



Universidad de Granada
Programa Oficial de Doctorado en Biomedicina

**Caracterización bioquímica y funcional del
factor de transcripción y *splicing* PRPF40B**

Soraya Becerra Ortiz

**Instituto de Parasitología y Biomedicina “López-Neyra”
Granada, 2015**

Editorial: Universidad de Granada. Tesis Doctorales

Autora: Soraya Becerra Ortiz

ISBN: 978-84-9125-072-2

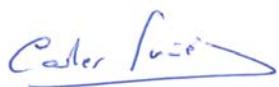
URI: <http://hdl.handle.net/10481/40006>

El doctorando SORAYA BECERRA ORTIZ y el director de la tesis CARLES M^a SUÑÉ NEGRE titulada “CARACTERIZACIÓN BIOQUÍMICA Y FUNCIONAL DEL FACTOR DE TRANSCRIPCIÓN Y SPLICING PRPF40B”:

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección del director de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 28 de Enero 2015

Director de la Tesis



Fdo.: Dr. Carles M^a Suñé Negre

Doctorando



Fdo.: Soraya Becerra Ortiz

Resumen

La variabilidad y complejidad característica de los organismos eucariotas viene dada por su expresión génica, un proceso que conlleva una serie de etapas altamente reguladas y conectadas entre sí que se inician con la transcripción del ADN para generar un ARNm que será traducido a proteínas. El procesamiento del ARNm mediante el *splicing* durante su maduración supone una etapa adicional de variabilidad, al generar diferentes isoformas de transcrito a partir de un mismo gen. Este proceso está muy regulado y errores en su control se asocian a diversas patologías humanas como las enfermedades neurodegenerativas y cáncer. La comunicación entre el *splicing* y su etapa anterior, la transcripción, aumentan el control de la regulación y mejoran la eficiencia global del proceso; por ello la relación e interconexiones existentes entre ambos procesos han sido ampliamente estudiados. Diversos factores celulares han sido propuestos como posibles acopladores entre la transcripción y el *splicing* al estar involucrados en ambos procesos y sus conexiones. Entre ellos destacan un grupo de proteínas que comparten unos motivos estructurales, los dominios WW y FF, capaces de establecer importantes interacciones para conectar las maquinarias de estos dos procesos.

Esta tesis aporta conocimientos básicos sobre la proteína con dominios WW y FF menos conocida hasta la fecha, PRPF40B, un homólogo humano del factor esencial de *splicing* en levaduras Prp40. Nuestros experimentos muestran la localización nucleoplasmática de PRPF40B y su enriquecimiento en los *speckles* nucleares, así como los dominios estructurales más relevantes responsables de su distribución. También hemos identificados importantes interacciones con factores de *splicing* que relacionan a PRPF40B con las etapas iniciales de este proceso. Asimismo, hemos realizado ensayos funcionales de *splicing* que muestran que PRPF40B es capaz de regular el *splicing* alternativo de genes apoptóticos, como *Fas*, y analizado secuencias necesarias para desempeñar dicho papel. Basándonos en los resultados con el gen *Fas*, hemos observado que el silenciamiento de PRPF40B incrementa el número de receptores de membrana Fas/CD95 y aumenta la muerte celular por apoptosis.

Mediante un análisis global del transcriptoma humano en células de neuroblastoma SH-SY5Y bajo condiciones de silenciamiento de PRPF40B hemos revelado que este factor está involucrado en rutas de cáncer, y en señalización y transporte en el sistema nervioso. Partiendo de los datos generados y realizando

ensayos funcionales, mostramos que PRPF40B es capaz asimismo de activar la transcripción. Además hemos demostrado que mutaciones asociadas al síndrome mielodisplásico disminuyen drásticamente la capacidad de activación transcripcional de PRPF40B.

Índice

I. ABREVIATURAS	1
II. INTRODUCCIÓN	7
1. LA EXPRESIÓN GÉNICA. LA TRANSCRIPCIÓN	9
1.1 La ARN polimerasa	9
1.2 El CTD de la RNAPII	9
1.3 Etapas de la transcripción	10
1.3.1 Iniciación	10
1.3.2 Elongación	10
1.3.2. Terminación	11
1.4 Regulación de la transcripción	12
2. EL PROCESAMIENTO DEL ARNm	14
2.1 El <i>spliceosoma</i>	14
2.2 La definición del exón	16
2.3 El <i>splicing</i> alternativo	17
2.4 Tipos de <i>splicing</i> alternativo	18
3. REGULACIÓN DEL <i>SPLICING</i>	19
3.1. Elementos reguladores <i>cis</i>	19
3.2. Elementos reguladores en <i>trans</i>	19
3.3 Mecanismos de regulación del <i>splicing</i>	22
4. ACOPLAMIENTO DE LA TRANSCRIPCIÓN Y EL <i>SPLICING</i>	25
4.1 Modelo cinético	25
4.2 Modelo de reclutamiento	28
4.2.1 Factores acopladores de la transcripción y el <i>splicing</i>	29
4.2.2 Proteínas con dominios WW y FF	31
5. EL NÚCLEO DE LA CÉLULA EUCARIOTA	38
5.1 Los <i>speckles</i> nucleares	39
5.1.1 Función de los <i>speckles</i> nucleares	40
5.1.2 Composición de los <i>speckles</i> nucleares	41
5.1.3 Dominios de señalización hacia los <i>speckles</i> nucleares	42
6. APOPTOSIS	44
6.1 Vía extrínseca	45

6.2 Vía intrínseca	45
6.3 Caspasas	46
7. APOPTOSIS Y <i>SPLICING</i>	48
7.1 Regulación de Fas	49
8. EL <i>SPLICING</i> ALTERNATIVO Y LA ENFERMEDAD	51
8.1 Tipos de alteraciones en el <i>splicing</i>	51
8.2 El <i>splicing</i> alternativo y el sistema nervioso	52
8.3 El <i>splicing</i> alternativo en el cáncer	54
III. OBJETIVOS	57
IV. MATERIALES Y MÉTODOS	61
1. PLÁSMIDOS	63
2. CULTIVOS CELULARES Y TRANSFECCIONES	65
2.1 Cultivos celulares	65
2.2 Transfecciones	66
2.3 Infección de cultivos celulares	67
3. PROCEDIMIENTOS DE BIOLOGÍA MOLECULAR	68
3.1 Inmunofluorescencia	68
3.2 Extracción de ARN y RT-PCR	69
3.3 Extracción de ARN total para los análisis por <i>array</i>	71
3.4 Mutagénesis dirigida de PRPF40B	72
3.5 Ensayos de actividad luciferasa	72
3.6 Ensayos de interacción proteína-proteína <i>in vivo</i>	73
3.7 Ensayos de interacción proteína-proteína <i>in vitro</i>	74
4. PROCEDIMIENTOS DE BIOLOGÍA CELULAR	74
4.1 Ensayos de muerte celular y ciclo celular	75
4.2 Ensayo de unión de Anexina V	75
4.3 Medida de la actividad de la caspasa-3	75
4.4 Análisis de expresión del receptor de membrana Fas/CD95	76

5. MANIPULACIÓN Y ANÁLISIS DE PROTEÍNAS	76
5.1 Análisis por <i>Western Blot</i>	76
5.2 Tinción de plata	77
5.3 Tinción de <i>Coomasie</i>	77
5.4 Producción y purificación de proteínas recombinantes en <i>E. coli</i>	77
5.5 Producción y purificación de proteínas recombinantes en células HEK293T	79
5.6 Producción de anticuerpo anti-PRPF40B en conejo	80
5.7 Purificación de IgG totales del suero anti-PRPF40B	80
6. ADN <i>Array</i>	81
7. OLIGONUCLEÓTIDOS Y ANTICUERPOS	82
7.1 Oligonucleótidos de ADN	82
7.2 Oligonucleótidos de ARN (siRNA)	85
7.3 Oligonucleótidos antisentido <i>Morpholino</i> (AMO)	85
7.4 Anticuerpos	86
V. RESULTADOS	89
1. ANÁLISIS ESTRUCTURAL DE PRPF40B	91
2. ANÁLISIS DE LA LOCALIZACIÓN CELULAR DE PRPF40B	93
2.1 PRPF40B es una proteína nuclear enriquecida en los <i>speckles</i> Nucleares	93
2.2 Identificación de las secuencias de PRPF40B necesarias para la localización nuclear y enriquecimiento en los <i>speckles</i> nucleares	95
2.3 La asociación de PRPF40B con los <i>speckles</i> nucleares no está afectada por el tratamiento con ARNasa o por la inhibición de la transcripción	100
3. ESTUDIO DE LA INTERACCIÓN DE PRPF40B CON LOS FACTORES DE <i>SPLICING</i> SF1 Y U2AF ⁶⁵	102
3.1 PRPF40B colocaliza con los factores de <i>splicing</i> SF1 y U2AF ⁶⁵ en el núcleo celular	102
3.2 Asociación de PRPF40B con el factor de <i>splicing</i> SF1 <i>in vitro</i>	104
3.3 Asociación de PRPF40B con los factores de <i>splicing</i> SF1 y U2AF ⁶⁵ <i>in vivo</i>	106

3.4 La asociación de PRPF40B con los factores de <i>splicing</i> SF1 y U2AF ⁶⁵ depende de ARN <i>in vivo</i>	108
4. ESTUDIO DEL EFECTO DE PRPF40B SOBRE EL <i>SPLICING</i> ALTERNATIVO	110
4.1 PRPF40B interviene en la selección del sitio de <i>splicing</i> alternativo	110
4.2 El efecto de PRPF40B en el <i>splicing</i> alternativo de <i>Fas</i> necesita secuencias del exón 6 y los sitios 5'SS y 3'SS	113
5. EFECTO DE PRPF40B EN APOPTOSIS	117
5.1 El silenciamiento de PRPF40B aumenta el número de receptores de Fas/CD95	117
5.2 La depleción de PRPF40B aumenta el porcentaje de muerte celular por apoptosis	118
5.3 El silenciamiento de PRPF40B aumenta la actividad de la caspasa-3	120
5.4 El silenciamiento de PTB aumenta el porcentaje de células en apoptosis	122
6. ANÁLISIS GLOBAL DEL SILENCIAMIENTO DE PRPF40B EN EL TRANSCRIPTOMA CELULAR	123
6.1 Análisis de los datos obtenidos usando el <i>Human Transcriptome Array (HTA)</i> en células SH-SY5Y	123
6.2 Validación de genes regulados a nivel de transcripción	127
6.3 Validación de genes regulados a nivel de <i>splicing</i>	131
7. ESTUDIO DE PRPF40B COMO ACTIVADOR DE LA TRANSCRIPCIÓN	134
7.1 PRPF40B activa la transcripción de promotores reporteros	134
7.2 Importancia de los dominios WW y FF de PRPF40B en la activación de la transcripción	136
8. IMPORTANCIA DE PRPF40B EN LA LEUCEMIA MIELOIDE AGUDA (AML)	138
VI. CONCLUSIONES	141
VII. DISCUSIÓN	151
1. PRPF40B es un factor nuclear enriquecido en los <i>speckles</i> nucleares	153

2. Caracterización de elementos estructurales necesario para la localización de PRPF40B	154
3. PRPF40B interacciona con los componentes del <i>spliceosoma</i> SF1 y U2AF ⁶⁵	156
4. PRPF40B modula el <i>splicing</i> alternativo	157
5. PRPF40B interviene en el proceso de apoptosis	159
6. PRPF40B se expresa diferencialmente en el tejido nervioso y funciona como activador de la transcripción	159
7. Implicación de PRPF40B en AML	162
8. Similitudes y diferencias entre PRPF40A y PRPF40B	164
9. PRPF40B como acoplador de los procesos de transcripción y <i>splicing</i>	165
VIII. BIBLIOGRAFÍA	169

Abreviaturas

3'SS	<i>3' splice site</i> , sitio de <i>splicing</i> 3'
5'SS	<i>5' splice site</i> , sitio de <i>splicing</i> 5'
ALPS	<i>autoimmune lymphoproliferative syndrome</i> , síndrome lifoproliferativo autoimmune
ALS	<i>amyotrophic lateral sclerosis</i> , esclerosis lateral amiotrófica
AMO	<i>antisense morpholino oligonucleotide</i> , oligonucleótidos antisentido Morpholino
AML	<i>acute myeloid leukemia</i> , leucemia mieloide aguda
ARNm	ARN mensajero
Bcl	<i>B cell lymphoma</i>
BPS	Branch-point sequence
CDK	<i>cyclin dependent kinase</i> , quinasa dependiente de ciclina
c-FLIP	<i>cellular FLICE inhibitory protein</i>
CF	citometría de flujo
ChIP-seq	<i>Chromatin immunoprecipitation sequencing</i> , secuenciación de inmunoprecipitación de cromatina
cIAP-1/2:	<i>Cellular Inhibitor of Apoptosis 1 y 2</i> / Inhibidores celulares de apoptosis 1 y 2.
CLIP-seq	<i>Cross-linked immunoprecipitation followed by next generation sequencing</i>
CMML	<i>chronic myelomonocytic leukemia</i> , leucemia mielomonocítica crónica
cSNP	<i>coding single-nucleotide polymorphism</i>
CTD	<i>carboxyl-terminal domain</i> , dominio carboxilo terminal
DISC	<i>death-inducing signaling complex</i>
DM1	distrofia miotónica 1
ED	<i>exon definition</i> , definición del exón
EGFP	<i>enhanced green fluorescent protein</i>
ESE/ESS	<i>exonic splicing enhancer/exonic splicing silencer</i>
FADD	<i>Fas-associated death domain</i>
FasL	Fas ligando
FBP	<i>Formin-binding protein</i>
FDTP-17	demencia frontotemporal y Parkinson asociado al cromosoma 17
FITC	<i>fluorescein isothiocyanate</i> , isotiocianato de fluoresceína
Fw	<i>forward</i>

Abreviaturas

h	hora
H3-K36me3	trimetilación de la histona 3 en la lisina 36
hnRNP	<i>heterogeneous nuclear ribonucleoprotein</i> , ribonucleoproteína heterogénea nuclear
HuR	Hu antígeno R
HYP A	<i>Huntigtin-Yeast Partner A</i>
HYP C	<i>Huntigtin-Yeast Partner C</i>
HTA	<i>Human transcriptome array</i>
IAP	<i>inhibitor of apoptosis proteins</i> , inhibidores de la apoptosis
IF	inmunofluorescencia
IGC	<i>interchromatin granule cluster</i> , gránulos intercromatínicos
ISE/ISS	<i>intronic splicing enhancer/intronic splicing silencer</i>
lncRNA	<i>long non-coding RNA</i> , ARN no codificante largo
MALAT1	<i>metastasis-associated lung adenocarcinoma transcript 1</i>
MDS	<i>myelodysplastic syndrome</i> , síndrome mielodisplásico
MOMP	permeabilización de la membrana externa mitocondrial, <i>mitochondrial outer membrane permeabilization</i>
min	minute
MPS	neoplasma mieloproliferativo
ncRNA	<i>non-coding RNA</i> , ARN no codificante
NIPP1	subunidad nuclear de la proteína fosfatasa 1
NLS	<i>nuclear localization signal</i> , señal de localización nuclear
PA	poliadenilación
PARP	<i>Poly ADP-ribose polymerase</i>
PFs	perichromatin fibrils, fibras pericromatínicas
PIC	<i>pre-initiation complex</i> , complejo preiniciación
PIP3	quinasa fosfatidil-inositol fosfato
PML	<i>promyelocytic leukemia</i> , leucemia promielocítica
PP1	proteína fosfatasa 1
PQB1	proteína de unión al tracto de poliglutaminas
PRPF40A	<i>PRP40 pre-mRNA processing factor 40 homolog A</i>
PRPF40B	<i>PRP40 pre-mRNA processing factor 40 homolog B</i>
PTB	<i>polypyrimidine tract binding protein</i>
Py	<i>polypyrimidine tract</i> , tracto polipirimidinas

Rev	<i>reverse</i>
RLU	unidades de luz relativa
RNAPII	ARN polimerasa II
RRM	<i>RNA recognition motif</i> , motivo de reconocimiento del ARN
s	segundo
SEM	error estándar de la media
Ser2	serina 2
Ser5	serina 5
SF1	<i>Splicing factor 1</i>
SMA	<i>spinal muscular atrophy</i> , atrofia muscular espinal
snoRNPs	<i>small nucleolar ribonucleoprotein</i> , ribonucleoproteínas pequeñas nucleolares
SNP	<i>single-nucleotide polymorphism</i>
SRE	<i>splicing regulatory elements</i> , elementos reguladores de <i>splicing</i>
TCERG1	<i>Transcription Elongation Regulator 1</i>
Th4	treonina 4
TIA-1	<i>T-cell intracellular antigen-1</i>
TNFR	<i>tumor necrosis factor receptor</i> , receptor del factor de necrosis tumoral
TSS	<i>transcription start site</i> , sitio de inicio de la transcripción
U snRNP	<i>uridine-rich small nuclear ribonucleoprotein</i>
U2AF	<i>U2 snRNP auxiliar factor</i>
URE6	<i>uridine-rich exonic 6 element</i>
URI6	<i>uridine-rich intronic 6 element</i>
WB	<i>western blot</i>
WT	<i>wild type</i> /salvaje
XIAP	<i>X-linked Inhibitor of Apoptosis</i>
YBX3	<i>Y box binding protein 3</i>

Introducción

1. LA EXPRESIÓN GÉNICA. LA TRANSCRIPCIÓN

La expresión génica es el proceso mediante el cual los organismos utilizan la información genética para sintetizar proteínas atendiendo a sus necesidades. Comprende etapas esenciales como la transcripción, donde se sintetiza ARN mensajero (ARNm) a partir del ADN, el procesamiento y la maduración del ARNm que será transportado desde el núcleo al citoplasma para ser traducido a proteína por los ribosomas.

1.1 La ARN polimerasa

La transcripción es llevada a cabo por la ARN polimerasa. En organismos eucariotas existen tres polimerasas que realizan este proceso, cada una con diferente especificidad. La ARN polimerasa I que transcribe genes que codifican para ARN ribosómico 18S y 28S. La ARN polimerasa II (RNAPII) que transcribe genes que codifican para proteínas (ARNm) y ARN no codificante como el ARN pequeño nuclear (snRNA). La ARN polimerasa III que transcribe genes cortos como el ARN de transferencia o el ARN ribosomal 5S. Destacar por su singularidad y relevancia funcional el extremo carboxilo-terminal (CTD, *carboxyl terminal domain*) de la subunidad mayor (Rbp1) de la RNAPII (Corden et al., 1985).

1.2 El CTD de la RNAPII

El CTD de la RNAPII está compuesto por un número variable de repeticiones del heptapéptido $Y_1S_2P_3T_4S_5P_6S_7$ (52 repeticiones en humanos) más un decapeptido carboxilo-terminal de secuencia ISPDDSDEEN que confiere estabilidad a la enzima (Bartolomei et al., 1988). EL CTD es esencial para la viabilidad de los organismos y el correcto funcionamiento del proceso de formación del ARNm (Fong et al., 2003; McCracken et al., 1997), aunque no es necesario para la iniciación de la transcripción *in vitro* (Zehring et al., 1988).

El CTD puede ser modificado por fosforilación y ello está estrechamente relacionado con cada etapa del ciclo de transcripción como se describirá a continuación, generando lo que se ha denominado como el “código CTD” (Egloff and Murphy, 2008). Los principales residuos modificados por fosforilación son las serinas 2 (Ser2) y 5 (Ser5), aunque también se ha descrito la fosforilación de la serina 7 (Ser7) o de la

treonina 4 (Thr4) (Chapman et al., 2007; Hsin, Sheth, and Manley, 2011; Patturajan et al., 1998). Además de la fosforilación, se han descrito otro tipo de modificaciones postraduccionales como la glicosilación, metilación o isomerización de la prolina, aunque la importancia funcional de estas modificaciones está todavía por investigar (Kelly, Dahmus, and Hart, 1993; Sims et al., 2011).

1.3 Etapas de la transcripción

La transcripción puede dividirse en tres etapas: iniciación, elongación y terminación. Todas ellas dependen de la secuencia del ADN, la estructura de la cromatina y de factores que son reclutados o liberados del complejo transcripcional dependiendo del patrón de fosforilación del CTD (Egloff, Dienstbier, and Murphy, 2012).

1.3.1 Iniciación

En la iniciación de la transcripción se ensambla el complejo de preiniciación (PIC, *pre-initiation complex*) en la secuencia promotora cuando existe una apertura de la cromatina adecuada, y donde tiene lugar la adición del residuo de metilguanosina en el extremo 5' del ARN (*5' capping*) (Buratowski, 2009; Panning and Taatjes, 2008). El PIC está formado por la RNAPII y factores generales de transcripción como TFIIA, TFIIB, TFIID, TFIIE, TFIIF y TFIIH (Roeder, 1996). La RNAPII es reclutada al promotor cuando su CTD está hipofosforilado, al interaccionar con componentes del PIC y con el complejo Mediador, que favorecen la el reconocimiento al promotor (Myers et al., 1998; Phatnani and Greenleaf, 2006). Además el reclutamiento del PIC está regulado por factores unidos a secuencias anexas al promotor, como NF- κ B, E2F o ARF-2, que activan o inhiben la activación del complejo transcripcional. También factores implicados en la remodelación de la cromatina, como SWI/SNF, o en la modificación de histonas, influyen en el inicio de la transcripción (Berk, 1999; Roeder, 2005)

1.3.2. Elongación

La elongación es la etapa en la que se forma el transcrito a partir del ADN molde. Comienza desde que la RNAPII se separa del complejo de iniciación y puede dividirse a su vez en dos etapas, temprana y tardía o procesiva, que se diferencian por el

estado de fosforilación del CTD. La fosforilación de la Ser5 por la quinasa CDK7 (*cyclin dependent kinase 7*) está asociada a la etapa temprana de transcripción y a la región 5' de los genes y parece desestabilizar las uniones de la RNAPII con el PIC (Lu et al., 1992). La fosforilación de la Ser2 por la quinasa CDK9 (subunidad catalítica del factor de elongación de la transcripción P-TEFb), convierte a la RNAPII a la forma productiva, y se asocia con la región 3' de los genes (Marshall and Price, 1995). Aunque tradicionalmente se ha asociado a CDK9 con la elongación de la transcripción, también se han descrito más recientemente fosforilaciones en Ser2 por las quinasas CDK12 y CDK13 (Bartkowiak et al., 2010). Además de la Ser2 y Ser5 también se ha visto que la Ser7 es susceptible de ser fosforilada, y esta modificación se ha asociado con el extremo 5' de los genes, sugiriendo que podría estar implicada también en la elongación de la transcripción (Kim et al., 2010).

1.3.3. Terminación

La terminación de la transcripción consiste en el reconocimiento del extremo 3' de los genes y la disociación del transcrito sintetizado de la hebra molde de ADN. Este proceso altamente regulado, está íntimamente relacionado con el procesamiento del extremo 3' y la poliadenilación (Yonaha and Proudfoot, 2000). La RNAPII también se separa del ADN y es reciclada para nuevos ciclos de transcripción (Yao et al., 2007)

Aunque el código de fosforilación del CTD parece ser general durante el ciclo transcripcional, existen modificaciones que afectan de manera específica a determinados genes. Es el caso de la fosforilación de la Ser7 por CDK7 y CDK9, relacionada con la transcripción del snRNA, o la fosforilación de la Thr4 probablemente por la quinasa Plk3, cuya función todavía no está clara pero parece estar asociada en mamíferos con la maduración de los extremos 3' del ARNm sin intrones (Egloff, Dienstbier, and Murphy, 2012; Hintermair et al., 2012).

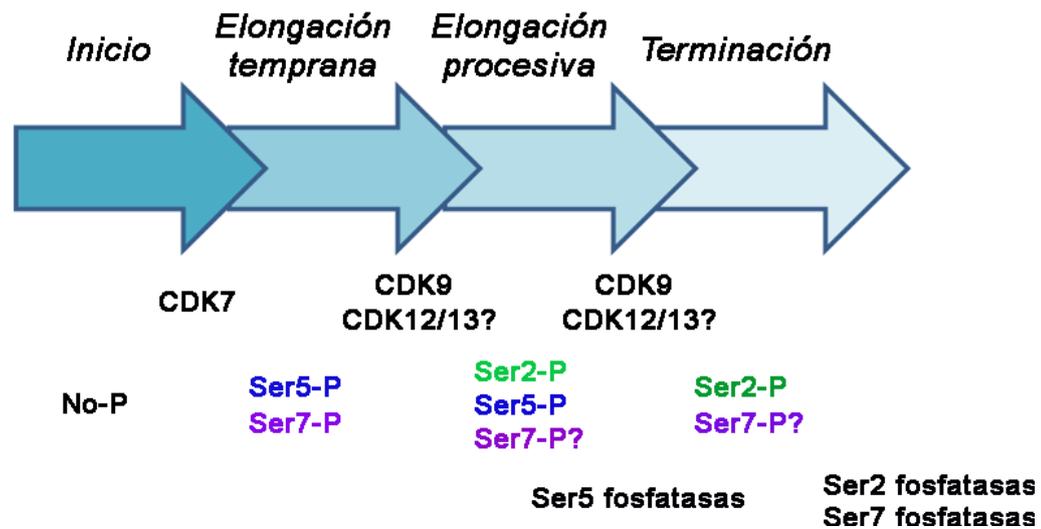


Figura I-1. Ciclo de fosforilación del CTD durante las fases de la transcripción. En la fase de inicio el CTD no está fosforilado (No-P). En el comienzo de la elongación temprana la enzima CDK7 es reclutada y fosforila a la Ser5 y la Ser7 del CTD de la RNAPII. Durante la elongación procesiva la Ser2 es fosforilada principalmente por CDK9. En la fase de terminación la Ser5 es desfosforilada. Finalmente cuando termina la transcripción los residuos de Ser2 y Ser7 son desfosforilados y la RNAPII es reciclada para un nuevo ciclo de transcripción.

1.4 Regulación de la transcripción

Aunque tradicionalmente se ha pensado que la iniciación de la transcripción era la etapa limitante de este proceso, estudios más recientes dan a la fase de la elongación un papel determinante en la regulación de la transcripción.

Un importante punto de control de la transcripción es la parada de la RNAPII en regiones próximas al promotor (*promoter proximal pausing*), a unos 30-100 nucleótidos del sitio de comienzo de la transcripción (TSS, *transcription start site*), una vez se ha ensamblado el PIC y se ha producido el *capping*, hasta la activación de señales que inducen la elongación. En un principio parecía que la parada de la RNAPII estaba asociada a determinados genes, pero estudios globales del genoma mediante ChIP-seq (*Chromatin immunoprecipitation sequencing*), han mostrado la generalidad de este mecanismo de regulación de la transcripción, sobre todo en genes de desarrollo (Levine, 2011; Nechaev et al., 2010). Todavía no se conoce con exactitud el mecanismo de parada y liberación de la RNAPII en el promotor, aunque se han descrito diversos

factores que están implicados en esta regulación. El caso más conocido es la inhibición por los factores NELF y DSIF que impiden el comienzo de la elongación transcripcional, al aumentar la afinidad de la RNAPII por el promotor (Core and Lis, 2008). La reactivación de la RNAPII se produce por el reclutamiento de P-TEFb, que fosforila el CTD en la Ser2, así como a DSIF, convirtiéndolo en activadores de la transcripción (Bres, Yoh, and Jones, 2008). P-TEFb está a su vez regulado por diversos mecanismos como modificaciones postraduccionales o activadores transcripcionales como NF- κ B o Brd4 (Barboric et al., 2001; Yang et al., 2005).

Además de la parada de la RNAPII en regiones próximas al promotor, dos publicaciones recientes midiendo el ARN naciente de forma global en levaduras han demostrado que la RNAPII se acumula transitoriamente en sitios específicos antes de la terminación y alrededor de los sitios 3' de splicing (Carrillo Oesterreich, Preibisch, and Neugebauer, 2010). Aunque la naturaleza y la posición de la parada observada difieren entre ambos estudios, se observó una fuerte correlación entre la RNAPII en esos lugares y el splicing co-transcripcional (ver más adelante). Estas observaciones están de acuerdo con el modelo en el que las paradas transcripcionales actúan como un punto de control regulatorio, en este caso para un evento de procesamiento situado tras el punto de parada.

