SUPT3H-less SAGA coactivator can assemble and function without significantly perturbing RNA polymerase II transcription in mammalian cells

Veronique Fischer, Vincent Hisler, Elisabeth Scheer, Elisabeth Lata, Bastien Morlet, Damien Plassard, Dominique Helmlinger, Didier Devys, László Tora, Stéphane D Vincent

Nucleic Acids Research, 2022 **1**
https://doi.org/10.1093/nar/gkac637

SUPT3H-less SAGA coactivator can assemble and function without significantly perturbing RNA polymerase II transcription in mammalian cells

Veronique Fischer[1,2,3,4,†], Vincent Hisler[1,2,3,4,†], Elisabeth Scheer[1,2,3,4,†], Elisabeth Lata[1,2,3,4], Bastien Morlet[1,2,3,4], Damien Plassard[1,2,3,4,5], Dominique Helmlinger[6], Didier Devys[1,2,3,4], László Tora[1,2,3,4,*] and Stéphane D. Vincent[1,2,3,4,*]

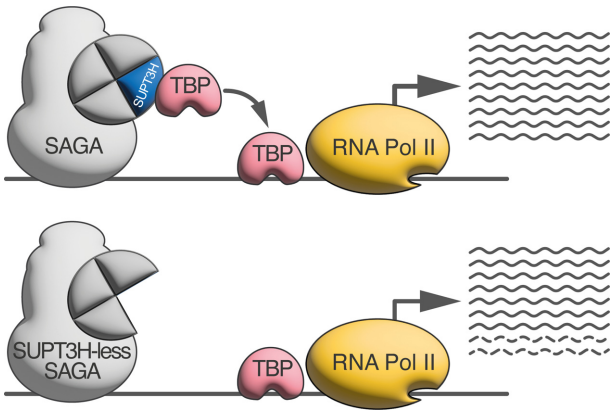
[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67404 Illkirch, France, [2]Centre National de la Recherche Scientifique (CNRS), UMR7104, 67404 Illkirch, France, [3]Institut National de la Santé et de la Recherche Médicale (INSERM), U1258, 67404 Illkirch, France, [4]Université de Strasbourg, 67404 Illkirch, France, [5]Plateforme GenomEast, infrastructure France Génomique, 67404 Illkirch, France and [6]CRBM, University of Montpellier, CNRS, Montpellier, France

Received May 30, 2022; Revised July 05, 2022; Editorial Decision July 05, 2022; Accepted July 12, 2022

ABSTRACT

Coactivator complexes regulate chromatin accessibility and transcription. SAGA (Spt-Ada-Gcn5 Acetyltransferase) is an evolutionary conserved coactivator complex. The core module scaffolds the entire SAGA complex and adopts a histone octamer-like structure, which consists of six histone-fold domain (HFD)-containing proteins forming three histone-fold (HF) pairs, to which the double HFD-containing SUPT3H adds one HF pair. Spt3, the yeast ortholog of SUPT3H, interacts genetically and biochemically with the TATA binding protein (TBP) and contributes to global RNA polymerase II (Pol II) transcription. Here we demonstrate that (i) SAGA purified from human U2OS or mouse embryonic stem cells (mESC) can assemble without SUPT3H, (ii) SUPT3H is not essential for mESC survival, but required for their growth and self-renewal, and (iii) the loss of SUPT3H from mammalian cells affects the transcription of only a specific subset of genes. Accordingly, in the absence of SUPT3H no major change in TBP accumulation at gene promoters was observed. Thus, SUPT3H is not required for the assembly of SAGA, TBP recruitment, or overall Pol II transcription, but plays a role in mESC growth and self-renewal. Our data further suggest that yeast and mammalian SAGA complexes contribute to transcription regulation by distinct mechanisms.

GRAPHICAL ABSTRACT



INTRODUCTION

Formation of the transcription preinitiation complex (PIC), containing RNA polymerase II (Pol II) and six general transcription factors (GTFs), is a major regulatory step in eukaryotic gene expression (1,2). PIC formation is mediated by TFIID and the accessibility of the transcription machinery to template DNA is controlled by co-activator complexes. The SAGA complex (Spt-Ada-Gcn5 Acetyltransferase) is an evolutionary conserved, multifunctional coactivator composed of about 18–20 subunits (3,4). Early genetic studies, predominantly performed in yeast, have established that SAGA is organized into distinct functional and structural entities, now called modules. These modules comprise a structural core, histone acetyltransferase (HAT), hi-

*To whom correspondence should be addressed. Tel: +33 3 88 65 34 44; Fax: +33 3 88 65 32 01; Email: laszlo@igbmc.fr
Correspondence may also be addressed to Stéphane D. Vincent. Tel: +33 3 88 65 34 25; Fax: +33 3 88 65 32 01; Email: vincent@igbmc.fr
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.
Present address: Elisabeth Lata, Bordeaux Institute of Oncology (BRIC), University of Bordeaux, INSERM U1312, 33076 Bordeaux, France.

stone deubiquitinase (DUB), TBP-binding, and transcription factor (TF)-binding activities (5).

Saccharomyces cerevisiae (y) *SPT3* was isolated as an allele-specific suppressor of a mutation in the *TBP* gene, which encodes the TATA binding protein (TBP), suggesting that Spt3 may play a role in recruiting yeast SAGA (ySAGA) to promoters (6–8). The TBP-binding module of ySAGA comprises Spt3 and Spt8, which both interact with TBP directly (9,10). Recent high resolution cryo-electron microscopy (cryo-EM) structures of ySAGA complexes indicated that the core structural module contains a histone octamer-like structure consisting of seven histone fold domain (HFD)-containing proteins, including Ada1/Taf12, Taf6/Taf9 and Taf10/Spt7, which form three HF pairs, and Spt3, which harbours two HF domains forming an intramolecular HF pair (10–12). Interestingly, Spt3 is homologous to the corresponding Taf11/Taf13 HF heterodimer in TFIID, and Spt3 of ySAGA binds the same side of TBP as does the Taf11/Taf13 HF pair (12–16). Purified recombinant TBP binds to the histone octamer-like structure of ySAGA through Spt3 and Spt8 (10,14), further arguing that yeast SAGA contributes to TBP delivery to core promoters. Indeed, structural and biochemical evidence suggest a model by which Spt3 prevents spurious TBP binding to DNA through steric hindrance, which is relieved by the synergistic binding of TFIIA and a cognate TATA element to TBP (10). In agreement, transcription of stress-inducible genes containing a TATA box in their promoters is more sensitive to SAGA subunit mutations than to TFIID subunit mutations (17). However, recent analyses in yeast based on quantification of newly synthesized mRNA, demonstrated that both TFIID and SAGA are required for the transcription of almost all Pol II genes, although each subunit may act through different mechanisms and thus to different extent (18–20). Interestingly, ySAGA can assemble in absence of Spt3 and Spt8 and *SPT3* mutant yeasts are viable and show a mild growth phenotype (9,21,22). Surprisingly, however, constitutive deletion of *SPT3* has a global effect on Pol II transcription indicating that Spt3 contributes to the expression of a large fraction of genes in *S. cerevisiae* (18).

The modular organization of mammalian SAGA complexes is very similar to that of their yeast counterparts (4). For example, their histone octamer-like structure is built up by three homologous HF pairs, TADA1/TAF12, TAF6L/TAF9/9b and TAF10/SUPT7L, and SUPT3H, the mammalian homologue of yeast Spt3 (23,24). However, metazoan SAGA complexes are lacking a Spt8 homologue (4) and there is currently no evidence that the SAGA mediated loading of TBP is conserved beyond *S. cerevisiae*. In addition, fly and mammalian SAGA complexes contain also a splicing module (4,24,25).

Here, we investigated the role of SUPT3H in SAGA complex assembly, Pol II transcription regulation and TBP recruitment to target gene promoters in human U2OS and mouse embryonic stem cells (ESCs). Our results indicate that, in both cell types, SAGA subunit composition is not affected by the absence of SUPT3H, similar to *S. cerevisiae* (9,21,22). In contrast, we observed a striking divergence between ySpt3 and mammalian SUPT3H orthologues in their regulatory roles. While we previously showed that yeast Spt3 contributes to global Pol II transcription (18), we show here that SUPT3H is required only for the expression of a limited number of genes in both cell types. TBP recruitment experiments further show that absence of SUPT3H does not affect TBP binding to selected gene promoters. Despite these limited regulatory roles, we found that *Supt3* gene inactivation in mESCs affects growth and self-renewal. Together these data suggest that the role of SUPT3H in gene transcription diverged substantially between yeast and mammals, suggesting either that mammalian SAGA complexes contribute to transcription regulation by distinct mechanisms, or that SAGA-dependent TBP delivery to promoters requires other subunits.

MATERIALS AND METHODS

Reagents

Reagents are described in Supplementary Table S1.

Biological resources

Biological resources are described in Supplementary Table S2, and see below:

a) **Generation of stable Flag-SUPT3H overexpressing U2OS cell lines**

The coding region (CDS) of the human *SUPT3H* gene was PCR amplified from the pREV-*SUPT3H* vector (26) using HA and 13662980/1 primers described in Supplementary Table S3, and Phusion™ High Fidelity polymerase (Thermo Fisher Scientific, Cat# F-530), following manufacturer's instructions. This first PCR product was purified on a gel using NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Cat# 740609.50). A second round of PCR using the same primer pair was performed to amplify the first amplicon and to eliminate pREV-*SUPT3H* template. The final product was purified on gel and cloned into the pSG5-puro A frame expression vector (27). First the *SUPT3H* PCR product was digested with Xho I and Sma I restriction enzymes (New England Biolabs, Cat# R0146 and R0141), and purified using Nucleospin® PCR clean-up kit (Macherey-Nagel, Cat# 740609.50) and ligated into a gel purified Xho I and Sma I fragment of the pSG5-puro expression vector using T4 DNA ligase (New England Biolabs, Cat# M0202). The ligation was transformed in DH5α competent bacteria and ampicillin resistant colonies were selected and validated by sequencing. The pSG5puro-*hSUPT3H* final plasmid was amplified and purified using Xtra Maxi NucleoBound kit (Macherey-Nagel, Cat# 740414.50).

U2OS cells were transfected with 5 μg of circular pSG5puro-*hSUPT3H* plasmid, using Lipofectamine 2000 following manufacturer's instructions (ThermoFisher Scientific, Cat# 11668019). One single cell per well was seeded in 96-well plates using the BD Biosciences FACSAria II (BD Biosciences) apparatus. After a 10-days puromycin selection, positive cell clones were selected by western blot analysis using the anti-FLAG and anti-SUPT3H antibodies.

b) **Generation of *Supt3*$^{-/-}$ mutant mESC lines**

Mouse E14 ESCs were transfected with a plasmid construct encoding for two sgRNAs (sequences in Supplementary Table S4) as well as a Cas9-GFP fusion protein at a confluency of 70–80% using Lipofectamine2000 (ThermoFisher Scientific, Cat# 11668019) following manufacturer's instructions. Two days after transfection, cells were selected for expression of the Cas9-GFP fusion protein by fluorescence activated sorting and five 96-well plates were seeded with one GFP positive cell per well using the BD FACSAriaTM II (BD Biosciences) apparatus. Clones were amplified and genomic DNA was extracted to select knockout clones by PCR, using the Phire direct PCR kit (Thermo Scientific, Cat# F-1265) following manufacturer's instructions. Primer sequences for PCR are shown in Supplementary Table S3. Homozygous mutant clones were sequenced using primers spanning the deletion site (Supplementary Table S3).

Cell culture conditions

Human HeLa cervical carcinoma cells (CCL-2; ATCC) and human U2OS osteosarcoma cells (HTB-96; ATCC) and U2OS Fl-SUPT3H cells were cultured using DMEM medium supplemented with 10% fetal calf serum (Sigma Aldrich, Cat# F7524) and 40 μg/ml gentamycin (KALYS, Cat# G0124-25). Cells were grown at 37°C in a humidified, 5% CO_2 incubator.

Mouse ES E14 cells were cultured on plates coated with 0.1% gelatin solution in 1× PBS (Dutcher, Cat# P06-20410) using DMEM medium supplemented with 15% fetal calf serum ES-tested (ThermoFisher Scientific, Cat# 10270-106), 2 mM L-glutamine (ThermoFisher Scientific, Cat# 25030-024), 0.1% β-mercaptoethanol (ThermoFisher Scientific, Cat# 31350-010), 100 UI/ml penicillin and 100 μg/ml streptomycin (ThermoFisher Scientific, Cat# 15140-122), 0.1 mM non-essential amino acids (ThermoFisher Scientific, Cat# 11140-035) and 1500 U/ml leukemia inhibitory factor (home-made). For medium described as FCS + LIF + 2i medium, 3 μM CHIR99021 (Axon Medchem, Cat# 1386) and 1 μM PD0325901 (Axon Medchem, Cat# 1408) were added freshly to the medium. Cells were grown at 37°C in a humidified, 5% CO2 incubator.

Schizosaccharomyces pombe cells were grown in autoclaved YES medium (yeast extract, adenine, histidine, uracil, leucine, lysine, 3% glucose) at 32°C.

Clonal assays of mouse ESCs

For clonal assay analyses, 1500–3000 cells, adapted to the respective media through at least three passages, were plated in wells of six-well plates. Medium was changed every other day. On the sixth day, colonies were washed twice with 1× PBS before fixation with 4% paraformaldehyde (Electron Microscopy Sciences, Cat# 15710, 16% solution) for 30 min followed by two washes with 1× PBS at room temperature. To assess the alkaline phosphatase (AP) activity of mESC colonies, Alkaline Phosphatase Kit (Vector Laboratories, Cat# SK-5100) was used following manufacturer's instructions. Colonies were stained with AP for 5–10 min. For clonal assay analyses in FCS + LIF medium, an additional staining with 0.1% crystal violet was performed after AP staining for at least 30 min.

