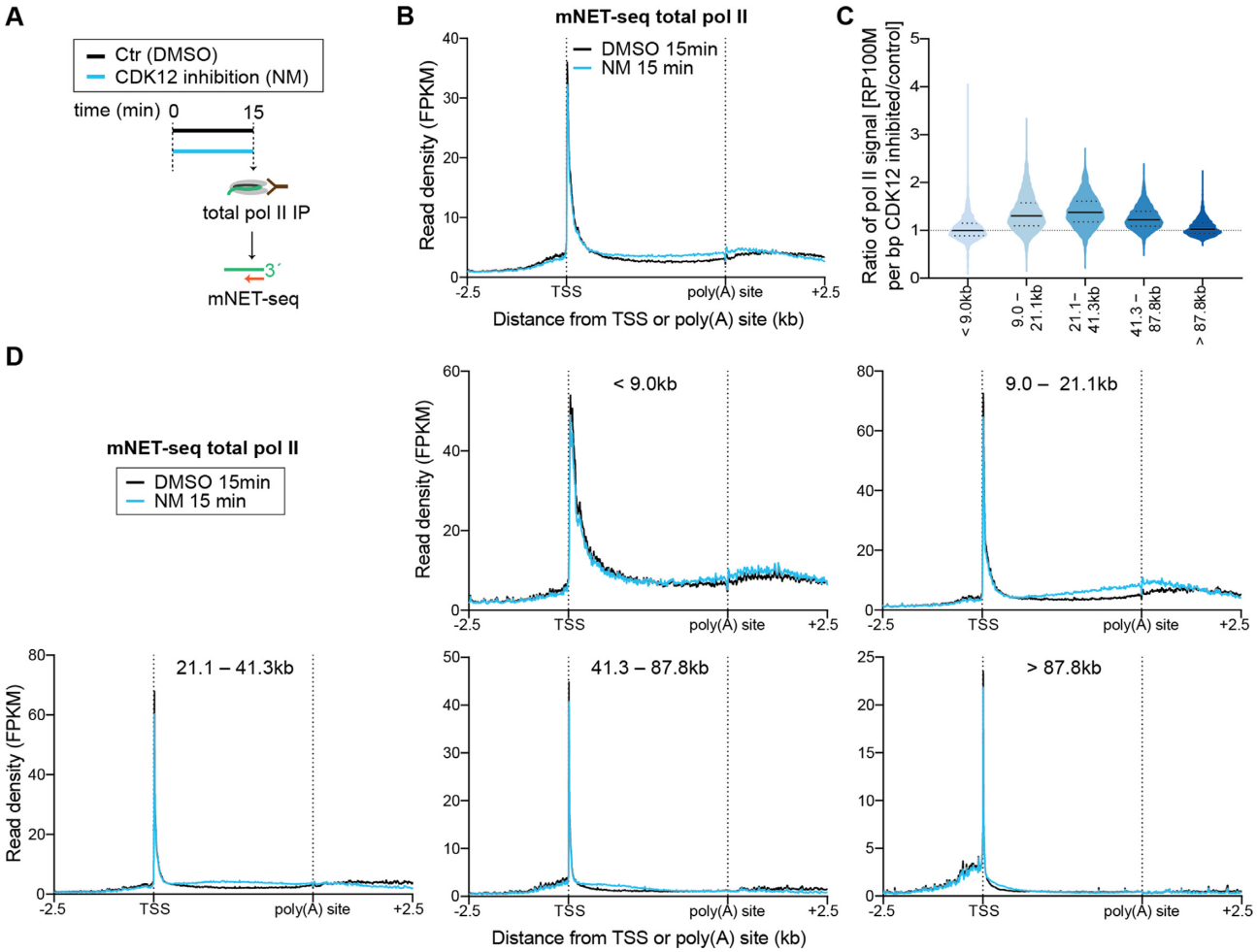

Figure 2. Inhibition of CDK12^{as} affects transcription elongation. (**A**) Experimental design. mNET-seq using an antibody against total pol II was performed in CDK12^{as} HEK293 cells after treatment with 15 min of DMSO solvent control or 7.5 μM NM. (**B**) Metagene analysis comparing the average mNET-seq signal before and after CDK12^{as} inhibition for 15 min of expressed genes. (**C**) Violin plots of the different length classes noted show the ratio of mNET-seq signal after 15 min inhibition of CDK12^{as} compared to control. (**D**) Metagene analysis for the different length classes noted comparing the average mNET-seq signal before and after CDK12^{as} inhibition for 15 min. The mNET-seq reads were averaged and aligned at their TSSs and (pA)-sites.

CDK12 phosphorylates transcribing pol II

mNET-seq can also be carried out with CTD phospho-site-specific antibodies to detect nascent RNA associated with particular pol II phospho-isoforms (29). Accordingly, we have used mNET-seq to investigate the primary effect of CDK12^{as} inhibition on CTD phosphorylation. As short-term inhibition of CDK12^{as} is sufficient to cause global changes in transcription (Figures 1 and 2), we treated CDK12^{as} with NM for 15 min and obtained mNET-seq profiles for pol II CTD phosphorylation at Ser2P and Ser5P using antibodies against each of these phospho-forms (Figure 3).

As transcription of histone genes is relatively unaffected by CDK12^{as} inhibition, we have investigated whether pol II CTD Ser2P and Ser5P levels are affected by CDK12^{as} inhibition. Ser2P and Ser5P ChIP-qPCR for five different histone genes indicates that Ser2P and Ser5P, ratioed and unratioed to pol II levels, is unaffected by NM treatment (Supplementary Figure S4C), indicating no requirement for CDK12 for CTD phosphorylation during transcription of these genes. Interestingly, the level of Ser2P relative to pol II on histone is generally much lower than on other protein-coding genes (a ratio of up to 0.6 on histone genes and a ratio of >40 at the end of *KPNB1* with ab5095) (Supplementary Figures S4 and S5), which is consistent with the relative lack of requirement for a Ser2 kinase. We have therefore used the levels of Ser2P and Ser5P on histone genes to normalize our mNET-seq Ser2P and Ser5P data (see Materials and Methods). We have displayed the results with and without normalisation of the mNET-seq signals to total pol II levels to take changes of pol II occupancy into account (Figure 3A and B). Figure 3C shows the genome browser track of the unratioed mNET-seq results for the *CCND2* gene, a ∼32 kb long gene displaying increased pol II signal in the gene body and a loss of pol II downstream of the poly(A) site. The mNET-seq signals for two different Ser2P antibodies and for the Ser5P antibody show a decrease after CDK12^{as} inhibition when ratioed to the total pol II level (Figure 3A and B). The same effect of CDK12^{as} inhibition

