12444–12461 *Nucleic Acids Research, 2022, Vol. 50, No. 21* *Published online 1 December 2022*
https://doi.org/10.1093/nar/gkac1055

SUMO conjugation regulates the activity of the Integrator complex

Laureano Bragado[1,2], Melina Magalnik[1,2], Pablo Mammi[1,2], Agustín Romero[1,2], Nicolás Gaioli[1,2], Berta Pozzi[1,2] and Anabella Srebrow[1,2,*]

[1]Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Fisiología, Biología Molecular y Celular, Buenos Aires, Argentina and [2]CONICET-Universidad de Buenos Aires, Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE), Buenos Aires, Argentina

Received February 07, 2022; Revised October 13, 2022; Editorial Decision October 17, 2022; Accepted November 29, 2022

ABSTRACT

RNA polymerase II (RNAPII) transcribes small nuclear RNA (snRNA) genes in close proximity to Cajal bodies, subnuclear compartments that depend on the SUMO isopeptidase USPL1 for their assembly. We show here that overexpression of USPL1 as well as of another nuclear SUMO isopeptidase, SENP6, alters snRNA 3′-end cleavage, a process carried out by the Integrator complex. Beyond its role in snRNA biogenesis, this complex is responsible for regulating the expression of different RNAPII transcripts. While several subunits of the complex are SUMO conjugation substrates, we found that the SUMOylation of the INTS11 subunit is regulated by USPL1 and SENP6. We defined Lys381, Lys462 and Lys475 as bona fide SUMO attachment sites on INTS11 and observed that SUMOylation of this protein modulates its subcellular localization and is required for Integrator activity. Moreover, while an INTS11 SUMOylation-deficient mutant is still capable of interacting with INTS4 and INTS9, its interaction with other subunits of the complex is affected. These findings point to a regulatory role for SUMO conjugation on Integrator activity and suggest the involvement of INTS11 SUMOylation in the assembly of the complex. Furthermore, this work adds Integrator-dependent RNA processing to the growing list of cellular processes regulated by SUMO conjugation.

INTRODUCTION

In eukaryotic cells, protein-coding genes are transcribed by RNAPII, generating pre-mRNAs that contain exons and introns. The splicing process removes introns and joins exons, giving rise to mature mRNAs, and is catalyzed by the spliceosome. This highly complex and dynamic macromolecular machine is composed of five small nuclear ribonucleoprotein particles (snRNPs) termed U1, U2, U4, U5 and U6, in the case of the so-called 'major spliceosome', and many non-snRNP splicing factors. Each snRNP comprises one small nuclear RNA (snRNA) associated with a specific set of proteins. While the function of snRNAs and the assembly of snRNPs during the splicing process are well characterized (1), the regulation of snRNA gene expression is still poorly understood.

Most snRNAs are transcribed by RNAPII, including U1, U2, U4 and U5. Unlike pre-mRNAs, snRNAs are intronless and non-polyadenylated. 3′-end processing of the nascent snRNA transcript depends on a sequence termed the 3′ box (2), which is recognized by the Integrator complex (3). In the absence of this complex, nascent snRNAs are cleaved and polyadenylated by the cleavage and polyadenylation machinery (CPA) that operates on pre-mRNAs (4).

The Integrator complex contains 14 subunits (INTS1–INTS14), where INTS9 and INTS11 have been identified as homologs of cleavage and polyadenylation specificity factors (CPSFs) 100 and 73, respectively (3), members of the CPA. Although INTS11 is responsible for the endonucleolytic activity, it has been shown that INTS4, INTS9 and INTS11 associate into a ternary complex, which constitutes the Integrator catalytic core. Furthermore, the structure of this core resembles that of the heterotrimeric complex CPSF73/100/Symplekin (5, 6).

In recent years, several reports have linked the activity of the Integrator complex with the biogenesis of a broader set of substrates, including enhancer RNAs (eRNAs), viral microRNAs and diverse non-polyadenylated long non-coding RNAs (lncRNAs) (7–12). In addition, this complex is involved in the fine regulation of promoter-proximal pausing of RNAPII on protein-coding genes (10,13–19). Moreover, the Integrator complex is necessary for the transcriptional response downstream of growth factor signaling, and INTS11 is phosphorylated by ERK1/2 (20). However, the relevance of post-translational modifications

*To whom correspondence should be addressed. Email: asrebrow@fbmc.fcen.uba.ar

for the activity of INTS11, or other subunits of the Integrator complex, has not yet been addressed.

SUMO conjugation (aka SUMOylation) is a rapid, reversible post-translational modification (PTM) consisting of the covalent attachment of a small ubiquitin-related modifier (SUMO) peptide to a lysine residue in the target protein. There are three well-characterized functional SUMO isoforms encoded by the human genome (SUMO1, 2 and 3), which modify distinct but overlapping sets of substrates. While it is still unclear whether SUMO4 is conjugated to cellular proteins, SUMO5 has been more recently identified as a novel, primate- and tissue-specific SUMO variant (21–25). Like ubiquitin, SUMO is conjugated to its targets by an isopeptide bond between its C-terminal glycine and the ε-NH_2 group of the target lysine residue. In general, SUMOylation substrates contain the consensus motif ΨKxD/E, where Ψ is a large, hydrophobic amino acid, K is the target lysine, x is any amino acid, D is aspartic acid and E is glutamic acid. However, many SUMOylated proteins deviate from this consensus sequence or even lack one (26, 27). The steps of the SUMO conjugation pathway resemble those of the ubiquitin pathway. Before being conjugated, SUMO is cleaved by specific proteases (SENPs), exposing its C-terminal Gly–Gly motif (28). Subsequently, mature SUMO is activated by the SUMO-activating enzyme E1, the heterodimer 'AOS1–UBA2', in an ATP-dependent manner and then transferred to the catalytic Cys residue of 'Ubc9', the SUMO-specific E2 conjugating enzyme. Finally, an isopeptidic bond is formed between the C-terminal Gly residue of SUMO and a Lys residue in the target protein. This step is generally aided by SUMO E3 ligases and, among those characterized so far, few display substrate specificity while others display SUMO isoform preferences (22). SUMOylated proteins are substrates of SENPs (SENP1, 2, 3, 5, 6 and 7) and other isopeptidases, which display differential subcellular localization and are able to deconjugate SUMO, ensuring the reversibility and dynamic nature of the process. Most frequently, SUMO conjugation regulates intra- or intermolecular interactions, altering either the conformation of the modified protein or the recruitment of its partners (22). In several cases, SUMOylation fosters new associations by non-covalent interaction of conjugated SUMO with proteins harboring SUMO interaction motifs (SIMs). The establishment of SUMO–SIM interactions exerts a variety of effects, ranging from intramolecular structural rearrangements, as reported for thymine DNA glycosylase, to the assembly and stabilization of multiprotein complexes, as described for PML nuclear bodies (29). In addition, SUMOylation can also interfere with protein stability by triggering ubiquitylation of poly-SUMO-modified proteins through the recruitment of SUMO-targeted ubiquitin ligases (STUbLs) (30).

The biological relevance of protein SUMOylation is clearly demonstrated by the fact that inactivation of SUMO in *Saccharomyces cerevisiae* or of the unique E2 SUMO-conjugating enzyme Ubc9 in mice is lethal (31, 32). Accordingly, multiple studies have shown that SUMOylation regulates a wide range of cellular functions, including intracellular transport, maintenance of genome integrity, formation of nuclear subdomains (21) and also some aspects of rRNA or small nucleolar RNA (snoRNA) metabolism (33–35). Furthermore, SUMO conjugation affects not only the stability, localization and activity of transcriptional regulators, but also the activity of DNA and histone modifiers, leading to changes in chromatin structure and hence gene expression (36).

Proteomic approaches have revealed that RNA-related proteins are predominant among SUMO substrates (37). In addition, Ubc9 has been found to reside in nuclear speckles (38), which are thought to coordinate splicing and gene expression, as they contain not only splicing factors, but also other proteins involved in mRNA metabolism, such as transcription factors, RNAPII subunits, cleavage and polyadenylation factors, and RNA export proteins (39). Furthermore, SUMO conjugation regulates different aspects of mRNA metabolism, such as pre-mRNA splicing, pre-mRNA 3′-end processing and RNA editing, by modifying the function of spliceosomal proteins, poly(A) polymerase, Symplekin and CPSF-73, and ADAR1 respectively (40–42).

RNAPII transcribes snRNA genes in close proximity to Cajal bodies (CBs) (43–47), and coilin, a protein known to function as a scaffold for CBs, has been shown to directly interact with snRNAs (48). The Little Elongation Complex (LEC), necessary for elongation of RNAPII during transcription of snRNAs, co-localizes with coilin at CBs (49). In addition, the more recently described SUMO protease USPL1 has been found to localize with coilin within these nuclear bodies and to interact with the LEC. Although USPL1 knockdown causes diminution of CB formation and a reduction of snRNA levels, the relevant substrates of USPL1 have not been identified and the molecular mechanisms underlying these effects are still unclear (50, 51).

In this study, we have investigated the possible involvement of SUMO conjugation in snRNA biogenesis. We verified several subunits of the Integrator complex as bona fide SUMO conjugation substrates in cultured human cells. In particular, we focused on the catalytic subunit of this complex, INTS11, mapping the target residues for this modification and demonstrating that its SUMOylation levels are regulated not only by the SUMO isopeptidase USLP1 but also by SENP6. In addition, we have shown that these two SUMO proteases affect snRNA 3′-end processing. By generating a SUMOylation-deficient mutant of INTS11, we found that SUMO conjugation to this subunit is crucial not only for proper snRNA biogenesis but also for the expression of different non-coding RNAs [eRNAs, promoter upstream transcripts (PROMPTs) and downstream-of-gene-containing transcripts (DoGs)] recently found to be regulated by the Integrator complex. We propose that while the assembly of the catalytic core is not dependent on INTS11 SUMOylation, this modification is required for the proper subcellular localization of this subunit as well as for its interaction with components of the Integrator complex other than INTS4 and INTS9. These results reveal a novel regulatory mechanism for Integrator activity and add Integrator-dependent RNA

processing to the list of cellular processes regulated by SUMO conjugation.

MATERIALS AND METHODS

Cell lines

HeLa (human cervix adenocarcinoma cell line, ATCC CCL-2) and HEK 293T cells (human embryonic kidney cell line, ATCC CRL-1573) were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 100 U/ml penicillin and 100 μg/ml streptomycin.

DNA plasmids

The expression vector for FLAG-INTS11 was a generous gift from Dr Ramin Shiekhattar (University of Miami Health Systems, USA), expression vectors for FLAG-INTS9, FLAG-INTS10, HIS-MYC-INTS9 and HIS-MYC-INTS11 were a generous gift from Dr Shona Murphy (University of Oxford, UK), expression vectors for HA-INTS4, HA-INTS13 and HA-INTS14 were a generous gift from Dr Stefanie Jonas (Institute of Molecular Biology and Biophysics, ETH Zurich), expression vectors for HA-USPL1 and HA-USPL1 C236S were a generous gift from Dr Frauke Melchior (ZMBH, Heidelberg, Germany) and Dr Angus Lamond (Dundee, UK), expression vectors for flag-SENP and the flag-SENP6 catalytically inactive mutant were generous gifts from Dr Ron Hay (Dundee, UK) and Dr Alfred Vertegaal (Leiden University, the Netherlands), and U7 snRNA–green fluorescent protein (GFP) reporter was a generous gift from Dr Omar Abdel-Wahab (Memorial Sloan Kettering Cancer Center, NY, USA).

The expression vector for GFP–FLAG-INTS11 was generated in our laboratory. Briefly, INTS11 cDNA was amplified from FLAG-INTS11 plasmid with the following primers: forward, AAAGAATTCTGACTACAAAGAC GAT; and reverse, AAAGGATCCTCTAGAGTCGACT GGT. The polymerase chain reaction (PCR) product was digested with EcoRI and BamHI restriction enzymes and subcloned into EcoRI/BamHI-digested pEGFP-C1 plasmid (Addgene #2487).

Transfection

Transfection of plasmid DNA and siRNA was carried out with Lipofectamine 2000 (Thermo Fisher) according to the manufacturer's instructions.

