Supplementary Note

Transient transcriptome sequencing (TT-seq)

Transient transcriptome sequencing (TT-seq) was performed in biological duplicates as described[1]. 150mio KBM7 MED14-dTAG cells were treated for 1-2h with 500nM dTAG7 or DMSO. For combined MED14/CDK9 perturbation experiments, 200mio cells were treated for 2h with 500nM dTAG7 or DMSO, 2h DMSO and adding 500nM NVP2 for the last 30min (total 2h), or 2h 500nM dTAG7 and adding 500nM NVP2 for the last 30min (total 2h). In the last 5min, newly synthesized RNA was labeled by addition of 500µM 4-thio-uridine (4sU; Jena Bioscience) from 1M DMSO stocks directly to the medium and incubation at 37°C. Cells were decanted into 50mL tubes and harvested by starting centrifugation at 3000g for 2min after exactly 5min of labeling. Supernatant was decanted and cells lysed by addition of 5mL Qiazol. After 10min incubation with occasional vortexing, lysates were stored at -80°C until further processing. This procedure was performed in staggered batches of two samples to ensure accurate treatment and labeling times.

TT-seq experiments in MED14-dTAG HCT-116 cells were performed in biological duplicates as described for KBM7 cells with minor changes. Treatments with 500nM dTAG^{V}-1 ligand were performed on five 15cm dishes per condition with 20mio cells per plate by direct addition of 500µM ligand stock to 15mL supernatant. In the last 10 minutes of treatment, 4sU labeling was performed as described above. After labeling, supernatant was aspirated and 5mL Qiazol added directly to each plate. Cells were lysed by vigorously pipetting up and down and lysates from the five plates pooled in a single 50mL falcon tube per condition. Lysates were stored at -80°C until further processing.

Before RNA isolation, spike-ins were added to the QIAzol lysates (240 ng of spike-ins per 150mio cells) (ref. [1]). Total RNA was sonicated to obtain fragments of up to 10kb using Covaris S220 (parameters: water level 13, treatment time 10 seconds, peak power 100 W, cycles 200, duty cycle 1%). Labeled RNA was purified from 240µg (DMSO), 480µg (dTAG7), or 960µg (NVP2; dTAG7+NVP2) of total sonicated RNA as described[1,2]. Labeled RNA was captured with streptavidin beads (Miltenyi Biotec) and eluted with 100mM DL-Dithiothreitol (Sigma-Aldrich). Eluted RNA was transferred to RNeasy MinElute spin columns (QIAGEN) for RNA cleanup and DNase I treatment according to the manual. RNA concentration was measured with a Qubit RNA HS Assay Kit. Library preparation was performed using 100ng of RNA per sample with the Ovation® Universal RNA-Seq System according to the manufacturer's instructions. For first strand cDNA synthesis random primers were used only. Quality of DNA fragments was checked by Fragment Analyzer. Libraries were pooled equimolar and sequenced in 75bp paired-end mode using an Illumina NextSeq 550 sequencer (see Supplementary Table 6).

TT-seq raw data were processed essentially as described[3]. Raw reads were aligned to the human hg38 genome assembly using STAR v2.3.0 with a maximum of two mismatches per 100bp and filtered

with *samtools view -q 7 -f 2* (ref. [4,5]). All further processing was performed using the R/Bioconductor framework. Antisense bias, sequencing depth and cross-contamination rates were calculated as described, based on exogenous RNA spike-ins[3]. Reads were mapped to transcription units, which represent the union of all annotated UCSC RefSeq isoforms per gene. The number of transcribed bases per transcription unit was calculated as the sum of the coverage of evident (sequenced) fragment parts (read pairs only) for all fragments in addition to the sum of the coverage of the inner mate interval if not entirely overlapping a Refseq annotated intron (UCSC RefSeq GRCh38). Cross-sample normalization was performed by balancing signal at the end of genes, which are longer than polymerase travels during the duration of treatment, assuming an elongation speed of 2.4kb/min (ref. [6]). To that end, we calculated size factors on the last 25kbp of transcription units exceeding 175kbp and used these to correct for library size and sequencing depth variations[7].

Precision nuclear run-on sequencing (PRO-seq)

Precision nuclear run-on sequencing (PRO-seq) was performed in biological duplicates to map the genome-wide distribution of transcriptionally engaged RNA polymerases following published protocols[8]. To account for global differences in transcriptional output, we employed an exogenous spike-in approach: $3x10^6$ (5%) fresh *Drosophila* S2 cells were added to $57x10^6$ MED14-dTAG KBM7 cells immediately before preparing nuclei for the biotin-NTP run-on reaction. Samples were kept cool to prevent premature run-on while making nuclei. Cells were harvested and washed once with ice-cold PBS, before incubating for 5 min on ice with 40mL dounce buffer (10mM Tris-HCl pH7.4, 300mM sucrose, 3mM $CaCl_2$, 2mM $MgCl_2$, 0.1% Triton-X-100, 0.5mM DTT). The cell suspension was gently dounced 25 times in a 15mL dounce homogenizer on ice, transferred to a fresh 50mL tube and nuclei harvested by centrifugation. Nuclei were washed with 2x 5mL dounce buffer, transferring them to 15mL tubes after the first wash to facilitate aspiration of the supernatant. Pellets were now taken up in 400µL ice-cold nuclei storage buffer (10mM Tris-HCl pH8.0, 25% glycerol, 5mM $MgCl_2$, 0.1mM EDTA, 5mM DTT), snap-frozen in liquid nitrogen and stored at -80°C.

Four parallel nuclear run-on reactions were performed per sample, equaling roughly $15x10^6$ input cells per run-on. Run-on was performed as published with two biotin nucleotides, bio-11-CTP and bio-11-UTP (PerkinElmer). After Trizol extraction, RNA was ethanol-precipitated for 30 min at -20°C. Total RNA pellets from the four parallel run-on reactions per sample were now pooled in a total volume of 20µL H_2O at step 17 in (ref. [8]). RNA was heat denatured, base-hydrolyzed and gel filtrated before the first streptavidin bead purification as published. 3' VRA adapters (see Supplementary Table 8) were ligated in a PCR cycler for 4h at 20°C and then cooled to 4°C overnight. Biotinylated RNA was purified the next morning in a second streptavidin bead enrichment and enzymatically prepared for 5' adapter ligation by removing 5' caps with RppH before phosphorylation with T4 PNK. RNA was purified with Trizol and 5' VRA adapters were ligated overnight with the same cycler settings. The third morning,

RNA was again subjected to streptavidin enrichment before reverse transcription with Superscript III RT enzyme as published (see Supplementary Table 8 for primer sequences). cDNA was stored at -20°C until performing PCR amplification. Test PCRs with dilution series were performed as published, including an extra dilution that equivalents to 9 full-scale PCR cycles. PCR products were separated on 2% agarose gels and visualized with post-run ethidium bromide staining. Full-scale PCRs were performed in 2x 50µL per sample with sample-specific RPI-x barcoded primers (see Supplementary Table 8) with 9-14 cycles as published. Sufficient amplification was verified by loading 7.5µL PCR product on an agarose gel before DNA precipitation with salt/ethanol and resuspending DNA in 30µL H_2O.

