Supporting Online Material

Threonine-4 of Mammalian RNA Polymerase II CTD is Targeted by Polo-like Kinase-3 and Required for Transcriptional Elongation

Corinna Hintermair,[1*] Martin Heidemann,[1*] Frederic Koch,[2*] Nicolas Descostes,[2] Marta Gut,[3] Ivo Gut,[3] Romain Fenouil,[2] Pierre Ferrier,[2] Andrew Flatley,[4] Elisabeth Kremmer,[4] Rob D. Chapman,[1] Jean-Christophe Andrau,[2§] Dirk Eick[1§]

[1]Department of Molecular Epigenetics, Helmholtz Center Munich, Center of Integrated Protein Science (CIPSM), Munich, Germany.

[2]Centre d'Immunologie de Marseille-Luminy, Université Aix-Marseille, Campus de Luminy, France.

[3]Centre Nacional D'Anàlisi Genòmica, Parc Cientific de Barcelona, Barcelona, Spain.

[4]Institute of Molecular Immunology, Helmholtz Center Munich, Munich, Germany.

[*]contributed equally

Materials and methods

ChIPseq

All experiments were performed using Protein-G coated Dynabeads (Invitrogen). Per ChIP, 400µl beads were washed 3x with 1ml and subsequently resuspended in 250µl of cold blocking solution (0.5% BSA in 1x DPBS). After the addition of 40µg of antibody, the beads were incubated at 4°C overnight on a rotating wheel. Unbound antibodies were removed through three further 1ml washes and resuspension in 100µl of cold blocking solution. Extracts equivalent to $1x10^8$ cells (3ml) were added and the mix was incubated overnight at 4°C on a rotating wheel. Beads were washed 8x with RIPA buffer (50mM Hepes pH 7.6, 500mM LiCl, 10mM EDTA pH 8.0, 1% NP-40, 0.7% Na-Deoxycholate) and 1x with TE+ (10mM Tris pH 8.0, 1mM EDTA pH 8.0, 50mM NaCl). Final concentrations of 1x phosphatase inhibitors, 1x EDTA-free protease inhibitors (Roche), 0.2mM PMSF and 1µg/ml pepstatin were added to both washing buffers. Immunoprecipitated chromatin was recovered from the beads in two elutions at 65°C for 15 and 10min in 110µl and 100µl of elution buffer (50mM Tris pH 8.0, 10mM EDTA pH 8.0, 1% SDS) respectively. The two eluates were combined and DNA was purified as described for the input. ChIP DNA was quantified using 20% of the material in High Sensitivity DNA chips on a 2100 Bioanalyzer (Agilent). At least 1ng of ChIP DNA was used for library preparation according to the Illumina ChIP-seq protocol. Input DNA was sequenced to serve as a control (see below).

Antibody (clone)	Reference	Approx. DNA quantity
total (N-20)	Santa Cruz (sc-899x)	22ng
Ser2P (3E10)	Chapman R. *et al*., 2007	1ng
Ser5P (3E8)	Chapman R. *et al*., 2007	3ng
Thr4P (6D7)	*this work*	2ng
HA	Abcam (ab9110)	6ng

After end-repair, adenylation and adapter ligation, fragments were size selected on an E-Gel SizeSelect gel (Invitrogen) and PCR-amplified. The prepared library was submitted to a single read Genome Analyzer flowcell for clustering on a c-Bot (Illumina) fluidics robot and subsequently sequenced on an Illumina GAIIx instrument with a standard 36-base pair single read run protocol using SCS2.6_RTA1.6 data collection and analysis software. The rough yield was 22 million tags per lane on the Illumina SR flow cell. One lane was used for each sample. Primary data analysis was carried out with the standard Illumina pipeline.

Data Processing and Wiggle File Generation

All samples were aligned against the UCSC 2006 human genome (hg18) using the integrated Eland software from Illumina. Only uniquely mapped tags were used for further processing and all duplicate tags with identical sequences/coordinates were filtered out at a threshold of one per 7 million tags to remove possible sequencing

and/or alignment artifacts. Remaining tags were processed using a custom R pipeline (Koch et al, 2011).

In brief, the Watson and Crick strand tags were merged after 3' elongation/size extension to the gel purified fragment size, estimated *in silico* for each experiment individually. Binned wiggle files for genome-wide scores were generated by using the average score every 50bp. Furthermore, for intra-experimental normalization, we used the genome-wide average binding score for each experiment to estimate the background level and rescaled the scores accordingly. Finally, to correct for possible (un-) favorable events during sonication and/or DNA sequencing, we employed an input subtraction step for each experiments using the normalized input control. Box plots were generated using a accustomed script in R integrating the non parametric Mann-Whitney-Wilcoxon test (Wilcox.test function).