Western blot assay and antibodies

Protein samples were resolved by sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and transferred to nitrocellulose membranes (BioRad). Membranes were blocked and then incubated with the corresponding primary antibody. After washing, membranes were incubated with IRDye® 800CW (LI-COR Biosciences) secondary antibody. Bound antibody was detected using an Odyssey imaging system (LI-COR Biosciences). Western blots (WBs) were performed at least three times, and representative images are shown in each case. The antibodies used were: mouse monoclonal anti-β-actin C4 (Santa Cruz Biotechnology), polyclonal rabbit anti-β-tubulin H-235 (Santa Cruz Biotechnology), monoclonal mouse anti-GFP B2 (Santa Cruz Biotechnology), monoclonal mouse anti-c-Myc 9E10 (Santa Cruz Biotechnology), monoclonal mouse anti-FLAG-M2 (Sigma-Aldrich), monoclonal mouse anti-HA MMS-101R (Covance), polyclonal rabbit anti-INTS11 ab75276 (Abcam), polyclonal rabbit anti-INTS11 A301-274A (Bethyl) and polyclonal rabbit anti-INTS9 13945 (Cell Signaling Technology)

Site-directed mutagenesis

Mutagenesis was performed by the DpnI method, based on Stratagene's QuickChange protocol. The primers used to mutate putative SUMO sites from Lys to Arg are listed in Supplementary Table S2. Mutations were always verified by sequencing.

Purification of HIS-SUMO conjugated proteins

HEK 293T cells were transfected in 35 mm culture wells with the indicated plasmids. After 48 h, 6×HIS-SUMO2 conjugates were purified under denaturing conditions using Ni-NTA–agarose beads according to the manufacturer's instructions (Qiagen). Briefly, transfected cells were harvested in ice-cold phosphate-buffered saline (PBS) plus 100 mM iodoacetamide (IAA). An aliquot was taken as input and the remaining cells were lysed in 6 M guanidinium-HCl containing 100 mM Na_2HPO_4/NaH_2PO_4, 10 mM Tris–HCl pH 8.0, 5 mM imidazole and 10 mM IAA. Samples were sonicated to reduce viscosity and centrifuged for 20 min at 12 000 *g*. Afterwards, proteins in the supernatants were purified using Ni-NTA beads (Qiagen) according to (52). Samples were subsequently washed with wash buffer I (8 M urea, 10 mM Tris–HCl, 100 mMNa_2HPO_4/NaH_2PO_4, 5 mM imidazole, 10 mM IAA, pH 8.0), wash buffer II (8 M urea, 10 mM Tris–HCl, 100 mM Na_2HPO_4/NaH_2PO_4, 0.2% Triton X-100, 5 mM imidazole, 10 mM IAA, pH 6.3) and wash buffer III (8 M urea, 10 mM Tris–HCl, 100 mM Na_2HPO_4/NaH_2PO_4, 0.1% Triton X-100, 5 mM imidazole). Samples were eluted in 2× Laemmli sample buffer containing 300 mM imidazole (pH 6.3) for 3 min at 95°C.

Quantitative reverse transcription–PCR (RT–qPCR) for cellular RNAs

Total cellular RNA was isolated by using 250 μl of Tri-Reagent (MRC) and measured with a NanoDrop 1000 spectrophotometer (Thermo Scientific). RNA was treated with DNase RQ1 (Promega) following the manufacturer's instructions. Then, 1 μg of each RNA sample was reverse transcribed to cDNA with random decaoligonucleotide or oligo-dT primer mix using MMLV Reverse Transcriptase (Invitrogen). Quantitative PCRs (qPCRs) were performed using SYBR Green dye, 1/20 dilution of cDNA sample and Taq DNA polymerase (Invitrogen) in a Mastercycler®

ep realplex PCR device (Eppendorf). The annealing temperature was 60°C and the elongation time at 72°C was 30 s. Relative RNA abundance from cDNA samples and no-reverse transcription controls was estimated employing internal standard curves with a PCR efficiency of 100 ± 10% for each set of primers, in each experiment. Realplex qPCR software was used to analyze the data. The specific primers used are listed in Supplementary Table S2.

CRISPR interference (CRISPRi)-mediated depletion

For CRISPRi, 50000 HEK 293T cells were seeded in 24-well plates. After 24 h, 400 ng of dCas9-KRAB expression vector (Addgene #60954), 170 ng of lentiGuide puro (Addgene #52963) and 30 ng of expression vector for INTS11 were transfected. Cells were harvested 72 h post-transfection. Sequences of single guide (sgRNAs) are listed in Supplementary Table S2

Chromatin immunoprecipitation (ChIP)

Cells were cross-linked with 1% (v/v) formaldehyde (final concentration), washed twice with cold PBS, scraped, collected and centrifuged. Cell pellets were resuspended in 2 ml of SDS lysis buffer (1% w/v SDS, 10 mM EDTA, 50 mM Tris–HCl, pH 8.1) containing Complete Protease Inhibitor Cocktail (Roche) and incubated for 10 min on ice. Cell extracts were sonicated with a Branson sonicator W-450 D at 30% amplitude with 15 bursts of 10 s each, resulting in ∼500 nt chromatin fragments, and then centrifuged for 10 min at 12 000 *g*. A 50 μl sample of the supernatant was saved as input DNA and the remainder was diluted 1:10 in ChIP dilution buffer (0.01% w/v SDS, 1.1% v/v Triton X-100, 1.2 mM EDTA, 16.7 mM Tris–HCl pH 8.1, 167 mM NaCl) containing protease inhibitors. The chromatin solution was pre-cleared at 4°C with Protein G Dynabeads® for 1 h before incubating overnight at 4°C with antibodies. Complexes were incubated with Dynabeads® Protein G beads for 1 h at 4°C. Beads were washed by rocking for 4 min, once in each of the following buffers: low salt immune complex wash buffer (0.1% w/v SDS, 1% v/v Triton X-100, 2 mM EDTA, 20 mM Tris–HCl, pH 8.1, 150 mM NaCl), high-salt immune complex wash buffer (same as low salt buffer, except with 500 mM NaCl) and LiCl immune complex wash buffer (0.25 M LiCl, 1% v/v NP-40, 1% w/v deoxycholic acid, 1 mM EDTA, 10 mM Tris–HCl, pH 8.1), and then twice in TE (10 mM Tris–HCl, 1 mM EDTA). Bound complexes were eluted in 1% (w/v) SDS and 50 mM $NaHCO_3$, and cross-linking was reversed by incubating overnight at 65°C. Samples were digested with proteinase K for 1 h at 45°C and DNA was extracted using a Qiagen PCR purification kit. DNA retrieved by ChIP was analyzed by quantitative real-time PCR with primers listed in Supplementary Table S2. Datasets were normalized to ChIP input values.

Immunoprecipitation of the Integrator complex

Cells were harvested and lysed in RIPA buffer [50 mM Tris–HCl pH 7.5, 1% (v/v) NP-40, 0.5% (w/v) sodium deoxycholate, 0.05% (w/v) SDS, 1 mM EDTA, 150 mM NaCl] containing Complete Protease Inhibitor (Roche). Extracts were sonicated at high amplitude with three 10 s bursts, and insoluble material was pelleted. Anti-FLAG M2 antibodies were added to the supernatant and incubated overnight. Then, complexes were incubated with Protein G Dynabeads for 1 h and washed three times in wash buffer [50 mM Tris–HCl (pH 7.5), 0.1% (v/v) NP-40, 1 mM EDTA, 125 mM NaCl]. For WB analysis, beads were resuspended in 2× Laemmli sample buffer. For GFP and GFP–INTS11 immunoprecipitations, cells were lysed in RIPA buffer (the same as above) containing Complete Protease Inhibitor (Roche). Extracts were sonicated at high amplitude with three 10 s bursts, and insoluble material was pelleted. Supernatants were incubated for 1 h with GFP-Trap coupled to magnetic agarose beads (Chromotek). Then, complexes were washed three times in wash buffer (the same as above). For WB analysis, beads were resuspended in 2× Laemmli sample buffer.

Microscopy confocal assays

HeLa cells were seeded into 24-well plates containing glass coverslips. Twenty-four hours later, cells were transfected with the indicated plasmids. After 24 h, coverslips were collected, and cells were fixed with paraformaldehyde (PFA) 4% in PBS pH 7.4 at room temperature for 10 min. PFA-fixed cells were then permeabilized with 1% Triton X-100 for 5 min at room temperature. Coverslips were incubated with RNase A for 30 min at 37°C. After three washes with PBS, incubation with propidium iodide was performed for 5 min followed by washing four times with PBS. Coverslips were mounted with a drop of mounting medium (Vectashield) and images were obtained with a Zeiss LSM900 confocal microscope (with a ×40 objective). Fluorescence intensity analysis of images was performed with Cell Profiler software (v 3.1.5) after generating a suitable pipeline for this purpose.

Statistics

Typically, three or four independent experiments in triplicate repeats were conducted for each condition examined. Average values are shown with standard deviation and *P*-values, measured with a paired two-tailed *t*-test or analysis of variance (ANOVA), as indicated for each case. Significant *P*-values are indicated by asterisks above the graphs (***P <0.001; **P <0.01; *P <0.05).

RESULTS

Overexpression of USPL1 affects the levels of nascent snRNAs

USPL1 is an essential SUMO isopeptidase that localizes to CBs and its knockdown leads to disruption of these subnuclear structures (50). Moreover, USPL1 has been shown to interact with components of the RNAPII-associated LEC, and knockdown of USPL1 leads to reduced RNAPII-mediated snRNA gene transcription, diminished production of snRNPs and altered pre-mRNA splicing (51). However, USPL1 substrates implicated in these effects have not yet been identified.

To further study the involvement of USPL1 in snRNA biogenesis, we overexpressed HA-USPL1 together with a U7–GFP readthrough reporter construct (53) in human cultured cells (Figure 1A). Under basal conditions, marginal EGFP expression from the reporter is observed due to 3′-end cleavage of U7 snRNA gene transcripts, which impairs expression of the downstream EGFP open reading frame (ORF). However, when snRNA 3′-end processing is perturbed, e.g. by tampering with the expression of Integrator complex subunits, RNAPII transcribes beyond the U7 cleavage site (3′ box) and recognizes a canonical cleavage and polyadenylation signal (pA) downstream of the EGFP ORF (Figure 1A) (53). Overexpression of HA-USPL1 markedly increases GFP expression levels compared with the levels observed upon co-transfection of the reporter with an expression vector for a catalytically inactive variant of the SUMO protease (HA-USPL1 C236S) or with empty plasmid (pcDNA) (Figure 1B). We then analyzed total and uncleaved levels of endogenous, RNAPII-dependent snRNA transcripts, by RT–qPCR (Figure 1C). The levels of uncleaved U1, U2, U4 and U5 snRNAs are increased upon overexpression of wild-type (WT) USPL1 but not upon overexpression of USPL1 C236S (Figure 1D). The SUMO peptidase activity of USPL1 variants was validated by co-expressing them with HA-SUMO2 followed by WB analysis of SUMO-conjugated proteins (Supplementary Figure S1A, B). To corroborate that the observed effects are indeed related to protein modification by SUMO conjugation, cultured cells were transfected with an expression vector for a different SUMO isopeptidase, SENP6. This protease is known to display a nucleoplasmic distribution (54) and decreases global SUMOylation levels even further than USPL1 (Supplementary Figure S1C). In accordance with the results obtained with WT USLP1, the levels of GFP derived from the U7–GFP reporter (Supplementary Figure S1D) as well as those of uncleaved U1, U2, U4 and U5 snRNAs are increased upon overexpression of SENP6 (Figure 1E), suggesting that SUMOylation is involved in snRNA maturation. Similar results were obtained when levels of uncleaved snRNAs were normalized against an RNA polymerase III-dependent transcript (U6 snRNA) (Supplementary Figure S1E).

Previous studies have shown that in the absence of proper snRNA 3′-end cleavage, snRNAs become polyadenylated (4,9). Thus, the levels of polyadenylated snRNAs can be considered as indicative of 3′-end processing failure. Consistent with our above-mentioned observations, overexpression of HA-USPL1 also leads to increased levels of polyadenylated snRNAs (Supplementary Figure S1F). In contrast, the levels of other polyadenylated and RNAPII-dependent transcripts are not altered by USPL1 overexpression (Supplementary Figures S1G).