Quantification of clonal assays

For clonal assay analyses in FCS + LIF + 2i medium, colony areas were measured automatically using ImageJ software (28). For clonal assay analyses in FCS + LIF medium, the colonies were counted manually using the ImageJ interface. We considered colonies as AP+ colonies if they either stained entirely red or if they possessed a centre of red cells surrounded by unstained cells. The total number of colonies as assessed by crystal violet staining, was used for normalization between replicates.

4sU metabolic labelling

Metabolic labelling of newly synthesized RNA was adapted from previously described protocols (29–31). In brief, cells were plated the day before to reach >90% of confluency before labelling. The nucleoside analogue 4-thiouridine (4sU) (Glentham Life Sciences, Cat# GN6085) was added to the cell culture medium at a final concentration of 500 μM for a 20-min pulse for mouse ES E14 and human U2OS cells. After the labelling period, the medium containing 4sU was removed, the cells were washed with ice cold 1× PBS and immediately lysed using TRI Reagent® (Molecular Research Center Inc., Cat# TR 188).

S. pombe cultures were grown to an OD_{600} of 0.8. Fourthiouracil (Sigma Aldrich, Cat# 440736) was freshly dissolved in DMSO and added to the cultures at a final concentration of 1 mM. Labelling was performed for 6 min. After this time period, yeast cells were pelleted, washed with ice-cold 1× PBS and aliquoted before being flash frozen in liquid nitrogen and stored at –80°C.

Total RNA extraction

Total RNA was extracted following TRI Reagent® (Molecular Research Center Inc., Cat# TR 188) manufacturer's instructions. To remove any potential genomic DNA contamination from the total RNA extracts, the TURBO DNA-*free*™ Kit (ThermoFisher Scientific, Cat# AM1907) was used following manufacturer's instructions for rigorous DNase treatment.

For total RNA extraction of yeast cells, the RiboPure™ RNA Purification Kit (ThermoFisher Scientific, Cat# AM1926) was used following manufacturer's instructions.

Reverse Transcription

Reverse Transcription (RT) was performed with 2 μg total RNA and using 0.4 μg random hexamer primers (ThermoFisher Scientific, Cat# SO142) and 10 U Transcriptor Reverse Transcriptase (Roche, Cat# 03531287001) following manufacturer's instructions.

qPCR

For qPCR, the cDNA samples were amplified using LightCycler® 480 SYBR® Green 2× PCR Master Mix I (Roche, Cat# 04887352001) and 0.3 or 0.6 μM of forward and reverse primer respectively. The primer pairs used for qPCR are listed in Supplementary Table S3. The qPCR was performed using a LightCycler® 480 (Roche). For the

assessment of mRNA levels, the obtained threshold-values were used to calculate the relative gene expression using the $2^{-\Delta\Delta CT}$ method and considering the individual primer pair efficiencies (32). For TBP ChIP-qPCR, the percentage of input was calculated.

Newly synthesized RNA purification

The purification of newly synthesized RNA was based on previously described protocols starting from 4sU-labelled total RNA (29–31). As spike-in, 4sU-labelled *S. pombe* total RNA was added in a ratio 1:10 to labelled mESC total RNA preparations; or in a ratio of 1:25 to labelled U2OS total RNA preparations, to a final amount of 200 μg of total RNA prior to newly synthesized RNA purification. The RNA was precipitated and resuspended in 130 μl RNase-free water (Sigma Aldrich, Cat# 95284) and sonicated on a E220 Focused-ultrasonicator (Covaris) using the following settings: 1% duty factor, 100 W, 200 cycles per burst, 80 s to reach fragment size range from 10 kb to 200 bp. For purification, the fragmented total RNA was incubated for 10 min at 60°C and immediately chilled on ice for 2 min to open secondary RNA structures. The 4sU-labelled RNA was thiol-specific biotinylated by addition of 200 μg EZ-link HPDP-biotin (ThermoFisher Scientific, Cat# 21341), biotinylation buffer (10 mM HEPES–KOH pH 7.5 and 1 mM EDTA) and 20% DMSO (Sigma Aldrich, Cat# D8418). Biotinylation was carried out for 3 h at 24°C in the dark and with gentle agitation. After incubation, biotin excess was removed by adding an equal volume of chloroform and centrifugation at 16 000 g for 5 min at 4°C. RNA was precipitated from the aqueous phase by adding 0.1 volumes of 5 M NaCl and an equal volume of 100% isopropanol followed by centrifugation at 16 000 g for 30 min at 4°C. After washing with 75% ethanol, the RNA pellet was resuspended in 100 μl of RNase-free water and denatured for 10 min at 65°C followed by immediate chilling on ice for 5 min. The samples were incubated with 100 μl of streptavidin-coated μMACS magnetic beads (Miltenyi Biotec, Cat# 130-074-101) for 90 min at 24°C under gentle agitation. The μMACS columns (Miltenyi Biotec, Cat# 130-074-101) were placed on a MACS MultiStand (Miltenyi Biotec) and equilibrated with washing buffer (100 mM Tris–HCl pH 7.5, 10 mM EDTA, 1 M NaCl, 0.1% Tween20) before applying the samples twice on the columns. The columns were then washed one time with 600 μl, 700 μl, 800 μl, 900 μl and 1 ml washing buffer before elution of the newly synthesized RNA with two washes of 100 μl 0.1M DTT. The isolated newly synthesized RNA was recovered either using RNeasy MinElute Cleanup Kit (Qiagen, Cat# 74204) following manufacturer's instructions or by precipitation.

Library preparation of 4sU RNA-seq

For U2OS cells, 4sU RNA-Seq libraries were generated from 15 ng of purified, newly synthesized RNA using Illumina Stranded Total RNA Prep, Ligation with Ribo-Zero Plus kit and IDT for Illumina RNA UD Indexes, Ligation (Illumina, San Diego, USA, Cat# 20040525 and 20040553/4, respectively), according to manufacturer's instructions. Briefly, abundant ribosomal RNAs were depleted by hybridization to specific DNA probes and enzymatic digestion. The depleted RNAs were purified and fragmented using divalent cations at 94°C for 2 min. After random hexamers annealing, fragmented RNAs were then reverse transcribed into first strand complementary DNA (cDNA). Second strand cDNA synthesis further generated blunt-ended double-stranded cDNA and incorporated dTTP in place of dUTP to achieve strand specificity by quenching the second strand during amplification. Following A-tailing of DNA fragments and ligation of pre-index anchors, PCR amplification was used to add indexes and primer sequences and to enrich DNA libraries (30 s at 98°C; [10 s at 98°C, 30 s at 60°C, 30 s at 72°C] × 12 cycles; 5 min at 72°C). Surplus PCR primers were further removed by purification using SPRIselect beads (Beckman-Coulter, Villepinte, France, Cat# B23319).

For mESCs, 4sU RNA-seq libraries were generated from 15 to 50 ng of purified, newly synthesized RNA using TruSeq Stranded Total RNA LT Sample Prep Kit with Ribo-Zero Gold (Illumina, San Diego, CA, Cat# RS-122-2301) according to the Illumina protocol with the following modifications. Four-thiouridine-labelled RNA was cleaned up using 1.8× RNAClean XP beads and fragmented using divalent cations at 94°C for 1 min without depletion of rRNA. While double stranded cDNA synthesis and adapter ligation were performed according to manufacturer instructions, the number of PCR cycles for library amplification was reduced to 10 cycles. After purification using SPRIselect beads (Beckman-Coulter, Villepinte, France, Cat# B23319).

All the libraries were sequenced with 1 × 50 bp on a HiSeq4000 System (Illumina).

Data analysis of 4sU RNA-seq

Reads were preprocessed using CUTADAPT 1.10 (33) in order to remove adaptors and low-quality sequences and reads shorter than 40 bp. rRNA sequences were removed for further analysis. For samples VQFR188, VQFR189, VQFR191, VQFR192, the remaining reads were aligned to a hybrid genome composed of hg38 and ASM294v2 assemblies of *H. sapiens* and *S. pombe* genomes, respectively, and for mESC samples VQFR25, VQFR26, VQFR29, VQFR30 and SNVT220, SNVT221, SNVT222, SNVT223, the remaining reads were aligned to a hybrid genome composed of mm10 and ASM294v2 assemblies of *M. musculus* and *S. pombe* genomes respectively with STAR 2.5.3a (34). Gene quantification was performed with htseq-count 0.6.1p1 (35), using 'union' mode and Ensembl 93 annotations for all organisms except for *S. pombe* where Ensembl Fungi 41 annotations were used. For 4sU-seq data, 'type' option was set to 'gene' in order to take also into account reads aligned onto introns. Differential gene expression analysis was performed using DESeq2 1.16.1 (36) Bioconductor R package on *H. sapiens* or *M. musculus* counts normalized with size factors computed by the median-of-ratios method proposed by Anders and Huber (37) based on the *S. pombe* spike-in counts (using the following options: cooksCutoff = TRUE, independentFiltering = TRUE, alpha = 0.05). *P*-values were adjusted for multiple testing using the Benjamini and Hochberg method (38). For subse-

quent data analyses and visualization, only protein-coding genes were considered. Further, a threshold of 2 reads per kb gene length was used to define expressed genes in the U2OS and mESC datasets.

Genome coverage files used for the UCSC genome browser snapshots were generated as follow. For each condition, the BAM files for the two replicates were merged, sampled, sorted and indexed using samtools 1.9 (39). BigWig files were generated using the deeptools 2.5.3 bamCoverage command (40).

For TATA-less and TATA-box promoter analysis (violin plots in Figures 2D, 5F and 6F), TATA-box containing promoters were extracted for the human (hg38) and mouse (mm10) genomes using the Eukaryotic Promoter Database (EPD, https://epd.epfl.ch/index.php, (41)).

EU metabolic labelling in mESCs

After 3 passages in FCS + LIF + 2i or FCS + LIF medium, mESCs were plated on gelatinized glass slides and incubated with 1 mM of 5-ethynyl-uridine (EU, Molecular Probes™ Click-iT™ RNA Alexa Fluor™ 488 kit,ThermoFischer Scientific, Cat# C10329) during 1 h at 37°C. After fixation in 4% PFA 15 min at room temperature, cells were treated with the Molecular Probes™ Click-iT™ RNA Alexa Fluor™ 488 kit (ThermoFischer Scientific, Cat# C10329) according to the manufacturer's instructions. Cells were imaged using a LSM 510 laser-scanning microscope (Leica) and a 63x Plan APO objective and EU signal was quantified using ImageJ. The pictures are shown with the LUT 'Green Fire Blue' scale to better show the variation in fluorescence intensity.

Whole cell extract preparation from U2OS cells

The required number of cells were trypsinized, transferred into 1.5 ml Eppendorf tubes, centrifuged 600 g 4°C for 2 min, and washed twice with 1 ml 1× PBS. Pellets were resuspended in 1 packed cells volume (PCV) extraction buffer (400 mM KCl, 20 mM Tris–HCl pH 7.5, 20% glycerol, 2 mM DTT, 1× cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)). Tubes were frozen and thawed subsequently 4 times (from liquid nitrogen to ice) and centrifuged at 14 000 g 4°C for 10 min. The supernatants were stored at –80°C.

Human cell nuclear extract preparation

To enrich extracts for nuclear proteins, cells were harvested and washed twice with 1× PBS. Cell pellets were resuspended in 4 times PCV of hypotonic buffer (50 mM Tris pH 7.9, 1 mM EDTA, 1 mM DTT and 1x cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)), left to swell for 30 min on ice, then dounced 10 times using a B dounce homogenizer to break cytoplasmic membrane. After a 10 min centrifugation at 1000–1800 g, at 4°C, supernatant was removed and the pellet resuspended in a high salt buffer (50 mM Tris pH 7.9, 25% glycerol, 500 mM NaCl, 0.5 mM EDTA, 1 mM DTT and 1x cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)). To break the nuclear membranes, suspension was homogenized by douncing 20 times using a B dounce, then incubated at 4°C (under stirring), for 30 min prior to centrifugation at 10 000 g for 20 min at 4°C. The supernatant was dialyzed overnight at 4°C against an isotonic salt buffer (50 mM Tris pH 7.9, 20% glycerol, 5 mM $MgCl_2$, 100 mM KCl, 1 mM DTT and 1x×protein inhibitor cocktail). The dialyzed fraction was kept as nuclear extract.

Mouse ESC nuclear extract preparation

To enrich extracts for nuclear proteins, cells were harvested and washed twice with 1× PBS. Cell pellet was resuspended in hypotonic buffer (10 mM Tris–HCl pH 8.0, 1.5 mM $MgCl_2$, 10 mM KCl and 1× cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) and dounced 10–20 times using a B dounce homogenizer to isolate the nuclei. After centrifugation at 10 000 g for 10 min at 4°C, supernatant was removed and pellet resuspended in high salt buffer (20 mM Tris–HCl pH 8.0, 25% glycerol, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 450 mM NaCl, 0.1% NP40 and 1× protein inhibitor cocktail). Suspension was homogenized by douncing as described before prior to centrifugation at 10 000 g for 10 min at 4°C. The supernatant was kept at –80°C as nuclear extract.

Antibodies

The list of antibodies used in this study is shown in Supplementary Table S5. Anti-TADA2B (3122) polyclonal antibody (pAb) was obtained by immunization of rabbits with the C-terminal region (amino acids 221–420) of hTADA2B (Q86TJ2). For this, the *TADA2b* cDNA fragment was PCR amplified and cloned in pET15b vector (Novagen) using Nde I and BamH I sites. The 6xHis-TADA2B protein fragment was expressed in *E. coli* (BL21), centrifuged, lysed in a buffer L (50 mM Tris–HCl pH 7.5, 400 mM NaCl, 10% glycerol, 1 mM DTT and 1× cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)). Recombinant proteins were solubilized in buffer L containing 8 M urea, purified under denaturing conditions using a Ni^{2+}-NTA column, eluted with 300 mM imidazole and dialyzed in 1× PBS. Rabbits were immunized with the purified proteins for 6 weeks as described in (42). The obtained crude rabbit sera were then purified on an Affi-Gel 10/15 column (Bio-Rad, Cat# 153-6098) on which the purified protein fragment was immobilized. Then the column was extensively washed with 1× PBS, pAb3122 eluted with 0.1 M glycine pH 2.5 buffer and immediately neutralized with 2 M Tris pH 8.8.