Contrary to the published protocol, unused primers and primer-dimers were removed by performing 2-3 consecutive 1.1:1 (V_{beads}/V_{sample}) AMPure XP bead purification according to standard procedures: 33µL bead slurry was added to 30µL sample DNA, incubated for 5 min at room temperature, magnetized and the supernatant discarded. Beads were washed with 2x 200µL 70% ethanol, briefly air dried and resuspended in 30µL fresh H_2O to elute DNA. This procedure was repeated for a total of 2-3 consecutive purifications and finally purified DNA eluted in 20µL H_2O. DNA concentration was measured with Qubit dsDNA Hs assay and fragment size checked with Bioanalyzer. Libraries were pooled to equimolar amounts for 50bp single-read sequencing on a HiSeq 3000/4000 platform to obtain 70-120 million raw reads per sample (see Supplementary Table 6).

Raw data were processed using a custom pipeline that builds on published workflows[8]. Raw reads were trimmed with cutadapt version 1.9.1 *'-a "TGGAATTCTCGGGTGCCAAGG" --minimum-length=10'* and reverse complemented using fastx_reverse_complement version 0.0.14 (ref. [9]). Reverse complement reads were aligned to a concatenated hg38_dm6_rDNA-U13369 index using bowtie version 2.2.9 with *'--very-sensitive'* setting and filtered with samtools version 1.7 *'view -q 20'* (ref. [5,10]). Reads originating from *Drosophila* S2 spike-in cells were counted with *"samtools view | awk '$3 ~ /dm6_/ {++dm6_count}'"* and normalization factor *alpha=1e6/dm6_count* was calculated as described[11]. Detection of 4.90±0.45% (mean±stdev) reads originating from spike-in genome in DMSO control conditions indicated sufficient run-on competence of *Drosophila S2* cells for robust inference of normalization factors[11]. Alternative normalization based on balancing signal at the end of very long genes led to similar normalization factors (performed as for TT-seq; data not shown) (ref. [7]). Signal originating from human genome scaffolds was now trimmed to a single 3' base, the active center of RNA polymerase during the run-on reaction, using bedtools v2.26.0 *"bamtobed | awk '($6 == "+") {print $1,$3-1,$3,$4,$5,$6}; ($6 == "-") {print $1,$2,$2+1,$4,$5,$6}'"* (ref. [12]). Single-base BAM files were then generated using bedtools v2.26.0 *BedToBam*.

Histone ChIP-seq

ChIP-seq was essentially performed as described with the following modifications[13,14]. 100mio KBM7 wild-type cells were crosslinked with 1.1% PFA (10x PFA mix: 11% formaldehyde, 50 mM HEPES pH 7.3, 100 mM NaCl, 1 mMEDTA pH 8.0, 0.5 mM EGTA pH 8.0) for 10min, washed 3x with cold PBS and snap-frozen. After isolating and washing nuclei, chromatin was sheared using a probe sonicator for a total of 3min (1s on, 4s off) on ice. Lysates were cleared by centrifugation for 10min at 20,000g and incubated with antibody pre-coated protein A/G magnetic beads (H3K27ac: Abcam ab4729, H3K4me3: Millipore 07-473) overnight at 4°C on a rotating wheel. Beads were washed 3x with sonication buffer, 1x with sonication buffer containing 500mM NaCl, 1x with LiCl buffer (20mM Tris pH8.0, 1mM EDTA, 250mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate), and 1x with TE buffer. DNA was eluted with elution buffer (50mM Tris-HCl pH8.0, 10mM EDTA, 1% SDS). Crosslinks were reversed by overnight incubation at 65°C. The next morning, RNA and protein were digested by incubation with 0.2mg/mL RNaseA for 2h at 37°C and 0.2mg/mL proteinaseK for 1h at 55°C. DNA was purified by phenol-chloroform extraction and ethanol precipitated. Libraries for next generation sequencing were prepared from 50ng ChIP-DNA using ThruPLEX DNA-seq kits (Rubicon). Libraries were sequenced in 40bp single-read mode on an Illumina HiSeq 2500 sequencer (see Supplementary Table 6).

Raw reads were trimmed with skewer v0.1.126 with parameters *"-f sanger -t 8 -m any"* and aligned to the hg38 genome assembly using bowtie v2.3.4 with *"--very-sensitive"* (ref. [10,15]). Aligned reads were filtered for duplicates and MAPQ > 30 using sambamba v0.5.5 before calling peaks using macs2 v2.1.0 *callpeak* with *"--fix-bimodal --extsize 180 --bw 200 -c $KBM7_input"* (ref. [16,17]). Publically available ChIP-seq data from HCT-116 cells (GSE72622; see Supplementary Table 7) were processed using the same strategy[18].

Super-enhancer calling and definition of a subset of auto-regulatory transcription factors

Super-enhancers (SEs) were called from TSS-distal KBM7 wild-type H3K27ac signal. High-confidence H3K27ac peaks (p < 1e-9), which did not overlap H3K4me3 peaks (p < 1e-9) over RefSeq TSS +/- 1kb, were used as input for ROSE with *"-c $KBM7_input -s 12500 -t 0"*, yielding 532 SE regions[19,20]. These regions were then assigned to the closest expressed protein coding genes (from SE center to TSS) located in the same topologically associating domain (TAD; from GSM1551624-GSM1551628; see Supplementary Table 7), yielding a set of 464 SE target genes (see Supplementary Table 2) (ref. [21,22]). Expressed genes were defined by counting merged, 100bp extended DMSO 1h and 2h TT-seq signal over "known mRNA" hg38 UCSC refGene annotations (NM_* identifiers). For each gene, a single annotation with highest sense TT-seq RPKM was chosen as the main isoform. Genes with RPKM larger than the median of all main isoforms were considered expressed.

To define subsets of SE-regulated transcription factor (TF) genes, we used a curated list of 1639 known and likely human TFs[23]. Of these, 1272 TFs had at least one directly annotated DNA binding motif in a large collection of 20,003 position weight matrices (PWMs) compiled from various

resources[24]. Intersection with the list of SE target genes yielded 53 SE-regulated TFs (see Supplementary Table 2), which were annotated for a total of 2033 DNA binding motifs.

We further refined the set of 53 SE-regulated TFs to define a subset of 24 auto-regulatory TFs, which strongly regulate their own expression by binding their proximal SE regions. The SE TF constituent DNA sequences were extended by 500bp on each side using bedtools *slop* and extracted using bedtools *getfasta* (ref. [12]). These regions were searched for TF binding motifs using *Cluster-Buster* with parameters *"cbust -c 5 -m 6 -G 0 -f 5"* (ref. [25]). Auto-regulatory TFs were defined as those TFs whose assigned SE constituents contained at least six motifs and at least six motif clusters that can be bound by their own protein products, yielding a set of 24 auto-regulatory TFs (see Supplementary Table 2).