Altogether, these results demonstrate that a proper balance of SUMO conjugation and de-conjugation cycles is critical for efficient snRNA 3′-end processing. Whether this is due to a direct effect of USPL1 or SENP6 on components of the 3′-end processing machinery or to an overload of this machinery by an increase in snRNA transcription remains to be elucidated.

Integrator subunits are modified by SUMO in cultured cells

Having observed that overexpression of two different SUMO isopeptidases alters snRNA biogenesis, we analyzed human SUMO proteome datasets to further explore the possible effect of SUMO conjugation on the activity of snRNA 3′-end processing factors. Different proteomic studies revealed several Integrator complex subunits as SUMO2/3 conjugation targets (Supplementary Table S1) (27).

To validate their modification by SUMO conjugation in our culture conditions, cell extracts from HEK 293T cells co-transfected with expression vectors for 6×His-SUMO2 and tagged Integrator subunits were subjected to nickel affinity purification, to enrich for SUMOylated proteins. Pulled-down proteins were analyzed by WB with specific antibodies against the different fused tags. As shown in Figure 2, we were able to detect SUMO conjugation to HA-INTS4 (Figure 2A), FLAG-INTS9 (Figure 2B), FLAG-INTS10 (Figure 2C), FLAG-INTS11 (Figure 2D), HA-INTS13 (Figure 2E) and HA-INTS14 (Figure 2F). It has been proposed that INTS4, INTS9 and INTS11 form the Integrator 'cleavage module' responsible for 3′-end processing of target RNAs. INTS11 contains the active site that is responsible for Integrator activity, but its association with INTS4 and INTS9 is crucial for catalysis (5). After identifying that the three subunits of the Integrator cleavage module are bona fide SUMOylation substrates, we investigated the possible regulation of this modification by USPL1. To do so, we performed a nickel affinity purification strategy followed by WB, as indicated above, but with total lysates derived from cells also overexpressing HA-USPL1. Remarkably, the levels of INTS11 SUMOylation, but not those of INTS4 and INTS9, are modulated by USPL1 (Figure 2G–I; Supplementary Figure S2A). In agreement with the aforementioned results, the levels of INTS11 SUMOylation are also regulated by SENP6 (Supplementary Figure S2B). These results are all consistent with the above-proposed involvement of SUMO conjugation in snRNA maturation.

INTS11 is SUMOylated at lysines 381, 462 and 475

Considering the observed regulation of snRNA processing as well as the levels of INTS11 SUMOylation by the two SUMO isopeptidases, together with the described key role of INTS11 in snRNA endonucleolytic cleavage, we focused on the modification of INTS11 by SUMO conjugation.

Initially, we took advantage of a bioinformatics prediction program (Sumo sp 2.0 www.sumosp.biocuckoo.org) to define potential canonical SUMO acceptor sites within this protein. This search revealed a putative SUMO conjugation site at Lys462 of human INTS11, within the ΨKxD/E SUMOylation consensus motif. Site-directed mutagenesis was performed in order to replace this Lys by Arg, a widely used strategy to map SUMOylation target residues. SUMOylation levels of the FLAG-INTS11 mutant were analyzed by WB after nickel affinity purification of SUMOylated proteins from cells co-expressing 6×His-SUMO2 and the different

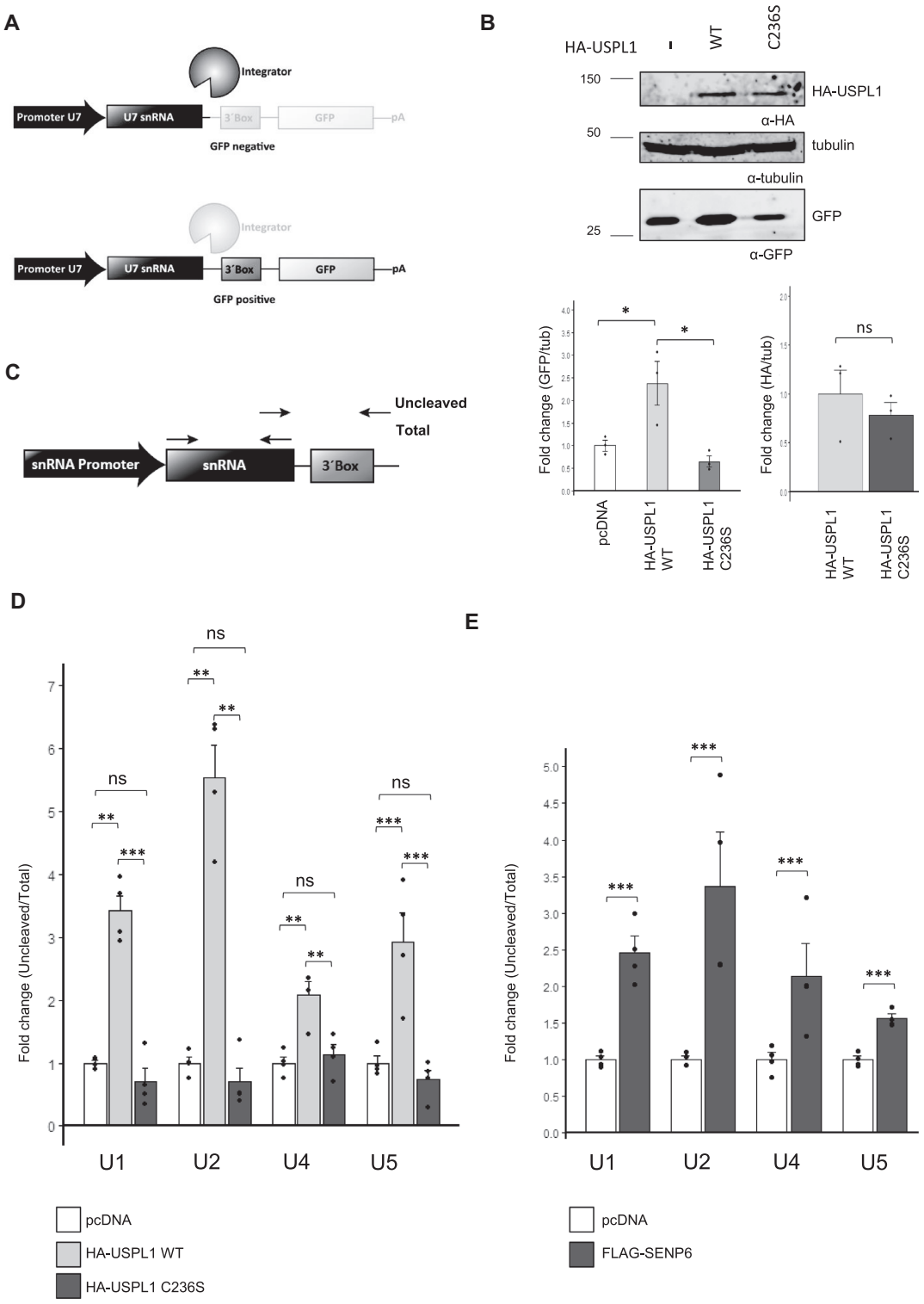

Figure 1. Overexpression of USPL1 and SENP6 affects nascent snRNA levels. (**A**) Schematic of the U7–GFP reporter construct that produces GFP upon abrogation of Integrator complex activity. (**B**) HEK 293T cells were co-transfected with the U7–GFP reporter mentioned in (A) and expression vectors for WT or C236S (catalytic mutant) HA-USPL1. Upper panel: levels of GFP, HA-USPL1 variants and tubulin were analyzed by WB. Lower panel: quantification corresponding to three independent experiments was performed with Image Studio Software (LI-COR Biosciences). Average values are shown with the standard error (SE) and *P*-values, determined using a *t*-test. Significant *P*-values are indicated by the asterisks above the graphs ($n = 3$, $^{*}P < 0.05$). (**C**) Schematic of human snRNA genes showing the position of the primers used to amplify total and uncleaved snRNAs by RT–qPCR. (**D**, **E**) RT–qPCR analyses of RNA samples corresponding to cell culture conditions indicated by different colors, using primer pairs targeting the snRNA genes indicated below each graph. Reverse transcription was performed with random decamer primers. Average values are shown with the SE and *P*-values, determined using one-way ANOVA and Tukey post-hoc test (D) or a paired two-tailed *t*-test (E). Significant *P*-values are indicated by the asterisks above the graphs ($n = 4$, $^{***}P < 0.001$; $^{**}P < 0.01$; $^{*}P < 0.05$).

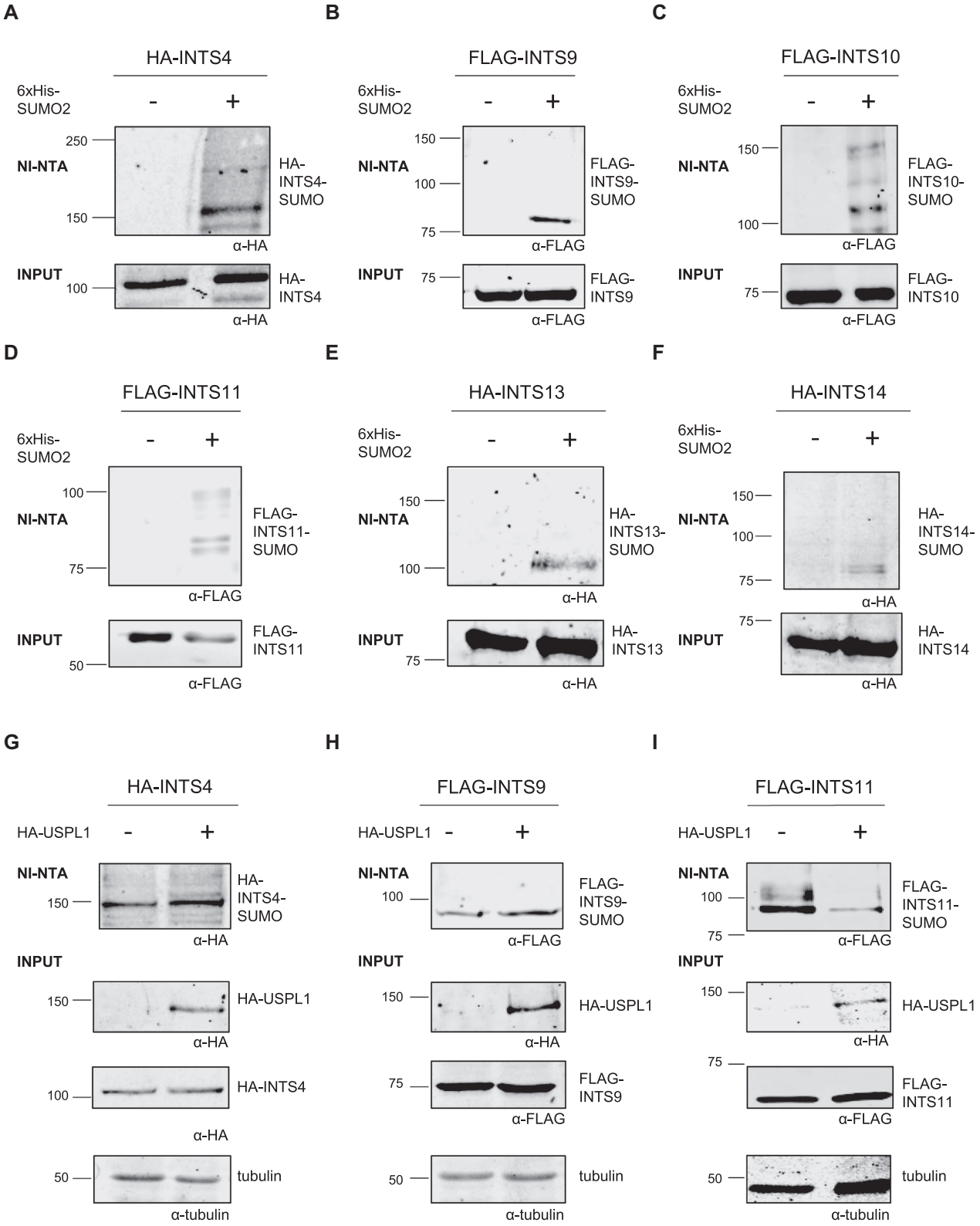

Figure 2. Integrator subunits are modified by SUMO conjugation in human cultured cells. (**A**–**I**) HEK 293T cells were transfected with expression vectors indicated at the top of each panel. After 48 h, cells were lysed, and whole-cell lysates were subjected to nickel affinity chromatography (Ni-NTA). Aliquots of cell lysates (input) and eluates (Ni-NTA) were analyzed by WB with the antibodies indicated below each panel.