Immunoprecipitation protocol from human cell extracts

Prior to immunoprecipitation, protein-A sepharose and ANTI-FLAG® M2 Affinity Gel beads (Sigma Aldrich, Cat# A2220) were washed three times with 1× PBS and two times with IP100 buffer (25 mM Tris–HCl 7.5, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP40, 100 mM KCl, 2 mM DTT and 1× cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)), prior to use. Nuclear extracts were pre-cleared with 1/10 volume of packed bead volume for 1 h at 4°C with agitation. For antibody binding, packed bead

volume corresponding to 1/10 volume of input extract was incubated with the antibody 1 h at room temperature, with agitation. Protein-A antibody bound sepharose beads were washed twice with IP500 buffer (25 mM Tris–HCl 7.5, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP40, 500 mM KCl, 2 mM DTT and 1× cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) buffer and then with IP100 buffer. Precleared extracts were then incubated overnight with washed protein-A antibody bound sepharose beads at 4°C. Protein complex bound beads were washed twice with IP500 buffer and twice with IP100 buffer. Complexes were subsequently eluted twice with one bead volume of 0.1 M glycine pH 2.8 buffer at room temperature and with agitation. Eluates were immediately neutralized to pH7.5 by adding the required quantity of Tris–HCl pH 8.8 buffer. Eluates were characterized by western blot or mass spectrometry analyses.

Immunoprecipitation protocol for mESC nuclear extracts

Prior to immunoprecipitation, Protein-A or Protein-G sepharose beads were washed three times with filtered 1× PBS and two times with IP100 buffer. Nuclear extracts were pre-cleared with 1/5 of 50% bead slurry for 2 h at 4°C with agitation. For antibody binding, the 50% bead slurry was incubated with 5–8 μg of antibody 2 h at 4°C with agitation. After incubation, beads were washed three times with IP500 buffer and twice with IP100 buffer before addition of 1/5 volume of the 50% antibody-bead slurry to the precleared nuclear extracts. Nuclear extracts were incubated with beads overnight at 4°C with agitation. After incubation, beads were washed three times with IP500 buffer and twice with IP100 buffer. Complexes were eluted from the beads using two subsequent 0.1 M glycine pH 2.8 elutions at room temperature and with agitation. Importantly, prior to anti-TAF10 IPs in mESC, nuclear extracts were depleted for TFIID by overnight incubation with beads coated with antibodies targeting the TFIID-specific subunit TAF7. This allowed to increase the purification efficiency for SAGA in anti-TAF10 IPs given that TAF10 is shared between the SAGA and TFIID complexes. All other IPs have been performed without pre-depletion. Eluates were then characterized by western blot or mass spectrometry analyses.

Mass spectrometry

Protein mixtures were trichloroacetic acid (Sigma Aldrich, Cat# T0699) -precipitated overnight at 4°C. Samples were then centrifuged at 16 000 g for 30 min at 4°C. Pellet were washed twice with 1 ml cold acetone and centrifuged at 16 000 g for 10 min at 4°C. Washed pellet were then urea-denatured with 8 M urea (Sigma Aldrich, Cat# U0631) in Tris–HCl 0.1 mM, reduced with 5 mM TCEP (tris(2-carboxyethyl)phosphine) for 30 min, and then alkylated with 10 mM iodoacetamide (Sigma Aldrich, Cat# I1149) for 30 min in the dark. Both reduction and alkylation were performed at room temperature and under agitation. Double digestion was performed with endoproteinase Lys-C (Wako, Cat# 125–05061) at a ratio 1/100 (enzyme/proteins) in 8 M urea for 4 h, followed by an overnight modified trypsin digestion (Promega, Cat# V5111) at a ratio 1/100 (enzyme/proteins) in 2 M urea. Both Lys-C and Trypsin digestions were performed at 37°C. Peptide mixtures were then desalted on C18 spin-column and dried on Speed-Vacuum before LC-MS/MS analysis. Samples were analyzed using an Ultimate 3000 nano-RSLC (Thermo Scientific, San Jose California) coupled in line with a LTQ-Orbitrap ELITE mass spectrometer via a nano-electrospray ionization source (Thermo Scientific, San Jose California). Peptide mixtures were loaded on a C18 Acclaim PepMap100 trap-column (75 μm ID × 2 cm, 3 μm, 100Å, Thermo Fisher Scientific) for 3.5 min at 5 μl/min with 2% Acetonitrile MS grade (Sigma Aldrich, Cat# 1207802), 0.1% formic acid (FA, Sigma Aldrich, Cat# 94318) in H_2O and then separated on a C18 Accucore nano-column (75 μm ID × 50 cm, 2.6 μm, 150Å, Thermo Fisher Scientific) with a 90 min linear gradient from 5% to 35% buffer B (A: 0.1% formic acid (FA) in H_2O / B: 99% acetonitrile MS grade, 0.1% FA in H_2O), then a 20 min linear gradient from 35% to 80% buffer B, followed with 5 min at 99% B and 5 min of regeneration at 5% B. The total duration was set to 120 min at a flow rate of 220 nl/min. The oven temperature was kept constant at 38°C. The mass spectrometer was operated in positive ionization mode, in data-dependent mode with survey scans from m/z 350–1500 acquired in the Orbitrap at a resolution of 120 000 at m/z 400. The 20 most intense peaks (TOP20) from survey scans were selected for further fragmentation in the Linear Ion Trap with an isolation window of 2.0 Da and were fragmented by CID with normalized collision energy of 35%. Unassigned and single charged states were rejected. The Ion Target Value for the survey scans (in the Orbitrap) and the MS2 mode (in the Linear Ion Trap) were set to 1E6 and 5E3 respectively and the maximum injection time was set to 100 ms for both scan modes. Dynamic exclusion was used. Exclusion duration was set to 20 s, repeat count was set to 1 and exclusion mass width was ±10 ppm. Proteins were identified by database searching using SequestHT (Thermo Fisher Scientific) with Proteome Discoverer 2.4 software (PD2.4, Thermo Fisher Scientific) on *Homo sapiens* database (UniProt, reviewed, release 2020_11_27, 20 309 entries) and *Mus musculus* database (UniProt, non-reviewed, release 2020_07_13, 55 428 entries). Precursor and fragment mass tolerances were set at 10 ppm and 0.6 Da respectively, and up to two missed cleavages were allowed. Oxidation (M) was set as variable modification, and carbamidomethylation (C) as fixed modification. Peptides were filtered with a false discovery rate (FDR) at 1%, rank 1 and proteins were identified with 1 unique peptide. Normalized spectral abundance factors (NSAF) were calculated for each protein as described earlier (43,44). First to obtain spectral abundance factors (SAF), spectral counts identifying a protein were divided by the protein length represented by the number of amino acids. Then to calculate NSAF values, the SAF values of each protein were divided by the sum of SAF values of all detected proteins.

Anti-TBP chromatin immunoprecipitation (ChIP)

Cells were washed with 1× PBS, fixed at room temperature with 1% PFA in 1× PBS for 20 min. Fixation was

stopped by adding glycine to a final concentration of 125 mM for 10 min at room temperature. Cells were washed twice with ice-cold 1× PBS and collected by scrapping. Cells were centrifuged at 2000 g for 5 min at 4°C and washed once with ice-cold 1x PBS. Cells were lysed in L1 buffer (50 mM Tris–HCl pH 8.0, 2 mM EDTA, 0.1% NP40, 10% glycerol and cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) and incubated on ice for 10 min before centrifugation at 2000 g for 10 min at 4°C. Nuclei were resuspended in L2 buffer (0.5% SDS, 10 mM EDTA ph 8.0, 50 mM Tris–HCl pH 8.0 and cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) and sonicated using a Focused Ultra-sonicator E210 (Covaris) to on average 300 bp fragments. Protein-G sepharose beads were washed twice with TE buffer (10 mM Tris–HCl pH 7.5, 1 mM EDTA) and blocked for 3 h with 1 μg/μl denatured yeast tRNA and 1.5% fish-skin gelatine. Beads were washed twice with TE buffer and stored at 4°C. For ChIP, 50 μg of chromatin was diluted with chromatin dilution buffer (0.01% SDS, 1.1% Triton-X 100, 1.2 mM EDTA, 167 mM NaCl, 16.7 mM Tris–HCl pH 8.0 and cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) to a volume of 800 μl and precleared in 30 μl of blocked bead slurry for 1 h at 4°C. Five microgram of anti-TBP antibody (Abcam, Cat# ab51841) was added to precleared chromatin and incubated overnight at 4°C. Fifty μl of bead slurry was added to samples and incubated for 2–4 h at 4°C with overhead shaking. After centrifugation for 2 min at 1000 g, beads were washed two times with low salt washing buffer (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris–HCl pH 8.0 and cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) for 10 min, two times with high salt washing buffer (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris–HCl pH 8.0 and cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) for 10 min, two times with LiCl washing buffer (250 mM LiCl, 1% NP40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris–HCl pH 8.0 and cOmplete protease inhibitor cocktail (Roche, Cat# 11836170001)) for 10 min and two times with TE buffer for 10 min. Bound fragments were eluted using with freshly prepared elution buffer (1% SDS and 0.1 M sodium bicarbonate) at room temperature. Chromatin was reverse crosslinked by addition of 0.2 M NaCl and 50 μg/ml RNase A. Samples were incubated at 37°C for 1 h. Twenty microgram of Proteinase K was added and samples were incubated at 65°C overnight in a thermomixer. DNA was isolated by phenol-chloroform purification and resuspended in water after precipitation.

Statistical analysis

All statistical tests were performed using R (version 3.6.0). Two-sided Wilcoxon rank sum test was used for 1 to 1 comparisons. The statistical details for individual experiments can be found in the figure legends or Results section. This includes number of replicates and statistical tests performed. The statistical threshold was defined with a *P*-value of <0.05.

RESULTS

Human U2OS cells do not express the SUPT3H subunit of the SAGA complex

When analysing RNA expression data available at the Human Protein Atlas (HPA) database for mRNAs expression of different SAGA subunits in a panel of 64 human cell lines, we noticed that *SUPT3H* is not expressed in U2OS cells, a human osteosarcoma cell line, while *SUPT3H* was expressed in all other cell lines analysed (Supplementary Figure S1A, S1B, S1C). Absence of active transcription was supported by the lack of H3K4me3 and H3K27ac marks as well as Pol II accumulation at the *SUPT3H* promoter, as compared to the neighbouring *RUNX2* gene expressed on the opposite strand (Supplementary Figure S1D). To verify this observation, we designed primer pairs to amplify the whole coding sequence (CDS) of *SUPT3H* from cDNA prepared from U2OS and HeLa cells. This experiment confirmed the lack of *SUPT3H* expression in U2OS cells (Figure 1A, Supplementary Figure S2A). Similar results were obtained when amplifying smaller fragments of the *SUPT3H* mRNA (Supplementary Figure S2A and S2B). As an additional control, we generated an U2OS cell line stably expressing FLAG-tagged wild type (WT) SUPT3H protein (Figure 1B, Supplementary Figure S2C), which we refer to as U2OS Fl-SUPT3H. Growth curve analyses revealed that U2OS Fl-SUPT3H cells proliferate at a comparable rate than the parental U2OS cells (Supplementary Figure S2D). In agreement with the HPA data, we could not detect the *SUPT3H* mRNA in U2OS cells, while it was clearly detectable in HeLa and U2OS Fl-SUPT3H cells (Figure 1A). Together these data indicate that the human U2OS osteosarcoma cell line does not express the SAGA subunit SUPT3H. Importantly, we could not detect major differences in the promoter sequence of the *SUPT3H* gene when comparing U2OS and HeLa cells, suggesting that the loss of *SUPT3H* expression is not caused by changes in the tested promoter sequence (Supplementary Figure S2E).

