



UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Fisiología, Biología Molecular y Celular.

Participación de la conjugación de SUMO en la biogénesis de los ARN pequeños nucleares

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el
área de CIENCIAS BIOLÓGICAS

Laureano Fabian Tomas Bragado

Directora de tesis: Dra. Anabella Srebrow

Consejero de estudios: Dr. Alberto Kornblihtt

Lugar de trabajo: IFIBYNE-CONICET; FCEyN-UBA

Buenos Aires, 2022.

Participación de la conjugación de SUMO en la biogénesis de los ARN pequeños nucleares

RESUMEN

Los ARN pequeños nucleares (snRNAs) son moléculas de ARN no codificantes, de menos de 350 nucleótidos, ricos en uracilos, que se asocian con un gran conjunto de proteínas formando ribonucleopartículas pequeñas nucleares conocidas como “snRNPs”. Las funciones biológicas de los snRNAs están bien caracterizadas desde hace años. Los snRNAs U1, U2, U4, U5 y U6 forman las partículas snRNPs que llevan el mismo nombre y que se ensamblan para dar lugar al spliceosoma “mayor”, mientras que U11, U12, U4 atac, U6 atac juntos con U5 participan de la conformación del spliceosoma “menor” o minoritario. Por su parte, el snRNA U7 es requerido para el procesamiento de los pre-ARNm que codifican para las histonas no replicativas y el U3 es necesario para el procesamiento del ARN ribosomal. El snRNA 7SK es responsable de regular los niveles activos de p-TEFb, un complejo quinasa responsable de fosforilar a la ARN polimerasa II. La ARN polimerasa III es la responsable de transcribir a U6, U6 atac y 7SK, mientras que la ARN polimerasa II transcribe los restantes snRNAs.

Teniendo en cuenta la relevancia de los snRNAs para la formación del spliceosoma y consecuentemente para el proceso de splicing, que la regulación de la expresión de estos ARNs no codificantes es aún escasamente comprendida, así como también trabajos previos de nuestro laboratorio que demostraron la importancia de la conjugación de SUMO (o “SUMOilación”) para el ensamblado y actividad catalítica del spliceosoma, decidimos explorar una posible conexión entre la maquinaria de SUMO y la biogénesis de los snRNAs transcritos por la ARN polimerasa II.

Hemos hallado que disminuir la SUMOilación global en células en cultivo, mediante la sobre-expresión de distintas SUMO proteasas, altera los niveles de snRNAs nacientes (no procesados). Además, pudimos identificar que diversas proteínas que forman parte de los complejos reguladores de las diferentes etapas de la biogénesis de los snRNAs son blanco de conjugación de SUMO. Nos enfocamos en dos de estas proteínas: SNAPC1, un factor de transcripción específico de snRNAs que pertenece al

complejo SNAPc; y la proteína INTS11, subunidad responsable del procesamiento 3' de los snRNAs nacientes dentro del complejo *Integrator*.

En el caso de SNAPC1, observamos que al mutar las lisinas 245 y 333 disminuye drásticamente su conjugación a SUMO y esta proteína mutante deficiente de SUMOilación (SNAPC1 K245/333R) es incapaz de activar la transcripción de U1 y U2, a pesar de no tener alterado su reclutamiento a los promotores correspondientes. Esto indicaría que la SUMOilación de SNAPC1 es necesaria para que pueda ensamblarse el complejo SNAPc o bien para que este complejo reclute a los factores generales de la transcripción. Además, observamos que en ciertas líneas celulares en cultivo, la sobre expresión de dicha mutante lleva a la muerte celular. Esto podría deberse a alguna tolerancia diferencial a las alteraciones en los niveles de snRNAs.

En el caso de INTS11, mediante mutagénesis dirigida pudimos determinar que las lisinas 381, 462 y 475 son blancos de conjugación de SUMO. Observamos que una versión triple mutante de esta proteína, INTS11 K381/462/475R o "3KR", es incapaz de clivar transcriptos nacientes de snRNAs, a diferencia de la versión salvaje. Además, INTS11 3KR tiene alterada su interacción con algunas subunidades del complejo *Integrator*, pero no con aquellas que conforman el módulo catalítico (INTS4/INTS9/INTS11). Estos resultados sugieren que la SUMOilación de INTS11 no participa en el ensamblado del módulo catalítico, pero es necesaria para la conformación de todo el complejo *Integrator*. Por otro lado, encontramos que INTS11 3KR presenta una localización sub-celular más citoplasmática que la versión salvaje. Se ha postulado al complejo *Integrator* como un complejo modular que se va ensamblado en forma co-transcripcional. Nuestros resultados nos llevan a proponer que el módulo catalítico se encuentra pre-ensamblado y que podría hallarse como tal en el citoplasma.

Los resultados presentados muestran la relevancia de la SUMOilación en la biogénesis de los snRNAs, tanto en el inicio de la transcripción como en el procesamiento 3' de dichos transcriptos.

Palabras claves:

Modificaciones post-traduccionales – conjugación de SUMO - snRNAs - Complejo Integrator - INTS11 - Complejo SNAPc - SNAPC1

INVOLVEMENT OF SUMO CONJUGATION IN SMALL NUCLEAR RNA BIOGENESIS

ABSTRACT

Small nuclear RNAs (snRNAs) are non-coding RNA molecules of less than 350 nucleotides, rich in uracil, which associate with a large set of proteins giving rise to small nuclear ribonucleoproteins known as "snRNPs". The biological functions of snRNAs have been well characterized. The snRNAs U1, U2, U4, U5 and U6 form the snRNPs that are recognized by the same name, and which composed the major spliceosome; while U11, U12, U4^{atac}, U6^{atac} together with U5 participate in the conformation of the minor spliceosome. U7 snRNA is required for the processing of pre-mRNAs that encode non-replicative histones and U3 is necessary for ribosomal RNA processing. 7SK snRNA is responsible for regulating the activity of p-TEFb, a kinase complex responsible for phosphorylating RNA polymerase II. RNA polymerase III is responsible for transcribing U6, U6^{atac} and 7SK, whereas RNA polymerase II transcribes the remaining snRNAs.

Taking into account the relevance of snRNAs for spliceosome composition and function and consequently for the splicing process, that the regulation of the expression of these non-coding RNAs is still poorly understood, together with previous work from our laboratory that demonstrated the importance of SUMO conjugation (or "SUMOylation") for spliceosome assembly and catalytic activity, we decided to explore a possible connection between the SUMO machinery and the biogenesis of snRNAs, in particular those transcribed by RNA polymerase II.

We have found that decreasing global SUMOylation in cultured cells, through the overexpression of different SUMO proteases, alters the levels of nascent (unprocessed) snRNAs. Furthermore, we have identified a variety of proteins that are part of the regulatory complexes of the different stages of snRNA biogenesis as *bonafide* SUMO conjugation targets. We focused on two of these proteins: SNAPC1, a specific snRNA transcription factor that belongs to the SNAPc complex; and the INTS11 protein, the subunit responsible for the 3' endonucleolytic cleavage of nascent snRNAs within the Integrator complex.

In the case of SNAPC1, we have observed that mutating lysine residues 245 and 333 drastically reduced SUMO conjugation to this protein. This obtained SUMOylation-deficient mutant (SNAPC1 K245/333R) is unable to activate the transcription of U1 and U2, despite the fact it is still recruited to the corresponding promoters. This would indicate that SUMOylation of SNAPC1 is required for SNAPc complex assembly or for this complex to recruit general transcription factors. Furthermore, we have observed that in certain mammalian cell lines, the overexpression of this mutant leads to cell death. This could be due to a differential sensitivity to alterations in snRNA levels.

In the case of INTS11, through site-directed mutagenesis we have determined that lysine residues 381, 462 and 475 are SUMO conjugation sites within this protein. We observed that a triple mutant version, INTS11 K381 / 462 / 475R ("3KR"), is unable to cleave nascent snRNA transcripts, unlike the wild type version. Furthermore, INTS11 3KR displays a diminished interaction with certain subunits of the Integrator complex, but not with those that compose the catalytic core (INTS4 / INTS9 / INTS11). These results suggest that INTS11 SUMOylation is not involved in the assembly of the Integrator catalytic core but is necessary for the conformation of the entire complex. On the other hand, we have found that INTS11 3KR displays a more cytoplasmic sub-cellular localization than the wild type version. The Integrator complex has been postulated as a modular complex that is assembled in a co-transcriptional manner. Our results lead us to suggest that the catalytic module is independently pre-assembled and that it could be found as such in the cytoplasm.

The results presented in this thesis uncover the relevance of SUMOylation for snRNA biogenesis, both for transcription initiation as well as for 3' end processing of these transcripts.

Key words:

Posttranslational Modifications - SUMO conjugation – snRNAs - Integrator complex – INTS11 - SNAP complex – SNAPC1

AGRADECIMIENTOS

A lo largo de los 6 años (5+1 por COVID) que duro mi doctorado muchas personas me ayudaron, creo que nada de lo que pueda escribir va a representar totalmente mi agradecimiento y sé que me voy a quedar corto con cada uno, pero bueno voy a hacer lo mejor que pueda, agradezco:

A Anabella Srebrow por enseñarme a ser crítico con el trabajo, por tener en cuenta mis ideas y darme libertad a la hora de trabajar. Gracias por la paciencia a la hora de corregir mis textos (sobre todo en inglés), explicarme como mejorar mis redacciones y los millones de consejos útiles a lo largo de estos años. Gracias por fomentar mi curiosidad y encontrar soluciones a todo tipo de problemas. Pero sobre todo por ser una buena persona, eso es algo que valoro muchísimo. Gracias Bella!

A Berta “la Gringa” Pozzi, por tenerme paciencia y enseñarme todo sobre este trabajo desde los primeros años hasta hoy que me ayuda con problemas técnicos y discusiones científicas. Gracias por ser tan generosa, buena compañera y amiga.

A Pablito Andres Mammmmi por ayudarme cada vez que se lo pedí (fueron muchas), ser critico a la hora de discutir mis resultados y darme ideas para mejorar. Gracias por los fernets compartidos, las frases memorables y por ser tan buen amigo. Realmente fue un gusto trabajar con vos.

A Melina Magalnik (no confundir con Melisa Magali) por la excelente predisposición al trabajar, por la buena onda y entusiasmo, por estar sonriendo casi siempre. Gracias por tenerme paciencia y la ayuda al final, hubiese sido imposible obtener los últimos resultados sin ella.

A Nicolas “el Pocho” Gaioli por la energía y la buena actitud que contagia en el día a día. Gracias por darme una mano siempre que te lo pedí, gracias ser un excelente compañero, persona, y por los millones de viajes en el pochimovil.

A Alberto Kornblihtt por contagiar la pasión por la ciencia, fomentar la unión del grupo grande “ARKAS”, marcar el camino y mostrar que se puede ser un excelente científico-docente y ser una buena persona al mismo tiempo. A Nicolas “el Niño pH” Nieto Moreno por enseñarme a ser crítico y perseverante, por compartir su

conocimiento, por las risas, por responder preguntas a cualquier hora, por ser un grande dentro y fuera del labo. A Luciano “el Gato de Chaca” Marasco por ser tan honesto y generoso, por los cientos de favores que le he pedido, los mil viajes en el luchomóvil, por compartir los horarios más raros para trabajar, por ser el culpable de todo el quilombo del labo (incluso ahora que no estas), un “distinto entre los distintos”. A Valeria Buggiano por enseñar desde la humildad, dar una mano siempre y ser un pilar fundamental en el labo.

A Ezequiel “Petry” Petrillo por entenderme cuando digo muchas palabras sin sentido, por darme mil tips en casi todos los protocolos, por charlar de cualquier cosa con los primeros mates del día, por los consejos para el trabajo y la vida en general, por estar siempre (sobre todo cuando pasa algo en el labo y no sé qué hacer). Manuel “Nolo” Muñoz por mostrar que la tenacidad es clave en este trabajo, por ser un compañero más, por aconsejarme cuando lo necesite. A Ignacio “Nacho” Schor y Manuel de la Mata por ayudarme y enseñarme cada vez que se los pedí, por mostrarme que la manera de crecer en un trabajo es siendo un buen compañero y trabajando en colaboración.

A Juan “Bigote” Muñoz por su compañerismo, por su curiosidad innata, por compartir cada idea nueva, por escuchar. A Guillermina Kubaczka Zoppi por tener infinitos temas de charlas, por traer música nueva, por ser tan generosa, por la energía que contagia. A Adri Botto por compartir su mirada crítica, por explicar con palabras simples temas complejos, por optimizar muchísimos de los protocolos que usamos. A Federico “Fefuwi” Fuchs por estar siempre con una buena actitud, por los aventones a villurca, por compartir desde algún reactivo hasta sus conocimientos en inglés. A Lucas “el oso Lukaku” Servi y Jose “el Tano” Stigliano por compartir su increíble sentido del humor y visión de la vida (los dos son tan distintos y originales). A Inu Beckerman y Florchi Rodriguez por preguntar y repreguntar, por compartir alegría y contagiar entusiasmo, por ser buenas contadoras de anécdotas random. A Mica Godoy Herz, Lu Giono, Eze Nazer y Marquitos Palavecino por compartir ideas, reactivos y ayudar con los problemas de la mesada. A CamiArcu, Maru Pinkasz y Peter por los intercambios de conocimiento sobre bioinformática/mesada, por sumar una buena onda y ganas de trabajar. A los demás miembros del labo que generan (o generaron) un excelente lugar

de trabajo: Rocha, Tincho lungman, Carchu, Rita, Lucha Gomez Acuña, Anita Fiszbein, Alejo Prego. No me quiero olvidar de agradecer especialmente a los altos del labo: Guilli, Fefuwi, Lucho y Marquitos, gracias por alcanzarme todo lo que se encuentra en los estantes superiores.

A muchas personas por fuera del grupo "ARKAS" que facilitan el día a día. A Amaranta Avendaño por su buena onda, por las charlas, por los miles de reactivos que me ha preparado, por el millón de maxi que preparo, por ser hiper generosa y buena gente. Al gran Silvio Traba por los mates increíbles (y los creíbles), por estar siempre con ganas de ayudar y dar una mano, por entender que cuando digo 5 minutos pueden ser 5 horas, por ser incondicional. A Vikuz Medina, Martu Garcia Sola, Daiu Sapochnik ("los Coso") por su generosidad excesiva, por los mil mates, por fomentar la unión entre grupos, por ayudar con cualquier detalle, por mirar los problemas desde otra perspectiva. A Agus Romero por ayudarme con el microscopio, por explicar todo desde la física, por contagiar entusiasmo. Eve Olszanowski, Adriana "Flavia" Camarano, Mica Stedile y Sabri Vallone por prestarme reactivos, protocolos, o ayudarme en general cuando se los pedí, por su buena onda conmigo en todo momento. A les Becaries Pecci, Kordon, Colman, Aguilar, por generar un excelente ambiente laboral y de compañerismo. A los demás integrantes del IFIBYNE pasando por seguridad, becaries, técnicos, responsable de limpieza, jefes de grupo. No quiero dejar de nombrar a: Orlandito Andrada, Jose "Pepe" Clemente, Esteban Cardinaux, Juana Ponzinibbio, Anto, Marie Veggeti, Ale Lescano, Paula Felman y Dani de Luca por las charlas de pasillo y por estar siempre que les pedí un favor.

A Sabri Beltrame y Luqui Bayones por enseñarme a hacer los primeros experimentos en un laboratorio, por explicarme como es la vida de un científico (lo bueno y lo no tan bueno) y por mostrarme que se puede tener de amigos a los compañeros de trabajo. A los integrantes de los laboratorios de Shona Murphy y Maria Carmo Fonseca por hacerme sentir uno más cuando estuve trabajando con ellos.

A la Universidad de Buenos Aires por brindarme educación de calidad. Al CONICET por financiarme el doctorado. A Alexandra "la Rusa de Sci-hub" Elbakyan por darme acceso a muchos artículos científicos.

A mis amigos de la facu, Mariano Martin, Nacho Muñoz, Tomi Righetti, Adri Botto, Juli Tarsitano y Mati Zanon, por generar catarsis en los momentos complicados, por las charlas de todo tipo (incluido científicas). A Seba Mildiner por ser un gran amigo, gracias por ser una persona que la distancia y el tiempo no influye en su forma de ser, gracias por los consejos. A mis amigas de la Facu, Agustina Birba, Daniela Ruiz, Patricia Gonzalez y Tatiana Bortolato, por aportar miradas distintas a cualquier tema y ser tan buenas personas conmigo.

A mis amigos de la vida, a esos que no saben bien cuál es mi trabajo, ni que estudie. Gracias por estar cuando los necesite, por su consejo, porque cuando nos vemos volvemos a charlar de cualquier cosa (no incluye científicas) como si nos hubiésemos visto el día anterior. Gracias por ser mis amigos a: Jonathan “el Jony” Romero, Pedro “el Chapa” Cantisani, Leonardo “el Gusa” Bauman, Francisco “Pacho” Vasquez, Dulce Rodriguez, Lautaro “el Tronco” Ledesma, Santiago “el Pera” Peranio y por último, pero no menos importante, Fede Buselli.

A mi familia que son un gran sostén. A Marta Pereyra, mi mamá, que me pudo bancar tantos años en Buenos Aires, me enseñó que es importante trabajar de lo que te gusta, y trabajar a pesar de todo, que nunca hay que dejar de estudiar. Gracias por ser el ejemplo de una persona optimista, justa, bondadosa, trabajadora y honesta. A mi hermano y mis hermanas, Fausto Bragado, Atali Bragado y Francina Bragado, por querer ayudarme siempre y tratar de entender que es lo que hago. No quiero dejar de agradecer a la Fran y Juanga Geuna por compartir al pequeño gigante Gino siempre que pueden. A Fabian Bragado (mi papa), Teo Bragado (mi hermanito) y Pia Bragado (mi hermanita), que me han facilitado innumerables cosas a lo largo de estos años.

A la persona que me ayudo en todo momento a lo largo de estos años, una mujer increíble e incondicional Daniela Galvan. Gracias por ser mi pareja, acompañarme, compartir la vida, por esperarme con la comida hecha (un millón de veces habrá ocurrido), bancarme en las mil horas del labo, por enseñarme a ver la vida de otra manera, bancarme de mal humor porque tal o cual experimento no salía. El trabajo no hubiese sido posible sin tu ayuda. Gracias por tanto amor y por estar siempre Dani!

INDICE

INTRODUCCION	13
Función de los ARN pequeños nucleares	13
Variantes de snRNAs	14
Características de los genes de los snRNAs	14
Proteínas involucradas en la transcripción de los snRNA	16
Factores de iniciación de la transcripción	17
Factores de elongación	19
Factores de procesamiento 3'	20
Ensamblado de las partículas ribonucleoproteicas pequeñas nucleares o snRNPs	25
Modificación post traduccional de proteínas: conjugación de péptidos de la familia de Ubiquitina	27
SUMO	29
Vía de conjugación de SUMO	30
Secuencias blanco de SUMOilacion	32
Proteasas específicas de SUMO	34
Consecuencias de la SUMOilacion	35
SUMO y biogénesis de los snRNAs	37
OBJETIVOS	39
MATERIALES Y METODOS	41
Cultivo Celular	41
Mantenimiento de líneas celulares:	41
Transfecciones de ADN	41
Transfecciones con ARN	42
Silenciamiento mediante CRISPRi	42
Ensayo de apoptosis	43
Tinción con yoduro de propidio y anexina	43
Tinción con cristal violeta	43
Técnicas de interacción entre macromoléculas	44
Purificación de proteínas conjugadas a 6xHis-SUMO	44
Co-inmuno precipitación de proteínas	45
Pull-down nativo con resina de Niquel	46
Inmunoprecipitación de la cromatina (CHIP)	47
Western Blots	48
Microscopía de Fluorescencia	49

Preparación de plásmidos	49
Transformación de bacterias.....	49
Preparación de plásmidos a pequeña escala o minipreps	50
Preparación de plásmidos a gran escala o maxipreps.....	50
Cuantificación de plásmidos	50
Mutagénesis dirigida	51
ARN, RT & PCR	52
Purificación de ARN	52
Transcripción reversa (RT).....	53
PCR en tiempo real (qPCR)	53
Lista de primers utilizados:	55
RESULTADOS	56
La disminución global de SUMOilación aumenta los niveles de snRNAs nacientes	56
Proteínas involucradas en la biogénesis de los snRNA son modificadas por conjugación de SUMO2	59
SNAP43.....	62
Las lisinas 245 y 333 son sitios de conjugación de SUMO dentro de la proteína SNAP43.....	62
La proteína mutante SNAP43 2KR es incapaz de estimular la transcripción de los snRNAs	65
La disminución de SUMOilación de SNAP43 no altera su reclutamiento a la cromatina	69
La sobre-expresión de SNAP43 2KR induce la muerte celular en células HeLa.....	70
INTS9	74
La lisina 607 es sitio de conjugación de SUMO en la proteína INTS9, miembro del complejo <i>Integrator</i>	74
La disminución de SUMOilación de INTS9 no modifica su capacidad para procesar el snRNA naciente en su región 3'	75
INTS11	77
Las lisinas 381, 462 y 475 son sitios de conjugación de SUMO en la proteína INTS11, miembro del complejo <i>Integrator</i>	77
La proteína mutante de SUMOilación INTS11 3KR es incapaz de clivar el snRNA naciente en su extremo 3'	80
La disminución de SUMOilación de INTS11 no impide su interacción con INTS9 e INTS483	
La disminución de la SUMOilación de INTS11 interfiere con su interacción con INTS13 e INTS14	86
La SUMOilación de INTS11 regula su localización sub-celular	88
DISCUSION.....	97
Resumen de los resultados obtenidos.....	97

Los niveles de SUMOilacion son importantes para la biogénesis de los snRNAs.....	98
Las proteínas involucradas en la biogénesis de los snRNAs son sustrato de conjugación de SUMO	99
¿Es posible obtener mutantes incapaces de ser SUMOiladas?.....	99
Consecuencias de la SUMOilación	101
Complejo SNAPc	101
Complejo <i>Integrator</i>.....	103
BIBLIOGRAFIA.....	106

INTRODUCCION

Función de los ARN pequeños nucleares

En las células animales existen tres ARN polimerasas, la ARN polimerasa I es responsable de transcribir el ARN ribosomal (Moss et al. 2007), mientras que la ARN polimerasa III transcribe diversos ARN no codificantes, entre ellos los ARN de transferencia, los ARN pequeños nucleares 7SK, U6 y U6 atac (White et al. 2011). Por otro lado, la ARN polimerasa II (RNAPII) es la encargada de transcribir los ARN mensajeros (mRNAs), los ARN largos no codificantes (lncRNAs), diversos micro ARN (microRNAs) y los restantes ARN pequeños nucleares (snRNAs) (Lee et al. 2004; Jawdekar & Henry 2008; Shi et al. 2013; Srivastava & Ahn 2015; Guiro & Murphy 2017; Chen et al. 2018).

Los snRNA tienen menos de 350 nucleótidos de longitud, son ricos en uracilos y están asociados con proteínas formando ribonucleopartículas pequeñas nucleares llamadas “snRNPs”. Las funciones biológicas de los snRNAs están bien caracterizadas desde hace años. Los snRNAs U1, U2, U4, U5 y U6 forman el spliceosoma mayor, una maquinaria ribonucleoproteica, que se ensambla sobre los intrones canónicos (conocidos por comenzar con secuencias 5' GT y terminar en su extremo 3' con AG) de los ARN mensajeros inmaduros (pre-mRNAs) posibilitando que sean escindidos de sus exones flanqueantes con la precisión de un nucleótido y los exones son posteriormente religados para dar lugar a ARNm maduros, proceso conocido como *splicing*. Existen otro tipo de intrones, minoritarios, que están flanqueados por secuencias AT/AC y son capaces de reclutar el spliceosoma menor compuesto por los snRNAs U11, U12, U4 atac, U5 y U6 atac snRNPs, para realizar el proceso de *splicing* (Turunen et al. 2013). El spliceosoma no se encuentra pre-formado como tal, sino que se ensambla de *novo* para cada intrón que deba ser escindido, lo cual va ocurriendo a medida que el pre-mRNA emerge de la maquinaria transcripcional comandada por RNAPII, en lo que se conoce como *splicing* co-transcripcional (Herzel et al. 2017).

Por otro lado, el snRNA U7 es requerido para el procesamiento 3' de los pre-ARNm que codifican para las histonas no replicativas, proceso que también ocurre en forma co-transcripcional, mientras que el sn(o)RNA U3 es necesario para el

procesamiento de ARN ribosomal (Culver 2002). La función del snRNA 7SK es regular los niveles de P-TEFb activo, un complejo proteico compuesto por la quinasa CDK9 y la ciclina T1, capaz de fosforilar la región C-terminal de la RNAPII, controlando así la elongación de la RNAPII (Michels et al. 2004).

En esta tesis nos enfocaremos en los snRNAs transcritos por la RNAPII y que conforman el splicesosoma mayoritario: U1, U2, U4 y U5.

Variantes de snRNAs

Los genes que codifican para snRNAs suelen estar en múltiples copias, por ejemplo, el gen de U1 tiene cuatro copias en tándem en el cromosoma 1 y el gen U2 poseen quince copias en tándem en el cromosoma 17. Además, se han detectados pseudogenes para snRNAs (Marz et al. 2008) pero en los últimos años, se observó en distintos tipos celulares que algunas variantes de U1, que habían sido descritas como pseudogenes inactivos, poseen marcas activas de transcripción en sus promotores, comprobándose además que se expresan, que son necesarios para el *splicing* de un determinado set de ARN mensajeros, (Kyriakopoulou et al. 2006; O'Reilly et al. 2013; Guiro & O'Reilly 2015; Vazquez-Arano et al. 2016). Por otro lado, han sido descritas distintas variantes de U5, cinco en humanos y siete en *Drosophila melanogaster*, expresadas diferencialmente a lo largo del desarrollo y cada variante es capaz de formar parte de snRNPs funcionales y participar en el proceso de *splicing* (Sontheimer et al 1992; Chen et al. 2005).

Características de los genes de los snRNAs

A pesar de que la función de los snRNAs está bien caracterizada, la regulación de su expresión no está completamente comprendida aún. Vale la pena mencionar que los niveles de snRNAs en los humanos varía entre los diferentes órganos, entre tejido sano y tejidos correspondientes a diferentes tipos de cáncer (Dvinge et al. 2019). Además, fue demostrado que es posible disminuir la abundancia de los snRNAs en células en cultivo, mediante la transfección de ASOs (oligonucleótido antisentido) que tienen como blanco los snRNAs, y al alterar la abundancia de los snRNAs se genera un cambio en los patrones globales de *splicing* alternativo (Dvinge et al. 2019).

La estructura del promotor que dirige la expresión de los snRNAs humanos presenta características que la distinguen de aquella correspondiente a los promotores a partir de los cuales se transcriben los ARN mensajeros y de otros ARN no codificantes transcritos por la RNAPII: posee una secuencia específica llamada *proximal sequence element* (PSE), comparable a la "TATA box", y una región distal llamada *distal sequence element* (DSE) que actúa como *enhancer* (Matera & Wang 2014). El PSE es reconocido por un complejo denominado *small nuclear activator complex* (SNAPc), y el DSE recluta a los factores de transcripción Oct-1/POU2F1 y ZNF143 (Murphy 1997; Yoon et al. 1995). Si la región promotora además posee una región TATA box río abajo del PSE, como es el caso de U6, el snRNA es transcrito por la ARN polimerasa III con sus factores accesorios para la elongación y procesamiento 3', lo que indica que las regiones PSE y DSE regulan la transcripción de snRNAs mediada por ambas ARN polimerasas (Hernandez 2001). Los transcriptos correspondientes a los snRNAs se encuentran modificados en su extremo 5' al igual que los ARNm, es decir poseen un "cap", pero no contienen intrones ni son poliadenilados. Los genes de los snRNAs que son transcritos por la RNAPII poseen en su extremo 3' una región llamada *3' box* que es necesaria para el procesamiento del ARN naciente y en el gen se encuentra ubicada a unas 9-19 bases río abajo de lo que correspondería a la secuencia madura del snRNA (CDS)(Figura I1). La secuencia del *3' box* es GUUUN₀₋₃AAARNNAGA, donde N es cualquier base y R es una purina, tiene una extensión entre 16-20 nucleótidos y durante su transcripción es reconocido por un complejo multiproteico conocido como *Integrator complex* que procesa el ARN naciente (Baillat et al. 2005), únicamente cuando el transcripto proviene de un promotor de snRNA. Este último dato indica la importancia de los factores de transcripción y elongación específicos para el reconocimiento del *3' box* y por lo tanto para el procesamiento 3' de los snRNAs (FIGURA I1).

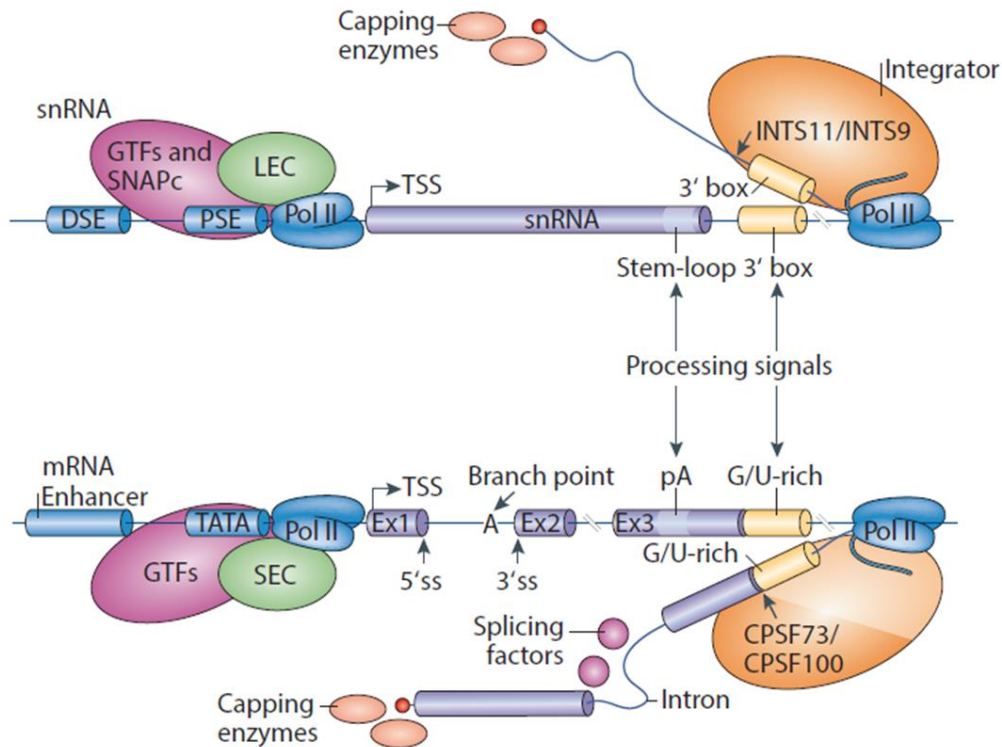


Figura I1. Esquema de transcripción de los snRNAs y mRNAs transcritos por ARN polimerasa II. El distal sequence element (DSE) actúa como un *enhancer* de la transcripción, el proximal sequence element (PSE) es el promotor específico de los snRNAs, comparable con el “TATA box”. Pol II es RNAP II, GTF indica los factores generales de la transcripción, SNAPc es un factor de transcripción específico de los snRNAs, LEC es el complejo de elongación necesario para la expresión de los snRNAs, mientras que el SEC es el complejo de elongación para la expresión de los mRNAs. Los snRNAs no poseen intrones, ni señal de corte y poliadenilación (pA), utilizan la secuencia conocida como 3' box, la cual es reconocida por el *Integrator complex* que, a través de su subunidad catalítica INTS11, procesa el extremo 3' de los snRNAs inmaduros o pre-snRNAs. Adaptado de Matera & Wang 2014.

Proteínas involucradas en la transcripción de los snRNA

La RNAPII es una enzima multimérica compuesta por 12 subunidades. Rpb1 es la subunidad mayor y en la que reside la actividad catalítica, cuenta con un dominio carboxilo terminal (CTD) que consiste en 52 repeticiones de la heptada YSPTSPS. Durante la transcripción, la fosforilación, entre otras modificaciones post traduccionales, del CTD favorecen o impiden el reclutamiento de factores, por lo cual se define el CTD como “plataforma de interacción” (Muñoz et al. 2010; Srivastava & Ahn 2015; Zaborowska et al. 2016). La serina 7, dentro de la heptada arriba

mencionada, se encuentra fosforilada durante la expresión de todos los genes, pero resulta esencial para la expresión de los snRNAs (Egloff et al. 2007). Tanto la serina 7 como la serina 5 son fosforiladas por la quinasa dependiente de ciclina 7 (CDK7), subunidad del complejo TFIIH (Akhtar et al. 2009), y luego del inicio de la transcripción, la serina 7 media el reclutamiento de la fosfatasa RPAP2 la cual desfosforila la serina 5. CDK9 fosforila la serina 2 de las heptadas, dentro del CTD, y tanto la fosforilación de la serina 7 como la de la serina 2 son necesarias para el reconocimiento del 3' *box* (Egloff et al 2010). Ciertas quinasas parecen cumplir algún rol durante la biogénesis de los snRNAs dado que inhibidores de CDK9 afectan el procesamiento 3' de los snRNAs pero no su expresión (Cuello et al. 1999; Medlin et al. 2005), mientras que la ausencia de CDK12, otra quinasa capaz de fosforilar la serina 2, afecta la transcripción de los snRNAs (Bartkowiak et al 2010; Liang et al 2015). Una vez iniciada la transcripción, tiene lugar el agregado del *cap* (*capping*) en el extremo 5' de los transcritos nacientes: en los ARNm la maquinaria involucrada es reclutada a dicho extremo en forma co-transcripcional y es activada alostéricamente al reconocer la fosforilación de la serina 5. Al igual que para los ARNm, el *capping* ocurre co-transcripcionalmente en la síntesis de snRNAs, posiblemente mediada por la fosforilación de la serina 5, (Guiro & Murphy 2017).

Factores de iniciación de la transcripción

Los factores de transcripción Oct1, SP1, NF1 y Staf se unen a las secuencias DSE, que actúa como *enhancer*. Oct1 estabiliza la unión del complejo SNAPc al PSE a través de una interacción directa (Jawdekar & Henry 2008). SNAPc recluta a la proteína *TATA-binding protein* (TBP) y a los factores asociados a TBP (TAFs), que incluyen a TAF5, TAF6, TAF8, TAF9, TAF11 y TAF13. Este subset de TAFs/TBP también se encuentra involucrado en la transcripción de genes codificantes dependiente de RNAPII. Sin embargo, no todos los genes de los snRNAs reclutan los mismos factores, aún cuando son transcritos por la misma ARN polimerasa. Por ejemplo, el gen U2 no recluta al TAF7, pero sí es reclutado a los genes U1, U4, U5 y U11 (Zaborowska et al. 2012). TAF7 es capaz de reclutar al complejo multi proteico *Mediator complex*, el cual es necesario para una transcripción eficiente de los snRNAs, al igual que de los genes codificantes. Se desconoce de qué manera el *Mediator complex* es reclutado al gen U2. Al igual que

para los genes codificantes, para la transcripción de los snRNAs son necesarios los factores generales de la transcripción TFIIA, TFIIB, TFIIE, TFIIF y TFIIH (Kuhlman et al. 1999; Glover-Cutter et al. 2009; O'Reilly et al 2014).

Small nuclear activator complex (SNAPc) es un complejo de cinco subunidades (SNAPC1 a SNAPC5) que actúa como factor de transcripción basal de todos los snRNAs, es capaz de reconocer la región PSE (Figura 11) y favorece el ensamblado de los complejos de iniciación y elongación (Cuello et al. 1999; Kuhlman 1999; Jawdekar & Henry 2008; Canella et al. 2010) (Figura 12). SNAPC4 (o SNAP190) la proteína más grande del complejo se une al PSE directamente a través de un dominio Myb de unión al ADN, aunque podría no ser la única dado que la subunidad SNAPC3 (o SNAP50) es capaz de unirse a ADN en ensayos *in vitro*, y en *Drosophila melanogaster*, el homólogo de SNAPC1 (o SNAP43) es capaz de interactuar en forma directa con el ADN (Kim et al. 2010; James Faresse et al. 2012). En el ensamblado del complejo se demostró que SNAPC1 y SNAPC5 (o SNAP19) se asocian directamente al extremo N-terminal de SNAPC4, mientras que SNAPC2 (o SNAP45) se asocia por su C-terminal (James Faresse et al. 2012). Por otro lado, SNAPC1 sirve como andamiaje para la interacción de SNAPC3 con SNAPC4, lo que permitió describir la formación de un complejo mínimo SNAPc que incluye a SNAPC1, SNAPC3 y la región N-terminal de SNAPC4, que es capaz de reclutar a TBP y de activar la transcripción (Mittal et al 1999; James Faresse et al. 2012). Por otra parte, SNAPC1 es capaz de interactuar con las proteínas Rb y p53, y se postula que estas interacciones regulan la expresión de los snRNAs durante el ciclo celular (Hirsch et al. 2000; Gridasova et al. 2005).

Más allá de su participación en la transcripción de los snRNA, por ensayos de inmunoprecipitación de la cromatina (ChIP) y secuenciación masiva, se reportó a SNAPC1 como un factor que se encuentra asociado a genes codificantes, en particular aquellos de respuesta temprana al factor de crecimiento epidermal (EGF) (Baillat et al. 2012). Sin embargo, otros trabajos que utilizaron las mismas metodologías y fueron publicados en el mismo año, lo han reportado principalmente asociado a los genes de los snRNAs (James Faresse et al. 2012).

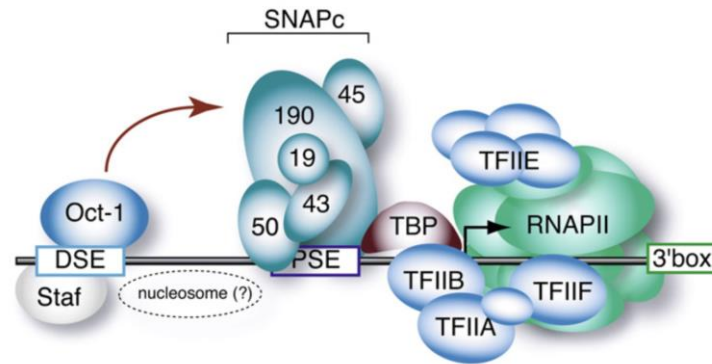


Figura 12. Factores involucrados en la transcripción de los snRNAs. En la región DSE se unen los factores de transcripción Oct-1 y Staf. En el PSE se une el complejo SNAPc, formado por cinco proteínas, encargado de reclutar a TBP y RNAPII con sus factores asociados. SNAPC1 = SNAP43, SNAPC2 = SNAP45, SNAPC3= SNAP50, SNAPC4= SNAP190 y SNAPC5= SNAP19. Obtenido de Jawdekar & Henry 2008.