Average Profile Generation

To generate average binding profiles, hg18 Refseq gene annotations were used to select all genes with a distance of at least 4kb from any other known annotation (mRNAs, snRNAs, snoRNAs, tRNAs). Based on this gene selection, we used custom R scripts to retrieve scores from wiggle files for 4kb or 1.5kb/3kb around the TSS and 3' end as well as across the gene bodies. Values were re-centered using the following linear interpolation function, resulting in 1000 points for each of these three regions of selected genes: $y = y_0 + (x - x_0)\frac{y_1 - y_0}{x_1 - x_0}$, where *y* refers to the interpolated values at coordinates *x* considering the neighboring values y_1 and y_0 at respective coordinates x_1 and x_0 in the dataset.

We normalized for varying gene sizes by interpolating coordinates on gene bodies from their respective TSS to the 3' ends. As a result, we obtained matrices with 1000

columns corresponding to "normalized coordinates" for every gene. The average of all columns returned the composite profile for each selected genomic region. Different profiling categories (top 5%, 5%-20%, 20%-80%, 80%-95% and bottom 5%) were based on the average ChIP signal from 4kb upstream of the TSS to 4kb downstream of the 3' end of genes for each experiment. The 7826 selected genes were thus divided into groups of 392, 1174, 4695, 1173 and 392 respectively, as represented by differentially colored lines.