To discern regulatory TF interactions, we applied the same criteria (at least 6 motifs and 6 motif clusters) to predict which of the 24 auto-regulatory TFs could bind SEs assigned to the other auto-regulatory TFs. This identified all possible inter-connected auto-regulatory loops. If different motifs, annotated for the same TF, were detected in a SE, then the TF-motif with the highest number of instances per SE was selected. If there were more than one SE assigned to a TF, then the TF-motif with the highest number of motif instances was selected per each SE and the numbers of instances across all SEs summed. An edge-weight force-directed network was assembled based on these regulatory interactions using cytoscape[26].

For each SE set, equal numbers of expression-matched (TT-seq RPKM median +/- 0.1 standard deviations), non-SE control genes were randomly picked from either all genes or only annotated TFs to investigate if SE-selectivity is explained by high expression alone.

The same strategy was followed to call SE-proximal genes in HCT-116 cells, using publically available ChIP-seq and TAD data (see Supplementary Table 7) (ref. [18,22,27]). The TAD containing the MYC gene on chromosome 8 was manually fused with the TAD immediately upstream, which contains the well-characterized MYC SE in colorectal cancer cells, such as HCT-116 (ref. [28]). This yielded a set of 290 SE-proximal genes, of which 23 were SE-proximal TFs. These TFs were classified auto-regulatory TFs, if their assigned SE constituents contained the at least three motifs and at least three motif clusters that can be bound by their own protein products, yielding 17 auto-regulatory TFs in HCT-116 cells (see Supplementary Table 2).

TT-seq differential expression analysis

TT-seq differentially expressed genes were identified from fused RefSeq annotations using the DESeq2 v1.14.1 package with standard settings using end-of-long-gene normalization factors (see above) (ref. [29]). SE gene sets were identified in transcription unit (TU) annotations by overlaying ENSEMBL gene IDs.

For gene set enrichment analysis, the annotations were expanded to yield only TUs overlapping a single ENSEMBL gene ID. Fold changes for this list of detected ENSEMBL gene IDs were now assembled from the DESeq2 result table. The table was cleaned up by exploding TUs with multiple ENSEMBL gene IDs, retaining the same log2FC values for both underlying genes. Now, ENSEMBL gene IDs were translated to HGNC gene names and the log2FC table grouped by HGNC name, assigning the lowest log2FC for genes occurring multiple times. If different HGNC genes featured the exact same log2FC, the first occurring gene was retained to allow unique sorting of the resulting pre-ranked table. The pre-ranked table was used for gene set enrichment analysis (GSEA) using GenePattern's GSEAPreranked with collections KEGG_pathway, GO_process, GO_function, and custom KBM7 SE gene sets with standard settings[30].

Genome browser snapshots of NGS data

Genome browser shots were generated with bamplot with manually set nBins=1000 and UCSC_refseq_hg38 annotations (https://github.com/linlabbcm/bamplot). For ChIP-seq tracks, plots were generated from filtered BAM files with parameters *'-y uniform -p multiple -s both -r'*. For normalized TT-seq tracks, plots were generated from polymerase signal BAM files (read1 only with *samtools view -b -f 0x40*) with additional flags *'-e 100 -s +/- --scale $SF'*, where *$SF* contains sample-specific normalization factors (see above). For normalized PRO-seq tracks, plots were generated from single-base BAM files with flags *'-y uniform -p multiple -s +/- --scale $alpha -r'*, where *$alpha* contains the exogenous spike-in normalization factors (see above). Thus, signal for unnormalized ChIP-seq tracks is reads per million per basepair (rpm/bp), for spike-in normalized PRO-seq tracks reads per million reference reads per basepair (rrpm/bp), and for end-of-long-gene normalized TT-seq tracks normalized reads per million per basepair (norm_rpm/bp). Biological replicates were merged for this visualization.

HiChIP analysis of chromatin contacts involving H3K27ac-marked regions

H3K27ac HiChIP was performed in bioglogical duplicates essentially as described[31]. 15mio KBM7 MED14-dTAG cells were treated for 2h with DMSO or 500nM dTAG7 and fixed by directly adding methanol-free PFA to 1% (v/v) final concentration. Cells were rocked for 10min at room temperature, quenched with 125mM glycine for 5min. Fixed cells were pelleted and washed 2x with cold PB, transferred to 2mL DNA LoBind tubes, snap-frozen in liquid nitrogen and stored at -80°C until further processing.

In the morning, pellets were resuspended in 500µL ice-cold Hi-C lysis buffer (10mM Tris-HCl pH7.4, 10mM NaCl, 0.2% Triton-X-100, 1x protease inhibitor) and rotated at 4°C for 30min. Nuclei were pelleted for 5min at 2500g, washed 1x with ice-cold Hi-C lysis buffer, and then permeabilized in 100µL 0.2% SDS for 10min at 62°C. 310µL water and 25µL 10% Triton-X-100 were added to quench SDS. Nuclei

were mixed well and incubated for 15min at 37°C. Chromatin was digested by addition of 50µL NEBuffer2 and 15µL MboI and rotation for 2h at 37°C. MboI was heat-inactivated for 20min at 62°C before biotin end-repair by addition of 52µL fill-in mix (37.5µL 0.4mM bio-dATP, 1.5µL each of 10mM dCTP/dGTP/dTTP, 10µL 5U/µL Klenow fragment) and incubation for 1h at 37°C. In situ proximity ligation was performed by addition of 948µL ligation mix (150µL 10x T4 DNA ligase buffer, 125µL 10% Triton-X-100, 3µL 50mg/mL BSA in water, 10µL 400U/µL T4 DNA ligase, 660µL water) and rotation for 4h at room temperature. Nuclei were then pelleted for 5min at 2500g, taken up in 880µL nuclear lysis buffer (50mM Tris-HCl pH7.4, 10mM EDTA, 1% SDS, 1x protease inhibitor) and transferred to 1mL AFA milliTubes (Covaris). Ligated chromatin was sheared for 5min using a Covaris S220 (4°C, fill level 10, duty cycle 5, PIP 140, cycles per burst 200). Lysates were cleared for 15min at 16,100g and split into two 2mL tubes before addition of 750µL ChIP dilution buffer (0.01% SDS, 1.1% Triton-X-100, 1.2mM EDTA, 16.7mM Tris-HCl pH7.4, 167mM NaCl) to each of the ~400µL lysate. Each lysate half was pre-cleared by rotation for 1h at 4°C with 30µL washed (3x 500µL ChIP dilution buffer) protein A dynabeads. Supernatants were transferred to new 2mL tubes and 4.5µL H3K27ac antibody (Abcam ab4729) was added to each of the two tubes for immunoprecipitation by overnight rotation at 4°C.

In the morning, another 30µL protein A dynabeads per half were washed and added to lysates to capture antibodies by rotation for 2h at 4°C. Beads were washed in 3x500µL of each low salt (0.1% SDS, 1% Triton-X-100, 2mM EDTA, 20mM Tris-HCl pH7.4, 150mM NaCl), high salt (0.1% SDS, 1% Triton-X-100, 2mM EDTA, 20mM Tris-HCl pH7.4, 500mM NaCl), and LiCl wash buffer (10mM Tris-HCl pH7.4, 250mM LiCl, 1% Triton-X-100, 1% Na-deoxycholate, 1mM EDTA), pooling the split samples back during the first washing step. DNA was eluted by 100µL fresh elution buffer (50mM NaHCO3, 1% SDS), rotation for 10min at room temperature and 3min shaking at 37°C. Eluates were transferred to fresh 1.5mL tubes and DNA eluted from beads a second time with fresh 100µL elution buffer to yield a total volume of 200µL eluate per sample. Protein was digested by addition of 5µL 20mg/mL proteinase K and shaking for 45min at 55°C before raising the temperature to 67°C and incubating 1.5h more. Ligated ChIP DNA was purified using Zymo ChIP DNA purification columns and elution in 11µL water. DNA was quantified using Qubit dsDNA Hs kits and stored at -20°C.