Factores de elongación

La función del complejo mediador o *Mediator complex*, (compuesto por 26 subunidades), ha sido bien caracterizada para los genes codificantes transcritos por la RNAPII: actúa como un nexo entre los factores de transcripción y la maquinaria basal de transcripción compuesta por los factores generales de la transcripción y la RNAPII, favoreciendo el ensamblado del complejo de pre-iniciación formado por la proteína TBP y sus factores asociados (Chen et al. 2018). Por su parte, el reclutamiento del *Mediator complex*, mediado por el factor TAF7, a los genes de snRNAs ha sido descrito más recientemente. Se observó que la subunidad Med26, en la región promotora de los snRNAs, intercambia a la proteína TAF7 por el factor de elongación *Little elongation complex* (LEC), permitiendo el inicio de la transcripción y la elongación de dichos genes (Takahashi et al. 2015).

Es importante mencionar que luego del inicio de la transcripción, en diversos genes codificantes opera un control temprano de la elongación (EEC por su sigla en inglés). Lo descrito hasta el momento es que los factores de elongación negativos DSIF (heterodímero compuesto por las proteínas Spt5 y Spt4) y NELF (compuesto por NELF-A, NELF-B, NELF-C/D y NELF-E) generan el pausado de la RNAPII. La quinasa CDK9, subunidad de P-TEFb, fosforila la serina 2 del CTD de la RNAPII, a NELF que al ser fosforilado deja de interactuar con la RNAPII, y a la región C-terminal de Spt5 (DSIF),

análoga al CTD, lo cual convierte al represor DSIF en un activador de la elongación, liberando así el pausado de la RNAPII. Posteriormente, es reclutado un complejo llamado *super elongation complex* (SEC) compuesto por los factores de elongación ELL, AFF1, AFF4, AF9, ENL y P-TEFb (Smith et al. 2011; Jonkers & Lis 2015; Chen et al. 2018). Los genes de los snRNAs, por su parte, tienen un factor de elongación alternativo, y específico llamado *little elongation complex* (LEC), contiene a ELL, ICE1, ICE2, y ZC3H8 (Figura I1) (Hu et al 2013).

Dentro de los componentes del LEC, el factor ELL es necesario para la expresión de los snRNAs. ICE1 actúa como andamiaje para la formación del LEC y es la proteína encargada de reclutar todo el complejo a los *loci* de los snRNAs. Hasta la fecha, se desconoce la función de ICE2 y ZC3H8. El snRNP 7SK, compuesto por el snRNA 7SK, Larp7 y MePCE, como fue mencionado anteriormente es capaz de secuestrar a P-TEFb e inactivarlo, razón por la cual es conocido principalmente como un regulador negativo de la elongación. Sin embargo, se observó que es necesario para el reclutamiento y ensamblado del LEC y la correcta expresión de los snRNAs, por lo tanto, ha sido propuesto como parte del LEC (Egloff et al. 2017).

Recientemente, se ha reportado que el complejo *Mediator* está involucrado no sólo en el reclutamiento del LEC a los genes de snRNAs, sino también a los genes codificantes de las histonas no replicativas, que al igual que los snRNAs no se poliadenilan y su procesamiento 3' es llevado a cabo por un complejo proteico específico llamado *heat-labile factor* (HLF). Además, fue demostrado que el complejo LEC está involucrado en el reclutamiento de los factores de procesamiento 3' tanto de los transcritos de histonas como de los snRNAs (Takahashi et al. 2020). Sin embargo, no está del todo esclarecido como el complejo LEC recluta dos complejos de procesamiento 3' diferente.

Factores de procesamiento 3'

El *Integrator complex* es un mega complejo multiproteico de más de un millón de daltons de peso molecular, compuesto por 14 subunidades (INTS1-INTS14), que van desde los 49 hasta los 244 KDa, y necesario para el procesamiento 3' de los snRNAs transcritos por la RNAPII (Baillat & Wagner 2015). La depleción de diversas

subunidades de este complejo genera acumulación de snRNAs nacientes sin procesar, llevando consecuentemente a una poliadenilación anómala de estos transcritos (Yamamoto et al. 2014). Durante el inicio de la transcripción, la quinasa CDK7 (perteneciente al factor general de la transcripción TFIIF) fosforila la serina 5 y la serina 7 del CTD de la RNAPII. La fosforilación en la serina 7 favorece el reclutamiento de la fosfatasa RPAP2, quien desfosforila la serina 5 recluta a las subunidades INTS1, INTS4, INTS5, INTS6 e INTS7. Durante la elongación transcripcional, la serina 2 del CTD de la RNAPII es fosforilada por CDK9. Se determinó que la fosforilación de serina 2 y serina 7 es necesaria para el reclutamiento de INTS9 e INTS11, postulándose que el *Integrator complex* es un complejo modular que se va ensamblando durante la transcripción mediada por el CTD de la RNAPII (Figura 13), y que algunas subunidades cumplirían funciones al inicio de la transcripción mientras que otras durante el procesamiento 3' (Egloff et al. 2010; Egloff et al. 2012). La idea de que el *Integrator* es realmente un complejo modular está respaldada por el hecho que distintas subunidades del mismo forman módulos independientes. El módulo catalítico fue el primero en describirse, demostrándose que INTS4 actúa como *scaffold* para la unión de INTS9 e INTS11 de manera estable sin necesidad de ninguna otra subunidad del complejo. Dicho módulo, formado por INTS4, INTS9 e INTS11 es el responsable del clivaje endonucleolítico de los snRNAs nacientes (Albrecht et al. 2018). Además, las subunidades INTS10, INTS13 e INTS14 forman un módulo independiente, que dentro del complejo es capaz de unirse al ARN naciente y estabilizar la interacción del módulo catalítico con el resto del complejo (Sabath et al. 2020). Por otro lado, las subunidades INTS3 e INTS6 pertenecen a otro complejo que reconoce daño al ADN, llamado sensor *of single stranded DNA* (SOSS) (Li et al. 2009; Zhang et al. 2013). Dado que los genes de los snRNAs están altamente expresados y que al menos el gen de U2 no posee nucleosomas (Egloff et al. 2009), se sugiere que INTS3 e INTS6 disminuyen el daño al ADN en la región correspondiente a estos genes, lo cual mantiene una cromatina laxa favoreciendo la expresión (Baillat & Wagner 2015).

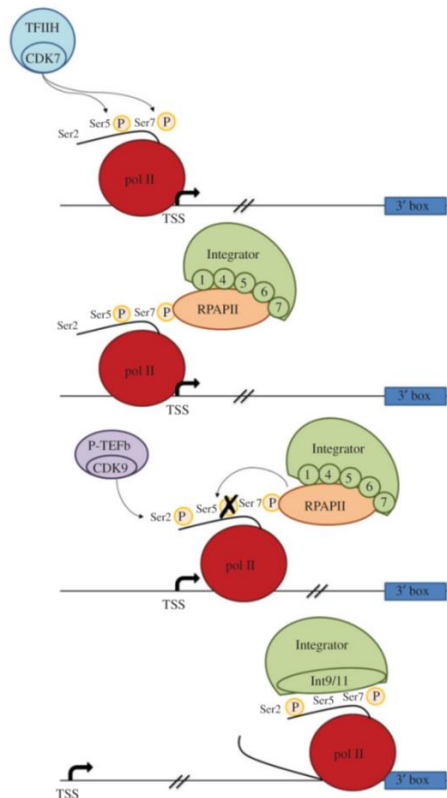


Figura 13. Reclutamiento del *Integrator* a los genes de snRNAs. Una vez iniciada la transcripción, el CTD de RNAPII es fosforilado en la serina 5 y la serina 7 por CDK7. La serina 7 fosforilada recluta la fosfatasa RPAPII, que desfosforila la serina 5 y favorece el reclutamiento de las subunidades INTS1, INTS4, INTS5, INTS6 e INTS7. Finalmente, CDK9 fosforila la serina 2, que junto a la serina 7 fosforilada es reconocida por las subunidades INTS9 e INTS11 que procesan el extremo 3' del snRNA naciente. Obtenido de Guiro & Murphy 2017.

La mayoría de las subunidades del *Integrator* comparten poca identidad con otras proteínas y se desconoce su función. Sin embargo, se encontró que INTS11 es un parálogo de CPSF-73, el cual forma parte del complejo de clivaje y poliadenilación (CPSF), responsable del clivaje 3' de ARNm poliadenilados. CPSF-73 e INTS11 pertenecen a la superfamilia de las proteínas metalo- β -lactamasa, el dominio metalo- β -lactamasa de cada proteína contienen aminoácidos conservados que coordinan la catálisis del ARN. Poseen un dominio β -CASP dentro del dominio metalo- β -lactamasa, el sitio activo se encuentra en la interfase entre ambos dominios. El dominio metalo- β -lactamasa (residuos: 1 al 450) y el β -CASP (residuos 207 al 390) de INTS11 poseen un 40% de identidad de secuencia con los dominios de CPSF-73, indicando la similitud e importancia de estos dominios (Baillat et al. 2005). La subunidad INTS9 es similar a la

proteína CPSF-100, y ambos poseen los mismos dominios que INTS11 y CPSF-73, pero no conservan los aminoácidos catalíticos y la función que ejerce cada una dentro de sus respectivos complejos aún permanece desconocida. Además de los dominios mencionados, las proteínas INTS9, INTS11, CPSF-73 y CPSF-100 poseen un dominio carboxilo terminal (CTD), pobremente conservado entre ellos. INTS11 e INTS9 forman un heterodímero mediante la interacción de sus respectivos CTDs y estabilizado por INTS4 (Albrecht et al. 2012; Yixuan Wua et al. 2017; Albrecht et al. 2018).

En los genes de los snRNAs, el complejo DSIF y NELF cumplen funciones distintas a las mencionadas anteriormente para genes codificante. En los genes de snRNAs, DSIF es capaz de unirse a la RNAPII antes de iniciar la transcripción, y en ausencia de la subunidad Spt5 disminuye el reclutamiento del resto del complejo, la RNAPII y el complejo *Integrator* a los promotores. Por otro lado, el complejo NELF es reclutado por el *Integrator*, y en su ausencia es reclutada la proteína Cstf64 a los genes, generando una poliadenilacion anomala. Se cree que la función del complejo DSIF es reclutar al *Integrator* al promotor de los snRNAs, mientras que NELF está encargado de favorecer el reconocimiento del 3' *box* y consecuentemente el adecuado procesamiento 3' de los transcritos (Yamamoto et al. 2014).

El complejo *Integrator* fue descubierto inicialmente como el encargado de procesar los snRNAs nacientes (Baillat et al. 2005). Sin embargo, es capaz de regular la expresión y procesamiento 3' de otros ARN no codificantes, diferentes a los snRNAs, tales como los transcritos provenientes de enhancers (eRNAs) y de transcritos que utilizan el mismo promotor que un mRNA pero se transcriben en dirección contraria (PROMPTs) (Lai et al. 2015; Nojima et al. 2018). Los eRNAs favorecen la interacción entre un *enhancer* y su promotor *target* (Hsieh et al. 2014; Lai et al. 2015). Además, un trabajo reciente mostro que la expresión del eRNA-SphK1 puede evitar el reclutamiento de factores negativos de la transcripción favoreciendo la expresión de su gen *target* SPHK1 (Blank-Giwojna et al. 2019). Tanto los PROMPTs como los eRNAs, han sido clasificados como ARN no codificantes largos (lncRNA), la mayoría son mono exónicos y capaces de formar híbridos ARN-ADN, llamado R-loops, los cuales pueden inducir daño al ADN e incluso pueden generar senescencia celular (Nojima et al. 2018).

Además, los R-loops son capaces de facilitar la síntesis de genes cercanos (Tan-Wong et al. 2019).

En los últimos años, se ha descrito al complejo *Integrator* regulando la expresión de distintos genes codificantes, entre ellos los de expresión inmediata temprana (immediate early genes o “IEGs”) inducidos por EGF. Luego del estímulo extracelular, INTS11 es fosforilado por ERK1/ERK2 y es reclutado, junto a las otras subunidades, a los promotores de los IEGs siendo necesario para la transcripción de estos genes (Gardini et al. 2014; Yue et al. 2017). Por otro lado, ha sido reportado que el complejo *Integrator* puede interactuar con NELF y esto regula su capacidad para detener la transcripción. En este mismo trabajo, se observó que algunos genes regulados por NELF presentan secuencias similares al 3' *box* en su región 3' y que en ausencia de INTS11 su elongación disminuye (Stadelmayer et al. 2014). Además del pausado proximal de la RNAPII, diversos genes codificantes (entre ellos los que codifican para factores reparadores del daño al ADN, reguladores transcripcionales, supresores de tumores, y otros), poseen un procesamiento 3' prematuro (cerca al inicio de la transcripción) siendo rápidamente degradados, limitando así la elongación de la RNAPII e impidiendo la producción de los ARN maduros correspondientes (Kamieniarz-Gdula & Proudfoot 2019). En *Drosophila*, el complejo que procesa prematuramente al ARN naciente es el *Integrator*. Llamativamente, en humanos se observó que muchos genes IEGs, en ausencia de EGF, son procesados prematuramente impidiendo su transcripción completa. De esta manera, en presencia del estímulo EGF, el *Integrator* es necesario para la expresión de estos genes mientras que, en ausencia del estímulo, impide la transcripción de los mismos (Elrod et al. 2019; Tatomer et al. 2019; Tatomer & Wilusz 2020). Además, el *Integrator* es capaz de reclutar a la fosfatasa PP2A que desfosforila el CTD de la RNAPII favoreciendo la atenuación transcripcional (Huang et al. 2020; Vervoort et al. 2021). Si bien se considera al complejo *Integrator* como un atenuador general de la transcripción, existen diversos genes codificantes que necesitan al *Integrator* para aumentar su expresión. Lo que ocurre en estos genes es que, una vez iniciada la transcripción, la RNAPII se detiene en los primeros nucleótidos y no elonga productivamente. El procesamiento de este RNA naciente por el complejo *Integrator* permite una elongación productiva por la RNAPII

(Beckedorff et al. 2020; Lykke-Andersen et al. 2021). En resumen, se ha postulado que por un lado, el complejo *Integrator* procesa el ARN naciente de genes que se expresan poco, manteniendo así su bajo nivel de expresión basal, como los PROMPTs, los eRNAs e IEGs en ausencia de estímulo extracelular. Por otro lado, el complejo *Integrator* favorece la transcripción de ciertos genes codificantes, al procesar RNA nacientes de RNAPII que no entraron en una elongación productiva (Figura I4) (Lykke-Andersen et al. 2021)

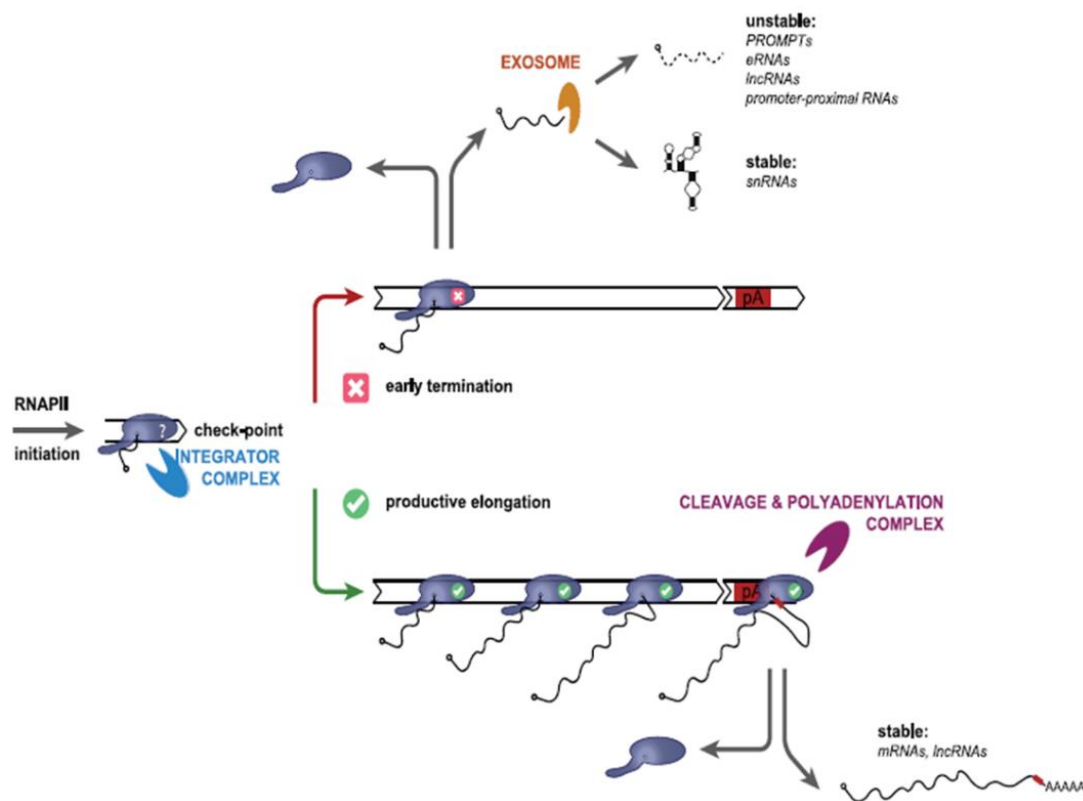


Figura I4. Rol dual del *Integrator* como atenuador y facilitador de la transcripción de la RNAPII. Una vez iniciada la transcripción, la RNAPII tiene un *check point* en el cual el Integrator procesa el transcrito naciente. Derivando en un desacople de la RNAPII del ADN y generando RNAs que luego serán degradados o generando una elongación productiva, y luego esos transcritos serán clivados y poliadenilados. Adaptado de Lykke-Andersen et al. 2021

Ensamblado de las partículas ribonucleoproteicas pequeñas nucleares o snRNPs

Los snRNAs luego de ser sintetizados y procesados por el complejo *Integrator* en su extremo 3', son exportados al citoplasma donde ocurren otros pasos de

maduración. Los snRNAs, al igual que los ARN mensajeros, poseen cap en el extremo 5'. Sin embargo, utilizan una maquinaria de exportación diferente a la utilizada por estos últimos, lo cual ha llevado a preguntarse cómo distingue la maquinaria de exportación los dos tipos de ARNs. Diversos trabajos han demostrado que no sólo importa el largo del transcripto, sino también con que proteínas de unión al ARN se asocian. Los transcriptos mayores a 300 nucleotidos son reconocidos por las proteínas hnRNP C1 y C2, las cuales reclutan a la proteína *nuclear RNA export factor 1* (NXF1, también conocida como *NTF2 related export1*) y asociados a NXF1 son exportados al citoplasma. (Masuyama et al. 2004; Fuke et al. 2008; McCloskey et al. 2012). Transcriptos más cortos, no son reconocidos por hnRNP C1 y C2 y son exportados por un set distinto de factores, entre ellos *cap binding protein 20* (CBP20) y *cap binding protein 80* (CBP80, *phosphorylated adapter RNA export* (PHAX30) y arsenite resistance 2 (ARS2). Dichas proteínas se unen al 5'cap de los snRNAs y la proteína *chromosome region maintenance 1* (CRM1; también conocida como *exportin 1*), permitiendo su exportación al citoplasma. Existen evidencias de que el ensamblado de este complejo de exportación ocurre en los cuerpos de Cajal (CBs por su sigla en ingles Cajal *bodies*) (Suzuki et al. 2010). Una vez en el citoplasma, se desensambla el complejo de exportación mediado por una desfosfrilación de PHAX30. El complejo SMN (*survival motor neuron*), el cual incluye a la proteína SMN y diversas proteínas asociadas de la familia GEMIN, regula la fase citoplasmática de los snRNAs (Massenet et al. 2002). El complejo SMN recluta siete proteínas Sm que forman un anillo alrededor de los snRNAs. El anillo de Sm estabiliza los snRNAs, permite un procesamiento exonucleolítico 3' a 5' en estos transcriptos, y los protege de nucleasas (Matera & Wang 2014). Luego, aún en el citoplasma, el complejo SMN recluta a la proteína *trimethylguanosine synthase 1* (TGS1), una metil transferasa que modifica el 5' del snRNA, cambiando la estructura del cap formando un 2,2,7 trimetil guanosina que sumado al anillo Sm actúa como una señal de localización nuclear. Esta señal es reconocida por *Snurportin* (SPN), que interactúa directamente con el receptor *importin-β* (Fischer & Luhrmann, R 1990; Fischer et al. 1993). Los snRNPs recién importados al núcleo se localizan transitoriamente en los CBs, donde intercambia la interacción con el complejo SMN por otras proteínas, dando lugar a snRNPs maduros,

competentes para realizar su función específica en el proceso de splicing (Nesic et al 2004; Schaffert et al 2004). El ensamblado de los snRNPs esta resumido en la figura 15.

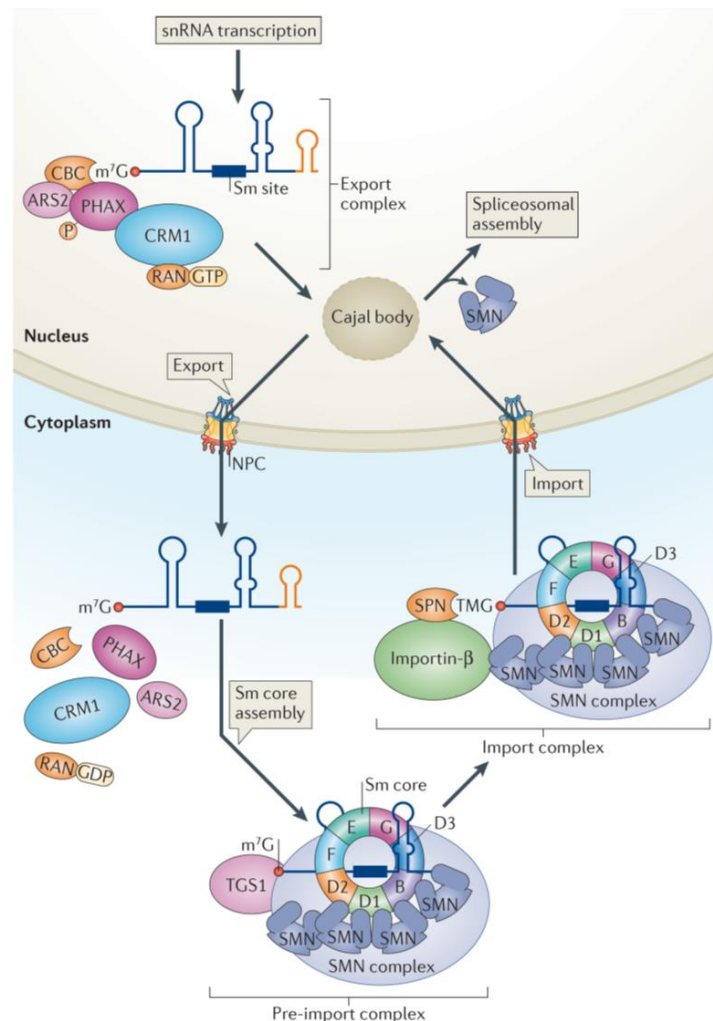


Figura 15. Ensamblado de snRNP. Luego de la transcripción, los snRNAs son reconocidos por un complejo proteico específico que media su exportación al citoplasma, donde sufren modificaciones y se asocian con las proteínas Sm, para luego ser re- importados al núcleo donde cumplirán su función en el proceso de splicing. Tomado de Matera & Wang 2014.

Modificación post traduccional de proteínas: conjugación de péptidos de la familia de Ubiquitina

Las proteínas pueden sufrir alteraciones químicas luego de ser traducidas, entre las cuales se encuentran: procesamientos proteolíticos, modificaciones de sus residuos por adición de grupos químicos, o bien la unión covalente (conjugación) de otros péptidos. En diversos contextos celulares estas modificaciones post-

traduccionales, o MPTs, pueden regular la localización, actividad, función y estabilidad de las proteínas blanco.

Dentro de la conjugación de otros péptidos, la MPT más conocida es la ubiquitinación, descubierta en 1975 como marca molecular para la degradación de proteínas a través de proteasoma, pero que hoy en día sabemos que puede regular diversas funciones proteicas. La ubiquitina (Ub), al igual que los miembros de su familia (Ubls por Ubiquitin-like proteins) son unidas covalentemente (conjugadas) a proteínas blanco por medio de cascadas enzimáticas, que involucra tres tipos de enzimas, E1 activadoras, E2 conjugadoras y E3 ligasas (Hay 2005). En humanos, se han identificado varias Ubls (Tabla I1), destacándose NEDD8 (Xirodimas 2008), ISG15 (Andersen & Hassel 2006; Pitha-Rowe & Pitha 2007; Ritchie & Zhang 2004), FAT10 (Groettrup et al. 2008), Atg8 y Atg12 (Geng & Klionsky 2008), y SUMO (Hochstrasser 2000).

Ub/Ubls	Saccharomyces cerevisiae	ID	Ct	E1	E2	E3	Protease	M/P	Substrate	Function/Comments
Ubiquitin	Ubiquitin	100%	Yes	Ube1/Uba6	> 37	> 600	~ 80	M/P	Thousands	Many, dependent on linkage
Nedd8	Rub1	58	Yes	NAE1-Uba3	Ubc12 (Ube2M), Ube2F	~10	CSN5, NEDP1	M/P?	Cullins, tumor suppressors, oncoproteins	Regulate Ub conjugation via cullinRING E3 ligases, cell cycle
MNSF8 (Fub1, Fau)	-	36	Yes						TCR α -like protein, Bcl-G, Endophilin II	Immunoregulatory role?
ISG15 (UCRP)	-	28/37	Yes	Ube1L	UbcH8, UbcH6	Herc5	UBP43	M	Viral and host proteins	Antiviral immunity, IFN-inducible
FAT10	-	27/36	No	Uba6	Use2				Use2	Ub-independent proteasomal degradation, immunoregulatory role?
Ufm1	-	23	Yes	Uba5	Ufc1	Ufi1	UfSP1, UfSP2		C20orf116	Erythroid and megakaryocyte development
SUMO1	Smt3	14	Yes	SAE1-SAE2	Ubc9	~15	SENP1,2	M	Hundreds	Alter interactions, localization, conformation
SUMO2	-	13	Yes	SAE1-SAE2	Ubc9	~15	SENP1,3,5,7	M/P		
SUMO3	-	13	Yes	SAE1-SAE2	Ubc9	~15	SENP1,3,5,7	M/P		
SUMO4	-	12							I κ B α ?	NF κ B signaling?, may be pseudogene or not processed
SUMO5										Nuclear bodies formation. Primates. Tissue-specific
Atg12	Atg12	12	No	Atg7	Atg10	-	-	M	Atg5, Atg3	Autophagy, mitochondrial homeostasis

Tabla I1. Lista de Ubls conocidas. Se detallan las proteínas de tipo Ubl junto con su % de identidad de secuencia aminoacídica y las enzimas E1, E2 y E3 específicas para la conjugación de cada una de ellas. Los porcentajes de identidad corresponden a los dos dominios de tipo Ub con respecto a ubiquitina humana madura. Ct se refiere al procesamiento C-terminal y M/P a la conjugación en forma de monómeros o polímeros (ver “Mecanismo de conjugación” más

abajo). Los espacios en blanco indican funciones o enzimas aún no asignadas. Tomado de van der Veen & Ploegh 2012.

SUMO

El gen de SUMO (Smt3 en levaduras) fue inicialmente identificado en *Saccharomyces cerevisiae* en un trabajo cuyo objetivo era la búsqueda de supresores de una proteína centromérica (Meluh & Koshland 1995). Se descubrió que SUMO era capaz de unirse covalentemente a otras proteínas, siendo la primera identificada RanGAP1 (Matunis et al. 1996; Mahajan et al. 1997). El proceso conocido como conjugación de SUMO, o “SUMOilación” es una modificación post-traducciona que consiste en la unión covalente del péptido denominado SUMO a una proteína blanco, en forma similar a la conjugación de Ub. Algunos organismos como *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Caenorhabditis elegans* y *Drosophila melanogaster* tienen un solo gen de SUMO. Otros como las plantas y animales vertebrados poseen varios genes de SUMO (Melchior 2000). Actualmente se conocen cinco genes en mamíferos que codifican para distintas isoformas de SUMO: SUMO1, SUMO2, SUMO3, SUMO4 y SUMO5. Todas ellas codifican para un polipéptido de aproximadamente 10 kDa, con una alta identidad de secuencia (Bayer et al. 1998; Mossessoiva et al. 2000). Las formas maduras de SUMO2 y SUMO3 comparten 97% de identidad de secuencia aminoacídica y ambas comparten sólo un 50% de identidad con SUMO1. La diferencia entre estas isoformas se ve reflejada en su funcionalidad, se cree que SUMO1 y SUMO2/3 cumplen funciones distintas *in vivo* porque son conjugadas a diferentes proteínas blanco (Wilson et al 2005; Vertegaal et al. 2006). A su vez, SUMO2/3 pueden formar cadena de poli SUMO *in vivo*, a diferencia de SUMO1 (Annemarthe et al. 2013) Con respecto a SUMO4, una mutación en su secuencia (sustitución M55V) ha sido asociada a diabetes tipo I (Guo et al. 2004). SUMO 5 fue recientemente descrito y es altamente homólogo a SUMO1, aunque difieren en varios aspectos. En primer lugar, SUMO5 se encuentra conservado en primates pero parece no expresarse en ratones. Segundo, es transcripcionalmente activo sólo en algunos tejidos, con alta expresión en testículos y células sanguíneas (Liang et al. 2016). A nivel molecular, es importante destacar que la estructura tridimensional de la

Ub y de las Ubls, en particular la de SUMO, está altamente conservada (Vijay-Kumar S et al. 1987; Bayer et al. 1998) (Figura I6).

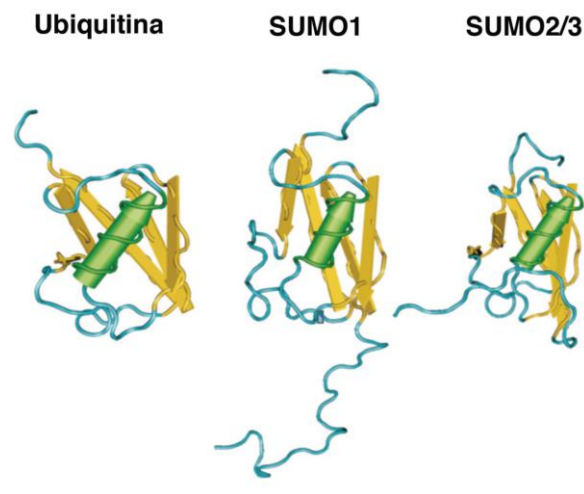


Figura I6. Estructura tridimensional de Ub y SUMO. Esquema comparativo de las estructuras tridimensionales de Ub, SUMO1 y SUMO2/3. Tomado de Martin et al. 2007.

Vía de conjugación de SUMO

Como la ubiquitinación, la SUMOilación resulta en la formación de una unión isopeptídica entre la glicina C-terminal de SUMO y el grupo amino ϵ de una lisina en la proteína blanco (Geissfriedlander & Melchior 2007). El péptido SUMO es sintetizado de manera “inmadura” y para ser conjugado a las proteínas blanco es necesario que atraviese un evento de maduración mediante el cual quedan expuestas dos glicinas en su extremo C-terminal. Para ello se requiere el clivado de los aminoácidos río abajo de este motivo di-glicina, proceso llevado a cabo por proteasas de SUMO a través de su actividad hidrolasa C-terminal. Una vez que SUMO sufrió esta proteólisis, se dice que el péptido se encuentra en estado maduro y es capaz de ser conjugado a los sustratos mediante la vía enzimática correspondiente. El primer paso de conjugación consiste en la activación de SUMO por la enzima activadora E1 específica de SUMO, el heterodímero de las proteínas SAE1-SAE2 en mamíferos (Uba2-Aos1 en levaduras). Esta reacción utiliza ATP para la formación de un intermediario SUMO-adenilato, en el que el extremo C-terminal de SUMO está covalentemente unido a AMP. La ruptura de la unión SUMO-AMP es seguida de la formación de un enlace tioester entre el grupo carboxilo del C- terminal de SUMO y el grupo -SH de la cisteína catalítica en SAE2

(C173). En el segundo paso de la reacción, SUMO es transferido desde SAE2 a la enzima conjugadora E2, Ubc9, formando un nuevo enlace tioéster, esta vez entre la cisteína catalítica de Ubc9 (C93) y el grupo carboxilo de la glicina C-terminal de SUMO. Una característica que diferencia a Ubc9 de las enzimas E2 del resto de las Ubls es su capacidad de reconocer al sustrato proteico directamente. De esta manera, el complejo entre SUMO y Ubc9 unidos mediante un enlace tioéster (Ubc9~S~SUMO) cataliza la formación de la unión isopeptídica entre SUMO y el sustrato. Si bien Ubc9 es capaz de transferir a SUMO directamente al sustrato *in vitro*, este proceso es facilitado *in vivo* por E3 ligasas de SUMO. Las E3 ligasas facilitan la transferencia de SUMO a la proteína blanco, regulan la selección de sustrato y el grado de SUMOilación *in vivo*. La reversibilidad de esta modificación es llevada a cabo por proteasas de SUMO que se encargan de de-conjugar este péptido de las proteínas blanco. Esta serie de pasos enzimáticos, así como también su comparación con la vía de conjugación de Ub, se resumen, en forma de esquema, en la figura 17.

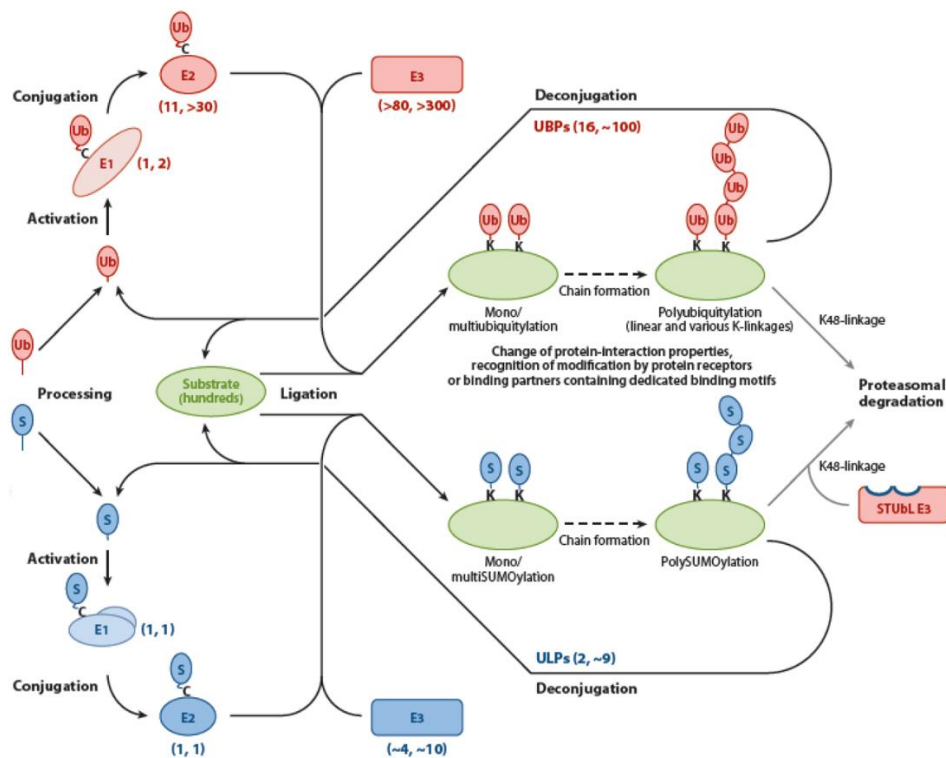


Figura I7. Comparación de la vía de conjugación de Ub y SUMO. Tanto ubiquitina (rojo), como SUMO (azul), pueden conjugarse a cientos de sustratos (verde), a pesar de la gran diferencia en el número de enzimas involucradas (E1, E2 y E3). Estos números de enzimas involucradas

para *Saccharomyces cerevisiae* y humanos se indican entre paréntesis para cada caso. Luego del procesamiento de los respectivos precursores, los péptidos son unidos covalentemente por sus residuos glicinas (G) C-terminal, al grupo ϵ -amino de las lisinasceptoras en los sustratos. Esta modificación es reversible gracias a la presencia de las enzimas de-conjugadoras (isopeptidasas de SUMO). Extraído de Jentsch & Psakhye 2013.