Además de las paradas de la RNAPII, otro importante mecanismo de regulación de la transcripción son las marcas epigenéticas. La presencia de nucleosomas puede actuar como una barrera dificultando la elongación transcripcional de la RNAPII. Por tanto, mecanismos que ayudan a desensamblar los nucleosomas mediante modificaciones de las histonas o remodeladores de la cromatina como FACT y SWI/SNF ayudan a descompactar la cromatina, facilitando la actividad transcripcional de la RNAPII (Li, Carey, and Workman, 2007).

2. EL PROCESAMIENTO DEL ARNm

El ARN generado en el proceso de la transcripción es sometido generalmente a un proceso de maduración que comprende la adición de un residuo de metilguanosa en su extremo 5' (*capping*), la eliminación de los intrones y unión de los exones (*splicing*) y el corte del extremo 3' seguido de la adición de una cola de adeninas (poliadenilación). Por interés para este trabajo, se hará a continuación una descripción más detallada sobre el proceso de *splicing*. En eucariotas superiores, la mayoría de genes están formados por secuencias que contienen segmentos codificantes (exones) de aproximadamente 50-250 nucleótidos, separados de secuencias no codificantes (intrones) que pueden llegar a tener un tamaño de hasta 1000 kb. El *splicing* consiste en la eliminación de estas secuencias intrónicas, seguida de la unión de los exones para dar lugar a un transcrito maduro.

2.1 El *spliceosoma*

El proceso de eliminación de intrones y unión de exones es llevado a cabo por el *spliceosoma*, un complejo macromolecular compuesto por cinco partículas ribonucleoproteicas U1, U2, U4, U5 y U6 snRNPs (*Uridine-rich small nuclear ribonucleoprotein*) que conforman el centro catalítico, y más de 200 proteínas adicionales. Cada partícula snRNP está compuesta por un snRNA y al menos siete proteínas auxiliares. Para el correcto desarrollo de la reacción de *splicing*, los snRNPs deben reconocer secuencias consenso características situadas en el precursor del ARNm (pre-ARNm) mediante la unión de sus snRNAs. Sin embargo, estas secuencias se encuentran poco conservadas en metazoos, siendo necesaria la asistencia de proteínas auxiliares que ayudan a la estabilización de cada snRNP (Ast, 2004).

Aunque no hay un acuerdo establecido sobre cómo se lleva a cabo el ensamblaje del *spliceosoma*, diversos trabajos realizados *in vivo* e *in vitro* coinciden en que es un proceso ordenado y secuencial, donde se pueden diferenciar varias fases determinadas por la formación de complejos característicos (Hoskins et al., 2011; Huranova et al., 2010).

Las primeras etapas de ensamblaje del *spliceosoma* comienza con el reconocimiento específico de los sitios que definen la unión intrón/exón, así como de

las secuencias consenso presentes en intrones y exones. La subunidad U1 snRNP se une por complementariedad de bases al sitio 5' de *splicing* (*5' splice site*, *5'SS*) del intrón. Sin embargo, esta unión ARN-ARN suele ser débil y necesita la presencia de proteínas propias de U1 snRNP y otros factores, como las proteínas SR, que ayudan a estabilizar esta asociación. A su vez, el factor de *splicing* SF1 reconoce la secuencia intrónica del *branch-point* (*Branch-Point Sequence*, *BPS*) y el heterodímero U2AF (*U2 snRNP auxiliary factor*) se une al tracto de polipirimidinas (Py) y a la secuencia AG del sitio 3' de *splicing* (*3' splice site*, *3'SS*), a través de la subunidad de 65 kDa y 35 kDa, respectivamente. Estas interacciones tempranas conforman el complejo E o complejo comprometido (*commitment complex*), el cual es crítico para el reconocimiento inicial de los sitios *5'SS* y *3'SS* así como para la continuidad del proceso de *splicing*. Seguidamente, se incorpora la partícula U2 snRNP al sitio *BPS* del pre-ARNm desplazando a SF1 del mismo en una reacción dependiente de ATP, dando lugar al complejo A. Esta unión entre ARNs es estabilizada por proteínas propias de la partícula U2 snRNP así como por U2AF⁶⁵. A continuación, es reclutado el triplete formado por las tres subunidades U4-U5-U6 snRNPs para formar el complejo B. Este complejo sufre una serie de reordenamientos conformacionales en el que U1 y U4 snRNP son liberados y las partículas U2 y U6 snRNP se aproximan dando lugar finalmente al complejo C, que es el encargado de llevar a cabo el proceso de *splicing* propiamente dicho, mediante una reacción de transesterificación dependiente de ATP. Finalmente, el pre-ARNm ya procesado es liberado del *spliceosoma* y las subunidades U2, U5 y U6 snRNP son recicladas para posteriores ciclos de *splicing*.

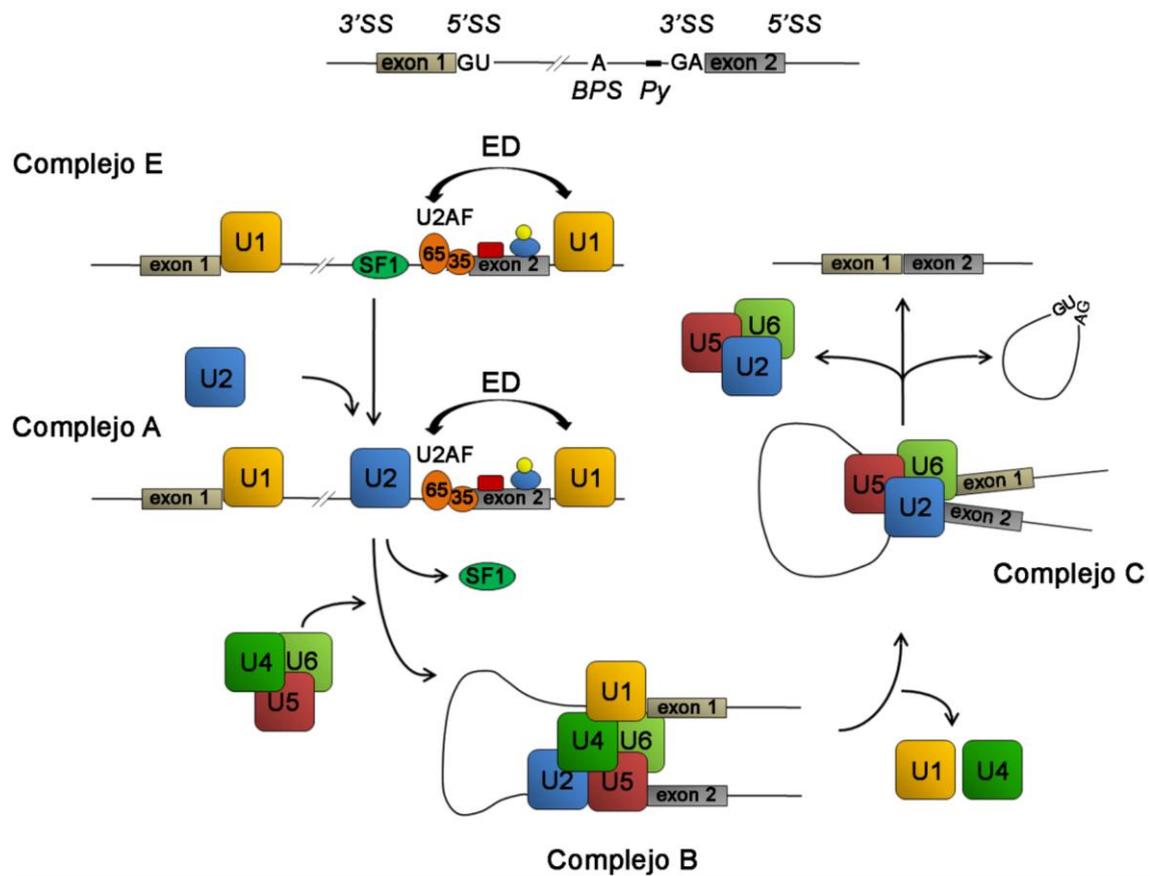


Figura I-2. Ensamblaje del *spliceosoma* y reacción de *splicing*. Diagrama esquemático de la formación de cada complejo (E, A, B, C) durante las etapas de ensamblaje del *spliceosoma*. Se representan las subunidades snRNPs U1, U2, U4, U5 y U6; así como los factores de *splicing* SF1 y U2AF. 3'SS, sitio de *splicing* 3' (GA); 5'SS, sitio de *splicing* 5' (GU); BPS, *branch point site* (A); Py, tracto de polipirimidinas; ED, definición del exón.

2.2 La definición del exón

El reclutamiento de factores a ambos lados del intrón durante el ensamblaje inicial del *spliceosoma* no ocurre de forma independiente, sino que existe una comunicación entre ellos que facilita el reconocimiento mutuo de los sitios 5'SS y 3'SS. Sin embargo, el extenso tamaño de los intrones en comparación con la longitud de los exones, hace que esta conexión tenga lugar a través de interacciones mediadas a lo largo del exón en un proceso conocido como definición del exón (*exon definition*) (Berget, 1995). Seguidamente, tiene lugar la definición del intrón, a través de interacciones establecidas a lo largo del intrón, permitiendo el acercamiento de las subunidades U1 y U2 snRNPs del *spliceosoma* y consecuentemente de los sitios 3'SS y 5'SS, dando lugar a un complejo funcional (De Conti, Baralle, and Buratti, 2013).

2.3 El *splicing* alternativo

El proceso de ensamblaje del *spliceosoma* y consecuente reacción de *splicing* es un proceso altamente regulado. Por eso, atendiendo a la eficiencia con que esto ocurre en un determinado exón se pueden distinguir dos tipos de eventos. El *splicing* es constitutivo cuando el reconocimiento de un exón y ensamblaje del *spliceosoma* se producen correctamente y de la misma forma en todos los pre-ARNm de un gen. El *splicing* alternativo es aquel en que los sitios 5' y 3' de *splicing* compiten por el reconocimiento y unión del *spliceosoma* con otros sitios 5'SS y 3'SS. Como consecuencia este exón no será procesado de la misma forma en todos los pre-ARNm de un gen, dando lugar a diferentes isoformas del transcrito maduro. Aunque en la mayoría de los casos descritos las primeras etapas de reconocimiento y ensamblaje de *spliceosoma* son cruciales para determinar el tipo de procesamiento que tendrá lugar, cada vez hay más ejemplos en los que esta decisión también se establece en etapas tardías de ensamblaje, incluso durante la reacción de transesterificación (Chen and Manley, 2009).

El *splicing* alternativo genera a partir de un gen distintos ARNm que codifican para proteínas diferentes, a veces incluso con funciones opuestas dependiendo del tejido o tipo celular (Wang et al., 2008). Por eso, el *splicing* alternativo constituye un importante mecanismo de diversidad proteica en los organismos eucariotas (Blencowe, 2006). Estudios de secuenciación masiva han estimado que en humanos más del 90% de los genes son regulados por *splicing* alternativo (Pan et al., 2008; Wang et al., 2008), corroborando la importancia de este proceso en la diversidad evolutiva. Además, aproximadamente el 15% de enfermedades humanas de origen hereditario, incluyendo enfermedades neurodegenerativas y cáncer, están asociadas a alteraciones en la regulación del *splicing* alternativo, bien por mutaciones en las secuencias consenso o por fallos en la propia maquinaria del *spliceosoma* (Graveley, 2000; Wang and Cooper, 2007; Ward and Cooper, 2010). Sin embargo y a pesar de la relevancia biológica del *splicing* alternativo, los mecanismos moleculares que controlan este proceso y las alteraciones del mismo que conducen a enfermedad todavía no se conocen por completo.

2.4 Tipos de *splicing* alternativo

La existencia del *splicing* alternativo implica el uso de sitios de *splicing* diferentes a los convencionales durante el procesamiento del transcrito para crear diversos patrones de ARNm a partir de un gen, favoreciendo la diversidad proteica. En un típico ARN multiexónico estos eventos pueden darse de diversas formas (Black, 2003):

- Inclusión o exclusión de un exón completo (*cassete exon*).
- Exones mutuamente excluyentes, basado en el caso anterior cuando determinados exones no coinciden en el ARNm maduro.
- Sitio aceptor alternativo de *splicing*, es aquel que tiene un sitio 5' de *splicing* alternativo dando lugar a un exón más largo.
- Sitio donador alternativo, se da cuando existe un sitio 3' de *splicing* alternativo, produciéndose un acortamiento del exón.
- Promotor alternativo o primer exón alternativo, tiene lugar cuando el transcrito puede empezar en exones diferentes.
- Señal de polyA o último exón alternativo, el ARNm puede terminar en exones distintos.
- Retención de intrón, tiene lugar cuando la secuencia de un intrón es incluida en el transcrito final.
- Combinación de varios de los eventos descritos anteriormente en un mismo pre-ARNm.

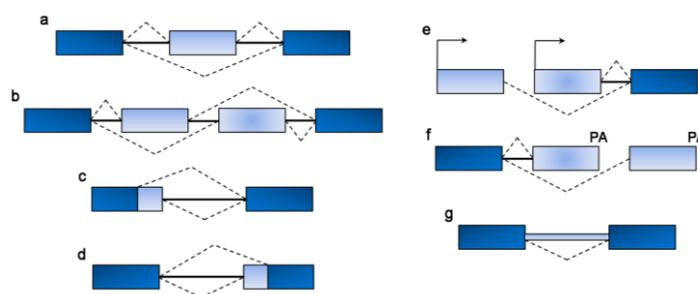


Figura I-3. Tipos de eventos de *splicing* alternativo. Diagrama esquemático de los diferentes patrones de procesamiento que puede sufrir un pre-ARNm. Se representan los exones constitutivos (azul oscuro), los exones alternativos (azul claro), intrones (negro), los patrones de *splicing* (líneas discontinuas), inicio del marco de lectura (flecha) y sitios de poliadenilación (PA). **(a)** *exon cassette*, **(b)** exones mutuamente excluyentes, **(c)** sitio 5' de *splicing* alternativo, **(d)** sitio 3' de *splicing* alternativo, **(e)** primer exón alternativo, **(f)** sitios de poliadenilación alternativos y **(g)** retención de intrones. Los patrones de *splicing* superiores indican el uso de los sitios de *splicing* convencionales y los patrones de *splicing* inferiores corresponderían al uso de sitios de *splicing* alternativos.

3. REGULACIÓN DEL *SPLICING*

3.1. Elementos reguladores en *cis*

Los procesos de *splicing* constitutivo y alternativo son dirigidos por la maquinaria del *spliceosoma*, por tanto, la regulación de su ensamblaje es crucial para asegurar la precisión y especificidad características de la reacción de *splicing*. Sin embargo, la identificación de los exones resulta difícil debido a la extensa longitud de los intrones (aproximadamente 5 kb de media) y el escaso nivel de conservación de las secuencias consenso (3'SS, 5'SS, BPS y Py). Como consecuencia, estas secuencias consenso son en general insuficientes para coordinar el correcto ensamblaje del *spliceosoma* y se necesitan secuencias reguladoras adicionales que garanticen el preciso funcionamiento de la maquinaria de *splicing* para la definición del exón y el intrón.

Estos elementos reguladores del *splicing* (*splicing regulatory elements, SRE*) son secuencias en *cis* situadas estratégicamente en el ARN. Se pueden clasificar en cuatro categorías en función de su localización (intrón o exón) y su efecto (potenciador o inhibidor) en el *splicing* como ISE/ISS, ESE/ESS (*intronic splicing enhancer/intronic splicing silencer, exonic splicing enhancer/ exonic splicing silencer*). Estas secuencias son reconocidas por proteínas con dominios de unión a ARN que se unen a estos elementos donde ejercerán un efecto potenciador o represor de la reacción de *splicing*. La concentración o actividad de estos factores reguladores determinarán el destino de los exones alternativos y por tanto, su inclusión o exclusión en el transcrito maduro. Dentro de los tipos de proteínas reguladoras del *splicing* que se conocen, destacan las proteínas SR que actúan como activadores del *splicing* al reconocer sitios ISE y ESE, y los hnRNP que ejercen un efecto inhibidor de la reacción de *splicing* al unirse a secuencias de tipo ISS y ESS.

3.2. Elementos reguladores en *trans*

Activadores del *splicing*

La familia de proteínas SR son un importante grupo de factores reguladores del *splicing* constitutivo y alternativo del pre-ARNm. Se trata de un conjunto de proteínas muy conservadas, caracterizadas por contener dominios RS, ricos en argininas y serinas,

de interacción proteína-proteína en su región carboxilo terminal y dominios de unión a ARN (RRM, *RNA recognition motif*) en su extremo amino terminal. Esta organización estructural las hace idóneas para desempeñar su función como reguladoras del *splicing*. Mediante sus dominios RRM reconocen elementos reguladores presentes en el pre-ARNm, como los ESE, cuya especificidad vendrá dada por cada proteína SR (Tacke and Manley, 1999) y a través de los motivos RS establecen puentes de unión con factores del propio complejo del *spliceosoma* u otras proteínas reguladoras dando lugar a un complejo activador que en última instancia favorecerá el reconocimiento del exón (Graveley, 2000). Existen dos modelos que explican el mecanismo de actuación de las proteínas SR como activadores del *splicing*: el “*modelo de reclutamiento*” sostiene que la unión de la proteína SR al sitio ESE recluta y estabiliza las interacciones de U1snRNP al sitio 5'SS y de U2AF⁶⁵ al 3'SS, favoreciendo la definición del exón (Graveley, Hertel, and Maniatis, 2001). El “*modelo de inhibición*” indica que la unión de las proteínas SR a los ESE puede antagonizar el efecto inhibitorio que las proteínas hnRNPs ejercen al unirse a los ESS.

También existen otras proteínas reguladoras del *splicing*, que de forma similar a como ocurre con las proteínas SR, reconocen las secuencias reguladoras consenso y promueven el reclutamiento y reconocimiento de otros miembros del *spliceosoma* facilitando su ensamblaje. El factor TIA-1 (*T-cell intracellular antigen-1*), como también se verá más adelante, reconoce una secuencia intrónica rica en uridinas situada cerca del sitio 5'SS del intrón 6 de *Fas*, y ayuda al reclutamiento de la subunidad U1snRNP al sitio 5'SS a través de su dominio rico en glutaminas en su región carboxilo terminal. Además parece que TIA-1 estaría bloqueando un elemento silenciador para hnRNP F/H impidiendo su acceso y por tanto, que inhiba el reconocimiento de dicho exón (Izquierdo et al., 2005). Otro ejemplo es la proteína de unión al ARN Sam68, cuya función como regulador del *splicing* varía dependiendo del contexto, pero en el caso del pre-ARNm del gen *CD44* colabora en el reclutamiento de U2AF al sitio 3'SS (Tisserant and König, 2008).

Represores del *splicing*

La inhibición del *splicing* alternativo a través de los elementos ISS y ESS es llevada a cabo por las ribonucleoproteínas heterogéneas nucleares (hnRNP), un amplio

y diverso grupo de proteínas de unión al ARN. El bloqueo de la reacción de *splicing* puede ser en general explicado mediante dos mecanismos moleculares:

El modelo de “unión y bloqueo” (“*bind and block*”) tiene lugar cuando las secuencias reguladoras se encuentran próximas a los sitios de *splicing*, generando una barrera física que limita el acceso de los componentes del *spliceosoma*. Es el caso de RBM5, que ejerce su efecto represor impidiendo la incorporación del triplete U4/U6-U5 snRNPs a los intrones adyacentes al exon 6 de *Fas* y bloqueando así la maduración del *spliceosoma*. En este caso, RBM5 además promueve la incorporación de las partículas U4/U5/U6 snRNPs en los exones adyacentes del exon 6 de *Fas* (Bonnal et al., 2008). Otra forma de inhibición por bloqueo consiste en competir con factores activadores del *splicing* por la unión a secuencias reguladoras. Por ejemplo Hu/ELAV inhibe la unión de U1snRNP compitiendo con TIA-1 por la unión a una secuencia rica en AU (Zhu et al., 2008); o la proteína SRSF7, que compete con hnRNP F y hnRNP H por la unión a una secuencia consenso para regular el *splicing* del exón 2 del gen de la α -tropomiosina (Crawford and Patton, 2006). También ocurre este mecanismo de inhibición con la proteína de unión al tracto de polipirimidinas PTB (*polypirimidine tract binding protein*), ya que debido a su afinidad por las secuencias ricas en uridinas, compete por la unión al tracto de polipirimidinas con U2AF⁶⁵ dificultando así su unión del mismo (Lin and Patton, 1995).

Además del bloqueo físico existe otro mecanismo de inhibición (“*looped out*”) que tiene lugar cuando las secuencias reguladoras se encuentran lejanas a los sitios de *splicing* y se producen interacciones proteína-proteína a lo largo del exón generando cambios conformacionales que dificultan el ensamblaje del *spliceosoma*, incluso cuando los sitios de reconocimiento de *splicing* no han sido bloqueados (Nasim et al., 2002). Esto ocurre con hnRNP-A1, que inhibe la inclusión del exon 7 del gen *SMN2* o PTB en el exon N1 de *c-SRC*, que se unen a secuencias reguladoras intrónicas y exónicas interaccionando entre ellas y generando una estructura que impide el reconocimiento de estos exones, y además se facilita el acceso de elementos de *spliceosoma* a exones adyacentes evitando finalmente la inclusión de estos exones en el transcrito maduro (Hutchison et al., 2002; Spellman and Smith, 2006). En el caso de hnRNP L, además de ejercer un bloqueo físico en el reclutamiento de U1 y U2 snRNP, el reconocimiento de

las secuencias reguladoras ESS puede provocar un cambio conformacional en el pre-ARNm que previene la unión al intrón de estos snRNPs (House and Lynch, 2006).

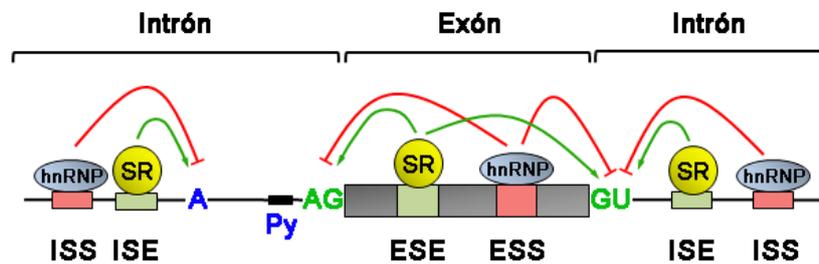


Figura I-4. Regulación del *splicing* alternativo mediante secuencias en *cis* y factores en *trans*. Diagrama esquemático de la regulación del *splicing* alternativo donde se representan: los elementos reguladores potenciadores ISE/ESE (*intrinsic/exonic splicing enhancer*) o inhibidores ISS/ESS (*intrinsic/exonic splicing enhancer*) de *splicing*; proteínas SR que favorecen el reconocimiento de los sitios de *splicing* y hnRNP que lo inhiben; *Branch Point Site* (A, azul), trazo de polipirimidina (Py, azul) y los sitios 5'SS y 3'SS. (3', AG, verde; 5', GU, verde).

3.3 Mecanismos de regulación del *splicing*

Ya hemos visto que el ensamblaje del *spliceosoma* y consecuente reacción de *splicing* es un proceso altamente controlado, en el que intervienen diversos elementos reguladores que aseguran el correcto funcionamiento. Sin embargo, este “código de *splicing*” no es suficiente para entender cómo se desarrolla completamente este proceso, ya que el nivel de complejidad es mayor de lo que en principio aparenta. Por eso cada vez hay más trabajos que intentan descifrar todos los factores que influyen en el correcto funcionamiento del *splicing*. Entre los mecanismos de control que se conocen se encuentran los siguientes:

- **Modificación de los elementos en *trans* reguladores del *splicing*.** Los casos más estudiados son los mecanismos de modificación postraducciona, principalmente por fosforilación, que determinan la actividad de un factor regulador de una forma espacio-temporal. Este hecho es habitual en las proteínas SR, altamente reguladas por procesos de fosforilación/defosforilación, de cuyo estado dependen las interacciones que establecen en el contexto del *splicing* (Black, 2003). También ocurre con Sam68, cuyo efecto modulador en *Bcl-x* viene dado por la fosforilación de un residuo de tirosina (Paronetto et al., 2007).

- **Modificación de los elementos en *cis* reguladores del *splicing*.** Un evento de *splicing* alternativo puede ser modificado por la presencia de mutaciones en los elementos reguladores en *cis*, al cambiar la afinidad de los reguladores en *trans*, con consecuencias patológicas importantes (Caceres and Kornblihtt, 2002). Un ejemplo se ha descrito en el exón 7 del gen *SMN2*, donde el cambio de C a T detectado en pacientes con atrofia muscular espinal genera un elemento ESS reconocido por Sam68 que modula su *splicing* alternativo (Pedrotti et al., 2010).

- **Especificidad de tejido y etapa del desarrollo.** Determinadas proteínas de unión al ARN pueden cambiar sus propiedades bioquímicas dependiendo del tejido en que se expresan o la etapa del desarrollo o diferenciación en que se encuentran, generando patrones de *splicing* diferentes dependiendo en cada tipo celular o estadio del desarrollo. Un ejemplo es el factor de *splicing* nPTB, cuya expresión y mecanismo de actuación se produce de forma exclusiva en neuronas y no en otro tipo celular (Coutinho-Mansfield et al., 2007).

- **Contexto de los elementos en *cis* del *splicing*.** La posición en que se encuentra un elemento regulador (intrón, exón, 5'SS, 3'SS) y su accesibilidad pueden determinar que un mismo factor de *splicing* pueda comportarse como un activador o un represor dependiendo del lugar en que lo reconozca. Existen diversos ejemplos en que proteínas SR o hnRNPs actúan inhibiendo o promoviendo el reconocimiento de un exón, según el contexto en el que se produce la unión. Uno de los casos más conocidos es el de NOVA, un factor de *splicing* neuronal que actúa como activador del *splicing* de subunidades de receptores neurotransmisores inhibidores como GABA_AR γ 2, al reconocer un elemento intrónico favoreciendo la inclusión del exón previo. Sin embargo, se han identificado secuencias de unión de NOVA en el exón E4 de Nova-1 que promueve la inhibición de la inclusión de este exón alternativo. En ambos casos los elementos consenso reconocidos por NOVA están formados por repeticiones de la secuencia YCAY. Sin embargo, la secuencia exónica represora presenta mayoritariamente cisteínas en las posiciones Y, así como en los tramos de separación entre las secuencias repetidas; mientras que en el *enhancer* intrónico se encuentran un mayor número de uridinas en esas posiciones (Dredge et al., 2005).

- **Equilibrio entre los elementos en *trans*.** Determinadas secuencias en el pre-ARNm pueden ser reconocidas por varias proteínas reguladoras a la vez, dando lugar a

complejos que pueden cooperar entre sí para lograr un objetivo común, ya sea el reconocimiento o la inhibición de un exón. Estos factores también pueden competir por la unión a una misma secuencia generando un equilibrio que estará influenciado por el contexto. La ausencia de alguno de estos factores desestabiliza el equilibrio existente y favorece el desarrollo de un evento de *splicing* concreto. Un ejemplo de cooperación son las proteínas SR, tradicionalmente relacionadas con el reconocimiento de secuencias ESE que facilitan el reconocimiento de un exón. En un estudio reciente mediante análisis de CLIP-seq (*cross-linked immunoprecipitation followed by next generation sequencing*), se ha mostrado que SRSF1 y SRSF2 reconocen secuencias solapantes compitiendo por el sitio de unión a un exón, aunque no suelen ejercer funciones redundantes. La ausencia de SRSF2 genera mecanismos compensatorios que tratarían de suplir la falta de la proteína mediante la unión de otros factores como SRSF1 a su sitio de unión. Cuando además se genera un SRSF1 mutante se observa un aumento en la exclusión de exones alternativos probablemente por el incremento de factores represores presentes. Por tanto, el equilibrio de los complejos de proteínas SR o desequilibrio circunstancial determinan el patrón de *splicing* (Pandit et al., 2013).