Supplementary figures

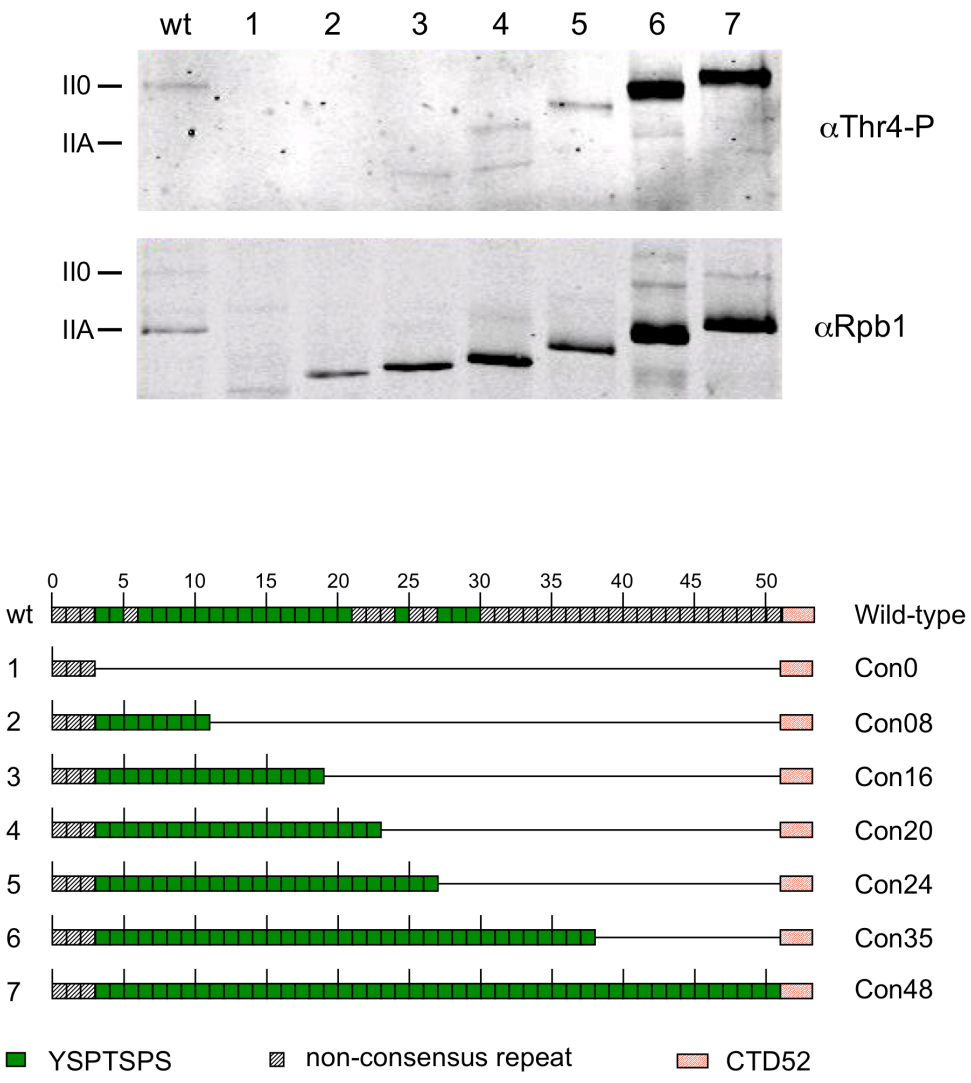


Figure S1

CTD phosphorylation is affected by its length. *(A)* Western blot of proteins extracted from Raji cell lines expressing wt, or CTD mutants of varying length (1-7). Membranes were screened with anti-Rpb1 to reveal exogenously expressed α-amanitin resistant, and any remaining endogenous polymerase, or with mAb (6D7) to reveal Thr4 phosphorylation. *(B)* Map of CTD length mutants. CTD required a minimal length of >24 repeats for proper Thr4 phosphorylation.