Secuencias blanco de SUMOilacion

Originalmente se consideraba que la proteína blanco de SUMOilación tenía que cumplir dos requisitos para que ocurra esta modificación: contener una secuencia consenso para la conjugación de SUMO definida como Ψ KXD/E donde Ψ es un aminoácido grande e hidrofóbico, K es la lisina blanco de la reacción, X es cualquier aminoácido, D es ácido glutámico y E, ácido aspártico; y contener una señal de localización nuclear (Rodriguez et al. 2001). En cuanto al requerimiento de localización nuclear, se ha descartado como requisito dado que depende de cada caso específico y de las redes de interacción de cada proteína (Geissfriedlander & Melchior 2007). Con el tiempo, se han ido identificando extensiones del consenso: la primera es un sitio de SUMOilación dependiente de fosforilación, PDSM (del inglés Phosphorylation-Dependent SUMOylation Motif) (Hietakangas et al. 2006). Este motivo consiste en el motivo convencional mencionado anteriormente seguido de un número variable de entre 2 y 5 aminoácidos cualquiera, seguidos de una serina fosforilada y una prolina (Ψ KXD/EX2-5pSP) (Hietakangas et al. 2006). Se especula que la carga negativa del grupo fosfato podría estimular la interacción entre Ubc9 y el sustrato (Mohideen et al. 2009). En términos más generales, un rol importante para residuos cargados negativamente en posiciones hacia el extremo carboxilo de la lisina aceptora ha sido revelado por el análisis de secuencias de muchos sustratos de SUMOilación (Blomster et al. 2009). Esto llevó a la identificación de un segundo motivo extendido, NDSM (del inglés Negatively charged amino-acid-Dependent SUMOylation Motif). Ambos motivos tienen en común el hecho de que una carga negativa cerca del sitio consenso mínimo estimula la SUMOilación. Esto se debe a que una región básica de la enzima Ubc9 debe interaccionar con una región ácida de la proteína blanco (Mohideen et al. 2009; Yang et al. 2006). Con el surgimiento de trabajos a gran escala que utilizaron la identificación sitio específica de lisinas SUMOiladas por espectrometría de masa (MS), fueron identificados nuevos sitios consenso, como por ejemplo el ICM (del inglés

“Inverted Consensus Motif”) o el HCSM (del inglés “Hydrophobiy Cluster SUMOylation Motif”) (Da Silva-Ferrada et al. 2012; Matic et al. 2010). Con el tiempo, se obtuvieron mejores rendimientos en purificación de proteínas SUMOiladas a niveles endógenos y analizados por MS, se detectaron nuevas proteínas, en las cuales el número de lisinas más SUMOiladas (más veces detectadas) se encontraban dentro del contexto K-x-E (Figura 18). Además, se comprobó que alrededor del 90% de las proteínas eran SUMOiladas en regiones desordenadas. A pesar de estos estudios, el número de lisinas SUMOiladas reportadas que se desvían de la secuencia consenso o que no poseen tal secuencia está en continuo crecimiento, de modo que actualmente no existen métodos confiables de predicción de secuencias blanco de SUMOilación (Hendriks 2018).

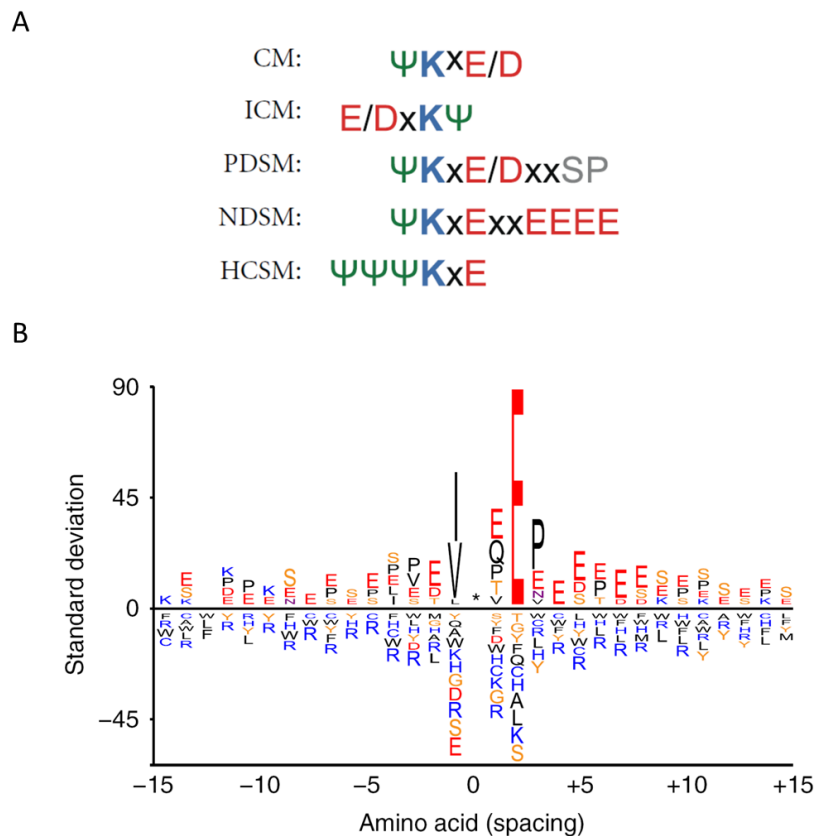


Figura 18. Motivos consenso de SUMOilación. (A) Se muestran las secuencias aminoácidas de los distintos motivos consenso de SUMOilación CM, ICM, PDSM, NDSM y HCSM. En la secuencia se representa Ψ como un aminoácido grande e hidrofóbico, K es la lisina blanco de la reacción, X es cualquier aminoácido, D es ácido glutámico, E es ácido aspártico, S es serina, P es prolina. Figura adaptada de (Da Silva-Ferrada et al. 2012). (B) Representación de los

aminoácidos enriquecidos en el 10% de las lisinas más SUMOiladas en células humanas con niveles endógenos de SUMO, la lisina SUMOilada está marcada con *. Los aminoácidos sobre el eje x corresponden a aquellos enriquecidos mientras que aquellos por debajo del eje a los depletados en torno a los sitios de SUMOilación. La altura de la letra corresponde al grado de enriquecimiento o depleción. Tomado de Hendriks 2018.

Proteasas específicas de SUMO

Las proteasas de SUMO son requeridas para dos tareas: la maduración de SUMO (actividad hidrolasa C-terminal) y la remoción de SUMO de sus blancos (actividad isopeptidasa). La de-SUMOilación o de-conjugación es crucial para el estado estacionario de los niveles de SUMOilación de un determinado blanco. Un número sorprendentemente pequeño de isopeptidasas de SUMO ha sido identificado, de las cuales todas son cisteín-proteasas. La primera identificada en levaduras fue Ulp1 (Takahashi et al. 2000) y le siguieron en mamíferos las llamadas SENPs (del inglés sentrin-specific proteases) en base a la similitud de secuencia (Drag & Salvesen 2008). La familia Ulp/SENp comprende a dos proteínas en levaduras (Ulp1, esencial, y Ulp2) y seis en mamíferos (SENp1, 2, 3, 5, 6 y 7), las cuales difieren entre sí en la actividad catalítica relativa y la preferencia por las distintas isoformas de SUMO (Mukhopadhyay & Dasso 2007). Sin embargo, sus dominios catalíticos no parecen distinguir entre blancos individuales. Un determinante para ello *in vivo* es la región N-terminal, el cual promueve la interacción con distintos blancos y regula la localización sub-celular. La segunda familia de proteasas de SUMO comprende a dos proteínas, DeSI-1 y 2 (del inglés deSUMOylating isopeptidase) (Yeh 2009), las cuales parecen exhibir mayor especificidad de sustrato que las de la familia SENp, de modo consistente con que su depleción no provoca cambios notorios en los niveles de SUMOilación global (Suh et al. 2012). La incorporación más reciente a la lista de isopeptidasas de SUMO es la de USPL1. Esta proteína pertenece a la familia de USPs (Ubiquitin specific proteases) pero cliva SUMO en lugar de ubiquitina. Se encuentra presente en baja abundancia y localizada en los CBs. Su silenciamiento en células de mamíferos no parece afectar los patrones de SUMOilación global y su función para la proliferación celular e integridad de los CBs no depende de su actividad catalítica (Schulz et al. 2012). En la tabla 12 se indica la localización y la preferencia de la SUMO petidasas

Name	Subcellular localization*	Isoform preference	Precursor processing	Deconjugation	Chain editing‡
<i>Saccharomyces cerevisiae</i>					
Ulp1	Nuclear pore	Smt3	Yes	Yes	No
Ulp2	Nucleoplasm	Smt3	No	Yes	Yes
<i>Mammals</i>					
SEN1	Nuclear pore and nuclear foci	SUMO1 and SUMO2/3	Yes	Yes	No
SEN2 [§]	Nuclear pore and nuclear foci; cytoplasm	SUMO2/3 > SUMO1	Yes	Yes	No
SEN3	Nucleolus	SUMO2/3	Unknown	Yes	No
SEN5	Nucleolus and mitochondria	SUMO2/3	Yes	Yes	No
SEN6	Nucleoplasm	SUMO2/3	No	Yes	Yes
SEN7	Nucleoplasm	SUMO2/3	No	Yes	Yes
DESI1	Cytoplasm and nucleus	SUMO1 and SUMO2/3	Weak	Yes	Weak
DESI2	Cytoplasm	Unknown	No	Unknown	Unknown
USPL1	Cajal bodies	SUMO2/3 > SUMO1	Weak	Yes	Yes

Tabla 12. SUMO isopeptidasas. SENP, sentrin-specific protease; Smt3, suppressor of the mitotic fidelity gene 3; SUMO, small ubiquitin-related modifier; Ulp, UBL-specific protease; USPL1, ubiquitin-specific protease-like 1. *Se nombran la localización subcelular principal. ‡Chain editing hace referencia a eliminar el péptido SUMO de una cadena de poliSUMO sin eliminar la conjugación de SUMO de la proteína blanco. §SEN2 tiene tres variantes de splicing que localizan diferencialmente en la célula. Tomado Hickey et al. 2012.

Consecuencias de la SUMOilación

La lista de funciones relacionadas a la conjugación de Ub fue ampliándose a medida que avanzó la investigación de esta modificación post-traducciona. En la actualidad, ya no es relacionada de manera exclusiva con el etiquetado de proteínas para la degradación, debido principalmente a la existencia de distintos motivos o dominios proteicos capaces de reconocer la mono Ub o cadenas de Ub, los cuales favorecen la interacción entre distintas proteínas, y consecuentemente su actividad. De esta manera, que una proteína se encuentre conjugada a Ub no necesariamente está relacionado con su degradación proteasomal (Hurley et al. 2006). De hecho, vale la pena citar un trabajo en el que demuestran que la mono-ubiquitinación de un factor de splicing es indispensable para la interacción de determinadas partículas ribonucleoproteicas en el correcto ensamblado del spliceosoma (Bellare et al. 2008). De manera análoga, la SUMOilación puede alterar la interacción entre macromoléculas, mediante la modificación de superficies proteicas o estructuras tridimensionales de las proteínas blanco de conjugación (Geissfriedlander & Melchior

2007). En este contexto, vale la pena mencionar la identificación de un motivo aminoacídico corto y de naturaleza hidrofóbica en proteínas capaces de interactuar con SUMO en forma no-covalente. Este motivo se denomina SIM (del inglés non-covalent SUMO-Interaction Motif) (Perry et al. 2008; Seu & Chen 2009). Se ha establecido un rol para los motivos SIM en el mecanismo básico de SUMOilación. Ubc9 es sustrato de SUMOilación y se ha demostrado que esto aumenta su afinidad –y por ende, su capacidad de conjugar SUMO- por un conjunto de sustratos, los cuales poseen dominios SIM (Knipscheer et al. 2008). De este modo, dichos sustratos se independizarían de la actividad de E3 ligasas, cuya función sería reemplazada por la interacción SIM-SUMO-Ubc9 (Knipscheer et al. 2008). También se ha demostrado que la actividad E3 ligasa de la proteína Pc2 depende de motivos SIM (Merrill et al. 2010). Más recientemente se ha reportado que tanto la actividad de la E3 ligasa PIASY como su interacción con el péptido SUMO dependen de dos motivos SIM en el extremo C-terminal de esta proteína (Kaur et al. 2017). A nivel funcional, las consecuencias de la conjugación de SUMO pueden ser variadas (Geissfriedlander & Melchior 2007). A diferencia de la conjugación de Ub que, en muchos casos, lleva a degradación por el proteasoma, la SUMOilación no lleva a sus proteínas blanco a un destino específico ni a la degradación directamente. La conjugación de SUMO altera la superficie de una proteína de manera tal que la misma puede perder o ganar interacciones tanto intra como inter-moleculares (Figura 19). A su vez, la conjugación de SUMO puede competir con otras modificaciones post-traduccionales que tengan como blanco al aminoácido lisina, como lo son la ubiquitinación y la acetilación. Si bien queda mucho por aprender acerca del mecanismo mediante el cual la conjugación de SUMO altera la función proteica, está claro que la SUMOilación regula la actividad y/o localización sub-celular de proteínas modificando, principalmente, interacciones proteína-proteína (Gill 2004). El número de procesos celulares reportados que son regulados por SUMOilación está en aumento e incluyen: transcripción (Sapetschnig et al. 2002; Verger et al. 2003; Yang & Sharrocks 2004; Savare et al. 2005; Valin & Gill 2007), splicing (Pelisch et al 2010; Pozzi et al 2017), estructura de la cromatina (Ouyang & Gill 2009; Rouvière et al. 2013), reparación del ADN y estabilidad genómica (Zhao & Blobel 2005; Lee & O'Connell 2006; Heideker et al. 2009; Jentsch & Psakhye 2013), transducción de señales (Yang et al. 2003; Miyazono et al. 2008). Además, un análisis exhaustivo de los

diversos estudios proteómicos realizados hasta el momento reveló que la SUMOilación puede ocurrir en proteínas de una misma vía metabólica, reforzando el concepto de “SUMOilación en grupo”, el cual sugiere que la conjugación a SUMO puede regular todo un proceso celular mediante su conjugación o de-conjugación en muchas proteínas que participan del mismo..

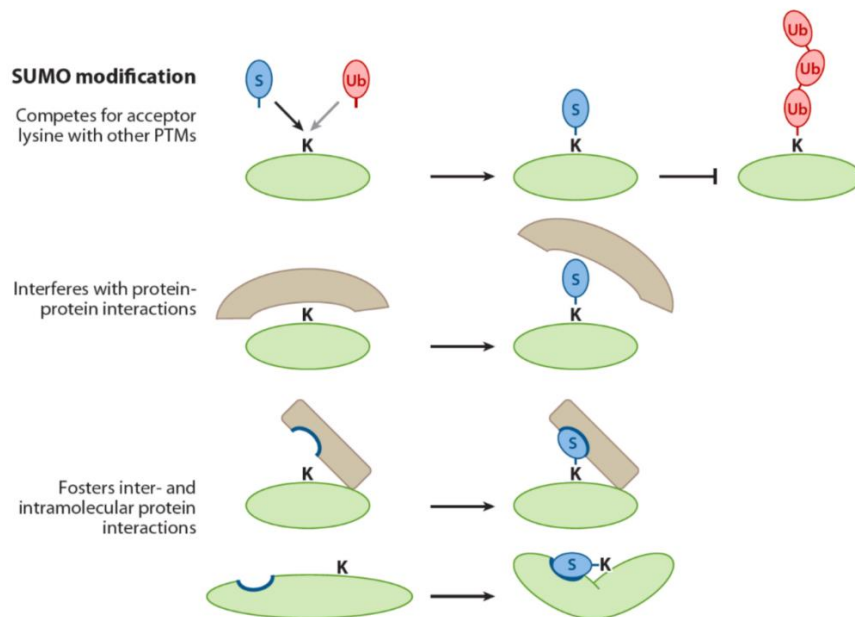


Figura 19. Consecuencias moleculares de la conjugación de SUMO. Al ser SUMOilada, la proteína blanco puede cambiar el repertorio de modificaciones postraduccionales que sufre, ya sea ocurridas o no en el mismo residuo lisina, perder o ganar interacciones y/o cambiar su conformación y consecuentemente, su función. Tomado de (Jentsch & Psakhye 2013).

SUMO y biogénesis de los snRNAs.

Los genes de los snRNAs se encuentran en regiones cercanas a los CBs -donde ocurre la maduración de los snRNPs- e incluso los complejos *mediator*, el LEC y NELF son reclutados a dichos cuerpos nucleares (Takahashi et al. 2020). Diversos trabajos indican la importancia de la proteína coilina para la formación y mantenimiento de los CBs y se la ha observado asociada a los genes de los snRNAs y sus transcritos (Liu 2009; Strzelecka et al. 2010; Machyna et al. 2013). Como se comentó anteriormente la SUMO proteasa USLP1 es necesaria para la integridad de estos cuerpos nucleares, pero independientemente de su actividad catalítica (Schulz et al. 2012). Un trabajo posterior demostró que en ausencia de USLP1 disminuye la abundancia de los snRNAs y en

particular la transcripción de U1 y U2 (Hutten et al. 2014). Sin embargo, se desconoce el mecanismo molecular por el cual USLP1 regula los niveles de snRNAs y si este efecto depende de su actividad catalítica como proteasa de SUMO. Más aún, hasta la fecha se desconoce una regulación directa de SUMO sobre la actividad de las proteínas involucradas en la biogénesis de los snRNAs dependientes de RNAPII.

OBJETIVOS

En esta tesis nos proponemos estudiar la participación de la conjugación de SUMO en la biogénesis de los snRNAs. Nuestra hipótesis de trabajo plantea que dicha modificación post-traducciona l podría influir sobre el ensamblado, localización y/o estabilidad de los complejos multi-proteicos involucrados en la transcripción y el procesamiento 3' de los snRNAs nacientes. Para abordar este objetivo general proponemos los siguientes objetivos específicos:

1) Explorar la participación de diferentes isopeptidasas de-conjugadoras de SUMO en la regulación de la biogénesis de los snRNAs

Evaluar si alteraciones en el balance entre conjugación y de-conjugación de SUMO mediante la sobre-expresión de isopeptidasas de SUMO producen cambios en los niveles de snRNAs nacientes o no procesados, en células en cultivo.

2) Analizar la modificación por conjugación de SUMO de proteínas involucradas en la biogénesis de snRNAs

Evaluar si diferentes componentes de los complejos multiproteicos responsables de regular la transcripción de los snRNAs son modificados por conjugación de SUMO en células en cultivo y en caso de serlo, si dicha modificación es alterada por la sobre-expresión de isopeptidasas de SUMO mencionadas en el objetivo 1.

Evaluar si diferentes subunidades del *Integrator complex*, complejo responsable del procesamiento 3' de los snRNAs nacientes, son modificadas por conjugación de SUMO en células en cultivo y en caso de serlo, si dicha modificación se altera por la sobre-expresión de isopeptidasas de SUMO mencionadas en el objetivo 1.

3) Analizar el posible rol regulatorio de la SUMOilación en la actividad de los complejos involucrados en la transcripción y/o procesamiento 3' de los snRNAs

En base a las proteínas sustrato de SUMOilación identificadas en el objetivo 2, generar versiones mutantes que sean incapaces de conjugarse a SUMO y con ellas evaluar si la pérdida de SUMOilación de dichas proteínas altera la actividad de las mismas en la biogénesis de los snRNAs.

4) Evaluar la influencia de la SUMOilación sobre el ensamblado de complejos regulatorios multiproteicos, reclutamiento a la cromatina y localización subcelular de los factores involucrados en la biogénesis de los snRNAs

Utilizando las mutantes deficientes de SUMOilación generadas en el objetivo 3, analizar si la SUMOilación es necesaria para la interacción entre la proteína en estudio y otras subunidades de su respectivo complejo.

Utilizando las mutantes deficientes de SUMOilación generadas en el objetivo 3, analizar si la SUMOilación está involucrada en el reclutamiento a la cromatina y/o en la localización sub-celular de las proteínas en estudio.

MATERIALES Y METODOS

Cultivo Celular

Mantenimiento de líneas celulares:

Las células HEK 293T (ATCC #CRL-11268) fueron mantenidas en medio DMEM suplementado con 10% suero fetal bovino, 4,5 g/L glucosa, 110 mg/L piruvato de sodio y antibióticos (100 U/ml final de penicilina y 100 µg/ml final de estreptomicina). Antes de que las monocapas alcancen la confluencia total, se aspiró el medio de cultivo y se colocó 1 ml de tripsina 2,5 % p/v. Luego de incubar a 37°C por cinco minutos, se detuvo la reacción con 9 ml de medio completo, se re-plaqueó entre 1/5 y 1/20 del volumen a una nueva placa de cultivo de 10 cm y se llevó a 12 ml de volumen final con medio completo. Las células fueron congeladas en N₂ líquido y descongeladas siguiendo protocolos convencionales. Las células HeLa (ATCC #CCL-2) fueron mantenidas de la misma manera, pero en medio carente de piruvato de sodio.

Transfecciones de ADN

Las células HEK 293T (5×10^5 células) o HeLa ($2,5 \times 10^5$ células) fueron transfectadas en placas de cultivo con Lipofectamina 2000 (Invitrogen) según el protocolo provisto por el fabricante. Cada pocillo fue transfectado en relación 1:3 de ADN total y Lipofectamina. Brevemente, el protocolo de transfección con Lipofectamina que utilizamos en nuestro laboratorio es el siguiente (se ejemplifica utilizando las cantidades de los reactivos adecuadas para pocillos de 35mm):

Se colocan 2 µg totales de ADN plasmídico en un tubo *ependorf* que contiene 250 µl de Opti-MEM (Invitrogen).

Se colocan 6 µl de lipofectamina 2000 en un tubo polipropileno de 5 ml con tapa que contiene 250 µl de Opti-MEM y se incuba por 5 min.

Se transfiere el contenido del tubo *ependorf* al tubo con lipofectamina 2000 y se mezcla vigorosamente.

Se incuba 20 minutos para que se formen los complejos entre el poli-cation y el ADN. Se agregan por goteo 500 µl de los complejos en cada pocillo.

Se dejan las células en el incubador durante 4 a 6 h aproximadamente para que ocurra la transfección.

Luego se aspira el medio para eliminar la Lipofectamina 2000, que de lo contrario tiene efectos tóxicos a largo plazo, y se agrega 2 ml de medio completo por pocillo.

Transfecciones con ARN

Para los experimentos de ARN de interferencia, las células HEK 293T (5×10^5) o HeLa ($2,5 \times 10^5$) fueron transfectadas en placas de cultivo de 35 mm con 3 ml Lipofectamina 2000 (Invitrogen) y 10-25 nM siRNA. El procedimiento es el mismo que el detallado para la transfección con ADN. Los siRNA utilizados en el trabajo fueron siRNA anti 3' UTR de INTS11: UCUUUUGUUCAGCUUUUACU; siRNA anti 3' UTR de INTS9: UCACUUUACCAUAUUUUGGGA; siRNA Accell Human SNAPC1 (6617) pre-diseñado por la empresa Dharmacon; como control utilizamos siRNA anti Luciferasa: CUUACGCUGAGUACUUCGA.

Silenciamiento mediante CRISPRi

Células HEK293T (5×10^4 células) fueron plaqueadas en cada pocillo de una placa *multiwell* de 24 pocillos y 24h más tarde, transfectadas de acuerdo con el protocolo arriba mencionado utilizando 400 ng del plásmido dCas9-KRAB (Addgene #60954), 170 ng del plásmido lentiGuide puro (Addgene #52963), en la cual se clono una guía cuyo blanco es el promotor de INTS11 o una guía que no tiene blanco en el genoma humano (NT) y 30 ng de los vectores codificantes para las distintas variantes de INTS11 o el vector de expresión EGFP-C1 (Addgene #2487) como control. Luego de 72 h, las células fueron cosechadas en TRIzol (Invitrogen) para extraer ARN o Laemmli sample buffer 2X para obtener lisados proteicos totales.

A continuación, se detalla la composición de Laemmli sample buffer 2X :

Laemmli sample buffer 2X: Tris HCl pH 6,8 125 mM, SDS 4% m/v, β -mercaptoetanol 10% v/v, glicerol 20% v/v y azul de bromofenol como colorante

Ensayo de apoptosis

Tinción con yoduro de propidio y anexina

Para determinar la inducción de la apoptosis celular ante distintos tratamientos utilizamos el kit FITC Annexin V Apoptosis Detection Kit I (BD Biosciences). Las células tratadas según lo indicado en cada caso fueron tripsinizadas, la tripsinización fue detenida con medio completo y las células transferidas a un tubo de microcentrifuga para su posterior centrifugación por 5 min a 500 g. Las células fueron lavadas con 1 ml de PBS, y centrifugadas nuevamente. El pellet de células fue resuspendido en 100 μ l de *Binding Buffer* 1X. Posteriormente, la muestra fue incubada 15 minutos con 5 μ l de Anexina V acoplada a FITC y 5 μ l de yoduro de propidio (tal como vienen preparados en el kit). Luego del agregado de 400 μ l de *Binding Buffer* 1X, la fluorescencia correspondiente al FITC y al yoduro de propidio fue evaluada por citometría de flujo. Este ensayo nos permite clasificar a las células en 4 subpoblaciones: las células negativas para anexina V y negativas para yoduro de propidio son células vivas y sanas, las células positivas para anexina V y negativas para yoduro de propidio se encuentran en apoptosis temprana y las células positivas para anexina V y positivas para yoduro de propidio se encuentran en apoptosis tardía. Las células negativas para anexina y positivas para yoduro de propidio son células muertas, pero por necrosis en lugar de apoptosis. Para cada ensayo agregamos controles de citometría de células no teñidas, teñidas solo con yoduro de propidio y solo con Anexina V-FITC para establecer las compensaciones necesarias.

Tinción con cristal violeta

Al momento de realizar el experimento para determinar la viabilidad celular lo primero que se realizó fueron dos lavados con PBS 1X para eliminar de los pocillos de cultivo las células muertas. Luego, se agregó cristal violeta 0,5% p/v diluido en metanol 20% v/v, se incubó por 20 min a temperatura ambiente. En este paso se fijaron y tiñeron únicamente las células vivas. Se realizaron 4 lavados con agua corriente para quitar el exceso de cristal violeta y se dejaron secando los pocillos con células, a temperatura ambiente durante toda la noche. Posteriormente, para desteñir, se agregó 200 μ l de metanol durante 20 min a temperatura ambiente y se procedió a cuantificar la cantidad de células finales midiendo la absorbancia a 595 nm.

Técnicas de interacción entre macromoléculas

Purificación de proteínas conjugadas a 6xHis-SUMO

Células HEK 293T (5×10^5 células) fueron plaquedas en placas de 35mm y 24h más tarde transfectadas de acuerdo con el protocolo antes mencionado. Luego de 48 h, las células fueron cosechadas en PBS con inhibidores de proteasas de SUMO (iodoacetoamida -IAA- 50mM) y se tomó el 5% del volumen total como alícuota *input* al cual se le agregó directamente Laemmli sample buffer 2X. El resto de la muestra se lisó en 1 ml de buffer de extracción. Los lisados fueron sonicados dos veces a una amplitud del 40% por 10 seg para disminuirles la viscosidad y posteriormente se les agregó 35 μ l de la resina Ni^{2+} -NTA (QIAGEN). Luego de incubar a temperatura ambiente con rotación durante 4 h, los tubos fueron centrifugados a 5.000 g por 5 min y se descartó el sobrenadante. La resina fue resuspendida en 0,5 ml de buffer Urea I. Luego de 5 min de incubación con rotación, las muestras fueron centrifugadas por 10 seg a 13.000 g y se descartó el sobrenadante. Luego, se repitió este procedimiento de lavado utilizando secuencialmente los buffers Urea II y Urea III. Luego, las proteínas unidas a la resina fueron eluidas mediante incubación por 20 min con 40 μ l de buffer de elución. En la Figura M1 se muestra la esencia de esta técnica, así como la visualización esperada de los resultados por western blot.

A continuación, se detalla la composición de todos los buffers mencionados:

Buffer de extracción (pH 8): guanidinio-HCl 8 M, Tris-HCl 10 mM, $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ 100 mM, imidazol 5 mM y iodoacetamida 10 mM.

Buffer Urea I (pH 8): Urea 8 M, Tris-HCl 10 mM, $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ 100 mM, imidazol 5 mM y iodoacetamida 10 mM.

Buffer Urea II (pH 6,3): Urea 8 M, Tris-HCl 10 mM, $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ 100 mM, tritón X- 100 0,2%, imidazol 5 mM y iodoacetamida 10 mM.

Buffer Urea III (pH 6,3): Urea 8 M, Tris-HCl 10 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ 100 mM, tritón X- 100 0,1%, imidazol 5 mM y iodoacetamida 10 mM.

Buffer de elución: Tris-HCl pH 6,7 125 mM, SDS 5%, glicerol 30%, β -mercaptoetanol 720 mM, imidazol 200 mM.

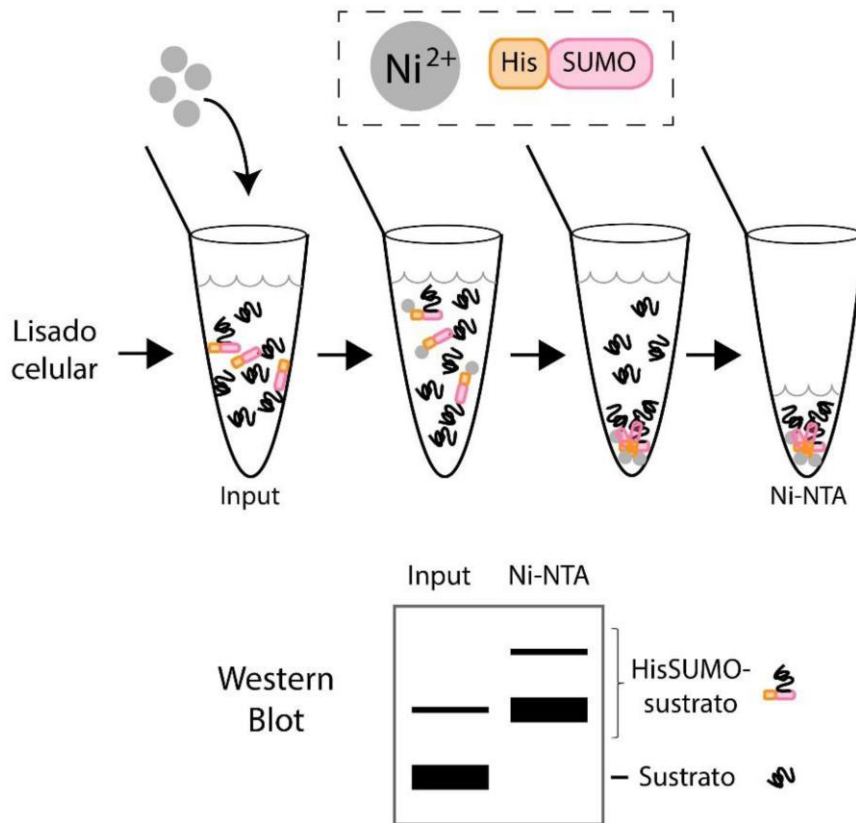


Figura M1. Purificación de proteínas SUMOiladas por resina de Niquel (Ni-NTA). Células en cultivo se transfectan con plásmidos para la expresión de una versión de SUMO etiquetada con 6 histidinas (His-SUMO). Una vez obtenidos los lisados celulares, donde algunas de las proteínas estarán conjugadas a His-SUMO, se toma un 5% como Input y al resto se les agrega microesferas con resina de Niquel (Ni^{2+}). Se incuban por el tiempo suficiente para permitir la interacción entre el Niquel y la etiqueta de histidinas. Finalmente, y luego de los lavados con buffer Urea, se resuspenden las microesferas en buffer de elución (fracción Ni-NTA). Ambas fracciones son analizadas por western blot con anticuerpos contra el sustrato o bien contra la etiqueta fusionada al mismo. Las bandas de menor corrida electroforética en la fracción Ni-NTA, evidencian el sustrato conjugado a His-SUMO.

Co-inmuno precipitación de proteínas

Células HEK 293T (3×10^6) fueron cosechadas en su propio medio de cultivo 24-36 h luego de ser transfectadas. Luego de centrifugarlas a 500g por 5 min a 4°C , fueron lavadas dos veces con PBS frío. El pellet de células fue resuspendido en 500 μl de buffer IP complementado con inhibidores de proteasas (*Complete*, Roche). Posteriormente, las muestras fueron sonicadas en hielo por tres pulsos de 20% de amplitud. Luego de una incubación de 30 min a 4°C , las muestras fueron centrifugadas

a 4000g por 20 min a 4°C. Se reservó el 5% del volumen de sobrenadante como *input*, y con el resto de las muestras (lisados celulares) se procedió con los ensayos de co-inmunoprecipitación según se detalla a continuación para cada caso.

Para inmunoprecipitación de proteínas etiquetadas con la secuencia FLAG, se utilizó 1 µl de anticuerpo anti-FLAG y se incubaron con los lisados a 4°C en rotación constante durante toda la noche. Al día siguiente, se agregó a cada muestra 20 µl de Proteína A/G Plus acopladas a microesferas magnéticas (Invitrogen), previamente equilibradas en buffer IP. Luego de incubar en rotación por 1-2 h a 4°C y centrifugar, el pellet de microesferas fue lavado cuatro veces con buffer Wash complementado con inhibidores de fosfatasas. Los inmunoprecipitados fueron resuspendidos en 30 µl de Laemmli sample buffer 2X.

Para inmunoprecipitación con *GPF-trap* (Chromotek), los lisados celulares fueron incubados con 10 µl de GPF-trap durante 1 h a 4°C. Luego de la incubación, las microesferas fueron lavadas cuatro veces con buffer Wash, centrifugadas y los inmunoprecipitados fueron resuspendidos en 50 µl de Laemmli sample buffer 2X.

A continuación, se detalla la composición de todos los buffers mencionados:

Buffer IP: Tris HCl pH 7,5 50 mM, NaCl 150 mM, Nonidet P-40 1%, Deoxicolato Na 0,5%.

Buffer Wash: Tris HCl 7,5 50 mM, NaCl 250 mM, Deoxicolato Na 0,05%, NP-40 0,1%.

Pull-down nativo con resina de Niquel

Células HEK 293T (3x10⁶) fueron cosechadas en su propio medio de cultivo 24-36 h luego de ser transfectadas. Luego de centrifugarlas a 500g por 5 min a 4°C, fueron lavadas dos veces con PBS frío. El pellet de células fue resuspendido en 500 µl de buffer de lisis complementado con inhibidores de proteasas (*Complete*, Roche). Posteriormente, las muestras fueron sonicadas en hielo por tres pulsos de 20% de amplitud. Luego de una incubación de 30 min a 4°C, las muestras fueron centrifugadas a 4000g por 20 min a 4°C. Se reservó el 5% del volumen de sobrenadante como *input*, y con el resto de las muestras (lisados celulares) se procedió con los ensayos de *pull-down*. Para ello, fueron agregados 50 µl de la resina Ni²⁺-NTA (QIAGEN) y las muestras incubadas por 1 h a temperatura ambiente. Luego de la incubación, la resina

fue lavada tres veces con buffer de lavado y los inmunoprecipitados fueron resuspendidos en 50 μ l de Laemmli sample buffer 2X.

A continuación, se detalla la composición de todos los buffers mencionados:

Buffer de lisis: 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM Imidazol, 0,5 % (v/v) Triton.

Buffer de lavado: 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 30 mM Imidazol.

Inmunoprecipitación de la cromatina (ChIP)

Las células fueron *crosslinkeadas* con 1% de formaldehído (v/v), lavadas dos veces con PBS frío y recolectadas. Los precipitados fueron resuspendidos en 2 ml de buffer de lisis conteniendo inhibidor de proteasas (*Complete*, Roche) e incubados durante 10 min en hielo. Los extractos fueron sonicados con un sonicador *Branson W-450* a una amplitud del 30% con 5 pulsos de 30 seg, resultando en fragmentos de cromatina de aproximadamente 300 pb, y posteriormente centrifugados a 12.000 g por 10 min. Una alícuota de 50 μ l fue reservada para *input* y el resto fue diluído 1/10 en buffer de ChIP. La solución de cromatina fue *pre-clareada* a 4°C con Protein G Dynabeads® (Invitrogen) durante 1 hora antes de incubarla durante toda la noche a 4°C en presencia de anticuerpo anti FLAG. Para inmunoprecipitar INTS11 e INTS9 endógenas fueron agregados 2 μ l del anticuerpo correspondiente y en paralelo, fue realizada una inmunoprecipitación control con IgG de conejo. Al día siguiente, los complejos fueron incubados con Protein G *Dynabeads*® (Invitrogen) durante 1 h a 4°C y luego las *beads* fueron lavadas en rotador a 4°C, una vez con cada uno de los siguientes buffers: buffer de lavado de baja sal, buffer de lavado de alta sal y buffer de lavado con LiCl, y luego dos veces con buffer TE. Los complejos fueron eluídos en buffer de elución y el *crosslinking* fue revertido llevando la solución a una concentración final de 200mM de NaCl y realizando una incubación a 65°C durante 6 h o toda la noche. Las muestras fueron digeridas con proteinasa K durante 1 h a 45°C y el ADN fue extraído con el kit de purificación de PCR de Qiagen. El ADN obtenido fue cuantificado por qPCR con los *primers* indicados en la lista y normalizado a los valores de *input* para tres experimentos independientes.

A continuación, se detalla la composición de todos los buffers mencionados:

Buffer de lisis: 1% p/v SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8,1.

Buffer de ChIP: 0,01% w/v SDS, 1,1% v/v Triton X-100, 1,2 mM EDTA, 16,7 mM Tris- HC pH 8,1, 167 mM NaCl e inhibidor de proteasas-Complete, Roche-.

Buffer de lavado de baja sal: 0,1% p/v SDS, 1% v/v Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8,1, 150 mM NaCl.

Buffer de lavado de alta sal: 0,1% p/v SDS, 1% v/v Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8,1, 500 mM NaCl.

Buffer de lavado con LiCl: 0,25 M LiCl, 1% v/v NP-40, 1% w/v deoxicolato, 1 mM EDTA, 10 mM Tris-HCl, pH 8,1.

Buffer TE: 10 mM Tris-HCl, 1 mM EDTA

Buffer de elución: 1% p/v de SDS y 50 mM NaHCO₃

Western Blots

Las muestras proteicas fueron corridas en geles de poliacrilamida con SDS, en buffer de corrida durante 50 min a 200V, y posteriormente transferidas por 1 h a 100V en buffer de transferencia, a membranas de Nitrocelulosa (GE Healthcare). Las membranas fueron bloqueadas por una hora en una solución conteniendo 5% leche en TBS. Posteriormente, se agregó la dilución apropiada del anticuerpo primario, en la misma solución de bloqueo y se incubó toda la noche a 4°C. Luego de 3 lavados de 10 min cada uno con TBS más 0,1% tween-20 (TTBS), las membranas fueron incubadas con los anticuerpos secundarios correspondientes, acoplados a fluoróforos (Licor). Luego de 4 lavados de 10 min cada uno con TTBS, las membranas fueron escaneadas en el equipo Odyssey para la detección de los anticuerpos Licor.