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

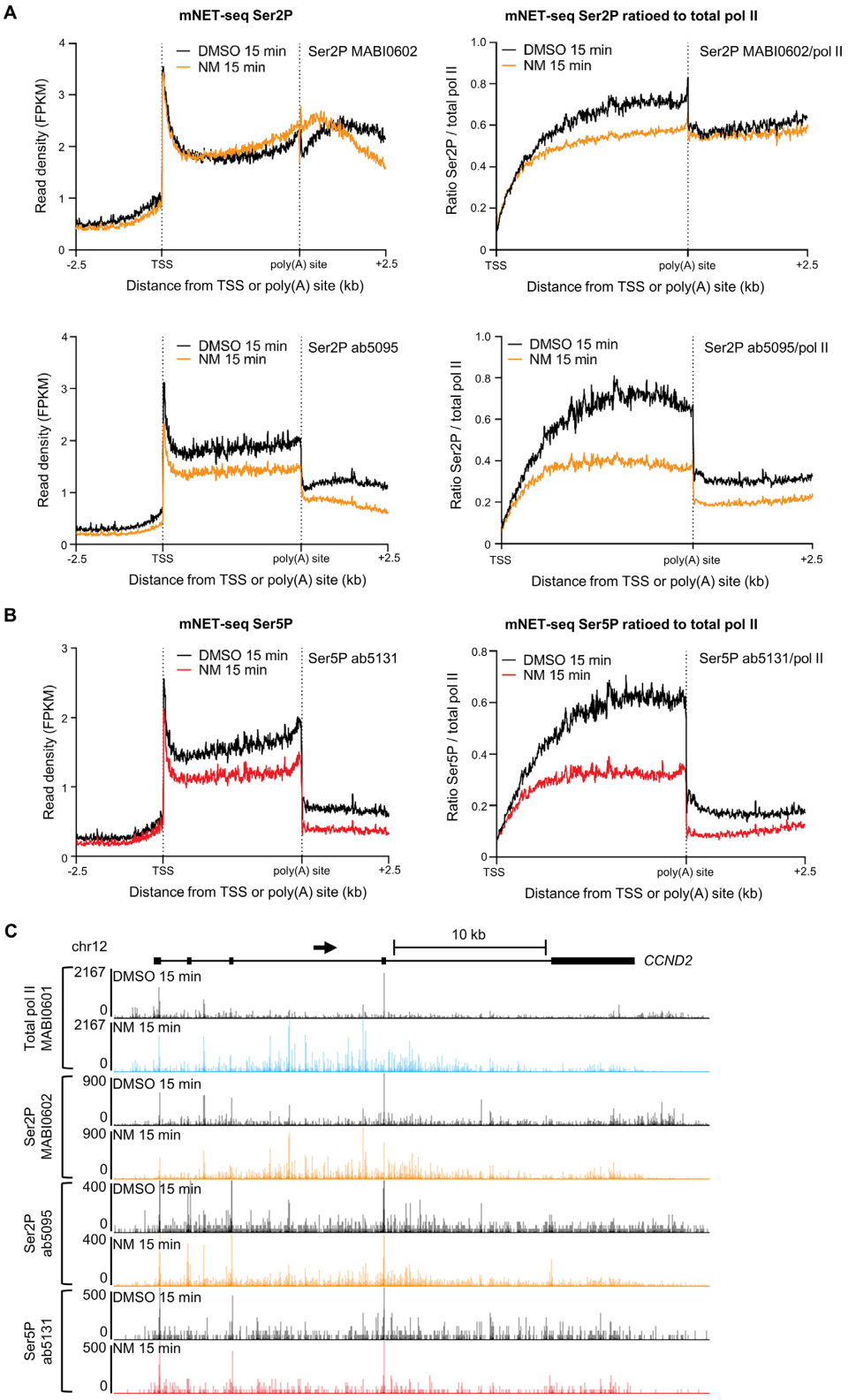

Figure 3. CDK12 phosphorylates transcribing pol II. (**A**) Meta-analyses of scaled expressed genes of mNET-seq for Ser2P with and without normalization to pol II after treatment of CDK12^{as} cells with DMSO (black) or 7.5μM NM (orange) for 15 min. (**B**) Meta-analyses of scaled expressed genes of mNET-seq for Ser5P with and without normalization to pol II after treatment of CDK12^{as} cells with DMSO (black) or 7.5 μM NM (red) for 15 min. (**C**) mNET-seq profiles across *CCND2* using a total pol II antibody, two Ser2P antibodies and one Ser5P antibody.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