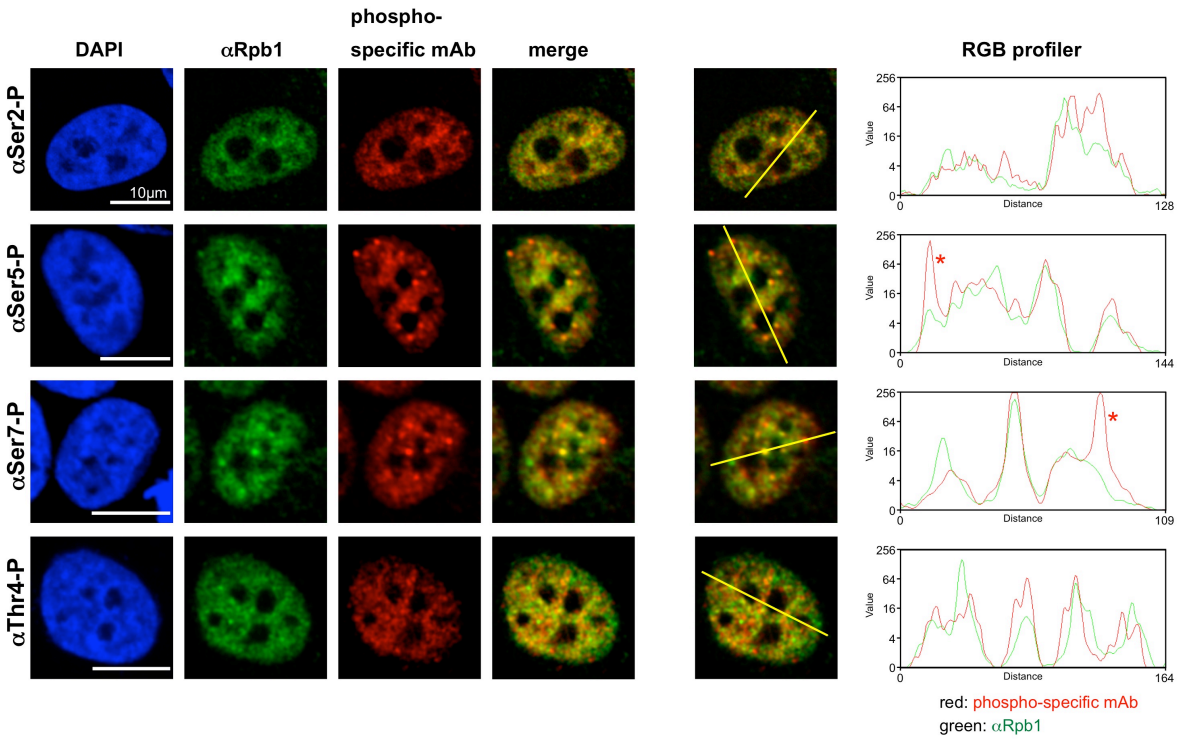


Figure S2

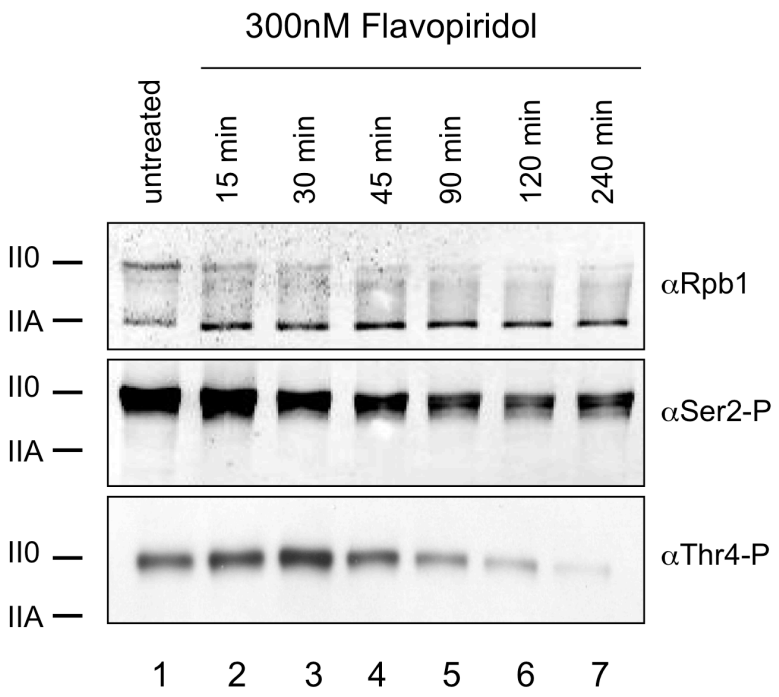
Nuclear distribution of Pol II0 populations. The POL3/3 mAb recognizes a Pol II epitope outside of the CTD and allows the visualization of the distribution of total Pol II in cells (green). In addition, cells were co-stained with a Ser2-P, Ser5-P, Ser7-P, or Thr4-P specific mAbs to show the abundance and nuclear distribution of individual CTD modifications (red). Both pictures were merged and the green and red signals were quantified in RGB profiles. Mergers revealed a co-localization of signals for Pol II and Pol II modifications, but also displayed nuclear areas were either the signal for Pol II (green) or the signal of a specific modification (red) prevailed. This probably reflects local differences in levels of Pol II modification, which are most apparent for Pol II populations with Ser5-P and Ser7-P modifications (marked by stars in RGB profiler) in distinct nuclear spots. The immunostains confirm the results of immunoprecipitation experiments that in addition to Ser2-P, Ser5-P and Ser7-P a large portion of Pol II is also modified by Thr4-P.

0132-0000-2	AKT1	0433-0000-1	MARK3
0300-0000-1	ARK5	0385-0000-1	MEK1 (SESE)
0166-0000-1	Aurora A	0550-0000-3	MEK1 wt
0190-0000-1	Aurora B	0380-0000-1	MAPKAPK5
0315-0000-1	Aurora C	0396-0000-1	MKK6
0403-0000-1	B-RAF	0286-0000-1	MST4
0181-0000-1	BRK	0445-0000-1	MYLK
0406-0000-1	Casein Kinase 1 alpha 1	0290-0000-1	NEK2
0124-0000-1	Casein Kinase 2 alpha 1	0209-0000-1	NEK6
0412-0000-1	Casein Kinase 2 alpha 2	0253-0000-1	NLK
0382-0000-1	CDC42BPB Kinase	0344-0000-1	NIK
0134-0135-1	CDK 1/Cyclin B	0443-0000-1	p38-alpha
0134-0055-1	CDK 1/Cyclin E	0357-0000-1	PAK1
0050-0054-1	CDK 2/Cyclin A	0185-0000-2	PBK
0050-0055-1	CDK 2/Cyclin E	0189-0000-1	PCTAIRE
0212-0055-1	CDK 3/Cyclin E	0353-0000-1	PDK1
0142-0143-1	CDK 4/Cyclin D1	0186-0000-1	PIM1
0142-0373-1	CDK 4/Cyclin D3	0223-0000-1	PIM2
0356-0389-1	CDK 5/p25NCK	0437-0000-1	PLK3
0356-0355-1	CDK 5/p35NCK	0183-0000-1	PLK1
0051-0143-1	CDK 6/Cyclin D1	0222-0000-1	PKC-alpha
0366-0360-4	CDK 7/Cyclin H	0232-0000-1	PKC-delta
0376-0390-1	CDK 8/Cyclin C	0420-0000-1	PKC-eta
0371-0345-1	CDK 9/Cyclin T	0204-0000-1	PKC-gamma
0282-0000-1	CHK1	0115-0000-1	PKC-mu
0447-0000-1	CLK1	0207-0000-1	PRK1 protein kinase N1
0180-0000-1	COT (MAP3K8)	0347-0000-1	ROCK2
0332-0000-1	DAPK1	0318-0000-2	SK6 S6 kinase
0634-0000-7	ERK2	0306-0000-1	SAK
0533-0000-1	FER	0199-0000-2	SGK1
0446-0000-1	FRK (RAK)	0198-0000-2	SGK2
0310-0000-1	GSK3-beta	0414-0000-1	SNARK
0444-0000-1	HRI	0200-0000-1	SRC
0258-0000-1	IKK-beta	0277-0000-1	SNK
0320-0000-1	IKK-epsilon	0434-0000-1	SRPK1
0268-0000-1	IARK4	0381-0000-1	SRPK2
0419-0000-1	JNK1	0187-0000-1	TSF1
0214-0000-1	JNK3	0220-0000-1	TSK2
0633-0000-1	MAPKAPK3	0284-0000-1	TTK
0432-0000-1	MARK1	0302-0000-1	VRK1