- **Regulación cruzada entre factores de *splicing*.** Dentro de esta intrincada red de regulación, los factores que forman parte de ella también son capaces de autorregularse. Uno de los ejemplos más claros es el descrito con PTB/nPTB, cuya inducción y silenciamiento es controlada de manera mutua (Boutz et al., 2007).

- **Regulación del *splicing* por el acoplamiento transcripcional.** Como se describirá más detalladamente en apartados posteriores, la comunicación existente entre las maquinarias de transcripción y *splicing*, permite que la tasa de velocidad de la RNAPII influya en el reconocimiento de determinados elementos reguladores del *splicing*.

- **Regulación del *splicing* alternativo por cambios epigenéticos.** Se ha descrito que las modificaciones postraduccionales en las histonas de los nucleosomas y la cromatina reclutan factores de *splicing* que determinan los patrones de procesamiento. Por ejemplo, el gen VEGFA es trimetilado en la lisina 9 de la histona 3 (H3-K9me3) por la metiltransferasa EHMT2 y esta modificación es reconocida por el modulador de la cromatina HP1- γ que recluta al factor de *splicing* SRSF1, generando un cambio en el procesamiento del transcrito (Salton, Voss, and Misteli, 2014).

4. ACOPLAMIENTO DE LA TRANSCRIPCIÓN Y EL *SPLICING*

La similitud entre el rango de tiempo necesario para que un pre-ARNm sea procesado (varios minutos) y el periodo promedio en que se transcribe un gen hacen posible la existencia de un *splicing* co-transcripcional. Se cree que el *splicing* puede empezar cuando la RNAPII ha elongado aproximadamente 5 kb desde el sitio 3'SS, coincidiendo con los datos obtenidos en el primer caso reportado de *splicing* co-transcripcional (Danko et al., 2013; Schmidt et al., 2011). Sin embargo, la primera evidencia de *splicing* co-transcripcional fue observada mediante imágenes de microscopía electrónica en embriones de *D. melanogaster*, donde se apreció la eliminación de intrones en transcritos nacientes que todavía se encontraban unidos al ADN (Beyer and Osheim, 1988). Posteriores trabajos han corroborado este hecho, utilizando por ejemplo sondas que hibridan en la región de unión intrón-exón para visualizar la eliminación de los intrones en lugares próximos a sitios de transcripción (Xing et al., 1993). Además existen elementos propios de cada maquinaria que influye de manera recíproca en el otro proceso. Así, se conocen componentes de la maquinaria de la transcripción como son la estructura del promotor, los factores de elongación transcripcional o los co-activadores específicos de la transcripción que afectan a los patrones de *splicing* (Cramer et al., 1997; Cramer et al., 1999; Nogues et al., 2002; Auboeuf et al., 2004).

El desarrollo de la transcripción y el *splicing* de forma simultánea, aporta ventajas para ambos procesos. Así, se facilita el reclutamiento de factores de *splicing* al lugar de transcripción, optimizando la eficiencia y precisión del procesamiento del transcrito naciente (Listerman, Sapra, and Neugebauer, 2006). Además, el acoplamiento entre la transcripción y el *splicing* añade un nivel más de regulación recíproca entre ambos procesos que se detallará a continuación atendiendo a dos modelos que no son exclusivos (Caceres and Kornblihtt, 2002).

4.1 Modelo cinético

El modelo cinético de la regulación del *splicing* propone que la selección del sitio de *splicing* viene determinada por la tasa de elongación transcripcional de la RNAPII. Este modelo se basa en que los intrones 5' que flanquean a cada exón son

eliminados antes que los intrones 3' puesto que son reconocidos y transcritos antes (Aebi et al., 1986). Como se describió previamente, la elongación de la transcripción es un proceso altamente regulado. Por tanto, los mecanismos que controlan el funcionamiento y paradas de la RNAPII durante la elongación, como la presencia de trans-activadores o la apertura de la cromatina (Sims, Belotserkovskaya, and Reinberg, 2004), están directamente implicados en la determinación de los patrones de procesamiento del ARNm durante el *splicing* co-transcripcional. Durante el transcurso de la elongación de la transcripción los componentes del *spliceosoma* se van uniendo al ARN naciente y de forma ordenada irán eliminando los intrones recién sintetizados. El espacio de tiempo disponible para llevar a cabo este procesamiento y competir por la unión al siguiente intrón es lo que se conoce como “la ventana de la oportunidad” (“*window of opportunity*”) (Perales and Bentley, 2009). Según este modelo, la velocidad de elongación de la RNAPII determina la duración del periodo en el que varios eventos de *splicing* competirán por el reconocimiento y ensamblaje del *spliceosoma*. Por tanto, si el sitio 3'SS débil de un exón alternativo compite con otro sitio 3'SS fuerte posterior, el primero será procesado cuando la tasa de elongación de la RNAPII sea baja y el tiempo sea suficiente para que el exón alternativo sea reconocido e incluido en el transcrito final.

El primer indicio de acoplamiento cinético fue descrito en un estudio donde se demostró que la tasa de elongación de la RNAPII afectaba la estructura secundaria del ARN naciente y como consecuencia al proceso de *splicing* (Eperon et al., 1988). Uno de los sistemas más estudiados es el exón alternativo ED-I del gen de la fibronectina. Este exón está precedido por un sitio 3'SS débil, el cual es procesado e incluido en el ARNm cuando la tasa de elongación de la RNAPII es baja y hay tiempo suficiente para que la maquinaria del *spliceosoma* reconozca esta secuencia (Nogues, Munoz, and Kornblihtt, 2003). En cambio, cuando la procesividad de la RNAPII es alta debido a elementos como la presencia de co-activadores transcripcionales, promotores fuertes o apertura de la cromatina es más probable que sitios 3' de *splicing* fuertes prevalezcan sobre el 3'SS degenerado del exón ED-I y finalmente no sea incluido en el transcrito maduro (Cramer et al., 1999; Kadener et al., 2001; Nogues et al., 2002).

Análisis globales del genoma en levaduras han identificado sitios de acumulación de la RNAPII en las regiones 3'SS de los intrones. Esta pausa constituye un punto de control para que tenga lugar el *splicing* co-transcripcional y parece estar

controlada por la fosforilación del CTD, donde la Ser5 marcaría la pausa de la RNAPII mientras que la Ser2 facilitaría la liberación (Alexander et al., 2010). Aunque estos datos todavía no están claros en metazoos, la existencia de posibles sitios de control que determinan el desarrollo de un *splicing* co-transcripcional, podría estar relacionado con la presencia factores “acopladores” que lleven a cabo esa regulación.

Entre los factores capaces de modular la tasa de elongación de la RNAPII se incluye la ocupación de nucleosomas, ya que podrían constituir barreras que dificulten en determinadas situaciones la actividad transcripcional (Schor et al., 2009; Schwartz, Meshorer, and Ast, 2009; Spies et al., 2009). Por otra parte, la acetilación/desacetilación de histonas influye en el estado de apertura de la cromatina. Un estado abierto de la cromatina debido a la acetilación de histonas, facilita el reclutamiento de factores de transcripción a la RNAPII (Lee, Smith, and Shilatifard, 2010). Asimismo, interacciones con la deacetilasa de histonas HDAC2, genera estados cerrados de cromatina que ralentiza a la RNAPII y favorece la inclusión de exones alternativos (Zhou et al., 2011). También existen otras proteínas capaces de modular la tasa de elongación de la RNAPII al interactuar con componentes de la maquinaria de transcripción, como SRSF2 que se une a P-TEFb o la subunidad Brm del complejo remodelador de cromatina SWI/SNF (también conocido como SMARCA2) que interacciona con Sam68 y provocan la acumulación de RNAPII en regiones exónicas concretas. Este último mecanismo de modulación de la RNAPII está muy relacionado con el *modelo de reclutamiento* que se detallará a continuación.

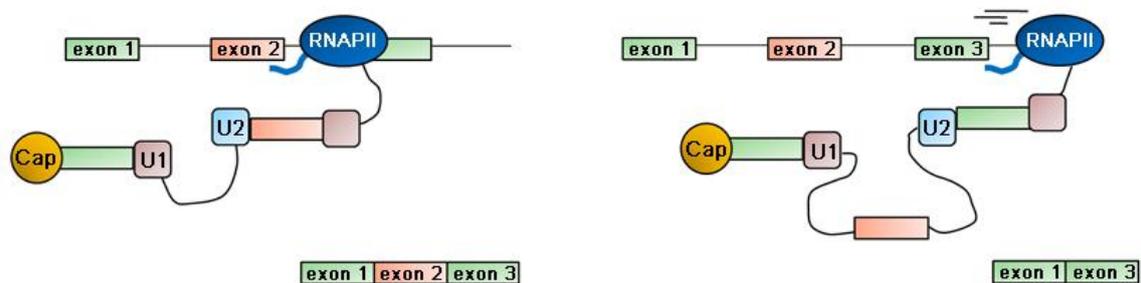


Figura I-5. Modelo de acoplamiento cinético. Una tasa de elongación baja permite el reclutamiento de factores de *splicing* a exones alternativos (exón 2 en rosa, panel izquierdo). Cuando la tasa de elongación es alta, no da tiempo a que el *spliceosoma* se ensamble en exones alternativos (panel derecho).

4.2 Modelo de reclutamiento

El modelo de reclutamiento se basa en que el acoplamiento entre los procesos de transcripción y *splicing* es facilitado por interacciones físicas establecidas entre componentes de ambas maquinarias. El CTD de la RNAPII tiene un papel fundamental en este acoplamiento, ya que sirve como plataforma para el reclutamiento de factores de procesamiento situándolos en las proximidades del ARN mientras se está llevando a cabo el proceso de transcripción.

Uno de los mecanismos más aceptados que explica el proceso de reclutamiento, se basa en los ciclos de fosforilación del CTD que caracteriza a cada una de las etapas de la transcripción (Buratowski, 2003), como ya se ha descrito previamente (apartado 1.3). Dependiendo del estado de fosforilación, se unirán diversos factores implicados en las distintas fases del proceso de transcripción (se detallarán más adelante), facilitando el desarrollo de cada una de ellas (Phatnani and Greenleaf, 2006).

En cuanto al *splicing*, existen numerosas observaciones que apoyan el requerimiento del CTD en el procesamiento del ARNm tanto *in vivo* como *in vitro*. Por ejemplo, la presencia de un CTD truncado reduce la eficiencia del *splicing in vivo* (Rosonina and Blencowe, 2004). La integridad del CTD también es importante para el reclutamiento de factores de *splicing* al sitio de transcripción *in vivo* (Misteli and Spector, 1999). Por otro lado, ensayos *in vitro* indican que la RNAPII en estado hiperfosforilado es capaz de estimular la reacción de *splicing*, mientras que la versión hipofosforilada lo inhibe (Hirose, Tacke, and Manley, 1999). Además se ha descrito que moléculas de CTD recombinantes puede activar el *splicing in vitro*, aunque esto sólo ocurre en el contexto de la definición del exón, cuando los sitios 5'SS y 3'SS están accesibles (Zeng and Berget, 2000).

Aparte del CTD, los factores de *splicing* pueden ser reclutados a sitios activos de transcripción a través de otros mecanismos. La cromatina constituye otra plataforma alternativa para este reclutamiento, ya que se han descrito interacciones directas con componentes de la maquinaria de *splicing* (Piacentini et al., 2009; Sims et al., 2007). Un claro ejemplo es el que se ha observado con la proteína de unión a PTB, la cual es reclutada al sitio de transcripción mediante su unión a la proteína de unión a la cromatina MRG15, que reconoce de manera específica la trimetilación en la lisina 36 en

la histona 3 (H3-K36me3). El reclutamiento de PTB mediado por MRG13 da lugar a una exclusión de exones alternativos (Luco et al., 2010).

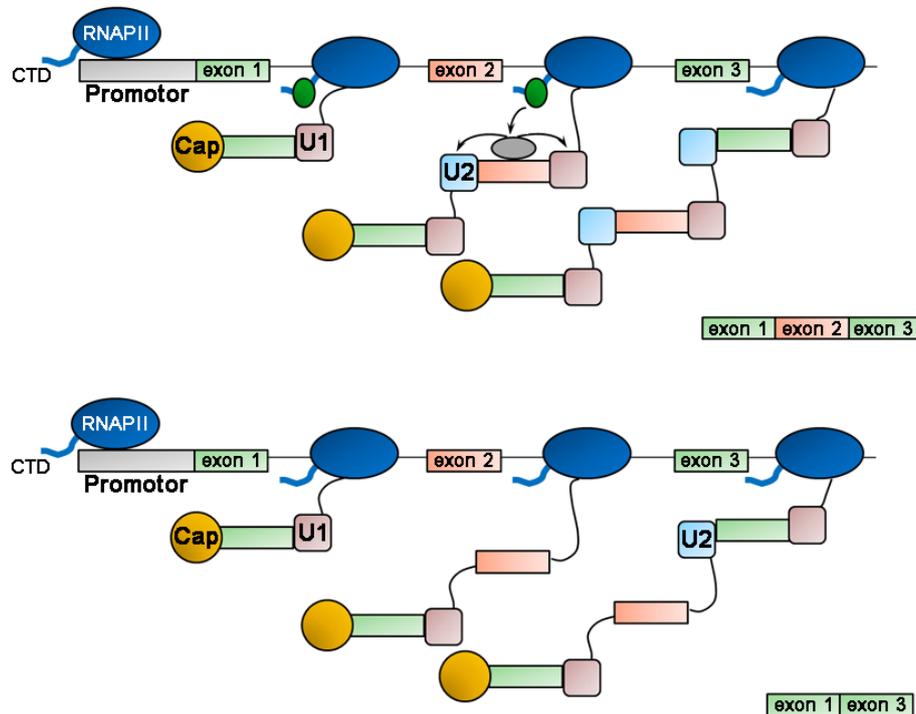


Figura I-6. Modelo de acoplamiento por reclutamiento. El CTD de la RNAPII (parte superior de la figura) sirve como plataforma para el reclutamiento de factores acopladores (verde) que directamente o a través de interacciones con otros factores (lila) favorecen el reconocimiento de exones alternativos (rosa). Cuando por determinadas circunstancias estos factores no son reclutados, los exones alternativos no son reconocidos y no se incluyen en el transcrito final (parte inferior de la figura).

4.2.1 Factores acopladores de la transcripción y el *splicing*

Tanto el modelo de reclutamiento e indirectamente el modelo cinético, apoyan la teoría de la existencia de un grupo de potenciales factores de acoplamiento, que comunican a las maquinarias de transcripción y *splicing*, facilitando el progreso del *splicing* co-transcripcional. Entre ellas podemos destacar las siguientes:

- **Proteínas SR.** Las proteínas SR, como ya se describió anteriormente, son un amplio grupo de factores relacionados con la regulación del *splicing*, pero que también se han relacionado con otros procesos relacionados con el metabolismo del ARNm. Se han descrito interacciones entre proteínas SR y el CTD de la RNAPII, aunque la relevancia de esta asociación aún está por determinar (Yuryev et al., 1996). Además, parece que esta unión podría estar relacionada con el reclutamiento de las proteínas SR

desde los *speckles* nucleares a los sitios activos de transcripción (Misteli and Spector, 1999). Además en sistemas de *splicing in vitro*, las proteínas SR podrían ser necesarias para el reclutamiento y ensamblaje del *spliceosoma* temprano en los sitios del ARNm naciente (Das et al., 2007). Asimismo, proteínas SR como SRp20 o SRSF1 dependen de la presencia del CTD de la RNAPII así como de su estado de fosforilación para llevar a cabo su función en el *splicing* alternativo de minigenes (Barboric et al., 2009; de la Mata and Kornblihtt, 2006). Es interesante señalar que un estudio reciente relaciona a SRSF1 con la activación de la transcripción. SRSF1 reconocería a la secuencia TAR del sitio de transcripción del gen del *HIV-1* y ayudaría al reclutamiento del complejo P-TEFb facilitando la transcripción durante los primeros estadios de la infección viral (Paz, Krainer, and Caputi, 2014). Como ya se describió en el modelo cinético, el factor de *splicing* SRSF2 participaría en la elongación de la transcripción al facilitar el reclutamiento del factor de transcripción P-TEFb al sitio de transcripción favoreciendo la procesividad de la RNAPII (Lin et al., 2008). Aunque el mecanismo por el cual las proteínas SR establecen estas interacciones importantes para el acoplamiento no está del todo claro, este hecho las relaciona con el ensamblaje del *spliceosoma* y la regulación de la elongación transcripcional en un contexto de *splicing* co-transcripcional.

- **SKIP.** Otros factores de *splicing* también participan en el acoplamiento entre los procesos de transcripción y *splicing*. El factor de *splicing* SKIP interacciona con el factor de transcripción P-TEFb y favorece la elongación transcripcional mediada por el activador Tat en el gen del *HIV-1* (Bres et al., 2005).

- **Tat-SF1.** A pesar de no tener una función clara en la célula, se han descrito diversas asociaciones del factor Tat-SF1 con componentes de la maquinaria de transcripción y *splicing*, por lo tanto podría estar conectando ambos procesos. Entre estas interacciones se encuentran la de Tat-SF1 con P-TEFb que a su vez recluta a factores de *splicing* al pre-ARNm naciente. Estudios *in vitro* han mostrado que la maquinaria de *splicing* reclutada por Tat-SF1 estimula la elongación transcripcional (Fong and Zhou, 2001; Kim et al., 1999; Maniatis and Reed, 2002). Sin embargo, un estudio más reciente del análisis genómico del silenciamiento de Tat-SF1, indica que este factor podría actuar en la célula como un activador transcripcional e intervenir en la regulación del *splicing*; sin embargo, solo un porcentaje muy pequeño de genes coincide entre los grupos de genes afectados a nivel de transcripción y *splicing*, poniendo en

duda su posible papel como acoplador general entre ambos procesos (Miller et al., 2011).

- **PSF y p54nrb/NonO.** Los factores PSF y p54nrb/NonO han sido implicados en procesos de transcripción y *splicing*. Estas dos proteínas están relacionadas entre sí, ya que comparten el 71% de identidad en sus dominios RRM (Dong et al., 1993). Ambas proteínas interactúan con el CTD hipofosforilado e hiperfosforilado de la RNAPII (Emili et al., 2002). Se ha mostrado la presencia de PSF y p54/NonO en complejos de alto peso molecular donde también se encuentran los componentes del *spliceosoma* U1 y U2 snRNPs, así como la RNAPII (Kameoka, Duque, and Konarska, 2004; Peng et al., 2006). Experimentos *in vivo* han mostrado que PSF es capaz de estimular la activación de la transcripción a través de su unión con el CTD de la RNAPII (Rosonina et al., 2005).

- **Complejo Mediador.** El complejo Mediador es un importante co-activador de la transcripción, además de tener un papel muy relevante como integrador de la transcripción por las numerosas interacciones que ejerce con otros activadores y represores transcripcionales (Malik and Roeder, 2010). Se ha descrito que la subunidad Med23, y no otras de este complejo, interactúa con el factor de *splicing* hnRNP L, alterando los patrones de *splicing* alternativo (Huang et al., 2012).

- **Proteínas con dominios WW y FF.** Establecen interacciones con factores implicados en los procesos de transcripción y *splicing* a través de sus dominios y son descritas con más detalle a continuación.

4.2.2 Proteínas con dominios WW y FF

Existe un conjunto de factores de particular interés para esta tesis doctoral, que han sido vinculados a procesos de transcripción y *splicing* por establecer interacciones físicas con componentes de ambas maquinarias, como se describirá a continuación. La mayoría de estas uniones se establecen a través de los dominios WW y FF que contienen en repeticiones en tándem en su estructura, sugiriendo que estos dominios podrían colaborar en la unión de diversos factores, aportando una mayor eficiencia en la interacción.

- **Dominios WW.** Son motivos de interacción pequeños, compuestos por aproximadamente 40 aminoácidos que se caracterizan por contener dos residuos de triptófano altamente conservados, separados entre sí por 20-22 aminoácidos (Sudol, 1996). Establecen interacciones de tipo proteína-proteína al reconocer secuencias peptídicas ricas en prolinas y sitios de serina/treonina fosforilados, presentes en sus ligandos (Lu et al., 1999). Su estructura se pliega de manera estable en forma de triple hélice antiparalela con estructura de lámina β en ausencia del ligando (Macias et al., 1996). Los dominios WW se encuentran en proteínas de diverso origen y son capaces de establecer gran variedad de interacciones, incluso un mismo dominio puede unirse a varias proteínas diferentes, lo que los convierte en grandes soportes de conexión entre diferentes procesos celulares (Ingham et al., 2005).

- **Dominios FF.** Son motivos estructurales de aproximadamente 50-60 aminoácidos, caracterizados por contener dos residuos de fenilalanina muy conservados situados cerca de sus extremos amino y carboxilo-terminal. Establecen interacciones del tipo proteína-proteína con distintas especificidades en ligandos con secuencias preferentemente fosforiladas. Su estructura está basada en tres hélices α y una hélice 3_{10} situados de forma ortogonal entre sí en el orden $\alpha 1-\alpha 2-3_{10}-\alpha 3$ (Allen et al., 2002). Aunque la función de los dominios FF no está muy clara, están presentes de forma bastante exclusiva en dos grupos de proteínas con funciones diferentes: 1) proteínas con dominios WW en su estructura, principalmente nucleares y relacionadas con procesos del metabolismo del ARN; 2) familia de proteínas p190 de las Rho/GTPasas, de localización citoplasmática e implicadas en la regulación de las vías de señalización del citoesqueleto (Bedford and Leder, 1999).

Prp40

La proteína de levaduras Prp40 fue caracterizada como un factor esencial de *splicing* asociada a la subunidad U1 snRNP del *spliceosoma*, implicada en el reconocimiento del sitio 5' de *splicing* en las primeras etapas de ensamblaje del *spliceosoma* (Kao and Siliciano, 1996). Prp40 también ha sido implicada en importantes interacciones que comunican ambos extremos del intrón, al unirse directamente con BBP, el ortólogo en levaduras del factor humano de *splicing* SF1, que a su vez se une a Mud2p, el homólogo de levaduras de U2AF⁶⁵. Además se ha demostrado que Prp40 es

capaz de asociarse con Prp8, proteína asociada a la subunidad U5 snRNP, siendo involucrada en el reclutamiento de factores durante el ensamblaje del *spliceosoma* (Abovich and Rosbash, 1997).

Prp40 es una proteína de 583 aminoácidos cuya estructura se caracteriza por contener dos dominios WW en la región amino terminal y cuatro dominios FF en tándem en la mitad carboxilo terminal. Las interacciones funcionales que definen el papel de la proteína se establecen a través de estos dominios. Por ejemplo, se ha sugerido que Prp40 se asocia con BBP al reconocer en su secuencia el motivo de prolina PPXY. Sin embargo, existe cierta controversia sobre el papel que pueden desempeñar los dominios WW en la definición del intrón, ya que un estudio ha mostrado que la ausencia de los motivos WW de Prp40 no afecta al reclutamiento de U1 y U2 snRNPs *in vivo* e *in vitro* (Gornemann et al., 2011). Una posible explicación sería que la participación de Prp40 a través de sus dominios WW en el reclutamiento durante las primeras etapas de ensamblaje del *spliceosoma* ocurre sólo en ciertos genes en un contexto específico. Por otra parte, Prp40 interactúa *in vitro* con el CTD fosforilado de la RNAPII a través de sus dominios WW y FF (Morris and Greenleaf, 2000). El hecho de que Prp40 sea capaz de establecer uniones con componentes de las maquinarias de transcripción y *splicing* lo hace un buen candidato para desempeñar una función de acoplador entre ambos procesos.

Prp40 tiene dos ortólogos en humanos: PRPF40A (nombre aceptado por la base de datos HUGO: *PRP40 pre-mRNA processing factor 40 homolog A*) también conocido como HYPA (*Huntigtin-Yeast Partner A*) o FBP11 (*Formin-binding protein 11*) y PRPF40B (nombre aceptado por la base de datos HUGO: *PRP40 pre-mRNA processing factor 40 homolog B*) o HYPC (*Huntigtin-Yeast Partner C*).

PRPF40A

PRPF40A, FBP11 o HYPA es una proteína de 957 aminoácidos cuya estructura contiene dos dominios WW en su región amino-terminal y cinco dominios FF en la parte carboxilo-terminal. PRPF40A fue identificada en ratón como una proteína de unión a la formina (FBP, *formin binding proteins*), proteínas implicada en el desarrollo de las extremidades y riñón murino (Chan, Bedford, and Leder, 1996). Posteriormente,

mediante ensayos de interacción *in vitro* con una librería de proteínas procedentes de muestras de extremidades embrionarias de ratón, se mostró que los dominios WW de FBP11 reconocían la secuencia consenso PPLP en sus ligandos. El hecho de que varios de los ligandos asociados a PRPF40A fueran nucleares, como MeCP2, una proteína de unión al ADN, la relacionaban con procesos metabólicos que tienen lugar en el núcleo (Bedford, Chan, and Leder, 1997). Consistente con su función nuclear, PRPF40A ha sido localizada en los *speckles* nucleares y se ha publicado que inhibe a la proteína N-WASP, implicada en la reorganización del citoesqueleto (Mizutani, Suetsugu, and Takenawa, 2004). Ensayos de interacción *in vitro* mediante espectrometría de masas mostraron la unión de PRPF40A a través de sus dominios WW con los factores de *splicing* SF1, U2AF⁶⁵, proteínas de la partícula U2 snRNP, algunos hnRNPs y factores remodeladores de la cromatina, entre otros. Estos datos vinculan a PRPF40A con el procesamiento del ARNm y más concretamente con el reconocimiento del sitio 3' de *splicing* (Lin, Lu, and Tarn, 2004). Ensayos más recientes de inmunoprecipitación utilizando extractos nucleares de células HeLa han mostrado que, a diferencia de su homólogo Prp40 en levaduras que forma parte de la partícula U1 snRNP, PRPF40A estaría asociado a la subunidad U2 snRNP del *spliceosoma* y proteínas relacionadas (Makarov et al., 2012). Por otra parte, una caracterización estructural de los dominios FF revelaron que PRPF40A, probablemente mediante su dominio FF1, podría reconocer la secuencia fosforilada del CTD de la RNAPII (Allen et al., 2002). Este dato contrasta parcialmente con el hecho de que el dominio FF1 de Prp40 no reconoce al CTD de la RNAPII, sino su dominio FF2, indicando que no todos los dominios FF son equivalentes y que la especificidad de estos motivos no están conservados entre ortólogos (Gasch et al., 2006). Este conjunto de datos relacionan a PRPF40A con los procesos de *splicing* y transcripción, acorde con un posible papel de acoplador de ambos procesos.

Otras interacciones de PRPF40A a través de sus dominios WW la han implicado en enfermedades neurodegenerativas. Es el caso de la enfermedad de Huntington, la cual se caracteriza por la expansión de una cola de poliglutaminas aberrante en el extremo carboxilo-terminal de la proteína huntingtina. La secuencia de esta proteína contiene un motivo de prolina que reconocen los dominios WW de PRPF40A (Faber et al., 1998). La interacción de PRPF40A con la huntingtina mutada no sólo cambia la

distribución de la primera, sino que podría estar alterando su función relacionada con el *splicing* (Jiang et al., 2011).

PRPF40B

PRPF40B o HYPC es el ortólogo menos conocido del factor de levadura Prp40. Es una proteína de 871 aminoácidos que contiene dos dominios WW y varios dominios FF en su estructura que no han sido previamente caracterizados. Por esa organización estructural y a pesar de la escasa información que existe sobre ella, se la ha relacionado con procesos de transcripción y *splicing*. Fue identificada inicialmente en un estudio de interacciones *in vitro* para identificar ligandos de la huntingtina, con el objetivo de detectar posibles dianas que pudieran estar alteradas en la enfermedad de Huntington. Al igual que ocurrió con PRPF40A, PRPF40B destacó en este ensayo por el reconocimiento de un motivo de prolinas a través de los dominios WW, mostrando PRPF40B mayor fuerza de interacción que PRPF40A (Faber et al., 1998).