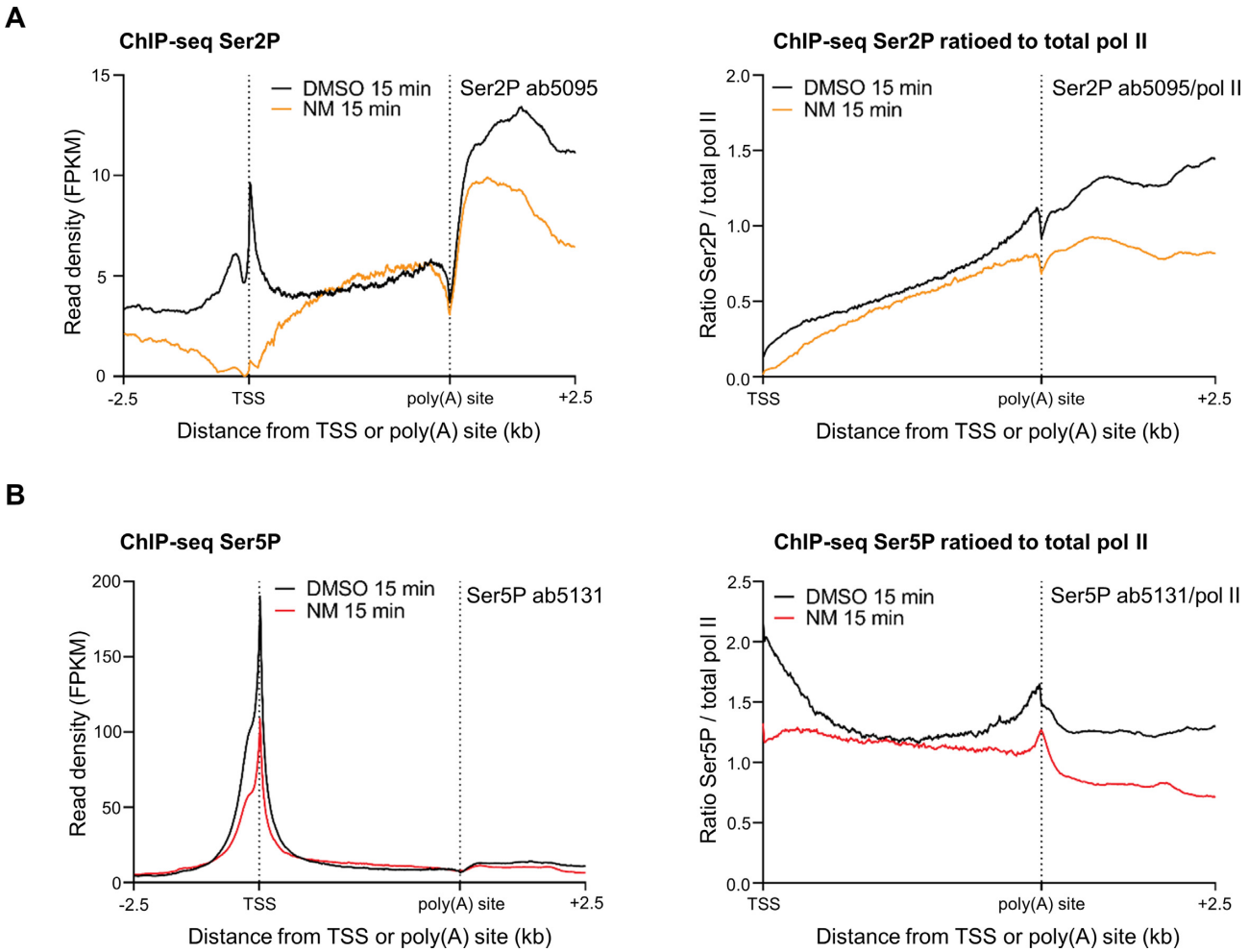

Figure 4. ChIP-seq analysis of CDK12 phosphorylation of transcribing pol II. (**A**) Meta-analyses of scaled expressed genes of ChIP-seq for Ser2P (ab5095) with and without normalization to pol II after treatment of CDK12^{as} cells with DMSO (black) or 7.5μM NM (orange) for 15 min. (**B**) Meta-analyses of scaled expressed genes of ChIP-seq for Ser5P (ab5131) with and without normalization to pol II after treatment of CDK12^{as} cells with DMSO (black) or 7.5 μM NM (red) for 15 min.

on Ser2P and Ser5P phosphorylation is seen using ChIP-seq of Ser2P and Ser5P and ChIP-qPCR on *KPNB1* (Figure 4A, B, Supplementary Figure S5A and B) Interestingly, we observed a loss of Ser5P phosphorylation at the TSS of genes after CDK12^{as} inhibition, indicating that CDK12 acts as a Ser5 kinase at this step and could affect mRNA capping, which relies on Ser5P (38) (Supplementary Figure S5B). We first confirmed the loss of Ser5P by ChIP-qPCR on the TSS of *KPNB1* and *CCND2* after 15- and 30-min inhibition of CDK12^{as} (Supplementary Figure S5C and D). To investigate a possible effect on mRNA capping, we performed ChIP-qPCR of RNGTT, the enzyme responsible for the first two catalytic steps of the cap formation, and found a decrease in signal at the TSS following CDK12^{as} inhibition consistent with the reduction in Ser5P relative to total pol II (Supplementary Figure S5C and D).

We also carried out western blot analyses on chromatin-associated pol II with antibodies against different pol II CTD phospho-isoforms after treating cells with NM for 15 and 30 min (Supplementary Figure S6A and B). In line with the results of mNET-seq analysis (Figure 3A and B), inhibition of CDK12^{as} markedly affects Ser2 and Ser5 phosphorylation within 15 min, as measured with two different antibodies against these modifications (Supplementary Figure S6A and B). Both Ser2P and Ser5P continue to be affected after 30 min, with a further decrease in phosphorylation observed for Ser2P MABI0602 and Ser5P MABI0603 (Supplementary Figure S6A and B). In contrast, no loss of Ser2P or Ser5P occurs when HEK293 cells are treated with NM for 15 or 30 min (Supplementary Figure S6C and D). Loss of Ser2P and Ser5P after 15 min of inhibition of CDK12^{as} was also confirmed by performing immunoprecipitation of total pol II followed by western blotting with a Ser2P (ab5095) and Ser5P (ab5131) antibody (Supplementary Figure S6E). These results are consistent with previous *in vitro* kinase assays and analysis of CDK12^{as} HeLa cells (14,15). Taken together, the western and mNET-seq analyses indicate that CDK12 contributes to phosphorylation of the CTD of engaged pol II with activity towards both Ser2 and Ser5.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

CDK12 activity is required for stable association of elongation and termination factors with chromatin