Figure S3

List of 80 commercial kinases and cyclin/kinase pairs expressed and purified from Baculovirus (Proqinase, Freiburg), DNA-PK was purchased from Promega (V5811). Each kinase or cyclin/kinase pair (each 100ng) was tested for specificity of CTD threonine-4 phosphorylation in ELISA experiments (in the recommended buffer of the supplier). High phosphorylation activity towards Thr4 was observed for Plk3 (Plk1 showed low activity), but not for other kinases. High activity of Plk3 for Thr4 phosphorylation was confirmed with purified Pol IIA as substrate (see Figure 4A, lane 4). Since the specific activities of kinases of the library were tested with individual substrates, a comparison of kinases activities is not possible.

A



B

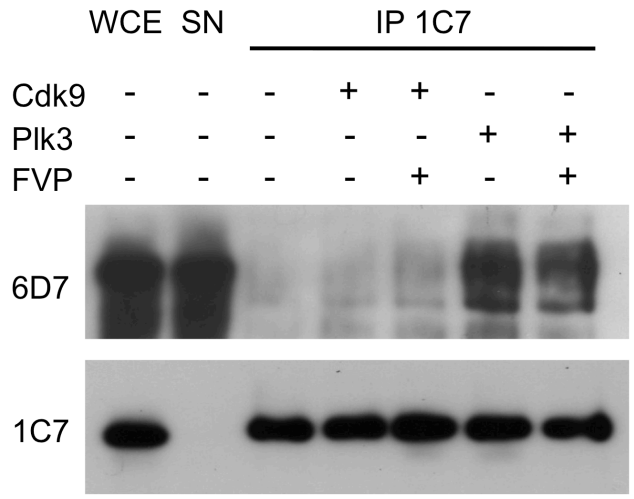


Figure S4

(A) Flavopiridol inhibits phosphorylation of CTD Ser2 and Thr4 residues in U2OS cells.Western blot analysis of USOS cell extract after flavopiridol treatment (300nM). Rpb1 (POL3/3), Ser2-P (3E10), and Thr4-P (6D7) specific mAb.
(B) In vitro kinase assay with purified Rpb1 (IIA). Flavopiridol (FVP, 1000 nM) does not inhibit threonine4 phosphorylation by Plk3.