Absence of SUPT3H does not affect human SAGA complex composition

As SUPT3H is a subunit of the core module of the SAGA complex interacting with several subunits of the complex, we were wondering if, and to which extent, SAGA is assembled in U2OS osteosarcoma cells lacking SUPT3H. To this end we prepared nuclear extracts from U2OS cells and carried out immunoprecipitation (IP) coupled to mass spectrometry using antibodies raised against SUPT20H (a SAGA core module subunit), TADA2B and TADA3 (two HAT module subunits). As a positive control we used human HeLa cell nuclear extracts (NEs) for IPs with the same antibodies. Quantitative mass spectrometry analyses of anti-SUPT20H, -TADA2B and -TADA3 affinity purifications showed that the three antibodies purified all SAGA subunits except SUPT3H from U2OS NEs, and all SAGA subunits from HeLa NEs (Figure 1C and D). Thus, the absence of SUPT3H does not affect the overall subunit composition of human SAGA. Importantly, anti-SUPT20H and anti-FLAG IPs performed using nuclear extracts from U2OS Fl-SUPT3H cells revealed that SUPT3H, when ex-

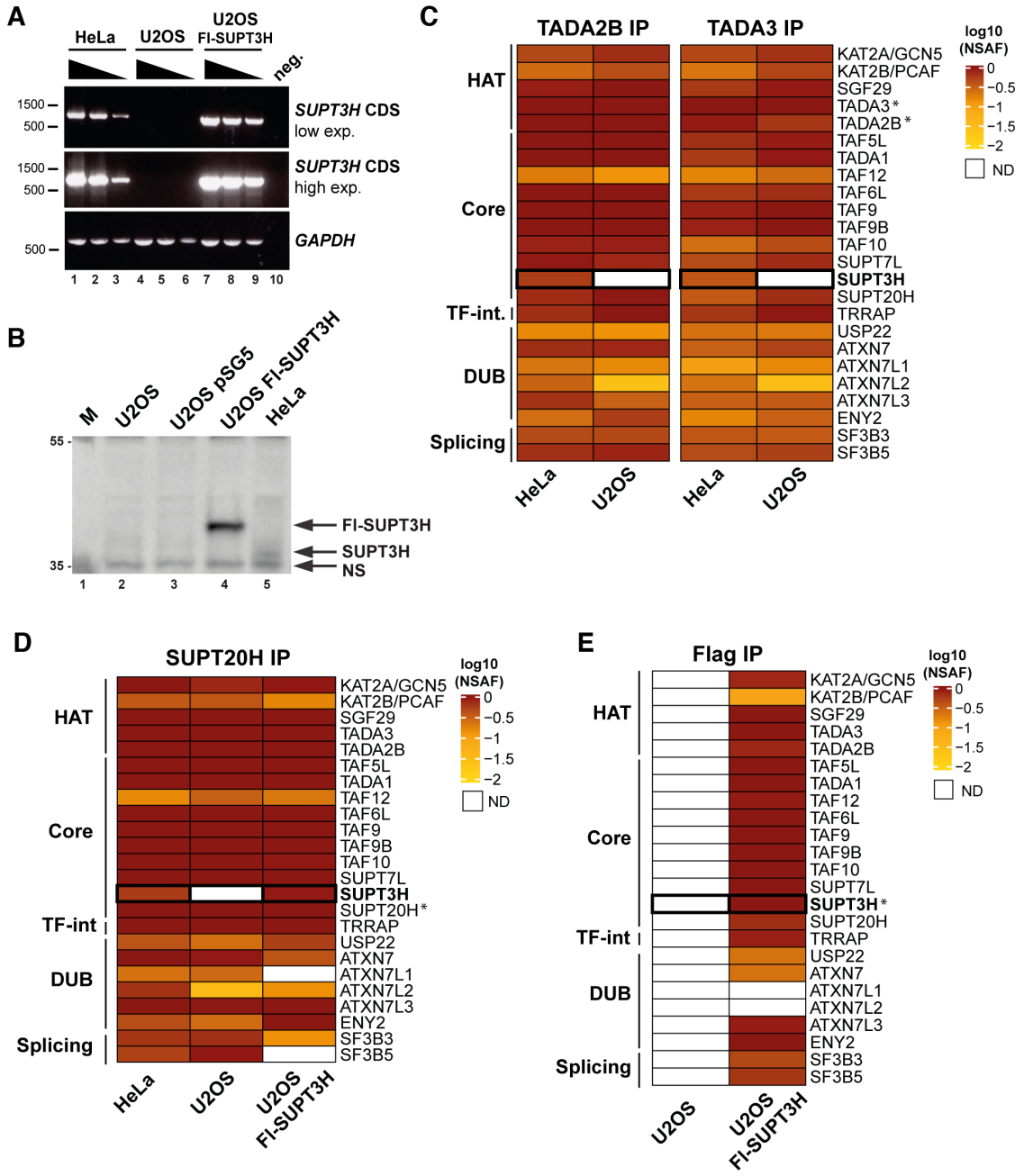


Figure 1. Absence of the SAGA subunit SUPT3H in human U2OS cells does not hinder assembly of the SAGA complex. (**A**) PCR of the coding sequence (CDS) of *SUPT3H* using cDNA obtained from HeLa cells, U2OS cells or U2OS cells re-expressing FLAG-SUPT3H (U2OS Fl-SUPT3H). Low and high exposures (exp.) are shown. *GAPDH* expression serves as loading control. $n = 3$ technical replicates. (**B**) Western blot analysis of U2OS parental cells, U2OS cells transfected with the empty expression vector pSG5, U2OS-cells over-expressing SUPT3H (U2OS-Fl-SUPT3H) and HeLa cells using anti-FLAG and anti-SUPT3H antibodies. The position of the proteins is indicated by arrows on the right of the blot. NS; non-specific signal. M: molecular weight marker (in kDa). (**C**) Log_{10}-transformed bait-normalized NSAF (Normalized Spectral Abundance Factor) values of mass spectrometry results from anti-TADA2B and anti-TADA3 immunoprecipitations (IPs) of the SAGA complex from HeLa and U2OS nuclear extracts. $n = 3$ technical replicates in each IPs. (**D**) Log_{10}-transformed bait-normalized NSAF values of mass spectrometry results from anti-SUPT20H IPs of the SAGA complex from HeLa, U2OS and U2OS Fl-SUPT3H nuclear extracts. $n = 3$ technical replicates. (**E**) Log_{10}-transformed bait-normalized NSAF values of mass spectrometry results from anti-FLAG IPs, also shown in Supplementary Figure S2F. $n = 3$ technical replicates. In (C, D, E), the stars (*) indicates the bait proteins to which the IPs were normalized. The distinct SAGA modules are indicated as HAT = histone acetyltransferase; TF-int = transcription factor-interacting; DUB = deubiquitylation and Splicing. ND = not detected.

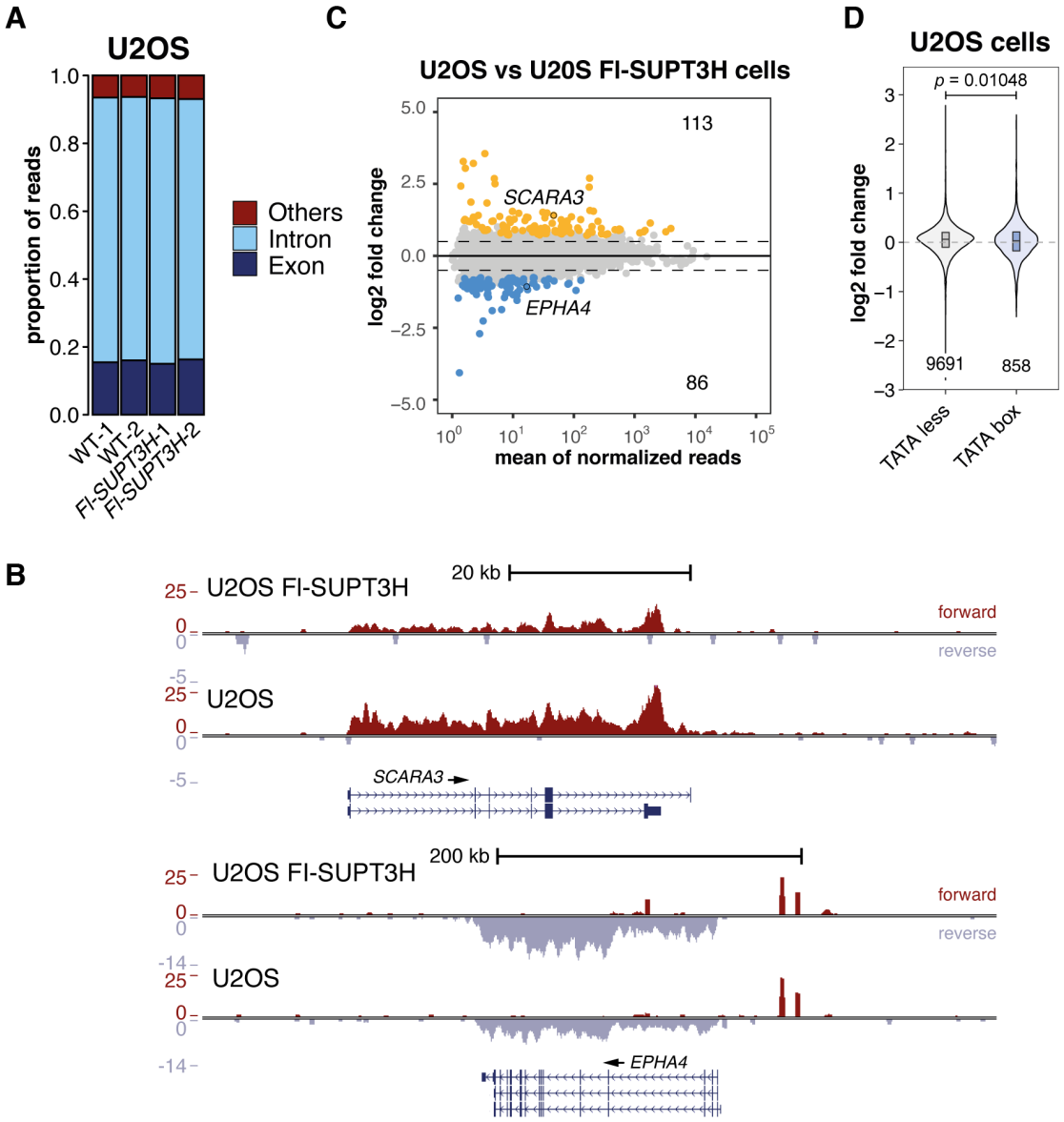


Figure 2. Only a small subset of Pol II-transcribed genes is responsive to the re-expression of SUPT3H in human U2OS cells. (**A**) Proportion of reads per genomic element for 4sU-seq experiments for two independent U2OS and Fl-SUPT3H clones. Exon = reads aligning to exons; Intron = reads aligning to introns, Others = reads aligning to intergenic regions, exon-intergenic junctions and exon-intron junctions. (**B**) UCSC genome browser views of 4sU-seq experiments on two differentially expressed genes (*SCARA3* and *EPHA4*) between U2OS Fl-SUPT3H and U2OS cells. Forward and reverse strands are shown. Y axes indicate the nascent mRNA sequencing coverage. Arrows indicate direction of transcription. (**C**) U2OS versus U2OS Fl-SUPT3H 4sU-seq comparison represented as MA plot with the numbers (on the right) of significantly down- (blue) or up-regulated (orange) genes, using two independent clones. A threshold of 2 normalized reads per gene length was used to define actively expressed genes. The position of the two genes shown in (**B**) are indicated by black circles. (**D**) Violin plot representation of the distribution of the nascent transcripts $\log_2$ fold changes in TATA-less and TATA-box containing gene classes. The number of genes per category are indicated at the bottom. The statistical test performed is two-sided Wilcoxon rank-sum test; *P* value is indicated on the graph.

ogenously expressed, can incorporate in the SAGA complex in the U2OS nucleus, with only minor, but not significant variations in the association of the DUB and splicing modules that do not directly interact with SUPT3H (24) (Figure 1D and E and Supplementary Figure S2E). These results indicate that the other subunits of the core module are intrinsically capable to incorporate SUPT3H in U2OS cells, if exogenously expressed. Furthermore, exogenous expression of SUPT3H did not impact SAGA HAT activity as levels of H3K9ac remain unaffected (Supplementary Figure S3A). We were not able to assess SAGA DUB activity because H2Bub1 is not detectable in both U2OS and U2OS Fl-SUPT3H cells (see below the DUB activity test in mESCs; Supplementary Figure S3B). To conclude, the absence of SUPT3H expression in U2OS cells allowed us to demonstrate that this subunit is dispensable for SAGA assembly and does not modulate the enzymatic activities of SAGA.

Restoring SUPT3H expression modifies RNA polymerase II transcription at a subset of genes in U2OS cells

The ability of SUPT3H to reconstitute an intact SAGA complex in U2OS Fl-SUPT3H allowed us to test the role of SUPT3H in Pol II transcription in human cells. For this, we performed 4-thiouridine (4sU) labelling of nascent RNA coupled with sequencing of the labelled RNA (4sU-seq) (29–31), comparing two independent U2OS Fl-SUPT3H clones with parental U2OS control cells (Supplementary Figure 4A). We observed a very similar enrichment in intronic reads in all samples (Figure 2A), confirming a reproducible enrichment of newly synthesized RNAs genome-wide. Differential expression analyses revealed that restoring *SUPT3H* expression has only a limited effect on nascent transcript levels (Figure 2B and C, Supplementary Figure S4B). Specifically, 199 genes were significantly deregulated in U2OS cells (lacking SUPT3H) compared to U2OS Fl-SUPT3H cells (re-expressing SUPT3H) at a threshold of 0.5 absolute $\log_2$ fold change and an adjusted *P*-value of 0.05. Eighty-six genes are downregulated and 113 genes upregulated in U2OS cells when compared to U2OS Fl-SUPT3H (Figure 2C, Supplementary Figure S4B). Only 43 and 52 genes are down-and upregulated, respectively, >2-fold. We note, however, that the variability between replicates might under-estimate these numbers (Supplementary Figure S4A). Overall, no GO categories were significantly enriched amongst genes deregulated by *SUPT3H* overexpression and only very few genes are responsive to the re-expression of SUPT3H in U2OS cells.

Earlier studies suggested that yeast SAGA predominantly regulates TATA-box containing genes, presumably through Spt3 (7,45,46). We therefore compared expression changes between U2OS and U2OS Fl-SUPT3H for genes with either a TATA-less or a TATA-box in their promoters, as defined by the Eukaryotic Promoter Database (Material and Methods) (Figure 2D). We observed no significant difference between the two gene classes, suggesting that TATA-box containing gene promoters do not show a stronger sensitivity to SUPT3H re-expression than TATA-less promoters in human U2OS cells. In conclusion, SUPT3H appears to regulate the transcription of only a few genes in U2OS Fl-SUPT3H cells, independently of the presence a TATA-box element in their promoter. This observation is in marked contrast with the strong, global decrease of nascent transcript levels observed in *S. cerevisiaeSPT3* deletion mutants.