FLAG-INTS11 versions. Mutation of Lys462 provoked only a slight decrease in INTS11 SUMO conjugation levels, compared with the WT version of this protein (Figure 3A, D). In addition to this canonical SUMO consensus sequence, we also identified a Lys residue within an inverted consensus motif D/ExKΨ, Lys381. A double mutant was then generated, and the combined replacement of Lys381 and Lys462 by Arg residues led to a marked reduction in INTS11 SUMOylation levels (Figure 3B, D). Site-specific SUMO proteomics data have identified Lys115, 289, 369, 462 and 475 as high score SUMO acceptor sites within INTS11 (27). Based on this information, several INTS11 triple mutants were generated in order to further reduce INTS11 SUMOylation. Only the combined replacement of Lys381, 462 and 475 by Arg residues (hereafter referred as the 'INTS11 3KR' mutant) led to a robust and drastic reduction of SUMO conjugation to INTS11 (Figure 3C–E). Remarkably, the single mutation of Lys475 only partially reduces INTS11 SUMOylation (Supplementary Figure S2C). It is worth noting that neither the FLAG-INTS11 catalytic-inactive mutant (E203Q) nor the GFP–FLAG-INTS11 fusion protein display any perturbation of their SUMOylation levels (Supplementary Figure S2D, E). The three Lys residues, and the SUMO consensus motif, are all conserved across different vertebrate species and some of them are also conserved in *Drosophila* and *Caenorhabditis*, suggesting that SUMO conjugation of these proteins may also be conserved (Supplementary Figure S2F).

Considering that the triple mutant 'INTS11 3KR' showed the lowest level of INTS11 SUMOylation in our experimental setting, it was used to further analyze the consequences of SUMO conjugation on different INST11-mediated processes.

Lack of INTS11 SUMOylation affects Integrator activity in cultured cells

To explore whether INTS11 SUMOylation indeed has any relevance for Integrator activity within cultured cells, we analyzed the levels of uncleaved as well as polyadenylated snRNA transcripts, indicative of 3′-end processing efficiency. For this purpose, endogenous INTS11 was knocked-down by co-transfecting an expression vector for dCAS9-KRAB with INTS11 promoter-specific guides (Figure 4A). This INTS11 depletion was then complemented either by the overexpression of a wild-type GFP–FLAG-INTS11 (WT) or the different INTS11 mutants, from a heterologous promoter (55). Levels of endogenous and transfected GFP–FLAG-INTS11 proteins were assessed by WB with an anti-INTS11 or anti-GFP antibody (Figure 4A). Endogenous INTS11 was efficiently knocked down (Figure 4A), leading to a clear increase in uncleaved (Figure 4B) and polyadenylated snRNAs (Figure 4C). As expected, transfected GFP–FLAG-INTS11 WT was able to restore basal levels of uncleaved and polyadenylated snRNAs, while the catalytically inactive mutant INST11 E203Q was unable to do so. Remarkably, the SUMOylation-deficient mutant, GFP–FLAG-INTS11 3KR, was also unable to rescue INTS11 depletion (Figure 4B, C, white bars). Consistently, similar results were obtained upon depletion of endogenous INTS11 by siRNA (Supplementary Figure S3A, B).

Over the last years, a growing body of experimental evidence has revealed that the Integrator complex also controls the processing and expression of other RNAPII-dependent transcripts beyond snRNAs, including eRNAs, PROMPTs, DoGs, different lncRNAs and certain mRNAs. We therefore selected several transcripts belonging to these different types of Integrator-regulated RNAs to further explore the involvement of SUMO in the diverse activities of the Integrator complex. To this end, we analyzed previously published TT-TimeLapse-sequencing (TT-TL-seq) data from INST11-depleted HEK 293T cells, rescued by the overexpression of INTS11 WT or the catalytically inactive mutant INTS11 E203Q (12) (Supplementary Figures S3C, S4A and S5A). In our hands, and in line with published TT-TL-seq data, the expression levels of the selected subset of transcripts increased upon depletion of endogenous INTS11. While overexpression of the WT version of this protein was able to restore transcript basal levels, neither the E203Q nor the GFP–FLAG-INTS11 3KR mutants was able to do so (Figure 4D; Supplementary Figures S4B and S5B). In agreement with these observations, overexpression of the SUMO isopeptidases HA-USPL1 or FLAG-SENP6 mirrored the effect of abrogating INTS11 SUMOylation (Supplementary Figure S5C). These results are consistent with the idea that SUMOylation of INTS11 plays an important role for the function of the Integrator complex on its wide spectrum of RNA substrates.

SUMOylation of INTS11 is relevant for its interaction with other Integrator subunits

To study a possible involvement of INTS11 SUMOylation in its interaction with protein partners, we co-transfected HEK 293T cells with expression vectors for GFP–FLAG-INTS11 or FLAG-INTS11, either WT or 3KR versions, plus different HA-tagged Integrator subunits. Whole-cell lysates were subjected to co-immunoprecipitation (co-IP) using anti-GFP nanobodies coupled to magnetic beads or an anti-FLAG antibody. The co-precipitation of selected Integrator subunits was assayed via WB with anti-HA antibody. INTS4 and INTS9, which associate with INTS11 forming the so-called 'catalytic core' of the Integrator complex co-precipitate with INTS11 WT and INTS11 3KR to a similar extent (Figure 5A, B). The interaction of different INTS11 variants with INTS9 was also confirmed by co-expression of 6xHis-INTS9 and FLAG-tagged INTS11 followed by a nickel-mediated pull-down of these two proteins (Supplementary Figure S5D). In contrast to what we observed with the catalytic module subunits, the levels of co-precipitated INTS13 and INTS14 with the INTS11 SUMOylation-deficient mutant were reduced in comparison with their co-precipitation with the WT version of INTS11 (Figure 5C–E). Taken together, these results suggest that SUMO conjugation to INTS11 is not necessary for its interaction with the other components of the catalytic core but is required for its interaction with other Integrator subunits, and consequently for proper assembly of this multimeric complex.

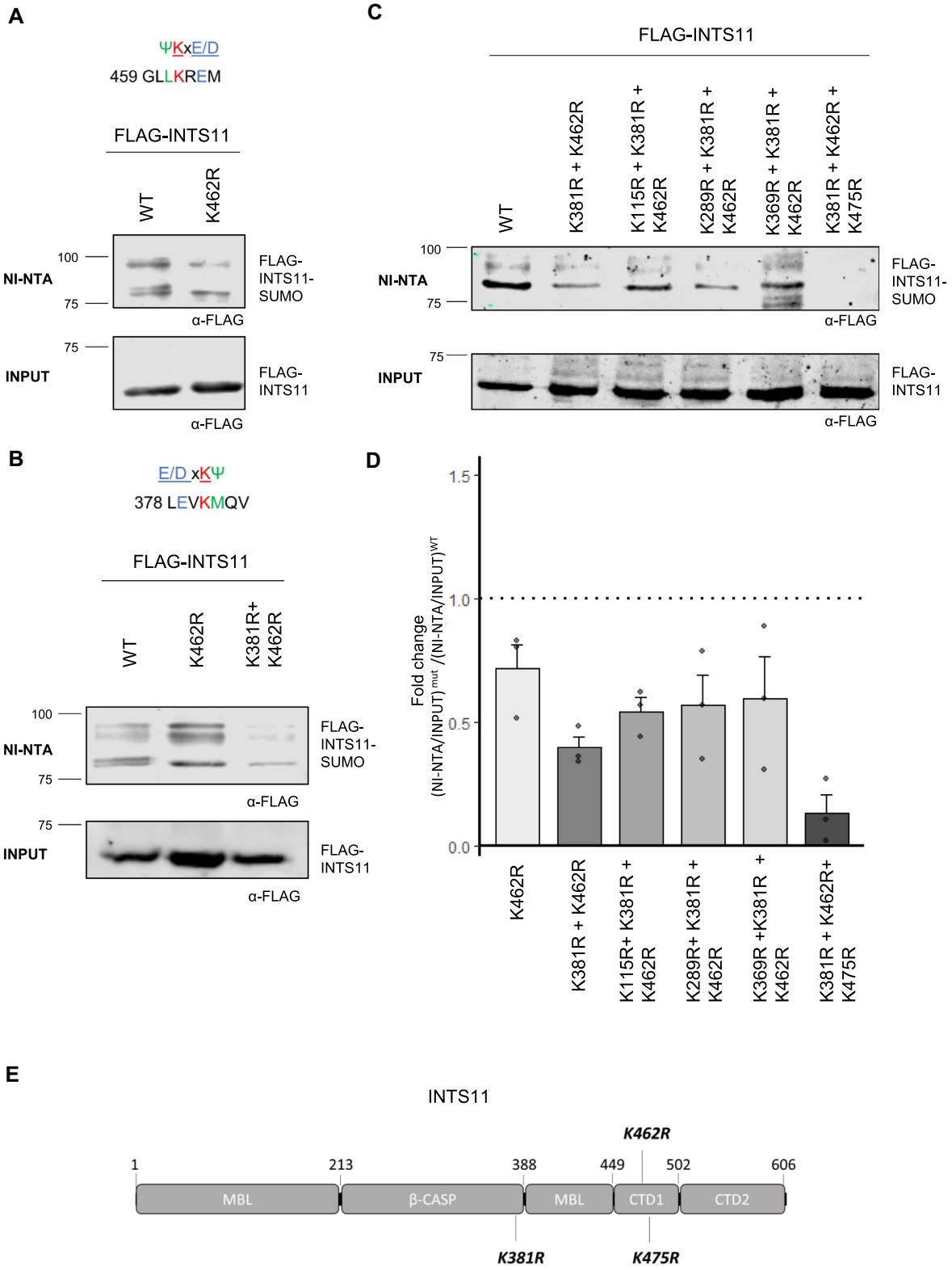

Figure 3. INTS11 is SUMOylated at lysine residues 381, 462 and 475. (**A–C**) HEK 293T cells were co-transfected with expression vectors encoding WT or mutated FLAG-INTS11, and 6×HIS-SUMO2, as indicated above each lane. After 48 h, cell lysates were subjected to nickel affinity chromatography (Ni-NTA). Aliquots of cell lysates (Input) and eluates (Ni-NTA) were analyzed by WB with an anti-FLAG antibody to detect the FLAG-INTS11 variants. (A, B) Upper panels show putative SUMO attachment sites in INTS11, predicted *in silico*, and the surrounding SUMO consensus motif. Ψ bulky, hydrophobic amino acid; x, any amino acid; E, glutamic acid; D, aspartic acid. (**D**) Quantification of SUMO conjugation levels corresponding to three independent experiments as shown in (A–C) was performed with Image Studio Software (LI-COR Biosciences), according to the following calculation: fold change is represented as mean (± SE) = $[\text{NI-NTA/INPUT}]^{\text{mut}}/[\text{NI-NTA/INPUT}]^{\text{WT}}$. (**E**) Domain architecture of INTS11 and localization of detected SUMO conjugation sites (bold).