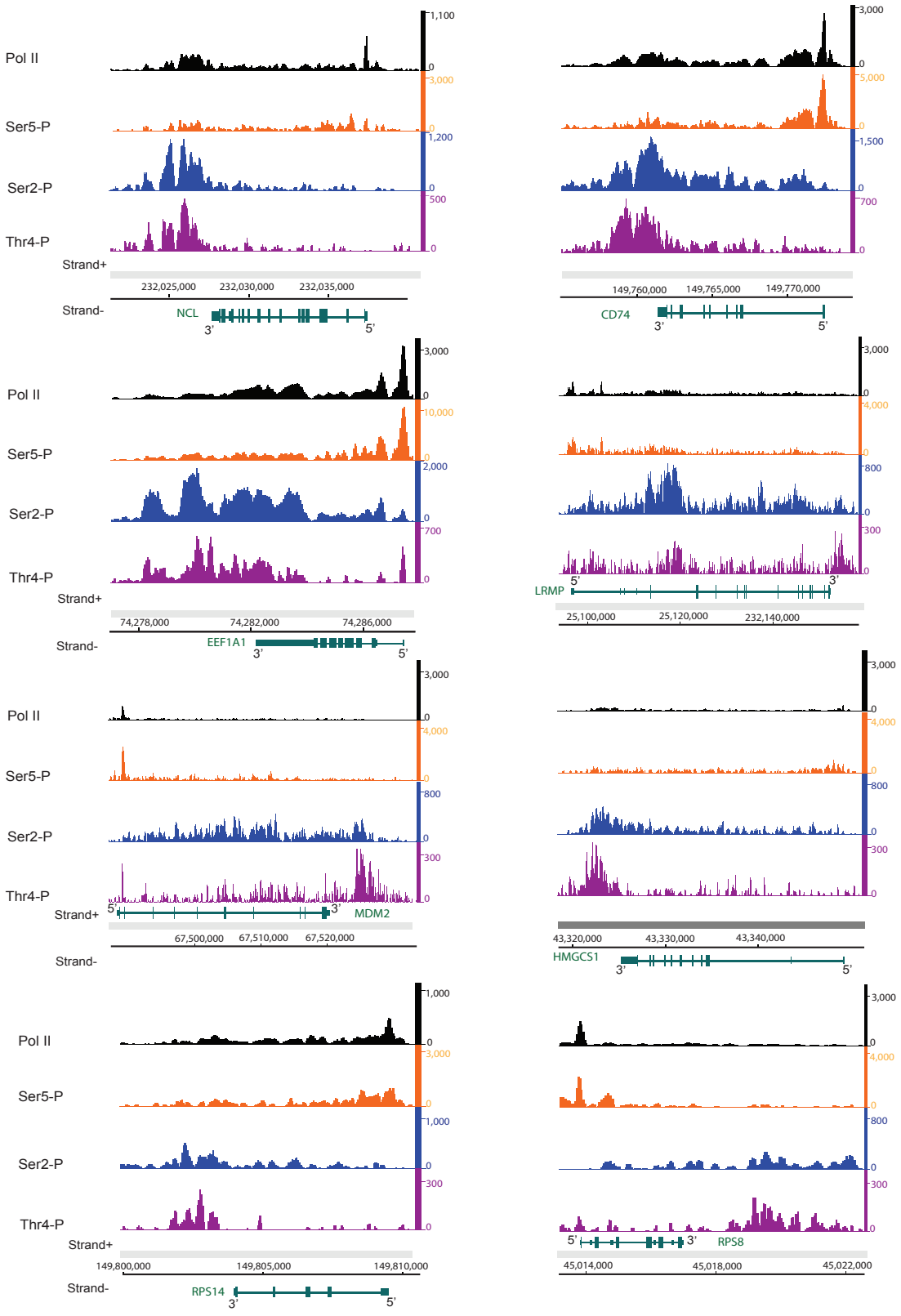


Figure S5

ChIPseq profiles for total Pol II, Thr4-P, Ser5-P and Ser2-P marks for selected genes. Normalized ChIPseq signals for each experiment are shown on the right. Genes on positive (+) and negative (-) strand are indicated below the ChIPseq lanes.