Our results indicate that CDK12 is required for normal transcription elongation of most pol II-transcribed genes (Figures 1 and 2). However, the mechanism remains unclear. CDK12 activity may be required for the recruitment of elongation factors to transcribing pol II or for the stabilization of their interactions with pol II. We have therefore tested by ChIP-seq and ChIP-qPCR whether the levels of the LEO1 and CDC73 subunits of the elongation factor PAF1 complex (PAF1C) and the elongation factor SPT6 detected in chromatin are affected by CDK12^{as} inhibition (Figure 5A–C and S7). We have displayed the LEO1 and SPT6 ChIP-seq results with or without normalisation to pol II, to take the changes in pol II levels into account (Figure 5A and B). The level of LEO1 and SPT6 associated with expressed genes appears to be affected genome-wide. In addition, LEO1 appears to be selectively lost from the newly-elongating pol II as long genes (e.g. >41.3 kb) have a reduction in the ratio of LEO1 to total pol II at the 5′ end, but not at the 3′ end, where the pol II that initiated before treatment is still elongating (Figure 5C). In contrast, SPT6 is reduced more globally as a decrease in signal is also observed at the 3′end of long genes, albeit less than at the 5′ end of genes (Figure 5D). In addition, the association of CDC73 and SPT6 by ChIP-qPCR with *KPNB1*, with or without normalisation to pol II, is reduced after 15 min inhibition of CDK12^{as} (Supplementary Figure S7C). Similarly to LEO1, the level of CDC73 on *KPNB1*, which is ∼34 kb long, is decreased in the gene body but not at the 3′ end by CDK12^{as} inhibition. Western analyses also indicate that inhibition of CDK12^{as} decreases the association of SPT6, LEO1, and CDC73 with chromatin (Figure 5E and F). The loss of SPT6, LEO1, and CDC73 from chromatin is mirrored by an increase in the level of these proteins in nucleoplasm fraction after CDK12^{as} inhibition, with a particularly marked increase in nucleoplasmic SPT6 (Supplementary Figure S8A and B). These results are in line with a loss of association of these factors with chromatin, although we cannot rule out that changes in post-translation modifications are affecting antibody reactivity. Importantly, no loss of these factors from chromatin is observed when HEK293 cells are treated with NM (Supplementary Figure S8C–F). Also, the decrease on chromatin of SPT6, LEO1, and CDC73 is not due to protein degradation, as the global protein level of these elongation factors remains stable after NM treatment, both in HEK293 and CDK12^{as} cells (Supplementary Figure S8G and H). The results of immunoprecipitation of total pol II from cell extracts followed by western analyses of total pol II and SPT6 after 15 min inhibition of CDK12^{as} also suggest that SPT6 association with pol II is reduced (Supplementary Figure S9). Taken together, these results support the notion that CDK12 activity plays a key role in ensuring that the critical elongation factors PAF1C and SPT6 are part of the pol II elongation complex (19,39,40). Loss of these factors could readily explain the elongation defect that we observe after CDK12 inhibition.

It was previously shown that CDK12 is required for the recruitment of polyadenylation/transcription termination factors (17,18). To investigate further whether the loss of Ser2P and PAF1C following CDK12^{as} inhibition affects the recruitment of termination factors, we performed ChIP-seq of CPSF73, the enzyme responsible for the cleavage of the pre-mRNA at the poly(A) site (Supplementary Figure S10). We observed a general loss of CPSF73 levels after 15 min of CDK12^{as} inhibition (Supplementary Figure S10A). Importantly, the level of CPSF73 is unaffected at the 3′end of long genes (> 41.3 kb), suggesting that, similarly to LEO1 and CDC73, CPSF73 is lost from the newly-elongating pol II (Supplementary Figure S10B). We confirmed the drop in CPSF73 levels detected at the same point as newly-initiated pol II after CDK12^{as} inhibition by ChIP-qPCR on the *KPNB1* gene (Supplementary Figure S10C). These results are in agreement with the finding that Ser2P is retained relative to pol II at the 3′ end of long genes (Supplementary Figure S11). In contrast Ser5P is reduced relative to pol II at the end of long genes (Supplementary Figure S11).

DISCUSSION

Here, we have established a protocol to rapidly and specifically inhibit CDK12 in human cells and immediately monitor changes in RNA synthesis and the occupancy of engaged pol II genome-wide. We have restricted our analysis to the first 30 min after inhibition to capture the primary effect of loss of CDK12 kinase activity. An additional benefit of our strategy is the potential to lose the activity of the kinase independently of losing the protein itself. In contrast to previous studies, our results indicate that CDK12 activity is required for efficient elongation of transcription of the vast majority of expressed pol II-dependent genes, indicating that it is a general transcription kinase, rather than playing a specific role in transcription of genes longer than 45 kb (25). Our results further suggest that CDK12 plays an increasing role in elongation after pol II has passed the CDK9-regulated early-elongation checkpoint (EEC) and entered productive elongation (Supplementary Figure S3C and D) (41). It has been shown that the pol II elongation rate increases during transcription (42) and CDK12 may therefore play a key role in pol II speeding up as it travels away from the EEC, explaining the increase in the elongation defect as pol II moves further from the TSS. Expression of long genes will therefore be disproportionately affected by an elongation defect and this will be amplified over time. This would therefore explain why CDK12 knockdown experiments show a loss of expression of long genes, in particular DDR genes. In addition, as CDK12 makes important contacts with RNA processing and elongation factors (12,18,19), the long-term loss of the protein itself may contribute to the gene-type-specific defects in the level of mRNA from long DDR genes. A recent study with CDK12^{as} HCT116 cells also indicates that CDK12 kinase activity is specifically required for transcription of long DDR genes (13). However, in this case, pol II ChIP-seq was performed after 4.5 h inhibition, which will result in a bias in the analysis towards detection of a defect on long genes.

The effect of CDK12^{as} inhibition on elongation of transcription is readily explained by the loss of association of LEO1 and CDC73, which are key components of the PAF1C elongation complex, and the elongation factor