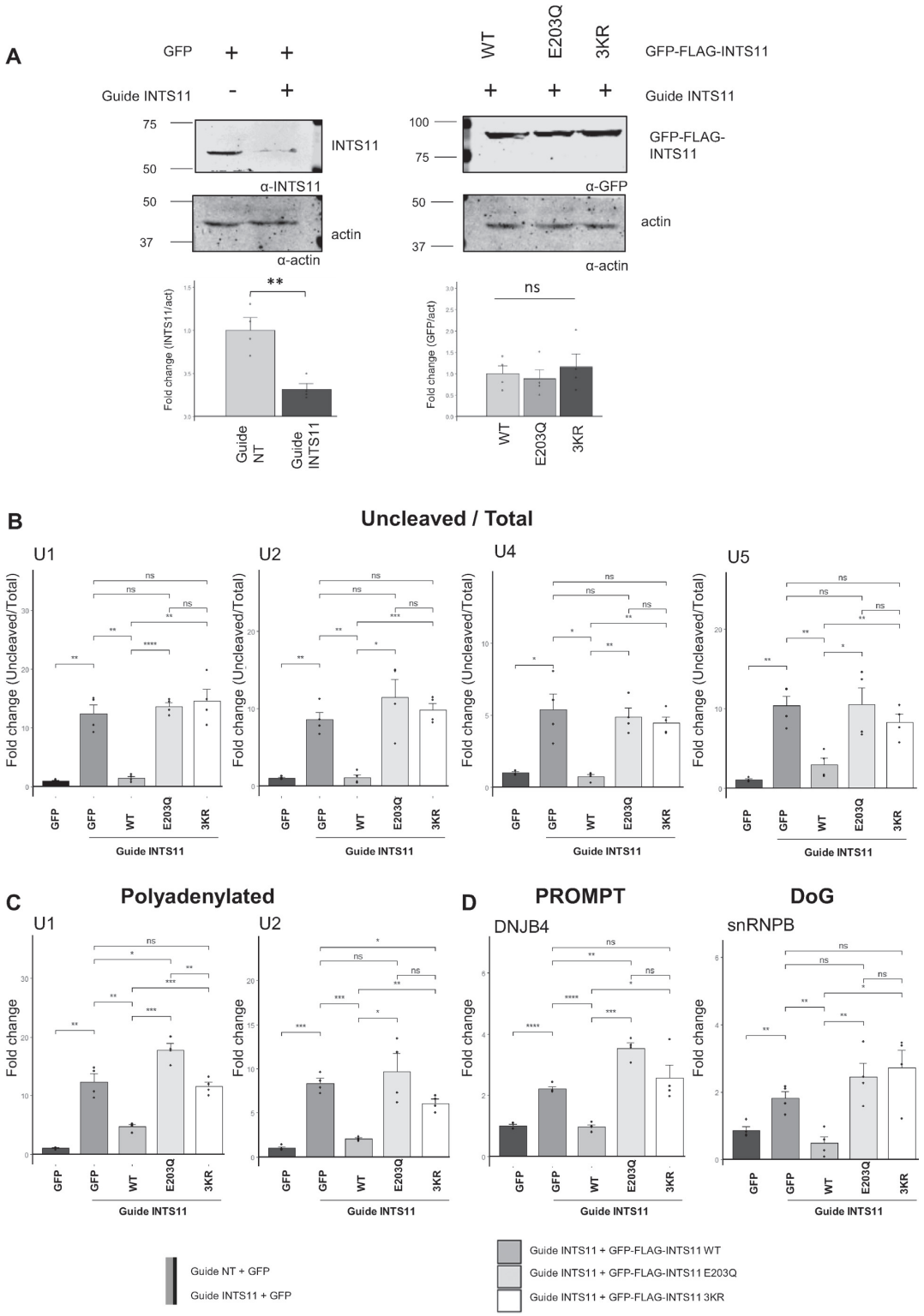

Figure 4. Lack of INTS11 SUMOylation affects Integrator activity in cultured cells. (**A**) HEK 293T cells were co-transfected with dCAS9-KRAB and a non-targeting guide (–) or a guide targeting the INTS11 promoter (+) together with an expression vector for GFP (control), GFP–FLAG-INTS11 WT or GFP–FLAG-INTS11 mutants, as indicated above each lane. Upper panels: whole-cell lysates were subjected to WB with the antibodies indicated below each panel. Lower panels: quantification corresponding to three independent experiments was performed with Image Studio Software (LI-COR Biosciences). Average values are shown with SE and *P*-values, determined using a *t*-test. Significant *P*-values are indicated by the asterisks above the graphs ($n = 3$, $^{**}P < 0.01$). (**B–D**) RT–qPCR analyses of RNA samples from cellular conditions indicated in (A) (and also below each graph and color-coded), using primer pairs targeting each of the indicated transcripts. Reverse transcription was performed with random decamer primers (B and D) or oligo(dT) primer (C). Transcripts levels were normalized to total snRNAs (B) or glyceraldehyde phosphate dehydrogenase (GAPDH) mRNA (C, D). Average values with SEs are shown. *P*-values were determined using a one-way ANOVA and Tukey post-hoc test ($n = 4$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$).

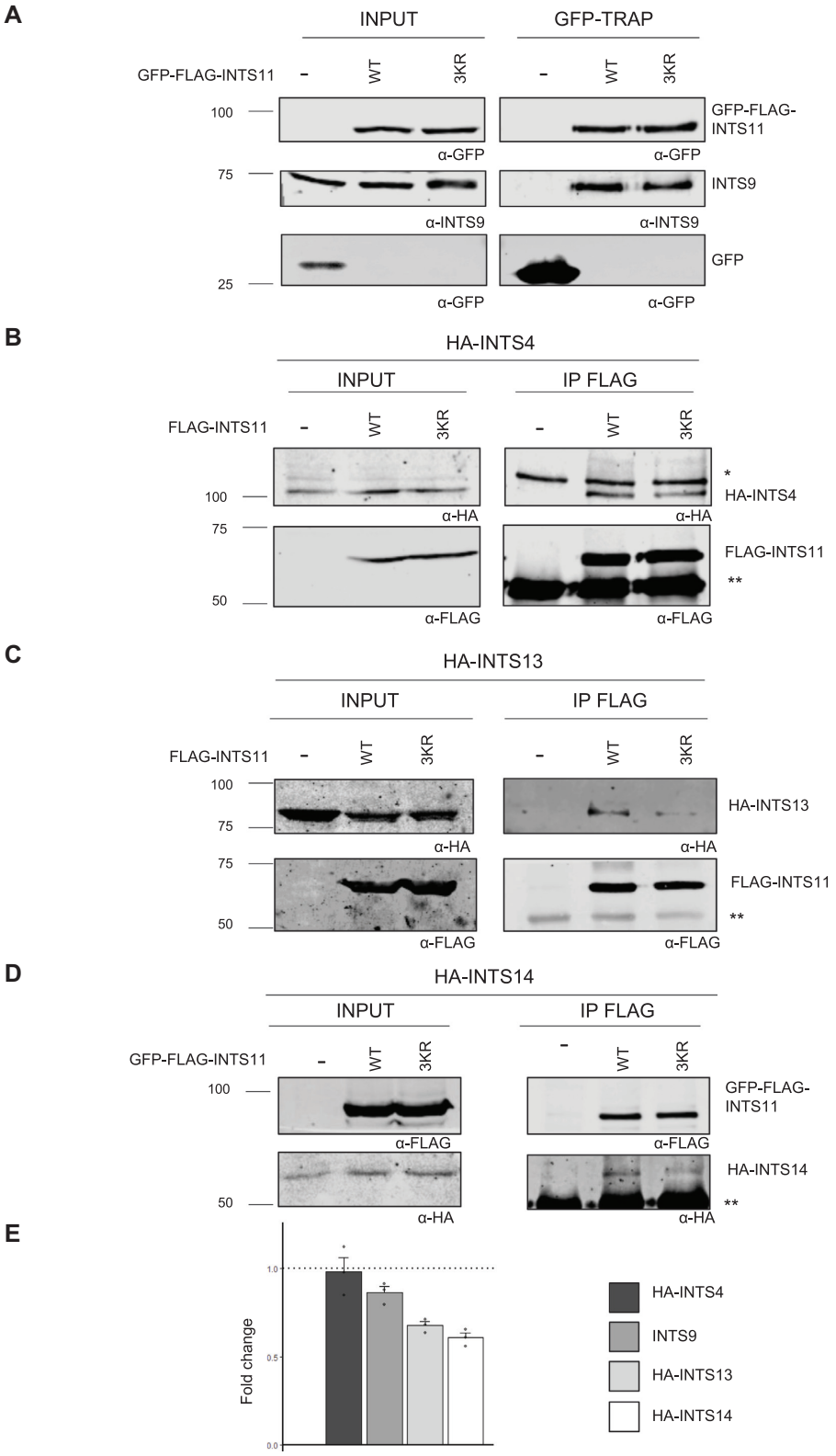

Figure 5. SUMOylation of INTS11 affects its interaction with other subunits of the Integrator complex. (**A**) Overexpressed GFP–FLAG-INTS11 WT or 3KR was immunoprecipitated from HEK 293T cell lysates using GFP–TRAP, and its association with INTS9 was assessed by WB with the antibodies indicated below each panel. (**B–D**) Overexpressed GFP–FLAG-INTS11 WT or 3KR was immunoprecipitated with anti-FLAG from HEK 293T cell lysates containing HA-INTS4 (B), HA-INTS13 (C) or HA-INTS14 (D). Aliquots of whole-cell lysates (input) and immunoprecipitates (IPs) were analyzed by WB with an anti-FLAG antibody to detect FLAG-INTS11 variants and anti-HA to detect enrichment of the other Integrator subunits. (**E**) Quantification corresponding to three independent experiments as shown in (A–D) was performed with Image Studio Software (LI-COR Biosciences), according to the following calculation: fold change is represented as mean (± SE) = $[(\text{IP INTSx/INPUT INTSx})/(\text{IP INTS11/ INPUT INTS11})]^{3KR}/[(\text{IP INTSx/INPUT INTSx})/(\text{IP INTS11/ INPUT INTS11})]^{WT}$ (*non-specific band, **immunoglobulin heavy chain).

INTS11 SUMOylation regulates Integrator subcellular localization

Considering the co-transcriptional nature of snRNA processing (56), we set out to determine whether SUMOylation affects INTS11 recruitment to chromatin. To this end, we performed a ChIP assay with an anti-FLAG antibody and measured by qPCR the levels of precipitated 3′-box region corresponding to the U2 gene. We observed significantly less association of the SUMOylation-deficient FLAG-INTS11 mutant with chromatin than of FLAG-INTS11 WT protein (Supplementary Figure S5E). Additionally, ChIP assays were performed using anti-INTS11 or anti-INTS9 antibodies, on extracts derived from INTS11-depleted cells rescued either with FLAG-INTS11 WT or with FLAG-INTS11 3KR. As expected, upon depletion of INTS11, less chromatin recruitment of this subunit was observed. In agreement with the functional heterodimerization of INTS11 and INTS9, less recruitment of INTS9 was also observed in INTS11-depleted cells (Figure 6A). Consistent with the results presented above, significantly less INTS9 and INTS11 associate with chromatin in the U2 3′-box region upon rescuing INTS11 depletion with FLAG-INTS11 3KR than with FLAG-INTS11 WT protein, even though similar expression levels of WT and 3KR FLAG-INTS11 were observed (Figure 6A; Supplementary Figure S6A, B).

Previous studies have examined the subcellular localization of Integrator subunits with respect to CBs (57). It has been demonstrated that depletion of several Integrator subunit, including the catalytic core INTS4/INTS9/INTS11, causes the disassembly of CBs although these proteins do not show a robust co-localization with these subnuclear bodies (5).

Based on the results mentioned in preceding sections, we set out to analyze the possible impact of SUMO on the subcellular distribution of INTS11. To do so, EGFP–INTS11 variants were overexpressed in HEK 293T and HeLa cells, and analyzed by fluorescence confocal microscopy. As already reported (5), EGFP–INTS11 WT displays a predominant nuclear localization (Figure 6B; Supplementary Figure S6C). Remarkably, EGFP–INTS11 3KR shows a perturbed localization as indicated by the decrease in its nuclear–cytoplasmic ratio (Figure 6B, C; Supplementary Figure S6C, D). Previous reports have demonstrated that the interaction of INTS11 with importins of the α type of nuclear transport receptors, as well as the nuclear localization signal (NLS) spanning from amino acids 460 to 479 within INTS11 were required for the nuclear localization of this protein (58).

To further explore whether INTS11 nuclear–cytoplasmic partitioning is regulated by SUMOylation, we first used bioinformatics prediction to search for putative NLSs within the INTS11 protein sequence. The software NLS mapper (http://nls-mapper.iab.keio.ac.jp/) recognizes the reported NLS (460–479) as well as two others: a monopartite signal between amino acids 469 and 480 and a bipartite signal spanning from amino acids 565 to 597. We found that Lys462, one of the residues that was replaced in our INTS11 3KR mutant, is strictly required for proper prediction of the bipartite NLS located at positions 460–479 (Supplementary Figure S7A). Based on this analysis, we generated a new INTS11 triple mutant termed ‘EA+2KR’. This INTS11 version keeps the Lys to Arg replacements at positions 381 and 475, as our previous SUMOylation-deficient mutant. However, SUMOylation at Lys462 is abolished not by the mutation of the target residue but instead by altering the SUMOylation consensus motif (ΨKxD/E), replacing Glu464 by Ala. It is noticeable that this latter substitution does not affect the prediction of the reported NLS (460–479) (Supplementary Figure S7A). Nevertheless, the subcellular localization of INTS11 EA+2KR is similar to that of INTS11 3KR (Figure 6B, C; Supplementary Figure S6C, 6D). In agreement with our previous observations, EA+2KR was unable to conjugate to SUMO (Supplementary Figure S7B) and to properly process the 3′-end of snRNAs (Supplementary Figure S7C, D). In order to validate these results, we overexpressed GFP–USPL1 or GFP–SENP6 and analyzed mCherry–FLAG-INTS11 WT subcellular localization. Consistently, the levels of SUMOylated mCherry–FLAG-INTS11 WT decreased by overexpressing GFP–USPL1 or GFP–SENP6 (Supplementary Figure S7E), increasing the INTS11 cytoplasmatic localization (Figure 7A, B).