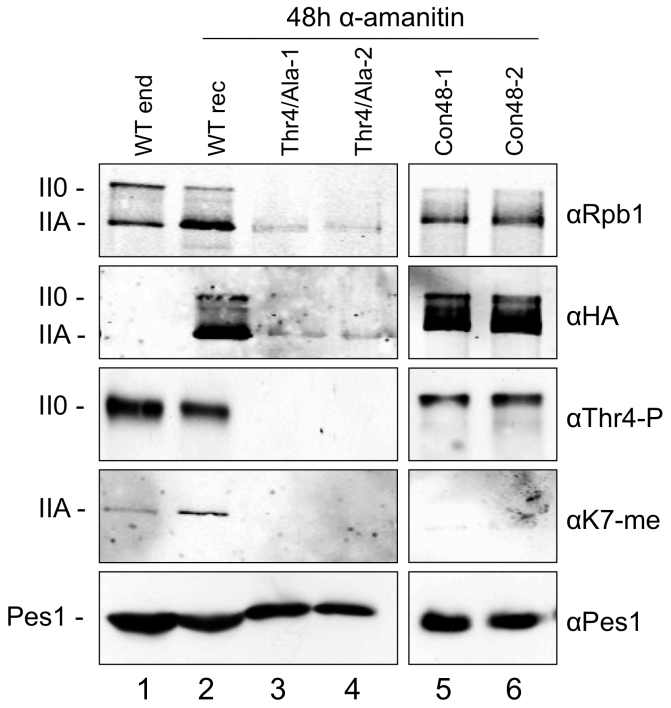


Figure S6

Conditional expression of wild-type, Con48, and Thr4/Ala mutants in Raji cells. Recombinant polymerase were induced by removal of tetracycline. After 24h α-amanitin was added (2µg/ml) and expression levels of endogenous and recombinant polymerases were analysed after additional 48h. While wild-type and the Con48 mutants are overexpressed (both support their own expression in the presence of α-amanitin), expression levels of the Thr4/Ala mutant is below the levels of the endogenous Pol II. The α-K7-me antibody recognizes methylated lysine residues in endogenous and wild-type polymerase, and serves as control for the degradation of endogenous Pol II. αHA detects all recombinant Rpb1, Pes1 serves as loading control.

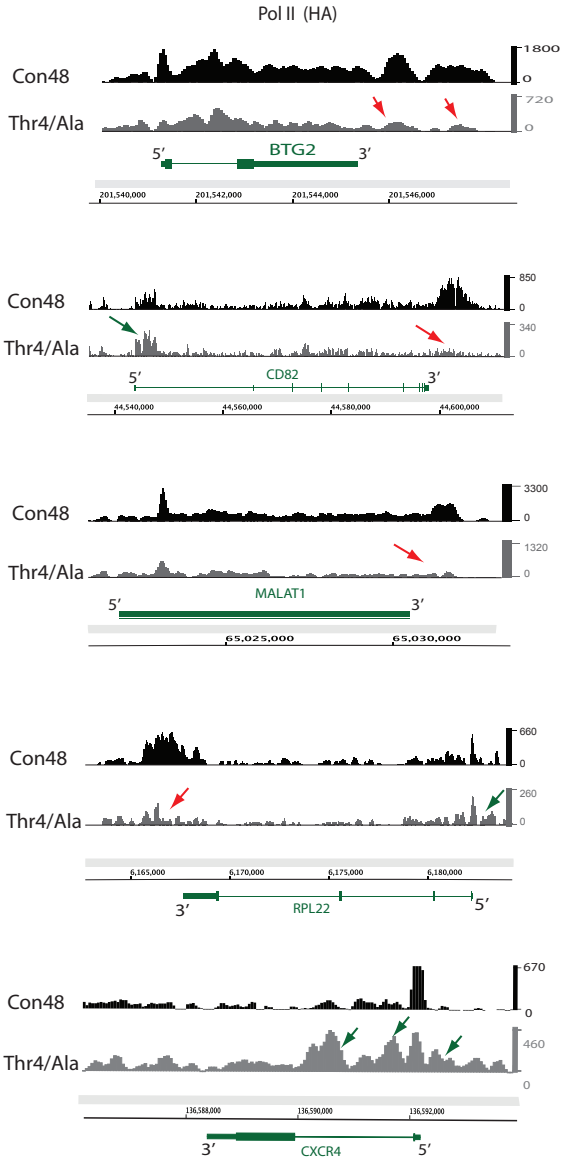


Figure S7

ChIPseq profiles for HA-tagged Pol II in mutants Con48 and Thr4/Ala for genes BTG2, CD82, MALAT1, RPL22, and CXCR4. Arrows indicate relative increase (green) and decrease (red) of signals.