Los anticuerpos utilizados en este trabajo fueron: ratón monoclonal anti- β -actin C4 (Santa Cruz Biotechnology), conejo policlonal anti β -tubulin H-235 (Santa Cruz Biotechnology), raton monoclonal anti GFP B2 (Santa Cruz Biotechnology), ratón monoclonal anti c-Myc 9E10 (Santa Cruz Biotechnology), ratón monoclonal anti FLAG-M2 (Sigma-Aldrich), ratón monoclonal anti HA MMS-101R (Covance), conejo policlonal

anti INTS11 ab75276 (Abcam), conejo policlonal anti INTS11 A301-274A (Bethyl), conejo polyclonal anti INTS9 13945 (Cell Signaling Technology) y conejo policlonal anti SNAPC1 (generosamente cedido por el laboratorio de Shona Murphy, Oxford, UK).

A continuación, se detalla la composición de todos los buffers mencionados:

Buffer de corrida: 25 mM Tris, 192 mM glycine, 0.1% SDS

Buffer transferencia: 25 mM Tris, 192 mM glycine, Metanol 20%

Microscopía de Fluorescencia

Células HeLa fueron plaqueadas en placas de cultivo *multiwell* sobre vidrios cubre-objetos para microscopía. Al día siguiente, las células fueron transfectadas con plásmidos de expresión para las variantes de INTS11 fusionadas a GFP. A las 24 h, las células fueron fijadas con paraformaldehído 4% por 10 min. Posteriormente, fueron permeabilizadas con Tritón X-100 1% en PBS por 10 min. Para la tinción de núcleos se utilizó ioduro de propidio por 5 min a temperatura ambiente, tratando previamente a las células con RNAsa 0,1 mg/μl a 37°C por 30 min. Finalmente, las células fueron lavadas con PBS y montadas en medio VECTASHIELD (Vector Labs). Las imágenes fueron adquiridas con un microscopio confocal Olympus FV1000, con un objetivo 60X 1,4 NA. Para cuantificar la señal de fluorescencia a partir de un número elevado de imágenes de microscopía confocal, se utilizó el *software* Cell Profiler (v3.1.5). Se elaboró una línea de comando para reconocer de manera automatizada los núcleos y citoplasmas. Para lograrlo, se aprovechó la tinción con ioduro de propidio y la fluorescencia de GFP. Estableciendo un diámetro mínimo y máximo para cada núcleo y citoplasma. La cuantificación se realizó analizando el cociente entre las intensidades media de los núcleos sobre las intensidades medias de los citoplasmas.

Preparación de plásmidos

Transformación de bacterias

Para generar bacterias competentes se utilizó la cepa DH5α de *Escherichia coli*, y un protocolo tradicional basado en CaCl₂ y otras sales. Con este protocolo, se generan bacterias con una eficiencia de aproximadamente 10⁶ UFC/μg. Estas bacterias

fueron transformadas según métodos convencionales basados en el protocolo de Hanahan (Hanahan D 1983)

Alternativamente, se prepararon bacterias electrocompetentes de la cepa STBL3 de *Escherichia coli* en los casos en que se necesitó una mayor eficiencia de transformación (10^7 - 10^9 UFC/ μ g ADN). El protocolo utilizado para desarrollar estas bacterias electrocompetentes fue extraído del trabajo Warren (Warren 2011)

Preparación de plásmidos a pequeña escala o minipreps

Se utiliza esta técnica para producir una pequeña cantidad de plásmido a partir de un cultivo bacteriano de unos pocos mililitros, con el objetivo de, por ejemplo, chequear mediante mapeo de restricción o secuenciación si los clonados realizados habían generado el plásmido de interés. Se utilizó un protocolo estándar adaptado y cuando la aplicación subsiguiente demandaba una más alta calidad de ADN plasmídico, se utilizó un kit de Invitrogen (PureLink HiPure Plasmid Miniprep).

Preparación de plásmidos a gran escala o maxipreps

Las columnas de intercambio aniónico de Qiagen permiten obtener unos 100 μ g de ADN plasmídico (columnas tip 100) o 500 μ g (columnas tip 500). También se utilizaron kits de Invitrogen (PureLink HiPure Plasmid Maxiprep) para la obtención de grandes cantidades de plásmido libres de endotoxinas. En todos los casos, se realizó el protocolo siguiendo las recomendaciones de los fabricantes.

Cuantificación de plásmidos

Para conocer la concentración de ADN o ARN en solución se midió absorbancia a 260 nm con un espectrofotómetro (NanoDrop). Como alternativa se midieron concentraciones por fluorimetría, utilizando el lector Qubit y los reactivos Quant-iT RNA y dsDNA BR (Invitrogen). Paralelamente, para cuantificar fragmentos de ADN se realizó una cuantificación por electroforesis, corriendo alícuotas en un gel de agarosa con bromuro de etidio y comparando con estándares de masa (High Mass y Low Mass Ladders, Invitrogen).

Mutagénesis dirigida

Fueron diseñados *primers forward* y *reverse* conteniendo la mutación deseada y fueron realizados 20 ciclos de amplificación por PCR a partir del plásmido molde. La reacción de amplificación contiene:

Buffer PCR para Pfu (con Mg): 5 μ l

Plásmido templado: 50 ng

Primer Fw: 125 ng

Primer Rev: 125 ng

dNTPs 10 mM: 1 μ l

PfuTurbo ADN polimerasa (2.5 U/ μ l): 1 μ l

H₂O c.s.p.: 50 μ l

La reacción de PCR consistió en:

Nº de ciclos	Temperatura (°C)	Tiempo
1	94	5'
	94	30"
	55	30"
20	72	1'/Kb
1	72	2'/Kb

Para la digestión, 0,5 μ l de la enzima de restricción Dpn I (NEB) fue agregada a cada reacción de amplificación e incubada durante 4 h a 37°C. Bacterias competentes químicas o electro de la cepa DH5 α /Stbl3 de *Escherichia coli* fueron transformadas con 5 μ l de la reacción de digestión. Finalmente, la mutagénesis fue corroborada por secuenciación de algunos de los clones obtenidos (Macrogen Inc.).

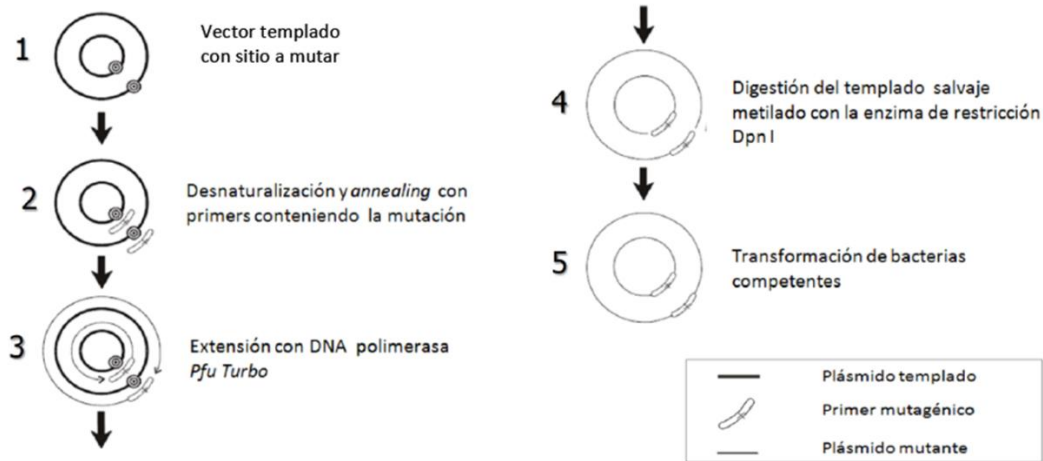


Figura M2: Esquema experimental del método de mutagénesis mediada por *Dpn I*. Adaptado de *Pfu ADN polymerase* Agilent Technologies.

ARN, RT & PCR

Purificación de ARN

Para obtener ARN de las células transfectadas, fue utilizado el protocolo de los reactivos TRIzol (Invitrogen) o TRI Reagent (MRC). Cabe aclarar que todos los reactivos utilizados deben estar libres de RNasas, lo que se logra autoclavando las soluciones o utilizando stocks nuevos, manipulados con todas las precauciones necesarias para no introducir RNasas presentes en el ambiente.

Se cosecha cada pocillo de células directamente con el agregado de 250 μl de reactivo y se las transfiere a un tubo eppendorf. Se agrega 1/5 de volumen (50 μl) de cloroformo y se mezcla vigorosamente. Se deja reposar a temperatura ambiente por 15 min. Posteriormente se centrifuga a 12000g por 15 min a 4°C. Se toma la fase acuosa (superior) y se la pasa a un nuevo tubo eppendorf. Se precipita con un volumen (100 μl) de isopropanol y se agita por inversión. Luego de incubar 20 min a -20°C, se centrifuga 15 min a 12000g a 4°C. Se descarta el sobrenadante por volcado y se agregan 500 μl de etanol 75% v/v y se agita por inversión para lavar el precipitado. Luego se centrifuga 5 min a 7600g a temperatura ambiente para disminuir la cantidad de sales precipitadas. Se vuelca el sobrenadante y se aspira bien para que no queden restos de etanol. Se seca el pellet al aire, se agrega agua (20 μl generalmente) se congela a -20°C. Para resuspender el ARN, se descongela en un bloque térmico a 55°C

durante 5 min y se agita vigorosamente con un vortex. Se puede controlar la integridad del ARN extraído corriéndolo en un gel de agarosa 1% p/v, a un voltaje relativamente alto y por poco tiempo, para disminuir la degradación durante la corrida. En una preparación buena de ARN, deben observarse 3 bandas correspondientes, de mayor a menor tamaño, al ARN ribosomal 28S, al ARN ribosomal 18S y a los ARN de transferencia.

Transcripción reversa (RT)

El protocolo para la obtención de ADN copia a partir de ARN se detalla a continuación. El mismo consiste en el protocolo provisto por el fabricante de la enzima Transcriptasa Reversa (MMLV-RT, Invitrogen) con leves modificaciones. Se desnaturalizan 5 µl de ARN colocándolo 5 min a 65°C y posteriormente se pasa el tubo inmediatamente a hielo para impedir la re-naturalización. Luego se agregan 15 µl de mezcla de reacción a cada tubo y se deja 10 min a temperatura ambiente para que el oligo(dT) pueda aparearse a las colas de poli(A) de ARN mensajeros poli-adenilados. Por último, se incuba 1 h a 37°C para que ocurra la reacción y se detiene la misma incubando 10 min a 70°C.

La mezcla de reacción está compuesta de la siguiente manera:

Buffer de RT 5X: 4µl

DTT 100 mM: 2µl

dNTPs (dATP, dCTP, dGTP y dTTP) 10 mM: 1 µl

Oligo Random (oligo de 10 nucleótidos) 100 µM: 0,5 µl

Enzima M-MLV RT 200 U/µl: 1,5 µl

H₂O c.s.p: 15 µl.

PCR en tiempo real (qPCR)

Para esta reacción de PCR se utilizan como molde, 2,5 µl de una dilución 1:20 de la reacción de RT que contiene el ADN copia. La curva estándar es realizada mediante 7 diluciones seriadas al medio partiendo de una dilución 1/2 de un pool de

las muestras. A cada pocillo con la dilución de ADN copia se le agrega 22,5 μ l de mix, compuesta de la siguiente manera:

Buffer PCR: 2,5 μ l

MgCl₂ 25 mM: 1,5 μ l

Primer Fw 20 μ M: 0,5 μ l

Primer Rev 20 μ M: 0,5 μ l

dNTPs 10 mM: 0,5 μ l

Taq ADN polimerasa (5 U/ μ l): 0,15 μ l

Sybr Green: 0,025 μ l

H₂O c.s.p: 20 μ l

La reacción de PCR consistió en:

Nº de ciclos	Temperatura (°C)	Tiempo
1	94	2'
20	94	15"
	60	15"
	72	20"
1	72	5'

Lista de primers utilizados:

	Nombre	Secuencia FW	Secuencia REV
Primers qPCR	U1 snRNA Total	CCATGATCACGAAGGTGGTTT	ATGCAGTCGAGTTCCACAT
	U2 snRNA Total	TTCTCGGCCTTTTGCTAAG	CTCCCTGCTCCAAAATCCA
	U4 snRNA Total	GCCAATGAGGTTTATCCGAGG	TCAAAAATTGCCAATGCCG
	U5 snRNA Total	GGTTTCTCTTCAGATCGTATAAATC	CTCAAAAATTGGTTTAAGACTCAGA
	GAPDH mRNA	ATCGTGGAAGGACTCATGAC	GCAGGGATGATGTTCTGGAG
	U1 snRNAs sin clivar	TACCTGGCAGGGGAGATAACC	GCGTACAGTCTACTTTTAAAACCTC
	U2 snRNAs sin clivar	CTCCACGCATCGACCTGG	TGCTACCGTCTCTCACCTC
	U4 snRNAs sin clivar	CGTAGCCAATGAGGTCTATCCG	CTCTTCAACCTCCAAAACC
	U5 snRNAs sin clivar	ATTTCCGTGGAGAGGAACAACCTC	GCACCATTGAACAGAAAAGGAA
Mutagénesis	INTS11 K115R	CGCCGTAGACAAGAGGGGCGAGGCCAACTTC	GAAGTTGGCCTCGCCCTCTGTCTACGGCG
	INTS11 K289R	CTGGACCAACCAGAGGATCCGCAAGACTTTC	GAAAGTCTTGCGGATCCTCTGTTGGTCCAG
	INTS11 K369R	CAGCGGGCAGCGGAGGCTCGAGATGGAGGG	CCCTCATCTCGAGCTCCGCTGCCGCTG
	INTS11 K381R	GGTGCTGGAGGTCAGGATGCAGGTGGAG	CTCCACCTGCATCCTGACCTCCAGCACC
	INTS11 K462R	CGCTGGGGCTGCTGAGGCGGGAGATGGCGC	GCGCCATCTCCCGCTCAGCAGCCCCAGCG
	INTS11 K475R	CCCTGAGGCCAAGAGGCCTCGGCTCCTGC	GCAGGAGCCGAGGCCTTGGCCTCAGGG
	INTS11 E464A	CTGCTGAAGCGGGCGATGGCGCAGGGGCT	AGCCCCCTGCGCCATCGCCCGCTTACAGCAG
	INTS11 E203Q	AACCTGCTCATCACACAGTCCACGTACGCCA	TGGCGTACGTGGACTGTGTGATGAGCAGGTT
	INTS9 K607R	GCTTCAGTGATATTAGGGTGGAGGACACAG	CTGTGTCTCCACCCTAATATCACTGAAGC
	SNAPC1 K245R	ATGATGGAGAAGAAAAGAAATGGAAGGAAATTC	ATGATGGAGAAGAAAAGAAATGGAAGGAAATTC
	SNAPC1 K286R	GCATCGTCAAGTCAGACTCGACTCTTCTGAC	GTCAGAAGAGTCGAGTCTGACTTGACGATGC
	SNAPC1 K302	GTC AAGGGCAAGTCAGAGCAACTAGGAAAAAAG	CTTTTTCTAGTTGCTGCTGACTTGCCCTTGAC
	SNAPC1 K333R	GTGCAGAATATACACAGGGAAGATAAACCTTTAAG	CTTAAAGGTTTATCTTCCCTGTGTATATTCTGCAC
	ChIP	PSE U1	GGAAAGGGCTCGGGAGTGCGCG
PSE U2		GATGAGAGTGGGACGGTGAC	CACTTGATCTTAGCCAAAAGG
Región Intergenica		GGAGCGGAGCGTTCTGTCTCCCC	AGAGTGTGAGCCCTCATTACGCCC
3' box U2		CTCCACGCATCGACCTGG	TGCTACCGTCTCTCACCTC
dCas9	Guía dCas9 NT	GCAGGGTTTTCCAGTCACGACGTTGTAAA	CTTTACAACGTCGTGACTGGGAAAACCTG
	Guía INTS11 promotor	CCACCCACAGCCCTCCCGG	CCGGGAGGGCTGTGGGGTGG

RESULTADOS

Resultados previos de nuestro laboratorio mostraron que en una reacción de *splicing, in vitro*, ciertas proteínas unidas al pre-ARN mensajero sustrato de la reacción se encuentran conjugadas a SUMO y que al disminuir los niveles de SUMOilación del extracto proteico utilizado para las reacciones mediante el agregado de una enzima de-conjugadora de SUMO se reduce la eficiencia de *splicing*. Por otro lado, comprobamos que la SUMOilación de la proteína spliceosomal Prp3 es requerida para el proceso de *splicing* en células de mamífero en cultivo. Habiendo demostrado la importancia de la conjugación SUMO en la regulación del proceso *splicing*, nos preguntamos entonces si la SUMOilación podría estar también regulando la biogénesis de los snRNAs, en particular de aquellos de clase Sm o dependientes de RNAP II.

La disminución global de SUMOilación aumenta los niveles de snRNAs nacientes

Nos propusimos estudiar qué ocurre con los snRNA nacientes o pre-snRNAs cuando se alteran los niveles globales de SUMOilación. Para ello, sobre-expresamos la proteasa de SUMO SENP1 en distintas líneas celulares mediante la transfección del plásmido de expresión correspondiente. Los snRNAs no poseen intrones, pero su procesamiento 3' mediado por el complejo *Integrator* es co-transcripcional lo cual nos permitió diseñar primers específicos para medir los niveles de snRNA inmaduros, no procesados o nacientes. En la figura 1 se muestra como al disminuir la SUMOilación global de la célula aumentan los niveles no procesados de un snRNA transcrito a partir de un plásmido reportero co-transfectado. Este plásmido contiene las regiones regulatorias del gen de U2, pero la unidad transcripcional de dicho gen se encuentra interrumpida por el exón 2 de la b-globina (Figura R1 A). Ante la sobre-expresión de SENP1, observamos un aumento en los niveles no procesados del transcrito quimérico U2-b-globina proveniente del reportero (Figura R1 B).

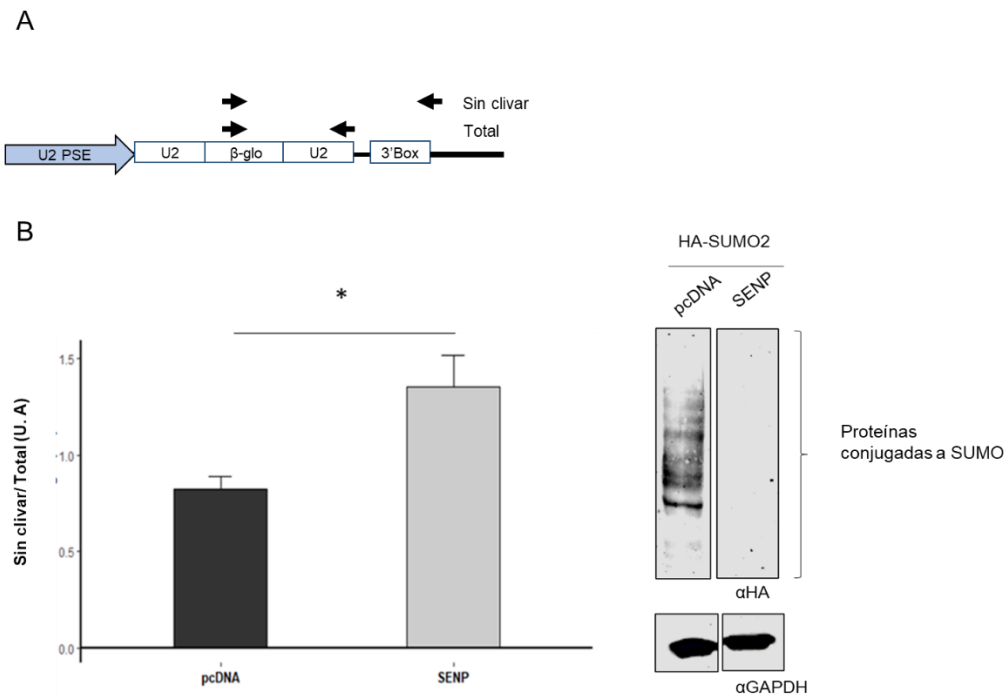


Figura R1: Niveles de snRNA no procesados proveniente de un reportero transcripcional aumentan cuando disminuye la SUMOilación global. A) Esquema del reportero transcripcional y los *primers* utilizados para distinguir los snRNAs (transcriptos quiméricos U2-b-globina) no procesados (sin clivar) y los totales (procesados + no procesados) derivados del mismo. B) Células HeLa fueron co-transfectadas con el plásmido reportero indicado en A, un plásmido que codifica para SUMO2 fusionada al epítipo HA y un plásmido que codifica para la SUMO proteasa SENP o, como control, un plásmido “vacío”. Luego de 48 h, se extrajo ARN total y se lo sometió a RT-qPCR con los *primers* indicados en A (panel izquierdo, n=3, * P<0,05, Student’s T). Los niveles de SUMOilación global fueron analizados mediante WB utilizando extractos celulares totales y un anticuerpo anti-HA (panel derecho).

Quisimos corroborar si el aumento en los niveles de transcriptos quiméricos de U2-b-globina no procesados como consecuencia de la disminución de la SUMOilación global detectado con el gen reportero ocurre también con los transcriptos de snRNA endógenos. Para ello, usamos una estrategia similar: sobre expresamos la proteasa de SUMO SENP1 en la línea celular HEK 293T, junto con un plásmido que codifica para SUMO2 fusionada a HA y procedimos a medir los niveles nacientes de los snRNA endógenos por RT-qPCR con los *primers* indicados en la figura R2 A. Al igual que para los transcriptos provenientes del gen reportero, observamos un aumento en los

niveles no procesados de los snRNA provenientes de los genes endógenos U1, U2, U4 y U5 al disminuir la SUMOilación global (Figura R2 B).

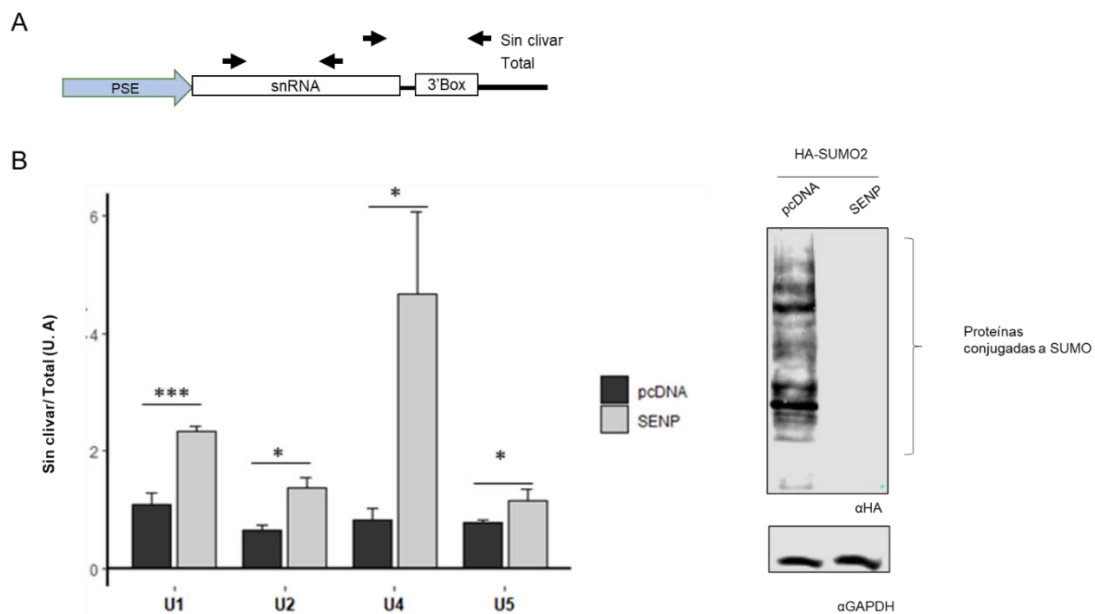


Figura R2: Niveles de snRNA no procesados endógenos aumentan ante la disminución de la SUMOilación global. A) Esquema de los *primers* utilizados para distinguir los snRNAs no procesados y los totales (procesados + no procesados), derivados de genes endógenos. B) Células HEK 293T fueron co-transfectadas con un plásmido que codifica para SUMO2 fusionada a HA y un plásmido que codifica para la SUMO proteasa SENP1 o, como control, un plásmido “vacío”. Luego de 48 h, se extrajo ARN total y se lo sometió a RT-qPCR con los *primers* indicados en A (panel izquierdo, n=3, *** P<0,001, * P<0,05, Student’s T). Los niveles de SUMOilación global fueron analizados mediante WB utilizando un anticuerpo anti-HA (panel derecho).

En los experimentos que dieron lugar a las figuras R1 y R2 es claro que la sobre-expresión de la proteasa SENP1 produjo una marcada alteración en los niveles globales de SUMOilación. Con el fin de alterar los niveles de conjugación de SUMO de manera más localizada usamos una estrategia similar pero esta vez sobre-expresando la SUMO proteasa USPL1, la cual se encuentra en compartimentos sub-nucleares denominados CBs, junto con un plásmido que codifica para SUMO2 fusionada a HA y procedimos a medir los niveles de snRNA nacientes (no procesados) provenientes de genes endógenos (Figura R3).

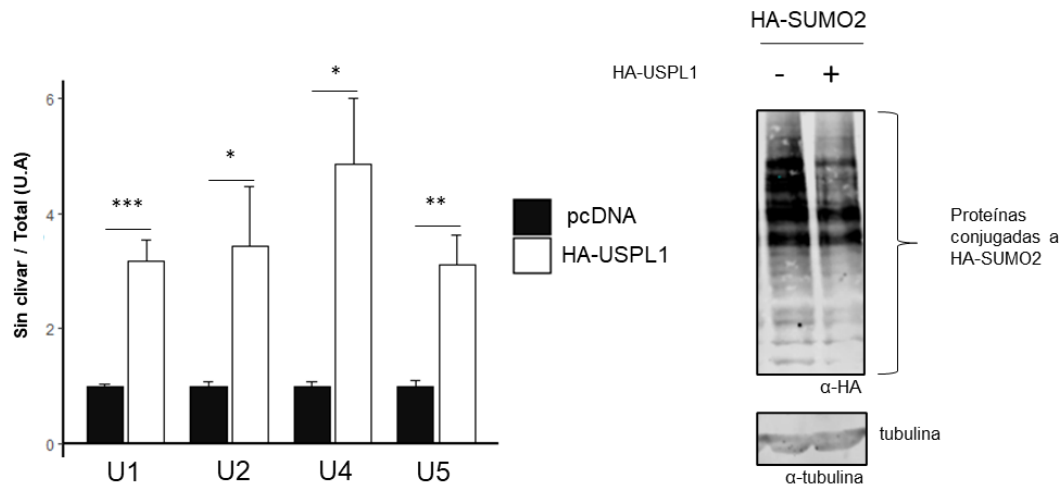


Figura R3: Niveles de snRNA no procesados endógenos aumentan ante la sobre-expresión de USPL1. Células HEK 293T fueron co-transfectadas con un plásmido que codifica para SUMO2 fusionada a HA y un plásmido que codifica para la SUMO proteasa USPL1 fusionada a HA o, como control, un plásmido “vacío”. Luego de 48 h, se extrajo ARN total y se lo sometió a RT-qPCR con los *primers* indicados en 2A (panel izquierdo, n=4, *** P<0,001, * P<0,05, Student’s T). Los niveles de SUMOilación global fueron analizados mediante WB utilizando un anticuerpo anti-HA (panel derecho).

Al igual que lo observado al sobre-expresar SENP1, la sobre-expresión de USPL1 produjo un aumento de los niveles sin procesar de los snRNAs, aún cuando los niveles globales de conjugación de SUMO no se vieron severamente disminuídos. Evaluamos qué ocurre cuando las células sobre-expresan USPL1 dado que la misma es una proteasa de SUMO que se localiza en los CBs, los cuales se ubican en las cercanías de donde ocurre la transcripción de los snRNAs (Schulz et al. 2012; Machyna et al. 2013). En conclusión, los resultados indican que alterar los niveles de SUMOilación global o bien en regiones cercanas a los genes de snRNAs altera el procesamiento 3’ de estos ARNs no codificantes.

Proteínas involucradas en la biogénesis de los snRNA son modificadas por conjugación de SUMO2

A partir de los resultados mostrados anteriormente decidimos analizar la SUMOilación de proteínas involucradas en la transcripción, elongación y el procesamiento 3’ de los snRNAs. Para ello, células HEK 293T fueron co-transfectadas con plásmidos de expresión para las distintas proteínas involucradas en la biogénesis

de snRNAs y un plásmido que codifica para 6xHis-SUMO2 (versión de SUMO-2 etiquetada con un epítipo de seis histidinas consecutivas) o, como control negativo, un plásmido vacío. Los lisados celulares obtenidos post-transfección fueron sometidos a cromatografía de afinidad en resina de níquel con el fin de enriquecer la muestra en proteínas que contengan el tracto de 6xHis, es decir, proteínas conjugadas a 6xHis-SUMO2. Ensayos de WB con anticuerpos específicos contra las etiquetas de las proteínas sobre-expresadas, revelaron que las mismas son susceptibles de conjugarse a SUMO2, evidenciado por la aparición de bandas de menor movilidad electroforética en la fracción “Ni-NTA” (material retenido por la columna) que la correspondiente al peso molecular de la proteína observada en los lisados celulares (Figura R4).

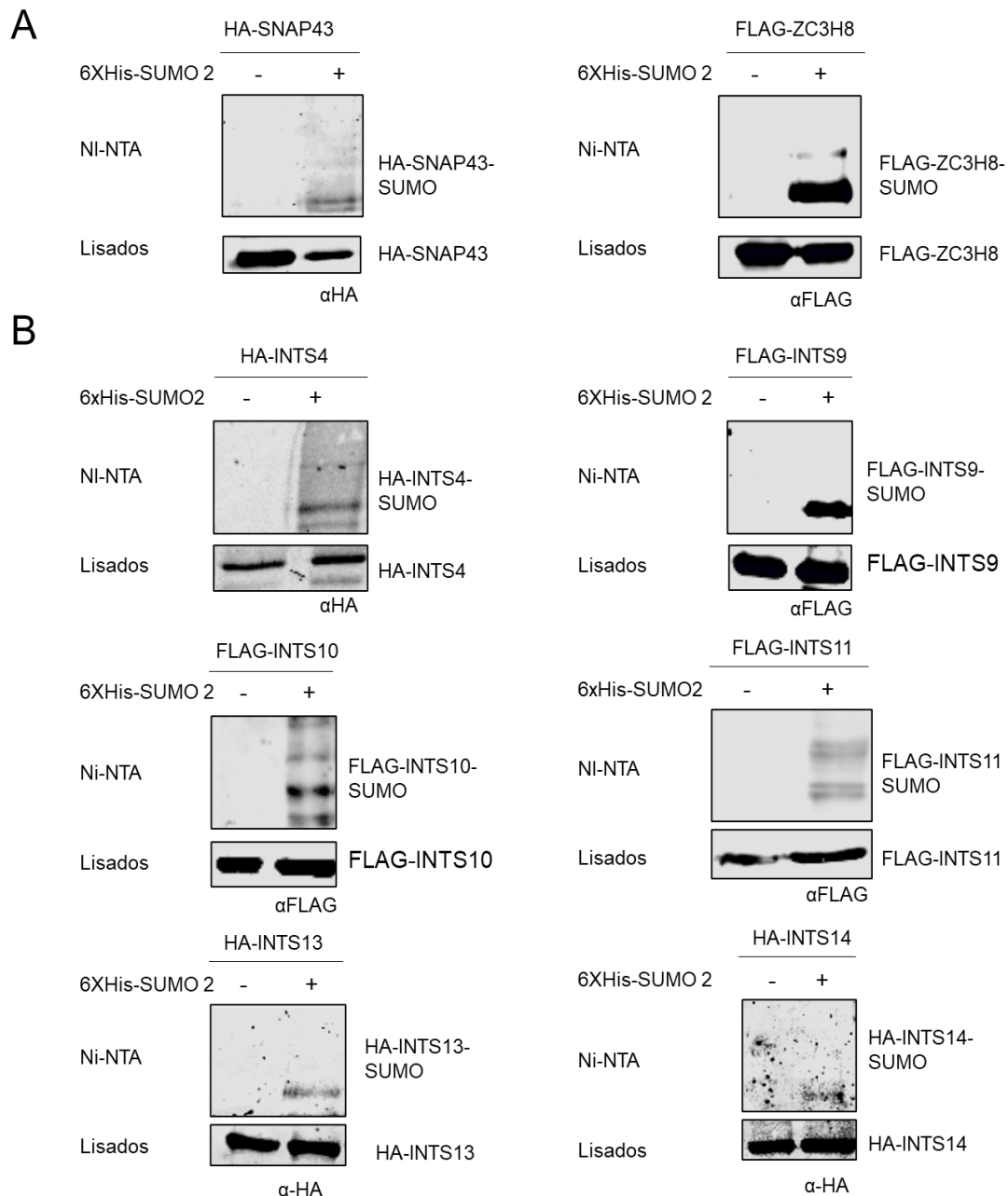


Figura R4: Diversas proteínas involucradas en la biogénesis de snRNAs son blanco de conjugación de SUMO. Células HEK 293T fueron co-transfectadas con plásmidos de expresión para Ubc9, 6XHis-SUMO2, o vector vacío, y un putativo sustrato de SUMOilación de interés etiquetado con HA o FLAG. Luego de 48 h, los lisados celulares fueron sometidos a purificación por resina de níquel para enriquecer en proteínas SUMOiladas. Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con los anticuerpos indicados. A) Proteínas involucradas en la transcripción (SNAP43) y elongación (ZC3H8) de los snRNAs. B) Proteínas del complejo *Integrator*, encargado del procesamiento 3' de lo snRNAs.

Como ya ha sido indicado en la sección Introducción, la proteína SNAP43 es parte de un complejo llamado SNAPc (*small nuclear activator complex*) que se une a la región promotora PSE (*proximal sequence element*). Esta proteína funciona como andamiaje para la interacción entre SNAP50 y SNAP190 y también es capaz de interactuar con Rb y p53 (Mittal et al. 1999; Hirsch et al. 2000; Gridasova et al. 2005; James Faresse et al. 2012). La proteína ZC3H8 es parte del complejo denominado LEC (*Little Elongation Complex*), su estructura muestra que ZC3H8 posee tres dominios *C3H1-zinc fingers*, que suelen estar presentes en proteínas de unión al RNA, pero se desconoce su función dentro del LEC. Observamos que efectivamente, SNAP43 y ZC3H8 son modificadas por conjugación de SUMO (Figura R4 A).

Además, detectamos la SUMOilación de seis proteínas del *Integrator complex*, el responsable de clivar el snRNA naciente en su región 3'. Este complejo está formado por 14 proteínas y se desconoce la función de la mayoría de ellas. Sin embargo, se ha descrito la formación de un heterodímero entre INTS9 e INTS11 responsable del procesamiento endonucleolítico del snRNA naciente. Como se observa en la figura R4 B, ambas proteínas son modificadas por conjugación de SUMO.

Vale la pena mencionar que hasta la fecha, ningún trabajo ha descrito la regulación de estas proteínas por modificaciones post traduccionales.

En las siguientes secciones, evaluamos la regulación de la conjugación SUMO sobre el factor de transcripción SNAP43, y los factores de procesamiento 3' INTS9 e INTS11. Para facilitar la lectura, hemos agrupado los resultados obtenidos según la proteína y el rol que cumple en la biogénesis de los snRNAs.

SNAP43

Las lisinas 245 y 333 son sitios de conjugación de SUMO dentro de la proteína SNAP43

Uno de los enfoques más utilizados para entender cómo regula una modificación post-traducciona a una proteína de interés, es generar mutantes de los residuos blanco de la modificación post-traducciona en estudio. En el caso de la SUMOilación, se suele realizar una sustitución de lisina por arginina, que es un

aminoácido con la misma carga neta pero no es sustrato de la conjugación SUMO, con el fin de no alterar significativamente la estructura de la proteína blanco.

Para poder estudiar cómo impacta la SUMOilación en la actividad de SNAP43, primeramente realizamos un análisis *in silico*, mediante un *software* especializado (SUMOsp2), que nos permitió predecir posibles lisinas blanco de SUMOilación, en principio, las más cercanas al consenso conocido (Ψ KxE/D) dentro de la estructura primaria de la proteína. Una lisina cumplía estrictamente con esta secuencia consenso, pero aun así el programa detectó éste y otro posible blanco (Figura R5 A). Consultando la bibliografía encontramos que SNAP43 fue detectada como blanco de conjugación de SUMO en ensayos de espectrometría de masa sitio-específicos para esta modificación (Hendriks 2018), reportándose 13 aminoácidos lisina blancos de conjugación. Decidimos entonces enfocarnos en las 4 lisinas mostradas en la Figura R5 B.

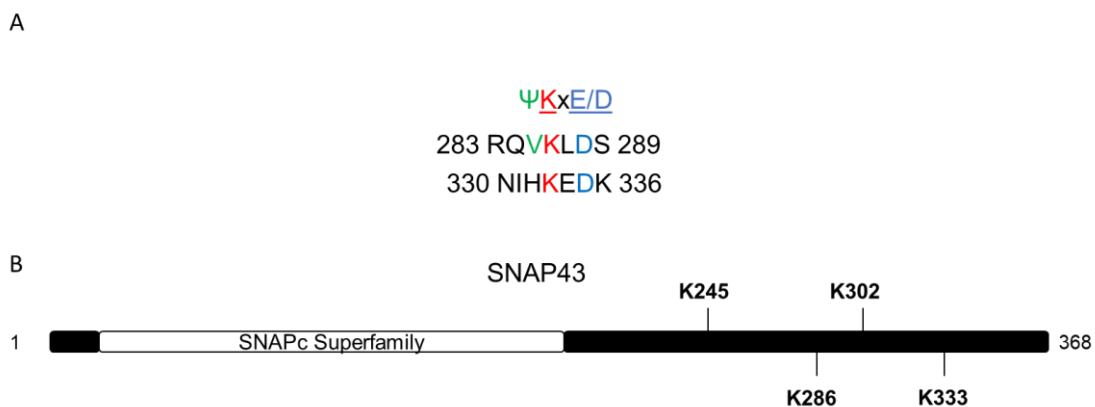


Figura R5: Secuencia consenso y sitios putativos de la proteína SNAP43. A) Sitios putativos de SUMOilación en SNAP43 detectados *in silico* y el motivo consenso de SUMOilación circundante. Ψ aminoácido voluminoso, hidrofóbico; X, cualquier aminoácido; E, ácido glutámico; D, ácido aspártico. B) Esquema de SNAP43, estructura de dominios, en blanco está indicado el dominio SNAPc *Superfamily* y en negro la secuencia que no posee ningún dominio definido y localización de los sitios de conjugación a SUMO detectados *in silico* y por bibliografía.