The next day, 5µL Streptavidin C-1 beads per sample were washed in 3x Tween wash buffer (5mM Tris-HCl pH7.4, 0.5mM EDTA, 1M NaCl, 0.05% Tween-20), resuspended in 10µL 2x biotin binding buffer (10mM Tris-HCl pH7.4, 1mM EDTA, 2M NaCl) and added to the 10µL ligated ChIP DNA. After 15min rotation at room temperature, beads were washed 2x with 500µL Tween wash buffer for 2min at 55°C, 3x with 100µL 1x (from 2x) fresh TD buffer (2x: 20mM Tris-HCl pH7.4, 10mM $MgCl_2$, 20% dimethylformamide) and then resuspended in 25µL 2x TD buffer. DNA was tagmented by addition of 2.5µL Tn5 (Nextera, Illumina) per 50ng post-ChIP DNA, filling up the reaction to 50µL with water, and incubation for 10min at 55°C with interval shaking. Beads were magnetized, discarding the

supernatant, and incubated 30min at 50°C in 100µL of 50mM EDTA. Beads were washed 2x with 100µL 50mM EDTA for 3min at 50°C, 2x with 100µL Tween wash buffer for 2min at 55°C, and 1x with 200µL 10mM Tris-HCl pH7.4. Beads were stored at -20°C until further processing.

Tagmented DNA was PCR amplified by resuspending in PCR master mix (1µL of 12.5µM universal Nextera Ad1_noMX primer, 25µL Phusion 2x MM, 23µL water, +1µL of 12.5µM sample-specific barcoded Ad2.x_N70x primer; see Supplementary Table 8 for primer sequences) and cycling for 7 cycles (72°C 5min; 98°C 1min; cycle 98°C 15sec; cycle 63°C 30sec; cycle 72°C 1min; final 72°C 5min; store 8°C infinite). The 50µL PCR reactions were magnetized and supernatants transferred to fresh 1.5mL tubes. Right side size-selection was performed by adding 25µL AMPure XP beads and transferring the supernatant to fresh 1.5mL tubes. For left side size-selection, 15µL fresh AMPure XP beads were added and the supernatant discarded. Final libraries were eluted in 12µL water and Qubit-quantified. Size distribution was quality-controlled using Bioanalyzer, pooled to equimolar amounts and sequenced as 4-plex pools on two Illumina HiSeq 3000/4000 lanes in 75bp paired-end mode (see Supplementary Table 6).

For analysis, raw reads were in-silico digested using the MboI restriction enzyme re-ligation sequence GATCGATC. Only the 5-prime half of cut reads was kept after in-silico cleavage. All non-cut reads were kept. Filtered reads were then mapped to the hg19 human genome assembly with bowtie2 v2.2.3 using options *--very-sensitive -L 30 --score-min L,-0.6,-0.2 --end-to-end --reorder* (ref. [10]). The human genome was in-silico digested with MboI enzyme and mapped reads assigned to the digested genome fragments. Read pairs were called valid if they had the correct orientation and mapped to different fragments. Biological duplicates were fused at this point and .hic contact matrices generated with juicer tools *'pre'* v1.8.9 (ref. [32]). The hichipper tool v0.7.3 was run on fused replicates to identify anchor sites and remove non-significant interactions (anchor-anchor sites supported by PETs) with a mango FDR < 0.01 (ref. [33]). KBM7 H3K27ac ChIP-seq peaks (see above) were used as peak input for hichipper. Filtered interactions for each fused sample were loaded into R and processed with the diffloop tool v1.12.0 (ref. [34]). Interactions less than 5kb apart and not present in all four samples were discarded. Differential loops were determined using the diffloop *'quickAssocVoom'* function, yielding interaction log2 fold-changes. Anchor sites were annotated by intersecting their coordinates with enhancers (H3K27ac; see above), promoters (H3K4me3; see above), super-enhancers (H3K27ac; see above), and CTCF sites (from HAP1; GSM2493878; see Supplementary Table 7) using bedtools intersect v2.27.1 (ref. [12]).

Acid-based preparation of Histone extracts

Histone extracts were prepared from $2x10^6$ KBM7 MED14-dTAG cells treated with DMSO or 500nM dTAG7 for up to 24h. Cells were harvested by centrifugation and washed with ice-cold PBS before resuspending in 500µL 0.5% Triton-X-100 in PBS and incubating on ice for 10min with occasional

vortexing. Nuclei were harvested by centrifugation at 1000g for 10min at 4°C and washed once in 250µL 0.5% Triton-X-100 in PBS. Nuclei were then resuspended in 50µL 0.2M HCl and shaken at 4°C overnight. Lysates were cleared by centrifugation at 20,000g for 10min at 4°C before determining protein content via BCA assay. The resulting histone extracts were complemented with 4x LDS sample buffer and 0.5µg of total protein subjected to western blot analysis.

Circularized chromosome conformation capture and sequencing (4C-seq)

In situ 4C-Seq was performed in biological triplicates as previously described[35]. Approximately 5mio KBM7 MED14-dTAG cells were fixed with 1% methanol-free PFA directly in medium. After crosslinking for 10 minutes at room temperature, PFA was quenched with 125mM glycine for 5min. Cells were harvested at 500g for 5min, washed 2x with cold PBS, snap frozen, and stored at 80°C until further processing.

Pellets were gently resuspended in 500µl Hi-C lysis buffer (10mM Tris-HCl pH8.0, 10mM NaCl, 0.2% Igepal) with 1x cOmplete protease inhibitors (Roche 11697498001). Cells were incubated on ice for 30min and washed 1x with 500µl of ice-cold Hi-C lysis buffer with protease inhibitors. Pellets were resuspended in 50µl of 0.5%SDS and incubated at 62°C for 7min. SDS was quenched by adding 145µl water and 25µl 10% Triton X-100 and incubation at 37°C for 15min. Chromatin was digested by adding 25µl of 10x Cutsmart buffer (New England Biolabs) and 200 units of NlaIII-HF (New England Biolabs) enzyme and incubation at 37°C in a thermomixer at 500rpm overnight.

The next morning, 200 more units of enzyme were added and the reaction was incubated another 4h at 37°C degrees in a thermomixer at 500rpm. Restriction enzyme was heat-inactivated at 62°C for 20min while shaking at 500rpm. Proximity ligation was performed in a total of 1200 µL with 2000 units of T4 DNA ligase (NEB M2020) for 6h at room temperature. After ligation, samples were spun down for 5min at 2500g and resuspended in 300µl 10mM Tris-HCl, 1% SDS and 0.5mM NaCl with 1000 units of Proteinase K. Crosslinks were reversed by overnight incubation at 65°C.