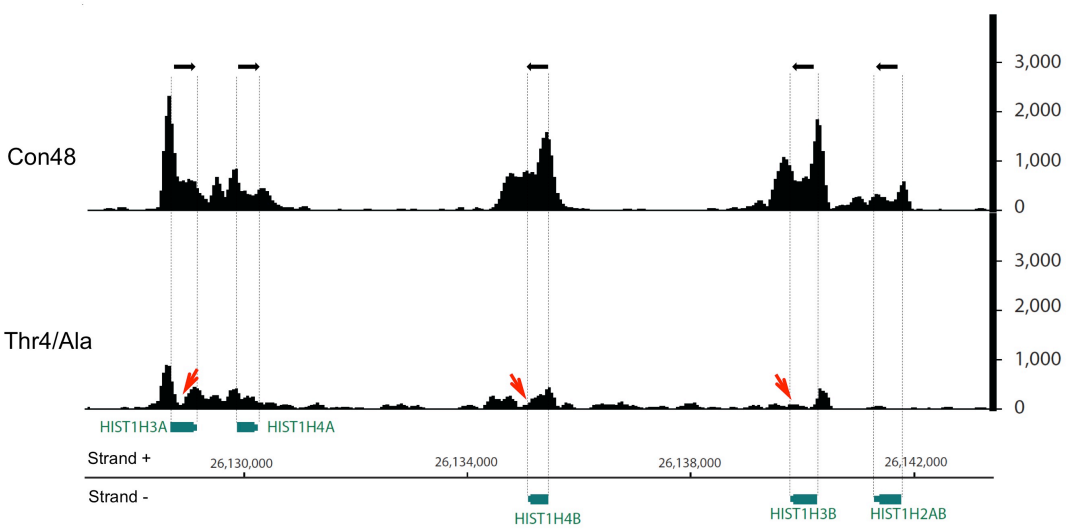


Figure S8

ChIPseq profiles for HA-tagged Pol II in mutants Con48 and Thr4/Ala on histone gene cluster HIST1H3A, HIST1H4A, HIST1H4B, HIST1H3B, and HIST1H2AB. In Con48 mutants the promoter/gene body ratio of Pol II is increased for HIST1H3A, HIST1H4B, and HIST1H3B (red arrows). Black arrows indicate direction of transcription.

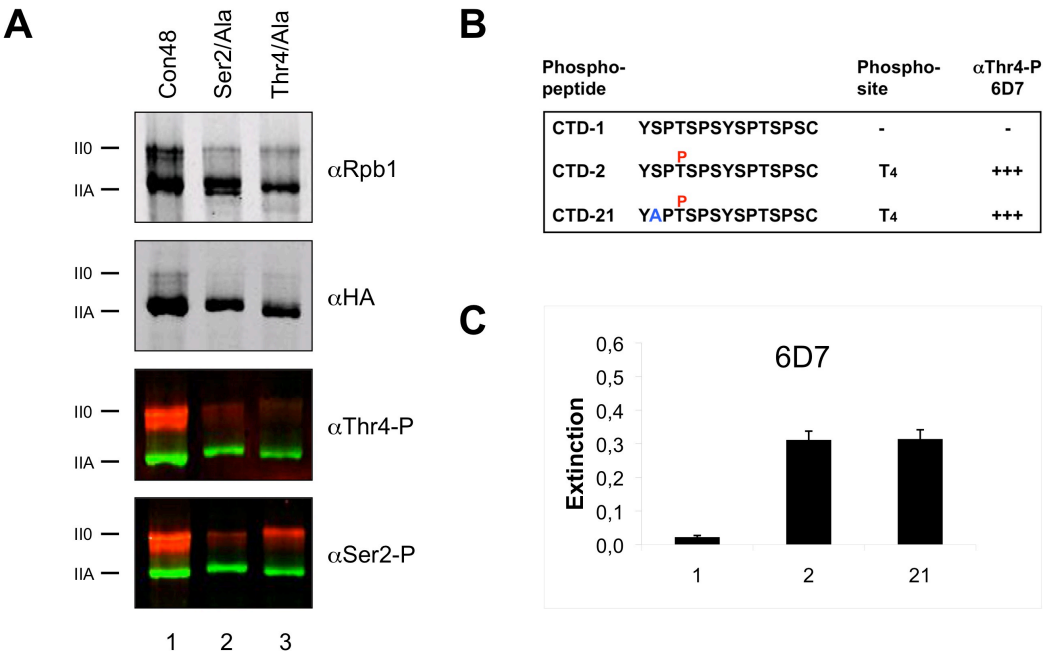


Figure S9

Ser2 is a prerequisite for Thr4 phosphorylation. (*A*) Ser2/Ala mutant was expressed in Raji cells (for details see Figures 1 and 6) and analyzed for Thr4 phosphorylation. Strong Thr4-P signals were detected in Con48 cells, while Ser2/Ala and Thr4/Ala show only background signals for Thr4P. (*B,C*) Antibody 6D7 shows the same reactivity for peptides CTD-2 and CTD-21, indicating that replacement of Ser2 by Ala does not inhibit antibody binding. We conclude that Thr4 is not phosphorylated in mutant Ser2/Ala in Raji cells. Green signals, αHA; red signals, modification specific mAbs.

References

Koch F, Fenouil R, Gut M, Cauchy P, Albert TK, Zacarias-Cabeza J, Spicuglia S, de la Chapelle AL, Heidemann M, Hintermair C, Eick D, Gut I, Ferrier P, Andrau JC (2011) Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol* **18:** 956-963