Loss of mouse *Supt3* does not affect SAGA subunit composition in mouse embryonic stem cells

Our results so far indicate that the function of ySpt3/hSUPT3H in RNA Pol II transcription has diverged substantially between *S. cerevisiae* and humans. To explore this further, we next turned to non-cancerous diploid mammalian cells and used CRISPR/Cas9 genome editing to inactivate the *Supt3* gene in mouse E14 ESCs. We obtained two individual clones with a homozygous deletion of exon 3 from the *Supt3* gene (Figure 3A), resulting in an out of frame stop codon. RT-qPCR analyses confirmed the deletion of the exon and the resulting decrease of remaining *Supt3* mRNA, presumably through non-sense mediated mRNA decay (Figure 3B). Western blot analyses confirmed the deletion as SUPT3H is not detectable in extracts from $Supt3^{-/-}$ mutant mESCs (Figure 3C). To determine the role of SUPT3H on mouse SAGA assembly, we purified SAGA using anti-SUPT20H or anti-TAF10 antibodies from WT and $Supt3^{-/-}$ mutant mESCs nuclear extracts and analysed their composition by mass spectrometry and western blot analyses. Both immunoprecipitation coupled to mass spectrometry analyses from $Supt3^{-/-}$ mESCs identified all core, HAT and TF-binding SAGA subunits, with the exception of SUPT3H, indicating that mouse SAGA can also assemble without SUPT3H (Figure 3D). We note that the association of the DUB and splicing modules with SAGA was weaker in both anti-SUPT20H or anti-TAF10 IPs, but independently of the genotype of mESCs, suggesting that our IP-MS experiments are not sensitive enough to detect reproducibly all the subunits of these modules. Moreover, the enzymatic activities of the related SAGA complexes were not impacted, as H3K9ac and H2Bub1 levels were unchanged in both WT and $Supt3^{-/-}$ mESCs (Supplementary Figure S3B). On the contrary, as described earlier, the loss of ATXN7L3, a critical subunit of the SAGA DUB module lead to the stabilization of H2Bub1 (Supplementary Figure S3B) (47). These results confirm that, as observed in human U2OS cells, mouse SUPT3H has no major role in SAGA integrity and HAT/DUB activities in mESCs. Thus, these data also suggest that the SAGA central HF module can assemble with a hexameric nucleosome-like structure without the intramolecular HF pair of SUPT3H.

Supt3 is required for mouse ESC growth and self-renewal

The derivation of two independent $Supt3^{-/-}$ mESC lines indicates that *Supt3* is not essential for mESC viability when cultured in medium containing fetal calf serum (FCS), leukemia inhibitory factor (LIF) and two small molecules that maintain efficient mouse ESC self-renewal (hereafter called FCS + LIF + 2i medium). However, we observed in clonal assays that $Supt3^{-/-}$ cells formed twice less and significantly smaller colonies compared to wildtype cells revealing that SUPT3H loss impairs mESC growth (Figure 4A and B). In agreement, growth curve analyses revealed decreased proliferation of $Supt3^{-/-}$ cells as compared to wildtype mESCs (Figure 4C).

To assess the importance of SUPT3H for the self-renewal capacities of mESCs, we performed clonal assays in medium without 2i (hereafter referred to as FCS + LIF medium) and observed about a 4-fold decrease in colony numbers in $Supt3^{-/-}$ mESCs grown in FCS + LIF medium, as compared to WT cells (Figure 4D). As a read-out of the self-renewal capacities of mESCs, we measured the proportion of alkaline phosphatase positive (AP+) colonies. When compared to WT mESCs, the $Supt3^{-/-}$ cells produced about 3-fold less AP + colonies (Figure 4E), demonstrating that SUPT3H absence has a major impact on the self-renewal capacities of mESCs. RT-qPCR analysis of pluripotency factors gene expression indicates that while the expression of *Tfcp2l1*, *Nanog* and *Esrrb* are downregulated, *Klf4*, *Pou5f1* and *Sox2* transcripts are only weakly, or not affected (Figure 4F). This suggests that the diminished self-renewal capacities of $Supt3^{-/-}$ mutant mESCs in the FCS + LIF medium is due to

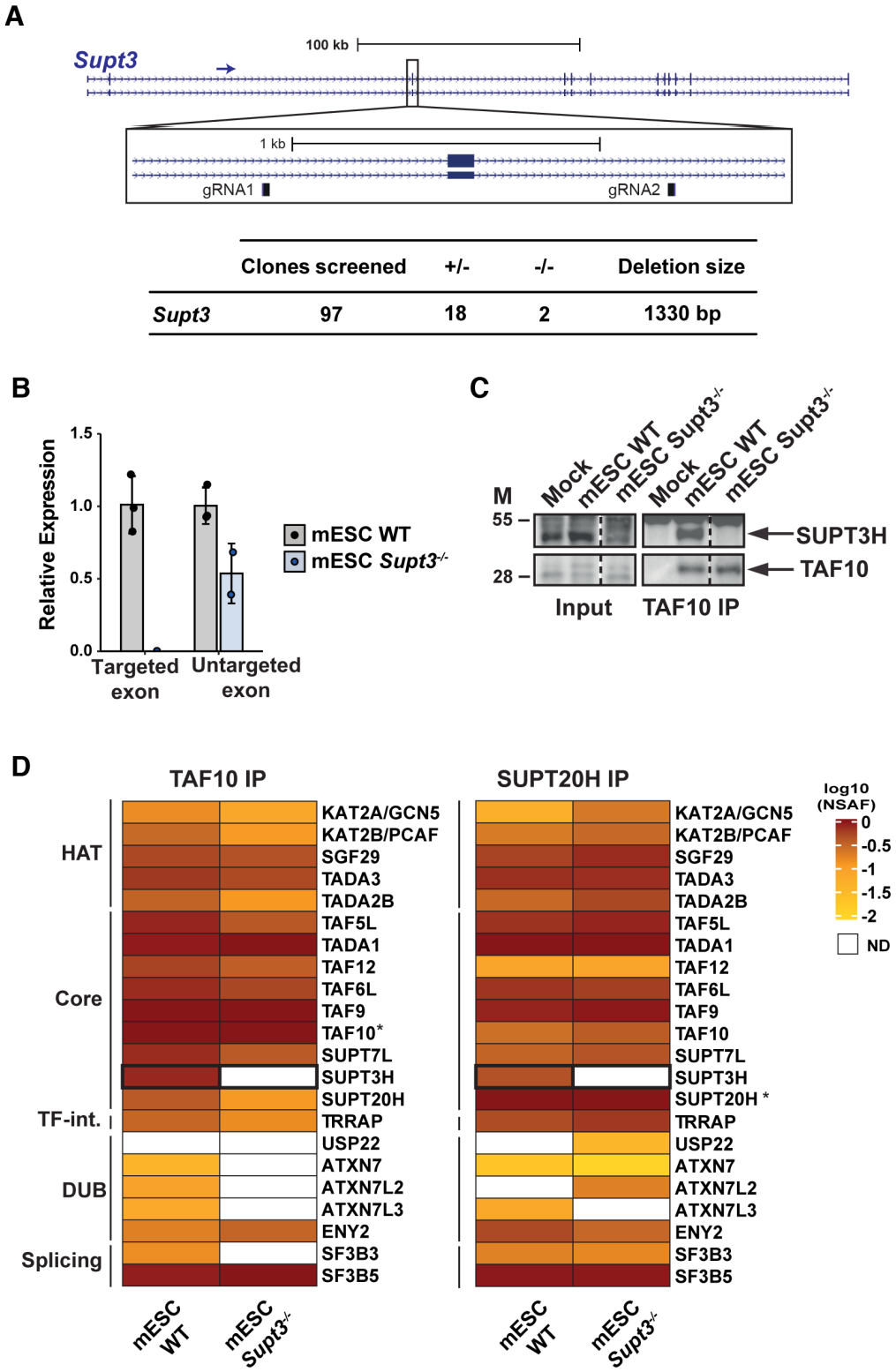


Figure 3. Loss of *Supt3* does not impair mouse ESC survival or formation of the SAGA complex. (**A**) Top, schematic representation of the mouse *Supt3* locus. The insert shows the position of the two gRNAs used to generate the *Supt3*$^{-/-}$ mESCs clones. Bottom, table showing the number of clones screened, the number of heterozygous (+/−) and of homozygous (−/−) clones obtained as well as the deletion size. (**B**) Validation of *Supt3*$^{-/-}$ cell lines by RT-qPCR revealing the absence of the targeted, out-of-frame exon 3 and overall reduced levels of the exon 3-deleted *Supt3* mRNA. Error bars show mean ± standard deviation (SD) of two independent clones, with each assessment of the cell lines being based on the mean of three technical replicates. (**C**) Western blot analysis of input and elution fractions of anti-TAF10 immunoprecipitations (IP) from wildtype (WT) and *Supt3*$^{-/-}$ mESCs. M = molecular weight markers (in kDa). (**D**) Log_{10}-transformed bait-normalized NSAF values of mass spectrometry results from anti-TAF10 IPs shown in (C) and anti-SUPT20H IPs. *n* = three technical replicates. Star (*) indicates the bait proteins, to which the IPs were normalized. The distinct SAGA modules are indicated as HAT = histone acetyltransferase; TF-int = transcription factor-interacting; DUB = deubiquitylation and Splicing. ND = not detected.

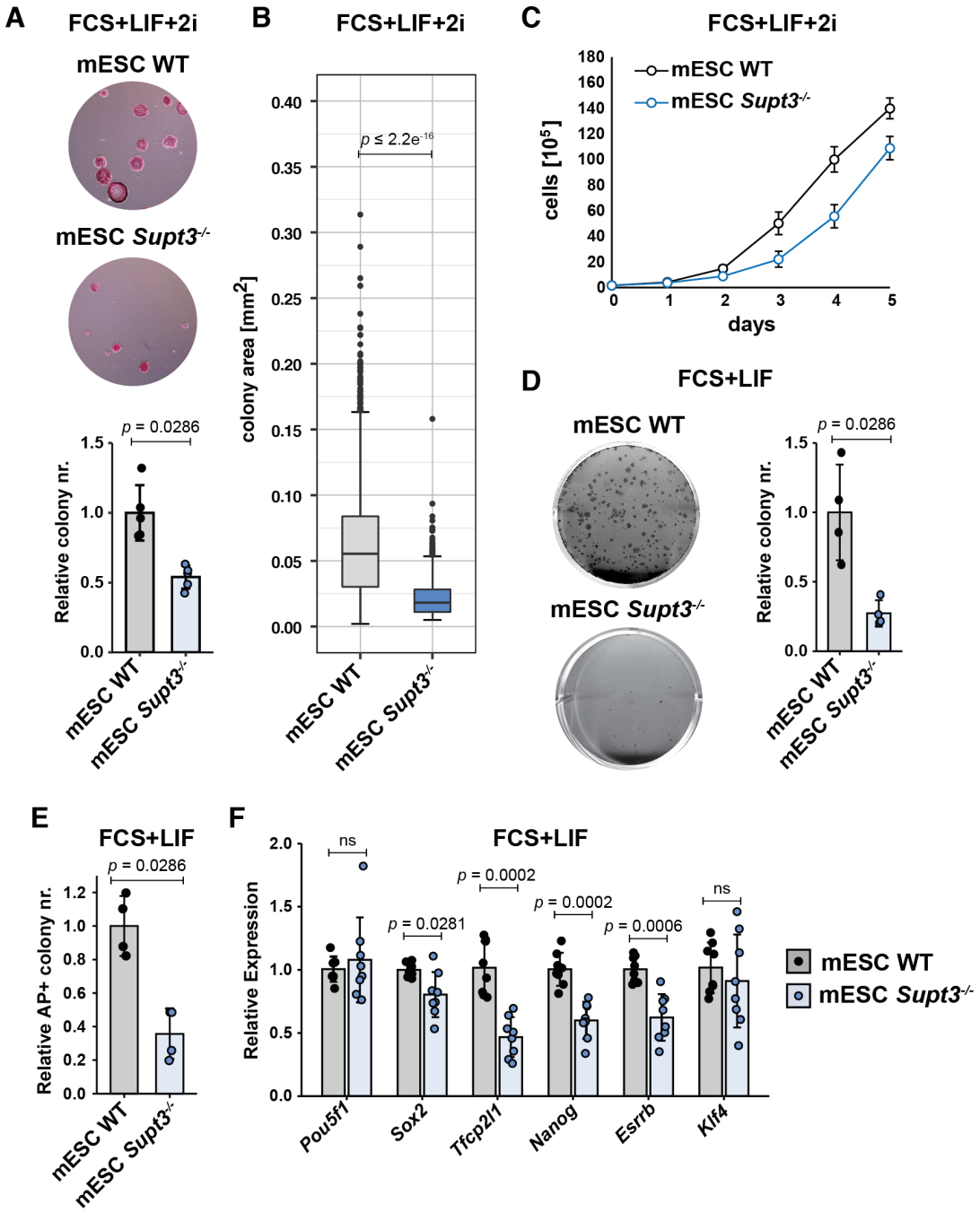


Figure 4. SUPT3H is required for mouse ESC growth and efficient self-renewal. (**A**) Clonal assay in FCS + LIF + 2i medium comparing WT and *Supt3*$^{-/-}$ mESCs. Colonies were stained for alkaline phosphatase activity. Top, representative images; bottom, quantification of colony numbers relative to WT condition. Error bars show mean ± standard deviation (SD) of four replicates from two biological replicates, with two independent clones each. (**B**) Quantification of colony areas using ImageJ. Error bars show mean ± standard deviation (SD) of four replicates from two biological replicates, with two independent clones each. (**C**) Growth curve analysis of viable cells comparing WT and *Supt3*$^{-/-}$ mESCs grown in FCS + LIF + 2i medium for 5 days. Error bars show mean ± standard deviation (SD) of two biological replicates using two independent clones. (**D**) Clonal assay in FCS + LIF medium comparing WT and *Supt3*$^{-/-}$ cells. Colonies were stained with crystal violet. Left, representative images; right, quantification of colony numbers relative to WT condition. Error bars show mean ± standard deviation (SD) of four replicates from two biological replicates, with two independent clones each. (**E**) Quantification of the number of colonies staining positive for alkaline phosphatase (AP+) in FCS + LIF medium normalized relative to WT cells. Error bars show mean ± standard deviation (SD) of four replicates from two biological replicates, with two independent clones each. (**F**) Relative mRNA levels of pluripotency transcription factors (as indicated) comparing WT and *Supt3*$^{-/-}$ mESCs grown in FCS + LIF medium. Expression was normalized to RNA polymerase III transcribed genes (*Rn7sk* and *Rpph1*) as well as to WT cells. Error bars show mean ± standard deviation (SD) of at least seven biological replicates (mean of three technical RT-qPCR replicates) using two independent clones each. The statistical tests performed are two-sided Wilcoxon rank-sum tests, *P* values are indicated on the graph. ns; not significant.

the deregulation of the expression of some genes associated with pluripotency maintenance.