Samples were then phenol-chloroform extracted and ethanol precipitated and the second digestion was performed overnight in 450µL with 50 units of restriction enzyme. CviQI (NEB R0639S) was used for MYB_SE_VP1, and DpnII (NEB R0543) was used for SATB1_SE_VP1. Samples were phenol-chloroform extracted and ethanol-precipitated, and the second ligation was performed in 14ml total volume with 6700 units of T4 DNA ligase (NEB M2020) at 16°C overnight. Samples were ethanol precipitated, resuspended in 500µl QIAGEN EB buffer, and purified with a QIAGEN PCR purification kit.

PCR amplification and indexing of samples were performed with two successive PCR reactions as previously described[36]. The first PCR amplification was performed with 4x 50µL PCR reactions using Roche Expand Long Template polymerase (Roche 11759060001). Reaction conditions were as follows: 5µL of 10 X Roche Buffer, 0.8µL of 10mM dNTPs (NEB), 10µl of the forward primer-reverse primer pair (5µM each; see Supplementary Table 8), 0.75µL Roche Expand Long Template Polymerase, adjusted to

50µl total volume with water. Cycling conditions were: 2min 94°C; cycle 15sec 94°C, 1min 55°C, 3min 68°C, repeating steps 2-4. After 30 total cycles, the reaction was held for 7min at 68°C and then stored at 4°C. Libraries were cleaned up using four columns of a QUIAGEN PCR purification kit per library. The DNA libraries were then barcoded in a second PCR reaction, which was performed as follows: 20µL of Phusion Polymerase Master mix (2x), 1µL of the indexing primer (10µM), 1µL of the adapter primer (10µM; see Supplementary Table 8) adjusted to 40µL total volume with water. Cycling conditions were: 30sec 98°C, 10sec 98°C, 30sec 60°C, 120sec 72°C, repeat steps 2-4. After 8 total cycles, the reaction was held for 5min at 72°C and then stored at 4°C. DNA was cleaned-up using using 1:1 (V:V) AMPure XP beads. Finally, libraries were Qubit quantified, pooled to equimolar amounts and sequenced on the Illumina NextSeq 500 platform in 75bp single-read mode (see Supplementary Table 6).

For analysis, raw reads were kept if the nucleotides 1-16 had a hamming distance of at most 2 with the read primer nucleotides 1 to 16 and if the restriction enzyme cut site (nt 17-20) also had a hamming distance of at most 2. The read primer was then removed from the raw reads and output. Duplicate reads were collapsed and the restriction enzyme recognition site was added back to the beginning of a collapsed read. The collapsed reads were then aligned to the hg19 human genome assembly with bowtie v1.2.3 and parameters *-n 2 -e 70 -m 1 -k 1 --best -l 200*, yielding only uniquely mapped reads with at most 2 mismatches[37]. The genome was in-silico digested with the two restriction enzymes of the respective 4C-seq experiment using the Hicup digester v0.6.1 (ref. [38]). Reads mapping to "blind" fragments were removed from further analysis[39], and only reads mapping to "nonblind" or "halfblind" fragments were kept: Nonblind fragments had both restriction enzyme cleavage sites present, one at each end. Halfblind fragments had the first restriction enzyme cut site present at both ends. Reads mapping within 1kb of the viewpoint were omitted. Bamtools v2.5.1 *merge* was used to merge all nonblind and halfblind files from the three biological replicates[40]. BigWig files were generated using bamCoverage v3.2.1 with options *--normalizeUsing CPM --binSize 500 --smoothLength 6500*.

PRO-seq enrichment heatmaps of enhancer regions

Active enhancer regions were identified by intersecting H3K27ac ChIP-seq and PRO-seq data. Cis regulatory elements were first classified using dREG (https://dreg.dnasequence.org/) on unnormalized, replicate-merged DMSO 1h PRO-seq signal and filtered for peaks scores > 0.5, yielding 37,790 regions[41]. To identify super-enhancer (SE) constituents, dREG peak centers were intersected with stitched SE regions, yielding 2,760 SE constituents. To identify other ("typical") enhancer regions, we intersected non-H3K4me3 (i.e. non-TSS) H3K27ac ChIP-seq peaks (input for SE calling; see above) with non-SE dREG peaks, yielding 6160 "typical" enhancers.

We used deeptools 3.3.0 to plot PRO-seq enrichment at these (super-) enhancer constituent regions[42]. Briefly, we visualized spike-in normalized PRO-seq signal +/- 1kb from peak centers for each

strand separately with *"computeMatrix reference-point -a 1000 -b 1000 --missingDataAsZero --skipZeros*" and *"plotHeatmap --sortUsing max --sortUsingSamples $DMSO_plus_signal --colorList '#050596,#ffffff00,#960505' --alpha 1 --zMax 0.2 --zMin '-0.2'"*. Signal from the respective strands was then graphically overlaid with missing data color being set to full transparency (#ffffff00 in the symmetric color scale).

PRO-/TT-seq metagene aggregate signal and enrichment heatmaps

We used deeptools 3.3.0 to plot PRO-seq metagene enrichment and aggregate transcription start site (TSS) coverages from replicate-merged signal files[42]. Genes were grouped into SE-regulated (464 genes), a subset of SE-auto-regulatory TFs (24 genes), and non-SE genes (9,276 genes; see above). To visualize strand-specific metagene signals, we first separately computed matrices for the genome's plus and minus strands using *"computeMatrix scale-regions -S $strand_specific_signal_bw -R $gene_regions_BED -b 2000 -m 10000 -a 5000 --unscaled5prime 500 --unscaled3prime 500 --missingDataAsZero --skipZeros -bs 25"*. The two resulting matrices were then filtered to only retain signal in the same sense as the respective gene annotation using *computeMatrixOperations filterStrand* (i.e. genomic plus with *"--strand +"* and genomic minus with "*--strand –"*). These matrices were then combined using computeMatrixOperations rbind, to yield a single sense matrix. Metagene profiles were now plotted using *"plotProfile --plotType fill --perGroup"*.

The analogous procedure was performed on TSS regions, but computing the first matrices with *"computeMatrix reference-point -a 500 -b 500 --missingDataAsZero --skipZeros"*. Due to the low number of underlying genes, meta-signal for 24 auto-regulatory TFs was smoothed by running means of +/- 3 bins before plotting.

To visualize spike-in normalized strand-specific fold changes over metagenes, we used single-base BAM files as input for "*bamCompare --scaleFactors $alpha --operation log2 --pseudocount 0.01 -bs 1 --samFlagExclude 16*" (or *--samFlagInclude 16* for genomic minus strand) and then followed the same strategy as above to compute strand-specific matrices and retain only same-sense signal.

To visualize read-through transcription, we defined a window from the last exon of RefSeq TUs (see above) to 25kb past the last annotated polyadenylation site. Signal was collected for all expressed genes and normalized in the visible region to a maximum of 1.0 and minimum of 0.