Taken together, these results indicate that INTS11 SUMOylation is necessary for its proper subcellular localization. In addition, considering that INTS11 3KR preserves its interaction with INTS4 and INTS9, at least in our co-IP analysis, the subcellular localization at which the assembly of the catalytic core takes place represents an intriguing question.

DISCUSSION

Despite the fact that RNA-related proteins are the most abundant group among SUMOylation substrates (59–61), including many snRNA-related factors, very little is known about the regulation of proteins involved in snRNA biogenesis by SUMO conjugation. Previous studies have shown that depletion of the SUMO protease USPL1 causes a reduction in nascent and mature snRNA levels, diminishes snRNP production and alters pre-mRNA splicing (51). However, these results could be linked to the relevance of USPL1 for assembly of CBs (50), subnuclear compartments where snRNA maturation occurs (1), which are in close proximity to where snRNA transcription takes place (48). To gain insight into the role of USPL1 in snRNA biogenesis, we overexpressed this SUMO protease in cultured cells and found increased levels of uncleaved snRNAs, suggesting that the balance between SUMOylation and de-SUMOylation is critical for snRNA maturation. This prompted us to investigate the impact of SUMO conjugation on the Integrator complex, which is responsible for snRNA 3′-end processing. Twelve out of 14 subunits of this complex have been detected as SUMO conjugation targets by proteomic studies (27). However, it was not clear what consequences this PTM could have for Integrator complex activity. We confirmed that several of these Integrator subunits are bona fide SUMO substrates in cultured cells (Figure 2) and asked whether SUMO conjugation could be regulating the activity and/or the assembly of this complex.

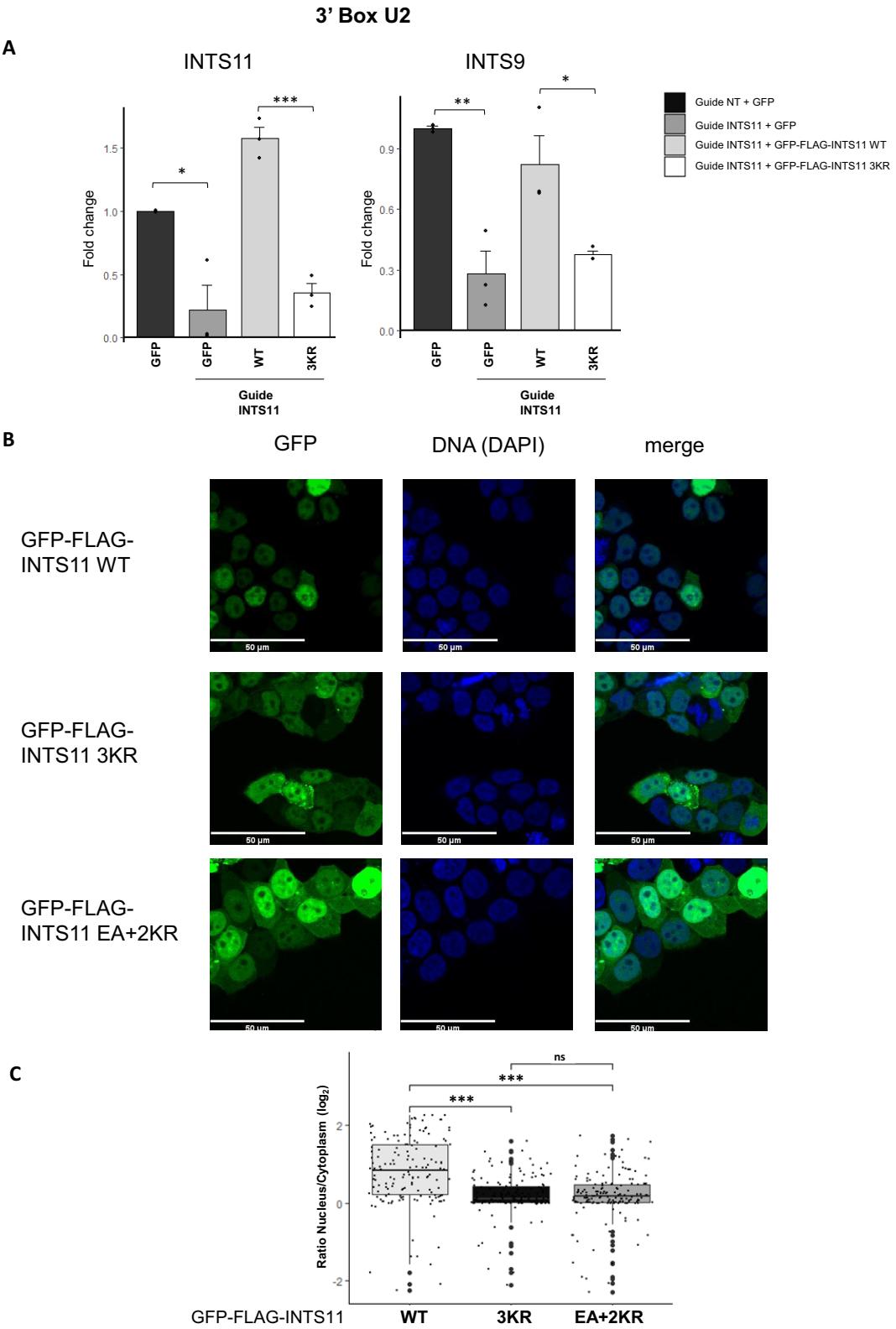

Figure 6. SUMOylation of INTS11 regulates its subcellular localization. (**A**) HEK 293T cells were co-transfected with dCAS9-KRAB and either a non-targeting guide (NT) or a guide targeting INTS11 promoter (Guide INTS11), plus GFP, GFP–FLAG-INTS11 WT or GFP–FLAG-INTS11 3KR, as indicated at the bottom of the panel and color-coded on the right side. After 72 h, ChIP analysis was performed with anti-INTS11 (left panel) or anti-INTS9 (right panel) antibodies. Quantification of immunoprecipitated DNA was assessed by qPCR with specific primers for the U2 3′ box. Average values are shown with the SE ($n = 3$, ***P <0.001; **P <0.01; *P <0.05; Student's t-test). (**B**) Representative confocal microscopy images of HEK 293T cells transfected with GFP–FLAG-INTS11 WT or mutated GFP–FLAG-INTS11. 4′,6-Diamidino-2-phenylindole (DAPI) nuclear staining is also shown. Scale bar, 50 µm. (**C**) Quantification of the nuclear–cytoplasmic ratio from confocal microscopy images ($n = 150$). P-values were determined using a Kruskal–Wallis test (***P <0.001).

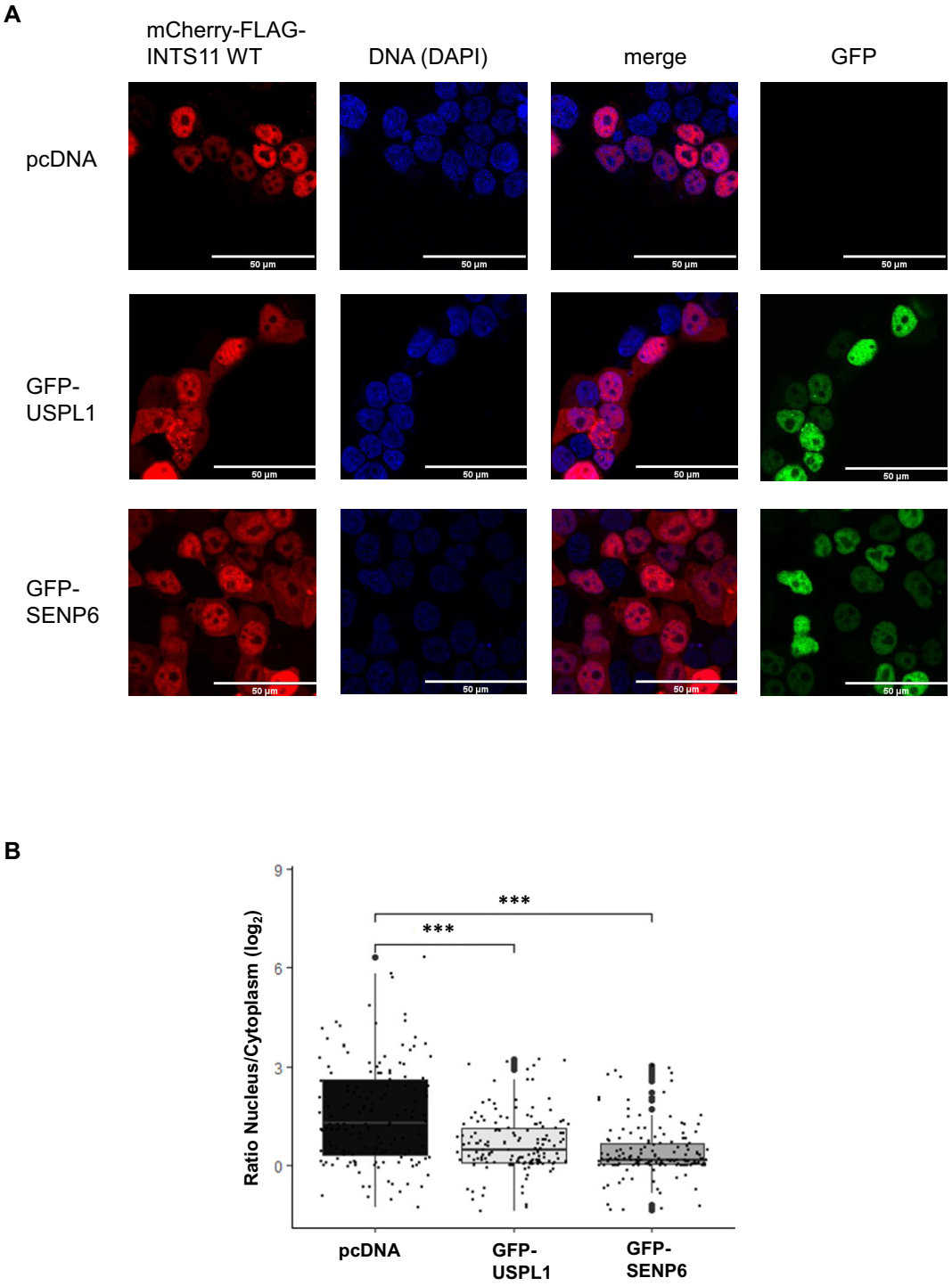

Figure 7. Regulation of INTS11 subcellular localization by SUMO isopeptidases. (**A**) Representative confocal microscopy images of HEK 293T cells co-transfected with mCherry–FLAG-INTS11 WT and GFP–USPL1, GFP–SENP6 or empty plasmid (pcDNA). DAPI nuclear staining is also shown. Scale bar, 50 µm. (**B**) Quantification of the nuclear–cytoplasmic ratio from confocal microscopy images ($n = 150$). *P*-values were determined using a Kruskal–Wallis test (***$P < 0.001$).

The cleavage module of the Integrator complex is formed by INTS4, INTS9 and INTS11 subunits, with INTS11 being responsible for 3′-end processing of snRNAs (5). This module is closely related to the cleavage module of the CPSF complex that is required for pre-mRNA cleavage at polyadenylation sites. Previous studies have not only shown that SUMO conjugation modulates pre-mRNA 3′-end processing but also that CPSF73, the subunit of the CPSF complex that catalyzes the endonucleolytic cleavage, is modified by SUMO. However, the consequences of this modification on CPSF73 catalytic activity are not completely understood (41). Taking into account that INTS11 shares sequence identity with CPSF73 (3), and that USPL1 (Figure 2I) and SENP6 (Supplementary Figure S2B) regulate the levels of INTS11 SUMO conjugation, we focused on INTS11 SUMOylation.