The loss of SUPT3H in mouse ESCs has no major effect on Pol II transcription in FCS + LIF + 2i condition

To determine the impact of SAGA lacking SUPT3H on Pol II transcription in mESCs, we analysed nascent RNA by 4sU-seq. We performed the experiment with cells grown in FCS + LIF + 2i medium (Supplementary Figure S4C) in which *Supt3*$^{-/-}$ mESCs can maintain their self-renewal capacities (as assessed by AP staining, Figure 4A), although they grow more slowly (Figure 4C). We could observe a similar enrichment of intronic reads in all samples (Figure 5A). We observed that only a few genes (173 among 10462 expressed transcripts, 76 up-regulated and 97 down-regulated), are differentially expressed between *Supt3*$^{-/-}$ and WT mESCs (Figure 5B and C and Supplementary Figure S4D). Specifically, the newly synthesized levels of 25 transcripts were significantly decreased when applying a threshold of 0.5 $\log_2$ fold change and an adjusted *P*-value of 0.05. Conversely, the nascent levels of 36 transcripts increased in *Supt3*$^{-/-}$ cells using the same thresholds (Supplementary Table S6), although the variability between replicates might under-estimate these numbers (Supplementary Figure S4B). We could not find specific GO categories to be affected among the differentially expressed genes, however, the deregulation of some of these genes may explain the growth and self-renewal defects observed in *Supt3*$^{-/-}$ mESCs.

To confirm the lack of global effect of the absence of SUPT3H in FCS + LIF + 2i conditions, we used an alternative imaging approach by performing 5-ethynyl uridine (EU) incorporation into nascent RNAs. Lack of SUPT3H in FCS + LIF + 2i conditions did not significantly affect the EU incorporation (Figure 5D and E) further suggesting that the loss of SUPT3H has no major effect on Pol II transcription in mESCs in FCS + LIF + 2i condition.

Finally, we compared transcription changes between genes possessing TATA-less and TATA-box containing promoters (Figure 5F). We observed no difference between the two gene classes suggesting that TATA-box containing promoters do not show a stronger sensitivity to the loss of *Supt3* than TATA-less promoters in mESCs. These data together show that the mouse SAGA complex lacking SUPT3H, as observed in human U2OS cells, affects the expression of a small subset of genes in FCS + LIF + 2i medium.

The loss of SUPT3H in mouse ESCs has no major effect on Pol II transcription in FCS + LIF condition

As *Supt3*$^{-/-}$ mutant mESCs in FCS + LIF + 2i medium have only a mild phenotype, and as mESC maintenance of pluripotency is affected in the absence of SUPT3H in FCS + LIF medium (Figure 4D and E), we analysed nascent transcription in cells grown in FCS + LIF. A similar enrichment of intronic reads was observed in all samples (Figure 6A). More variation in gene expression was observed compared to the FCS + LIF + 2i condition (Figure 6B) as 898 transcripts are differentially expressed among 12 045 expressed transcripts, with more transcripts (703) being downregulated than upregulated in the absence of SUPT3H. However, due to the variability between the samples (Supplementary Figure S4E and F), statistical significance was reached only for 2 transcripts (Figure 6C). In order to confirm our observation, we carried out EU incorporation assay in these cells, in FCS + LIF conditions. Using this *in vivo* imaging approach, no major effect on global nascent transcription could be observed (Figure 6D), as quantification did not reveal a statistically significant difference between WT and *Supt3*$^{-/-}$ mutant mESCs (Figure 6E). Similarly to the FCS + LIF + 2i condition, we did not observe any difference between the TATA-box containing and TATA-less promoters transcription activity (Figure 6F). However, analysis of nascent transcripts of different loci encoding for pluripotency genes (Supplementary Figure S5) confirmed the reduced levels observed at steady-state mRNA levels as measured by RT-qPCR (Figure 4F) showing that only a subset of these transcripts is downregulated in absence of SUPT3H. These data suggest that in FCS + LIF condition, absence of SUPT3H does not have a major effect on nascent Pol II transcription, but contributes to activate the transcription of a subset of genes, including genes promoting pluripotency.

Loss of SUPT3H in human U2OS cells or mouse ESCs does not significantly affect TBP binding

Genetic, biochemical and cryo-EM studies of yeast SAGA indicate that the Spt3 and Spt8 subunits bind to TBP and are involved in delivering TBP at specific promoters (6,7,10,11,48,49). We therefore wanted to assess whether SUPT3H has a role in recruiting TBP to mammalian gene promoters. For this, we performed TBP chromatin immunoprecipitation (ChIP) coupled to qPCR at selected gene promoters, in both U2OS and mESC cells (Figure 7A and B). We observed no difference in TBP occupancy at the tested gene promoters when comparing cells with and without SUPT3H. This suggests that TBP binding at these promoters does not require SUPT3H in human and mouse cells.

DISCUSSION

Previous SAGA subunit deletion and purification studies in *S. cerevisiae* demonstrated that Spt3 is not essential for viability and that yeast SAGA can assemble in the absence of Spt3, whereas constitutive deletion of *SPT3* globally reduces Pol II transcription (9,18,21,22).

In this study, we report that U2OS cells do not express *SUPT3H*. No mutation in the basal promoter could be identified and as the neighbouring genes (*RUNX2*, for example) are actively expressed in U20S cells, a major chromosomal rearrangement can probably be ruled out. We showed that the lack of *SUPT3H* expression correlates with the absence of accumulation of active epigenetic marks. An interesting hypothesis would be that U2OS cells bear mutations in specific *SUPT3H* enhancers.

SUPT3H is not required for mammalian cell survival and we showed that mammalian SAGA complexes lacking SUPT3H can stably assemble, and that TBP recruitment at several TATA-containing and TATA-less promoters is

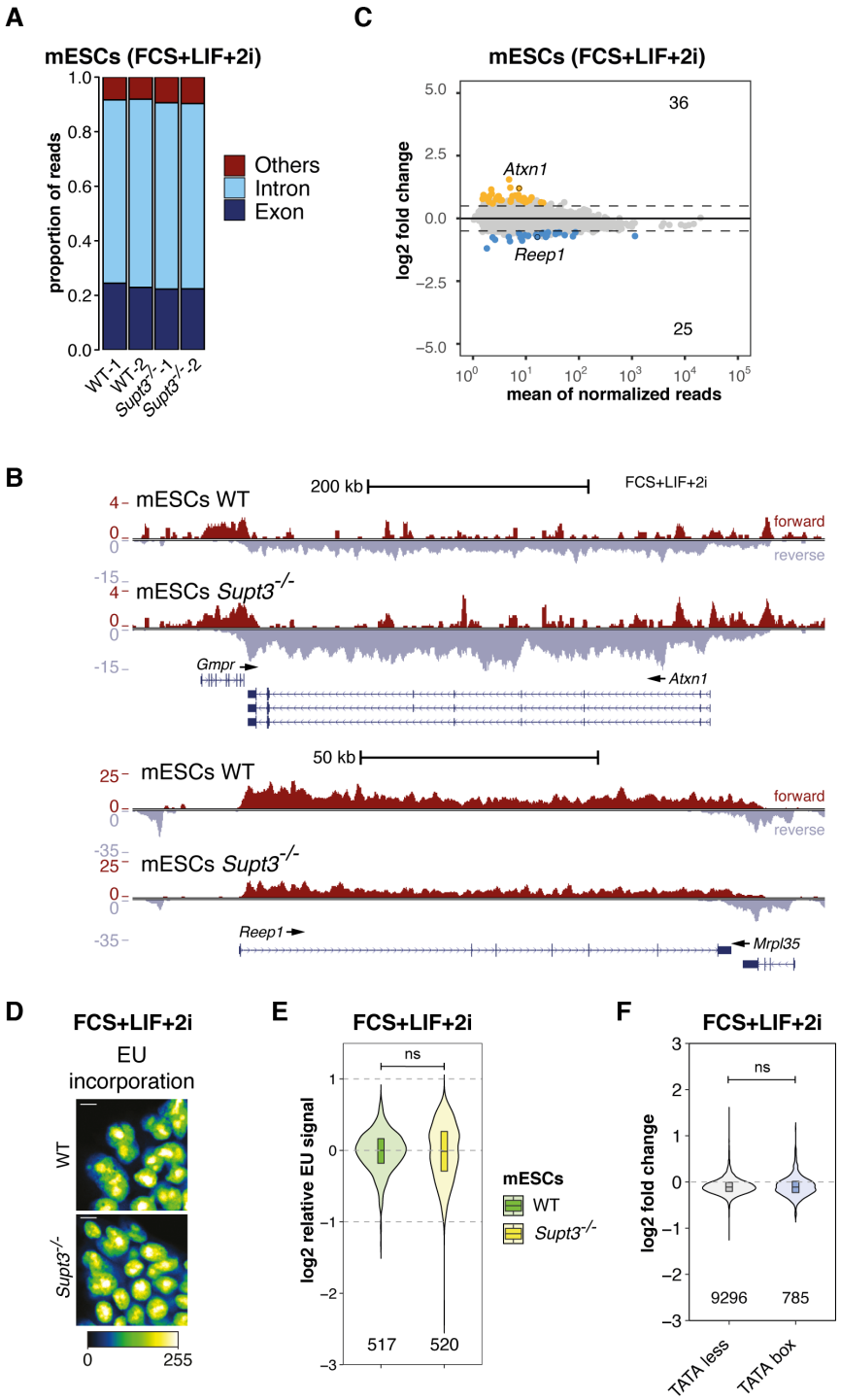


Figure 5. Loss of *Supt3* in mouse ESC has no major effect on Pol II transcription regulation in FCS + LIF + 2i medium. (**A**) Proportion of reads per genomic element for 4sU-seq experiments in two independent wildtype (WT) and *Supt3*$^{-/-}$ mESC clones in FCS + LIF + 2i medium. Exon = reads aligning to exons; Intron = reads aligning to introns, Others = reads aligning to intergenic regions, exon-intergenic junctions and exon-intron junctions. (**B**) UCSC genome browser views of 4sU-seq experiments on two differentially expressed genes (*Atxn1* and *Reep1*) between WT and *Supt3*$^{-/-}$ mESCs. Forward and reverse strands are shown. Y axes indicate the nascent mRNA sequencing coverage. Arrows indicate direction of transcription. (**C**) *Supt3*$^{-/-}$ versus WT mESCs 4sU-seq comparison represented as MA plot with the numbers (on the right) of significantly down- (blue) or up-regulated (orange) genes, in two WT and *Supt3*$^{-/-}$ mESC clones. A threshold of 2 normalized reads per kb gene length was used to define actively expressed genes. The position of the two differentially expressed genes shown in (B) are indicated by black circles. (**D**) Representative images of 5-ethynyl uridine (EU) incorporation in WT and *Supt3*$^{-/-}$ mESCs in FCS + LIF + 2i medium. Colour scale (Green Fire Blue LUT scale) is indicated at the bottom. Scale bar; 10 μm. (**E**) Quantification of the EU incorporation in WT and *Supt3*$^{-/-}$ mESCs in FCS + LIF + 2i medium, normalized to the WT condition and expressed in $\log_2$. The number of cells quantified is indicated at the bottom. (**F**) Violin plot representation of the distribution of the nascent transcripts $\log_2$ fold changes in TATA-less and TATA-box containing gene classes. The number of genes per category is indicated at the bottom. The statistical test performed is two-sided Wilcoxon rank-sum test. *P* values are indicated on the graphs. ns; not significant.