Reemplazamos las lisinas seleccionadas anteriormente por argininas mediante mutagénesis dirigida y luego evaluamos la SUMOilación de cada uno de estos mutantes respecto de la proteína salvaje (*wild type* o WT). Utilizamos nuevamente cromatografía de afinidad a níquel, para enriquecer nuestras muestras en proteínas

conjugadas a His-SUMO, a partir de extractos de células que sobre-expresaban 6xHis-SUMO2 y las distintas versiones mutante o WT de la proteína SNAP43 fusionadas al epítopo HA, seguido de un ensayo de WB (Figura R6 A) utilizando un anticuerpo anti HA. Del análisis de las simples mutantes, observamos que mientras que las mutaciones en las lisinas 286 y 302 no tuvieron consecuencias detectables, las mutaciones en la lisina 333 y en la 245 produjeron una reducción de la SUMOilación al compararlas con la proteína WT. Sin embargo, la reducción no fue total razón por la cual optamos por generar una doble mutante. La combinación de ambas mutaciones en la misma proteína (K245R/K333R) produjo una marcada reducción de los niveles de SUMOilación de la misma (Figura R6 B). A esta mutante la llamamos “SNAP43 2KR” y fue la utilizada para continuar nuestros estudios.

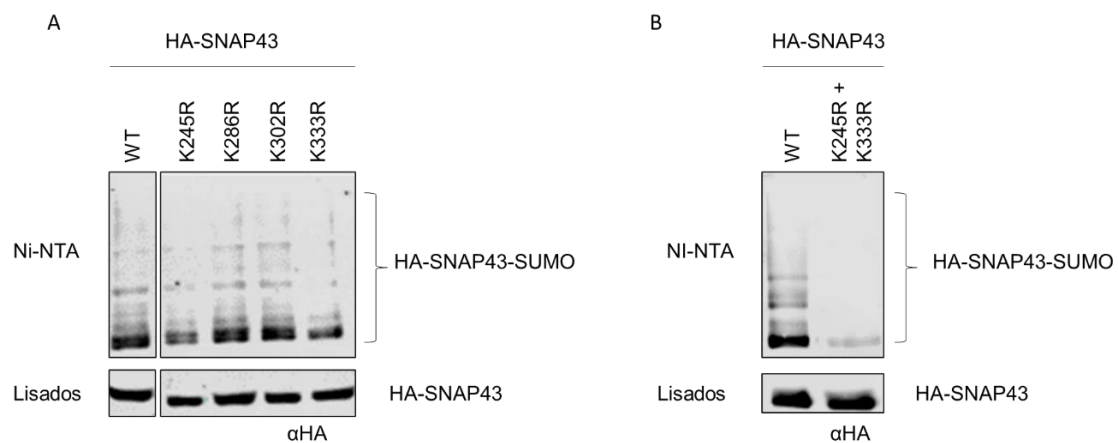


Figura R6: SNAP43 es SUMOilada en las lisinas 245 y 333. Células HEK 293T fueron co-transfectadas con vectores de expresión para V5-Ubc9, HA-SNAP43 WT o mutantes y 6xHis-SUMO2. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizadas por WB con un anticuerpo anti-HA.

Los residuos lisina son blanco de otras modificaciones post-traduccionales, como por ejemplo ubiquitinación, metilación y acetilación. Por lo tanto, hay que considerar que, al generar una versión mutante, podríamos impedir otras modificaciones en los residuos mutados, o incluso en otras lisinas, si existiese algún tipo de regulación cruzada. Dada la similitud entre la SUMOilación y la ubiquitinación, decidimos evaluar si SNAP43 es sustrato de ubiquitinación y, en caso de serlo, si en la mutante SNAP43 2KR esta modificación post-traduccionales se veía alterada. En la figura

R7 se observa que la versión WT es modificada por conjugación de ubiquitina, y que la mutante de SUMOilación si bien no tiene modificado el patrón de ubiquitinación, parecería estar levemente disminuida. Este resultado sugiere que las lisinas mutadas no serían blanco de ubiquitinación, pero no podemos descartar algún tipo de regulación cruzada.

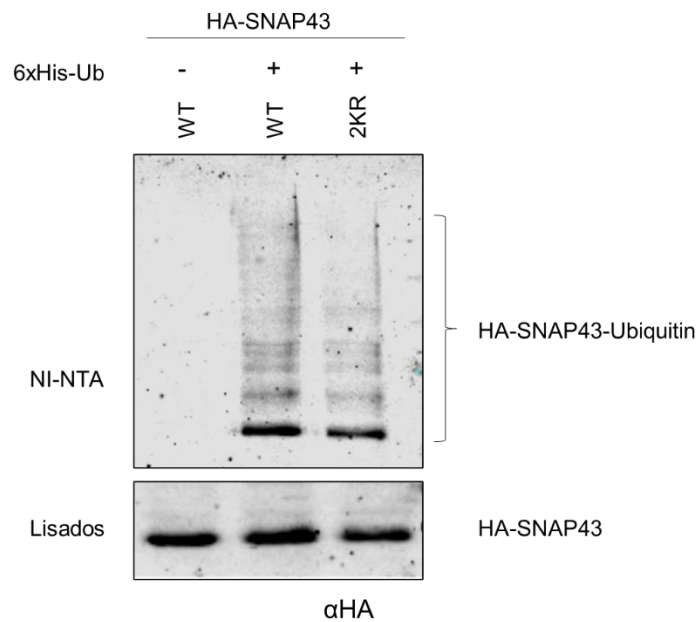


Figura R7: La mutación de las lisinas 245 y 333 no afecta el patrón de ubiquitinación de SNAP43. Células HEK 293T fueron co-transfectadas con vectores de expresión para HA-SNAP43 WT o 2KR, 6xHis-Ubiquitina y V5-Ubc9. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con un anticuerpo anti-HA.

La proteína mutante SNAP43 2KR es incapaz de estimular la transcripción de los snRNAs

Para evaluar si la disminución de la SUMOilación de SNAP43 afecta su función, lo primero que hicimos fue evaluar la transcripción de snRNAs en células HEK 293T co-transfectadas con versiones WT, mutante o plásmido vacío (control) junto con ARN de interferencia (siRNA) específico para depletar a la proteína SNAP43 endógena o un siRNA diseñado contra el mRNA de luciferasa que no está presente en las células utilizadas (control). Los niveles de proteína SNAP43 endógena fueron evaluados

mediante WB (Figura R8 A), mientras que los niveles de snRNAs naciente fueron cuantificados, al igual que en la Figura R2, por RT-qPCR. Dado que los genes de snRNAs de la clase Sm conservan las mismas regiones regulatorias, en el resto de esta tesis nos enfocaremos en los transcritos de U1 y U2.

El silenciamiento de SNAP43, como era previsible, disminuyó los niveles de snRNAs naciente, dado que disminuyó los niveles de una proteína del complejo SNAPc, el cual actúa como factor de transcripción para estos genes. Por su parte, al rescatar dicha depleción con la versión WT de SNAP43 se observó una recuperación de los niveles nacentes de snRNA, mientras que, la versión mutante deficiente de SUMOilación de SNAP43 (SNAP43 2KR) no sólo no logró este efecto, sino que incluso se observaron niveles menores de snRNA naciente que los observados en la condición de silenciamiento sin rescate (Figura R8 B). Estos resultados sugieren por un lado que la SUMOilación de SNAP43 sería necesaria para la participación de esta proteína en la transcripción de snRNAs y, por otro lado, que el mutante SNAP43 2KR podría estar actuando como un dominante negativo.

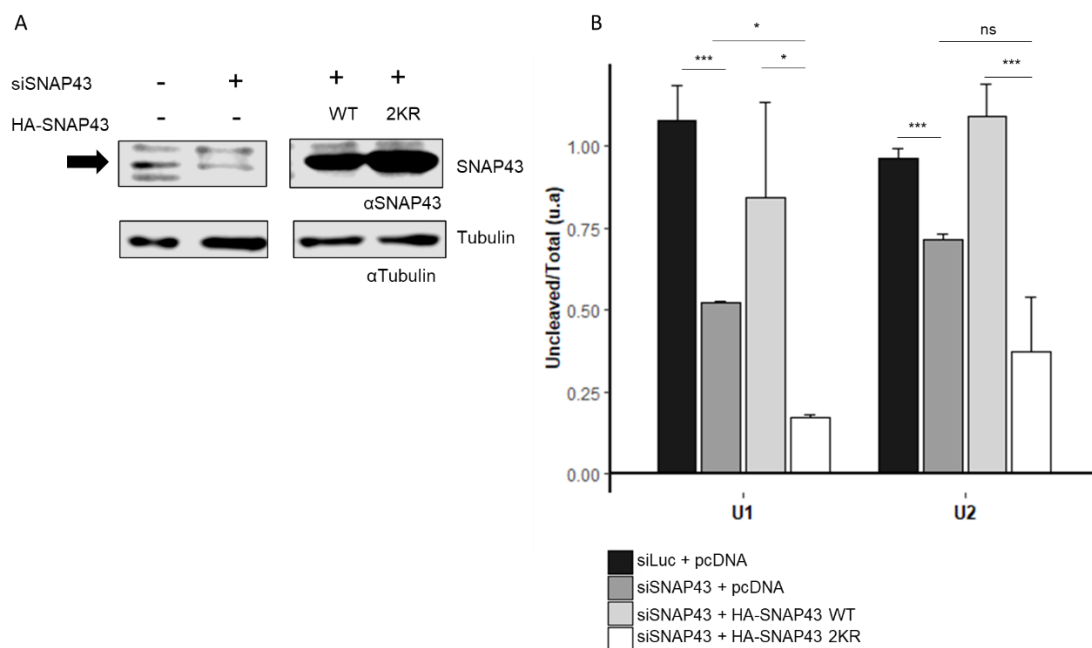


Figura R8: La disminución de SUMOilación de SNAP43 afecta la transcripción de snRNAs en células en cultivo. Células HEK 293T fueron co-transfectadas con un siRNA dirigido contra SNAP43 endógeno (siSNAP43) o siLuc como control (-) junto con vectores de expresión para HA-SNAP43 WT, HA-SNAP43 2KR, o como control, un vector vacío (pcDNA). Luego de 72 h, se

lisaron las células y se extrajeron proteínas y ARN total. A) Los niveles proteicos de SNAP43 fueron analizados mediante WB con anticuerpo anti-SNAP43, la flecha indica la banda correspondiente a SNAP43 según lo observado a la proteína sobre expresada. B) Los niveles de snRNAs nacientes fueron evaluados mediante RT-qPCR (n= 3, *P < 0,05; ***P <0,001; Student's t).

Una manera alternativa de tener un indicio sobre la transcripción es midiendo la cantidad de RNAPII reclutada al promotor de los genes correspondientes, y tal como fue mencionado anteriormente los genes U1 y U2 son transcritos por la RNAPII. Procedimos entonces a transfectar células HEK 293T con plásmidos que codifican para SNAP43 en su versión WT, mutante deficiente de SUMOilación, o como control un plásmido vacío, y luego de 48 h, evaluamos el reclutamiento de RNAPII mediante un ensayo de inmunoprecipitación de la cromatina (ChIP). En la Figura R9 se puede observar que el reclutamiento de la RNAPII a la región promotora de la transcripción de U1 y U2 es similar cuando sobre-expresamos la versión WT de SNAP43 respecto al control. Este resultado podría explicarse por el hecho de que SNAP43 es parte de un complejo multi-proteico y por ende, para aumentar la transcripción de los snRNAs no alcanzaría con sobre-expresar una sola proteína de dicho complejo. Sin embargo, cuando sobre- expresamos la versión deficiente de SUMOilación, en niveles similares a los de la versión WT (Figura R9 A), detectamos una disminución significativa en el reclutamiento de la RNAPII a los promotores en estudio (Figura R9 B). En resumen, los resultados de las Figuras R8 y R9 muestran como disminuye la transcripción de snRNAs en células en cultivo cuando SNAP43 2KR está siendo expresada ya sea en presencia o ausencia de la proteína endógena correspondiente. Por lo tanto, podemos concluir que la versión de SNAP43 deficiente de SUMOilación actúa como dominante negativa.

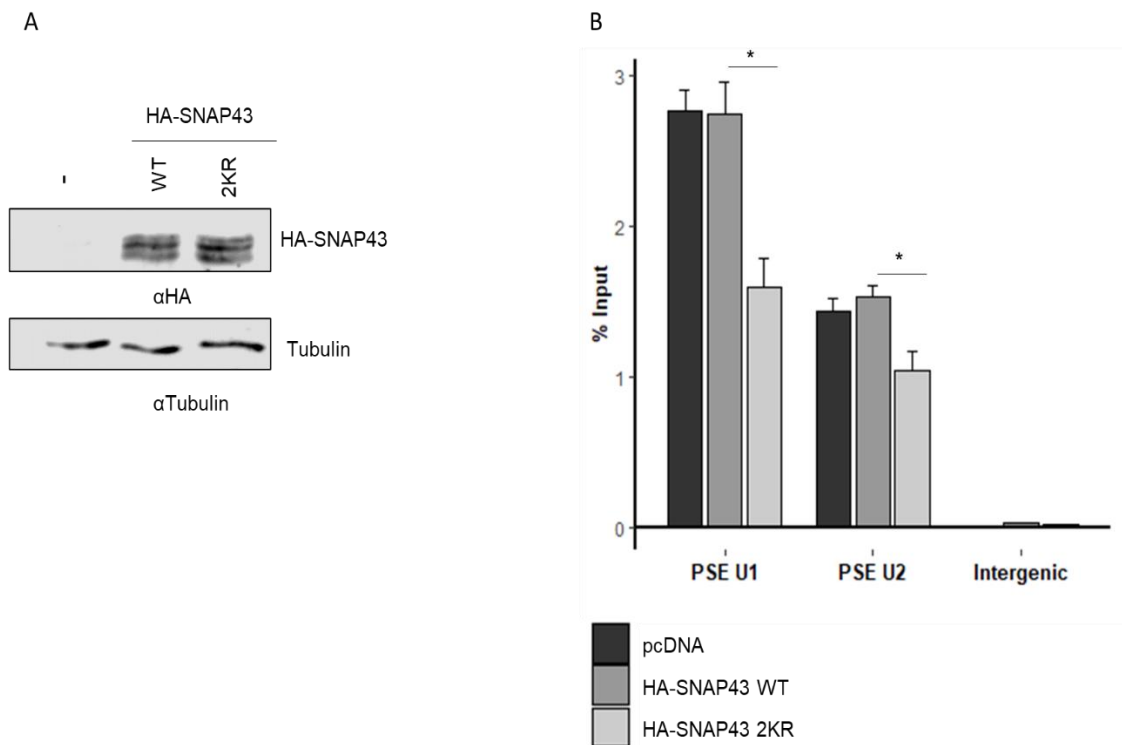


Figura R9: La sobre-expresión de SNAP43 2KR disminuye la abundancia de la RNPII en los promotores de los genes U1 y U2. Celulas HEK293 fueron transfectadas con un plásmido control (pcDNA) o vectores de expresión que codifican para SNAP43 WT o SNAP43 2KR. Luego de 48 h, se realizó una inmunoprecipitación de la cromatina utilizando anticuerpos anti RNAPII. A) Los niveles de HA-SNAP43 expresados a partir de los plásmidos transfectados fueron analizados mediante WB. B) La cuantificación del ADN precipitado fue realizada mediante qPCR con *primers* específicos contra la región correspondiente al *proximal sequence element* (PSE) de U1 y U2, como control analizamos una región intergénica transcripcionalmente inactiva (n=3, *P<0,05; student's t test).

Cabe destacar que en la figura R1, R2 y R3 observamos un aumento de los niveles de snRNAs nacientes al disminuir los niveles de SUMOilación, dichos resultados parecen contradictorios con los mencionados para las figuras R8 y R9. Sin embargo, cuando sobre-expresamos la proteasa de SUMO SENP1 o USPL1 estamos observando muy probablemente una sumatoria de efectos sobre la transcripción de snRNAs, como consecuencias de la disminución de la conjugación a SUMO de una gran variedad de proteínas celulares, mientras que en la figura R8 y R9 observamos la consecuencia en la transcripción de dichos genes de la disminución de SUMOilación de la proteína SNAP43 exclusivamente.

La disminución de SUMOilación de SNAP43 no altera su reclutamiento a la cromatina

Luego de evaluar la transcripción de snRNAs, nos preguntamos si la mutante de SNAP43 deficiente de SUMOilación estaba siendo reclutada a la región promotora de los snRNAs, o si por el contrario, su imposibilidad de ser reclutada a la cromatina era lo que impedía que actuase como factor de transcripción. Para responder este interrogante, células HEK 293T fueron transfectadas con plásmidos que codifican para la versión WT, la mutante deficiente de SUMOilación, o como control un plásmido vacío, y luego de 48 h evaluamos el reclutamiento de SNAP43 mediante un ensayo de inmunoprecipitación de la cromatina (ChIP) utilizando un anticuerpo contra HA. No detectamos un reclutamiento diferencial entre la versión WT o la mutante deficiente de SUMOilación en la región promotora de los genes de U1 y U2 (Figura R10).

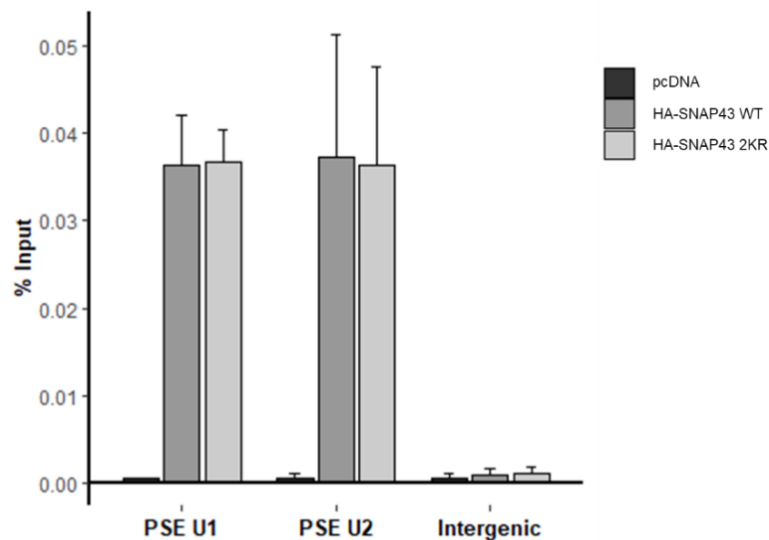


Figura R10: SNAP43 es reclutada a los genes de U1 y U2 independientemente de su nivel de SUMOilación. Células HEK 293T fueron transfectadas con un plásmido control (pcDNA) o vectores de expresión para SNAP43 WT o SNAP43 2KR fusionadas al epítipo HA. Luego de 48 h, se realizó un ChIP con anticuerpo anti HA. La cuantificación del ADN precipitado fue realizada mediante qPCR con *primers* específicos para la region correspondiente al *proximal sequence element* (PSE) de U1 y U2, y como control analizamos una región intergénica (n=3, *P<0,05; student's t test).

En base a estos resultados, proponemos que el reclutamiento de SNAP43 2KR a la cromatina podría estar disruptiendo el complejo SNAP lo que podría ocurrir si esta proteína mutante fuese incapaz de interaccionar con alguna otra proteína del

complejo. Alternativamente, el complejo conteniendo a la proteína SNAP43 2KR podría ser incapaz de reclutar a la RNAPII y/o podría estar reclutando factores represores de la transcripción. Resultará necesario realizar experimentos adicionales para corroborar alguna/s de estas hipótesis.

La sobre-expresión de SNAP43 2KR induce la muerte celular en células HeLa

Quisimos validar los resultados obtenidos anteriormente en otras líneas celulares, optamos entonces por utilizar la línea celular HeLa. En dichas células transfectamos plásmido vacío, plásmidos que codifican para la versión WT o la mutante deficiente de SUMO, y luego de 48 h observamos al microscopio que las células transfectadas con SNAP43 2KR mostraban un fenotipo compatible con apoptosis y muerte celular. Dicho comportamiento se observó reiteradas veces lo que llevó a preguntarnos si la sobre-expresión de SNAP43 2KR podría estar guiando a la muerte celular. Para analizar el efecto observado al microscopio de manera cuantitativa, transfectamos las células con plásmidos que codifican para la versión WT o mutante deficiente de SUMO y realizamos una co-tinción con yoduro de propidio y anexina V-FITC. La anexina V se une a la fosfatidil-serina, que se encuentra en la cara interna de la membrana plasmática en células sanas pero pasa a ocupar la cara externa por “flip-flop” de fosfolípidos durante el proceso de apoptosis. El yoduro de propidio es un colorante que ingresa a las células si éstas se encuentran muertas, por lo tanto, es un colorante de exclusión vital. Esta co-tinción nos permite clasificar a las células en 4 subpoblaciones: las células negativas para propidio y anexina son células vivas y sanas, las células positivas para anexina y negativas para propidio se encuentran en apoptosis temprana y las células positivas para anexina y propidio se encuentran en apoptosis tardía. Por su parte, las células negativas para anexina y positivas para propidio son células muertas, pero por necrosis en lugar de apoptosis (ver cuadrantes en la Figura R 11). Como se observa, la condición de sobre-expresión de SNAP43 WT presenta mayoritariamente células vivas, mientras que la condición de sobre-expresión de SNAP43 2KR muestra un claro aumento en los niveles de muerte celular.

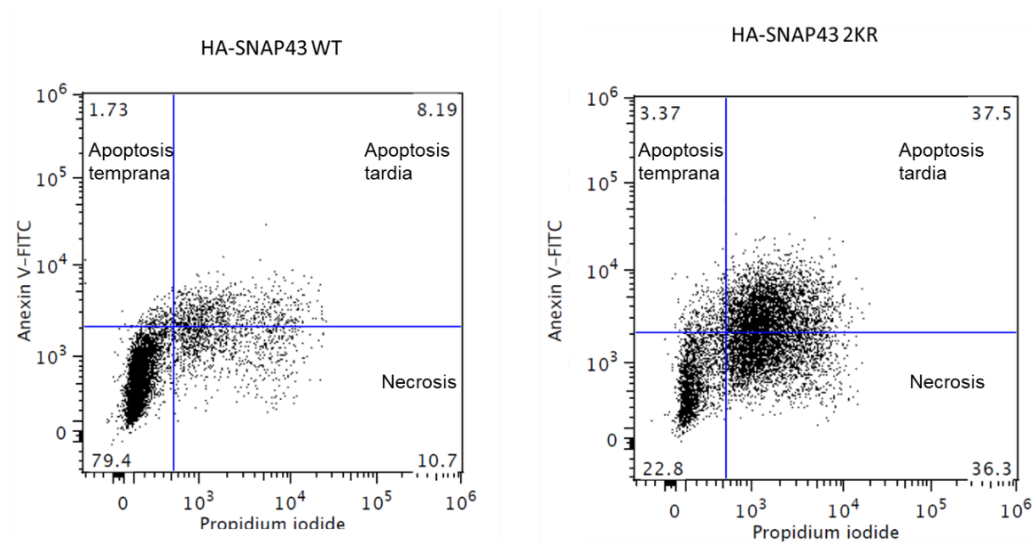


Figura R11: La sobre-expresión de SNAP43 2KR induce la muerte celular en células HeLa.

Transfectamos células HeLa con plásmidos que codifican para SNAP43 WT o mutante deficiente de SUMOilación (2KR). A las 48 h, teñimos las células con ioduro de propidio y Anexina V acoplada a FITC para determinar la proporción de células vivas (ioduro y anexina negativas), células necróticas (ioduro positivas y anexina negativas), células en apoptosis temprana (ioduro negativas y anexina positivas) y en apoptosis tardía (ioduro y anexina positivas).

Considerando que las células adherentes se despegan del sustrato al morir o bien permanecen muy lábilmente unidas al mismo o a las células que las rodean, es posible usar el parámetro de “adhesión” para cuantificar la viabilidad celular. Nos propusimos entonces cuantificar la cantidad de células adheridas a la placa de cultivo luego de sobre-expresar SNAP43 2KR durante 48 horas, como indicativo de cuántas células vivas hay. Para ello utilizamos el método de tinción con cristal violeta el cual interacciona con proteínas y ADN celular, tiñendo todas las células presentes. Iguales cantidades de células HeLa fueron transfectadas con plásmidos que codifican para SNAP43 WT, deficiente de SUMOilación o como control un plásmido vacío y luego de 48 h se lavaron las monocapas respectivas para eliminar todo desecho celular o células muertas y se procedió a realizar la tinción con cristal violeta. Posteriormente, se midió la absorbancia a 595 nm. Como se observa en la figura R12, la sobre-expresión de SNAP43 2KR disminuyó la absorbancia a 595 nm respecto a la sobre-expresión de SNAP43 WT, lo cual indica una menor tinción con cristal violeta en el caso de las

células transfectadas con SNAP43 2KR, es decir, una disminución de la cantidad de células vivas.

Para corroborar que el efecto observado se debía a la expresión de la proteína mutante SNAP43 2KR y no al tipo de plásmido utilizado para el clonado del ADNc correspondiente o bien a una mutación presente en alguna región del plásmido diferente a la región del ADNc sub-clonado, generamos nuevos vectores de expresión, haciendo uso de un esqueleto plasmídico diferente al utilizado anteriormente, y clonando en dicho nuevo esqueleto plasmídico el ADNc correspondiente a SNAP43 tanto WT como mutante deficiente de SUMOilación pero en este caso fusionadas a mCherry, y repetimos el ensayo realizado anteriormente. Los efectos observados fueron los mismos con ambos esqueletos plasmídicos (pCDNA y mCherry C-1) y con las dos formas de “etiquetar” a nuestras proteínas de interés (fusionada a HA vs mCherry) (Figura R12).

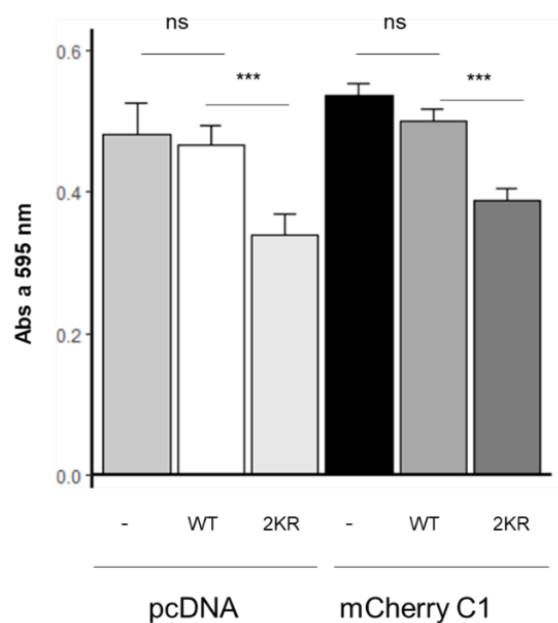


Figura R12: La sobre-expresión de SNAP43 2KR induce la muerte celular en células HeLa independientemente del plásmido de expresión. Células HeLa fueron transfectados con plásmido vacío (pcDNA) o plásmido que codifica para mCherry como control, o bien plásmidos pcDNA conteniendo el ADNc de SNAP43 WT fusionado a HA (WT), plásmido pcDNA conteniendo el ADNc de SNAP43 deficiente de SUMOilación fusionado a HA (2KR), plasmido mCherry (mCherry-C1) conteniendo el ADNc de SNAP43 WT fusionado a mCherry o plásmido mCherry C1 conteniendo el ADNc de SNAP43 2KR fusionado a mCherry. Luego de 48 h, las

monocapas celulares fueron lavadas y sometidas a tinción con cristal violeta, cuantificándose luego la cantidad de cristal violeta remanente en la placa mediante absorbancia a 595 nm (n=3, ***P<0,001; student's t test).

A diferencia del método de tinción con yoduro de propidio y anexina V, que mide directamente la muerte celular tiñendo diferencialmente las células muertas, el método de tinción por cristal violeta tiñe células vivas, es decir es una medición indirecta de la muerte celular. Aun así, hay que destacar que es un método muy reproducible, sencillo y económico que nos permitió corroborar el efecto observado anteriormente (Figura R11). Por otro lado, comprobamos que la inducción de la muerte celular se debe específicamente a la sobre-expresión de la proteína SNAP43 2KR y no al esqueleto plasmídico en el cual el ADN que codifica para esta proteína se encuentra sub-clonado.

Además, corroboramos que al igual que lo que ocurría con SNAP43 2KR fusionada a HA y clonada en pCDNA, SNAP43 2KR fusionada a mCherry y clonada en el vector mCherry C1 es efectivamente deficiente de SUMOilación. Para ello, utilizamos nuevamente la estrategia de cromatografía de afinidad a níquel para purificar proteínas conjugadas a 6xHis-SUMO2 a partir de células transfectadas con el vector de expresión para dicho péptido y las distintas versiones mutante o WT de SNAP43, seguido de un ensayo de WB (Figura R13).

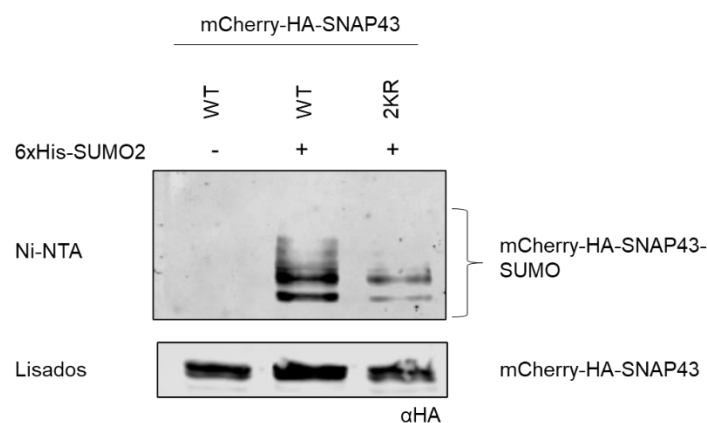


Figura R13: La mutación de las lisinas 245 y 333 de SNAP43 disminuye los niveles de SUMOilación de esta proteína aún fusionada a mCherry. HEK 293T fueron co-transfectadas con vectores de expresión para mCherry-HA-SNAP43 WT o 2KR, V5-Ubc9 y 6xHis-SUMO2, o como control, plásmido vacío. Luego de 48 h, las células fueron lisadas y sometidas a

cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con un anticuerpo anti-HA.

Este ensayo nos permitió validar la disminución de los niveles de SUMOilación de SNAP43 2KR expresada a partir de un contexto plasmídico diferente.

En base a los resultados presentados en figuras R8, R9, R10 y R11 podemos concluir que la sobre-expresión de una versión de SNAP43 deficiente de SUMOilación lleva a una menor transcripción de snRNA, actuando como mutante dominante negativa sobre la proteína endógena WT, y esto podría inducir la muerte celular de manera dependiente del contexto celular.

INTS9

La lisina 607 es sitio de conjugación de SUMO en la proteína INTS9, miembro del complejo *Integrator*

Utilizando la misma estrategia que fue mencionada anteriormente para SNAP43, decidimos estudiar las posibles lisinas que podían ser blanco de SUMOilación en la proteína INTS9. En este caso, el *software* SUMOsp2 detectó una única lisina que cumplía con el motivo consenso canónico de SUMOilación dentro de la estructura primaria de la proteína (Figura R14 A). Procedimos entonces a reemplazar la lisina 607 por arginina mediante mutagénesis dirigida, y luego de corroborar dicha mutación mediante secuenciación, evaluamos los niveles de SUMOilación de la proteína mutante en comparación con su contraparte salvaje (WT). Utilizamos nuevamente cromatografía de afinidad a níquel para purificar proteínas conjugadas a SUMO2 a partir de extractos proteicos de células que sobre-expresaban 6xHis-SUMO2 y las distintas versiones mutante o WT de INTS9 fusionadas al epítopo FLAG. Posteriormente, las muestras fueron analizadas por ensayo de WB utilizando un anticuerpo anti FLAG (Figura R14 B).

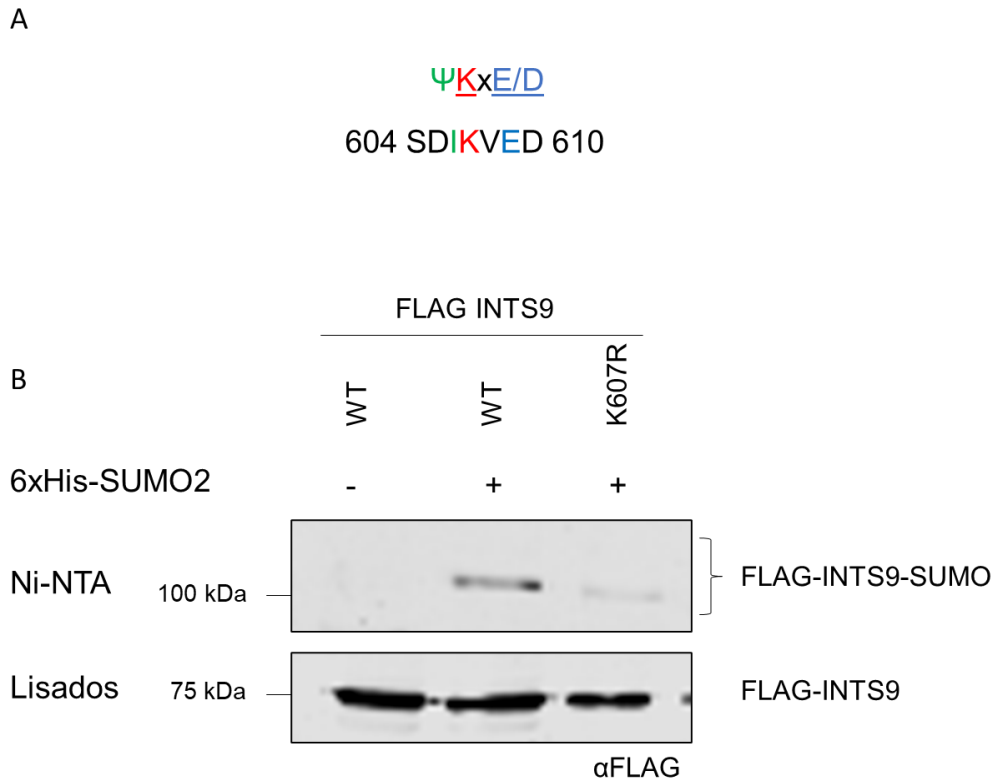


Figura R14: INTS9 es conjugada a SUMO en la lisina 607. A) Sitio putativo de SUMOilación en INTS9 detectado *in silico* y el motivo consenso de SUMOilación circundante. Ψ , aminoácido voluminoso, hidrofóbico; X, cualquier aminoácido; E, ácido glutámico; D, ácido aspártico. B) Células HEK 293T fueron co-transfectadas con vectores de expresión para V5-Ubc9, FLAG-INTS9 WT o mutante (K607R) y 6xHis-SUMO2, o plásmido vacío como control. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con un anticuerpo anti-FLAG.

Como se observa en la figura R14B, la mutante presenta una reducción parcial de la SUMOilación, lo que podría indicar la presencia de alguna otra lisina blanco de conjugación de SUMO aún en ausencia del consenso.

La disminución de SUMOilación de INTS9 no modifica su capacidad para procesar el snRNA naciente en su región 3'

Como fue comentado anteriormente, los snRNA son procesados en el extremo 3' por el complejo denominado *Integrator* y no son poliadenilados en condiciones normales. Sin embargo, la ausencia de alguna proteína del complejo y por ende la imposibilidad de un procesamiento 3' apropiado lleva a la poliadenilación de estas

moléculas de ARN no codificantes, detectándose una acumulación de snRNAs poliadenilados (Yamamoto et al. 2014).

En base a este conocimiento, y con el fin de determinar si la disminución de la SUMOilación de INTS9 altera su función y/o participación en el complejo *Integrator*, evaluamos los niveles de snRNA poliadenilados frente a la sobre-expresión de las distintas variantes de esta proteína, WT o mutante, en células depletadas de la proteína INTS9 endógena. Para ello, células HEK 293T fueron co-transfectadas con un siRNA específico para INTS9 endógeno o un siRNA control (para luciferasa) y plásmidos de expresión para versiones WT o mutante de INTS9, insensibles al silenciamiento. Luego de 72 h, extrajimos el ARN total y realizamos la retrotranscripción del mismo utilizando oligo-dT como *primer*, para obtener únicamente ADNc provenientes de ARNs poliadenilados. Luego, mediante qPCR analizamos los niveles de snRNAs.

Tanto la depleción de INTS9 endógena mediante siRNA como la sobre-expresión de las variantes WT o mutante de SUMOilación de INTS9 fueron confirmadas mediante ensayo de WB usando un anticuerpo anti INTS9 (Figura R15 A).

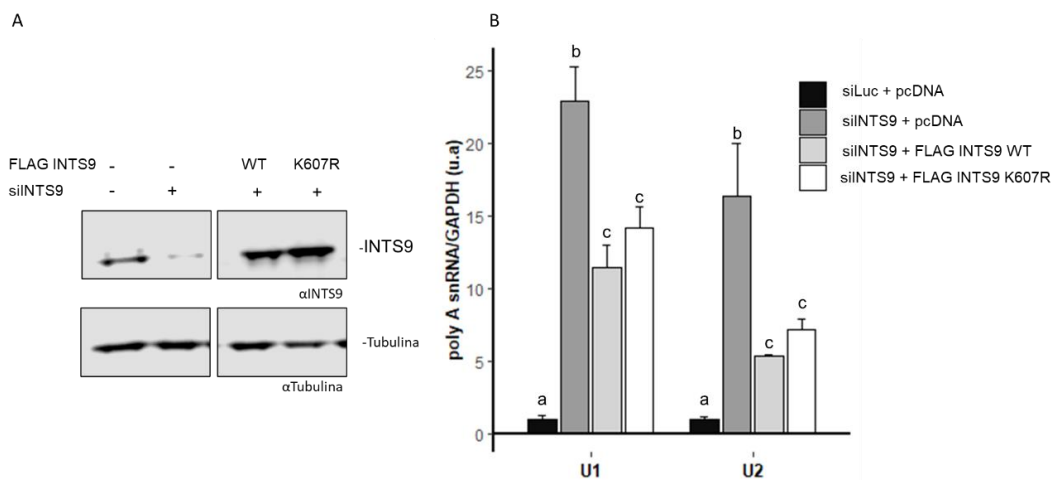


Figura R15: La disminución de SUMOilación de INTS9 no afecta el procesamiento 3' de snRNAs en células en cultivo. Células HEK 293T fueron co-transfectadas con siRNA dirigido contra INTS9 endógeno (siINTS9) o siLuc como control (indicado con signo menos en panel A) y con vectores de expresión para FLAG-INTS9 WT, FLAG INTS9 K607R (ambos insensibles al silenciamiento), o un vector vacío como control (pcDNA). Luego de 72 h, se extrajeron proteínas y ARN total. A) Los niveles proteicos de INTS9 fueron analizados mediante WB con anticuerpo anti-INTS9. B) Los niveles de snRNAs poliadenilados fueron evaluados mediante RT

(oligo dT)-qPCR (n= 3, *P < 0,05; ***P <0,001; ANOVA de un factor seguido de una prueba de Tukey). Las barras rotuladas en la parte superior con la misma letra minúsculas no presentan diferencias significativas, mientras que las indicadas con letras minúsculas diferentes sí.