Al igual que con PRPF40A, se ha sugerido que el aumento de la interacción de manera aberrante con la huntingtina mutada podría ser un desencadenante de la enfermedad. Aunque ambas proteínas guardan una gran similitud de secuencia de aminoácidos, PRPF40A tiene mayores niveles de expresión que PRPF40B (Passani et al., 2000). PRPF40B también ha sido relacionada con el síndrome de Rett, un desorden neurológico asociado al cromosoma X causado por mutación y consecuente pérdida de función de MeCP2, una proteína de unión a islas CpG del ADN metiladas. Ensayos de interacción *in vitro*, han mostrado la interacción directa de PRPF40B a través de sus dominios WW con MeCP2 al reconocer en ésta última el motivo de prolinas PPLP. Esta secuencia coincide con el fragmento donde se concentra un gran número de mutaciones asociadas al síndrome. De hecho, ensayos de coimmunoprecipitación de MeCP2 con deleciones en la región de poliprolinas y los dominios WW de PRPF40B, revelaron una disminución en la especificidad de interacción. Estos datos sugieren que esta asociación podría ser importante para el desarrollo del fenotipo del síndrome de Rett (Buschdorf and Stratling, 2004).

En un estudio más reciente, se ha relacionado a PRPF40B con el síndrome mielodisplásico (MDS), un conjunto de desórdenes hematológicos con una producción desregulada de producción de glóbulos rojos que muestra una predisposición a leucemia

mieloide aguda (AML). Un estudio de secuenciación del exoma completo de pacientes con MDS, ha revelado un conjunto de mutaciones asociadas a componentes del *spliceosoma*, implicados principalmente en el reconocimiento del sitio 3'SS, lo que podría estar alterando patrones de procesamiento. Entre los genes afectados se ha encontrado a PRPF40B. En este factor se detectaron seis mutaciones que generan un cambio de aminoácido y aunque tres de ellas parecen SNPs, las otras tres mutaciones se han relacionado con predisposición a enfermedad mediante estudios bioinformáticos, por lo que podrían estar alterando la función de la proteína con consecuencias patológicas para la enfermedad (Yoshida et al., 2011).

FBP21

FBP21 es una proteína de 376 aminoácidos con dos dominios WW en su estructura. A pesar de no tener dominios FF, ha sido relacionada con este conjunto de factores porque sus motivos WW, al igual que el resto de proteínas del grupo, contienen tres residuos de tirosina centrales y una separación entre los dominios WW muy parecida, que no tienen otras proteínas con estos motivos. Además FBP21 ha sido relacionada con el proceso de *splicing*, ya que se ha mostrado su interacción con las proteínas U1C de la subunidad U1 snRNP, así como con los factores de *splicing* SmB, SmB' y también con SF1, al reconocer el motivo PGM a través de sus dominios WW. Además, ensayos de inmunoprecipitación han mostrado la asociación de FBP21 a la partícula U2 snRNP. Asimismo, ensayos de inmunofluorescencia han mostrado la localización de FBP21 en los *speckles* nucleares, orgánulos relacionados con el procesamiento del ARNm (Bedford, Reed, and Leder, 1998). Consistente con este último dato, un estudio más reciente ha revelado la interacción de FBP21 con el factor de *splicing* SF3B4 en los *speckles* nucleares (Klippel et al., 2011). Además, los dominios WW de FBP21 parecen ejercer importantes interacciones funcionales, ya que a través de ellos es capaz de activar el *splicing in vivo* al unirse y activar a la proteína SIPP1, un factor de *splicing* que se asocia a la proteína de unión al tracto de poliglutaminas (PQBP1) y a la proteína fosfatasa 1 (PP1) (Huang et al., 2009).

Todos estos datos sugieren que FBP21 participa en el procesamiento del ARNm conectando ambos extremos del intrón y reclutando factores, afectando en última instancia al desarrollo de la reacción de *splicing*.

TCERG1

TCERG1 (nombre aceptado por HUGO: *Transcription Elongation Regulator 1*), también conocido como CA150, es una proteína de 1098 aminoácidos con tres dominios WW en su región amino-terminal y seis dominios FF en la parte carboxilo-terminal. Fue inicialmente purificada en fracciones nucleares de células HeLa procedentes de una cromatografía de afinidad a Tat e identificada como co-factor de la trans-activación mediada por Tat del gen del *HIV-1* (Sune et al., 1997). Se han descrito numerosas interacciones a través de sus dominios WW como los factores de *splicing* SF1, U2AF⁶⁵, proteínas del U2 snRNP y también mediante los dominios FF como Tat-SF1 o proteínas SR (Carty et al., 2000; Goldstrohm et al., 2001; Lin, Lu, and Tarn, 2004; Sanchez-Alvarez et al., 2006; Smith, Kulkarni, and Pawson, 2004). Los dominios FF4 y FF5 son también muy importantes para la localización de TCERG1 en la periferia de los *speckles* nucleares (Sanchez-Hernandez et al., 2012).

TCERG1 es capaz de regular la activación transcripcional de genes reporteros (Sanchez-Alvarez et al., 2010; Sune and Garcia-Blanco, 1999). Además se ha mostrado que TCERG1 regula la expresión del gen de apoptosis *Bcl-x* modulando la tasa de elongación de la RNAPII (Montes et al., 2011). Estos datos han sido corroborados en un estudio más reciente donde TCERG1 regula el patrón de expresión del gen del *HIV-1* al incrementar la tasa de elongación de la RNAPII mediante la fosforilación en la Ser2 del CTD (Coiras et al., 2013). Asimismo, numerosos estudios muestran que TCERG1 es capaz de regular el *splicing*. Se han descrito diversos trabajos donde TCERG1 modula el *splicing* de minigenes reporteros de la β -globina, β -tropomiosina, *CD44*, *fibronectina* y *Bcl-x* (Cheng et al., 2007; Lin, Lu, and Tarn, 2004; Montes et al., 2011; Pearson et al., 2008; Sanchez-Alvarez et al., 2010). Además, la identificación de genes diana mediante *microarrays* revelan que al silenciar TCERG1 están afectados genes implicados en el procesamiento del pre-ARNm (Pearson et al., 2008) (datos del laboratorio sin publicar).

TCERG1 ha sido también implicada en la enfermedad de Huntington, al interaccionar con sus dominios WW con la región amino-terminal de la proteína huntingtina (Holbert et al., 2001) indicando que en el contexto de la enfermedad, esta asociación puede alterar la función de TCERG1 y la restauración de la proteína podría general un efecto neuroprotector (Arango et al., 2006).

5. EL NÚCLEO DE LA CÉLULA EUCARIOTA

La estructura de la célula eucariota se caracteriza por tener dos compartimentos claramente definidos: el citoplasma y el núcleo. A su vez cada uno contienen diversos orgánulos, pero a diferencia de lo que ocurre en el citoplasma, los compartimentos nucleares no están delimitados por membranas, indicando diferentes mecanismos de funcionamiento. A pesar de la ausencia de membranas, el núcleo se encuentra altamente organizado en compartimentos definidos, existiendo una gran dinámica entre ellos. En el núcleo existen diversos orgánulos nucleares distribuidos en el escaso espacio intercromatínico que existe, ya que la mayor parte se encuentra ocupado por la cromatina, en territorios cromosómicos. Por este motivo, una de las principales características del núcleo es que contiene el material genético y además es donde tienen lugar procesos esenciales para la célula, como la replicación, reparación del ADN, transcripción y procesamiento del pre-ARNm. Por tanto, los orgánulos nucleares participan de una forma directa o indirecta en el desarrollo de importantes procesos biológicos.

Entre estos orgánulos nucleares se encuentran: el *nucléolo*, implicado en la biogénesis de los ribosomas, así como el ciclo celular y la respuesta al estrés celular; los *cuerpos de Cajal*, implicados en la modificación y ensamblaje de los U snRNPs y partículas ribonucleoproteínas pequeñas nucleolares (snoRNPs) implicadas en el procesamiento del ARNr; *Gems*, estructuras adyacentes o coincidentes con los cuerpos de Cajal caracterizados por la presencia de las proteínas SMN y *Gemins2-7* y relacionados con el ensamblaje de los snRNPs; *PML*, implicado en el ensamblaje de la proteína supresora de tumores leucemia promielocítica (PML), además de otras funciones celulares como el secuestro y liberación de proteínas, modificaciones postraduccionales y respuesta al estrés celular; *paraspeckles*, son importantes para el control de la expresión génica mediante la retención de ARN de doble cadena sujetos a la edición A-I y han sido involucrados procesos de diferenciación, infección viral o respuesta a estrés; *speckles*, como se detallará a continuación, son estructuras nucleares relacionadas con el procesamiento del ARNm (Bernardi and Pandolfi, 2007; Emmott and Hiscox, 2009; Fox and Lamond, 2010; Gubitzi, Feng, and Dreyfuss, 2004; Morris, 2008).

5.1 Los *speckles* nucleares

Los *speckles* nucleares fueron identificados por primera vez mediante microscopía electrónica como partículas intercromatínicas situadas en el núcleo en forma de nubes, cuya composición se basaba en ARN (Swift, 1959). Sin embargo, su vinculación con el *splicing* del pre-ARNm se produjo posteriormente al examinar el patrón nuclear de los snRNPs en las proximidades de estos orgánulos (Spector, Schrier, and Busch, 1983). Su nombre “*speckle*” se debe al aspecto punteado que presenta, aunque por su relación con el *splicing* del ARNm también se les conoce como “compartimentos de factores de *splicing*” o “dominios SC35” debido a que el marcaje de la célula con el anticuerpo SC35 (ahora SFSR2) coincide casi exclusivamente con estos orgánulos nucleares (Phair and Misteli, 2000; Wansink et al., 1993). Aunque otras estructuras nucleares pueden contener factores de *splicing*, se consideran *speckles* a los gránulos intercromatínicos (IGC, *interchromatin granule cluster*) que presentan un enriquecimiento de factores de *splicing*. Los IGCs son estructuras de naturaleza ribonucleoproteica (20-50 por célula) de 20-25 nm localizadas en el espacio intercromatínico, rodeados de fibras pericromatínicas (PFs) de 3-5 nm, donde se sitúan los transcritos nacientes (Cmarko et al., 1999; Fakan, 1994).

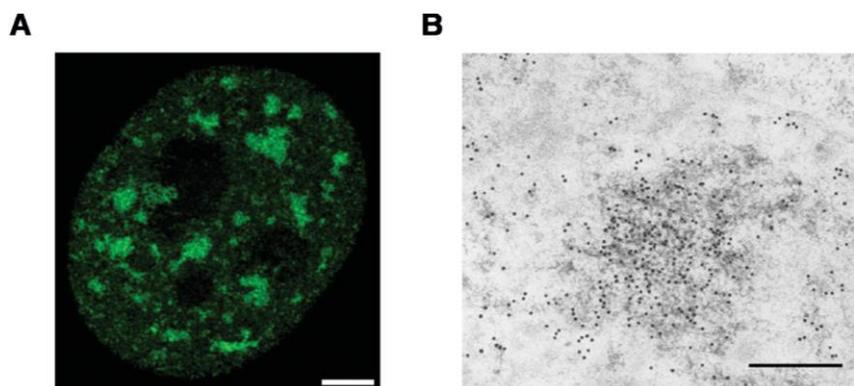


Figura I-8. Los *speckles* nucleares. **A.** Imagen de microscopía de alta resolución de una célula HeLa expresando la proteína de fusión YFP-SRSF2 donde se observa la distribución y morfología de los *speckles* nucleares. Barra = 2 μ m. Imagen tomada de (Spector and Lamond, 2011). **B.** Imagen de microscopía inmunoeléctronica donde se observan los gránulos intercromatínicos compuestos por partículas de 20-25 nm de diámetro conectadas por las fibrillas pericromatínicas. Barra = 500 nm. Imagen tomada de (Spector and Lamond, 2011).

5.1.1 Función de los *speckles* nucleares

La función mayormente aceptada de los *speckles* nucleares es la de compartimento de almacenamiento, ensamblaje y modificación de factores de *splicing* para ser suplementados a los sitios de activación de la transcripción (Lamond and Spector, 2003). Este reclutamiento parece depender de la integridad del CTD de la RNAPII, indicando la importancia de la elongación transcripcional en la incorporación de factores de procesamiento (Misteli and Spector, 1999).

A pesar de la ausencia de actividad transcripcional en los *speckles* nucleares, la proximidad a sitios activos de transcripción sugiere una función en la regulación de la expresión génica. Además, se han localizado preferentemente algunos genes específicos en los *speckles* nucleares (Brown et al., 2008; Moen et al., 2004). También se ha observado con mayor frecuencia la presencia de regiones cromosomales ricas en genes (bandas-R) cerca de los *speckles*, que regiones con pocos genes (bandas G), indicando que los *speckles* podrían actuar como centros organizativos para aproximar genes activos y facilitar la expresión génica (Shopland et al., 2003).

Como ya se mencionó previamente, en la periferia de los *speckles* nucleares, concretamente en las PFs se sitúa el pre-ARNm recién sintetizado. Por tanto, es probable que en estas PFs tenga lugar el *splicing* co-transcripcional. Sin embargo, hay intrones que son procesados de forma postranscripcional, como ocurre con el intrón 26 del pre-ARNm del gen *COL1A1*. Se ha mostrado que tras la mutación de sitios de *splicing* en este intrón se aumenta la acumulación de este transcrito en los *speckles* nucleares (Johnson et al., 2000), indicando que estos compartimentos pueden estar ejerciendo un control postranscripcional o de postmaduración relacionado con la supervivencia del ARN o con su exporte (Lawrence and Clemson, 2008).

Los *speckles* nucleares pueden también actuar como centros de secuestro e inhibición para controlar la dinámica de transcripción y factores de procesamiento. Este es el caso del factor de *splicing* SRSF1, el cual es reclutado a los *speckles* bajo condiciones de estrés a través de un mecanismo dependiente de la interacción de la quinasa SRPK con las chaperonas Hsp70 y Hsp90 (Zhong et al., 2009). De forma parecida, el ARN no codificante largo (lncRNA) MALAT1 (*metastasis-associated lung adenocarcinoma transcript 1*) está directamente vinculado con el reclutamiento de

proteínas SR desde los *speckles* nucleares al sitio de transcripción (Bernard et al., 2010), así como con la modulación del *splicing* alternativo mediante la alteración de los patrones de fosforilación de las proteínas SR (Tripathi et al., 2010).

5.1.2 Composición de los *speckles* nucleares

Los *speckles* nucleares están formados principalmente por factores de *splicing*, incluyendo snRNP y proteínas SR (Fu 1995), que suponen más del 50% de la composición proteica. También están presentes otros factores implicados en procesos de transcripción, poliadenilación y traducción, entre otros procesos relacionados con el metabolismo del ARNm (Saitoh, Spahr, 2004). Además, se ha descrito la presencia de numerosas quinasas como Clk/STY, PSKHI o DYRK1A (Alvarez, Estivill, and de la Luna, 2003; Brede, Solheim, and Prydz, 2002; Colwill et al., 1996), y fosfatasas como PP1 (Trinkle-Mulcahy et al., 1999), que fosforilan y desfosforilan los componentes de la maquinaria de *splicing*. Corroborando la función de los *speckles* como compartimentos de almacenaje y modificación de factores de *splicing*.

A pesar de la ausencia de actividad transcripcional en los *speckles* nucleares (Cmarko et al., 1999), experimentos de microscopia de fluorescencia han mostrado la presencia de una población de RNAPII fosforilada en la Ser2 del CTD, implicada en la elongación de la transcripción (Mortillaro et al., 1996). Otro factor implicado en la elongación de transcripción localizado en los *speckles* nucleares ha sido P-TEFb, formado por la ciclina T1 y la quinasa Cdk9. Ambas se encuentran distribuidas de forma difusa por el citoplasma excepto en los nucléolos y con un enriquecimiento en los *speckles* nucleares, principalmente la ciclina T1 (Herrmann and Mancini, 2001).

Además de estas proteínas de procesamiento y transcripción, en los *speckles* nucleares existe una población de moléculas de ARNm poly(A), que al contrario de lo que cabría esperar, no es transportada al citoplasma en condiciones de bloqueo de la transcripción con α -amanitina. Esto sugiere que este grupo de ARN podría participar en procesos de regulación en el núcleo, como el transporte del ARN hacia el citoplasma (Huang et al., 1994). Consistente con el papel regulador del ARN en los *speckles* nucleares, se ha observado un enriquecimiento en los *speckles* nucleares del ya mencionado lncRNA, MALAT1, cuya función reguladora de las proteínas SR se describió en el apartado anterior.

Además de elementos reguladores, en los *speckles* nucleares aparecen otras proteínas que podrían desempeñar una función estructural. Aunque todavía no se conoce con detalle cómo se organiza estructuralmente un *speckle*, se han descrito varias proteínas que podrían estar implicados dicha función, como la laminina A (Jagatheesan et al., 1999) y la actina asociada a los snRNPs (Nakayasu and Ueda, 1984). Además se han encontrado en los *speckles* nucleares fosfatidil-inositol-(4,5)-bisfosfato (Zhao et al., 1998), un lípido que regula la unión de proteínas a la actina, así como varias isoformas de la quinasa fosfatidil-inositol fosfato (PIP2K) (Boronenkov et al., 1998). Un trabajo más reciente muestra que Son, una proteína con una región que presenta repeticiones de serinas en tándem y un dominio RS, podría estar implicada en la hipotética estructura de los *speckles* nucleares. Experimentos de silenciamiento de Son apoyan su papel en la sujeción de esta estructura (Sharma et al., 2010). A pesar de la presencia de este grupo de proteínas en los *speckles* nucleares con una clara función estructural, existen diversos trabajos que comprometen dicho papel. Por ejemplo, se ha mostrado que la ausencia de Laminina A/C en las células no altera la formación de los *speckles* (Vecerova et al., 2004).

5.1.3 Dominios de señalización hacia los *speckles* nucleares

La localización de una proteína en los *speckles* nucleares puede explicarse por dos modelos. El mecanismo de localización directa, explica que la presencia de una señal de localización en una proteína es suficiente para establecer interacciones que le permiten distribuirse correctamente; en el mecanismo indirecto, las proteínas carecen de este tipo de señal y necesitan interactuar con otros factores que contienen estas señales de localización para poder dirigirse a los *speckles* nucleares (Caceres et al., 1997). Se han identificado varias señales capaces de dirigir a una proteína hacia los *speckles* nucleares.

- **Dominios RS.** Estos dominios ricos en argininas y serinas, que caracterizan a las proteínas SR, han mostrado ser necesarios y suficientes para dirigir a varios de estos factores de *splicing* hacia los *speckles* nucleares, como SRp20, SRSF2 y SRSF7 (anteriormente conocido como 9G8) (Caceres et al., 1997). Además de estos factores, el dominio RS presente en la quinasa CrkRS también es responsable de la localización en los *speckles* nucleares de esta proteína (Ko, Kelly, and Pines, 2001).

- **Dominios RRM.** Los dominios de unión a ARN presentes en SRSF1 o SRp40 son capaces de dirigir a las proteínas a los *speckles* nucleares junto con el dominio SR que funciona como señal de localización nuclear (Caceres et al., 1997). Este dominio presente en la proteína PSF, es el responsable de su localización en los *speckles*. Sin embargo, la fusión de este motivo a otras proteínas heterólogas no las dirige a estos compartimentos nucleares, sugiriendo un mecanismo indirecto de localización (Dye and Patton, 2001).

- **Tracto de histidinas.** En un estudio genómico a gran escala se han identificado 86 proteínas que contienen un tracto homopolimérico de histidinas. Dentro de este grupo, 15 de los 22 factores nucleares se localizan en los *speckles*, y esa secuencia es necesaria para dicha distribución. La variación de la carga en función del pH de estos trectos de histidinas le confiere propiedades óptimas para interaccionar con otras proteínas en función del medio en que se encuentren (Salichs et al., 2009). Un ejemplo es la quinasa DYRK1A, cuyo extremo carboxilo contiene una secuencia rica en histidinas necesaria y suficiente para dirigir por sí misma a la proteína hacia los *speckles* nucleares, que actúa de forma independiente de la señal NLS (Alvarez, Estivill, and de la Luna, 2003).

- **Dominios FF.** Los dominios FF4 y FF5 de la proteína TCERG1 han mostrado ser necesarios y suficientes para la correcta localización de la proteína en los *speckles* nucleares. Las propiedades similares en cuanto a estructura secundaria y la estabilidad de los dos dominios cuando se expresan de forma conjunta y no con otros dominios FF parece ser relevante para dirigir a la proteína a los *speckles* nucleares (Sanchez-Hernandez et al., 2012).

- **Dominios TP.** La presencia de repeticiones de treonina-prolina en el factor SF3b¹⁵⁵, componente del *spliceosoma*, es suficiente para dirigir a esta proteína a los *speckles* (Eilbracht and Schmidt-Zachmann, 2001).

- **Dominio “forkhead-associated”.** Este dominio se encuentra en la subunidad nuclear de la proteína fosfatasa 1 (NIPP1). Aunque se trata de un motivo independiente de la señal de localización nuclear, es necesario para el enriquecimiento de esta proteína en los *speckles* nucleares (Jagiello et al., 2000).

- **Región rica en arginina, prolina y serina.** El comportamiento de las señales de localización de la proteína SRm160 es diferente al resto de proteínas SR. Carente de dominios de unión a ARN, los dominios RS que contiene son cortos y se encuentran dispersos a lo largo de la secuencia, por lo que no son suficientes para su localización nuclear o su localización en los *speckles*. Parece que la secuencia rica en arginina, prolina y serina facilita la correcta localización de esta proteína en los *speckles* nucleares. Resulta interesante la distribución de la SRm160 alrededor de la matriz nuclear al deletar la región de argininas, prolinas y serinas (Wagner, Chiosea, and Nickerson, 2003).

Como ya se comentó previamente, las modificaciones postraduccionales son un importante mecanismo de regulación de la distribución de proteínas situadas en los *speckles* nucleares. El ejemplo mejor conocido son los ciclos de fosforilación y desfosforilación de las proteínas SR. Estos factores de *splicing* se encuentran en estado hipofosforilado en los *speckles*, y acuden a los sitios de activación de transcripción cuando son fosforiladas por quinasas presentes en estos mismo compartimentos, como la Clk (Colwill et al., 1996)

6. APOPTOSIS

La apoptosis es el proceso de muerte celular programada de organismos multicelulares caracterizado por cambios bioquímicos y morfológicos como la condensación del citoplasma y núcleo de la célula, fragmentación del ADN por la activación de endonucleasas, fragmentación de la membrana en cuerpos apoptóticos y una fagocitosis inmediata por las células vecinas. Por tanto, se diferencia del otro tipo de muerte celular, la necrosis, en que el contenido celular no es liberado al exterior (Kerr, Wyllie, and Currie, 1972; Wyllie, 1980). La muerte celular programada es un proceso fisiológico esencial que garantiza el equilibrio homeostático entre la proliferación y muerte celular, por tanto, es de gran relevancia para el mantenimiento de tejidos y procesos biológicos, como la diferenciación o en respuesta a procesos citotóxicos, como la progresión tumoral (Cotter et al., 1990). Alteraciones en el proceso de apoptosis conlleva anomalías que pueden desencadenar enfermedades autoinmunes, neurodegenerativas, cáncer o infecciones virales entre otras (Thompson, 1995).

El desarrollo de la muerte celular por apoptosis puede darse por dos rutas: la vía extrínseca o ruta de los receptores de la familia del receptor del factor de necrosis tumoral, TNF-R; y la vía intrínseca o mitocondrial.

6.1 Vía extrínseca

La vía extrínseca es activada por estímulos extracelulares que inducen la unión de un ligando a alguno de los receptores extracelulares anclados en la superficie celular, Fas/APO-1/CD95, TNFR-1 y TRAIL-R. La ruta de apoptosis mejor caracterizada en la vía extrínseca es la de Fas/CD95. Cuando se produce la unión ligando-receptor, se desencadena la respuesta de apoptosis que comienza con la oligomerización de los dominios citoplasmáticos de los receptores, dando lugar al complejo de señalización inducido por muerte, DISC (*death-inducing signaling complex*) (Wajant, 2002). Este complejo recluta proteínas adaptadoras como FADD (*Fas-associated death domain*), que contiene sitios de unión para las procaspasas 8 y 10, que son activadas e inician una cascada de procesamientos proteolíticos que son responsables de ejecutar el proceso de apoptosis propiamente dicho. FADD también es capaz de unir a c-FLIP (*cellular FLICE inhibitory protein*), en cualquiera de sus isoformas: c-FLIP_S y c-FLIP_R, que no contienen dominios proteolíticos y actúan como inhibidores de la apoptosis al ocupar el sitio de unión de la procaspasa-8; c-FLIP_L que contiene un dominio proteolítico homólogo al de la caspasa 8 y puede funcionar como activador o inhibidor de la apoptosis (Fricker et al., 2010; Peter and Krammer, 2003).

6.2 Vía intrínseca

La vía intrínseca o mitocondrial es inducida por estímulos intracelulares o algún daño celular que provoca la liberación de proteínas solubles mitocondriales hacia el citoplasma, como el citocromo c, Smac/Diablo o HtrA2/Omi, debido a la permeabilización de la membrana externa mitocondrial (MOMP, *mitochondrial outer membrane permeabilization*) (Vaux, 2011). En el caso del citocromo c, recluta a la proteína Apaf-1 y la pro-caspasa 9 dando lugar a un complejo funcional, el apoptosoma, que activa a la caspasa 9 e inicia una cascada proteolítica llevada a cabo por caspasas, que termina con el corte de proteínas esenciales para la actividad celular (Li et al., 1997; Rodriguez and Lazebnik, 1999).

La liberación de citocromo c está regulada por miembros de la familia de proteínas Bcl-2 (*B cell lymphoma*), que a su vez puede dividirse en tres subfamilias dependiendo del número de dominios BH (*Bcl-2 homolog*): proteínas anti-apoptóticas con cuatro dominios BH (BH1, 2, 3, 4), incluye proteínas como Bcl-2 y Bcl-x_L; proteínas pro-apoptóticas multidominio sin el BH4, incluye Bax, Bak o Bcl-x_S; proteínas pro-apoptóticas que sólo contienen un dominio BH3 (*BH3-only*), incluye a Bim, o Bid entre otras (Youle and Strasser, 2008). Las proteínas proapoptóticas de multidominio facilitan el MOMP al oligomerizar en la membrana mitocondrial, permitiendo la liberación de citocromo c y parece que inhiben a proteínas anti-apoptóticas como Bcl-2 o Bcl-x_L (Letai et al., 2002). Las proteínas anti-apoptóticas intentan prevenir esa oligomerización mediante el secuestro de estos factores. Cuando se produce un estímulo apoptótico, las proteínas pro-apoptóticas con un solo dominio BH3 son activadas y favorecen la oligomerización de Bax y Bak en la membrana mitocondrial por una activación directa o eliminando la inhibición que ejercen las proteínas anti-apoptóticas (Cheng et al., 2001; Kuwana et al., 2005).

El proceso de apoptosis puede tener lugar independientemente de la actividad de las caspasas, mediante la liberación de factores que inducen la apoptosis (AIF, apoptosis inducing factor) de la mitocondria y su translocación al núcleo. Allí llevan a cabo una serie de eventos como la condensación de la cromatina y fragmentación del ADN (Lorenzo et al., 1999), aunque la morfología de la célula es diferente a la de la apoptosis inducida por caspasas (Norberg, Orrenius, and Zhivotovsky, 2010).