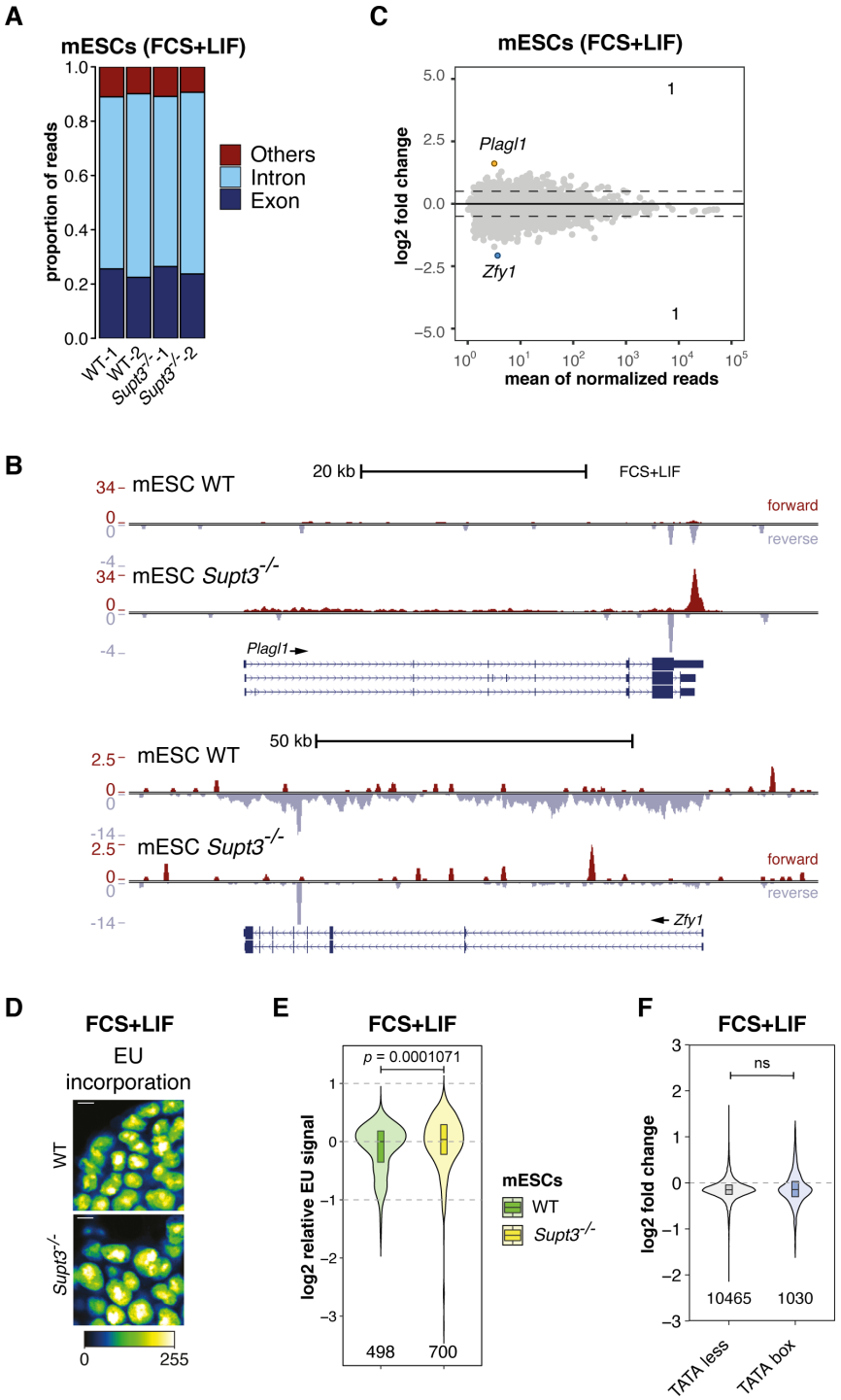


Figure 6. Loss of *Supt3* in mouse ESC has no major effect on Pol II transcription regulation in FCS + LIF medium. (**A**) Proportion of reads per genomic element for 4sU-seq experiments in two independent wildtype (WT) and *Supt3*$^{-/-}$ mESC clones. Exon = reads aligning to exons; Intron = reads aligning to introns, Others = reads aligning to intergenic regions, exon-intergenic junctions and exon-intron junctions. (**B**) UCSC genome browser views of 4sU-seq experiments on two differentially expressed genes (*Plagl1* and *Zfy1*) between WT and *Supt3*$^{-/-}$ mESCs. Forward and reverse strands are shown. Y axes indicate the nascent mRNA sequencing coverage. Arrows indicate direction of transcription. (**C**) *Supt3*$^{-/-}$ versus WT mESCs 4sU-seq comparison represented as MA plot with the numbers (on the right) of significantly down- (blue) or up-regulated (orange) genes, in two WT and *Supt3*$^{-/-}$ mESC clones. A threshold of 2 normalized reads per kb gene length was used to define actively expressed genes. The position of the two differentially expressed genes shown in (B) are indicated by black circles. (**D**) Representative images of 5-ethynyl uridine (EU) incorporation in WT and *Supt3*$^{-/-}$ mESCs in FCS + LIF medium. Colour scale (Green Fire Blue LUT scale) is indicated at the bottom. Scale bar; 10 μm. (**E**) Quantification of the EU incorporation in WT and *Supt3*$^{-/-}$ mESCs in FCS + LIF medium, normalized to the WT condition and expressed in $\log_2$. The number of cells quantified is indicated at the bottom. (**F**) Violin plot representation of the distribution of the nascent transcripts $\log_2$ fold changes in TATA-less and TATA-box containing gene classes. The number of genes per category is indicated at the bottom. The statistical test performed is two-sided Wilcoxon rank-sum test. *P* values are indicated on the graphs. ns; not significant.

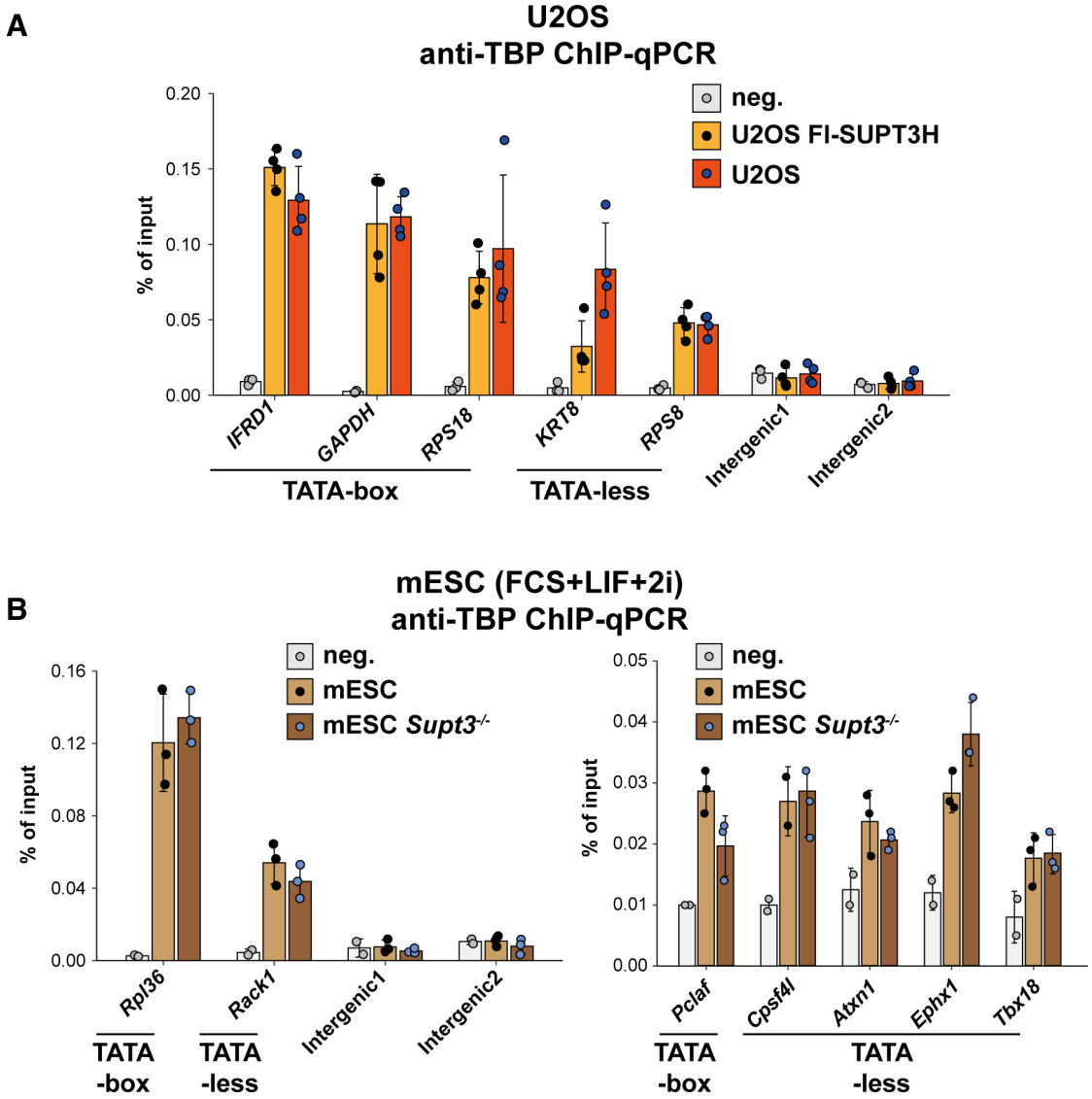


Figure 7. Loss of *SUPT3H* in human U2OS cells or mouse ESCs does not significantly affect TBP binding at promoters. (A, B) Quantification of anti-TBP ChIP-qPCR experiments in U2OS and U2OS Fl-SUPT3H cells (**A**) or in WT and *Supt3*$^{-/-}$ mESCs (**B**) at promoters of selected genes (as indicated) and at two independent intergenic regions. TATA-less and TATA-box-containing promoters are indicated. ChIP-qPCR experiments without antibody were used as negative control (neg.). Error bars show mean ± standard deviation (SD) of at four and three biological replicates using two independent clones for human and mouse data, respectively, in each sample, the mean shows at least four technical RT-qPCR replicates. Statistical test performed is two-sided Wilcoxon rank-sum test. Statistical significance ($P < 0.05$) was not reached.

not influenced by the absence of SUPT3H in mammalian SAGA. Nevertheless, mESCs lacking SUPT3H show impaired growth potential. Moreover, *Supt3*$^{-/-}$ mESCs grown in FCS + LIF medium display diminished self-renewal capacities. Our experiments further show that constitutive loss of SUPT3H in both human U2OS and mESCs grown in FCS + LIF + 2i or FCS + LIF medium does not globally impairs Pol II transcription.

Our data indicate that absence of SUPT3H only affects a subset of genes (Supplementary Table S6) but these genes are different between the different cell types and culture conditions. However, the observed very low (−0.1) $\log_2$ fold change median observed in the TATA-containing and TATA-less promoter categories comparisons from mESCs grown in FCS + LIF + 2i or FCS + LIF medium, would suggest a very weak tendency towards down-regulation of gene expression in the absence of SUPT3H (Figures 5F and 6F). The lack of common specific GO categories affected is in favour of a role in fine-tuning distinct gene expression programs within each cellular context. In particular, we observed that expression of some pluripotency factors (*Sox2*, *Tfcp2l1*, *Nanog* and *Esrrb*) is affected in *Supt3*$^{-/-}$ mESCs grown in FCS + LIF condition, but not in FCS + LIF + 2i condition. Further studies will be necessary to identify which effects are directly caused by the loss of SUPT3H and how these changes contribute to the growth and self-renewal defects observed in mESCs.

Importantly, *Supt3* deletion in mice indicated that SUPT3H is important for mouse embryogenesis as embryos die between E9.5 and E14.5 (50). This observation suggests

that SUPT3H is not essential for early mouse development, similarly to other SAGA subunits, such as GCN5, PCAF, SUPT20H, USP22 and ATXN7L3 (47,51–55).

Cryo-EM structural studies suggested that yeast Spt3, or human SUPT3H, are an important structural subunit of SAGA, as the double HF-containing subunit assembles with three pairs of other HFD-containing SAGA subunits to form a deformed histone octamer-like structure (10,11,24). Importantly, it has been described in ySAGA that while the six histone-fold pairs in the histone-fold octamer-like structure are oriented similarly as in the canonical nucleosome, the intramolecular HF pair of Spt3 is tilted by 20 degrees compared with its analogous nucleosome histone H2A–H2B histone pair (10). This tilt therefore could almost completely free Spt3 from its association with the histone-fold octamer. However, ySAGA subunit composition is not affected by the loss of Spt3 (9,21,22), which we confirmed for mammalian SAGA in both human and mouse cells. Interestingly, plants are probably also lacking a Spt3 orthologue (5,56). Moreover, no ortholog of Spt8 was detected in mammalian SAGA (5), further suggesting that the TBP-loading activity described in ySAGA, would be affected in mammalian cells lacking SUPT3H. Note however, that it is not known whether the SAGA mediated TBP loading to core promoters suggested in yeast exists in metazoans and plants. The minimal effect of SUPT3H on Pol II transcription and TBP recruitment suggests that the SAGA/Spt3-dependent promoter delivery of TBP mechanism may not be conserved in mammalian cells. Alternatively, TFIID may be able to compensate the constitutive loss of SUPT3H, masking the potential *in vivo* contribution of SAGA to TBP delivery in our experimental conditions. Finally, it is conceivable that SUPT3H may be important for TBP delivery and transcription of genes that are induced in response to specific developmental and/or environmental signals.

DATA AVAILABILITY

The 4sU-seq data generated for this publication have been deposited in NCBI's Gene Expression Omnibus (57) and are accessible through GEO Series accession numbers GSE175901 and GSE202823. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (58) partner repository with the dataset identifier PXD026991. BED and BigWIG datasets used in Supplementary Figure S1D were downloaded from GEO repository: GSM1901948 (U2OS; Pol II ChIP), GSM4194653 (U2OS; H3K4me3), GSM4194660 (U2OS; H3K27ac ChIP), GSM803533 (HeLa; Pol II ChIP), GSM733682 (HeLa; H3K4me3 ChIP) and GSM733684 (HeLa; H3K27ac ChIP) (59–62).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank all members of the Tora lab for protocols, thoughtful discussions and suggestions, A. Ben Shem for comments and critically reading the original version of the manuscript, the GenomEast platform [France Génomique consortium (ANR-10-INBS-0009)] for library preparation and sequencing; C. Ebel and M. Philipps for help with FACS sorting, B. Reina San Martin and N. Jung for help with the CRIPR/Cas9 *Supt3* KO experiment, designing and establishing the corresponding vectors, and the IGBMC cell culture facility for help with U2OS, HeLa and E14 ES cells.