Working with a mutant version of INTS11 that displays severely diminished SUMO conjugation levels, INTS11 3KR, we observed that INTS11 SUMOylation is necessary for efficient Integrator activity in the 3′-end processing of snRNAs (Figure 4; Supplementary Figure S3). Previous reports proposed additional roles for the Integrator complex within global transcription. It acts as an attenuator of expression for transcripts derived from weak promoters, such as PROMPTS, eRNAs and certain lncRNAs (7,10,14,15). However, in some cases, Integrator activity allows RNAPII to enter productive elongation and hence increases transcription (10,16). Here we show that INTS11 3KR is not capable of attenuating transcription of a diverse set of non-coding RNAs (Figure 4D; Supplementary Figures S4B and S5B). Taken together, these results suggest that SUMOylation of INTS11 is necessary for regulating the expression of RNAPII-dependent transcriptional units.

The formation of the INTS9–INTS11 heterodimer requires two specific domains in the C-terminal region of each of these proteins: CTD1 and CTD2 (5,6,53,62). It has been postulated that INTS9/INTS11 assembly requires an interaction between the CTD2 of INTS9 and the CTD2 of INTS11, followed by a gradual formation of multiple interactions between the CTD1 domains of both subunits. As a result, the dimerized CTD1s allow the recruitment of INTS4 to form the catalytic module (6). In this context, it has been shown that deletion of CTD1 from INTS11, where according to our work two SUMO acceptor residues are positioned (Figure 3E), does not impair the interaction between INTS9 and INTS11 (6). Consistent with these results, we observed that INTS11 3KR retains its interaction with INTS9 (Figure 5A; Supplementary Figure S5D). Furthermore, we show that the three mutations present within INTS11 3KR (K381/462R/475R) do not affect its interaction with INTS4 (Figure 5B). This indicates that SUMOylation of INTS11 is not required for the formation of the trimeric catalytic module and suggests that these mutations do not alter, or at least not considerably, the CTD1 domain. Recently, INTS10/INTS13/INTS14 has been characterized as an independent module, which binds to DNA and RNA and possibly stabilizes the cleavage module to target RNAs (63). We observed that INTS11 3KR displays a decreased interaction with two subunits of this module in comparison with the WT version (Figure 5C, D). Although further experimental evidence is still required, these results may suggest that INTS11 SUMOylation is necessary to assemble the cleavage module into the whole Integrator complex.

ChIP analysis revealed a diminished recruitment of INTS11 3KR to chromatin (Figure 6A; Supplementary Figure S5E). We further observed that INTS9 is less recruited to U2 gene loci in cells depleted of endogenous INTS11 than in cells with normal levels of this subunit. Furthermore, we detected less recruitment of INTS9 to chromatin when INTS11-depleted cells were rescued with INTS11 3KR than when they were rescued with INTS11 WT (Figure 6A). Consistent with the finding that SUMOylation of INTS11 does not alter its interaction with INTS9, these results suggest that INTS9 recruitment to chromatin is dependent on INTS11.

Considering the requirement of SUMO conjugation for INTS11 chromatin recruitment as well as the notion that SUMOylation is able to regulate protein intracellular transport, we analyzed whether SUMO conjugation to INTS11 could alter its subcellular localization. Indeed, GFP–FLAG-INTS11 3KR displays a decreased nuclear–cytoplasmic distribution ratio as compared with GFP–FLAG-INTS11 WT (Figure 6B, C; Supplementary Figure S6C, D). INTS11 possesses three putative NLSs (Supplementary Figure S7A). In particular, the signal encompassing residues 460–479 has been validated in cultured cells (58). By *in silico* analysis, we identified that the replacement of Lys462 by Arg abrogates this validated NLS. An additional INTS11 mutant (INTS11 EA+2KR) in which the SUMO consensus motif has been altered without affecting the NLS score is unable to conjugate to SUMO and fails to process nascent snRNAs (Supplementary Figure S7). Remarkably, INTS11 EA+2KR displays the same subcellular distribution as INTS11 3KR, pointing to the relevance of SUMO conjugation for proper INTS11 nuclear–cytoplasmic localization (Figure 6A, B; Supplementary Figure S6C, D). In addition, a previous study showed that Lys462 is necessary for the correct 3′-end processing of snRNAs (6). Consistent with our results, this study reported that mutation of this residue does not alter the formation of the catalytic module. Furthermore, it has been demonstrated that a highly positively charged composite tunnel formed by INTS9/INTS11/INTS4 is necessary for appropriate snRNA processing. These conclusions were achieved by altering the charges involved in this tunnel formation by mutating different residues in the involved proteins, including Lys462 within INTS11 (6). More recently, binding sites for inositol hexakisphosphate (IP_6) have been identified in the positively charged tunnel of the Integrator catalytic module in *Drosophila*. To study the role of IP_6 binding, the positively charged Lys462 was replaced by a negative residue (Glu) within the intS11 subunit. This change impaired the interaction with other subunits of the Integrator complex, affecting its function (64) . Nevertheless, here we show that even without altering the charge of residue 462 (replacing Lys by Arg) or by abrogating the SUMO consensus motif surrounding Lys462, SUMOylation disruption in that

position shifts the subcellular localization of INTS11. Thus, the results reported by these two different studies could also be interpreted in the context of the importance of INTS11 SUMOylation for its subcellular distribution and consequently for its efficient activity. Whether SUMO conjugation to INTS11 modulates its binding to IP_6 awaits further investigation. Strengthening the relevance of SUMO conjugation, overexpression of two SUMO isopeptidases, USPL1 or SENP6, mirrored the changes in subcellular localization of WT INTS11 observed upon abrogating the SUMOylation of this protein (Figure 7).

It has been shown that INTS6/INTS8 form a module that recruits protein phosphatase 2 (PP2A) to active genes. This recruitment dephosphorylates CDK9 substrates, including RNAPII, and regulates the transcription cycle by pausing RNAPII (18, 19). Furthermore, the structure of the Integrator complex bound to PP2A and paused RNAPII has been recently elucidated. This report helps to understand the mechanism by which the Integrator complex regulates transcription (65). However, almost nothing is known about the regulation of the function of this multimeric complex. Here, we demonstrate that INTS11 is SUMOylated and this post-translational modification is necessary to allow its proper subcellular localization, regulating the activity of the Integrator complex and possibly altering its assembly. The consequences of SUMO conjugation for other subunits of the complex remain to be explored.

DATA AVAILABILITY

The raw data underlying this article are available in its online supplementary material.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Valeria Buggiano and Amaranta Avendaño for technical help; Shona Murphy, Joana Guiro, Maria Carmo-Fonseca, Joao Abreu Pessoa and other members from the Murphy and Fonseca laboratories for support, experimental suggestions and helpful discussion; Ezequiel Petrillo, Nicolas Nieto Moreno, and members of the Kornblihtt, de la Mata, Petrillo, Schor and Muñoz laboratories for valuable discussion; and Shona Murphy for critical reading of the manuscript.

FUNDING

This work was supported by grants from the Agencia Nacional de Investigaciones Científicas y Tecnológicas of Argentina (ANPCyT) [2014-2888, 2015-1731, 2017-0111 and 2019-00263 to A.S.] and from the University of Buenos Aires, Argentina (UBACyT) [20020170100045BA to A.S.]. L.B., M.M., A.R. and N.G. are recipients of doctoral fellowships from the CONICET. P.M. has been a doctoral fellow from CONICET (2015–2019) and a postdoctoral fellow (2019–2020) supported by H2020-Marie Sklodowska-Curie Research and Innovation Staff Exchanges (734825-LysoMod). L.B.'s short-term visits to the Murphy laboratory at Oxford University (Oxford, UK) and to the Fonseca laboratory at Instituto de Medicina Molecular (Lisbon, Portugal) have been supported by the same program. A.S. and B.P. are career investigators from CONICET.

Conflict of interest statement. None declared.

REFERENCES

1. Matera,A.G. and Wang,Z. (2014) A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.*, **15**, 108–121.
2. Cuello,P., Boyd,D.C., Dye,M.J., Proudfoot,N.J. and Murphy,S. (1999) Transcription of the human U2 snRNA genes continues beyond the 3′ box in vivo. *EMBO J.*, **18**, 2867–2877.
3. Baillat,D., Hakimi,M.A., Näär,A.M., Shilatifard,A., Cooch,N. and Shiekhattar,R. (2005) Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell*, **123**, 265–276.
4. Yamamoto,J., Hagiwara,Y., Chiba,K., Isobe,T., Narita,T., Handa,H. and Yamaguchi,Y. (2014) DSIF and NELF interact with Integrator to specify the correct post-transcriptional fate of snRNA genes. *Nat. Commun.*, **5**, 4263.
5. Albrecht,T.R., Shevtsov,S.P., Wu,Y., Mascibroda,L.G., Peart,N.J., Huang,K.L., Sawyer,I.A., Tong,L., Dundr,M. and Wagner,E.J. (2018) Integrator subunit 4 is a 'Symplekin-like' scaffold that associates with INTS9/11 to form the Integrator cleavage module. *Nucleic Acids Res.*, **46**, 4241–4255.
6. Pfleiderer,M.M. and Galej,W.P. (2021) Structure of the catalytic core of the Integrator complex. *Mol. Cell*, **81**, 1246–1259.
7. Skaar,J.R., Ferris,A.L., Wu,X., Saraf,A., Khanna,K.K., Florens,L., Washburn,M.P., Hughes,S.H. and Pagano,M. (2015) The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res.*, **25**, 288–305.
8. Cazalla,D., Xie,M. and Steitz,J.A. (2011) A primate herpesvirus uses the integrator complex to generate viral microRNAs. *Mol. Cell*, **43**, 982–992.
9. Lai,F., Gardini,A., Zhang,A. and Shiekhattar,R. (2015) Integrator mediates the biogenesis of enhancer RNAs. *Nature*, **525**, 399–403.
10. Lykke-Andersen,S., Žumer,K., Molska,E.Š., Rouvière,J.O., Wu,G., Demel,C., Schwalb,B., Schmid,M., Cramer,P. and Jensen,T.H. (2021) Integrator is a genome-wide attenuator of non-productive transcription. *Mol. Cell*, **81**, 514–529.
11. Nojima,T., Tellier,M., Foxwell,J., Almeida,C.R. de, Tan-Wong,S.M., Dhir,S., Dujardin,G., Dhir,A., Murphy,S. and Proudfoot,N.J. (2018) Deregulated expression of mammalian lncRNA through loss of SPT6 induces R-loop formation, replication stress, and cellular senescence. *Mol. Cell*, **72**, 970–984.
12. Rosa-Mercado,N.A., Zimmer,J.T., Apostolidi,M., Rinehart,J., Simon,M.D. and Steitz,J.A. (2021) Hyperosmotic stress alters the RNA polymerase II interactome and induces readthrough transcription despite widespread transcriptional repression. *Mol. Cell*, **81**, 502–513.
13. Gardini,A., Baillat,D., Cesaroni,M., Hu,D., Marinis,J.M., Wagner,E.J., Lazar,M.A., Shilatifard,A. and Shiekhattar,R. (2014) Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell*, **56**, 128–139.
14. Elrod,N.D., Henriques,T., Huang,K.L., Tatomer,D.C., Wilusz,J.E., Wagner,E.J. and Adelman,K. (2019) The integrator complex attenuates promoter-proximal transcription at protein-coding genes. *Mol. Cell*, **76**, 738–752.
15. Tatomer,D.C., Elrod,N.D., Liang,D., Xiao,M.S., Jiang,J.Z., Jonathan,M., Huang,K.L., Wagner,E.J., Cherry,S. and Wilusz,J.E. (2019) The Integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes Dev.*, **33**, 1525–1538.
16. Beckedorff,F., Blumenthal,E., daSilva,L.F., Aoi,Y., Cingaram,P.R., Yue,J., Zhang,A., Dokaneheifard,S., Valencia,M.G., Gaidosh,G. *et al.* (2020) The human integrator complex facilitates transcriptional elongation by endonucleolytic cleavage of nascent transcripts. *Cell Rep.*, **32**, 107917.
17. Zheng,H., Qi,Y., Hu,S., Cao,X., Xu,C., Yin,Z., Chen,X., Li,Y., Liu,W., Li,J. *et al.* (2020) Identification of Integrator–PP2A complex