6.3 Caspasas

Las caspasas son enzimas cisteína proteasas (*cistein-aspartic proteases*) responsables de la mayoría de cambios morfológicos efectuados durante el proceso de apoptosis. Son expresadas de forma ubicua como precursores inactivos o pro-caspasas (zimógenos) con escasa o inexistente actividad. Las caspasas son activadas en respuesta a un estímulo de muerte celular, mediante un corte en residuos de ácido aspártico específicos, dando lugar a la eliminación de un dominio amino-terminal y produciendo una subunidad mayor (p20) y otra menor (p10). Estas subunidades se unen entre sí para formar heterodímeros o heterotetrámeros que serán los encargados de llevar a cabo los cortes proteolíticos (Earnshaw, Martins, and Kaufmann, 1999; Shi, 2002). La familia de

las caspasas ha sido tradicionalmente dividida en dos grupos: caspasas iniciadoras y caspasas efectoras.

- **Caspasas iniciadoras.** Las caspasas iniciadoras incluyen las caspasas-1, -2, -4, -5, -8, -9, -10, -11 y -12. Se caracterizan porque contienen un prodominio en su extremo amino-terminal de larga longitud que se unen a las moléculas adaptadoras que promueven la multimerización y activación de las caspasas. Por ejemplo, la caspasa-8 y -10 se activan al interactuar con FADD, mientras que la caspasa-9 se activa en presencia de ATP al unirse a Apaf-1 para formar el apoptosoma junto al citocromo c (Earnshaw, Martins, and Kaufmann, 1999).

- **Caspasas efectoras.** Las caspasas efectoras incluyen a las caspasas-3, -6 y -7. Se caracterizan por tener un prodominio de corta longitud que es cortado por las caspasas iniciadoras dando lugar a la forma activa de las caspasas efectoras, que actuarán activando o inactivando a una multitud de sustratos como el citoesqueleto, proteínas reparadoras del daño celular, del metabolismo del ARNm, del ciclo celular o reguladoras de la apoptosis, entre otras, para producir los cambios morfológicos característicos del proceso de apoptosis (Fischer, Janicke, and Schulze-Osthoff, 2003). La caspasa-3 es activada al dimerizar los fragmentos generados por el corte proteolítico de otras caspasas como la 8 o la 9. La caspasa-3 inicia un proceso de degradación de proteínas esenciales para la integridad celular, como PARP (*Poly ADP-ribose polymerase*) (Tewari et al., 1995). A diferencia de la caspasa-7, que parece tener funciones redundantes a la caspasa-3, la caspasa-6 está implicada en la degradación de la laminina A, proteína estructural de la membrana nuclear, que implica la fragmentación nuclear y condensación de la cromatina (Ruchaud et al., 2002).

- **Inhibidores de caspasas.** Paralelamente a la activación de las caspasas, la inactivación de estas proteasas es igualmente importante para mantener el equilibrio homeostático. Por ello, existe una familia de inhibidores de caspasas, que funcionan como inhibidores de la apoptosis a la vez (IAP, *inhibitor of apoptosis proteins*), capaces de unirse a las caspasas e inhibir su activación. Los IAPs más conocidos son XIAP (*X-linked Inhibitor of Apoptosis*), cIAP-1 (*Cellular Inhibitor of Apoptosis*) y cIAP-2, que reprimen a las caspasas 3, 7 y 9 (Deveraux et al., 1998). Las IAPs son inhibidas a su vez por las caspasas o bien secuestradas por proteínas mitocondriales como Smac/DIABLO (Salvesen and Duckett, 2002).

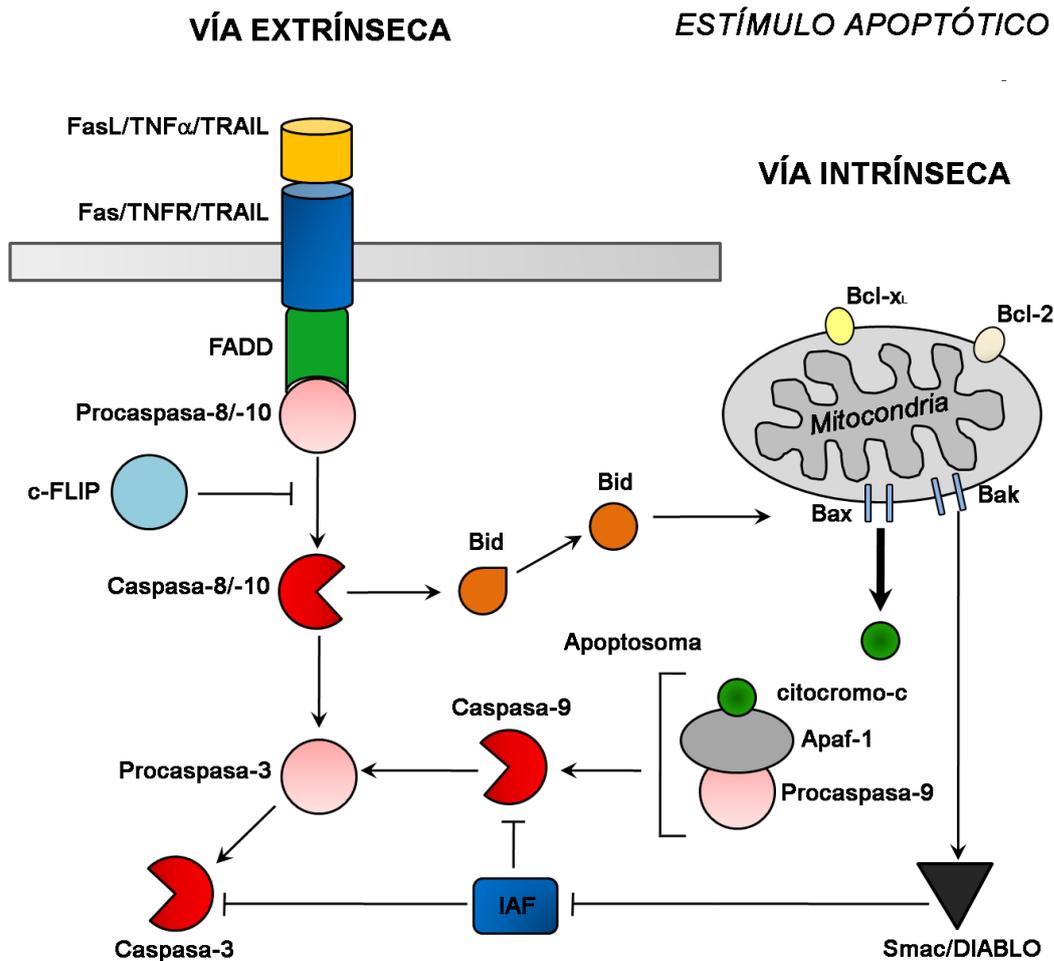


Figura I-9. Representación esquemática de las rutas de apoptosis. La vía extrínseca es activada por estímulos extracelulares y se desencadena por la unión de un ligando (FasL/TNF α /TRAIL) a su receptor de membrana extracelular (Fas/TNFR/TRAIL) y activación de las caspasas. La vía intrínseca se activa por señales intracelulares que reorganizan la membrana mitocondrial permitiendo la salida de citocromo c al citoplasma activando la ruta de las caspasas. Estas caspasas pueden ser inhibidas por moléculas IAF, que a su vez son reguladas por factores como Smac/DIABLO al salir de la mitocondria al citoplasma. Ambas vías están conectadas por el corte proteolítico de Bid.

7. APOPTOSIS Y *SPLICING*

Muchas de las proteínas implicadas en el proceso de apoptosis son reguladas mediante *splicing* alternativo, generando isoformas que tienen funciones opuestas, como ocurre con el gen *Bcl-x* y *Fas*. Por el interés para esta tesis doctoral, a continuación se describirán algunos aspectos sobre la regulación del *splicing* alternativo de *Fas*.

7.1 Regulación de Fas

El gen *FAS/APO-1/CD95* codifica para un receptor de superficie celular de la familia de receptores TNFR (*tumor necrosis factor receptor*) que participa en la iniciación de la respuesta de apoptosis por la vía extrínseca. Este gen puede ser procesado alternativamente, mediante la exclusión del exón 6, que codifica para el dominio transmembrana, dando lugar a una isoforma soluble anti-apoptótica, ya que compite con el receptor de Fas por la unión de Fas ligando (FasL) (Cascino et al., 1995; Cheng et al., 1994; Papoff et al., 1996). El cambio de la isoforma soluble a la transmembrana tiene lugar durante la activación de los linfocitos T (Liu, Cheng, and Mountz, 1995) en un proceso altamente regulado, y defectos en la regulación en el cambio de isoformas se han asociado con ciertas anomalías, como el síndrome linfoproliferativo autoinmune (ALPS, *autoimmune lymphoproliferative syndrome*) (Roesler et al., 2005). Un trabajo más reciente relaciona el síndrome del sarcoma de Ewing con una disminución de la isoforma del receptor, al alterar el reconocimiento de los sitios de *splicing* por U1 snRNP y U2AF⁶⁵ durante la definición del exón (Paronetto et al., 2014).

El receptor de Fas no sólo participa en la respuesta de apoptosis, si no que es capaz de activar otras rutas como las de MAPK quinasas, p38, JNK1/2 y ERK1/2, y la del factor de transcripción NF- κ B participando así en procesos de proliferación celular, migración e inflamación (Brint, O'Callaghan, and Houston, 2013).

Existen diversos mecanismos que modulan la respuesta del receptor de Fas, entre los que se encuentran el *splicing* alternativo. Por eso, la regulación del *splicing* alternativo de *Fas* ha sido ampliamente estudiada. La proteína de unión a ARN TIA-1 (*T-cell intracellular antigen 1*), se une a una región intrónica próxima al sitio 5'SS del exón 6 de *Fas* (URI6, *uridine-rich intronic 6 element*), promoviendo el reclutamiento de U1 snRNP y facilitando a su vez la unión de U2AF⁶⁵ y U2 snRNP con el intrón 5 durante la definición del exón. Además de facilitar la unión de U1 snRNP, parece que TIA-1 podría estar favoreciendo el reconocimiento del exón 6 de *Fas* por el bloqueo físico de una secuencia rica en guanidinas situada en la unión exón6/intrón6, donde se une el represor hnRNP F/H (Izquierdo et al., 2005). El efecto de TIA-1 en el *splicing* de *Fas* podría verse incrementado por la fosforilación de la quinasa FAST K (Izquierdo and Valcarcel, 2007). Por el contrario, la proteína represora PTB, reconoce en el exón 6

de *Fas* un elemento silenciador (URE6, *uridine-rich exonic 6 element*) y actúa de forma antagonica a TIA-1 al interrumpir los eventos moleculares que tienen lugar durante la definición del exón y por tanto, la comunicación entre ambos extremos del exón, inhibiendo la unión de U2AF⁶⁵ y U2 snRNP en el otro extremo del exón 6. Sin embargo, no altera el reconocimiento del sitio 5' de *splicing*, sino que es necesario que se produzca para ejercer su efecto represor (Izquierdo et al., 2005). De forma silimar, el regulador anti-apoptótico Hu antígeno R (HuR) promueve la exclusión del exón 6 al inhibir la asociación de U2AF⁶⁵ al sitio 3' de *splicing*, sin alterar la unión de TIA-1 aunque el reconocimiento del sitio 5' de *splicing* por U1 snRNP es necesario para el efecto represor de HuR sobre la definición del exón (Izquierdo, 2008).

Además de estos factores que regulan el *splicing* de *Fas* durante la definición del exón, también se han descrito reguladores que participan en las etapas de maduración del *spliceosoma*. La proteína RBM5/Luca-15/H37, implicada en el control de la proliferación celular y apoptosis, promueve la exclusión del exón 6 de *Fas* al inhibir el reclutamiento del triplete U4/U5/U6 snRNP a través de su dominio OCRE, al complejo pre-spliceosomal, impidiendo la maduración del *spliceosoma* a ambos lados del exón 6 (Bonnal et al., 2008). Un reciente estudio del genoma ha identificado 200 reguladores del *splicing* de *Fas*, donde se han encontrado además de componentes del *spliceosoma*, factores implicados en transcripción y remodelación de la cromatina, apoyando el *splicing* co-transcripcional como mecanismo de regulación en la selección del sitio de *splicing*. También se han identificado factores de señalización celular y control metabólico, que corroboran la relación del *splicing* de *Fas* con procesos celulares. Además este estudio revela la relación de la homeostasis del hierro y la modulación del *splicing* del exón 6 de *Fas*. Concretamente, SRSF7 podría estar realizando un papel mediador entre los niveles de hierro en la célula y la inclusión del exón 6 (Tejedor, Papasaikas, and Valcarcel, 2014).

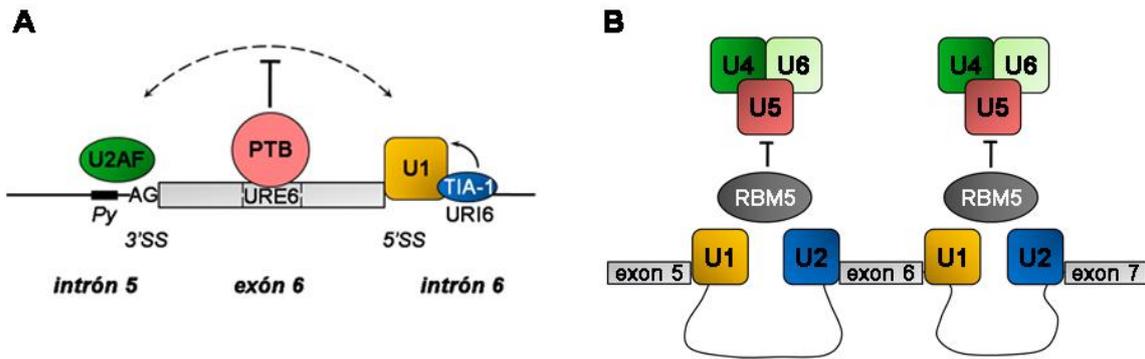


Figura I-10. Representación esquemática de la regulación del *splicing* del exón 6 de Fas. **A.** La unión de TIA-1 al sitio URI6 del intrón 6 promueve la unión de U1 snRNP al sitio 5'SS. Esta unión a su vez favorece el reclutamiento de U2AF al sitio 3'SS y al tracto de polipirimidinas (Py). El reconocimiento de la secuencia URE6 por PTB inhibe la comunicación entre ambos extremos del exón. **B.** RBM5 inhibe la unión del triplete U4/U5/U6 snRNP a los intrones 5 y 6 que flanquean el exón 6 de Fas.

8. EL *SPLICING* ALTERNATIVO Y LA ENFERMEDAD

El *splicing* es un proceso esencial para la célula por lo que se encuentra altamente regulado. En una situación normal, existe un equilibrio entre las isoformas de ARNm generadas a partir de un gen de acuerdo a las condiciones celulares existentes. Sin embargo, mutaciones o polimorfismos que dan lugar a un cambio en las variantes de ARNm pueden alterar la homeostasis celular llegando a desencadenar enfermedades. Un ejemplo del equilibrio en los patrones de *splicing* ocurre con la familia de proteínas de apoptosis Bcl-2. Muchas de ellas generan mediante *splicing* alternativo, isoformas que promueven o previenen la apoptosis, como el gen *Bcl-x*, cuyo desequilibrio en las dos isoformas está implicado en algunos tipos de cáncer (Tu et al., 1998).

8.1 Tipos de alteraciones en el *splicing*

Se estima que el 15% de las mutaciones que causan enfermedades genéticas afectan al procesamiento del pre-ARNm (Krawczak, Reiss, and Cooper, 1992). La mayoría de estas mutaciones se dan en regiones no codificantes como los sitios 3'SS y 5'SS, el tracto de polipirimidinas o el BPS, que generan errores en el proceso de *splicing* y conlleva cambios en el patrón de procesamiento, incluyendo exones alternativos o reteniendo intrones, entre otros eventos (Krawczak, Reiss, and Cooper, 1992).

Mutaciones en las regiones codificantes del ARNm también pueden generar anomalías cuando crean un cambio de aminoácido que modifica la función de la proteína. Estos cambios también pueden ser silenciosos cuando tienen lugar en las secuencias reguladoras ESE o ESS, alterando la función reguladora de las mismas por una pérdida de función o incluso, generando nuevas secuencias *enhancer* o silenciadores que en última instancia alteran el proceso de *splicing*. Ya se ha comentado previamente el caso de la enfermedad SMA, donde se produce un cambio de base en el exón 7 del gen *SMN1* que genera una mutación silenciosa. Sin embargo, este cambio de base crea un nuevo ESS que inhibe la inclusión del exón en el transcrito final, generando una proteína truncada (Monani et al., 1999). Esto también ocurre con polimorfismos de cambio de un solo nucleótido en regiones codificantes (cSNP, *coding single-nucleotide polymorphism*). Generalmente son cambios silenciosos puesto que no da lugar a un cambio de aminoácido, pero sí afecta al *splicing* alternativo al modificar los elementos reguladores ESE y ESS, crear nuevos sitios de *splicing*, fortalecer o debilitar los sitios de *splicing* ya existentes, alterar la estructura secundaria del pre-ARNm que afectan a la definición del exón o modificar la tasa de elongación de la RNAPII que a su vez afecta al proceso de *splicing*. Estos defectos no son exclusivos de los cSNPs, si no también están relacionados con mutaciones sin sentido (aparece un nuevo codón de terminación) con cambio de sentido (genera un cambio de aminoácido), deleciones o inserciones exónicas.

Además de mutaciones en elementos *cis*, también existen mutaciones en los factores reguladores de *splicing* en *trans* que ya no sólo afectan a un exón concreto de un gen, sino que se altera por completo el *splicing* de los genes que regulan, con consecuencias graves. Hay descritos un menor número de casos de este tipo de mutaciones, probablemente por la pérdida de función de la proteína que puede tener consecuencias letales durante etapas relevantes como el desarrollo, aunque se conocen algunos ejemplos de enfermedades como SMA, la esclerosis lateral amiotrófica (ALS, *amyotrophic lateral sclerosis*) y la retinitis pigmentaria (Cooper, Wan, and Dreyfuss, 2009). La proteína SMN es requerida para la formación de los snRNPs, por tanto, deficiencias en esta proteína genera defectos en el *splicing* de varios tejidos, principalmente neuronales, generando SMA (Zhang et al., 2008). La proteína TDP-43 es una hnRNP implicada en la exclusión del exón 9 del gen CFTR. Se ha descrito su

anormal ubiquitinación y agregación en neuronas y células de la glía en enfermedades neurodegenerativas como la demencia frontotemporal y ALS (Neumann et al., 2006).

Otro tipo de alteración que afecta al *splicing* son los desórdenes causados por la extensión de microsatélites, secuencias repetitivas de un nucleótido. Estas repeticiones pueden causar una pérdida o ganancia de función tanto de la proteína cuando se encuentran en regiones codificantes, como del ARN cuando se encuentran en regiones no codificantes (Gatchel and Zoghbi, 2005). Existen casos muy conocidos como la enfermedad de Huntington, donde la proteína huntingtina contiene una cola de poliglutaminas que provoca una ganancia de función de la proteína. Sin embargo uno de los casos mejor caracterizados es en la distrofia miotónica (DM1). La DM de tipo 1 (DM1) está causada por la extensión de un triplete CTG en la región 3' UTR del gen *DMPK*, mientras que la de tipo 2 (DM2), se debe a la expansión de cuatro bases CCTG en el primer intrón del gen *ZNF9* (Liquori et al., 2001; Mahadevan et al., 1992). Estas secuencias son reconocidas con una alta afinidad por el regulador de *splicing* MBLN1, secuestrando a la proteína del nucleoplasma. En cambio el factor de *splicing* CUGBP1 es activado de forma aberrante al reconocer la repetición CUG del ARN. La alteración de la función de estos dos factores parece afectar el desarrollo normal de los patrones de *splicing* en los tejidos con DM (Kuyumcu-Martinez, Wang, and Cooper, 2007; Savkur, Philips, and Cooper, 2001).

8.2 El *splicing* alternativo y el sistema nervioso

El *splicing* alternativo tiene especial relevancia en el sistema nervioso, ya que es donde se expresan el mayor número de isoformas de ARNm, aumentando la diversidad proteica, así como la complejidad en la regulación génica. Por eso, el sistema nervioso es uno de los tejidos que presenta mayor número de factores de *splicing* y con una expresión diferencial (Grosso et al., 2008). Además, el *splicing* alternativo tiene un papel fundamental en la regulación de procesos neuronales esenciales como la diferenciación y la migración neuronal ya que participa en el mantenimiento de sus componentes, como las neurexinas, implicados en la sinapsis. También procesos del sistema nervioso, como la despolarización, activan el *splicing* alternativo aunque todavía no se conoce con claridad el mecanismo (Norris and Calarco, 2012). Por tanto, es evidente la relación existente entre los fallos producidos en el proceso de *splicing* y el

desarrollo de enfermedades asociadas al sistema nervioso como SMA, ALS, enfermedad de Huntington, entre otras.

Una de las primeras enfermedades neurológicas donde se estudiaron alteraciones en elementos de la maquinaria de *splicing* fue la demencia frontotemporal y Parkinson asociado al cromosoma 17 (FDTP-17). Esta enfermedad es causada por mutaciones en el gen *MAPT* que codifica la proteína Tau, un factor implicado en el ensamblaje de microtúbulos que tiene varias isoformas generadas por *splicing* alternativo. Estudios sobre la enfermedad revelaron que mutaciones silenciosas (que no generan cambio de aminoácido) en el exón 10 de *MAPT*, tenían una mayor importancia en el desarrollo de la enfermedad que alteraciones presentes en la secuencia de la proteína Tau. Este hecho era debido a la inclusión del exón 10 en los transcritos maduros de una forma aberrante (D'Souza et al., 1999).

8.3 El *splicing* alternativo en el cáncer

Es frecuente encontrar diferentes variantes de *splicing* en tejidos cancerígenos comparado con tejidos normales. La presencia de estas isoformas aberrantes puede deberse a mutaciones presentes en el pre-ARNm o bien a alteraciones producidas en las proteínas reguladoras (Grosso, Martins, and Carmo-Fonseca, 2008). Sin embargo, todavía queda por determinar si cambios en el *splicing* alternativo promueven o inhiben la progresión del tumor. También es necesario profundizar en la detección de alteraciones del *splicing* como dianas terapéuticas, aunque cada vez hay más estudios que intentan buscar marcadores de cáncer en el spliceosoma (Bonnal, Vigevani, and Valcarcel, 2012). De hecho, como se describe a continuación, se han publicado recientemente numerosos estudios en relación a la mielodisplasia que está ligada a la leucemia y recurrentes mutaciones en factores de *splicing* como el factor SF3B1.

El síndrome mielodisplásico (MDS) y desórdenes relacionados (mielodisplasia) son un conjunto de neoplasmas mieloides caracterizados por la desregulación de la producción de células sanguíneas, morfología de células displásica y la predisposición a la leucemia aguda mieloide (AML, *acute myeloid leukaemia*) (Corey et al., 2007). Estudios de secuenciación del genoma completo de pacientes con diferentes subtipos de mielodisplasia, han identificado mutaciones en seis genes de factores relacionados con el *spliceosoma*: *U2AF35*, *ZRSR2*, *SRSF1*, *SF3A1*, *SF3B1* y *PRPF40B*, donde *SF3B1*

presenta el mayor índice de incidencia, seguido de *U2AF35*. Todos estos factores están implicados en las primeras etapas del ensamblaje del *spliceosoma*, excepto *PRPF40B* del cual se disponía de escasa información antes del desarrollo de esta tesis. Estas mutaciones que aparecían de una forma mutuamente exclusiva en el 55,2% de los pacientes, se confirmaban en el 57% de casos de MDS, 54.5% de leucemia mielomonocítica crónica (CMML), 25.8% de AML y un porcentaje muy bajo en mutaciones *de novo* en AML (6.6%) y neoplasma mieloproliferativo (MPS) (9.4%). (Papaemmanuil et al., 2011; Yoshida et al., 2011). Este conjunto de mutaciones representan un ejemplo de funciones celulares esenciales como el *splicing* afectadas en cáncer. Estos estudios también proporcionan algunas pistas sobre el mecanismo mediante el cual la alteración en el *splicing* puede intervenir en el desarrollo de cáncer, aunque la relación entre las mutaciones y el fenotipo MDS aún no está clara. En el caso de *U2AF35* (también conocido como *U2AF1*) parece que las alteraciones en el *splicing* se deben a la pérdida de especificidad por el sitio 3' de *splicing* (Okeyo-Owuor et al., 2014). También se han hecho ensayos con *U2AF35* mutado donde se observa un aumento de la retención de intrones y exclusión de exones alternativos, pero al contrario de lo que cabría esperar estas isoformas generadas parecen reducir la viabilidad celular y promover la apoptosis (Graubert et al., 2012). En cambio se ha mostrado que Fas es un regulador clave de la apoptosis en células neoplásicas, aunque su actividad parece inhibirse por la hipermetilación del promotor al pasar de un nivel de MDS leve a uno grave o AML (Karlic et al., 2014). Esta modificación epigenética en el promotor parece ocurrir también en el factor relacionado con el *splicing* *YBX3* (*Y box binding protein 3*), en la línea celular mieloide U937 (Wong et al., 2014).

Objetivos

Los objetivos de esta tesis doctoral son:

1. Caracterización bioquímica de PRPF40B
 - a. Estudio de la localización de PRPF40B
 - b. Análisis general de los elementos estructurales necesarios para su distribución y función.
 - c. Identificación de importantes interacciones para la funcionalidad de PRPF40B.

2. Caracterización funcional de PRPF40B
 - a. Determinar el papel de PRPF40B en el proceso de *splicing* alternativo.
 - b. Investigación de la participación de PRPF40B en la transcripción
 - c. Estudio general de la función de PRPF40B en el transcriptoma humano.

Materiales y Métodos

1. PLÁSMIDOS

El vector comercial que incluye la secuencia de ADN codificante de PRPF40B (clon MGC:71266 IMAGE:664467, Source Bioscience) fue utilizado como intermediario para generar las construcciones basadas en PRPF40B que se describirán a continuación. El vector de expresión eucariota pEFBOST7-PRPF40B (1-871) y las versiones truncadas pEFBOST7-PRPF40B (1-687), pEFBOST7-PRPF40B (263-871), pEFBOST7-PRPF40B (394-871), pEFBOST7-PRPF40B (474-871) y pEFBOST7-PRPF40B (618-871) se obtuvieron por técnicas de PCR convencional con las parejas de oligonucleótidos HYPC1Fw/T7HYPCRev, HYPC1Fw/HYPC687Rev, HYPC263Fw/T7HYPCRev, HYPCGFPCt/T7HYPCRev, HYPC474Fw/T7HYPCRev y HYPC618Fw/T7HYPCRev, respectivamente. Los fragmentos generados fueron digeridos con las enzimas de restricción *XbaI* y *ClaI* e insertados en el vector pEFBOST7 usando las dianas *XbaI* y *BstBI*; este vector codifica el epítipo T7 de 11 aminoácidos en su extremo amino. pEFBOST7 fue obtenido al retirar, mediante la digestión con *XbaI* y *BstBI*, el fragmento de TCERG1 (1-662) que contenía el vector pEFBOST7-TCERG1 (1-662), generado previamente en el laboratorio. Todas estas secuencias que incluían la región codificante para PRPF40B fusionadas al epítipo T7 fueron extraídas por digestión con *EcoRI* e insertadas en las mismas dianas del vector de expresión eucariota pEFBOS-EGFP/T7TCERG1 generado anteriormente en el laboratorio, que codifica para la proteína fluorescente EGFP. El vector de expresión eucariota pEFBOS/EGFP/T7-PRPF40B (474-553/672-871) fue generado a partir de la digestión del vector pEFBOS/EGFP/T7-PRPF40B (474-871) con las enzimas de restricción *XbaI*/*BamHI*, y por inserción del fragmento amplificado por PCR PRPF40B (474-553) con dianas para *XbaI*/*BamHI*.

El vector de expresión pEFBOST7-SF1 fue construido por PCR utilizando los oligonucleótidos SF1FLT7Fw y SF1FLT7Rv usando como molde el vector pGEM-SF1-Bo, proporcionado por la Dra. Angela Kramer (Université de Genève, Suiza). Los fragmentos obtenidos fueron cortados con las enzimas *XbaI* y *BstBI*, aprovechando la presencia de estas dianas en el vector pEFBOST7.