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

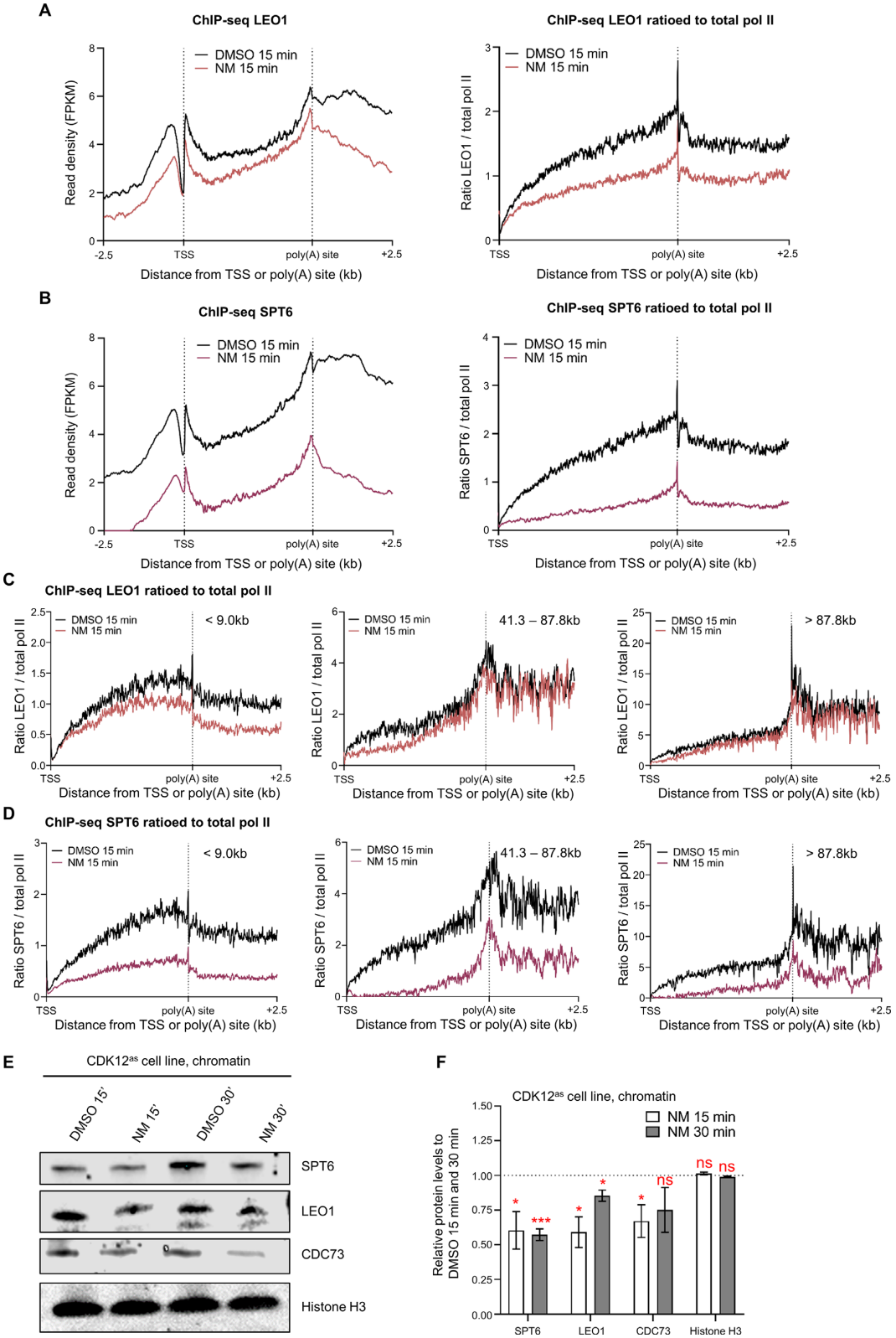

Figure 5. CDK12 activity is required for stable association of elongation factors with chromatin. (**A**) Meta-analysis of LEO1 subunit of the PAF1 complex (PAF1C) ChIP-seq across scaled expressed genes. (**B**) Meta-analysis of SPT6 ChIP-seq across scaled expressed genes. (**C**) Meta-analyses of LEO1 ChIP-seq ratioed to the total pol II signal before and after CDK12^{as} inhibition for 15 min across scaled expressed genes for the different length classes noted. (**D**) Meta-analyses of SPT6 ChIP-seq ratioed to the total pol II signal before and after CDK12^{as} inhibition for 15 min across scaled expressed genes for the different length classes noted. (**E**) Western blots of CDK12^{as} cell chromatin extracts. Cells are either treated with 7.5 μM 1-NM-PP1 or DMSO for 15 and 30 min as noted. The antibodies used are indicated on the right. Histone H3 was used as a loading control. (**F**) Quantitation of chromatin protein levels of SPT6, LEO1, CDC73, and histone H3 in CDK12^{as} cells relative to DMSO controls. Error bars = s.e.m. ($n = 3$ biological replicates). Statistical test: two-tailed unpaired t test, ns = not significant, * $P < 0.05$, *** $P < 0.001$.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