PRO-seq pausing and read-through indices

Pausing and read-through indices were essentially calculated as previously described with following modifications[13,43]. First, only expressed main isoforms (see above) of at least 2kb were considered. From these, sub-genic windows were defined as the following: TSS-windows are *TSS-50bp to TSS+250bp*; genebodies are *TSS+500bp to gene end*; pre-polyA-windows are *polyA-1000bp to gene end*; termination-windows are *polyA+3000bp to polyA+6000bp*.

Strand-specific, replicate-merged, spike-in normalized coverages over these different sub-genic windows were computed from single-base PRO-seq signal using bedtools *"coverageBed -s | awk '{$7 = $7 * $alpha}'"* and normalized for annotation length to yield RPKM. To calculate pausing indices, genes with the lowest 10% read-coverage in either TSS-window or genebody were excluded. Pausing indices represent the ratio of TSS-window over genebody RPKM. In analogy, read-through indices represent ratios of filtered termination-window over pre-polyA-window RPKM.

TT-/PRO-seq-based modeling of pause duration and productive initiation rates

Modeling of Pol II kinetic parameters was performed as described, with minor adaptations[6,44]. Briefly, we used PRO-seq instead of mNET-seq data to infer genome-wide polymerase density distributions. For information on polymerase activity, we used TT-seq data to infer the number of (newly produced) RNA molecules per cell. For each transcription unit (TU), the productive initiation frequency I_i [cell^{-1}min^{-1}] was inferred using the antisense bias-corrected number of transcribed bases tb_i, normalizing for labeling duration and TU length L_i, and calibrated to literature estimates with size factor κ (ref. [6,44]).

$$I_i = \kappa \cdot \frac{tb_i}{t \cdot L_i}$$

Note that tb_i and L_i were restricted to the first 10 kbp of the transcription unit in order to reveal initiation frequency changes given the short treatment times. Note also that this parameter does not directly reflect how many polymerase molecules initiate transcription at the TSS, but rather informs on those initiation events, which successfully overcome pausing and processively elongate a transcript in the genebody. If the fraction of prematurely terminating polymerase is low, this also corresponds to the number of polymerases loaded at the promoter per unit time.

To calculate pause durations, information on polymerase location was obtained from spike-in normalized single-base PRO-seq signal. First, pause sites (i.e. the single base where Pol II pauses) were defined from RefSeq TUs with a single annotated TSS in a window from TSS to the end of the first exon (excluding the last five bases). The most likely pause site ρ_i was inferred as the base with the highest PRO-seq signal in this window via maximizing the function below, which was additionally required to exceed five times that window's median signal strength (calculated on all non-negative antisense bias corrected PRO-seq coverage values p_{im}).

$$\rho_i = \max_m p_{im}$$

Pause durations d_i [min] were then calculated as the time that polymerases reside in a +/-100bp window around the pause site ρ_i by integrating information from the TT-seq-informed productive initiation I_i rates, and calibrated to literature estimates with size factor σ (ref. [6,44]).

$$d_i = \sigma \cdot \frac{\sum_{+/-\ 100} p_{im}}{I_i}$$

For pause sites within the first 100bp of a TU, the pause window was defined as the first 200bp of that TU. Absolute values for pause durations were obtained by integrating a previously measured polymerase elongation velocity of 2.4kb/min (ref. [6]).

Chromatin fractionation and mass spectrometry

Subcellular fractionation was performed in biological triplicates and chromatin-bound proteins subjected to unbiased mass spectrometry analysis, similar to previously described procedures[45,46]. Briefly, 40x10^6 washed and snap-frozen MED14-dTAG KBM7 cells were thawed on ice and incubated with 350µL hypotonic buffer SB (10mM HEPES pH7.5, 10mM KCl, 1.5mM $MgCl_2$, 340mM sucrose, 0.1% Triton-X-100, 1x Halt protease inhibitor cocktail) for 30 min at 4°C on a rotating wheel. Nuclei were harvested by 3 min centrifugation at 4°C and 2000g, keeping the supernatant as cytoplasmic fraction. Pellets were taken up in 500µL buffer SB, layered onto 2.5mL sucrose buffer SC (10mM HEPES pH7.5, 10mM KCl, 1.5mL $MgCl_2$, 2.1M sucrose, 1x Halt protease inhibitor cocktail) and ultracentrifuged for 3h at 4°C and 150,000g. Supernatant was kept as the nuclear fraction and pellets taken up in 500µL buffer SB. Pellets were collected by 45 min centrifugation at 4°C and 20,000g and washed twice with 500µL wash buffer (10mM HEPES pH7.5, 1x Halt protease inhibitor cocktail). The chromatin pellets were transferred to 1mL Covaris AFA Fiber milliTUBEs with 200µL Benzonase digest buffer (15mM HEPES pH7.5, 1mM EDTA, 1mM EGTA, 0.1% NP-40, 1x Halt protease inhibitor cocktail) and sonicated for 120s at 4°C in a Covaris S220 focused ultrasonicator with following settings: peak power 140, duty factor 10, cycles/burst 200. Samples were transferred to fresh 1.5mL microcentrifuge tubes and chromatin was digested by addition of 2.5U Benzonase (Merck; 25U/µL stock) and 20µg RNase A (Thermo Fisher; 10mg/mL stock) and 30 min incubation on a rotating wheel at 4°C. Samples were incubated for 10 min at room temperature after addition of equal volume 2x SDS lysis buffer (100mM HEPES pH7.5, 4% SDS, 2mM PMSF, 1x Halt protease inhibitor cocktail) and then heated to 99°C for 5 min. Lysates were cleared by 10 min centrifugation at 16,000g and 20°C and transferred to fresh microcentrifuge tubes. Protein concentration was determined by BCA (Pierce Biotechnology, Rockford, IL).

Filter aided sample preparation (FASP) of 200µg total protein was performed using a 30 kDa molecular weight cutoff filter (VIVACON 500; Sartorius Stedim Biotech GmbH, 37070 Goettingen, Germany), essentially following described procedures[47]. 50µL sample were directly mixed in the filter unit with 200 µL of freshly prepared 8M urea in 100mM Tris-HCl pH8.5 buffer (UA buffer) and centrifuged at 14,000g for 15min at 20°C to remove SDS. Residual SDS was washed out by two washing steps with 200µL UA buffer. Proteins were alkylated by incubation with 100µL of 50mM iodoacetamide in the dark for 30min at RT. After washing 3x with 100µL UA buffer and 3x with 100µL of 50mM TEAB

buffer pH 8 (SIGMA-Aldrich Chemie GmbH, Germany), proteins were digested with 4µg trypsin overnight at 37 °C. Peptides were recovered from the filter by centrifugation, applying 40µL of 50mM TEAB buffer followed by 50µL of 0.5M NaCl. Eluted peptides were acidified with TFA, desalted using C18 solid phase extraction spin columns (The Nest Group, Southborough, MA), organic solvent removed in a vacuum concentrator at 45°C and reconstituted in 5% formic acid for analysis by LC-MS/MS.