El vector de expresión eucariota pcDNA3/HA-PRPF40B fue obtenido por métodos de PCR convencional utilizando la pareja de oligonucleótidos GSTHYPCFw/GSTHYPCRv con dianas *BglIII* y *EcoRI*. El fragmento amplificado fue

insertado en las dianas *BamHI* y *EcoRI* del vector de expresión pcDNA3/HA cedido por el Dr. Javier Oliver (IPBLN-CSIC, Granada). Este vector fue utilizado como molde para generar los vectores de expresión eucariota con las mutaciones puntuales pcDNA3/HA-PRPF40B:P383L, pcDNA3/HA-PRPF40B:D442N y pcDNA3/HA-PRPF40B:P540S, siguiendo las instrucciones del kit de mutagénesis dirigida *Quickchange II XL-Site-Directed* (Stratagene).

Los vectores de expresión eucariota pmCherry-PRPF40B y pPRPF40B-mCherry fueron generados mediante procedimientos estándar de PCR utilizando las parejas de oligonucleótidos HYPCBsrGIFw/HYPCBsrGIRv y HYPCAgeIFw/HYPCAgeIRv, respectivamente. Los fragmentos de PCR obtenidos fueron digeridos con las enzimas BsrGI o AgeI e insertados en las mismas dianas del vector pmCherry-C1 que codifica para la proteína fluorescente roja mCherry. El vector pmCherry-C1, así como los vectores de expresión YFP-SF1, YFP-U2AF⁶⁵, U2AF⁶⁵-YFP fueron cedidos por el Dr. José Rino (Instituto de Medicina Molecular, Lisboa, Portugal) y han sido descritos previamente (Rino et al., 2008).

Los vectores de expresión en bacterias pGEX2TK-PRPF40B (1-871), pGEX2TK-PRPF40B (1-333) y pGEX2TK-PRPF40B (446-871), basados en el vector comercial pGEX2TK (Pharmacia Biotech) fueron obtenidos mediante técnicas convencionales de PCR con los siguientes pares de oligonucleótidos: GSTHYPCFw/GSTHYPCRv, GSTHYPCFw/GSTHYPC333Rv y GSTHYPCtFw/GSTHYPCRv. Los fragmentos de PCR fueron digeridos con las enzimas *BglIII* y *EcoRI*, e insertados en las dianas *BamHI* y *EcoRI* presentes en el vector pGEX2TK. Del mismo modo fue generado el vector pGEX2TK-U2AF⁶⁵, utilizando la pareja de oligonucleótidos U2AF65BamHIFw/U2AF65EcoRIRv para amplificar la secuencia de U2AF⁶⁵ a partir del vector de expresión eucariota GFP-U2AF⁶⁵ proporcionado por el Dr. Juan Valcárcel (Centro de Regulación Genómica, Barcelona). Los vectores de expresión en bacterias pGEX2TK-TCERG1 (234-662) y pGEX2TK (631-1098) han sido descritos previamente (Carty et al., 2000).

El vector de expresión eucariota pcDNA3/FLAG/MS2-PRPF40B fue obtenido por PCR utilizando la pareja de oligonucleótidos HYPCMS2Fw/HYPCMS2Rv con dianas *AvrII* y *MluI*. El fragmento amplificado fue insertado en el vector de expresión pcDNA3/FLAG/MS2 con las mismas dianas. Este último vector así como

pcDNA3/FLAG/MS2-PTB fueron cedidos por el Dr. Juan Valcárcel (Centro de Regulación Genómica, Barcelona).

Los plásmidos reporteros p3xMHC-LUC y pRL-TK-REN fueron cedidos por el Dr. Mariano García-Blanco (Duke University Medical Center, USA).

Los plásmidos reporteros de luciferasa pG5HIV-LUC y pG5NFI+II-LUC fueron generados en el laboratorio y descritos en anteriores trabajos (Montanuy et al., 2008).

Los minigenes reporteros de *splicing* del gen *Fas*: WT, Py, U1C, e1, e2, e3, e23, e123, m0, m1, m2, Long exón y MS2 fueron proporcionados por el Dr. Juan Valcárcel (Centro de Regulación Genómica, Barcelona).

El minigen reportero de *splicing* del gen *Bcl-x* (*HIV-X2*) fue cedido por el Dr. Benoit Chabot (University of Sherbrooke, Canadá).

Los vectores lentivirales que codifican para la cápsida (pHR'82ΔR) y envuelta (pVSV.G) del virus, así como los vectores pLKO.1-puro que producen las partículas lentivirales y contienen las secuencias shRNA (*small hairpin RNA*) de PRPF40B fueron proporcionados por la Dra. Carme Gallego (Instituto de Biología Molecular de Barcelona, Barcelona). Los shRNAs codifican un fragmento de secuencia de PRPF40B de 21 nucleótidos sentido y antisentido separadas por el espaciador CTCGAG y son de la colección *Mission* de Sigma.

2. CULTIVOS CELULARES Y TRANSFECCIONES

2.1 Cultivos celulares

Todas las líneas celulares que se detallarán a continuación se cultivaron en condiciones estándares idénticas: 37°C en atmósfera de CO₂ del 5% y humedad del 95%.

Las líneas celulares HEK293T y HeLa se cultivaron en medio de cultivo *Dulbecco's modified Eagle Medium* (DMEM) con una concentración 1 g/L de D-glucosa (Gibco), suplementado con 10% (v/v) de suero fetal bovino (Gibco), penicilina/estreptomicina (100U y 0,1 mg/ml, respectivamente) (Gibco), 4 mM L-glutamina (Gibco) y una solución de aminoácidos no esenciales (Gibco).

Las células HEK293T utilizadas como productoras de virus se cultivaron en medio DMEM con una concentración de glucosa de 4,5 g/L, suplementado de la misma forma previamente descrita y adicionando Myco kit (Labclinics) como agente antimicoplasma.

La línea celular de neuroblastoma SH-SY5Y fue proporcionada por la Dra. Carme Gallego (Instituto de Biología Molecular de Barcelona, Barcelona). Se cultivaron en medio DMEM con una concentración de 4,5 g/L de glucosa, 10% (v/v) de suero fetal bovino y penicilina/estreptomicina (100U y 0,1 mg/ml, respectivamente). Las células SH-SY5Y establemente silenciadas se mantuvieron con el mismo medio de cultivo adicionando puromicina (1 µg/ml) (Invitrogen) como antibiótico de selección.

2.2 Transfecciones

En los estudios de inmunofluorescencia se utilizaron las líneas celulares HEK293T y HeLa. Se sembraron aproximadamente 1×10^5 células sobre cubreobjetos de 12 mm de diámetro colocados en pocillos de 22 mm de placas M12 (Nunc) 20 h antes de la transfección. Cuando las células alcanzaron aproximadamente un 50-60% de confluencia se transfectaron con las correspondientes concentraciones de plásmidos utilizando el método de precipitación de Ca_2PO_4 . Para pocillos de 22 mm de diámetro, se mezclan 100 µl de Solución A (250mM CaCl_2) con el ADN. A continuación se añaden 100 µl de Solución B (50mM HEPES, 1,5mM NaPO_4 , 140mM NaCl, ajustar pH a 7,5) y tras 2 min de incubación, la mezcla se adiciona al correspondiente pocillo. 24 h después de la transfección se recogieron las células para ser procesadas.

Para experimentos de inmunoprecipitación se sembraron aproximadamente $1,5 \times 10^6$ células HEK293T en placas de 100 mm de diámetro (BD Falcon) 24 h antes de la transfección. Cuando las células alcanzaron aproximadamente un 80-90% de confluencia fueron transfectadas con las cantidades de ADN plasmídico mediante métodos estándar de Ca_2PO_4 . Después de 40 h de la transfección las células fueron recogidas y procesadas.

Para ensayos de actividad luciferasa y experimentos de sobreexpresión en ensayos de RT-PCR, se sembraron aproximadamente $2,5 \times 10^5$ células HEK293T o HeLa en pocillos de 35 mm (BD Falcon) 20 h antes de la transfección. Cuando la

confluencia de células fue del 80% para ensayos de luciferasa o del 60-70% para análisis por RT-PCR, se transfectaron con las correspondientes cantidades de plásmidos utilizando Ca_2PO_4 . Las células fueron recogidas para su análisis 40 h después de la transfección.

En el caso de los experimentos de RT-PCR con siRNAs se sembraron aproximadamente 9×10^4 células por pocillo de 35 mm. La transfección de los siRNAs se realizó cuando las células alcanzaron aproximadamente un 30-40% de confluencia utilizando el reactivo *Lipofectamine 2000* (Invitrogen) siguiendo las indicaciones del fabricante. La concentración final de los siRNAs fue de 60nM. Para los experimentos de *splicing*, las células se recogieron 48 h después de la transfección para su procesamiento, o bien, se transfectaban las correspondientes cantidades de ADN plasmídico de los respectivos minigenes descritos, utilizando el reactivo *LipoD 293T* (SignaGen Laboratories) siguiendo las indicaciones del fabricante. Después de 24 h de la transfección de ADN, las células se recogieron para su posterior análisis. Para los experimentos de apoptosis, las células se recogieron 72 h después de la transfección.

La inhibición transcripcional con α -amanitina se hacía añadiendo 25 $\mu\text{g}/\text{ml}$ de este compuesto durante 6 h a 37°C. Posteriormente, se continuaba con el análisis por inmunofluorescencia.

2.3 Infección de cultivos celulares

Las partículas de lentivirus utilizadas para la infección de la línea celular SH-SY5Y fueron producidas en células HEK293T.

Para la transfección de los plásmidos virales, se sembraron 8×10^5 células HEK293T en pocillos de 35 mm (BD Falcon) 20 h antes de la transfección. Cuando se alcanzó una confluencia de aproximadamente el 80-90%, las células fueron transfectadas con la cantidad adecuada de los plásmidos lentivirales de la envuelta, cápsida y portadores de las secuencias de shRNA en una relación 1:1:2 utilizando como reactivo de transfección *LipoD 293T* (SignaGen Laboratories) según las indicaciones del fabricante. 72 h después de la transfección, se recogieron los sobrenadantes de cada pocillo y se centrifugaron. Tras eliminar los sobrenadantes, 500 μl de las soluciones virales se diluyeron en el medio de cultivo adecuado, al cual se añadió *Polibrene*

(Sigma-Aldrich) a una concentración final de 8 µg/ml, para facilitar el contacto del virus con la célula receptora. Las soluciones de partículas de lentivirus resultantes se añadieron a las células receptoras SH-SY5Y, sembradas el día anterior sobre pocillos de 35 mm ($2,5 \times 10^5$ células/pocillo). Aproximadamente 16 h después de la primera infección se repitió el proceso de recogida y centrifugación de los sobrenadantes virales y se realizó una segunda infección, en este caso utilizando 4 µg/ml de *Polibrene*. Tras 5 h de infección, se cambió el medio y 48 h después, se añadió puromicina a 1 µg/ml como antibiótico de selección.

3. PROCEDIMIENTOS DE BIOLOGÍA MOLECULAR

3.1 Inmunofluorescencia

Para los análisis por inmunofluorescencia, las células situadas sobre cubreobjetos fueron fijadas con 3,5% de paraformaldehído (PFA) en PBS 1X, 20mM HEPES (pH 7,4) durante 45 min a 4°C. Tras tres lavados con PBS 1X, las células se permeabilizaron con una solución al 0,5% de Triton-X100 durante 5 min a temperatura ambiente y se lavaron tres veces con PBS 1X. A continuación, las células se bloquearon con una solución al 2,5% (p/v) de albúmina de suero bovino (BSA) en PBS 1X a 4°C, durante al menos 16 h. Tras dos lavados con PBS 1X conteniendo 0,1% de BSA y un lavado con PBS 1X a temperatura ambiente, se procedió al tratamiento con anticuerpos. Para el lavado e incubación de anticuerpos se utilizó una misma solución de 0,1% BSA en PBS 1X. Se añadieron 100 µl de la dilución adecuada del anticuerpo sobre el correspondiente cubreobjeto y se incubaron en una cámara húmeda en oscuridad durante 1 h a temperatura ambiente. Tras cinco lavados, se procedió del mismo modo a la incubación con los correspondientes anticuerpos secundarios. Tras los lavados con 0,1% BSA en PBS 1X, se realizaron tres lavados con PBS 1X y se prepararon las muestras utilizando el medio de montaje *Prolong Gold Antifade* (Molecular Probes).

El tratamiento con ARNasa se llevó a cabo, tras la permeabilización y posterior lavado con PBS 1X, añadiendo ARNasa (Roche) a 0,1 mg/ml en PBS 1X durante 2 h a 37°C. Después se continuó con el bloqueo y el resto del procesamiento descrito de inmunofluorescencia.

Para el estudio de la localización de los factores de splicing SF1 y U2AF⁶⁵ se siguió el protocolo que se detalla a continuación. Las células se lavaron dos veces con PBS 1X y se permeabilizaron con 0,5% de Triton-X100 en el tampón CSK (100mM NaCl, 300mM sacarosa, 10nM PIPES, 3mM MgCl₂, 1mM EGTA, pH 6,8) durante 1 min a 4°C. Seguidamente las células se fijaron con 3,5% PFA en una solución de CSK 5 min a temperatura ambiente. Tras lavar con PBS 1X, las células se bloquearon con 1% de suero fetal bovino (FBS) en PBS 1X durante 10 min a temperatura ambiente. La solución de bloqueo se utilizó para incubar los anticuerpos primarios y secundarios. Se añadieron 100 µl de la correspondiente dilución de anticuerpo primario sobre cada cubreobjetos y se incubó durante 1 h a temperatura ambiente en una cámara húmeda en oscuridad. Tras tres lavados con PBS 1X, se procedió del mismo modo al tratamiento con el anticuerpo secundario. Tras lavar con PBS 1X las muestras se montaron con *Prolong Gold Antifade*.

Las imágenes fueron adquiridas con el microscopio confocal invertido Leica SP5. Las imágenes eran tomadas simultáneamente en los casos en que se realizaba un doble marcaje con EGFP o el fluorocromo 647, excitados con el láser de helio-neón (He-Ne) de 633 nm, y el fluorocromo Alexa 488, excitado con el láser de argón a 488 nm al 20% de potencia. En el caso en que el doble marcaje se llevó a cabo con mCherry, excitado con el láser HeNe 543 y EGFP o el fluorocromo 647, las imágenes se tomaron de forma secuencial. En ambos casos el espectro de detección para cada canal se ajustó para evitar el paso de señal de un canal a otro. Las imágenes se tomaron en condiciones de no saturación con un objetivo 63x, resolución 1024 x 1024 y un diámetro de ranura confocal (*pinhole*) de 1 µm. Las imágenes se analizaron con el software Las AF (Leica). Posteriormente las imágenes fueron procesadas con el software *Adobe Photoshop CS3 v10.0*.

3.2 Extracción de ARN y RT-PCR

Para los análisis por RT-PCR, aproximadamente 20 h antes de la transfección se sembraron $2,5 \times 10^5$ células o 1×10^5 células por pocillo de 35 mm, para los estudios de sobreexpresión o silenciamiento, respectivamente.

Para analizar el efecto de la sobreexpresión de PRPF40B en el *splicing*, las células fueron transfectadas con 1 µg de PRPF40B o un plásmido vacío y las cantidades

de ácidos nucleicos se igualaron utilizando ARNt (Sigma). Aproximadamente 40 h después de la transfección, las células fueron recogidas y procesadas. Para los experimentos con minigenes reporteros, se seguía el mismo procedimiento pero las células eran co-transfectadas con las cantidades adecuadas de cada minigen: 0,25 µg de *Fas* o 0,6 µg de *HIV-X2*. Para analizar el efecto del silenciamiento de PRPF40B en el *splicing*, las células eran transfectadas con 60nM de cada siRNA utilizando como control siEGFP. 48 h después de la transfección de los siRNAs, las células fueron procesadas para el análisis de genes endógenos. Para estudiar el efecto del silenciamiento de PRPF40B en minigenes reporteros, 48 h después de la transfección de los siRNAs, se transfectaron las cantidades indicadas anteriormente de los minigenes reporteros. Las células eran procesadas 24 h más tarde para los experimentos del minigen *Fas* o 48 h después en el caso del minigen *HIV-X2*. Se recogían las células en su mismo medio de cultivo y un décimo del volumen se lisaba y mezclaba con tampón Laemli 2x para su análisis por *Western Blot*.

El ARN total era extraído con *peqGOLD Trifast* (peQlab) según las indicaciones del fabricante. 800 ng de ARN derivado de cada minigen se retro-transcribía con *M-MLU retrotranscriptase* (Invitrogen) durante 1 h a 37°C utilizando el oligo *csnPT2* para el minigen *Fas* y el oligo *RT-Sveda* para el minigen *HIV-X2*. Un cuarto de la reacción se amplificó mediante PCR convencional. El minigen *Fas* se amplificó utilizando los oligonucleótidos *csnPT1/csnPT2* tras 30 ciclos de PCR (Tm 61,5°C). El minigen *HIV-X2* se amplificó tras 30 ciclos de PCR (Tm 60°C) con los oligonucleótidos *X34/XAgeI*. Los productos de la PCR fueron resueltos en geles de agarosa al 2,5%. Las intensidades de las bandas fueron cuantificadas con el programa *Quantity One* v4.6.5 (Bio-Rad). Todos los experimentos se realizaron como mínimo por triplicado.

Para el estudio del efecto de PRPF40B en el *splicing* de *Fas* endógeno, se siguió el mismo procedimiento de extracción del ARN descrito anteriormente. Seguidamente, 1 µg de ARN se retro-transcribía utilizando *qScript cDNA Supermix* (Quanta Bioscience) siguiendo las indicaciones de la casa comercial, que contenía una mezcla con oligo (*dT*), *random primers* y la retro-transcriptasa *qScript*. El producto de reacción resultante se diluía cinco veces para su análisis. En el análisis del gen *Fas* endógeno por PCR convencional, un veintavo de la muestra se amplificó utilizando los oligonucleótidos *FasE5Fw/FE7Rv* (Tm 55°C), que comprenden la región situada entre

los exones 5 y 7 incluyendo las isoformas pro-apoptótica y antiapoptótica del gen *Fas*. Los productos de PCR amplificados fueron resueltos en geles de agarosa al 3,5%. Las intensidades de las bandas fueron cuantificadas con el programa *Quantity One* v4.6.5 (Bio-Rad). Los experimentos se realizaron por triplicado. La cuantificación del gen *Fas* por PCR a tiempo real, se llevó a cabo utilizando *iQ SYBR Green Supermix* (Bio-Rad) en la estación *iCycler thermal cycler station* (Bio-Rad) con los oligonucleótidos *FE6/FE7* (T_m 53,5°C) que amplifican la región comprendida entre los exones 6 y 7 del gen *Fas*, incluida en la isoforma pro-apoptótica. Como control interno de cada experimento se midieron los niveles de *Glyceraldehyde-3-phosphate Dehydrogenase* (GAPDH) utilizando los oligonucleótidos *GAPDH-F/GAPDH-R*. Los experimentos se realizaron cinco veces como mínimo. Los valores están representados como $(E_{Fas})^{\Delta CT(Fas)} / (E_{GAPDH})^{\Delta CT(GAPDH)}$, donde E representa la eficiencia de la PCR y ΔCT : CT muestra control-CT muestra problema.

Para el análisis estadístico se usó el programa *Prism 5.0 software* (GraphPad). Se aplicó el test *t-Student* para comparar las medias entre muestras. Los valores P están representados en las gráficas mediante asteriscos (*, $P < 0,05$; **, $P < 0,01$; ***, $P < 0,005$).

3.3 Extracción de ARN total para los análisis por *array*

Para los análisis por *array*, se sembraron aproximadamente $2,5 \times 10^5$ células en pocillos de 35 mm de las líneas SH-SY5Y silenciadas establemente con un shRNA control o dirigido contra PRPF40B hasta alcanzar una confluencia del 60%.

Las células se lavaron dos veces con PBS 1X para eliminar restos celulares. Para obtener ARN total con alto grado de pureza se utilizó el kit comercial *RNeasy Mini Kit* (Qiagen). Para ello las células se lisaron con el tampón RLT del kit. A continuación se pasó el extracto celular de cada muestra diez veces por agujas de 0,6 mm y 25 mm para fragmentar el ADN. Seguidamente se adicionó etanol al 70% al ARN para que adquiriese las condiciones apropiadas de unión a las columnas proporcionadas por el kit. Se continuó con una serie de lavados y eluciones de acuerdo con las indicaciones del fabricante. Para los análisis posteriores se utilizó 1 μ g de ARN de tres muestras independientes de las dos líneas establemente silenciadas.

3.4 Mutagénesis dirigida de PRPF40B

Para realizar las mutaciones puntuales de PRPF40B asociadas al síndrome mielodisplásico descritas anteriormente (Yoshida et al., 2011), se utilizó el kit de mutagénesis dirigida *Quickchange II XL-Site-Directed* (Stratagene). Se siguieron las recomendaciones indicadas por el fabricante con ligeras variaciones, donde se utilizó el plásmido pcDNA-PRPF40B como molde y parejas de oligonucleótidos que generaban cambios puntuales en los aminoácidos P383L, D442N y P540S. La amplificación se llevó a cabo en 18 ciclos con un gradiente de temperaturas de elongación, comenzando en 45°C y aumentando 5°C en cada ciclo hasta alcanzar una temperatura de 68°C. A continuación, el producto de la PCR fue digerido con 1 µl de DpnI, enzima que degrada el ADN metilado, quedando sólo el producto mutado amplificado. Finalmente, se transformó la mezcla en células competentes proporcionadas por el *kit* y se comprobaron las mutaciones puntuales mediante secuenciación.

3.5 Ensayos de actividad luciferasa

Para ensayos de actividad luciferasa se sembraron $2,5 \times 10^5$ células en pocillos de 35 mm. Al día siguiente cuando las células alcanzaron un 80% aproximado de confluencia, se cotransfectaron usando el método de Ca_2PO_4 . Las cantidades de ADN transfectadas por punto fueron las siguientes: 200 ng del plásmido reportero de la activación de la transcripción p3xMHC-LUC, 25 ng del plásmido reportero pG5HIV o 200 ng del plásmido reportero pG5NFI+II; 5 ng del vector de expresión de *Renilla* pRL-TK-REN como normalizador de la transfección y 2 µg de las correspondientes construcciones de PRPF40B. Las cantidades de ADN de todos los puntos se igualaron con ARNt.

Aproximadamente 40 h después de la transfección las células se recogieron en PBS 1X pre-enfriado y se lisaron en 200 µl del tampón T7 frío (20mM HEPES, pH 7.9; 150mM NaCl, 5mM EDTA; 1% NP-40, 1mM ditioneitol (DTT), 1mM PMSF y mezcla de inhibidores de proteasas (Complete, Roche)) durante 30 min a 4°C. Tras eliminar los restos celulares por centrifugación, un décimo de la muestra se mezcló con el mismo volumen de tampón Laemli 2x para su análisis por *Western Blot*. La actividad luciferasa y *renilla* se midió utilizando el kit comercial *Dual Luciferase Assay* (Promega)

siguiente las instrucciones del fabricante y un luminómetro (Berthold Detection Systems).

3.6 Ensayos de interacción proteína-proteína *in vivo*

Para ensayos de inmunoprecipitación se sembraron 2×10^6 células en placas de 100 mm de diámetro (BD Falcon). Al día siguiente y cuando la confluencia alcanzaba el 90%, las células se transfectaron con 6 μg de PRPF40B o SF1 con el método de Ca_2PO_4 . Las cantidades de ADN se igualaban en todos los puntos utilizando ARNt.

Aproximadamente 40 h después de la transfección las células fueron recogidas en 2 ml de PBS 1X frío (4°C) y lisadas en 500 μl de tampón T7 frío (20mM HEPES, pH 7,9; 150mM NaCl, 5mM EDTA; 1% NP-40, 1mM DTT, 1mM PMSF e inhibidores de proteasas (Complete, Roche)) durante 30 min a 4°C. Tras eliminar los restos celulares por centrifugación se mezcló un décimo de cada muestra con tampón Laemli 2x y se guardó convenientemente para su análisis por *western blot* como muestra del extracto celular total (*whole cellular extract, WCE*). A continuación, los lisados fueron diluidos hasta 1 ml con tampón T7 e incubados durante 4-6 h en rotación a 4°C con 50 μl de agarosa unida covalentemente a un anticuerpo monoclonal anti-T7 tag (Novagen). Tras cinco lavados con el tampón T7, las resinas fueron hervidas durante 7 min a 95°C mezcladas con 50 μl de tampón Laemli 2x en agitación. La mitad de la muestra se analizó por *Western Blot*.

Para el tratamiento con ARNasa, los lisados celulares se trataron con 0,1 mg/ml de ARNasa A (Roche) durante 1 h a 4°C en rotación antes de la incubación con la agarosa.

Para el bloqueo de los U1 snRNA utilizando oligonucleótidos antisentido Morpholino (AMO, Gene Tools), las células fueron transfectadas aproximadamente 24 h después de la transfección del ADN con 3 μM del U1 AMO o del control CTRL AMO, utilizando el reactivo de transfección *EndoPorter* (Gene Tools), siguiendo las indicaciones del fabricante. Aproximadamente 16 h después, las células se recogieron, utilizando un quinto de la muestra para la extracción de ARN y el resto para su análisis por inmunoprecipitación, como se describió anteriormente.

3.7 Ensayos de interacción proteína-proteína *in vitro*

Para el estudio de la interacción *in vitro* de la proteína PRPF40B con el factor de *splicing* SF1 se siguió un protocolo basado en trabajos anteriores (Rain et al., 1998). Se unió 1 µg de cada proteína GST recombinante purificada (GST, GST-PRPF40B, GST-PRPF40B (1-333), GST-PRPF40B (446-871), GST-TCERG1Nt (235-662), GST-TCERG1Ct (631-1098) y GST-U2AF⁶⁵) a 50 µl de agarosa acoplada a glutatión (Sigma) durante 30 min a 4°C en 300 µl de tampón NETN (20mM Tris-HCl, pH 8; 100mM NaCl; 0,5% NP-40; 0,5mM EDTA) en rotación. Tras dos lavados con 500 µl de NETN, se añadió 1 µg de proteína T7-SF1 purificada y se incubó durante 30 min a 4°C en un volumen de 200 µl de NETN en rotación. Seguidamente las muestras fueron lavadas seis veces con 200 µl de NETN y hervidas durante 5 min a 99°C con agitación en 50 µl de tampón Laemli 2x, para la elución de los complejos proteicos unidos. Parte de la muestra se analizó por *Western Blot* y parte por tinción de plata.

Para el tratamiento con ARNasa, las resinas de agarosa fueron incubadas durante 30 min a temperatura ambiente con 10 µg/ml de ARNasa A en 200 µl de tampón NETN en rotación antes de los tres últimos lavados. Posteriormente se lavó otras tres veces y se procedió a la elución por calor durante 5 min a 99°C en agitación junto con el tampón Laemli.

Para el estudio de la fuerza de unión proteína-proteína, el ensayo se hizo aumentando la concentración de NaCl a 150mM, 200mM o 300mM del tampón NETN durante todo el proceso.

4. PROCEDIMIENTOS DE BIOLOGÍA CELULAR

Para los ensayos de biología celular y análisis mediante citometría de flujo, se sembraron 9×10^4 células por pocillo de M6 y se transfectaron con 60 nM de siRNAs. 72 h después de la transfección, las células eran procesadas para su análisis.

Para el análisis estadístico se usó el programa *Prism 5.0 software* (GraphPad). Se aplicó el test *t-Student* para comparar las medias entre muestras. Los valores *P* están representados en las gráficas mediante asteriscos (*, $P < 0,05$; **, $P < 0,01$).