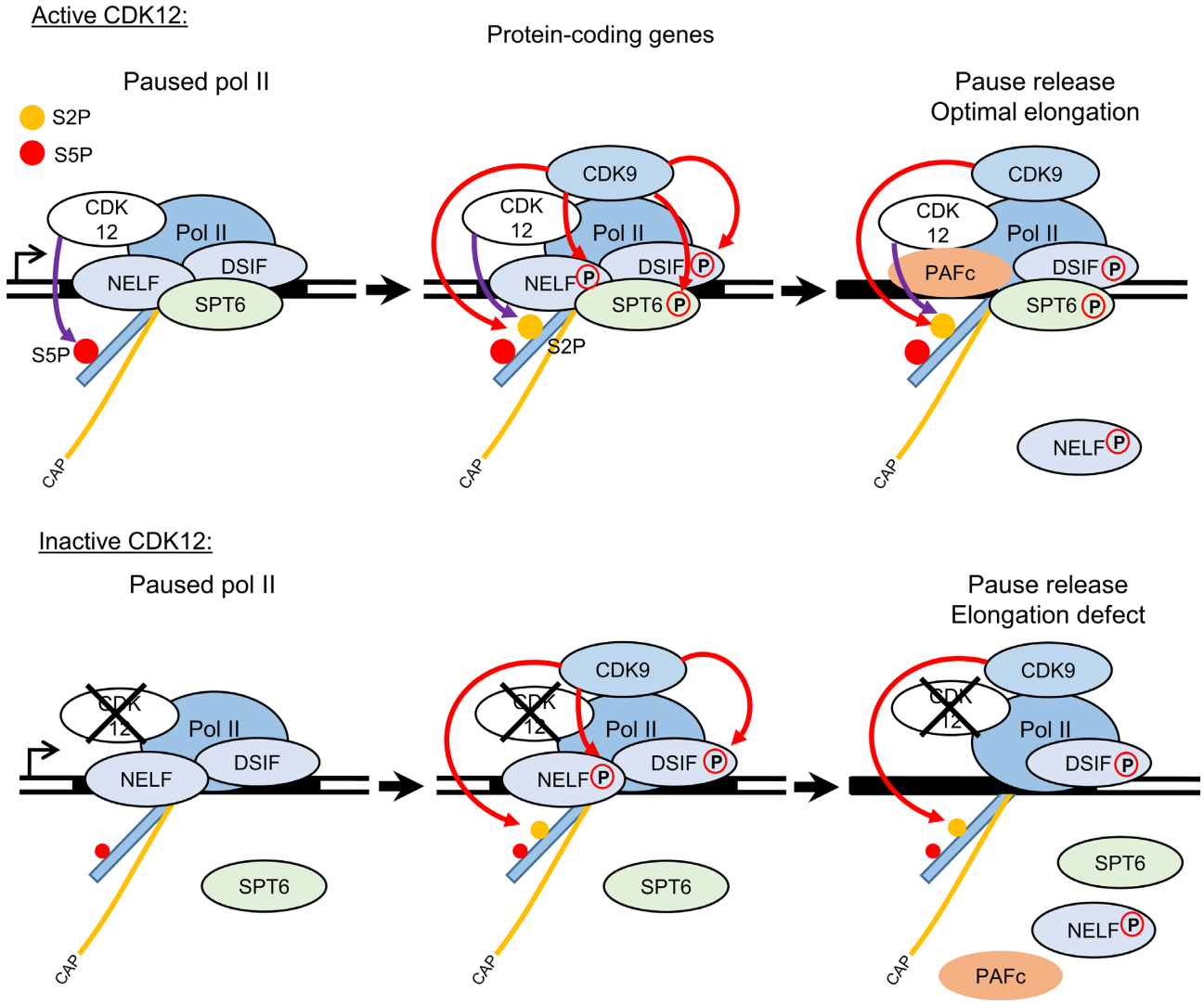

Figure 6. Model of CDK12 function in transcription. During the transcription cycle, CDK12 phosphorylates Ser2 and Ser5 of the pol II CTD and possibly other factors involved in transcription such as SPT6. Inhibition of CDK12 affects the pol II elongation rate by affecting the recruitment and/or stability on the chromatin of LEO1 and CDC73 from the PAF1 complex (PAF1C) and SPT6. CDK12 inhibition does not affect pol II pause release, which is mediated by CDK9. Orange line with 'cap': capped mRNA; orange and red dot: Ser2P and Ser5P, respectively. Red and violet arrows: phosphorylation mediated by CDK9 and CDK12, respectively.

SPT6. We cannot rule out that CDK12^{as} inhibition causes rapid nuclear export of these factors or that their association with chromatin appears reduced due to epitope masking caused by changes in post-translational modifications. However, CDK12as inhibition not only causes a decrease in the level of elongation factors on chromatin but an increased level in the nucleoplasm as would be expected if CDK12 activity is required for their effective recruitment to and/or stable association with pol II. CDK12 has already been shown to interact with the PAF1 complex (19). It has also recently been shown that the CDK12/CDK13 inhibitor, THZ531, affects phosphorylation of SPT6 (25). Phosphorylation of SPT6 by CDK12 could therefore play a role in contacts between SPT6 and other components of the elongation complex, including pol II and PAF1C (40). In turn, loss of SPT6 association may further destabilize the elongation complex (43). We therefore favour the notion that CDK12 inhibition causes reduced association of elongation factors with pol II through the loss of phosphorylation of one or more of its targets. Both loss of CDK12 and inhibition of CDK12 and CDK13 by THZ531 lead to the activation of gene-internal poly(A) sites and premature termination of transcription (10,25), which may be the direct result of slower elongation as suggested (25) (Figure 6).

We also observed a loss of the polyadenylation factor CPSF73 from newly-elongating pol II (Supplementary Figure S10). It was previously shown that CDC73 can interact with CPSF73 (44) and our data support this observation as both PAF1C and CPSF73 are lost from the newly-elongating pol II. CTD Ser2P also plays a role in CPSF73 recruitment (17). At the 3′ end of long genes, the retention of CDC73 and Ser2P would therefore be sufficient to recruit CPSF73 to/stabilize its interaction with the pol II elongation complex, in the absence of CDK12 activity. As Ser5P

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

is somewhat lost from the 3′ end of long genes, this mark is likely to be less important for PAF1C/CPSF recruitment at this point.

Notably, the effect of inhibiting CDK12 differs markedly from the effect of CDK9 inhibition, which causes a drastic loss of elongating pol II downstream of the early-elongation checkpoint (EEC) (31,45). CDK9 and CDK12 therefore play non-redundant roles in ensuring efficient transcription of pol II-dependent genes. However, these two kinases are intimately connected as CDK9 activity is required for PAF1C recruitment and PAF1C in turn recruits CDK12 (1,19). In addition, inhibition of both CDK12 and CDK9 causes premature termination close to poly(A) sites where Ser2P is peaking (45,46), indicating that the activity of both kinases is required for the correct transition between elongation and termination, through their phosphorylation of Ser2 or other targets.

CDK12 knockdown and THZ531 inhibition studies indicate that CDK12 phosphorylates Ser2 of the human pol II CTD (3,4,7), whereas *in vitro* CDK12 has robust Ser5 kinase activity (15,16). We find that detectable losses of Ser2P and Ser5P occur rapidly after CDK12^{as} inhibition, with the biggest effect on Ser5P. Relevant to this, specific inhibition of CDK12^{as} in HeLa cells results in a decrease in detection of pol II by both the H14 and H5 antibodies, which recognize Ser5P and Ser2P, respectively (14,47,48). A recent study with a CDK12^{as} HCT116 cell line observed a decrease in Ser5P but not of Ser2P following CDK12 inhibition (13). We have shown here that association of a Ser5P-binding capping factor with genes is affected by CDK12^{as} inhibition. Our results therefore support the notion that CDK12 is a major Ser5 kinase. However, based on previous findings, we would have expected inhibition of CDK12^{as} to cause a greater loss of Ser2P. Thus, in at least some human cells, the Ser2 kinase activity of CDK12 may be somewhat redundant with other Ser2 kinases, such as CDK9.

In conclusion, we have shown that inhibition of CDK12 causes a genome-wide pol II defect in pol II velocity and processivity and loss of PAF1C components and SPT6 from the newly-elongating pol II. Our findings also indicate that CDK12 is both a CTD Ser2 and Ser5 kinase *in vivo*. It remains to be determined whether CDK12 function in transcription elongation is mediated through pol II CTD phosphorylation and/or through phosphorylation of other targets such as SPT6.

DATA AVAILABILITY

mNET-seq and ChIP-seq data have been deposited in GEO under accession number GSE115290. TT-seq data have been deposited in GEO under accession number GSE133156.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics for sequencing and Professor David Greaves and Dr Daniel Lucy (Sir William Dunn School of Pathology) for the use of and assistance with the xCelligence cell analysis system.