En concordancia con lo reportado en bibliografía, el silenciamiento de INTS9 aumentó los niveles de snRNAs poliadenilados, mientras que el rescate de este silenciamiento mediante la sobre-expresión de la versión WT o la mutante de SUMOilación de esta proteína logró una reversión parcial similar de dicho aumento (Figura R15 B). Cabe mencionar que la imposibilidad de lograr una reversión completa, es decir de recuperar los niveles basales correspondientes a la condición “siLuc + pcDNA”, podría deberse al tiempo de vida media de los ARNs analizados.

Los resultados mostrados en la figura R15 nos llevan a concluir que la disminución de la SUMOilación de INTS9 lograda mediante la mutación de la lisina 607 no altera la función del complejo *Integrator*, posiblemente por no alterar la participación de esta proteína en dicho complejo. No es posible descartar que una anulación completa de la SUMOilación de esta proteína mediante la anulación de otro/s residuos lisina adicionales tenga consecuencias diferentes.

INTS11

Las lisinas 381, 462 y 475 son sitios de conjugación de SUMO en la proteína INTS11, miembro del complejo *Integrator*.

Desde el descubrimiento del complejo *Integrator* se lo ha relacionado funcionalmente al complejo CPSF (*cleavage polyadenylation specificity factor*) que es el responsable del procesamiento 3' y la poliadenilación de los ARN mensajeros, entre otros ARNs. En particular se describió la similitud entre INTS11 y el factor CPSF-73: ambas proteínas cumplen la misma función en cuanto a que son capaces de clivar el ARN naciente, y poseen un 40% de identidad de secuencia. Teniendo en cuenta dicha similitud se logró desarrollar una mutante catalítica de INTS11, reemplazando el ácido glutámico en la posición 203 por una glutamina (E203Q) (Baillat et al. 2005). Por otro lado, hace muchos años fue reportado que la proteína CPSF-73 era blanco de conjugación SUMO y que esta modificación podría estar regulando su función (Vethantham et al. 2007). Todo esto llevó a preguntarnos qué ocurre con la proteína INTS11 en cuanto a su SUMOilación y la relación de esta MPT con la función de dicha

proteína. Procedimos nuevamente a buscar lisinas blanco de conjugación de SUMO utilizando el *software* SUMOsp2, el cual detecto un único residuo (lisina 462) que se encontraba inmerso en el consenso de SUMOilación (Figura R16 A). En base a esto, y utilizando la estrategia ya mencionada anteriormente, generamos una versión mutante de INTS11 reemplazando la lisina 462 por una arginina y medimos su susceptibilidad a conjugarse a SUMO (Figura R16 B). Como se puede observar, se detecta una disminución de una única banda en el patrón obtenido. Teniendo en cuenta que el algoritmo utilizado por el *software* mencionado sólo identifica lisinas que cumplen con el consenso estricto y canónico de SUMOilación “ΨKxD/E” dentro de la estructura primaria, y que mutar una lisina dentro de dicho consenso no necesariamente anula por completo la SUMOilación, tal como lo mostrado anteriormente para INTS9, optamos entonces por realizar búsquedas manuales de lisinas que cumplan con otro tipo de consensos reportados más recientemente en el campo de la SUMOilación (Hendriks et al. 2018). Observamos que la lisina 381 se encuentra dentro de lo que se conoce como “consenso invertido” (Figura R16 C). Utilizando nuevamente la técnica de mutagénesis dirigida, generamos una doble mutante K381R/K462R, la cual disminuyó notoriamente su nivel de conjugación a SUMO (Figura R16 D).

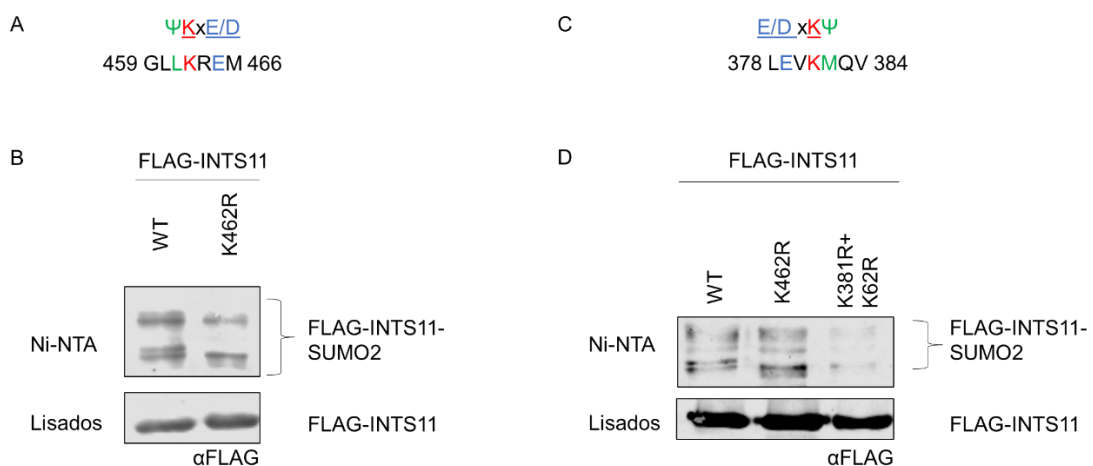


Figura R16: INTS11 es blanco de conjugación a SUMO en sus residuos lisina 381 y 462. A,C) Sitio putativo de SUMOilación en INTS11 detectado *in silico* (A) o en forma manual (C) y el consenso de SUMOilación circundante. Ψ aminoácido voluminoso, hidrofóbico; X, cualquier aminoácido; E, ácido glutámico; D, ácido aspártico. B,D) Células HEK 293T fueron co-

transfectadas con vectores de expresión para FLAG-INTS11 WT o mutantes (según se indica en la parte superior de las imágenes), 6xHis-SUMO2 y v5-Ubc9. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con un anticuerpo anti-FLAG.

Con el objetivo de obtener una mutante que no pueda ser modificada por conjugación de SUMO, decidimos mutar distintas lisinas que fueron descritas por espectrometría de masa sitio-específica para sustratos de conjugación de SUMO. Vale la pena mencionar que la lisina 462, la cual cumple el consenso clásico, fue identificada en diversas publicaciones compiladas en Hendriks et al. 2018. Usando estas referencias generamos distintas triple mutantes y evaluamos si seguían siendo blanco de SUMOilación. La triple mutante que logró disminuir completamente la SUMOilación fue la K381R/K462R/K475R a la cual llamamos “INTS11 3KR”. Además, se puede observar que la mutante catalítica (E203Q) es susceptible de ser modificada por conjugación de SUMO (Figura R17).

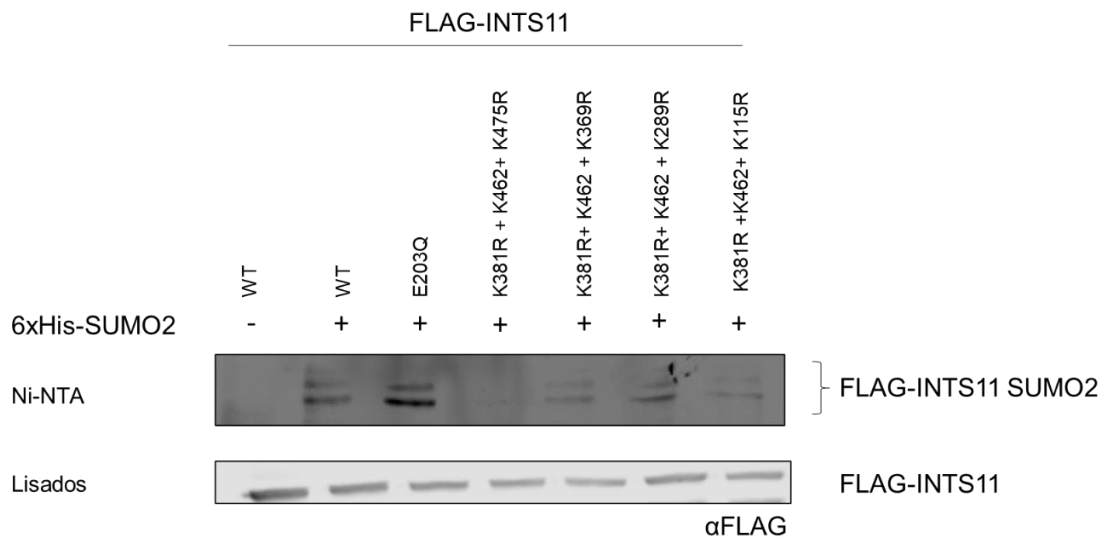


Figura R17: Las lisinas 381, 462 y 475 en la proteína INTS11 son residuos blanco de conjugación de SUMO. Células HEK 293T fueron co-transfectadas con vectores de expresión para V5-Ubc9, FLAG-INTS11 WT o mutantes (indicados en la parte superior de la figura) y 6xHis-SUMO2 o plásmido vacío como control. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con un anticuerpo anti-FLAG.

Cabe destacar que la región cercana a lisina 462 de INTS11 es similar a la de la lisina 462 de CPSF-73, la cual fue reportada como sustrato de conjugación de SUMO. Además, llama la atención que las regiones aledañas a las 3 lisinas seleccionadas presentan similitud con las correspondientes de CPSF-73 (Figura R18 A) y también observamos que se encuentran conservadas en distintos organismos (Figura R18 B). Esto podría indicarnos que dichas lisinas y su conjugación a SUMO cumplen algún rol regulatorio.

A

```

368 L E V K M Q V E Y 486 Ints11
382 L P L K M S V D Y 390 CPSF-73

462 K R E M A Q G L L P 471 Ints11
462 K L A K V M G F L A 471 CPSF-73

468 G L L P E A K K P 476 Ints11
468 G F L A D - K K P 475 CPSF-73
    
```

B

	K381	K462	K475
Homo sapiens (Human)	LEM <small>E</small> GRQVLE <small>V</small> KM <small>Q</small> VEYMSFS	IPVGISLGLIK <small>R</small> EMA--QGLLPEAKKPRL-LHGTLMKDNSFRLV	
Rattus norvegicus (RAT)	LEM <small>E</small> GRQMLE <small>V</small> KM <small>Q</small> VEYMSFS	IPVGISLGLIK <small>R</small> EMV--QGLLPEAKKPRL-LHGTLMKDNFRLV	
Mus musculus (Mouse)	LEM <small>E</small> GRQMLE <small>V</small> KM <small>Q</small> VEYMSFS	IPVGISLGLIK <small>R</small> EMV--QGLLPEAKKPRL-LHGTLMKDNSFRLV	
Gallus gallus (Chicken)	LEM <small>E</small> GRQILE <small>V</small> KM <small>Q</small> VEYMSFS	IPVDISLGLIK <small>R</small> ETA--IGLLPDAKKPKL-MHGTLMKDNSFRLV	
Drosophila melanogaster (Fruit fly)	VEFENRQVVE <small>V</small> KM <small>Q</small> VEYMSFS	IPVDASVSLIK <small>A</small> EARSYNAQPPDPKRRRL-IHGVLMKDNRIMLQI	
Caenorhabditis elegans	IEIDQ-KM <small>E</small> IRL <small>G</small> VEYMSFS	LDIRVPLEKIDRSLS----LDPNP AKSECPFVAELVYDQENERLN	
	:*:: : : *::: *****	: : : : : * : : : : * : : *	

Figura R18: Alineamiento de secuencia aminoacídica entre INTS11 y CPSF-73 humanos y entre INTS11 de diversas especies. A) conservación de las lisinas seleccionadas para INTS11 en CPSF-73. B) Conservación de las lisinas blanco y la secuencia consenso de SUMOilación de INTS11 entre las distintas especies.

La proteína mutante de SUMOilación INTS11 3KR es incapaz de clivar el snRNA naciente en su extremo 3'

Utilizando la misma estrategia que usamos anteriormente para la proteína INTS9, evaluamos que ocurría con la poliadenilación de los snRNAs cuando expresábamos las distintas versiones de INTS11 en células deplegadas de la proteína endógena correspondiente. En este caso, utilizamos dos métodos distintos para generar depleción de INTS11 endógeno: en primer lugar, trabajamos con el mismo método mencionado anteriormente de ARN de interferencia (siRNA); luego pusimos a punto un método de depleción que utiliza la tecnología CRISPR-Cas9 llamada CRISPR de interferencia o "CRISPRi". Este método consiste en expresar en células en cultivo

una Cas9 inactiva o “dCas9” fusionada a un dominio llamado *Krüppel-associated box* (KRAB). El dominio KRAB recluta complejos que modifican la cromatina incluyendo al co-represor KAP1 y al *nucleosome remodeling and deacetylase* (NuRD). Cuando la proteína KRAB-dCas9 es reclutada a la región promotora del gen de interés a silenciar por medio de una guía específica, las histonas son deacetiladas (por acción de NuRD) y además, se favorece la trimetilación de la lisina 9 de la H3 (por acción de KAP1). Todo esto genera un silenciamiento transcripcional específicamente en el promotor blanco (Gilberts et al. 2014). En nuestro caso, diseñamos guías que aparean con el promotor de INTS11. El sistema CRISPRi genera una depleción robusta y es más específico y económico que el sistema de ARN de interferencia por siRNA.

Células HEK 293T en cultivo depletadas de INTS11 por siRNA o por CRISPRi fueron co-transfectadas con plásmidos de expresión para las distintas versiones de INTS11: WT, mutante catalítico “E203Q” o la mutante deficiente de SUMOilación “3KR” (Figura R19)

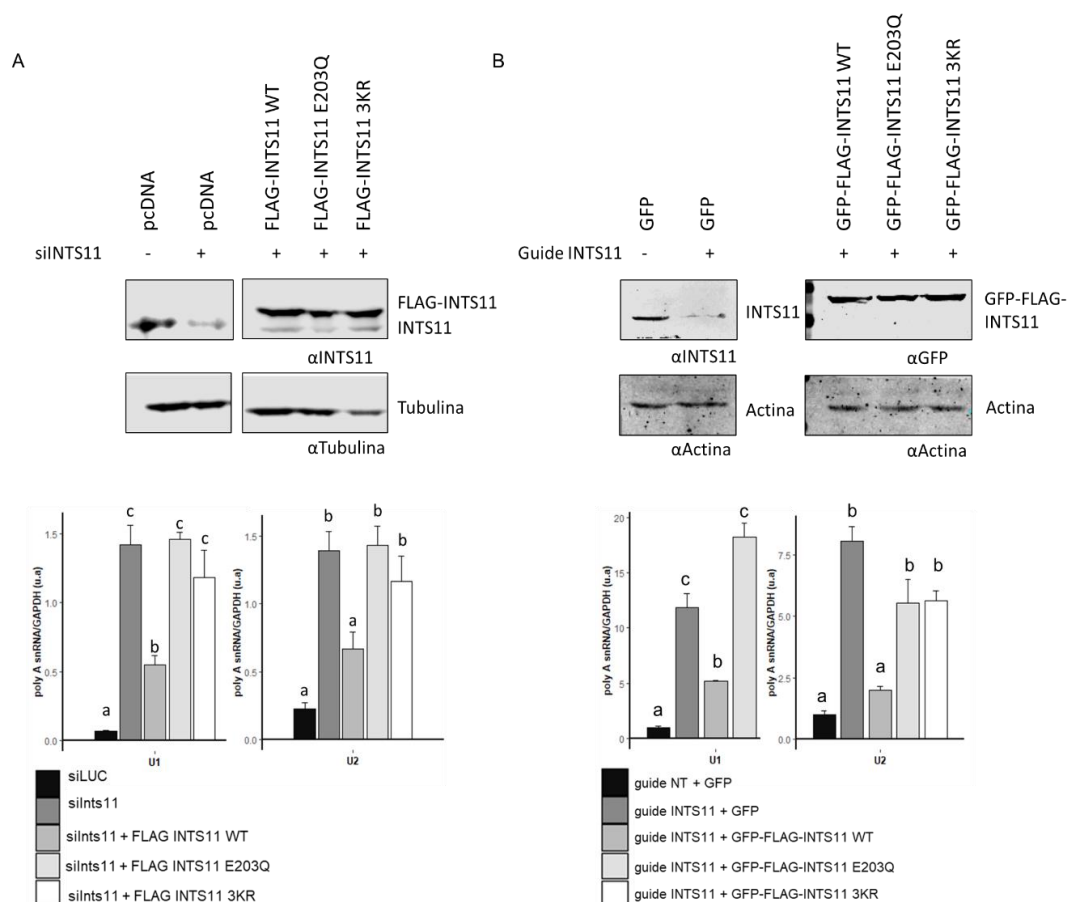


Figura R19: La disminución/falta de SUMOilación de INTS11 afecta el procesamiento 3' de snRNA en células en cultivo. A) Células HEK 293T fueron co-transfectadas con un siRNA dirigido contra INTS11 endógeno (o siLuc como control) y con vectores de expresión insensibles al silenciamiento para FLAG-INTS11 WT, mutantes, o como control, un vector vacío. Luego de 72 h, los niveles de INTS11 fueron analizado mediante WB con anticuerpo anti-INTS11, y los niveles de snRNAs poliadenilados fueron evaluados mediante RT (oligo dT)-qPCR. B) Células HEK 293T fueron co-transfectadas con un vector de expresión para dCas9-KRAB y una guía dirigida hacia el promotor de INTS11, o como control una guía dirigida hacia una región que no se encuentra presente en el genoma humano, y con vectores de expresión para GFP-FLAG-INTS11 WT, mutante de SUMOilación (3KR) o mutante catalítico (E203Q), o como control, la proteína GFP sola. Luego de 72 h, los niveles proteicos de INTS11 fueron analizados mediante WB con anticuerpo anti-INTS11 y los niveles de snRNAs poliadenilados fueron evaluados mediante RT (oligo dT)-qPCR. (n= 4, letras diferentes indican diferencias significativas, $P < 0,05$; ANOVA de un factor seguido de una prueba de Tukey).

En la Figura R19 se observa que en ambas condiciones de silenciamiento de INTS11 se produce un aumento en los niveles de poliadenilación de los snRNAs, acorde a lo esperado según la bibliografía. En ambas condiciones de silenciamiento también la sobre-expresión de la versión WT es capaz de disminuir dicho aumento de la poliadenilación de los snRNAs. Por otro lado, la mutante catalítica al igual que la mutante deficiente de SUMOilación fueron incapaces de procesar correctamente los snRNAs, es decir de revertir el aumento de esta poliadenilación “anómala”. Es importante mencionar que en el caso del silenciamiento con CRISPRi, los rescates fueron realizados con versiones de INTS11 fusionadas a GFP dado que éstas fueron utilizadas en secciones posteriores, principalmente para análisis por microscopía. Tal como muestra la Figura R20 la mutante 3KR fusionada no sólo a FLAG sino también a GFP es incapaz de conjugarse a SUMO, resultado que junto con el de la figura 19B sirvió para corroborar que la fusión a GFP no altera ni la susceptibilidad a conjugarse a SUMO ni la función de la proteína INTS11.

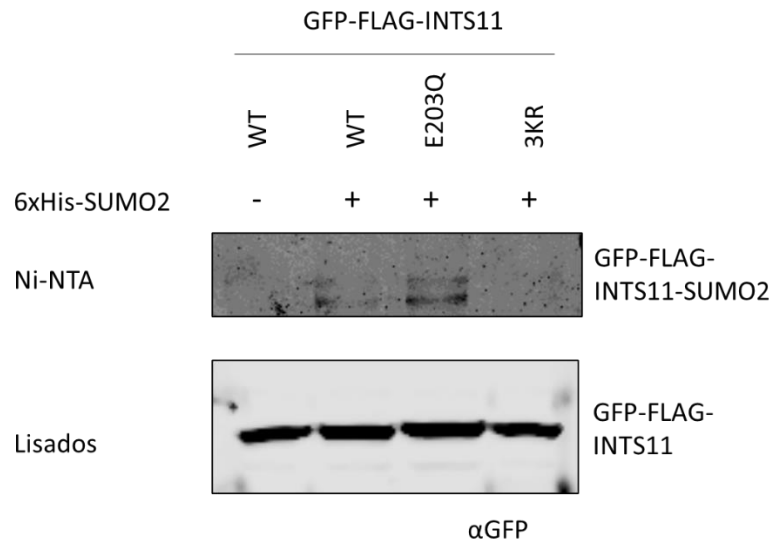


Figura R20: las lisinas 381, 462 y 475 son blanco de SUMOilación en la proteína INTS11 fusionada a GFP. Celulas HEK 293T fueron co-transfectadas con vectores de expresión para V5-Ubc9, GFP-FLAG-INTS11 WT o mutantes indicadas en la parte superior de la figura y 6xHis-SUMO2, o como control plásmido vacío. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con un anticuerpo anti-GFP.

La disminución de SUMOilación de INTS11 no impide su interacción con INTS9 e INTS4

Está extensamente reportada en la bibliografía la importancia de la SUMOilación para la interacción entre distintas proteínas, así como también la interacción entre las proteínas INTS9 e INTS11. Habiendo observado la pérdida de función del complejo *Integrator* ante la deficiencia de SUMOilación de la proteína INTS11 a través del uso de la versión mutante INTS11 3KR (Figura R19), nos preguntamos si esto podía deberse a una pérdida de interacción entre INTS11 e INTS9. Para evaluar esta hipótesis realizamos una cromatografía de afinidad a níquel a partir de lisados totales provenientes de células que expresan 6xHis-Myc-INTS9 y las distintas versiones de FLAG-INTS11. Vale la pena mencionar que las condiciones en las que fue realizada esta cromatografía o *pull down* fueron tales que garantizaron el mantenimiento de interacciones proteína-proteína. La fracción eluida de la resina de níquel fue analizada por WB. Como se observa, todas las versiones de INTS11

analizadas son capaces de interactuar con INTS9 en las condiciones de nuestro ensayo (Figura R21).

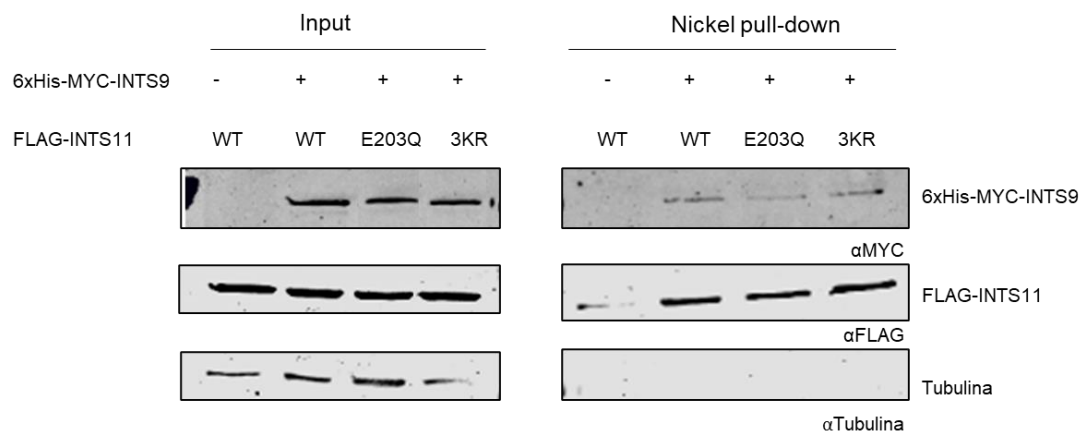


Figura R21: Interacción entre 6xHis-Myc-INTS9 y las distintas variantes de INTS11. HEK 293T fueron co-transfectadas con vectores de expresión para FLAG-INTS11 WT o las mutantes indicadas en la parte superior de la figura y 6xHis-Myc-INTS9. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel no desnaturizante. Alícuotas de las células (input) y los eluatos (Níquel pull-down) fueron analizadas por WB con los anticuerpos indicados debajo de cada panel.

Para corroborar este resultado, realizamos un ensayo de co-inmunoprecipitación utilizando microesferas magnéticas acopladas a anticuerpos anti-GFP (*GFP-trap*, Chromotek) a partir de lisados totales provenientes de células que expresan las distintas versiones de INTS11 fusionadas a GFP. Luego analizamos por WB la presencia de la proteína INTS9 en el precipitado. Dado que en este experimento medimos la interacción de INTS11 con la proteína INTS9 endógena, y no una versión sobre-expresada de esta última, nos parece una aproximación más adecuada para entender qué ocurre a nivel molecular con la interacción entre INTS11 3KR y sus proteínas *partner*. La versión WT de INTS11 fusionada a GFP es capaz de interactuar con INTS9, este resultado es acorde a lo esperado ya que la misma es también capaz de procesar el snRNA naciente según los resultados mostrados en la Figura R19. Por otra parte, con este procedimiento de co-inmunoprecipitación observamos una disminución parcial de co-inmunoprecipitación de INTS9 con la deficiente de SUMOilación (Figura R22). Sin embargo, al observar que la precipitación de la versión GFP-INTS11 3KR también fue levemente menor a la de la versión WT, y teniendo en

cuenta el resultado de la Figura R21, podemos concluir que las distintas variantes de INTS11 no pierden su capacidad de interaccionar con INTS9.

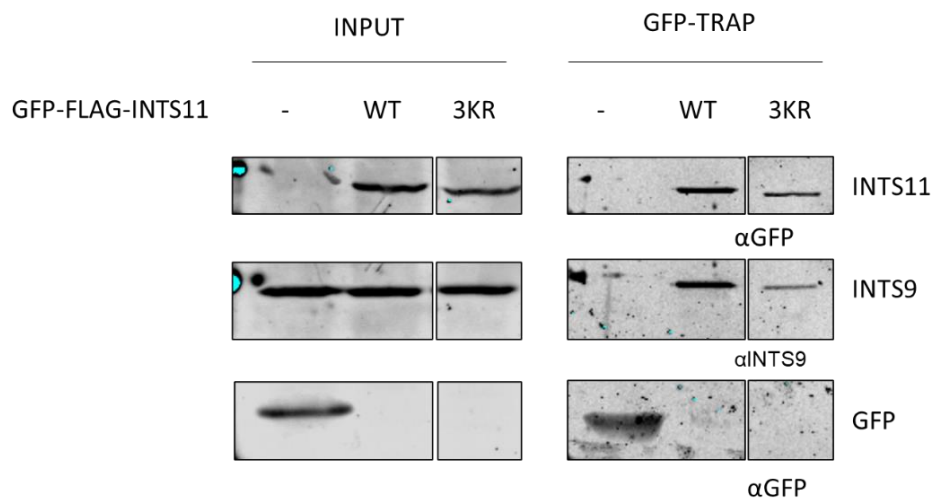


Figura R22: Interacción de variantes de INTS11 con INTS9 endógena. Células HEK293T fueron transfectadas con plásmidos de expresión para las variantes de GFP-INTS11 indicadas en la parte superior de la figura o para GFP sola. Los lisados celulares fueron inmunoprecipitados con GFP-trap y analizados por WB con los anticuerpos indicados debajo de cada panel.

Recientemente se ha descrito que distintas subunidades del complejo *Integrator* son capaces de formar módulos independientes. El módulo catalítico formado por INTS4/INTS9/INTS11 fue el primero en identificarse (Albrecht et al. 2018). Como INTS11 3KR es capaz de interaccionar con INTS9, nos preguntamos si era capaz de ensamblarse en el módulo catalítico. Para evaluarlo realizamos una co-inmunoprecipitación a partir de lisados totales provenientes de células que expresan HA-INTS4 y las distintas versiones de FLAG-INTS11 (Figura R23).

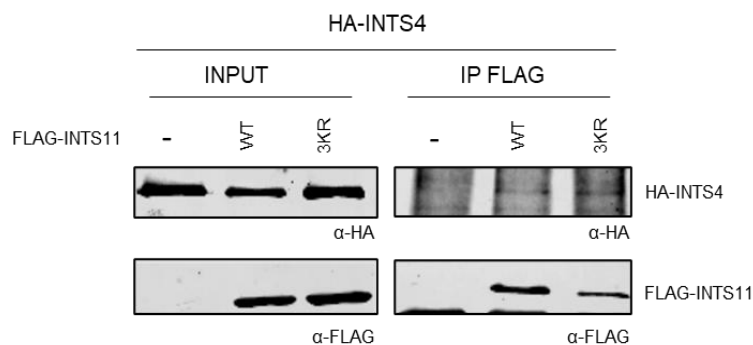


Figura R23: Interacción entre HA-INTS4 y las distintas variantes de INTS11. Células HEK 293T fueron co-transfectadas con vectores de expresión para FLAG-INTS11 WT o las mutantes

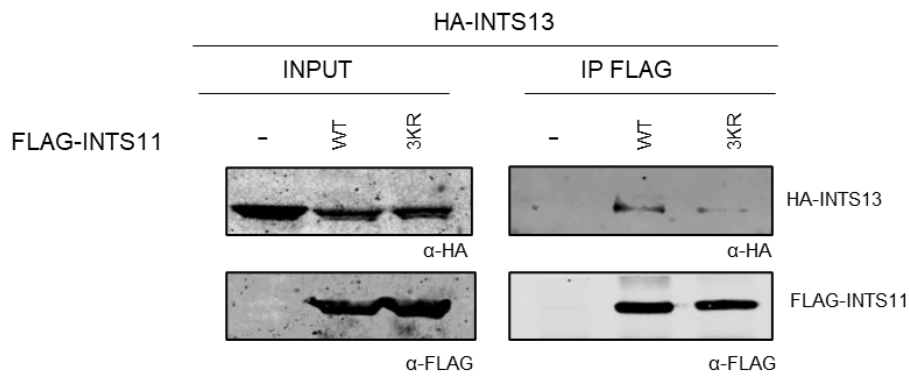
indicadas en la parte superior del panel y HA-INTS4. Luego de 48 h, los lisados celulares fueron inmunoprecipitados con anticuerpo anti FLAG y analizados por WB con los anticuerpos indicados debajo de cada panel.

Observamos que la variante incapaz de SUMOilarse puede interaccionar con INTS4. Por lo cual, a partir de las figuras R21, R22 y R23 podemos concluir que la SUMOilación de INTS11 no regula la interacción entre las subunidades del modulo catalítico ni, probablemente, el ensamblado dicho módulo.

La disminución de la SUMOilación de INTS11 interfiere con su interacción con INTS13 e INTS14

Debido a que lo observado en la figura R19 no se debe a una imposibilidad de INTS11 3KR para ensamblarse en el módulo catalítico, nos preguntamos si el efecto observado a nivel del procesamiento de los pre-snRNAs podría deberse a una pérdida de interacción entre esta variante de INTS11 y otras proteínas del complejo *Integrator*. Para ello, realizamos una estrategia similar a la mencionada anteriormente, co-inmunoprecipitación a partir de lisados totales provenientes de células que expresan las distintas versiones de FLAG-INTS11 y HA-INTS13 (Figura R24 A) o GFP-FLAG-INTS11 y HA-INTS14 (Figura R24 B).

A



B

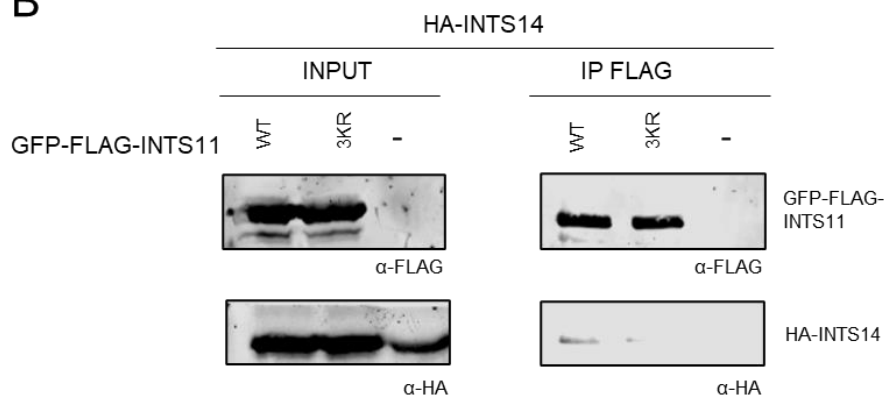


Figura R24: Interacción entre HA-INTS13 o HA-INTS14 y las distintas variantes de INTS11. (A) Células HEK 293T fueron co-transfectadas con vectores de expresión para FLAG-INTS11 WT o las mutantes indicadas en la parte superior del panel y HA-INTS13. Luego de 48 h, los lisados celulares fueron inmunoprecipitados con anticuerpo anti FLAG y analizados por WB con los anticuerpos indicados debajo de cada panel. (B) Células HEK 293T fueron co-transfectadas con vectores de expresión para GFP-FLAG-INTS11 WT o las mutantes indicadas en la parte superior del panel y HA-INTS14. Luego de 48 h, los lisados celulares fueron inmunoprecipitados con anticuerpo anti FLAG y analizados por WB con los anticuerpos indicados debajo de cada panel.

A diferencia de lo observado con las proteínas que forman el módulo catalítico, la variante INTS11 3KR muestra una disminución en la interacción con las proteínas INTS13 e INTS14 respecto a la variante WT. Vale la pena destacar que recientemente se describió un módulo formado por las subunidades INTS10/INTS13/INTS14, el cual se puede unir al ARN naciente y es el responsable de estabilizar la interacción del módulo catalítico con el resto del complejo, lo cual es necesario para que ocurra el procesamiento del ARN naciente (Sabath et al. 2020). Por lo observado en la figura

R24, podemos concluir que la SUMOilación de INTS11 es necesaria para su ensamblado dentro del complejo *Integrator*.

La SUMOilación de INTS11 regula su localización sub-celular

Con el objetivo de entender las consecuencias de la SUMOilación sobre la actividad de INTS11, y teniendo en cuenta la co-transcripcionalidad del procesamiento de los snRNAs, estudiamos el reclutamiento a la cromatina de las distintas variantes de la proteína mencionada. En células HEK 293T transfectamos plásmidos que codifican para INTS11 en su versión WT o mutante deficiente de SUMOilación, o como control un plásmido vacío. Posteriormente, evaluamos el reclutamiento de INTS11 mediante un ensayo de inmunoprecipitación de la cromatina (ChIP) utilizando un anticuerpo contra FLAG. Para la versión mutante analizada, detectamos una disminución de su reclutamiento a la region 3' box del gen de U2 (Figura R25).

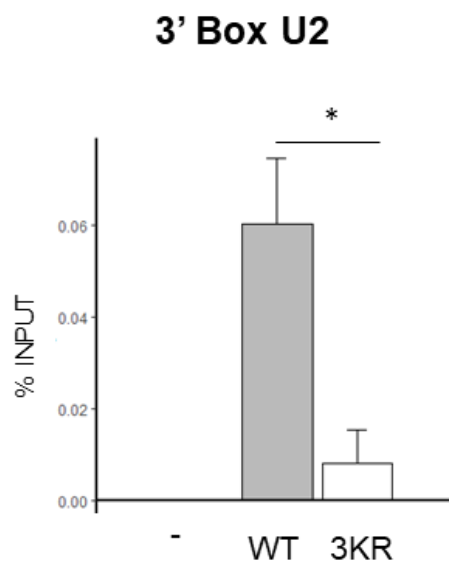


Figura R25: La versión mutante de INTS11 deficiente de SUMOilación no es reclutada a la region 3' box del gen U2. Células HEK 293T fueron transfectadas con un plásmido control (pcDNA) o vectores de expresión para INTS11 WT, o la mutante deficiente de SUMOilación (3KR). Luego de 48 h, se realizó un ChIP con anticuerpo anti FLAG. La cuantificación del ADN precipitado fue realizada mediante qPCR con *primers* específicos para la region 3' Box del gen U2 (n=3, *P<0,05; student's t test).

Al observar que la mutante deficiente de SUMOilación de INTS11 no es reclutada al 3' box de U2, nos preguntamos qué ocurre con el heterodímero

INTS9/INTS11. Para responder esta pregunta realizamos ensayos de ChIP usando anticuerpo anti INTS9 o anti INTS11 a partir de células HEK 293T depletadas de la proteína endógena INTS11 mediante CRISPRi y que sobre-expresan la variante WT o 3KR (Figura R26 A).