4.1 Ensayos de muerte celular y ciclo celular

Las células HEK293T fueron tripsinizadas, recolectadas en tubos de poliestireno y lavadas dos veces con PBS 1X centrifugando a 1600 rpm 4°C. Para los ensayos de muerte celular, la mitad de las células fueron incubadas en una solución de PBS 1X con yoduro de propidio (PI, Sigma) 40 µg/ml a 37°C durante 20 min en oscuridad. Para los análisis de ciclo celular, la otra mitad de células fueron fijadas con etanol 70% a 4°C durante al menos 15 min. Tras un par de lavados con PBS 1X, las células fueron incubadas durante 20 min a 37°C en oscuridad, en una solución de PBS 1X que contenía 100 µg/ml de ARNasa A (Roche) y 40 µg/ml de PI. Las muestras fueron analizadas por citometría de flujo (FACScalibur, BD). Los datos fueron procesados utilizando el software FlowJo v7.6.5.

4.2 Ensayo de unión de Anexina V

Las células HEK293T fueron tripsinizadas, recolectadas en tubos de poliestireno y se lavaron con PBS 1X centrifugando a 1600 rpm. Aproximadamente 1×10^6 células/ml se resuspendieron en 100 µl de una solución de tampón de unión Anexina V que contenía 5 µl de Anexina V conjugado a isotiocianato de fluoresceína (FITC) y 5 µl de PI durante 15 min a temperatura ambiente en oscuridad. Posteriormente, se añadió 400 µl del tampón de unión Anexina V. Las muestras fueron analizadas por citometría de flujo utilizando el software FlowJo v7.6.5. El análisis estadístico fue realizado utilizando el software Prism 5.0.

4.3 Medida de la actividad de la caspasa-3

Aproximadamente 1×10^4 de células HEK293T fueron recogidas para medir la actividad de la caspasa-3. La cuantificación se realizó utilizando el kit comercial *Caspase-Glo 3/7 assay* (Promega) siguiendo las indicaciones del fabricante. 40 µl de cultivo se mezclaron con 40 µl de reactivo en placas de 96 pocillos y se incubó durante 10 min a temperatura ambiente y en oscuridad. A continuación se midió la señal de quimioluminiscencia (unidades de luz relativa, RLU), que es directamente proporcional a la activación de la caspasa, en un luminómetro (Infinite F200 TECAN Luminometer).

4.4 Análisis de expresión del receptor de membrana Fas/CD95

Las células HEK293T fueron tripsinizadas, recogidas en tubos de poliestireno y lavadas con PBS 1X centrifugando a 1600 rpm. A continuación, las células fueron incubadas con el anticuerpo primario anti-Fas/CD95 en PBS 1X durante 30 min a 4°C en oscuridad. Tras dos lavados con PBS 1X, las células fueron incubadas con el anticuerpo secundario correspondiente durante 30 min a 4°C en oscuridad. Después las células fueron lavadas con PBS 1X y la fluorescencia se midió en el citómetro de flujo FACSCalibur. Los datos fueron analizados utilizando el software Cell Quest Pro v1.b.1f4b.

5. MANIPULACIÓN Y ANÁLISIS DE PROTEÍNAS

5.1 Análisis por *Western Blot*

Las distintas muestras a analizar por Western Blot fueron diluidas en tampón de carga Laemli 2x (0,2% azul de bromofenol, 10% glicerol, 0,1M Tris-HCl pH 6,8, 2% dodecil sulfato sódico (SDS) y 50mM DTT) y hervidas a 95°C durante 5 min. Después de cargar las muestras en geles al 10% o 12.5% SDS-PAGE, la electroforesis se llevó a cabo en tampón Laemli (25mM Tris, 192mM de glicina y 0,1% SDS) a un voltaje continuo de 150 V. Como marcador del peso molecular se utilizó una preparación comercial con un rango 175kD-7kD (New England Biolabs).

Una vez las proteínas fueron separadas, se transfirieron a membranas de nitrocelulosa (Hybond, Amersham) con un sistema de transferencia húmedo en una misma solución para el equilibrado de geles y la transferencia (20% (v/v) metanol, 25mM Tris, 192mM glicina) durante 1 h a 4°C con una intensidad constante de 200-300mA. Las membranas se bloquearon durante 2 h en una solución de bloqueo (5% (p/v) leche desnatada, 0,1% (v/v) Tween-20, PBS 1X). La incubación y lavados de los anticuerpos se realizaron con la misma solución de bloqueo. Las diluciones y condiciones de incubación se ajustaron a cada anticuerpo primario. Los anticuerpos secundarios conjugados con peroxidasa (PerkinElmer Life Science) fueron utilizados a una dilución 1:5000 durante 1h a temperatura ambiente. La señal de quimioluminiscencia fue detectada usando el reactivo comercial *Western Lightning Chemiluminescence Reagent Plus* (PerkinElmer Life Science).

El análisis densitométrico se llevó a cabo utilizando el software *Quantity One* v4.6.5 (Bio-Rad). La intensidad de las bandas se calculó normalizando respecto a un control interno (CDK-9) para cada muestra en cada condición.

5.2 Tinción de plata

Las proteínas a analizar por tinción de plata fueron separadas en geles al 10% SDS-PAGE. Posteriormente, los geles se incubaron en una serie de soluciones a temperatura ambiente en agitación. Los geles se sometieron a un proceso de fijación en una solución que contenía 40% (v/v) metanol, 10% (v/v) ácido acético durante 30 min, seguido del lavado del gel con una solución A (10% etanol (v/v), 5% (v/v) ácido acético) durante 5 min. A continuación el gel se sometió a un proceso de sensibilización con una solución de 0,01% tiosulfato de sodio ($\text{Na}_2\text{O}_3\text{S}_2$) durante 5 min realizándose varios lavados con la solución A durante 2 min, 10% etanol durante 10 min y agua destilada 10 min. Para la tinción con plata, el gel se incubó con una solución al 0,1% de nitrato de plata (AgNO_3) durante 10 min en oscuridad. Finalmente, para el revelado se incubó el gel con una solución al 10% de carbonato sódico (K_2CO_3) durante 5 min, seguido de otra solución al 2% de carbonato sódico y 0,01% formaldehído hasta la aparición de bandas. El revelado se detuvo sumergiendo el gel en solución de fijado.

5.3 Tinción de *Coomasie*

Las proteínas a identificar por tinción de *Coomasie* fueron separadas en geles al 10% SDS-PAGE. A continuación los geles se sumergieron durante 20 min en una solución de tinción/fijación TC (0,0025% (p/v) Coomassie Blue R250, 45% (v/v) metanol y 10% (v/v) ácido acético) y se lavaron repetidas veces con una solución LC (25% (v/v) metanol, 7% (v/v) ácido acético). Finalmente los geles se secaron en papel celofán (Bio-Rad) o bien en un aparato de vacío/calor para su posterior análisis y almacenamiento.

5.4 Producción y purificación de proteínas recombinantes en *E. coli*

La expresión heteróloga de proteínas recombinantes para su posterior purificación se realizó en bacterias *E. coli* deficientes en proteasas (BL21 pLys). Para ello se incubó un pre-inóculo en LB 1/20 del volumen final en presencia de ampicilina y

cloranfenicol a 37°C durante 16 h aproximadamente. Al día siguiente, el pre-cultivo era diluido hasta un volumen final de 500 ml en LB fresco con ampicilina e incubado a 37°C hasta que el cultivo alcanzaba una $DO_{600}=0,6$ UA. Entonces se indujo la producción de proteína recombinante al añadir isopropil- β -1-D-tiogalactopiranosido (IPTG; Sigma) a una concentración final de 0,03mM durante 6 h a 25°C para la producción de PRPF40B y de 0,4mM durante 2 h a 37°C para el resto de proteínas. Transcurrido el tiempo de inducción, las bacterias se centrifugaron a 6000 rpm para ser procesadas para la purificación de proteínas. Se tomaron muestras de 40 μ l antes y después de la centrifugación, se centrifugaron a 6000 rpm y se resuspendieron en tampón de carga Laemli 2x para comprobar el resultado de la inducción mediante una tinción con Coomasie.

La purificación de proteínas recombinantes se realizó con un sistema de cromatografía por afinidad en condiciones nativas de glutatión-S-transferasa (GST)/glutatión. Las bacterias se resuspendieron en PBS 1X suplementado con una mezcla de inhibidores de proteasas (Complete, Roche) en un volumen 1/100 del volumen del cultivo de origen. Para la ruptura de la pared bacteriana se adicionó 6 mg/ml de lisozima (USB biochemicals), se mezcló con vórtex y se incubó durante 10 min a temperatura ambiente. A continuación se digirió el ADN con 130 U de ADNasaI (Sigma) en presencia de 10mM $MgCl_2$ y 1mM $MnCl_2$ durante 30 min a 37°C en agitación. Tras centrifugar para eliminar los restos celulares, el lisado se aclaró adicionando una mezcla de 1% (v/v) TritonX-100, 1% (v/v) Tween-20, 1mM DTT, 0,5mM PMSF y PBS 1X y centrifugando a 14000 rpm. A continuación se preparó un volumen de 500 μ l por construcción de resina liofilizada de agarosa acoplada a glutatión (Sigma) según las indicaciones del fabricante. Los lisados bacterianos obtenidos previamente se incubaron con las resinas de agarosa durante 2 h a 4° en rotación. Transcurrido este tiempo, la agarosa se lavó tres veces a 4°C con una solución que contenía 1% (v/v) TritonX-100, 1% (v/v) Tween-20, 0,5mM DTT, 1mM PMSF y PBS 1X. A continuación la resina se equilibró con una solución de 100mM de Tris pH8, 100mM de NaCl, 0,5% de TritonX-100 y 20 μ g/ml de PMSF. La elución de las proteínas unidas se realizó por competición con 20mM de glutatión reducido (Sigma) en una solución que contenía 50mM Tris-HCl, pH8 durante 20 min en rotación a temperatura ambiente. Por último las proteínas se dializaron en dos pasos

durante al menos 16 h a 4°C en rotación, en un volumen 1000 veces mayor al obtenido en la muestra de PBS 1X, 1mM PMSF.

5.5 Producción y purificación de proteínas recombinantes en células HEK293T

Para la producción y purificación de proteína recombinante a partir de células HEK293T se siguió un protocolo basado en un trabajo descrito previamente (Cazalla et al., 2005).

Se sembraron seis placas de 100 mm de diámetro (BD Falcon) de células HEK293T aproximadamente 20 h antes de la transfección. Cuando las células alcanzaron el 80-90% de confluencia, se transfectó 4 µg de pEFBOST7-SF1 por placa mediante el método de Ca₂PO₄. Aproximadamente 48 h después de la transfección las células se recogieron en PBS 1X frío, se lisaron en un tampón frío (50mM tampón NaP, pH 8, 0,5M NaCl, 5mM β-glicerolfosfato, 5mM KF, 0,1% (V/V) Tween-20 y mezcla de inhibidores de proteasas (Complete, Roche)) y se sonicaron cinco veces durante 30 s en hielo (*vibra-cell*, Sonics materials). Para descartar los restos celulares, los lisados fueron centrifugados a la máxima velocidad (13000 g) durante 20 min. A continuación se preparó 1 ml de resina de agarosa unida covalentemente a un anticuerpo monoclonal anti-T7 tag (Novagen) en una columna de cromatografía (Bio-Rad) para efectuar la purificación de la proteína a temperatura ambiente. La columna se equilibró con 10 ml de tampón de lisis. A continuación se pasó el lisado celular dos veces, seguido de un lavado con 10 ml de tampón de lisis. La proteína se eluyó con 8 ml de tampón de elución (0,1M ácido cítrico pH 2,2, 5mM β-glicerofosfato, 5mM KF) de la cual se fueron recogiendo en 10 fracciones diferentes de 0,8 ml en tubos donde previamente se habían añadido 0,2 ml de solución Tris-HCl pH 8,8 para neutralizar el pH ácido del eluido. Se tomaron muestras de 30 µl del lisado y solución de lavado tras pasar por la columna, así como de cada una de las fracciones del eluido. Se mezclaron con tampón de carga Laemli 2x y se resolvieron en geles al 10% SDS-PAGE para identificar la proteína purificada con tinción de *Coomasie* y posteriormente analizarla por *Western blot*. Finalmente, las fracciones con mayor concentración de proteína se dializaron en un volumen 1000 veces mayor al obtenido en la muestra en un tampón de diálisis BC100 (20mM Tris pH 8, 100mM KCl, 0,2mM EDTA pH 8 y 20% glicerol) en dos pasos durante al menos 16 h a 4°C en rotación. La proteína dializada se analizó de nuevo por tinción de *Coomasie*.

5.6 Producción de anticuerpo anti-PRPF40B en conejo

La producción de sueros reactivos contra PRPF40B en conejos se realizó en colaboración con la Unidad de Experimentación Animal del IPBLN-CSIC.

Como antígeno se utilizó la proteína recombinante GST-PRPF40B, que incluía los aminoácidos (1-871) de la secuencia de PRPF40B en una solución de PBS 1X. Los conejos se inmunizaron cuatro veces en los días 0, 14, 35 y 56 con una emulsión de 0,5 mg de la proteína en una relación 1:1 con una solución de Adyuvante de Freund completo en la primera inmunización e incompleto en las otras tres inmunizaciones, para aumentar la reactividad del antígeno.

Para testar la presencia de anticuerpos policlonales contra la proteína PRPF40B, se tomaron muestras de sangre al día 69 y se compararon con las muestras pre-inmunes tomadas el día 0 de inmunización. Para la obtención del suero a partir de la sangre se inactivó el complemento por incubación a 37°C durante 30 min y las muestras se centrifugaron durante 15 min a 5000g para eliminar los restos celulares. Se tomaron muestras de los sobrenadantes y se mezclaron con tampón de carga Laemli 2x para su posterior análisis por *Western Blot*. El día 75 se procedió al sangrado final de los conejos para obtener todo el suero reactivo producido contra PRPF40B. La obtención y análisis del suero se realizó del mismo modo que las muestras tomadas anteriormente. Finalmente los sueros fueron testados por *Western Blot* para comprobar su eficiencia y especificidad con extractos celulares de células HEK293T, extractos nucleares de células HeLa y las proteínas recombinantes.

5.7 Purificación de IgG totales del suero anti-PRPF40B

Fracciones de suero reactivo contra PRPF40B fueron purificadas mediante cromatografía de afinidad en una columna con sepharosa conjugada a proteína A (GE Healthcare Life Science), para obtener un concentrado de IgGs totales. Para ello se preparó 1 ml de resina de agarosa lavada previamente en 0,1M Tris-HCl pH 8, en una columna de cromatografía (Bio-Rad) equilibrada con 10 ml de la misma solución de Tris-HCl. A continuación se cargaron dos veces a una velocidad inferior a 0,3 ml/min, 1 ml de suero de PRPF40B diluido un décimo en Tris 1M pH 8 para igualar el pH y la concentración salina con la columna. Después se realizaron dos lavados secuenciales

con 10 ml de 0,1M Tris-HCl pH 8 y 0,01M Tris-HCl pH 8, respectivamente. Las IgGs se eluyeron al añadir 3 ml de 0,1M Glicina pH 3 a una velocidad menor a 0,3 ml/min, y se recogieron en fracciones de 270 μ l en tubos donde previamente se habían puesto 30 μ l de 1M Tris-HCl pH 8 para neutralizar el pH ácido del eluido. Finalmente, se midió la concentración de proteínas de las fracciones a partir de su absorbancia a 280 nm, y se dializaron en un volumen 1000 veces mayor que la cantidad de la muestra de PBS 1X y 0,1% de azida sódica en dos pasos secuenciales de 2 y 16 h respectivamente.

6. ADN Array

El análisis de los cambios de expresión a nivel de ARN se realizó en la *Genome Core Facility-EMBL (Heidelberg)*. Tres muestras de 1 μ g de ARN procedente de células SH-SY5Y establemente silenciadas para PRPF40B o con un shRNA control se procesaron para medir cambios de expresión de forma global en el transcriptoma humano utilizando el *Human Transcriptome Array (HTA) 2.0* de Affymetrix.

Los cambios de expresión detectados en el HTA entre las muestras problema y control, fueron analizados por Pierre de la Grange (GenoSplice, Paris, [www.genosplice.com]), utilizando la interfaz EASANA, un módulo de visualización basado en las anotaciones disponibles en la base de datos genómica FAST DB. Este procesamiento conlleva una serie de pasos que se detallan a continuación (de la Grange et al., 2010): los datos procedentes del *array* son normalizados y seleccionados atendiendo a criterios de calidad de intensidad o su anotación como genes en FAST DB. A continuación, se identifican los exones expresados diferencialmente donde se comparan las muestras problemas o las muestras control entre ellas, utilizando la estrategia estadística del llamado índice de *splicing*, basado en el cociente de la intensidad de la sonda y el nivel de expresión de un gen. Se considera que un exón tiene una expresión diferencial cuando sus sondas muestran una expresión similar. Después se comparan los valores de intensidad asociados a cada exón entre las muestras control y las silenciadas utilizando de nuevo la estrategia del índice de *splicing*. Finalmente, el punto de corte establecido para comparar la expresión diferencial entre los dos grupos de muestras se basa en el valor *P*.

7. OLIGONUCLEÓTIDOS Y ANTICUERPOS

7.1 Oligonucleótidos de ADN

Los oligonucleótidos descritos en la siguiente tabla fueron sintetizados en la *Unidad de Genómica y Síntesis de DNA* del IPBLN-CSIC.

Nombre del oligonucleótido	Secuencia 5'-3'
Clonajes de PRPF40B en pEFBOST7	
<i>HYPC1Fw</i>	5'-GGGTCTAGAATGATGCCACCACCCTTC-3'
<i>T7HYPCRv</i>	5'-GGGATCGATTCACTGGTGATCATCCAGCTG-3'
<i>HYPC687Rev</i>	5'-GGGATCGATTCAATTCAGTCTGCTCCAG-3'
<i>HYPC263Fw</i>	5'-GGGTCTAGAATGAAGCCAGAACCAGAG-3'
<i>HYPGFPct</i>	5'-GGGTCTAGAATGCTCTTCTTCCTGGCC-3'
<i>HYPC474Fw</i>	5'-GGGTCTAGAATGGAAGAGGAGGAGGAACGG-3'
<i>HYPC618Fw</i>	5'-GGGTCTAGAATGCGGGAGAAGGAGGAG-3'
Clonaje de PRPF40B en pcDNA3/FLAG/MS2	
<i>HYPCMS2Fw</i>	5'-GGGCCTAGGATGATGCCACCACCCTTC-3'
<i>HYPCMS2Rv</i>	5'-GGGATCGATTCACTGGTGATCATCCAGCTG-3'
Clonajes de PRPF40B en pmCherry	
<i>HYPCBSrGIFw</i>	5'-GGGCTGTACATTATGATGCCACCACCCTTC-3'
<i>HYPCBSrGIRv</i>	5'-GGGCTTGTACATCTGGTGATCATCCAGCTG-3'
<i>HYPCAgeIFw</i>	5'-GGGACCGGTATGATGCCACCACCC-3'
<i>HYPCAgeIRv</i>	5'-GGGGACGGGTGCCTGGTGATCATC-3'
Clonajes de PRPF40B en pGEX-2TK	
<i>GSTHYPCFw</i>	5'-GGGAGATCTATGCCACCACCCTTCATGCC-3'
<i>GSTHYPCRv</i>	5'-GGGGAATTCTCACTGGTGATCATCCAGCTG-3'
<i>GSTHYPC333Rev</i>	5'-GGGGAATTCTCACTTCTCCCGCTGCGC-3'
<i>GSTHYPCctFw</i>	5'-GGGAGATCTATGTTTGCTCAGGACCATCAG-3'
Clonaje de SF1 en pEFBOST7	
<i>SF1FLT7Fw</i>	5'-GGGTCTAGAATGGCGACCGGAGCGAAC-3'
<i>SF1FLT7Rv</i>	5'-GGGTTCGAATCAGGGTGCTTTGCCGAA-3'
Mutagénesis dirigida de PRPF40B en pcDNA3-PRPF40B	
<i>P383LFw</i>	5'-CTGGGCTGTGGTCCTTGAGAGGGATCGAA-3'
<i>P383LRv</i>	5'-TTCGATCCCTCTCAAGGACCACAGCCCAG-3'
<i>D442NFw</i>	5'-CCAGCAGTACCTCATGAATAACCCAGCTTTGC-3'
<i>D442NRv</i>	5'-GCAAAGCTGGGGTTATTCATGAGGTACTGCTGG-3'
<i>P540SFw</i>	5'-AGCCGGGCTCCACCTCTCTGGACTTATTCA-3'
<i>P540SRv</i>	5'-TGAATAAGTCCAGAGAGGTGGAGCCCGGCT-3'

Retro-transcripción del ARN y PCR convencional del minigen *Fas*

<i>csnPT1</i>	5'-GTTCGACACTTGCTCAAC-3'
<i>csnPT2</i>	5'-AAGCTTGCATCGAATCAGTAG-3'

Retro-transcripción del ARN del minigen *HIV-X2*

<i>RT-Sveda</i>	5'-GGGAAGCTAGAGTAAGTAG-3'
-----------------	---------------------------

PCR convencional del minigen *HIV-X2*

<i>X34</i>	5'-AGGGAGGCAGGCGACGGCGACGAGTTT-3'
<i>XAgeIR</i>	5'-GTGGATCCCCCGGGCTGCAGGAATTCGAT-3'

PCR convencional del gen *Fas*

<i>FasE5Fw</i>	5'-GTGAACATGGAATCATCAAGG-3'
<i>FE7</i>	5'-TCCTTTCTGTGCTTTCTGCAT-3'

PCR cuantitativa del gen *Fas*, isoforma 6-7

<i>FE6</i>	5'-TAACTTGGGGTGGCTTTGTC-3'
<i>FE7</i>	5'-TCCTTTCTGTGCTTTCTGCAT-3'

PCR cuantitativa del gen *CACNG2*

<i>CACNG2Fw</i>	5'-CATCTTCTTCGTGTCTGCAG-3'
<i>CACNG2Rv</i>	5'-GACCAGCCGTATGAGTAACT-3'

PCR cuantitativa del gen *TNFAIP6*

<i>TNFAIP6Fw</i>	5'-AGGAGTGAAAGATGGGATGC-3'
<i>TNFAIP6Rv</i>	5'-ACGCTGACCATACTTGAGTC-3'

PCR cuantitativa del gen *SVC2*

<i>SVC2Fw</i>	5'-CTGCATGTTCTGGATGATCG-3'
<i>SVC2Rv</i>	5'-CCAACTGTGAACTGGTAGG-3'

PCR cuantitativa del gen *LEF1*

<i>LEF1Fw</i>	5'-GATCTCTTTCTCCACCCATC-3'
<i>LEF1Rv</i>	5'-GACATCTGATGGGATGTGTG-3'

PCR cuantitativa del gen *CCNA1*

<i>CCNA1Fw</i>	5'-TCACTGATAGCTGCAGCAGC-3'
<i>CCNA1Rv</i>	5'-TGAAGCTCACTCAGGCAAGG-3'

PCR cuantitativa del gen *RUNX1T1*

<i>RUNX1T1Fw</i>	5'-GTCTGGATACGTGCCAGAGG-3'
<i>RUNX1T1Rv</i>	5'-CCTTCTGCAGCTCCGTCATC-3'

PCR cuantitativa del gen *TCF7*

<i>TCF7Fw</i>	5'-GTTCTGGTGTACCTGGTCAC-3'
<i>TCF7Rv</i>	5'-CCTTGGACTCTGCTTGTGTC-3'

PCR cuantitativa del gen *LHX9*

LHX9Fw 5'-CTGCCTTACTTCAACGGTAC-3'
LHX9Rv 5'-CAAGTGGTCTGCCTCATTC-3'

PCR cuantitativa del gen *GBP4*

GBP4Fw 5'-GTGAGAGAACTCTTCACGCTG-3'
GBP4Rv 5'-CACTGTCAGCTGCTCTTCCT-3'

PCR cuantitativa del gen *PIK3R2*

PIK3R2Fw 5'-GATCCGAGACCAGTACCTCG-3'
PIK3R2Rv 5'-GTCCTCGTCCCTCCATGAGTG-3'

PCR cuantitativa del gen *NNAT*

NNATFw 5'- ACCATTCTTGGAAACCATGG-3'
NNATRv 5'- TGCAGCATTCCAGGAACAC-3'

PCR cuantitativa del gen *KCNMA1*

KCNMA1E21Fw 5'-CTACTTGGAAGGAGTCTCA-3'
KCNMA1E22Rv 5'-GGCAGACTTGTACTCAATG-3'
KCNMA1E38Fw 5'-CTACTTGGAAGGAGTCTCA-3'
KCNMA1E39Rv 5'-GGCAGACTTGTACTCAATG-3'

PCR cuantitativa del gen *NRG1*

NRG1E4Fw 5'-CAAGTGGTTCAAGAATGGG-3'
NRG1E5Rv 5'-CAGAATCAGCCAGTGATGC-3'
NRG1E17Fw 5'-GTAACAGCTTCCTCAGGCA-3'
NRG1E19Rv 5'-TGGTCATGGCTGACACATAC-3'

PCR cuantitativa del gen *SLC6A11*

SLC6A11E1Fw 5'-CAACAAGGTGGAGTTCGTG-3'
SLC6A11E2Rv 5'-CACCACGTAGGGAATCAGG-3'
SLC6A11E5Fw 5'-CCTGGACCATCTGTTACTTC-3'
SLC6A11E6Rv 5'-GGATCAGCAGCATGATGTAG-3'

PCR cuantitativa del gen *SMARCA2*

SMARCA2E26Fw 5'-CATCTTGGAGCATGAGGAG-3'
SMARCA2E27Rv 5'-CTTCTTCTCGTCGAGCAATC-3'
SMARCA2E36Fw 5'-CAGCTCAGTGAAGTCTTCA-3'
SMARCA2E37Rv 5'-GAAGCATGACATCCTTCTC-3'

7.2 Oligonucleótidos de ARN (siRNA)

Los siRNAs que se describen a continuación han sido sintetizados por Sigma Aldrich.

siRNAs para el silenciamiento transitorio de PRPF40B

<i>siPRPF40B</i>	5'-GCAGUUCUGGACAGCAUCA-3'
<i>siPRPF40B_as</i>	5'-UGAUGCUGUCCAGAACUGC-3'
<i>siPRPF40B2</i>	5'-AAUCAGAGACCACCAGCUAUCUU-3'
<i>siPRPF40B2_as</i>	5'-AAGAUAGCUGGUGGUCUCUGAUU-3'

siRNAs para el silenciamiento transitorio de PRPF40A

<i>siPRPF40A1</i>	5'-GUGUGGAAGUAUCCAGUAA-3'
<i>siPRPF40A1_as</i>	5'-UUACUGGAUACUCCACAC-3'
<i>siPRPF40A2</i>	5'-CAGCUCAGUGAAGUCUUCA-3'
<i>siPRPF40A2_as</i>	5'-GAAGCAUGACAUCUUCUC-3'

siRNA para el silenciamiento transitorio de PTB

<i>siPTB1</i>	5'-AAGCGGGGAUCUGACGAGCUU-3'
<i>siPTB1_as</i>	5'-AAGCUCGUCAGAUCCCCGCUU-3'

siRNA control para el silenciamiento transitorio

<i>siEGFP</i>	5'-CUACAACAGCCACAACGUC-3'
<i>siEGFP_as</i>	5'-GACGUUGUGGCUGUUGUAG-3'

7.3 Oligonucleótidos antisentido *Morpholino* (AMO)

Los oligonucleótidos antisentido *Morpholino* utilizados han sido descritos anteriormente (Kaida et al., 2010) y fueron sintetizados por *Gene Tools*.

AMO para el bloqueo de U1 snRNA

<i>U1 AMO</i>	5'-GGTATCTCCCCTGCCAGGTAAGTAT-3'
---------------	---------------------------------

AMO control

<i>CRTL AMO</i>	5'-CCTCTTACCTCAGTTACAATTTATA-3'
-----------------	---------------------------------

7.4 Anticuerpos

Los anticuerpos utilizados en este trabajo son los que se detallan a continuación.