Author contributions: M.T. and J.Z carried out most of the experimental analyses, M.T. carried out all the bioinformatic analysis of the mNET-seq and ChIP-seq data, L.C. and T.V. performed TT-seq, E.M. and B.S. carried out bioinformatic analysis of TT-seq data. I.F-V. and D.B. aided some experimental analyses and T.N. provided support with mNET-seq. S.M. produced materials, carried out experimental analyses and supervised most of the research. P.C. supervised the TT-seq analyses and helped with data interpretation. S.M., M.T. and J.Z. wrote the paper with contributions from all the authors.

FUNDING

Wellcome Trust Investigator Awards [WT106134AIA and WT210641/Z/18/Z to S.M.]; Cancer Research UK Oxford Centre [CRUKDF 0317-SMJZ to J.Z. and S.M.]; EMBO Long-Term Postdoctoral Fellowship [ALTF-1261-2014 to L.C.]; Deutsche Forschungsgemeinschaft [SFB860 and SPP1935 to P.C.]; European Research Council Advanced Investigator Grant TRANSREGULON [693023 to P.C.]; Volkswagen Foundation. Funding for open access charge: Wellcome Trust.

Conflict of interest statement. None declared.

REFERENCES

1. Zaborowska,J., Egloff,S. and Murphy,S. (2016) The pol II CTD: new twists in the tail. *Nat. Struct. Mol. Biol.*, **23**, 771–777.
2. Cho,E.J., Kobor,M.S., Kim,M., Greenblatt,J. and Buratowski,S. (2001) Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev.*, **15**, 3319–3329.
3. Bartkowiak,B., Liu,P., Phatnani,H.P., Fuda,N.J., Cooper,J.J., Price,D.H., Adelman,K., Lis,J.T. and Greenleaf,A.L. (2010) CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.*, **24**, 2303–2316.
4. Blazek,D., Kohoutek,J., Bartholomeeusen,K., Johansen,E., Hulinkova,P., Luo,Z., Cimermancic,P., Ule,J. and Peterlin,B.M. (2011) The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. *Genes Dev.*, **25**, 2158–2172.
5. Bowman,E.A., Bowman,C.R., Ahn,J.H. and Kelly,W.G. (2013) Phosphorylation of RNA polymerase II is independent of P-TEFb in the C. elegans germline. *Development*, **140**, 3703–3713.
6. Cheng,S.W., Kuzyk,M.A., Moradian,A., Ichu,T.A., Chang,V.C., Tien,J.F., Vollett,S.E., Griffith,M., Marra,M.A. and Morin,G.B. (2012) Interaction of cyclin-dependent kinase 12/CrkRS with cyclin K1 is required for the phosphorylation of the C-terminal domain of RNA polymerase II. *Mol. Cell. Biol.*, **32**, 4691–4704.
7. Zhang,T., Kwiatkowski,N., Olson,C.M., Dixon-Clarke,S.E., Abraham,B.J., Greifenberg,A.K., Ficarro,S.B., Elkins,J.M., Liang,Y., Hannett,N.M *et al.*2016) Covalent targeting of remote cysteine residues to develop CDK12 and CDK13 inhibitors. *Nat. Chem. Biol.*, **12**, 876–884.
8. Schuller,R., Forne,I., Straub,T., Schreieck,A., Texier,Y., Shah,N., Decker,T.M., Cramer,P., Imhof,A. and Eick,D. (2016) Heptad-specific phosphorylation of RNA polymerase II CTD. *Mol. Cell*, **61**, 305–314.
9. Marshall,N.F. and Price,D.H. (1995) Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J. Biol. Chem.*, **270**, 12335–12338.
10. Dubbury,S.J., Boutz,P.L. and Sharp,P.A. (2018) CDK12 regulates DNA repair genes by suppressing intronic polyadenylation. *Nature*, **564**, 141–145.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020