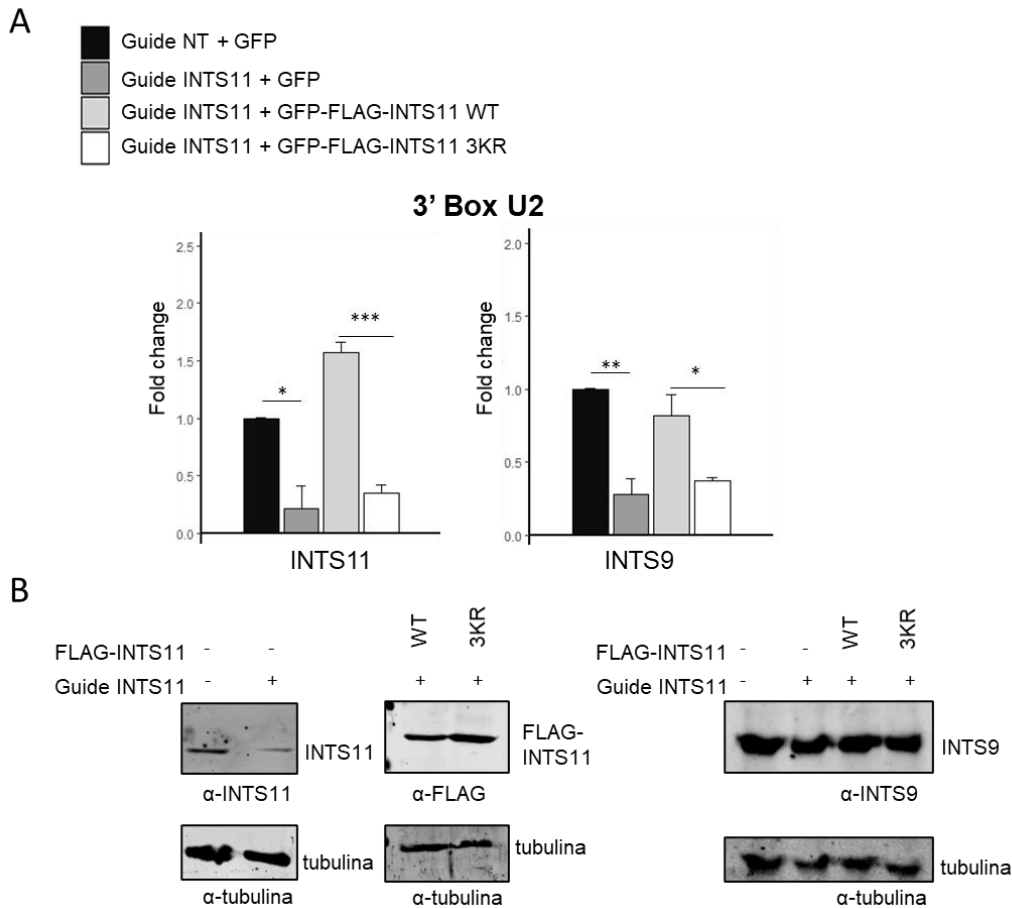


Figura R26: INTS9 no es reclutadas a la region 3' box del gen U2 en ausencia de INTS11 o en presencia de INTS11 3KR. Células HEK 293T fueron co-transfectadas con un vector de expresión para dCas9-KRAB y una guía dirigida hacia el promotor de INTS11, o como control una guía dirigida hacia una región que no se encuentra presente en el genoma humano, y con vectores de expresión para FLAG-INTS11 WT, mutante de SUMOilación (3KR) o como control de la depleción GFP. A) Luego de 72 h, se realizó un ChIP con anticuerpo anti INTS9 o anti INTS11. La cuantificación del ADN precipitado fue realizada mediante qPCR con *primers* específicos para la region 3' Box de U2 (n=3, *P<0,05; student's t test). B) Los niveles proteicos de INTS11 e INTS9 fueron analizados mediante WB con anticuerpos anti INTS11 e INTS9 respectivamente. Los niveles de las variantes WT o 3KR sobre-expresadas fueron analizados mediante WB con anticuerpos anti FLAG

Los resultados obtenidos muestran que ante la depleción de INTS11 detectamos menor reclutamiento de la proteína al 3' box del gen de U2 (Figura R26 A), lo cual es esperable porque partimos de menos proteína presente en las células (Figura R26 B). Por otro lado, detectamos un menor reclutamiento de INTS9 en ausencia de INTS11 endógeno (Figura R26 A), lo que nos llevó a concluir que para el correcto reclutamiento de INTS9 se necesita la presencia de INTS11. Además, observamos un menor reclutamiento de INTS9 en células que sobre-expresan la variante deficiente de SUMOilación de INTS11 respecto a aquellas que sobre expresan la variante WT (Figura R26 A). Al observar una disminución del reclutamiento de INTS11 3KR a la cromatina (Figura 25 y 26 A) es coherente que INTS9 tampoco este siendo reclutada, dado que INTS11 3KR mantiene la interacción con INTS9 (Figuras R21 y R22). Esto nos permite concluir que la SUMOilación de INTS11 regula el reclutamiento del heterodímero INTS9/INTS11. Vale la pena destacar que controlamos que los niveles de sobre-expresión de las variantes de INTS11 (WT o 3KR) sean similares, y que no haya alteraciones significativas en los niveles proteicos de INTS9 entre las distintas condiciones (Figura R26 B).

Observando este resultado decidimos indagar qué ocurre con la localización sub-celular de INTS11 WT y la mutante deficiente de SUMOilación. Para ello, transfectamos células HeLa con las variantes de INTS11 fusionadas a GFP y analizamos por microscopía confocal su localización sub-celular. Como se puede observar, la versión WT se encuentra principalmente en el núcleo. Sin embargo, observamos una distribución más citoplasmática de la variante deficiente de SUMOilación (Figura R27 A). Por medio de una cuantificación llevada a cabo con el *software Cell Profiler* pudimos concluir que INTS11 3KR está significativamente menos localizada en el núcleo, aumentando su localización en el citoplasma (Figura R27 B).

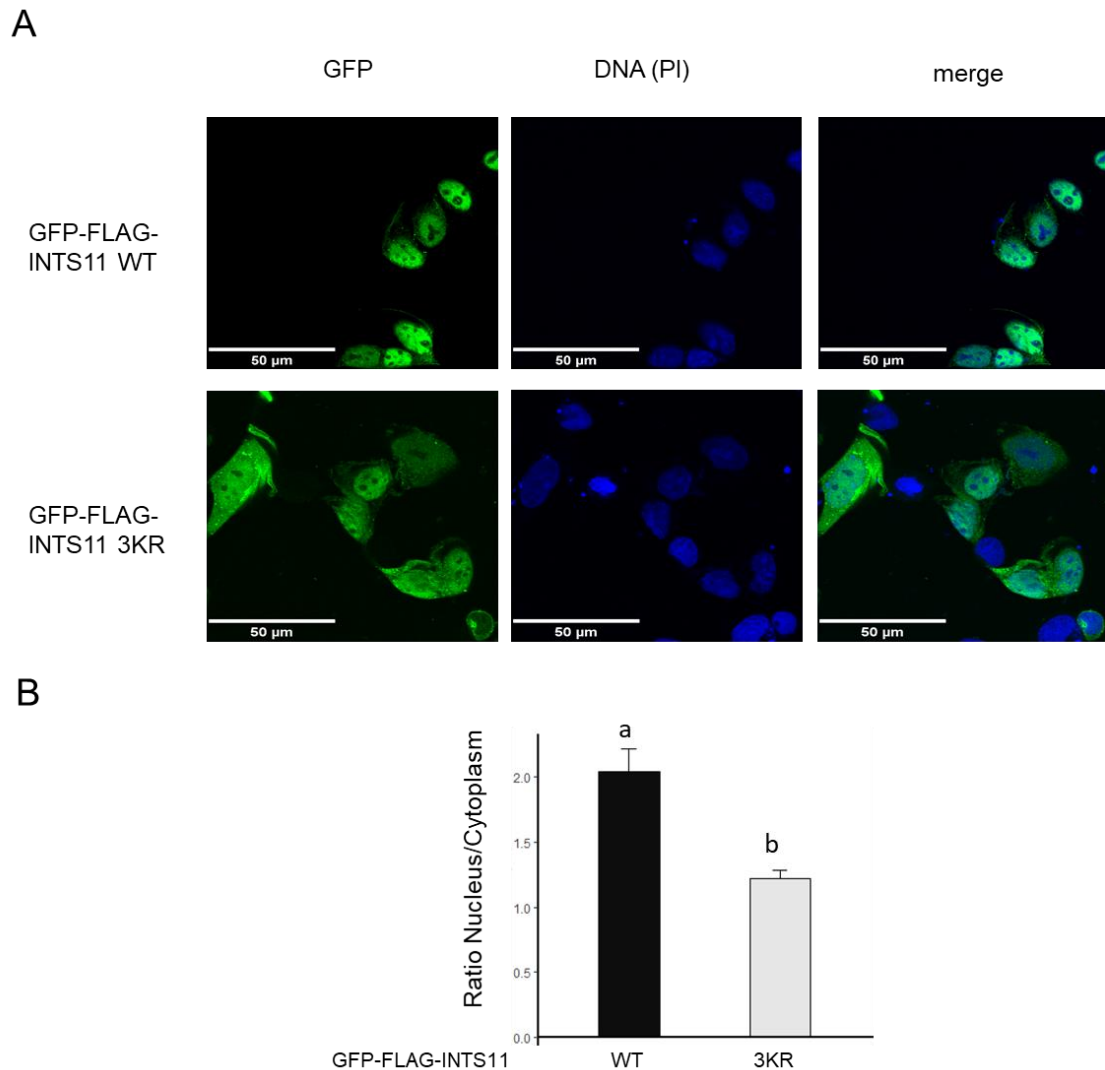


Figura R27: Localización sub-celular de las variantes de INTS11. A) Células HeLa fueron transfectadas con las variantes GFP-FLAG-INTS11 WT o 3KR, y 24h más tarde fueron fijadas y analizadas por microscopía de fluorescencia confocal. Se muestran las imágenes obtenidas por esta técnica, donde se observan los núcleos teñidos con Ioduro de Propidio (azul) (PI) y marca fluorescente verde correspondiente a la localización de cada variante de INTS11 fusionada a GFP B) Las imágenes tomadas por microscopía confocal fueron analizadas con *Cell Profiler*. Se desarrolló una línea de comando para identificar de manera automática los núcleos de todas las células y se cuantificó las intensidades medias de los núcleos y citoplasmas. (n=150, ***P<0,001; ANOVA de un factor seguido de una prueba de Tukey).

En la bibliografía existen distintos ejemplos de proteínas que al alterarse su SUMOilación cambian su localización sub-celular. Sin embargo, como se discutió con anterioridad, al realizar mutaciones puntuales, en este caso el reemplazo de lisinas por argininas, podemos alterar la conformación, estabilidad o incluso “señales”

importantes como la señal de localización nuclear (NLS) de la proteína en cuestión. Según datos bibliográficos, INTS11 contiene una NLS entre los residuos 460-479 (Mackmull et al. 2017), ante lo cual decidimos evaluar mediante un *software* llamado *cNLS Mapper* (<http://nls-mapper.iab.keio.ac.jp/>) si las mutaciones de la lisina 462 y 475 presentes en el INTS11 3KR alteraban dicha NLS. El predictor mencionado encontró diversos NLS en la proteína INTS11 WT, entre los cuales se encuentra el reportado por la bibliografía (460-479). Llamativamente, dicha señal no fue predicha (detectada) para la variante INTS11 3KR (Figura R28). El análisis de las simples mutantes nos llevó a concluir que la mutación de la lisina 462 por una arginina impidió la detección del NLS por parte del *software* específico (Figura R28). Vale la pena mencionar que los primeros trabajos tendientes a establecer un consenso para la conjugación SUMO fueron los que mostraron una relación entre NLS y SUMO. Dichos estudios definieron que la lisina blanco debía enmarcarse en una secuencia aminoacídica específica nombrada anteriormente (Ψ KxE/D) y encontrarse en un NLS aunque, como se mencionó en la introducción, la importancia de la NLS fue lo primero en ser quitado del consenso.

Nos preguntamos entonces si la mera mutación de la lisina 462, la cual se encuentra dentro del consenso y ha sido reportada en análisis proteómicos específicos para la conjugación de SUMO como lisina blanco de SUMOilación (Hendriks et al. 2018), genera un cambio en el NLS explicando así los efectos observados, o si los mismos son consecuencia de anular específicamente la SUMOilación en dicha lisina.

Ints11 WT				Ints11 K381R			
Pos	Sequence	NLS	Score	Pos	Sequence	NLS	Score
469-480	LLPEAKKPRLL	Monopartite	8,5	469-480	LLPEAKKPRLL	Monopartite	8,5
460-479	LLKREMAQGLLPEAKPRL	Bipartite	12	460-479	LLKREMAQGLLPEAKPRL	Bipartite	12
565-597	EDPGTKVLLVSWTYQDEELGSFLTSLKKGLP	Bipartite	5,2	565-597	EDPGTKVLLVSWTYQDEELGSFLTSLKKGLP	Bipartite	5,2

Ints11 3KR				Ints11 K462R			
Pos	Sequence	NLS	Score	Pos	Sequence	NLS	Score
469-480	LLPEAKRPRLL	Monopartite	10	469-480	LLPEAKKPRLL	Monopartite	8,5
460-479	LLRREMAQGLLPEAKRPRLL	Bipartite	ND	460-479	LLRREMAQGLLPEAKPRL	Bipartite	ND
565-597	EDPGTKVLLVSWTYQDEELGSFLTSLKKGLP	Bipartite	5,2	565-597	EDPGTKVLLVSWTYQDEELGSFLTSLKKGLP	Bipartite	5,2

Ints E464A 2KR				Ints11 K475R			
Pos	Sequence	NLS	Score	Pos	Sequence	NLS	Score
469-480	LLPEAKRPRLL	Monopartite	10	469-480	LLPEAKRPRLL	Monopartite	10
460-479	LLKRAMAQGLLPEAKRPRLL	Bipartite	15,3	460-479	LLKREMAQGLLPEAKRPRLL	Bipartite	13,5
565-597	EDPGTKVLLVSWTYQDEELGSFLTSLKKGLP	Bipartite	5,2	565-597	EDPGTKVLLVSWTYQDEELGSFLTSLKKGLP	Bipartite	5,2

Figura R28: Señales de localización nuclear obtenidas a partir del *software* cNLS. Las versiones indicadas en la parte superior de cada panel fueron analizadas *in silico* con el

software cNLS Mapper. Se muestran las distintas señales de localización nuclear con sus valores obtenidos. Mayor *score* indica mayor probabilidad de ser un NLS, ND indica la no detección de un NLS.

Para estudiar el efecto de la SUMOilación pero sin alterar el score predicho para una dada NLS, generamos una mutante en la que remplazamos el ácido glutámico en la posición 464 por una alanina. Es sabido tanto por la literatura como por resultados previos de nuestro laboratorio, que esta mutación del motivo consenso de SUMOilación es impide la conjugación de SUMO a la proteína en cuestión aún sin encontrarse alterada la lisina blanco de conjugación presente en dicho consenso (Riso et al., 2013, Pozzi et al., 2017). De esta manera generemos un nuevo triple mutante K381R/E464A/K475R a la que denominamos “EA + 2KR”. Como era de esperar dicha mutante no modifica su NLS (Figura R28). Corroboramos luego la imposibilidad de dicha mutante de ser modificada por conjugación de SUMO (Figura R29), lo que refuerza la idea de que la lisina 462 es un sitio *bona fide* de SUMOilación ya que no sólo la mutación de dicha lisina sino la mutación de las secuencias consenso circundantes aun dejando intacta las lisinas disminuye drásticamente la conjugación de SUMO a INTS11. Por otro lado, corroboramos que esta nueva triple mutante deficiente de SUMOilación (EA+2KR) que no tiene alterada la correspondiente NLS tampoco logra un procesamiento 3' adecuado de los snRNAs nacientes (Figura R30).

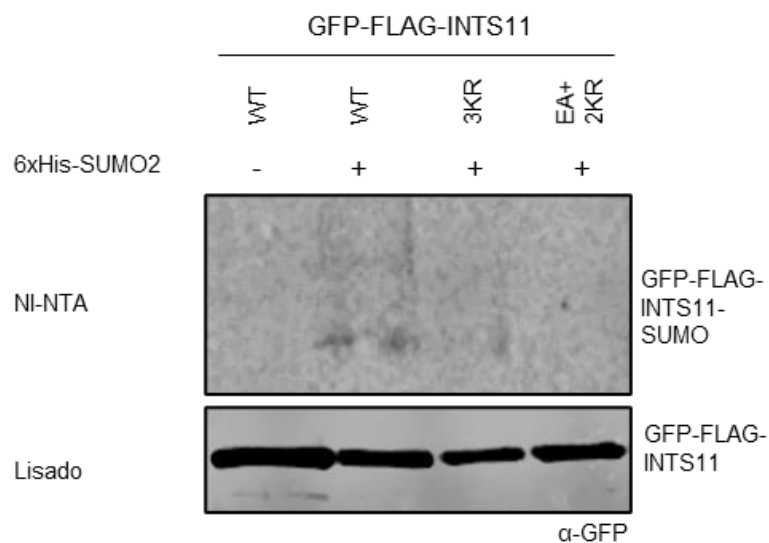


Figura R29: INTS11 EA+2KR no es susceptible de conjugarse a SUMO. Células HEK 293T fueron co-transfectadas con vectores de expresión para V5-Ubc9, GFP-FLAG-INTS11 WT, o las

mutantes indicadas y 6xHis-SUMO2, o como control plásmido vacío. Luego de 48 h, las células fueron lisadas y sometidas a cromatografía de afinidad a Níquel (Ni-NTA). Alícuotas de los lisados y los eluatos (Ni-NTA) fueron analizados por WB con un anticuerpo anti-GFP.

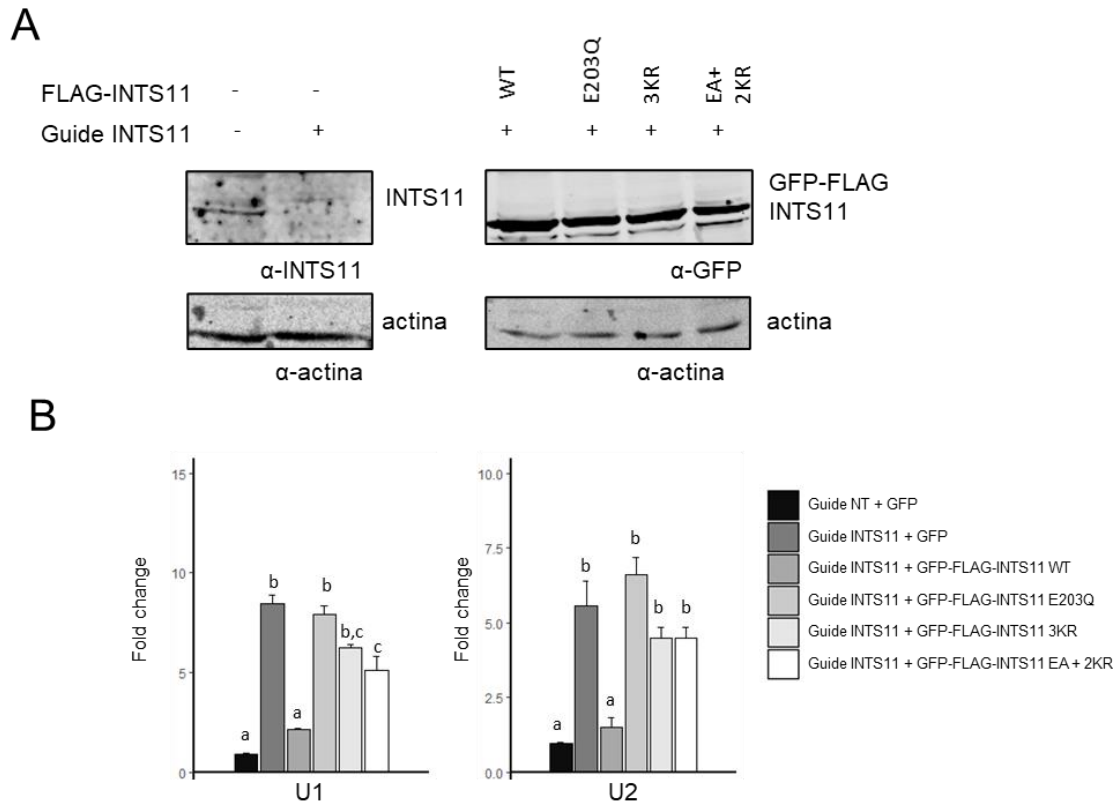
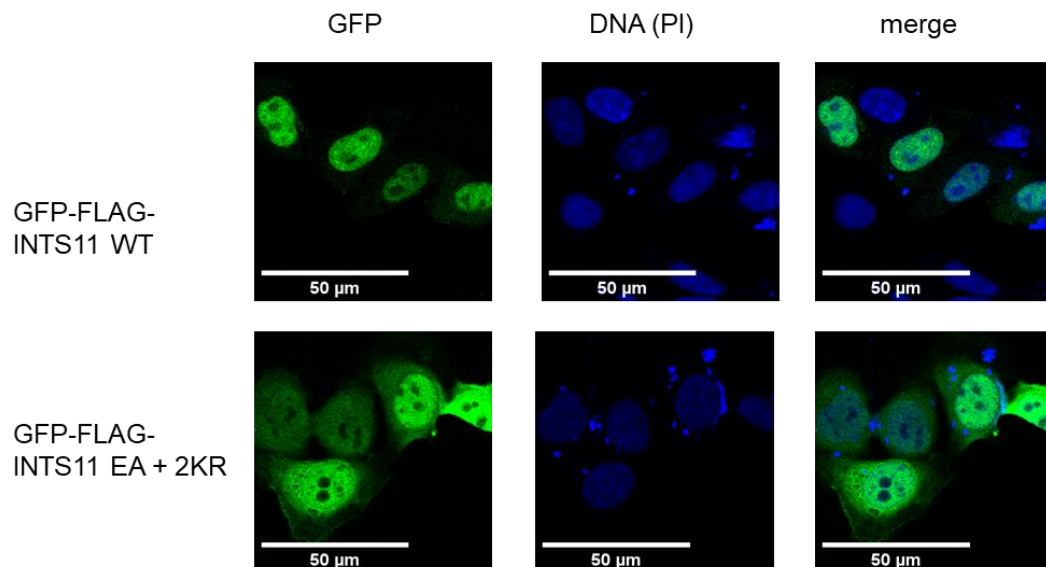


Figura R30: INTS11 EA+2KR es incapaz de llevar a cabo el procesamiento 3' de snRNA en células en cultivo. Células HEK 293T fueron co-transfectadas con dCas9-KRAB y una guía dirigida hacia el promotor de INTS11, o como control una guía contra una región que no se encuentra presente en el genoma humano, y con vectores de expresión para GFP-FLAG-INTS11 WT, mutantes, o como control, GFP solo. A) Luego de 72 h, los niveles de INTS11 fueron analizados mediante WB utilizando los anticuerpos indicados en cada panel. B) Los niveles de snRNAs poliadenilados fueron evaluados mediante RT (oligo dT)-qPCR. (n= 4, letras diferentes indican diferencias significativas, $P < 0,05$; ANOVA de un factor seguido de una prueba de Tukey).

Con la misma estrategia mencionada anteriormente procedimos a estudiar la localización sub-celular de la mutante EA+2KR (Figura R31 A). A diferencia de la versión WT, la nueva mutante no se encuentra principalmente en el núcleo sino que presenta una localización más citoplasmática (Figura R31 A y B), mostrando una distribución

similar a la al mutante 3KR (Figura R31 B). Podemos concluir que la SUMOilación de INTS11 es necesaria para su correcta localización sub-celular.

A



B

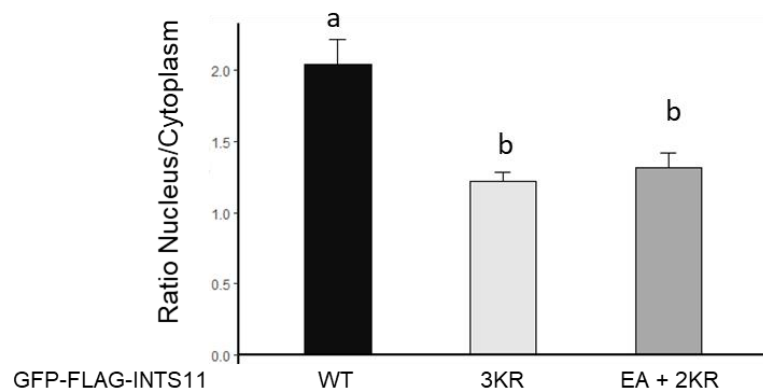


Figura R31: Localización sub-celular de las variantes de INTS11. A) Células HeLa fueron transfectadas con las variantes GFP-FLAG-INTS11 WT, 3KR o EA+2KR, y 24h más tarde fueron fijadas y analizadas por microscopía de fluorescencia confocal. Se muestran las imágenes obtenidas mediante esta técnica, donde se observan los núcleos teñidos con Ioduro de Propidio (azul) (PI) y marca fluorescente verde indicando la localización de cada variante de INTS11 fusionada a GFP. B) Las imágenes tomadas por microscopía confocal fueron analizadas con *Cell Profiler*. Se desarrolló una línea de comando para identificar de manera automática los

núcleos de todas las células y se cuantificó las intensidades medias de los núcleos y citoplasmas. (n=150, letras diferentes indican diferencias significativas, ***P<0,001; ANOVA de un factor seguida de una prueba de Tukey).

En resumen, en base a los resultados presentados hasta aquí, proponemos que INTS11 3KR puede ensamblarse al módulo catalítico, pero no al complejo *Integrator*, y que, a pesar de estar parcialmente en el núcleo, no es reclutado a la cromatina correctamente, generando alteraciones en el reclutamiento de INTS9.

DISCUSION

Resumen de los resultados obtenidos

El trabajo realizado en la presente tesis doctoral demostró que modificar los niveles globales de conjugación SUMO, o bien el balance entre conjugación y deconjugación, altera diferentes etapas de la biogénesis de los snRNAs, consecuentemente alterando la proporción entre snRNAs maduros e inmaduros (o sin procesar). Diversas proteínas, tales como factores de transcripción, factores de elongación de la RNAPII y factores de procesamiento 3' de los snRNAs fueron reportadas como blanco de SUMOilación mediante espectrometría de masa (Hendriks, 2018); a lo largo de este trabajo logramos validar algunas de ellas en nuestro modelo experimental. Enfocamos nuestro estudio en dos complejos proteicos involucrados en la biogénesis de snRNAs, el factor de transcripción o complejo "SNAPc" (por *snRNA activating protein complex*) y el complejo responsable del procesamiento 3' denominado *Integrator complex*. Mediante mutagénesis dirigida, obtuvimos una mutante de la proteína SNAPC1, componente del SNAPc, que presenta una fuerte disminución en sus niveles de SUMOilación. A pesar de ser reclutada a los promotores de los genes de snRNAs, esta mutante es incapaz de activar la transcripción de los mismos e incluso en algunas líneas celulares, y por motivos que aún desconocemos, induce la muerte celular. Por otra parte, generamos una mutante de la proteína INTS11, la subunidad catalítica del *Integrator complex*, la cual es incapaz de conjugarse a SUMO y es a su vez incapaz de rescatar la depleción de la proteína INTS11 endógena en la función relacionada con el procesamiento 3' de los snRNAs así como también en otras actividades regulatorias de la expresión genética recientemente descritas para este complejo proteico. Análisis de co-inmunoprecipitación nos permitieron concluir que la mutante deficiente de SUMOilación de INTS11 no presenta alteraciones en su interacción con las subunidades INTS9 e INTS4, con quienes forma el *core* o módulo catalítico del complejo, pero presenta menor interacción con otras subunidades del complejo con respecto a la versión WT. Además, esta mutante es más citoplasmática que la proteína salvaje, indicando la importancia de la conjugación de SUMO en la localización sub-celular de INTS11. Todos estos resultados demuestran que la conjugación SUMO tiene un rol fundamental en la biogénesis de los snRNAs.

Los niveles de SUMOilacion son importantes para la biogénesis de los snRNAs

Observamos que los niveles de snRNAs nacientes aumentan cuando las células en cultivo presentan menores niveles de SUMOilación global, debido a la sobre-expresión de la isopeptidasa de SUMO SENP1 (Figura R1-R2). En función de estos resultados, decidimos estudiar el efecto de sobre-expresar otras isopeptidasas de SUMO, diferentes a SENP1, enfocándonos en USPL1 dado que ésta se localiza principalmente en los CBs y es necesaria para el ensamblado de estos gránulos o compartimientos sub-nucleares sin membrana (Schulz et al. 2012), fundamentales para los primeros pasos de maduración de los snRNAs así como también para la maduración final de las partículas snRNPs (Matera & Wang 2014). Más aún, en la cercanía de estos gránulos ocurre la transcripción de los genes de snRNAs (Machyna et al. 2014). En concordancia con estos datos, también fue reportado que la depleción de USPL1 disminuye los niveles de snRNAs totales y nacientes. Además, esta enzima interacciona con el complejo de elongación LEC, requerido para la transcripción de snRNAs (Hutten et al. 2014). Vale la pena mencionar que, hasta el momento, no han sido reportados blancos específicos de de-conjugacion de SUMO de la isopeptidasa USPL1 y nada se sabe aún acerca del mecanismo molecular responsable de todos estos resultados reportados años atrás por diferentes laboratorios. En nuestro estudio, la sobre-expresión de USPL1 llevó a una disminución parcial de los niveles de SUMOilación, en contraposición a la drástica disminución producida por la sobre-expresión de SENP1. Sin embargo, los cambios producidos por ambas sobre-expresiones sobre la proporción de snRNA nacientes respecto a los niveles totales de estos transcritos fueron similares (Figura R3). Los resultados obtenidos muestran que los niveles de snRNAs nacientes son modificados por los niveles de SUMOilación celular ya sea que esta sea afectada más globalmente o de forma parcial, lo cual indicaría que diversas proteínas que participan en la biogénesis de los snRNAs son modificadas por conjugación de SUMO y que un desequilibrio entre la conjugación y la de-conjugación genera alteraciones en su actividad. Hay que tener en cuenta que los niveles de snRNA nacientes los obtenemos evaluando los transcritos sin procesar (sin clivar), por lo que resulta difícil técnicamente diferenciar si el efecto observado se debe a un aumento de la tasa de transcripción que genera una sobrecarga de la maquinaria de procesamiento 3' y consecuentemente una disminución en el

procesamiento 3' en relación a la cantidad de transcritos producidos, o bien a una genuina disminución o impedimento en el procesamiento 3' por un efecto directo sobre el funcionamiento de la maquinaria responsable del mismo. Por lo tanto, no descartamos ninguna de estas posibilidades.

Las proteínas involucradas en la biogénesis de los snRNAs son sustrato de conjugación de SUMO

En los últimos años ha habido un mejoramiento en algunas técnicas de detección masiva de PTM, en particular distintos trabajos donde fueron analizados a gran escala los sustratos de SUMOilación han demostrado que diversas proteínas involucradas en la biogénesis de los snRNAs son capaces de ser modificadas por conjugación de SUMO (Hendriks et al. 2018). En este trabajo de tesis, logramos validar que SNAPC1 del complejo SNAPc; ZC3H8 del complejo LEC; INTS4, INTS9, INTS10, INTS11, INTS13 e INTS4 del complejo *Integrator* son efectivamente blanco de conjugación de SUMO (Figura R4). Si bien estos blancos de SUMOilación habían sido ya detectados por análisis proteómicos, se desconocía por completo la consecuencia de la conjugación de SUMO sobre la función de dichas proteínas, así como también los mecanismos regulatorios de esta MPT para los sustratos mencionados.

¿Es posible obtener mutantes incapaces de ser SUMOiladas?

La SUMOilación es una modificación altamente reversible, y altamente inestable lo cual dificulta mucho su detección (Da Silva-Ferrada et al. 2012). En un principio, la mayoría de los estudios clásicos sobre esta modificación fueron hechos con los sustratos con más altos niveles de SUMOilación de la célula. Generalmente, el abordaje experimental para determinar la relevancia de la conjugación a SUMO consiste en la mutación de los residuos lisina a arginina. De todos los factores relacionados con la biogénesis de los snRNAs que hemos podido validar como sustratos de SUMOilación en nuestro laboratorio, nos enfocamos en tres proteínas: SNAPC1 del complejo SNAPc, INTS9 e INTS11 del complejo *Integrator*. En todos los casos utilizamos mutagénesis dirigida para obtener versiones mutantes, deficientes de SUMOilación.

En el caso de SNAPC1, observamos una drástica disminución en su conjugación a SUMO cuando mutamos las lisinas 245 y 333, generando así una variante deficiente

de SUMOilación “SNAPC1 2KR”. Dicho hallazgo no fue tarea fácil, dado que varias de las lisinas analizadas, aún encontrándose inmersas en un motivo consenso de SUMOilación estricto y habiendo sido detectado en análisis proteómicos SUMO-específicos, al ser mutadas no afectaron los niveles de SUMOilación de la proteína SNAPC1 (Figura R6).

Cuando empezamos a trabajar con proteínas pertenecientes al complejo *Integrator*, decidimos estudiar INTS9 e INTS11, ya que forman parte del módulo catalítico de dicho complejo. Parte de nuestro interés en estas dos subunidades, radicó en la ya conocida identidad de secuencia entre la proteína INTS11 y la proteína CPSF73, la cual forma parte del módulo catalítico del complejo CPSF, responsable del corte y poliadenilación de los pre-mRNAs, un proceso que ya conocíamos que es modulado por SUMOilación. Más aún, CPSF73 no solo es la responsable del clivaje endonucleolítico de los pre-mRNAs sino que también ha sido descrito como sustrato de conjugación de SUMO (Vethantham et al. 2007). En el caso de INTS9, por mutagénesis dirigida reemplazamos la lisina 607, la cual fue detectada por proteómica sitio-específica de sustratos de SUMOilación y se localiza dentro de un consenso canónico de conjugación de SUMO (Figura R14). A pesar de obtener una variante con una marcada disminución en los niveles de conjugación a SUMO, no observamos diferencias entre dicha variante y la versión salvaje en los ensayos funcionales de procesamiento del 3' de snRNAs (Figura R15). Este resultado nos llevó a buscar una variante incapaz de ser modificada por conjugación de SUMO para la subunidad INTS11. Para ello, fue necesario reemplazar las lisinas 381, 462 y 475, por argininas. Las lisinas 381 y 462 se encuentran enmarcadas dentro de un motivo consenso de SUMOilación invertido y un motivo consenso canónico, respectivamente, mientras que la lisina 475 no forma parte de ninguna secuencia consenso de SUMOilación descrita hasta la fecha (Figura R16-R17). A diferencia de lo observado con INTS9, la variante deficiente de SUMOilación de INTS11, INTS11 3KR, fue incapaz de llevar a cabo el procesamiento 3' de los snRNAs nacientes (Figura R19).

Hay que tener en cuenta que el uso de una proteína mutante con sustitución de aminoácidos para investigar la importancia de la SUMOilación sobre su funcionalidad tiene sus ventajas y desventajas. Su ventaja es permitir la realización de experimentos donde se puede comparar fácilmente con su contraparte salvaje. Por

otro lado, su debilidad es basarse justamente en el uso de una proteína mutante a la cual se le podrían atribuir efectos debido a la mutación en sí misma, pero independientes de la regulación de la función proteica por SUMOilación. Por eso mismo, el cambio de aminoácido realizado en las mutantes es de aminoácido lisina a arginina. Estos dos aminoácidos tienen tamaño y estructura similar, y la misma carga neta con lo cual no se espera que afecten severamente las estructuras secundaria y terciaria de la proteína.

Vale la pena destacar que la existencia de múltiples lisinas de conjugación a SUMO en un sustrato se ha reportado reiteradamente en la literatura, y pareciera ser aún más común que la existencia de un sitio único de SUMOilación. Un ejemplo de esto es la proteína Scc1, en la cual fue necesario mutar 15 lisinas de su dominio central de unión a la proteína SA2, para lograr una disminución significativa en sus niveles de SUMOilación (Wu et al., 2012). El hallazgo de los sitios blanco de conjugación de SUMO de distintas proteínas no ha sido tarea simple durante esta tesis doctoral, dado que muchas veces las lisinas no se presentaban dentro de secuencias consenso para la conjugación de SUMO y otras veces, fue necesario la combinación de varias mutaciones puntuales para lograr una disminución en la modificación post-traducciona estudiada. Más allá de ello, disminuir drásticamente la conjugación de SUMO no siempre fue suficiente para alterar la actividad de alguna de las proteínas de interés.

Consecuencias de la SUMOilación

Complejo SNAPc

SNAPC1 es una de las cinco proteínas del SNAPc, el complejo necesario para iniciar la transcripción de todos los snRNAs. En la línea celular HEK 293T, la mutante de SNAP43 deficiente de SUMOilación (SNAP43 2KR) es reclutada a los promotores de U1 y U2 (Figura R10), pero es incapaz de reestablecer los niveles basales de snRNA nacientes en células depletadas de la proteína SNAP43 endógena (Figura R8), cosa que sí logra la versión salvaje de dicha proteína. A su vez, SNAP43 2KR altera el reclutamiento de la RNAPII a los promotores de los snRNAs aún en presencia de la proteína endógena (Figura R9), lo cual sugiere un comportamiento de tipo dominante negativo. Sorprendentemente, observamos que algunas líneas celulares, como por

ejemplo HeLa, eran más sensibles a la presencia de SNAP43 2KR, induciendo su muerte celular (Figura R11-R12). La función de SNAP43 dentro del complejo es la de favorecer la interacción de SNAP50 y SNAP190, permitiendo la formación un complejo mínimo de SNAPc (Mittal et al 1999; James Faresse et al. 2012). Además, SNAP43 es capaz de interaccionar con las proteínas Rb y p53 lo cual serviría para regular la expresión de los snRNAs durante el ciclo celular (Hirsch et al. 2000; Gridasova et al. 2005). Los resultados observados en esta tesis podrían deberse a que la SUMOilación de SNAP43 es necesaria para la formación del complejo mínimo y al no formarse dicho complejo, se disrumpan todas las interacciones necesarias para el inicio de la transcripción. Los experimentos se realizaron en cultivos celulares no sincrónicos, dejando abierta la posibilidad de que exista alguna desregulación de la transcripción de los snRNAs durante el ciclo celular en el contexto de sobre-expresión de la mutante deficiente de SUMOilación de SNAP43

Vale la pena mencionar que los genes que codifican para snRNAs poseen múltiples copias. Si bien muchas de estas copias habían sido descritas como “pseudo genes”, recientemente se observó que algunos se expresan en forma dependiente de tipos celular y regulan el *splicing* de ciertos genes codificantes (Kyriakopoulou et al. 2006; Marz et al. 2008; O'Reilly et al. 2013; Guiro & O'Reilly 2015). Esto nos permite hipotetizar que SNAPC1 2KR podría ser capaz de estimular la transcripción de algunas variantes no canónicas de los snRNAs que al formar parte de snRNP maduros, alterarían la composición del spliceosoma y consecuentemente modificarían patrones de *splicing* de diferentes ARN mensajeros. La validación o refutación de esta hipótesis requiere tarea experimental que va más allá del contexto de la presente tesis.