Mass spectrometry was performed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific, San Jose, CA) coupled to a Dionex U3000 RSLC nano system (Thermo Fisher Scientific, San Jose, CA) via nanoflex source interface. Approximately 2µg tryptic peptides were loaded onto a trap column (Acclaim™ PepMap™ 100 C18, 3µm, 5 × 0.3 mm, Fisher Scientific, San Jose, CA) at a flow rate of 10µL/min using 2% acetonitrile in 0.1% TFA as loading buffer. After loading, the trap column was switched in-line with a 40cm, 75µm inner diameter analytical column (packed in-house with ReproSil-Pur 120 C18-AQ, 3 µm, Dr. Maisch, Ammerbuch-Entringen, Germany). Mobile-phase A consisted of 0.4% formic acid in water and mobile-phase B of 0.4% formic acid in a mix of 90% acetonitrile and 10% water. The flow rate was set to 230nL/min and a ~206min gradient used (4 to 24% solvent B within 191min, 24 to 36% solvent B within 10min and, 36 to 100% solvent B within 5min, 100% solvent B for 10min before re-equilibrating at 4% solvent B for 18min). Analysis was performed in a data-dependent acquisition mode using the integrated universal method. MS^1 spectra were acquired with a mass range of 375-1650m/z in the orbitrap at a resolution of 120,000 (at 200Th). Automatic gain control (AGC) was set to a target of 2×10^5 and a maximum injection time of 50msec. MS^2 scans were acquired in the linear ion trap (IT) at rapid scan rates using collision-induced dissociation (CID) and a normalized collision energy of 35%. Peptide monoisotopic precursor selection (MIPS) was enabled and an intensity threshold for precursor selection set to 1×10^4 applying a quadrupole isolation window of 1.6 Da and a maximum injection time of 250msec. Inject ions for all available parallelizable time was enabled. Charge states of 2-6 were included for MS^2 selection. Dynamic exclusion for selected ions was set to 90sec. A single lock mass at m/z 445.120024 was employed[48]. XCalibur version 4.0.0 and Tune 2.1.1565.24 were used to operate the instrument.

Raw files were processed using MaxQuant (version 1.6.2.3) using default settings unless noted otherwise[49]. Spectra were searched against the UniProtKB/Swiss-Prot human database (release 2018_07 containing 20386 sequences) concatenated with MaxQuant's contaminant database. Methionine oxidation and N-terminal acetylation were set as variable modifications and cysteine carbamidomethylation as fixed modification. In silico cleavage was performed with trypsin/P, allowing 2 missed cleavages. Match between runs was additionally enabled. Peptide-to-spectrum matches (PSMs) were filtered at 1 % FDR and label free quantification (LFQ) calculations were performed.

For further processing, both peptide- and protein-based quantification methods were used and combined. Protein summaries reported in MaxQuant's proteinGroups.txt file were processed using Perseus[49], where firstly reverse hits and those only identified by a single site were excluded from further analysis. LFQ and intensity-based absolute quantification (iBAQ) intensities were log2-transformed and only proteins with reported intensities in all biological replicates of at least one treatment group were kept for statistical analysis. Missing values were replaced by random numbers drawn form a normal distribution of 1.8 standard deviation down shift and a width of 0.3 of each sample. Significance of differential enrichment was calculated by comparing LFQ and iBAQ protein intensities of untreated and dTAG-treated samples in a two-sided t-test at 5 % FDR corrected via the Benjamini-Hochberg FDR procedure.

As a peptide-based quantification method, we used the MSqRob software with standard settings (version 0.7.5) (ref. [50]), where the 36,424 peptides reported in MaxQuant's peptide.txt file were used as input. Intensities were log2 transformed and quantile normalized. Contaminants and reverse identifications were removed as well as protein groups that contained one or more proteins present also in a smaller protein group. Protein quantification was performed on 18760 remaining proteins, where dTAG treatment was specified as a fixed effect and biological replicates as well as peptide sequences were set as random sample effects.

Significantly differentially chromatin-bound proteins ($p < 0.1$) from all three analysis methods were intersected to define a set of analysis-independent core hits (see Supplementary Table 3). The resulting proteins were mapped to the StringDB[51] protein-protein interaction database, only retaining the largest connected component with “experimental evidence” > 0.4 and “confidence” >= 1.

Salt-based fractionation of 7SK/P-TEFb complexes

Differential salt extraction to separate 7SK RNPs was performed similar to published procedures[52]. $10x10^6$ washed MED14-dTAG KBM7 cells were resuspended in 100µL cold Buffer A (10mM Tris-HCl pH7.4, 10mM KCl, 10mM $MgCl_2$, 1mM EDTA, 0.5% Triton-X-100, 1mM DTT, 1x Halt protease inhibitor cocktail) and incubated 20 min on ice with occasional vortexing. Nuclei were harvested by 5 min centrifugation at 4°C and 1000g and 75µL supernatant kept as cytosolic extract, while aspirating the rest. Nuclei were washed twice with 500µL Buffer A and then resuspended in 100µL Buffer C (20mM Tris-HCl pH7.4, 450mM NaCl, 1.5mM $MgCl_2$, 0.5mM EDTA, 1mM DTT, 1x Halt protease inhibitor cocktail). After 20 min incubation on ice, lysates were clarified by 10 min centrifugation at 4°C and 20000g. 75µL supernatant were kept as nuclear extract. After adding 4x LDS sample buffer, 20µL cytosolic and 8µL nuclear extract were loaded on 4-12% Bis-Tris gels for western blot analysis.

HEXIM1-based immunoprecipitation of 7SK/P-TEFb complexes

For HEXIM1 immunoprecipitation, 20µL Pierce Protein A/G bead slurry per condition was washed with 3x 400µL IP lysis buffer (50mM Tris-HCl pH7.4, 150mM NaCl, 0.1% Triton-X-100, 1mM EDTA, 5mM $MgCl_2$, 5% glycerol, 1x Halt protease inhibitor cocktail). Beads were then resuspended in 200µL IP lysis buffer with 4µg anti-HEXIM1 antibody per condition (Abcam ab25388) and incubated for 1h at 4°C on a rotating wheel. Meanwhile, $20x10^6$ washed MED14-dTAG KBM7 cells were resuspended in 250µL IP lysis buffer and incubated 20 min on ice. Lysates were clarified by 20 min centrifugation at 4°C and 20000g. Pre-incubated beads were washed with 3x 400µL IP lysis buffer supplemented with 0.5% BSA. Aliquots amounting to 20µL original bead slurry were then resuspended in 200µL cell lysate and incubated for 1.5h at 4°C on a rotating wheel. The remaining lysate was kept as input. After the incubation, 50µL flow-through was kept for western blot analysis. The beads were washed with 3x 400µL IP lysis buffer, ultimately taken up in 50µL 4x LDS sample buffer, heated to 95°C for 10 min and diluted to 1x with IP lysis buffer. 5µL (input and flow-through) or 25µL (IP: HEXIM1) were loaded on 4-12% Bis-Tris gels for western blot analysis.