11. Booth,G.T., Parua,P.K., Sanso,M., Fisher,R.P. and Lis,J.T. (2018) Cdk9 regulates a promoter-proximal checkpoint to modulate RNA polymerase II elongation rate in fission yeast. *Nat. Commun.*, **9**, 543.
12. Liang,K., Gao,X., Gilmore,J.M., Florens,L., Washburn,M.P., Smith,E. and Shilatifard,A. (2015) Characterization of human cyclin-dependent kinase 12 (CDK12) and CDK13 complexes in C-terminal domain phosphorylation, gene transcription, and RNA processing. *Mol. Cell. Biol.*, **35**, 928–938.
13. Chirackal Manavalan,A.P., Pilarova,K., Kluge,M., Bartholomeeusen,K., Rajecky,M., Oppelt,J., Khirsariya,P., Paruch,K., Krejci,L., Friedel,C.C. *et al.* (2019) CDK12 controls G1/S progression by regulating RNAPII processivity at core DNA replication genes. *EMBO Rep.*, **20**, e47592.
14. Bartkowiak,B., Yan,C. and Greenleaf,A.L. (2015) Engineering an analog-sensitive CDK12 cell line using CRISPR/Cas. *Biochim. Biophys. Acta*, **1849**, 1179–1187.
15. Bosken,C.A., Farnung,L., Hintermair,C., Merzel Schachter,M., Vogel-Bachmayr,K., Blazek,D., Anand,K., Fisher,R.P., Eick,D and Geyer,M. (2014) The structure and substrate specificity of human Cdk12/Cyclin K. *Nat. Commun.*, **5**, 3505.
16. Bartkowiak,B. and Greenleaf,A.L. (2015) Expression, purification, and identification of associated proteins of the full-length hCDK12/CyclinK complex. *J. Biol. Chem.*, **290**, 1786–1795.
17. Davidson,L., Muniz,L. and West,S. (2014) 3′ end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes Dev.*, **28**, 342–356.
18. Eifler,T.T., Shao,W., Bartholomeeusen,K., Fujinaga,K., Jager,S., Johnson,J.R., Luo,Z., Krogan,N.J. and Peterlin,B.M. (2015) Cyclin-dependent kinase 12 increases 3′ end processing of growth factor-induced c-FOS transcripts. *Mol. Cell. Biol.*, **35**, 468–478.
19. Yu,M., Yang,W., Ni,T., Tang,Z., Nakadai,T., Zhu,J. and Roeder,R.G. (2015) RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. *Science*, **350**, 1383–1386.
20. Juan,H.C., Lin,Y., Chen,H.R. and Fann,M.J. (2016) Cdk12 is essential for embryonic development and the maintenance of genomic stability. *Cell Death Differ.*, **23**, 1038–1048.
21. Joshi,P.M., Sutor,S.L., Huntoon,C.J. and Karnitz,L.M. (2014) Ovarian cancer-associated mutations disable catalytic activity of CDK12, a kinase that promotes homologous recombination repair and resistance to cisplatin and poly(ADP-ribose) polymerase inhibitors. *J. Biol. Chem.*, **289**, 9247–9253.
22. Ekumi,K.M., Paculova,H., Lenasi,T., Pospichalova,V., Bosken,C.A., Rybarikova,J., Bryja,V., Geyer,M., Blazek,D. and Barboric,M. (2015) Ovarian carcinoma CDK12 mutations misregulate expression of DNA repair genes via deficient formation and function of the Cdk12/CycK complex. *Nucleic Acids Res.*, **43**, 2575–2589.
23. Cancer Genome Atlas Research, N. (2011) Integrated genomic analyses of ovarian carcinoma. *Nature*, **474**, 609–615.
24. Choi,S.H., Martinez,T.F., Kim,S., Donaldson,C., Shokhirev,M.N., Saghatelian,A. and Jones,K.A. (2019) CDK12 phosphorylates 4E-BP1 to enable mTORC1-dependent translation and mitotic genome stability. *Genes Dev.*, **33**, 418–435.
25. Krajewska,M., Dries,R., Grassetti,A.V., Dust,S., Gao,Y., Huang,H., Sharma,B., Day,D.S., Kwiatkowski,N., Pomaville,M. *et al.* (2019) CDK12 loss in cancer cells affects DNA damage response genes through premature cleavage and polyadenylation. *Nat. Commun.*,**10**, 1757.
26. Blethrow,J., Zhang,C., Shokat,K.M. and Weiss,E.L. (2004) Design and use of analog-sensitive protein kinases. *Curr. Protoc. Mol. Biol.*, Chapter 18, Unit 18 11.
27. Schwalb,B., Michel,M., Zacher,B., Fruhauf,K., Demel,C., Tresch,A., Gagneur,J. and Cramer,P. (2016) TT-seq maps the human transient transcriptome. *Science*, **352**, 1225–1228.
28. Nojima,T., Gomes,T., Grosso,A.R.F., Kimura,H., Dye,M.J., Dhir,S., Carmo-Fonseca,M. and Proudfoot,N.J. (2015) Mammalian NET-Seq reveals genome-wide Nascent transcription coupled to RNA processing. *Cell*, **161**, 526–540.
29. Nojima,T., Gomes,T., Carmo-Fonseca,M. and Proudfoot,N.J. (2016) Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. *Nat. Protoc.*, **11**, 413–428.
30. Laitem,C., Zaborowska,J., Tellier,M., Yamaguchi,Y., Cao,Q., Egloff,S., Handa,H. and Murphy,S. (2015) CTCF regulates NELF, DSIF and P-TEFb recruitment during transcription. *Transcription*, **6**, 79–90.
31. Gressel,S., Schwalb,B., Decker,T.M., Qin,W., Leonhardt,H., Eick,D. and Cramer,P. (2017) CDK9-dependent RNA polymerase II pausing controls transcription initiation. *Elife*, **6**, e29736.
32. Dolken,L., Ruzsics,Z., Radle,B., Friedel,C.C., Zimmer,R., Mages,J., Hoffmann,R., Dickinson,P., Forster,T., Ghazal,P. *et al.* (2008) High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA*, **14**, 1959–1972.
33. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
34. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
35. 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.
36. Lopez,M.S., Kliegman,J.I. and Shokat,K.M. (2014) The logic and design of analog-sensitive kinases and their small molecule inhibitors. *Methods Enzymol.*, **548**, 189–213.
37. Tien,J.F., Mazloomian,A., Cheng,S.G., Hughes,C.S., Chow,C.C., Canapi,L.T., Oloumi,A., Trigo-Gonzalez,G., Bashashati,A., Xu,J. *et al.* (2017) CDK12 regulates alternative last exon mRNA splicing and promotes breast cancer cell invasion. *Nucleic Acids Res.* **45**, 6698–6716.
38. Ho,C.K. and Shuman,S. (1999) Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell*, **3**, 405–411.
39. Yang,Y., Li,W., Hoque,M., Hou,L., Shen,S., Tian,B. and Dynlacht,B.D. (2016) PAF complex plays novel subunit-specific roles in alternative cleavage and polyadenylation. *PLos Genet.*, **12**, e1005794.
40. Vos,S.M., Farnung,L., Boehning,M., Wigge,C., Linden,A., Urlaub,H. and Cramer,P. (2018) Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature*, **560**, 607–612.
41. Jonkers,I. and Lis,J.T. (2015) Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.*, **16**, 167–177.
42. Jonkers,I., Kwak,H. and Lis,J.T. (2014) Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife*, **3**, e02407.
43. Nojima,T., Tellier,M., Foxwell,J., Ribeiro de Almeida,C., Tan-Wong,S.M., Dhir,S., Dujardin,G., Dhir,A., Murphy,S. and Proudfoot,N.J. (2018) Deregulated expression of mammalian lncRNA through loss of SPT6 induces R-loop formation, replication stress, and cellular senescence. *Mol. Cell*,**72**, 970–984.
44. Rozenblatt-Rosen,O., Nagaike,T., Francis,J.M., Kaneko,S., Glatt,K.A., Hughes,C.M., LaFramboise,T., Manley,J.L. and Meyerson,M. (2009) The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3′ mRNA processing factors. *PNAS*, **106**, 755–760.
45. Laitem,C., Zaborowska,J., Isa,N.F., Kufs,J., Dienstbier,M. and Murphy,S. (2015) CDK9 inhibitors define elongation checkpoints at both ends of RNA polymerase II-transcribed genes. *Nat. Struct. Mol. Biol.*, **22**, 396–403.
46. Tellier,M., Ferrer-Vicens,I. and Murphy,S. (2016) The point of no return: the poly(A)-associated elongation checkpoint. *RNA Biol*, **13**, 265–271.
47. Chapman,R.D., Heidemann,M., Albert,T.K., Mailhammer,R., Flatley,A., Meisterernst,M., Kremmer,E. and Eick,D. (2007) Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science*, **318**, 1780–1782.
48. Heidemann,M., Hintermair,C., Voss,K. and Eick,D. (2013) Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim. Biophys. Acta*, **1829**, 55–62.

Downloaded from https://academic.oup.com/nar/article/48/14/7712/5858455 by MPI Biophysical Chemistry user on 28 September 2020