(INTAC), an RNA polymerase II phosphatase. *Science*, **370**, eabb5872.
18. Vervoort,S.J., Welsh,S.A., Devlin,J.R., Barbieri,E., Knight,D.A., Offley,S., Bjelosevic,S., Costacurta,M., Todorovski,I., Kearney,C.J. *et al.* (2021) The PP2A–Integrator–CDK9 axis fine-tunes transcription and can be targeted therapeutically in cancer. *Cell*, **184**, 3143–3162.
19. Huang,K.L., Jee,D., Stein,C.B., Elrod,N.D., Henriques,T., Mascibroda,L.G., Baillat,D., Russell,W.K., Adelman,K. and Wagner,E.J. (2020) Integrator recruits protein phosphatase 2A to prevent pause release and facilitate transcription termination. *Mol. Cell*, **80**, 345–358.
20. Yue,J., Lai,F., Beckedorff,F., Zhang,A., Pastori,C. and Shiekhattar,R. (2017) Integrator orchestrates RAS/ERK1/2 signaling transcriptional programs. *Genes Dev.*, **31**, 1809–1820.
21. Geiss-Friedlander,R. and Melchior,F. (2007) Concepts in sumoylation: a decade on. *Nat. Rev. Mol. Cell Biol.*, **8**, 947–956.
22. Gareau,J.R. and Lima,C.D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.*, **11**, 861–871.
23. Becker,J., Barysch,S.V., Karaca,S., Dittner,C., Hsiao,H.-H., Diaz,M.B., Herzig,S., Urlaub,H. and Melchior,F. (2013) Detecting endogenous SUMO targets in mammalian cells and tissues. *Nat. Struct. Mol. Biol.*, **20**, 525–531.
24. Impens,F., Radoshevich,L., Cossart,P. and Ribet,d,D. (2014) Mapping of SUMO sites and analysis of SUMOylation changes induced by external stimuli. *Proc. Natl Acad. Sci. USA*, **111**, 12432–12437.
25. Liang,Y.-C., Lee,C.-C., Yao,Y.-L., Lai,C.-C., Schmitz,M.L. and Yang,W.-M. (2016) SUMO5, a novel poly-SUMO isoform, regulates PML nuclear bodies. *Sci. Rep.*, **6**, 26509.
26. Matic,I., Schimmel,J., Hendriks,I.A., Santen,M.A. van, Rijke,F. van de, Dam,H. van, Gnad,F., Mann,M. and Vertegaal,A.C.O. (2010) Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol. Cell*, **39**, 641–652.
27. Hendriks,I.A., Lyon,D., Su,D., Skotte,N.H., Daniel,J.A., Jensen,L.J. and Nielsen,M.L. (2018) Site-specific characterization of endogenous SUMOylation across species and organs. *Nat. Commun.*, **9**, 2456.
28. Hickey,C.M., Wilson,N.R. and Hochstrasser,M. (2012) Function and regulation of SUMO proteases. *Nat. Rev. Mol. Cell Biol.*, **13**, 755–766.
29. Jentsch,S. and Psakhye,I. (2013) Control of nuclear activities by substrate-selective and protein-group SUMOylation. *Annu. Rev. Genet.*, **47**, 167–186.
30. Geoffroy,M.-C. and Hay,R.T. (2009) An additional role for SUMO in ubiquitin-mediated proteolysis. *Nat. Rev. Mol. Cell Biol.*, **10**, 564–568.
31. Johnson,P.R. and Hochstrasser,M. (1997) SUMO-1: ubiquitin gains weight. *Trends Cell Biol.*, **7**, 408–413.
32. Nacerddine,K., Lehembre,F., Bhaumik,M., Artus,J., Cohen-Tannoudji,M., Babinet,C., Pandolfi,P.P. and Dejean,A. (2005) The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev. Cell*, **9**, 769–779.
33. Panse,V.G., Kressler,D., Pauli,A., Petfalski,E., Gnädig,M., Tollervey,D. and Hurt,E. (2006) Formation and nuclear export of preribosomes are functionally linked to the small-ubiquitin-related modifier pathway. *Traffic*, **7**, 1311–1321.
34. Westman,B.J., Verheggen,C., Hutten,S., Lam,Y.W., Bertrand,E. and Lamond,A.I. (2010) A proteomic screen for nucleolar SUMO targets shows SUMOylation modulates the function of Nop5/Nop58. *Mol. Cell*, **39**, 618–631.
35. Finkbeiner,E., Haindl,M., Raman,N. and Muller,S. (2011) SUMO routes ribosome maturation. *Nucleus*, **2**, 527–532.
36. Cubeñas-Potts,C. and Matunis,M.J. (2013) SUMO: a multifaceted modifier of chromatin structure and function. *Dev. Cell*, **24**, 1–12.
37. Hendriks,I.A. and Vertegaal,A.C.O. (2016) A comprehensive compilation of SUMO proteomics. *Nat. Rev. Mol. Cell Biol.*, **17**, 581–595.
38. Ihara,M., Stein,P. and Schultz,R.M. (2008) UBE2I (UBC9), a SUMO-conjugating enzyme, localizes to nuclear speckles and stimulates transcription in mouse oocytes. *Biol. Reprod.*, **79**, 906–913.
39. Spector,D.L. and Lamond,A.I. (2011) Nuclear speckles. *Cold Spring Harb. Perspect. Biol.*, **3**, a000646.
40. Pozzi,B., Bragado,L., Will,C.L., Mammi,P., Risso,G., Urlaub,H., Lührmann,R. and Srebrow,A. (2017) SUMO conjugation to spliceosomal proteins is required for efficient pre-mRNA splicing. *Nucleic Acids Res.*, **45**, 6729–6745.
41. Vethantham,V., Rao,N. and Manley,J.L. (2007) Sumoylation modulates the assembly and activity of the pre-mRNA 3′ processing complex. *Mol. Cell. Biol.*, **27**, 8848–8858.
42. Desterro,J.M.P., Keegan,L.P., Jaffray,E., Hay,R.T., O'Connell,M.A. and Carmo-Fonseca,M. (2005) SUMO-1 modification alters ADAR1 editing activity. *Mol. Biol. Cell*, **16**, 5115–5126.
43. Frey,M.R., Bailey,A.D., Weiner,A.M. and Matera,A.G. (1999) Association of snRNA genes with coiled bodies is mediated by nascent snRNA transcripts. *Curr. Biol.*, **9**, 126–136.
44. Gao,L., Frey,M.R. and Matera,A.G. (1997) Human genes encoding U3 snRNA associate with coiled bodies in interphase cells and are clustered on chromosome 17p11.2 in a complex inverted repeat structure. *Nucleic Acids Res.*, **25**, 4740–4747.
45. Jacobs,E.Y., Frey,M.R., Wu,W., Ingledue,T.C., Gebuhr,T.C., Gao,L., Marzluff,W.F. and Matera,A.G. (1999) Coiled bodies preferentially associate with U4, U11, and U12 small nuclear RNA genes in interphase HeLa cells but not with U6 and U7 genes. *Mol. Biol. Cell*, **10**, 1653–1663.
46. Frey,M.R. and Matera,A.G. (2001) RNA-mediated interaction of Cajal bodies and U2 snRNA genes. *J. Cell Biol.*, **154**, 499–509.
47. Smith,K.P. and Lawrence,J.B. (2000) Interactions of U2 gene loci and their nuclear transcripts with Cajal (coiled) bodies: evidence for PreU2 within Cajal bodies. *Mol. Biol. Cell*, **11**, 2987–2998.
48. Machyna,M., Kehr,S., Straube,K., Kappei,D., Buchholz,F., Butter,F., Ule,J., Hertel,J., Stadler,P.F. and Neugebauer,K.M. (2014) The coilin interactome identifies hundreds of small noncoding RNAs that traffic through cajal bodies. *Mol. Cell*, **56**, 389–399.
49. Smith,E.R., Lin,C., Garrett,A.S., Thornton,J., Mohaghegh,N., Hu,D., Jackson,J., Saraf,A., Swanson,S.K., Seidel,C. *et al.* (2011) The little elongation complex regulates small nuclear RNA transcription. *Mol. Cell*, **44**, 954–965.
50. Schulz,S., Chachami,G., Kozaczkiewicz,L., Winter,U., Stankovic-Valentin,N., Haas,P., Hofmann,K., Urlaub,H., Ovaa,H., Wittbrodt,J. *et al.* (2012) Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions. *EMBO Rep.*, **13**, 930–938.
51. Hutten,S., Chachami,G., Winter,U., Melchior,F. and Lamond,A.I. (2014) A role for the cajal-body-associated SUMO isopeptidase USPL1 in snRNA transcription mediated by RNA polymerase. *J. Cell Sci.*, **127**, 1065–1078.
52. Tatham,M.H., Rodriguez,M.S., Xirodimas,D.P. and Hay,R.T. (2009) Detection of protein SUMOylation in vivo. *Nat. Protoc.*, **4**, 1363–1371.
53. Albrecht,T.R. and Wagner,E.J. (2012) snRNA 3′ end formation requires heterodimeric association of Integrator subunits. *Mol. Cell. Biol.*, **32**, 1112–1123.
54. Nayak,A. and Müller,S. (2014) SUMO-specific proteases/isopeptidases: SENPs and beyond. *Genome Biol.*, **15**, 422.
55. Gilbert,L.A., Horlbeck,M.A., Adamson,B., Villalta,J.E., Chen,Y., Whitehead,E.H., Guimaraes,C., Panning,B., Ploegh,H.L., Bassik,M.C. *et al.* (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell*, **159**, 647–661.
56. Guiro,J. and Murphy,S. (2017) Regulation of expression of human RNA polymerase II-transcribed snRNA genes. *Open Biol.*, **7**, 170073.
57. Takata,H., Nishijima,H., Maeshima,K. and Shibahara,K.I. (2012) The integrator complex is required for integrity of Cajal bodies. *J. Cell Sci.*, **125**, 166–175.
58. Mackmull,M.-T., Klaus,B., Heinze,I., Chokkalingam,M., Beyer,A., Russell,R.B., Ori,A. and Beck,M. (2017) Landscape of nuclear transport receptor cargo specificity. *Mol. Syst. Biol.*, **13**, 962.
59. Vertegaal,A.C.O., Ogg,S.C., Jaffray,E., Rodriguez,M.S., Hay,R.T., Andersen,J.S., Mann,M. and Lamond,A.I. (2004) A proteomic study of SUMO-2 target proteins. *J. Biol. Chem.*, **279**, 33791–33798.
60. Blomster,H.A., Hietakangas,V., Wu,J., Kouvonen,P., Hautaniemi,S. and Sistonen,L. (2009) Novel proteomics strategy brings insight into the prevalence of SUMO-2 target sites. *Mol. Cell. Proteomics MCP*, **8**, 1382–1390.
61. Golebiowski,F., Matic,I., Tatham,M.H., Cole,C., Yin,Y., Nakamura,A., Cox,J., Barton,G.J., Mann,M. and Hay,R.T. (2009)

System-wide changes to SUMO modifications in response to heat shock. *Sci. Signal.*, **2**, ra24.
62. Wu,Y., Albrecht,T.R., Baillat,D., Wagner,E.J. and Tong,L. (2017) Molecular basis for the interaction between Integrator subunits IntS9 and IntS11 and its functional importance. *Proc. Natl Acad. Sci. USA*, **114**, 4394–4399.
63. Sabath,K., Stäubli,M.L., Marti,S., Leitner,A., Moes,M. and Jonas,S. (2020) INTS10–INTS13–INTS14 form a functional module of Integrator that binds nucleic acids and the cleavage module. *Nat. Commun.*, **11**, 3422.
64. Lin,M.-H., Jensen,M.K., Elrod,N.D., Huang,K.-L., Welle,K.A, Wagner,E.J. and Tong,L. (2022) Inositol hexakisphosphate is required for Integrator function. *Nat. Commun.*, **13**, 5742.
65. Fianu,I., Chen,Y., Dienemann,C., Dybkov,O., Linden,A., Urlaub,H. and Cramer,P. (2021) Structural basis of Integrator-mediated transcription regulation. *Science*, **374**, 883–887.