Chromatin purification for immunoblotting of CDK9 targets

Purification of chromatin fractions was performed as previously described[53]. 15mio KBM7 MED14-dTAG cells were treated with 500nM dTAG7 for 1-2h in T25 flasks, harvested and washed 1x with cold PBS. Cells were resuspended in 150µL ice-cold cytoplasmic lysis buffer (10mM Tris-HCl pH7.0, 150mM NaCl, 0.15% NP-40, 1x protease inhibitor) and incubated on ice for 5min. The lysate was carefully layered onto 400µL cold sucrose buffer (10mM Tris-HCl pH7.0, 150mM NaCl, 25% sucrose, 1x protease inhibitor) in fresh 1.5mL tubes using cut pipette tips and centrifuged for 10min at 18,000g and 4°C (supernatant is cytoplasmic fraction). The pellet was washed with 500µL cold nuclei wash buffer (1mM EDTA, 0.1% Triton-X-100, 1x PBS, 1x protease inhibitor), spun for 1min at 3500rpm and supernatant discarded. Nuclei were gently resuspended in 200µL cold glycerol buffer (20mM Tris-HCl pH8.0, 75mM NaCl, 0.5mM EDTA, 50% glycerol, 0.85mM DTT, 1x protease inhibitor) using cut pipette tips before adding 200µL cold nuclei lysis buffer (20mM HEPES pH7.5, 300mM NaCl, 1% NP-40, 0.2mM EDTA, 1M urea, 1mM DTT, 1x protease inhibitor), mixing by pulse vortexing and incubation for 2min on ice. Supernatant was harvested as the nucleoplasmic fraction by centrifugation for 2min at 14,000rpm and 4°C. The chromatin pellet was then resuspended in 100µL cold RIPA buffer (50mM Tris-HCl pH8.0, 150mM NaCl, 1% Triton-X-100, 0.5% Na-deoxycholate, 0.1% SDS, 1x Halt protease inhibitor cocktail, 25U/mL Benzonase) and solubilized by sonication for 3x 1s at 25% intensity using a probe sonicator. 4x LDS sample buffer was added and chromatin samples heated to 95°C for 10min before loading equal volumes onto 4-12% BisTris gels for western blot analysis. To ensure accurate quantification of phosphorylation status, phospho- and total antibodies were combined in the same milk.

Supplementary Note References:

1. Schwalb, B. *et al.* TT-seq maps the human transient transcriptome. *Science* **352**, 1225-8 (2016).
2. Dolken, L. *et al.* High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *Rna* **14**, 1959-72 (2008).
3. Villamil, G., Wachutka, L., Cramer, P., Gagneur, J. & Schwalb, B. Transient transcriptome sequencing: computational pipeline to quantify genome-wide RNA kinetic parameters and transcriptional enhancer activity. *bioRxiv*, 659912 (2019).
4. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
5. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-9 (2009).
6. Gressel, S. *et al.* CDK9-dependent RNA polymerase II pausing controls transcription initiation. *Elife* **6**(2017).
7. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106 (2010).
8. Mahat, D.B. *et al.* Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). *Nat Protoc* **11**, 1455-76 (2016).
9. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3 (2011).
10. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-359 (2012).
11. Orlando, D.A. *et al.* Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep* **9**, 1163-70 (2014).
12. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
13. Rahl, P.B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432-45 (2010).
14. Winter, G.E. *et al.* BET Bromodomain Proteins Function as Master Transcription Elongation Factors Independent of CDK9 Recruitment. *Mol Cell* **67**, 5-18 e19 (2017).
15. Jiang, H., Lei, R., Ding, S.-W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
16. Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032-2034 (2015).
17. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
18. McNamara, R.P. *et al.* KAP1 Recruitment of the 7SK snRNP Complex to Promoters Enables Transcription Elongation by RNA Polymerase II. *Mol Cell* **61**, 39-53 (2016).
19. Whyte, W.A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-19 (2013).
20. Loven, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320-34 (2013).
21. Rao, S.S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-80 (2014).
22. Wang, Y. *et al.* The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *Genome Biology* **19**, 151 (2018).
23. Lambert, S.A. *et al.* The Human Transcription Factors. *Cell* **172**, 650-665 (2018).
24. Imrichová, H., Hulselmans, G., Kalender Atak, Z., Potier, D. & Aerts, S. i-cisTarget 2015 update: generalized cis-regulatory enrichment analysis in human, mouse and fly. *Nucleic Acids Research* **43**, W57-W64 (2015).
25. Frith, M.C., Li, M.C. & Weng, Z. Cluster-Buster: finding dense clusters of motifs in DNA sequences. *Nucleic Acids Research* **31**, 3666-3668 (2003).
26. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-504 (2003).
27. Rao, S.S.P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320 e24 (2017).
28. Schuijers, J. *et al.* Transcriptional Dysregulation of MYC Reveals Common Enhancer-Docking Mechanism. *Cell Rep* **23**, 349-360 (2018).
29. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
30. Reich, M. *et al.* GenePattern 2.0. *Nat Genet* **38**, 500-1 (2006).
31. Mumbach, M.R. *et al.* HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* **13**, 919-922 (2016).

32. Durand, N.C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell systems* **3**, 95-98 (2016).
33. Lareau, C.A. & Aryee, M.J. hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. *Nature Methods* **15**, 155 (2018).
34. Lareau, C.A. & Aryee, M.J. diffloop: a computational framework for identifying and analyzing differential DNA loops from sequencing data. *Bioinformatics (Oxford, England)* **34**, 672-674 (2018).
35. Weintraub, A.S. *et al.* YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **171**, 1573-1588.e28 (2017).
36. Franke, M. *et al.* Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265-269 (2016).
37. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25 (2009).
38. Wingett, S. *et al.* HiCUP: pipeline for mapping and processing Hi-C data. *F1000Research* **4**, 1310-1310 (2015).
39. van de Werken, H.J.G. *et al.* Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nature Methods* **9**, 969-972 (2012).
40. Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P. & Marth, G.T. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**, 1691-1692 (2011).
41. Wang, Z., Chu, T., Choate, L.A. & Danko, C.G. Identification of regulatory elements from nascent transcription using dREG. *Genome Res* **29**, 293-303 (2019).
42. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic acids research* **42**, W187-W191 (2014).
43. Nojima, T. *et al.* Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell* **161**, 526-540 (2015).
44. Gressel, S., Schwalb, B. & Cramer, P. The pause-initiation limit restricts transcription activation in human cells. *Nature Communications* **10**, 3603 (2019).
45. Dutta, B. *et al.* Elucidating the temporal dynamics of chromatin-associated protein release upon DNA digestion by quantitative proteomic approach. *J Proteomics* **75**, 5493-506 (2012).
46. Sdelci, S. *et al.* MTHFD1 interaction with BRD4 links folate metabolism to transcriptional regulation. *Nat Genet* **51**, 990-998 (2019).
47. Wisniewski, J.R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat Methods* **6**, 359-62 (2009).
48. Olsen, J.V. *et al.* Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics* **4**, 2010-21 (2005).
49. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**, 2301 (2016).
50. Goeminne, L.J.E., Gevaert, K. & Clement, L. Experimental design and data-analysis in label-free quantitative LC/MS proteomics: A tutorial with MSqRob. *Journal of Proteomics* **171**, 23-36 (2018).
51. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-d613 (2019).
52. Biglione, S. *et al.* Inhibition of HIV-1 replication by P-TEFb inhibitors DRB, seliciclib and flavopiridol correlates with release of free P-TEFb from the large, inactive form of the complex. *Retrovirology* **4**, 47 (2007).
53. Mayer, A. *et al.* Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **161**, 541-554 (2015).