FUNDING

Agence Nationale de la Recherche (ANR) [ANR-19-CE11-0003-02, ANR-PRCI-19-CE12-0029-01, ANR-20-CE12-0017-03]; NIH [MIRA grant R35GM139564]; NSF [NSF 1933344 to L.T., ANR-18-CE12-0026 to D.D., ANR-20-CE12-0014-02 to D.H. and D.D.]; IdEx-University of Strasbourg PhD program; 'Fondation pour la Recherche Médicale' (FRM) association [FDT201904008368 to V.F.]; [ANR-10-LABX-0030-INRT], under the frame program Investissements d'Avenir [ANR-10-IDEX-0002-02]. Funding for open access charge: Agence National de la Recherche (ANR).
Conflict of interest statement. None declared.

REFERENCES

1. Roeder,R.G. (2019) 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.*, **26**, 783–791.
2. Chen,X., Qi,Y., Wu,Z., Wang,X., Li,J., Zhao,D., Hou,H., Li,Y., Yu,Z., Liu,W. *et al.* (2021) Structural insights into preinitiation complex assembly on core promoters. *Science*, **372**, eaba8490
3. Spedale,G., Timmers,H.T. and Pijnappel,W.W. (2012) ATAC-king the complexity of SAGA during evolution. *Gene Dev*, **26**, 527–541.
4. Helmlinger,D. and Tora,L. (2017) Sharing the SAGA. *Trends Biochem. Sci.*, **42**, 850–861.
5. Helmlinger,D., Papai,G., Devys,D. and Tora,L. (2021) What do the structures of GCN5-containing complexes teach us about their function?*Biochim. Biophys. Acta Gene Regul. Mech.*, **1864**, 194614.
6. Eisenmann,D.M., Arndt,K.M., Ricupero,S.L., Rooney,J.W. and Winston,F. (1992) SPT3 interacts with TFIID to allow normal transcription in saccharomyces cerevisiae. *Genes Dev.*, **6**, 1319–1331.
7. Laprade,L., Rose,D. and Winston,F. (2007) Characterization of new Spt3 and TATA-binding protein mutants of saccharomyces cerevisiae: spt3 TBP allele-specific interactions and bypass of Spt8. *Genetics*, **177**, 2007–2017.
8. Grant,P.A., Winston,F. and Berger,S.L. (2021) The biochemical and genetic discovery of the SAGA complex. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1864**, 194669.
9. Sermwittayawong,D. and Tan,S. (2006) SAGA binds TBP via its Spt8 subunit in competition with DNA: implications for TBP recruitment. *EMBO J.*, **25**, 3791–3800.
10. Papai,G., Frechard,A., Kolesnikova,O., Crucifix,C., Schultz,P. and Ben-Shem,A. (2020) Structure of SAGA and mechanism of TBP deposition on gene promoters. *Nature*, **577**, 711–716.
11. Wang,H., Dienemann,C., Stutzer,A., Urlaub,H., Cheung,A.C.M. and Cramer,P. (2020) Structure of the transcription coactivator SAGA. *Nature*, **577**, 717–720.
12. Birck,C., Poch,O., Romier,C., Ruff,M., Mengus,G., Lavigne,A.C., Davidson,I. and Moras,D. (1998) Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell*, **94**, 239–249.
13. Anandapadamanaban,M., Andresen,C., Helander,S., Ohyama,Y., Siponen,M.I., Lundstrom,P., Kokubo,T., Ikura,M., Moche,M. and Sunnerhagen,M. (2013) High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. *Nat. Struct. Mol. Biol.*, **20**, 1008–1014.

14. Han,Y., Luo,J., Ranish,J. and Hahn,S. (2014) Architecture of the saccharomyces cerevisiae SAGA transcription coactivator complex. *EMBO J.*, **33**, 2534–2546.
15. Patel,A.B., Louder,R.K., Greber,B.J., Grunberg,S., Luo,J., Fang,J., Liu,Y., Ranish,J., Hahn,S. and Nogales,E. (2018) Structure of human TFIID and mechanism of TBP loading onto promoter DNA. *Science*, **362**, eaau8872.
16. Gupta,K., Watson,A.A., Baptista,T., Scheer,E., Chambers,A.L., Koehler,C., Zou,J., Obong-Ebong,I., Kandiah,E., Temblador,A. *et al.* (2017) Architecture of TAF11/TAF13/TBP complex suggests novel regulation properties of general transcription factor TFIID. *Elife*, **6**, e30395.
17. Huisinga,K.L. and Pugh,B.F. (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in saccharomyces cerevisiae. *Mol. Cell*, **13**, 573–585.
18. Baptista,T., Grunberg,S., Minoungou,N., Koster,M.J.E., Timmers,H.T.M., Hahn,S., Devys,D. and Tora,L. (2017) SAGA is a general cofactor for RNA polymerase II transcription. *Mol. Cell*, **68**, 130–143.
19. Warfield,L., Ramachandran,S., Baptista,T., Devys,D., Tora,L. and Hahn,S. (2017) Transcription of nearly all yeast RNA polymerase II-Transcribed genes is dependent on transcription factor TFIID. *Mol. Cell*, **68**, 118–129.
20. Donczew,R., Warfield,L., Pacheco,D., Erijman,A. and Hahn,S. (2020) Two roles for the yeast transcription coactivator SAGA and a set of genes redundantly regulated by TFIID and SAGA. *Elife*, **9**, e50109.
21. Wu,P.Y. and Winston,F. (2002) Analysis of Spt7 function in the saccharomyces cerevisiae SAGA coactivator complex. *Mol. Cell. Biol.*, **22**, 5367–5379.
22. Lee,K.K., Sardiu,M.E., Swanson,S.K., Gilmore,J.M., Torok,M., Grant,P.A., Florens,L., Workman,J.L. and Washburn,M.P. (2011) Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Mol. Syst. Biol.*, **7**, 503.
23. Yu,J., Madison,J.M., Mundlos,S., Winston,F. and Olsen,B.R. (1998) Characterization of a human homologue of the saccharomyces cerevisiae transcription factor spt3 (SUPT3H). *Genomics*, **53**, 90–96.
24. Herbst,D.A., Esbin,M.N., Louder,R.K., Dugast-Darzacq,C., Dailey,G.M., Fang,Q., Darzacq,X., Tjian,R. and Nogales,E. (2021) Structure of the human SAGA coactivator complex. *Nat. Struct. Mol. Biol.*, **28**, 989–996.
25. Stegeman,R., Spreacker,P.J., Swanson,S.K., Stephenson,R., Florens,L., Washburn,M.P. and Weake,V.M. (2016) The spliceosomal protein SF3B5 is a novel component of drosophila SAGA that functions in gene expression independent of splicing. *J. Mol. Biol.*, **428**, 3632–3649.
26. Umlauf,D., Bonnet,J., Waharte,F., Fournier,M., Stierle,M., Fischer,B., Brino,L., Devys,D. and Tora,L. (2013) The human TREX-2 complex is stably associated with the nuclear pore basket. *J. Cell Sci.*, **126**, 2656–2667.
27. Green,S., Issemann,I. and Sheer,E. (1988) A versatile in vivo and in vitro eukaryotic expression vector for protein engineering. *Nucleic Acids Res.*, **16**, 369.
28. Schneider,C.A., Rasband,W.S. and Eliceiri,K.W. (2012) NIH image to imagej: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
29. Rabani,M., Levin,J.Z., Fan,L., Adiconis,X., Raychowdhury,R., Garber,M., Gnirke,A., Nusbaum,C., Hacohen,N., Friedman,N. *et al.* (2011) Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat. Biotechnol.*, **29**, 436–442.
30. Radle,B., Rutkowski,A.J., Ruzsics,Z., Friedel,C.C., Koszinowski,U.H. and Dolken,L. (2013) Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. *J. Vis. Exp.*, **78**, 50195.
31. 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.
32. Pfaffl,M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, **29**, e45.
33. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, **17**, 10–12.
34. 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.
35. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq–a python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
36. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
37. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
38. Benjamini,Y. and ochberg,H, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B*, **57**, 289–300.
39. Danecek,P., Bonfield,J.K., Liddle,J., Marshall,J., Ohan,V., Pollard,M.O., Whitwham,A., Keane,T., McCarthy,S.A., Davies,R.M. *et al.* (2021) Twelve years of SAMtools and BCFtools. *Gigascience*, **10**, giab008.
40. Ramirez,F., Ryan,D.P., Gruning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dundar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
41. Dreos,R., Ambrosini,G., Perier,R.C. and Bucher,P. (2015) The eukaryotic promoter database: expansion of EPDnew and new promoter analysis tools. *Nucleic Acids Res.*, **43**, D92–D96.
42. Muratoglu,S., Georgieva,S., Papai,G., Scheer,E., Enunlu,I., Komonyi,O., Cserpan,I., Lebedeva,L., Nabirochkina,E., Udvardy,A. *et al.* (2003) Two different drosophila ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Mol. Cell. Biol.*, **23**, 306–321.
43. Zybailov,B., Mosley,A.L., Sardiu,M.E., Coleman,M.K., Florens,L. and Washburn,M.P. (2006) Statistical analysis of membrane proteome expression changes in saccharomyces cerevisiae. *J. Proteome Res.*, **5**, 2339–2347.
44. Zhang,Y., Wen,Z., Washburn,M.P. and Florens,L. (2010) Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal. Chem.*, **82**, 2272–2281.
45. Basehoar,A.D., Zanton,S.J. and Pugh,B.F. (2004) Identification and distinct regulation of yeast TATA box-containing genes. *Cell*, **116**, 699–709.
46. de Jonge,W.J., O'Duibhir,E., Lijnzaad,P., van Leenen,D., Groot Koerkamp,M.J., Kemmeren,P. and Holstege,F.C. (2017) Molecular mechanisms that distinguish TFIID housekeeping from regulatable SAGA promoters. *EMBO J.*, **36**, 274–290.
47. Wang,F., El-Saafin,F., Ye,T., Stierle,M., Negroni,L., Durik,M., Fischer,V., Devys,D., Vincent,S.D. and Tora,L. (2021) Histone H2Bub1 deubiquitylation is essential for mouse development, but does not regulate global RNA polymerase II transcription. *Cell Death Differ.*, **28**, 2385–2403.
48. Eisenmann,D.M., Chapon,C., Roberts,S.M., Dollard,C. and Winston,F. (1994) The saccharomyces cerevisiae SPT8 gene encodes a very acidic protein that is functionally related to SPT3 and TATA-binding protein. *Genetics*, **137**, 647–657.
49. Belotserkovskaya,R., Sterner,D.E., Deng,M., Sayre,M.H., Lieberman,P.M. and Berger,S.L. (2000) Inhibition of TATA-binding protein function by SAGA subunits Spt3 and Spt8 at Gcn4-activated promoters. *Mol. Cell. Biol.*, **20**, 634–647.
50. Perez-Garcia,V., Fineberg,E., Wilson,R., Murray,A., Mazzeo,C.I., Tudor,C., Sienerth,A., White,J.K., Tuck,E., Ryder,E.J. *et al.* (2018) Placentation defects are highly prevalent in embryonic lethal mouse mutants. *Nature*, **555**, 463–468.
51. Bu,P., Evrard,Y.A., Lozano,G. and Dent,S.Y. (2007) Loss of Gcn5 acetyltransferase activity leads to neural tube closure defects and exencephaly in mouse embryos. *Mol. Cell. Biol.*, **27**, 3405–3416.
52. Zohn,I.E., Li,Y., Skolnik,E.Y., Anderson,K.V., Han,J. and Niswander,L. (2006) p38 and a p38-interacting protein are critical for downregulation of E-cadherin during mouse gastrulation. *Cell*, **125**, 957–969.
53. Yamauchi,T., Yamauchi,J., Kuwata,T., Tamura,T., Yamashita,T., Bae,N., Westphal,H., Ozato,K. and Nakatani,Y. (2000) Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11303–11306.

54. Atanassov,B.S., Mohan,R.D., Lan,X., Kuang,X., Lu,Y., Lin,K., McIvor,E., Li,W., Zhang,Y., Florens,L. *et al.* (2016) ATXN7L3 and ENY2 coordinate activity of multiple H2B deubiquitinases important for cellular proliferation and tumor growth. *Mol. Cell*, **62**, 558–571.
55. Koutelou,E., Wang,L., Schibler,A.C., Chao,H.P., Kuang,X., Lin,K., Lu,Y., Shen,J., Jeter,C.R., Salinger,A. *et al.* (2019) USP22 controls multiple signaling pathways that are essential for vasculature formation in the mouse placenta. *Development*, **146**, dev174037.
56. Antonova,S.V., Boeren,J., Timmers,H.T.M. and Snel,B. (2019) Epigenetics and transcription regulation during eukaryotic diversification: the saga of TFIID. *Genes Dev.*, **33**, 888–902.
57. Edgar,R., Domrachev,M. and Lash,A.E. (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, **30**, 207–210.
58. Perez-Riverol,Y., Csordas,A., Bai,J., Bernal-Llinares,M., Hewapathirana,S., Kundu,D.J., Inuganti,A., Griss,J., Mayer,G., Eisenacher,M. *et al.* (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.*, **47**, D442–D450.
59. Pradhan,S.K., Su,T., Yen,L., Jacquet,K., Huang,C., Cote,J., Kurdistani,S.K. and Carey,M.F. (2016) EP400 deposits H3.3 into promoters and enhancers during gene activation. *Mol. Cell*, **61**, 27–38.
60. Kang,H., Shokhirev,M.N., Xu,Z., Chandran,S., Dixon,J.R. and Hetzer,M.W. (2020) Dynamic regulation of histone modifications and long-range chromosomal interactions during postmitotic transcriptional reactivation. *Genes Dev.*, **34**, 913–930.
61. Gertz,J., Savic,D., Varley,K.E., Partridge,E.C., Safi,A., Jain,P., Cooper,G.M., Reddy,T.E., Crawford,G.E. and Myers,R.M. (2013) Distinct properties of cell-type-specific and shared transcription factor binding sites. *Mol. Cell*, **52**, 25–36.
62. Encode Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57–74.