A lo largo de este trabajo no pudimos terminar de comprender de qué manera se desencadena la muerte celular frente a la sobre-expresión de la versión deficiente de SUMOilación de SNAP43, y en qué medida este evento depende del tipo celular. De acuerdo con nuestros experimentos en realizados en dos líneas celulares (HEK 293T y HeLa), podemos especular que las células HeLa son más sensibles a una disminución de la transcripción de los snRNAs dependientes de RNAPII. Podría ocurrir, sin ser excluyente con lo mencionado anteriormente, que exista una tolerancia diferencial, dependiente de la línea celular estudiada, a la composición alterada del spliceosoma

que se produce como consecuencia de la expresión de SNAP43 2KR, tal como se explica en el párrafo anterior.

Complejo *Integrator*.

El complejo *Integrator* está compuesto por 14 subunidades (INTS1-INTS14). En particular las subunidades INTS4, INTS9 e INTS11 forman el módulo catalítico (Albrecht et al., 2018; Pfeleiderer & Galej 2021). Existe un concepto emergente en el campo de SUMO que se conoce como “SUMOilación en grupo”, lo cual se refiere al requerimiento de la modificación simultánea de múltiples proteínas blanco involucradas en un mismo proceso celular (Flotho & Melchior 2013). Este podría ser perfectamente el caso del complejo *Integrator*, donde observamos varias subunidades susceptibles de ser modificadas por conjugación de SUMO (Figura R4). En este trabajo, nos enfocamos en estudiar las consecuencias de la conjugación de SUMO en las subunidades INTS9 e INTS11, las cuales fueron las primeras descritas como necesarias para la actividad catalítica del complejo y son aquellas que tienen homología con subunidades de CPSF, el complejo proteico responsable del corte y poliadenilación 3' de pre-mRNAs (Baillat et al. 2005; Albrecht et al. 2012).

En el caso de INTS9, obtuvimos una mutante cuyos niveles de SUMOilación se encuentran drásticamente disminuidos a pesar de lo cual el procesamiento de los snRNAs no se ve afectado (Figura R14-R15). Sería necesario generar nuevas mutantes, las cuales disminuyan aún más su nivel de conjugación a SUMO para conocer, si es que existe, el efecto que ejerce la ausencia de esta MPT sobre la función de esta proteína.

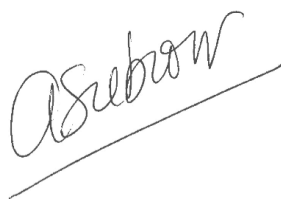
En el caso de INTS11, la mutante INTS11 3KR es prácticamente incapaz de ser modificada por conjugación de SUMO y su actividad catalítica se vió severamente alterada, a pesar de tener su sitio activo intacto (Figura R16-R17). Comprobamos que es capaz de ensamblarse en el módulo catalítico, es decir, de interactuar con INTS9 e INTS4 (Figura R21-R23). Sin embargo, esta variante de INTS11 presenta una interacción disminuída con otras proteínas componentes del complejo (Figura R24). El hecho de mantener su interacción con INTS4 e INTS9 y sugiere que las mutaciones generadas no alteraron en gran medida su estructura. Resultados de otros laboratorios han postulado que el complejo *Integrator* está conformado por distintos módulos, los cuales se van ensamblando a medida que ocurre la transcripción mediada por la

RNAPII (Figura I3) (Egloff et al. 2010; Egloff et al. 2012). Observamos que INTS11 3KR no es reclutado a la cromatina y que el reclutamiento a la cromatina de INTS9 es dependiente de INTS11 (Figura R25-R26). Por otro lado, según nuestras observaciones, la SUMOilación de la proteína INTS11 en la lisina 462, que se encuentra dentro de un NLS (Mackmull et al. 2017), resulta necesaria para su correcta localización sub-celular (Figura R27 y R31). Vale la pena mencionar que un trabajo reciente publicado por Pfleiderer & Galej demuestra que INTS4/INTS9/INTS11 forman un túnel cargado positivamente, el cual es necesario para el clivaje del snRNA naciente. Llegaron a dicha conclusión mutando distintos residuos cargados positivamente de las tres proteínas que forman el módulo catalítico, entre ellos reemplazaron la lisina 462 de INTS11 por un ácido glutámico que posee carga negativa. Al realizar dicha mutación observaron que era incapaz de realizar el procesamiento 3' de los snRNAs nacentes (Pfleiderer & Galej 2021). Sin embargo, nosotros reemplazamos la lisina 462 por una arginina, manteniendo la carga positiva en esa posición. Además, al mutar el ácido glutámico en la posición 464, que es parte del consenso de la conjugación a SUMO de la lisina 462, por una alanina observamos que su distribución núcleo/citoplasma es igual a la mutante 3KR, lo cual indica la importancia de la conjugación a SUMO para la localización de INTS11 (Figura R31). Por lo tanto, lo reportado por estos investigadores podría re-interpretarse a la luz de nuestros resultados, como una alteración en la SUMOilación y la localización sub-celular de INTS11, que consecuentemente afecte la actividad del complejo a nivel del procesamiento 3' de los snRNAs.

A partir de los resultados presentados en la figura R22, R23, R26, R27 y R31, es posible especular que el módulo catalítico podría encontrarse pre-ensamblado en el citoplasma y translocar como tal al núcleo. Dado que la variante de INTS11 deficiente de SUMOilación es más citoplasmática que la versión salvaje de esta proteína, pero aun así presenta cierto grado de localización nuclear, la disminución de la interacción de INTS11 3KR con algunas proteínas del complejo sugiere que aún la fracción nuclear de esta proteína no es capaz de ensamblarse correctamente en el complejo, mostrando la relevancia de su conjugación a SUMO. Experimentación adicional será necesaria para un análisis más detallado sobre la relevancia de la SUMOilación en la dinámica de distribución sub-celular de INTS11, así como también la dinámica de distribución de las otras subunidades con las que INTS11 interacciona y de ensamblado

del complejo *Integrator*. Claramente, la conjugación de SUMO, o incluso el balance entre conjugación y de-conjugación está jugando un papel importante en el tránsito núcleo-citoplasma de la proteína INTS11.

En los últimos años han ido emergiendo nuevas funciones para el complejo *Integrator* más allá de su ya conocido rol en el procesamiento de los snRNAs, y se ha podido establecer una conexión entre la actividad catalítica del complejo con la expresión y el procesamiento de diversos tipos de ARNs, tales como ARN transcriptos a partir de enhancers o “eRNAs” (Lai et al. 2015), diversos ARN no codificantes largos y no poliadenilados (Nojima et al. 2018), ARN virales (Cazalla et al., 2011), entre otros. Además, ha sido reportado que el complejo *Integrator* está involucrado en la regulación fina del pausado de la RNAPII en regiones proximales al promotor de diferentes genes codificantes de proteínas (Gardini et al. 2014; Skaar et al. 2015; Elrod et al. 2019; Tatomer et al. 2019; Beckedorff et al. 2020; Lykke-Andersen et al., 2021). Sin embargo, no existen hasta el momento reportes acerca de mecanismos de regulación de la actividad del complejo. Nuestro trabajo demuestra que la modificación por conjugación de SUMO es necesaria para la correcta actividad de INTS11, y consecuentemente para la actividad catalítica del complejo no sólo a nivel del procesamiento de los snRNAs sino también para su función regulatoria de la expresión de otros transcriptos como eRNAs y PROMPTS.



Anabella Srebrow



Laureano Bragado

BIBLIOGRAFIA

Akhtar MS, Heidemann M, Tietjen JR, Zhang DW, Chapman RD, Eick D, Ansari AZ. 2009 TFIIF kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol. Cell* 34.

Albrecht TR, Wagner EJ. 2012. snRNA 3' end formation requires heterodimeric association of integrator subunits. *Mol Cell Biol* 32:1112–1123.

Albrecht TR, Shevtsov SP, Wu Y, Mascibroda LG, Peart NJ, Huang KL, Sawyer IA, Tong L, Dundr M, Wagner EJ. 2018. Integrator subunit 4 is a 'Symplekin-like' scaffold that associates with INTS9/11 to form the Integrator cleavage module. *Nucleic acids research*. 46(8), 4241–4255.

Andersen JB, Hassel BA. 2006. The interferon regulated ubiquitin-like protein, ISG15, in tumorigenesis: Friend or foe? *Cytokine and Growth Factor Reviews*, 17:411–421.

Annemmarthe G. van der Veen¹ and Hidde L. Ploegh. 2013. Sumoylation: A Regulatory Protein Modification in Health and Disease. *Annu Rev Biochem*.82:357-85.

Baillat D, Gardini A, Cesaroni M, Shiekhattar R. 2012. Requirement for SNAPC1 in Transcriptional Responsiveness to Diverse Extracellular Signals. *Mol Cell Biol*; 32: 4642-50.

Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, 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.

Baillat D, Wagner EJ. 2015. Integrator: Surprisingly diverse functions in gene expression. *Trends Biochem Sci* 40:257–264.

Bartkowiak B, Liu P, Phatnani HP, Fuda NJ, Cooper JJ, Price DH, Adelman K, Lis JT, Greenleaf AL. 2010 CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev*. 24, 2303–2316.

- Bayer P, Arndt A, Metzger S, Mahajan R, Melchior F, Jaenicke R, Becker J. 1998. Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* 280, 2: 275–286.
- Beckedorff F, Blumenthal E, daSilva LF, Aoi Y, Cingaram PR, Yue J, Zhang A, Dokaneheifard S, Valencia MG, Gaidosh G, Shilatifard A, Shiekhattar R. 2020. The Human Integrator Complex Facilitates Transcriptional Elongation by Endonucleolytic Cleavage of Nascent Transcripts. *Cell Rep.* Jul 21;32(3):107917.
- Bellare P, Kutach AK, Rines AK, Guthrie C, Sontheimer EJ 2006. Ubiquitin binding by a variant Jab1/MPN domain in the essential premRNA splicing factor Prp8p. *RNA.* 12; 2: 292-302.
- Bellare P, Small EC, Huang X, Wohlschlegel JA, Staley JP, Sontheimer EJ. 2008. A role for ubiquitin in the spliceosome assembly pathway. *Nat Struct Mol Biol.* 15; 5: 444-51.
- Blank-Giwojna A, Postepska-Igielska A, Grummt I. 2019. lncRNA KHPS1 Activates a Poised Enhancer by Triplex-Dependent Recruitment of Epigenomic Regulators. *Cell reports.* 26(11), 2904–2915.e4.
- Blomster HA, Hietakangas V, Wu J, Kouvonen P, Hautaniemi S, Sistonen L. 2009. Novel proteomics strategy brings insight into the prevalence of SUMO-2 target sites. *Mol Cell Proteomics.* 8; 6: 1382-90.
- Burglin TR, Affolter M. 2016 Homeodomain proteins: an update. *Chromosoma* 125, 497–521.
- Canella D, Praz V, Reina JH, Cousin P, Hernandez N. 2010. Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res* 20: 710–721.
- Cazalla D, Xie M, Steitz JA. 2011. A primate herpesvirus uses the integrator complex to generate viral microRNAs. *Mol Cell.* 2011;43(6):982-992.
- Chen L, Lullo DJ, Ma E, Celniker SE, Rio DC, Doudna JA. 2005 Identification and analysis of U5 snRNA variants in *Drosophila*. *RNA.* 11, 1473–1477.

Chen, FX, Smith, ER & Shilatifard, A. 2018. Born to run: control of transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* 19, 464–478.

Cuello P, Boyd DC, Dye MJ, Proudfoot NJ, Murphy S. 1999 Transcription of the human U2 snRNA genes continues beyond the 30 box in vivo. *EMBO J.* 18

Da Silva-Ferrada E, Lopitz-Otsoa F, Lang V, Rodríguez MS, Matthiesen R. 2012. Strategies to identify recognition signals and targets of SUMOylation. *Biochem Res Int.* 875148.

Drag M, Salvesen, GS 2008. DeSUMOylating enzymes--SENPs. *IUBMB Life.* 60; 11: 734–742.

Dvinge H, Guenthoer J, Porter PL, Bradley RK. 2019. RNA components of the spliceosome regulate tissue- and cancer-specific alternative splicing. *Genome Res.* 29(10):1591-1604

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. 2007. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318.

Egloff S, Al-Rawaf H, O'Reilly D, Murphy S. 2009. Chromatin structure is implicated in 'late' elongation checkpoints on the U2 snRNA and beta-actin genes. *Mol. Cell. Biol.* 29, 4002–4013.

Egloff S, Szczepaniak SA, Dienstbier M, Taylor A, Knight S, Murphy S. 2010. The Integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J. Biol. Chem.* 285, 20 564–20 569.

Egloff S, Zaborowska J, Laitem C, Kiss T, Murphy S. 2012. Ser7 phosphorylation of the CTD recruits the RPAP2 Ser5 phosphatase to snRNA genes. *Mol. Cell* 45, 111–122.

Egloff S, Vitali P, Tellier M, Raffel R, Murphy S, Kiss T. 2017. The 7SK snRNP associates with the little elongation complex to promote snRNA gene expression. *EMBO J.* 36, 934–948.

Elrod ND, Henriques T, Huang KL, Tatomer DC, Wilusz JE, Wagner EJ, Adelman K. 2019. The Integrator Complex Attenuates Promoter-Proximal Transcription at Protein-Coding Genes. *Mol Cell*. 76:738-752.e7.

Fischer U, Sumpster V, Sekine M, Satoh T, Luhrmann R. 1993. Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *EMBO J*. 12, 573–583.

Fischer U, Luhrmann R. 1990. An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus. *Science* 249, 786–790.

Flotho A, Melchior F. 2013. Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem*. 2013;82:357-85.

Fuke H, Ohno M. 2008. Role of poly (A) tail as an identity element for mRNA nuclear export. *Nucleic Acids Res*. 36, 1037–1049.

Gardini A, Baillat D, Cesaroni M, Hu D, Marinis JM, Wagner EJ, Lazar MA, Shilatifard A, Shiekhhattar R. 2014. Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell* 56, 128–139.

Geissfriedlander R, Melchior F. 2007. Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol*. 12: 947-56.

Geng, J, Klionsky, DJ. 2008. The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. “Protein modifications: beyond the usual suspects” review series. *EMBO reports*, 9: 859–64.

Gill. 2004. SUMO and Ubiquitin in the Nucleus: Different Functions, Similar Mechanisms?. *Genes Dev*.1;17: 2046-59.

Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. 2014. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell*, 159, 647–661.

Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, Bentley DL. 2009. TFIIF-associated Cdk7 kinase functions in phosphorylation of C-terminal domain

Ser7 residues, promoter- proximal pausing, and termination by RNA polymerase II. *Mol. Cell. Biol.* 29, 5455–5464.

Gridasova AA, Henry RW. 2005. The p53 tumor suppressor protein represses human snRNA gene transcription by RNA polymerases II and III independently of sequence-specific DNA binding. *Mol Cell Biol.* Apr;25(8):3247-60.

Groettrup M, Pelzer C, Schmidtke G, Hofmann K. 2008. Activating the ubiquitin family: UBA6 challenges the field. *Trends in Biochemical Sciences*, 33: 230–237

Guiro J, Murphy S. 2017. Regulation of expression of human RNA polymerase II-transcribed snRNA genes. *Open Biol.* 7: 170073.

Guiro J, O'Reilly D. 2015. Insights into the U1 small nuclear ribonucleoprotein complex superfamily. *Wiley interdisciplinary reviews RNA.* 6:79-92.

Guo D, Li M, Zhang Y, Yang P, Eckenrode S, Hopkins D, Zheng W, Purohit S, Podolsky RH, Muir A, Wang J, Dong Z, Brusko T, Atkinson M, Pozzilli P, Zeidler A, Raffel LJ, Jacob CO, Park Y, Serrano-Rios M, Larrad MT, Zhang Z, Garchon HJ, Bach JF, Rotter JI, She JX, Wang CY. 2004. A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. *Nat Genet.* 36:837-41.

Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol.* Jun 5;166(4):557-80.

Hay RT. 2005. SUMO: a history of modification. *Mol Cell.* 18: 1-12.

Heideker J, Perry JJP, Boddy MN. 2009. Genome stability roles of SUMO-targeted ubiquitin ligases. *DNA Repair.* 8(4) 517–524.

Hirsch HA, Gu L, Henry RW. 2000. The retinoblastoma tumor suppressor protein targets distinct general transcription factors to regulate RNA polymerase III gene expression, *Mol. Cell Biol.* 20: 9182–9191

Hsieh CL, Fei T, Chen Y, Li T, Gao Y, Wang X, Sun T, Sweeney CJ, Lee GS, Chen S, Balk SP, Liu XS, Brown M, Kantoff PW. 2014. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(20), 7319–7324.

Hendriks IA, Lyon D, Su D, Skotte NH, Daniel JA, Jensen LJ, Nielsen ML. 2018. Site-specific characterization of endogenous SUMOylation across species and organs. *Nat Commun.* 9; 1: 2456.

Hernandez N. 2001. Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *J. Biol. Chem.* 276, 26733–26736.)

Herzel L, Ottoz D, Alpert T. 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat Rev Mol Cell Biol* **18**, 637–650.

Hickey CM, Wilson NR, Hochstrasser M. 2012. Function and regulation of SUMO proteases. *Nat Rev Mol Cell Biol.* 13; 12:755-66

Hochstrasser M. 2000. Biochemistry. All in the ubiquitin family. *Science.* 289: 563–564.

Hu D, Smith ER, Garruss AS, Mohaghegh N, Varberg JM, Lin C, Jackson J, Gao X, Saraf A, Florens L, Washburn MP, Eisenberg JC, Shilatifard A. 2013. The little elongation complex functions at initiation and elongation phases of snRNA gene transcription. *Mol. Cell* 51, 493–505.

Hurley JH, Lee S, Prag G. 2006. Ubiquitin-binding domains. *The Biochemical journal.* 399; 3: 361–72.

Hutten S, Chachami G, Winter U, Melchior F, Lamond AI. 2014 A role for the Cajal-body-associated SUMO isopeptidase USPL1 in snRNA transcription mediated by RNA polymerase II. *J. Cell Sci.* 127, 1065–1078.

James Faresse N, Canella D, Praz V, Michaud J, Romascano D, Hernandez N. 2012. Genomic Study of RNA Polymerase II and III SNAPc-Bound Promoters Reveals a Gene Transcribed by Both Enzymes and a Broad Use of Common Activator. *PLoS Genet.* 2012;8:e1003028

Jawdekar GW, Henry RW. 2008. Transcriptional regulation of human small nuclear RNA genes. *Biochimica et biophysica acta.* 1779, 295-305.

Jentsch S, Psakhye I. 2013. Control of nuclear activities by substrate-selective and protein-group SUMOylation. *Annual review of genetics,* 47: 167–86.

Jodoin JN, Sitaram P, Albrecht TR, May SB, Shboul M, Lee E, Reversade B, Wagner EJ, Lee LA. 2013. Nuclear-localized Asunder regulates cytoplasmic dynein localization via its role in the integrator complex. *Molecular biology of the cell*. 24(18), 2954–2965.

Jonkers I, Lis JT. 2015 Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 16, 167–177.

Kamieniarz-Gdula K, Proudfoot NJ. 2019. Transcriptional control by premature termination: a forgotten mechanism. *Trends Genet* 35: 553–564.

Kaur K, Park H, Pandey N, Azuma Y, De Guzman RN. 2017. Identification of a new small ubiquitin-like modifier (SUMO)-interacting motif in the E3 ligase PIASy. *J Biol Chem.* 292; 24: 10230-10238

Kim MK, Kang YS, Lai HT, Barakat NH, Magante D, Stumph WE. 2010. Identification of SNAPc subunit domains that interact with specific nucleotide positions in the U1 and U6 gene promoters. *Mol. Cell. Biol.* 30:2411–2423.

Knipscheer P, Flotho A, Klug H, Olsen JV, van Dijk WJ, Fish A, Johnson ES, Mann M, Sixma TK, Pichler A. 2008. Ubc9 sumoylation regulates SUMO target discrimination. *Mol Cell.* 31; 3: 371-82.

Koidl S, Eisenhardt N, Fatouros C, Droscher M, Chaugule VK, Pichler A. 2016. The SUMO2/3 specific E3 ligase ZNF451-1 regulates PML stability. *The international journal of biochemistry & cell biology.* 79, 478–487.

Kuhlman TC, Cho H, Reinberg D, Hernandez N. 1999. The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter. *Mol. Cell. Biol.* 19:2130–2141.

Kyriakopoulou C, Larsson P, Liu L, Schuster J, Söderbom F, Kirsebom LA, Virtanen A. 2006. U1 like snRNAs lacking complementarity to canonical 5' splice sites. *RNA.* 12:1603-11.

Lai F, Gardini A, Zhang A, Shiekhattar R. 2015. Integrator mediates the biogenesis of enhancer RNAs. *Nature* 525, 399–403.

- Laitem C, Zaborowska J, Tellier M, Yamaguchi Y, Cao Q, Egloff S, Handa H, Murphy S. 2015. CTCF regulates NELF, DSIF and P-TEFb recruitment during transcription. *Transcription* 6, 79–90.
- Lee KM, O’Connell MJ. 2006. A new SUMO ligase in the DNA damage response. *DNA Repair*. 5(1). 138–141.
- Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, et al. 2004. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal* 23,4051-60.
- Li Y, Bolderson E, Kumar R, Muniandy PA, Xue Y, Richard DJ, Seidman M, Pandita TK, Khanna KK, Wang W. 2009. HSSB1 and hSSB2 form similar multiprotein complexes that participate in DNA damage response. *J. Biol. Chem.* 284, 23 525– 23 531.
- Liang K, Gao X, Gilmore JM, Florens L, Washburn MP, Smith E, Shilatifard A. 2015 Characterization of human cyclin-dependent kinase 12 (CDK12) and CDK13 complexes in C-terminal domain phosphorylation, gene transcription, and RNA processing. *Mol. Cell. Biol.* 35, 928–938.
- Liang YC, Lee CC, Yao YL, Lai CC, Schmitz ML3, Yang WM. 2016. SUMO5, a Novel Poly-SUMO Isoform, Regulates PML Nuclear Bodies. *Sci Rep.* 23; 6: 26509
- Lykke-Andersen S, Žumer K, Molska EŠ, Rouvière JO, Wu G, Demel C, Schwalb B, Schmid M, Cramer P, Jensen TH. 2021. Integrator is a genome-wide attenuator of non-productive transcription. *Mol Cell.* Feb 4;81(3):514-529.e6.
- Liu JL, Wu Z, Nizami Z, Deryusheva S, Rajendra TK, Beumer, KJ, Gao H, Matera AG, Carroll D, Gall JG. 2009. Coilin is essential for Cajal body organization in *Drosophila melanogaster*. *Mol. Biol. Cell* 20, 1661–1670.
- Machyna M, Heyn P, Neugebauer KM. 2013. Cajal bodies: where form meets function. *Wiley Interdiscip. Rev. RNA* 4, 17–34.
- Mackmull MT, Klaus B, Heinze I, Chokkalingam M, Beyer A, Russell RB, Ori A, Beck M. 2017. Landscape of nuclear transport receptor cargo specificity. *Mol Syst Biol.* Dec 18;13(12):962.

Mahajan R, Delphin C, Guan T, Gerace L, Melchior F. 1997. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell*, 88: 97–107.

Martin S, Wilkinson KA, Nishimune A, Henley JM. 2007. Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction. *Nature reviews. Neuroscience*, 8; 12: 948–959

Marz M, Kirsten T, Stadler PF. 2008. Evolution of spliceosomal snRNA genes in metazoan animals. *Journal of molecular evolution*; 67:594-607.

Massenet S, Pellizzoni L., Paushkin S, Mattaj IW, Dreyfuss G. 2002. The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol. Cell. Biol.* 22, 6533–6541.

Masuyama K, Taniguchi I, Kataoka N, Ohno M. 2004. RNA length defines RNA export pathway. *Genes Dev.* 18, 2074–2085.

Matera A, Wang Z. 2014. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* 15, 108 121.

Matic I, Schimmel J, Hendriks IA, van Santen MA, van de Rijke F, van Dam H, Gnad F, Mann M, Vertegaal AC. 2010. Site-Specific Identification of SUMO-2 Targets in Cells Reveals an Inverted SUMOylation Motif and a Hydrophobic Cluster SUMOylation Motif. *Mol Cell.* 39; 4: 641-52.

Matunis MJ, Coutavas E, Blobel G. 1996. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J. Cell Biol.*, vol. 135, 6:1457–1470.

McCloskey A, Taniguchi I, Shinmyozu K, Ohno M. 2012. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science* 335, 1643–1646.

Medlin J, Scurry A, Taylor A, Zhang F, Peterlin BM, Murphy S. 2005 P-TEFb is not an essential elongation factor for the intronless human U2 snRNA and histone H2b genes. *EMBO J.* 24, 4154–4165.

- Melchior F. 2000. Sumo --N Onclassical U Biquitin. *Yeast*. 591–626.
- Meluh PB, Koshland D. 1995. Evidence that the MIF2 gene of *Saccharomyces cerevisiae* encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol. Biol. Cell*, 6:793–807.
- Michels AA, Fraldi A, Li Q, Adamson TE, Bonnet F, Nguyen VT, Sedore SC, Price JP, Price DH, Lania L, Bensaude O. 2004. Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J* 23: 2608.
- Mittal V, Ma B, Hernandez N. 1999. SNAP(c): a core promoter factor with a built-in DNA-binding damper that is deactivated by the Oct-1POU domain. *Genes Dev*. 13:1807–1821.
- Miyazono K, Kamiya Y, Miyazawa K. 2008. SUMO amplifies TGF- β signalling. *Nature Cell Biology*. 10(6). 635–637.
- Mohideen F, Capili AD, Bilimoria PM, Yamada T, Bonni A, Lima CD .2009. A molecular basis for phosphorylation-dependent SUMO conjugation by the E2 UBC9. *Nat Struct Mol Biol*. 16; 9:945-52.
- Moss T, Langlois F, Gagnon-Kugler T, Stefanovsky V. 2007. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. *Cell Mol. Life Sci*. 64:29–49
- Mossessova E, Lima CD. 2000. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell*. 5, 5: 865–876
- Mukhopadhyay D, Dasso, M. 2007. Modification in reverse: the SUMO proteases. *Trends in Biochemical Sciences*. 32: 6; 286–295.
- Muñoz MJ, de la Mata M, Kornblihtt AR. 2010. The carboxy terminal domain of RNA polymerase II and alternative splicing. *Trends Biochem Sci*.
- Murphy S. 1997. Differential in vivo activation of the class II and class III snRNA genes by the POU-specific domain of Oct-1. *Nucleic Acids Res*. 25, 2068–2076
- Nesic D, Tanackovic G, Kramer A. 2004. A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *J. Cell Sci*. 117, 4423–4433.

Nojima T, Tellier M, Foxwell J, Ribeiro de Almeida C, Tan-Wong SM, Dhir S, Dujardin G, Dhir A, Murphy S, Proudfoot NJ. 2018. Mammalian lncRNA through Loss of SPT6 Induces R-Loop Formation, Replication Stress, and Cellular Senescence. *Mol Cell*. 72(6):970-984.e7.

O'Geen H, Ren C, Nicolet CM, Perez AA, Halmai J, Le VM, Mackay, JP, Farnham PJ, Segal, DJ. 2017. dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic acids research*. 45(17), 9901–9916.

O'Reilly D, Kuznetsova OV, Laitem C, Zaborowska J, Dienstbier M, Murphy S. 2014 Human snRNA genes use polyadenylation factors to promote efficient transcription termination. *Nucleic Acids Res*. 42, 264–275.

O'Reilly D, Dienstbier M, Cowley SA, Vazquez P, Drozd M, Taylor S, James WS, Murphy S. 2013. Differentially expressed, variant U1 snRNAs regulate gene expression in human cells. *Genome Res*. Feb;23(2):281-91

Ong CT, Corces VG. 2014. CTCF: an architectural protein bridging genome topology and function. *Nature Reviews Genetics* volume. *Nat Rev Genet*. 15:234-46.

Ouyang J, Gill G. 2009. SUMO engages multiple corepressors to regulate chromatin structure and transcription. *Epigenetics*, 4(7),440–444.

Pelisch F, Gerez J, Druker J, Schor IE, Muñoz MJ, Risso G, Petrillo E, Westman BJ, Lamond AI, Arzt E, Srebrow A. 2010. The serine/arginine-rich protein SF2/ASF regulates protein sumoylation. *Proc Natl Acad Sci U S A*. Sep 14;107(37):16119-24.

Perry JJP, Tainer JA, Boddy MN. 2008. A SIM-ultaneous role for SUMO and ubiquitin. *Trends in Biochemical Sciences*, 33;5: 201–208.

Pfleiderer MM, Galej WP. 2021. Structure of the catalytic core of the Integrator complex. *Mol Cell*. Mar 18;81(6):1246-1259.e8.

Pitha-Rowe IF, Pitha PM. 2007. Viral defense, carcinogenesis and ISG15: Novel roles for an old ISG. *Cytokine and Growth Factor Reviews*, 18: 409–417.

Pozzi B, Bragado L, Will CL, Mammi P, Risso G, Urlaub H, Lührmann R, Srebrow A. 2017. SUMO conjugation to spliceosomal proteins is required for efficient pre-mRNA splicing. *Nucleic acids research*. 45(11), 6729–6745.

Rodriguez MS, Dargemont C, Hay RT. 2001. SUMO-1 Conjugation in Vivo Requires Both a Consensus Modification Motif and Nuclear Targeting. *Journal of Biological Chemistry*, 276: 16; 12654–12659.

Rouvière, JO, Geoffroy MC, Palancade B. 2013. Multiple crosstalks between mRNA biogenesis and SUMO. *Chromosoma*. 122(5), 387–399.

Sapetschnig, Rischitor, Braun, Doll, Schergaut, Melchior, Suskeet al., 2002. Transcription factor Sp3 is silenced through SUMO modification by PIAS1. *EMBO Journal*, 21(19), pp.5206–5215.

Savare J, Bonneaud N, Girard F. 2005. SUMO Represses Transcriptional Activity of the *Drosophila* SoxNeuro and Human Sox3 Central Nervous System–specific Transcription Factors. *Molecular biology of the cell*.1–13.

Schaffert N, Hossbach M, Heintzmann R, Achsel T, Luhrmann R. 2004 RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *EMBO J*. 23, 3000–3009.

Schulz S, Chachami G, Kozackiewicz L, Winter U, Stankovic-Valentin N, Haas P, Hofmann K, Urlaub H, Ovaa H, Wittbrodt J, Meulmeester E, Melchior F. 2012. Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions. *EMBO Rep*. 2012. 10: 930-8.

Seu CS, Chen Y. 2009. Identification of SUMO-binding motifs by NMR. *Methods in Molecular Biology*. 497: 121–138.

Shi X, Sun M, Liu H, Yao Y, Song Y. 2013. Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett*. 339, 159-66.

Skaar JR, Richard DJ, Saraf A, Toschi A, Bolderson E, Florens L, Washburn MP, Khanna KK, Pagano M. 2009 INTS3 controls the hSSB1-mediated DNA damage response. *J. Cell Biol*. 187, 25–32.

Smith E, Lin C, Shilatifard A. 2011 The super elongation complex (SEC) and MLL in development and disease. *Genes Dev.* 25, 661–672.

Sontheimer EJ, Steitz JA. 1992. Three novel functional variants of human U5 small nuclear RNA. *Mol. Cell. Biol.* 12, 734–746.

Srivastava R, Ahn SH. 2015. Modifications of RNA polymerase II CTD: Connections to the histone code and cellular function. *Biotechnol Adv* 33:856–872.

Stadelmayer B, Micas G, Gamot A, Martin P, Malirat N, Koval S, Raffel R, Sobhian B, Severac D, Rialle S, Parrinello H, Cuvier O, Benkirane M. 2014. Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat. Commun.* 5, 5531.

Strzelecka, M, Trowitzsch, S, Weber, G, Lührmann, R, Oates, AC, Neugebauer KM. 2010. Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. *Nature structural & molecular biology*, 17(4), 403–409.

Suh HY, Kim JH, Woo JS, Ku B, Shin EJ, Yun Y, Oh BH. 2012. Crystal structure of DeSI-1, a novel deSUMOylase belonging to a putative isopeptidase superfamily. *80;8:2099-104.*

Suzuki, T, Izumi H, Ohno M. 2010. Cajal body surveillance of U snRNA export complex assembly. *J. Cell Biol.* 190, 603–612

Takahashi H, Takigawa I, Watanabe M, Anwar D, Shibata M, Tomomori-Sato C, Sato S, Ranjan A, Seidel CW, Tsukiyama T, Mizushima W, Hayashi M, Ohkawa Y, Conaway JW, Conaway RC, Hatakeyama S. 2015. MED26 regulates the transcription of snRNA genes through the recruitment of little elongation complex. *Nature communications.* 6, 5941.

Takahashi H, Ranjan A, Chen S, Suzuki H, Shibata M, Hirose T, Hirose H, Sasaki K, Abe R., Chen K, He Y, Zhang Y, Takigawa I, Tsukiyama T, Watanabe M, Fujii S, Iida M, Yamamoto J, Yamaguchi Y, Suzuki Y,..., Hatakeyama, S. 2020. The role of Mediator and Little Elongation Complex in transcription termination. *Nature communications*, 11(1) 1063.

Takahashi Y, Mizoi J, Toh-E A, Kikuchi Y. 2000. Yeast Ulp1, an Smt3-specific protease, associates with nucleoporins. *J Biochem.* 128; 5:723-5.

- Tan-Wong SM, Dhir S, Proudfoot NJ. 2019. R-Loops Promote Antisense Transcription across the Mammalian Genome. *Molecular cell*. 76(4), 600–616.e6.
- Tatomer DC, Elrod ND, Liang D, Xiao MS, Jiang JZ, Jonathan M, Huang KL, Wagner EJ, Cherry S, Wilusz JE. 2019. The Integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes Dev*. 33(21-22):1525-1538
- Tatomer DC, Wilusz JE. 2020. Attenuation of Eukaryotic Protein-Coding Gene Expression via Premature Transcription Termination. *Cold Spring Harb Symp Quant Biol*. pii: 039644.
- Turunen J, Niemel E, Verma B, Frilander M. 2013. The significant other: splicing by the minor spliceosome. *WIREs RNA* 2013, 4:61–76.
- Valin a, Gill G. 2007. Regulation of the dual-function transcription factor Sp3 by SUMO. *Biochemical Society transactions*, 35; 1393–6.
- van der Veen AG, Ploegh HL. 2012. Ubiquitin-Like Proteins. *Annual Review of Biochemistry*, 81:.323–357
- Verger A, Perdomo J, Crossley, M. 2003. Modification with SUMO. A role in transcriptional regulation. *EMBO reports*, 4.137–142.
- Vertegaal ACO, Andersen JS, Ogg SC, Hay RT, Mann M, Lamond AI. 2006. Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Mol. Cell. Proteomics*. 5; 12:2298–2310.
- Vethantham V, Rao N, Manley JL. 2007. Sumoylation modulates the assembly and activity of the pre-mRNA 3' processing complex. *Molecular and cellular biology*. 27(24), 8848–8858.
- Vijay-Kumar S, Bugg CE, Cook WJ. 1987. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol*. 194; 3: 531-44.
- Warren DJ. 2011. Preparation of highly efficient electrocompetent *Escherichia coli* using glycerol/mannitol density step centrifugation. 2011. *Anal Biochem*. Jun 15;413(2):206-7.

White RJ. 2011. Transcription by RNA polymerase III: more complex than we thought. *Nat. Rev. Genet.* 12:459–63

Wilson G. 2005. Proteins of the PIAS family enhance the sumoylation of the papillomavirus E1 protein. *Virology.* 331; 1: 190-203.

Wu N, Kong X, Ji Z, Zeng W, Potts PR, Yokomori K, Yu H. 2012. Scc1 sumoylation by Mms21 promotes sister chromatid recombination through counteracting Wapl. *Genes Dev.* Jul 1;26(13):1473-85.

Xirodimas D. 2008. Novel substrates and functions for the ubiquitin-like molecule NEDD8. *Biochemical Society Transactions,* 36: 802–806.

Yamamoto J, Hagiwara Y, Chiba K, Isobe T, Narita T, Handa H, Yamaguchi T. 2014. DSIF and NELF interact with Integrator to specify the correct post- transcriptional fate of snRNA genes. *Nat. Commun.* 5, 4263.

Yang SH, Jaffray E, Senthinathan B, Hay RT, Sharrocks, AD. 2003. SUMO and transcriptional repression: dynamic interactions between the MAP kinase and SUMO pathways. *Cell cycle.* 2(6), 528–530.

Yang, SH, Sharrocks AD. 2004. SUMO promotes HDAC-mediated transcriptional repression. *Molecular Cell,* 13, 611–617.

Yang SH, Galanis A, Witty J, Sharrocks AD. 2006. An extended consensus motif enhances the specificity of substrate modification by SUMO. *EMBO J.* 25; 21: 5083-93.

Yeh ET. 2009. SUMOylation and De-SUMOylation: Wrestling with life's processes. *J Biol Chem.* 284; 13:8223-7

Yixuan Wua, Todd R. Albrechtb, David Baillatb, Eric J. Wagnerb and Liang Tonga. 2017. Molecular basis for the interaction between Integrator subunits IntS9 and IntS11 and its functional importance. *Proc Natl Acad Sci U S A,* 25; 114: 4394-4399

Yoon JB, Murphy S, Bai L, Wang Z, Roeder RG. (1995) Proximal sequence element binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. *Mol. Cell. Biol.* 15, 2019–2027

Yue J, Lai F, Beckedorff F, Zhang A, Pastori C, Shiekhattar R. 2017. Integrator orchestrates RAS/ERK1/2 signaling transcriptional programs. *Genes Dev.* Sep 1;31(17):1809-1820.

Zaborowska J, Egloff S, Murphy S. 2016. The pol II CTD: new twists in the tail. *Nat. Struct. Mol. Biol.* 23, 771–777.

Zaborowska J, Taylor A, Roeder RG, Murphy S. 2012. A novel TBP-TAF complex on RNA polymerase II- transcribed snRNA genes. *Transcription* 3, 92–104.

Zhang F, Ma T, Yu X. 2013. A core hSSB1-INTS complex participates in the DNA damage response. *J. Cell Sci.* 126, 4850–4855.

Zhao X, Blobel G. 2005. A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proceedings of the National Academy of Sciences of the United States of America.* 102(13).4777–82