Anticuerpo	Especie	Origen	Dilución
α PRPF40B	Conejo	Suero de conejo	WB 1:10,000
α PRPF40B	Conejo	IgG purificado	IF 1:1000
α T7	Conejo	Comercial: Bethyl (#A190-116A)	WB 1:20,000 IF 1:1000
α ERK2	Conejo	Comercial: Santa Cruz Biotechnology (#sc-154)	WB 1:2000
α SF1	Conejo	Cedido por el Dr. Juan Valcárcel (CRG, Barcelona)	WB 1:2000
α SF1	Ratón	Comercial: Abnova (#H00007536-M01A)	WB 1:100,000 IF 1:40000
α U2AF ⁶⁵	Ratón	Cedido por el DR. Juan Valcárcel (CRG, Barcelona)	WB 1:2000 IF 1:100
α U2AF ⁶⁵	Conejo	Comercial: Santa Cruz Biotechnology (#sc-48804)	WB 1:2000
α CDK9	Conejo	Comercial: Santa Cruz Biotechnology (#sc-484)	WB 1:2000
α SC-35	Ratón	Comercial: Sigma (#S4045)	IF 1:4000
α NOLC1	Conejo	Comercial: Abcam (#ab113217)	IF 1:500
α PTB	Conejo	Cedido por el DR. Juan Valcárcel (CRG, Barcelona)	WB 1:2000
α Flag	Conejo	Comercial: Sigma (#F7425)	WB 1:2000
α Caspasa-3	Conejo	Comercial: Santa Cruz Biotechnology (#sc-7148)	WB 1:500

α FLIP	Ratón	Comercial: Santa Cruz Biochechnology (#sc-5276)	WB 1:500
α Bcl2	Ratón	Comercial: Santa Cruz Biochechnology (#sc-7682)	WB 1:500
α Fas/CD95	Ratón	Comercial: BioLegend (#305613). Cedido por M ^a Carmen Ruiz (CIBM, Granada)	CF 1:100
α IgGRabbit-HRP	Cabra	Comercial: Perkin Elmer Life Science (#NEF812001EA)	WB 1:5000
α IgGMouse-HRP	Cabra	Comercial: Perkin Elmer Life Science (#NEF822001EA)	WB 1:5000 CF 1:5000
α IgGRabbit-Alexa488	Cabra	Comercial: Molecular Probes (#A-11008)	IF 1:500
α IgGMouse-Alexa647	Cabra	Comercial: Molecular Probes (#A-21235)	IF 1:500
α IgGRabbit-Alexa647	Cabra	Comercial: Life Technologies (#A-21244)	IF 1:500

WB: Western blot

IF: Inmunofluorescencia

CF: Citometría de flujo

Resultados

1. ANÁLISIS ESTRUCTURAL DE PRPF40B

Uno de los objetivos de este trabajo era contribuir a la caracterización bioquímica de la proteína PRPF40B, identificando los principales motivos estructurales que condicionan su localización subcelular y funcionalidad. Como ya se describió en la *Introducción* (Apartado 4.2.2), PRPF40B fue inicialmente identificada como miembro de una familia de factores que contienen repeticiones en tándem de dominios WW y FF (Figura R-1A). Para estudiar estos motivos de la secuencia de PRPF40B en profundidad se realizaron alineamientos múltiples utilizando diversas herramientas bioinformáticas (*InterPro*, *Pfam*, *SMART*, *PROSITE*), identificando en la secuencia primaria de PRPF40B la presencia de dos dominios WW en su extremo amino-terminal y cinco dominios FF en la mitad carboxilo-terminal (Figura R-1B).

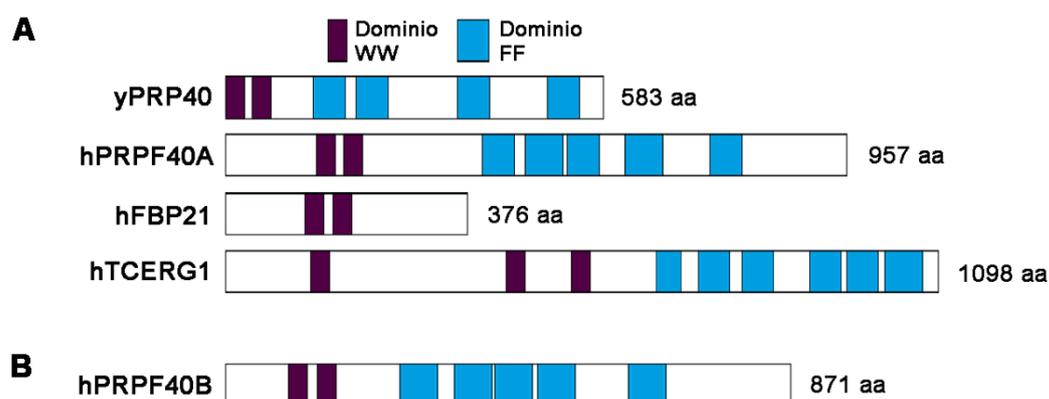


Figura R-1. Los dominios WW y FF de PRPF40B se encuentran en una proporción y distribuidos de forma semejante al resto de miembros de la sugerida familia. A. Diagrama esquemático de la familia de proteínas con dominios WW y FF, que incluye las proteínas PRP40 (levadura), PRPF40A (humano), FBP21 (humano) y TCERG1 (humano). En morado se indican los dominios WW y en azul los dominios FF. **B.** Diagrama esquemático de PRPF40B donde se representan los dominios WW (morado) y FF (azul).

Si observamos la estructura de PRPF40B a lo largo de la evolución, se aprecia un gran nivel de conservación en especies más cercanas al humano, como el ratón. Sin embargo, a medida que nos alejamos, se van reduciendo el número de dominios FF y la longitud de la proteína (Figura R-2A). La comparación de las secuencias de PRPF40B humana con su ortólogo en ratón muestra un alto grado de conservación, con porcentajes del 95% de identidad y 96.5% de similitud (Figura R-2A). Al repetir el procedimiento con su ortólogo en levaduras, PRP40, comprobamos que la disposición

de los dominios se mantenía aunque había una disminución en los valores de identidad, con un 10% y 30%, y similitud, con un 22% y 45% dependiendo si el alineamiento se realizaba de forma global o local, respectivamente (Figura R-2B). Esta divergencia evolutiva se refleja además en un aumento de la complejidad de la proteína PRPRF40B humana, que tiene una mayor longitud en su secuencia de aminoácidos y contiene cinco dominios FF frente a los cuatro motivos FF en PRP40.

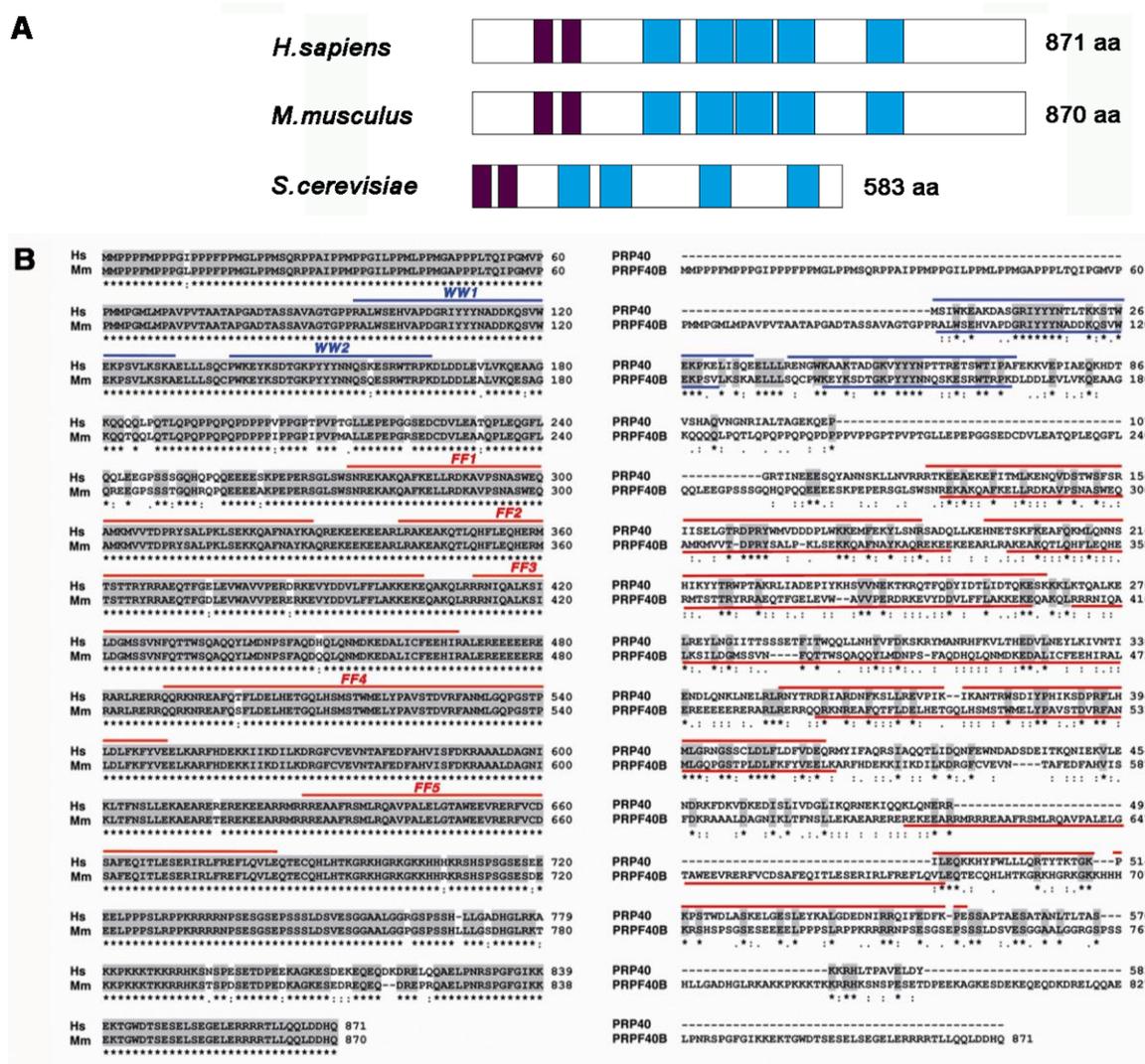


Figura R-2. La distribución y secuencia de los dominios WW y FF de PRPF40B se han conservado gradualmente a lo largo de la evolución. A. Diagrama estructural de conservación de PRPF40B en humano (*H. sapiens*), ratón (*Mus musculus*) y levadura (*S. cerevisiae*). **B.** Alineamiento de la secuencia proteica de PRPF40B de humano con ratón (izquierda) y levadura (derecha). Los aminoácidos idénticos se indican con asterisco (*), aquellos que tienen gran similitud aparecen con dos puntos (:), y los de menor similitud están señalados con un punto (·). Las líneas azules indican las secuencias que comprenden los dominios WW y las líneas rojas señalan las secuencias que corresponden a los dominios FF.

2. ANÁLISIS DE LA LOCALIZACIÓN CELULAR DE PRPF40B

2.1 PRPF40B es una proteína nuclear enriquecida en los *speckles* nucleares

El primer objetivo experimental que nos propusimos sobre la caracterización bioquímica de PRPF40B fue determinar su localización celular. Para ello, purificamos a partir de un suero reactivo de conejo de PRPF40B una fracción altamente enriquecida en IgG totales, cuyo análisis por *western blot* mostraba reactividad hacia una proteína de aproximadamente 100 KDa, correspondiente al tamaño de PRPF40B. El marcaje por inmunofluorescencia de las células HEK293T y HeLa mostró una tinción exclusivamente nuclear, en un patrón difuso nucleoplasmático ausente en el nucléolo y en regiones granuladas donde había un aumento de la señal (Figura R-3A). Esta distribución de la proteína recordaba al patrón de tinción de los *speckles* nucleares, que como ya se describió en la *Introducción* (Apartado 5.1) son orgánulos nucleares situados en el espacio intercromatínico del nucleoplasma de células de mamíferos enriquecidos en factores implicados en el metabolismo del pre-ARNm.

Para determinar si el patrón observado para la proteína PRPF40B endógena coincidía con los *speckles* nucleares, llevamos a cabo ensayos de inmunofluorescencia utilizando anticuerpos específicos dirigidos contra la forma fosforilada del factor esencial de *splicing* SRSF2 (anteriormente conocido como SC35), el cual se utiliza frecuentemente como marcador de los *speckles* nucleares. El marcaje simultáneo de PRPF40B y SRSF2 mostró un solapamiento de las señales en los *speckles* nucleares, tanto en las células HEK293T como en células HeLa (Figura R-3A). Para confirmar estos datos y determinar el grado de colocalización, se realizó un análisis semicuantitativo de la relación espacial entre las distribuciones de PRPF40B y SRSF2 (Figura R-3A).

El antisuero policlonal utilizado se generó inmunizando animales de experimentación con una proteína recombinante (ver *Materiales y métodos*) y detecta bandas de mayor tamaño molecular que PRPF40B por *western blot*. Debido al alto grado de homología de PRPF40B con la proteína PRPF40A, el otro ortólogo de Prp40, existe una alta probabilidad que el anticuerpo esté reconociendo ambas. De hecho según los resultados de los alineamientos globales y locales utilizando EMBOSS, ambas proteínas comparten porcentajes de identidad y similitud comprendidos entre 51-54% y

65-68%, respectivamente. Sin embargo, la obtención de anticuerpos más específicos utilizando en la inmunización péptidos exclusivos de PRPF40B o generando anticuerpos monoclonales en ratón, han sido fallidos (datos no mostrados). Por lo tanto y para corroborar los resultados previos de localización subcelular, se llevaron a cabo ensayos de inmunofluorescencia transfectando con un vector de expresión de la proteína fluorescente EGFP (*enhanced green fluorescent protein*) fusionada a PRPF40B (EGFP-PRPF40B), tanto en células HEK293T como en células HeLa. Las imágenes obtenidas mostraron un patrón de localización de la proteína de fusión similar a los observados anteriormente (Figura R-3B). Estos datos confirman la presencia de una población de PRPF40B que colocaliza con los *speckles* nucleares, sugiriendo una posible implicación de PRPF40B en el metabolismo del pre-ARNm.

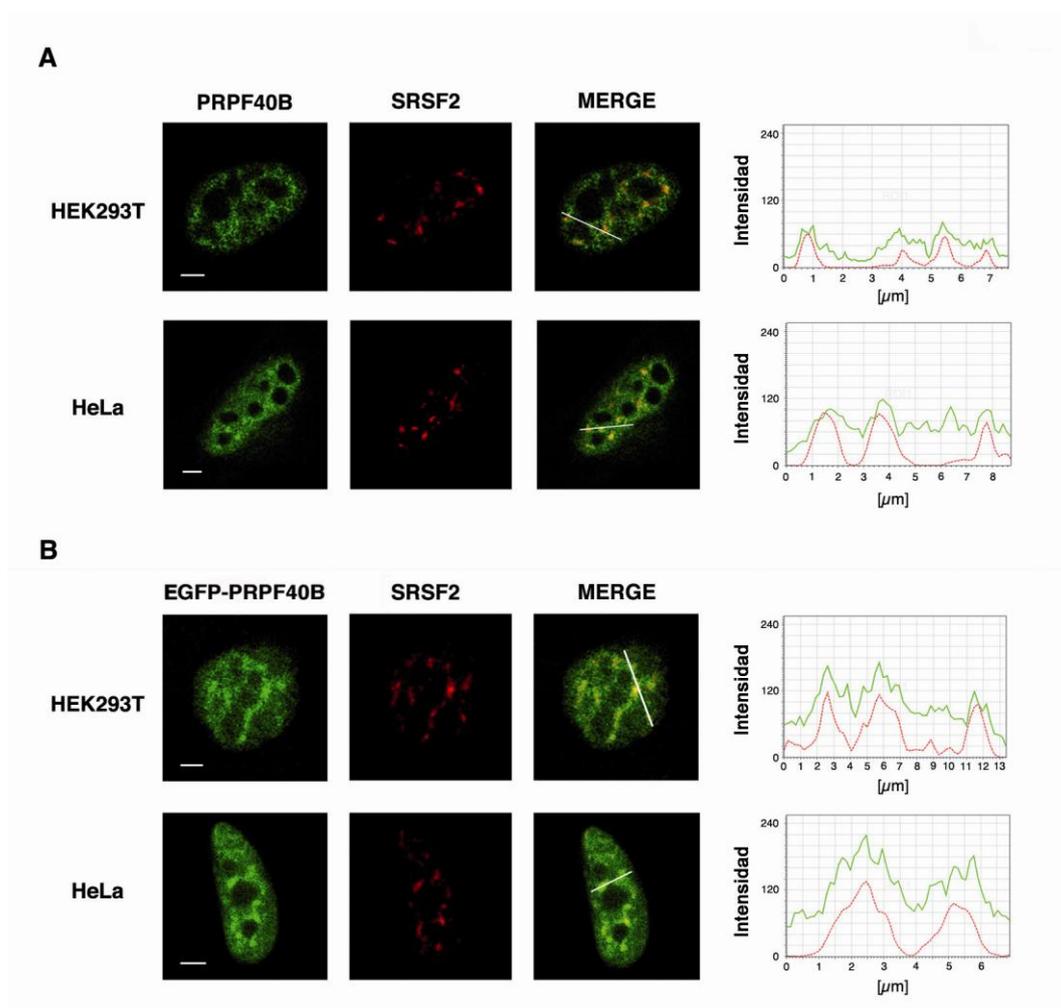


Figura R-3. PRPF40B se encuentra enriquecido en los speckles nucleares. A. PRPF40B endógeno fue detectado utilizando anticuerpos policlonales purificados por IgG. Se muestran imágenes de células HEK293T y HeLa marcadas doblemente con anticuerpos que detectan PRPF40B (verde) y SRSF2 (rojo),

así como la superposición de ambas (*merge*). A la derecha se representa el análisis semi-cuantitativo de la distribución de las intensidades. Las líneas dibujadas en las imágenes superpuestas indican la zona de cuantificación para el análisis. Barra = 3 μm . **B.** Análisis de la localización de PRPF40B mediante ensayos de inmunofluorescencia. Células HEK293T y HeLa fueron transfectadas con la proteína de fusión EGFP-PRPF40B y se tiñeron con anticuerpos que detectaban el marcador de *speckles* nucleares SRSF2 (rojo). Se muestran las imágenes individuales de cada marcaje así como la superposición de ambas (*merge*). Las líneas dibujadas en las imágenes superpuestas representan la región utilizada para el análisis semi-cuantitativo de la intensidad de las señales que se muestran en los paneles de la derecha. Barra = 3 μm .

2.2 Identificación de las secuencias de PRPF40B necesarias para la localización nuclear y enriquecimiento en los *speckles* nucleares

Para determinar el elemento que dirige a PRPF40B hacia el núcleo, se analizó en su secuencia de aminoácidos la presencia de potenciales señales de localización nuclear (NLS, *nuclear localization signal*), utilizando programas bioinformáticos como *WoLF PSORT* (http://www.genscript.com/psort/wolf_psort.html), *NucPred* (<http://www.sbc.su.se/~maccallr/nucpred/>) y *cNLS Mapper* (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi). A partir del estudio realizado se identificó, con la mayor puntuación, una NLS consenso KRRRR comprendida entre los residuos aminoacídicos 732-736 de PRPF40B. Para comprobar la funcionalidad de esta NLS, generamos una construcción truncada en la región carboxilo de PRPF40B fusionada a EGFP, en la cual no se encontraba la posible NLS [Figura R-4, EGFP-PRPF40B (1-687)]. El análisis de la expresión de esta construcción por inmunofluorescencia en las células HEK293T mostró mayoritariamente una localización citoplásmica de la proteína de fusión (Figura R-4A), indicando que la región delecionada contiene la señal responsable de la localización nuclear.

Con el objetivo de identificar de una forma más detallada los dominios estructurales necesarios para la correcta localización de PRPF40B, se crearon un conjunto de construcciones truncadas de PRPF40B fusionadas a EGFP, con deleciones seriadas en el extremo amino-terminal (Figura R-4A). Además, todas las construcciones contenían la NLS identificada en su extremo carboxilo-terminal. Los resultados observados para la construcción PRPF40B sin los dos dominios WW en su región amino-terminal [Figura R-4A, EGFP-PRPF40B (263-871)] reprodujeron el patrón nucleoplasmático, con un aumento de intensidad en los *speckles* nucleares similar a lo

observado con la proteína PRPF40B completa [Figura R-4A, EGFP-PRPF40B (1-871)]. Este resultado indica que los dominios WW no tienen un papel significativo en la localización de PRPF40B en los *speckles* nucleares.

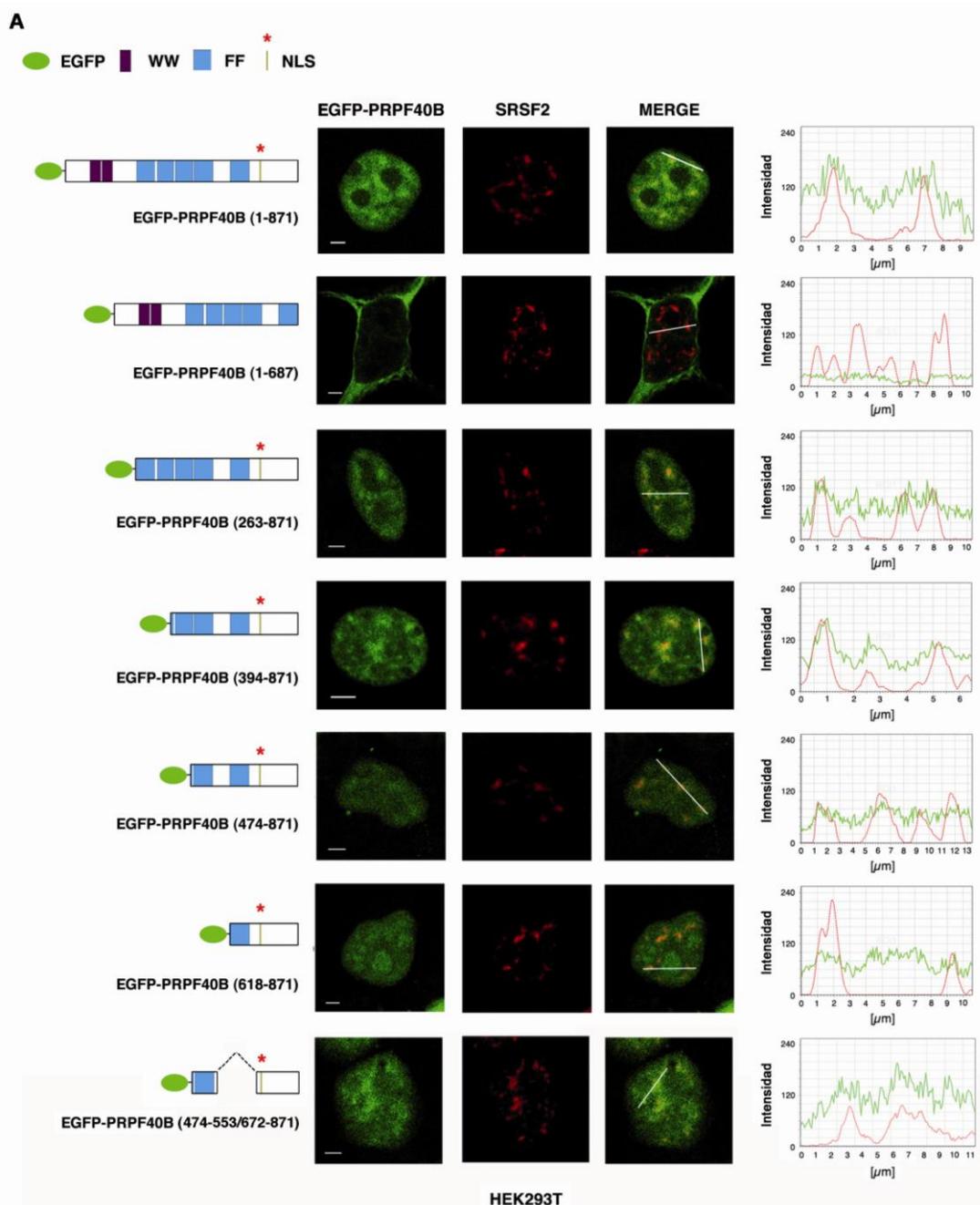
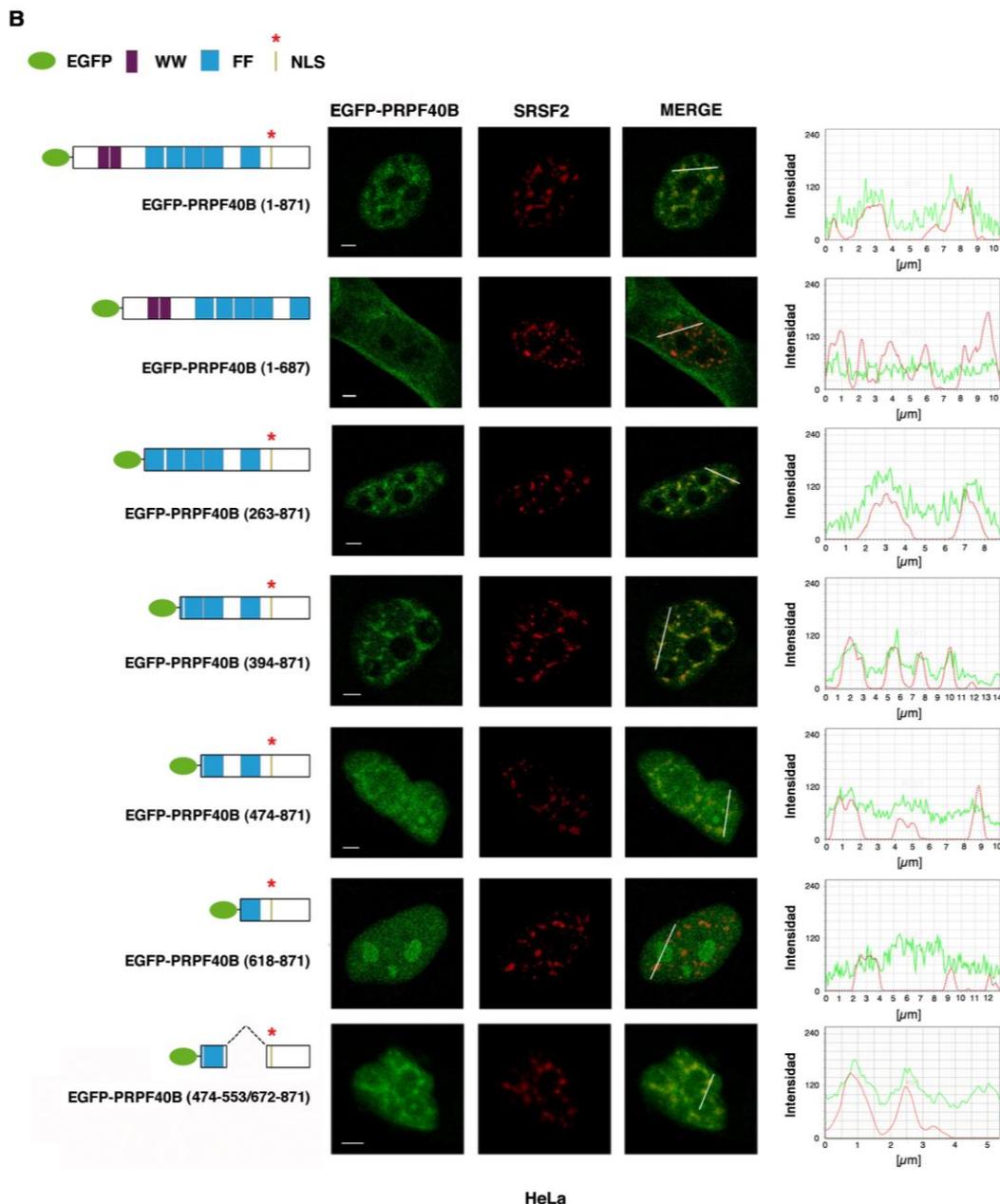


Figura R-4. Identificación de la señal de localización nuclear y a los *speckles*. Las imágenes muestran células HEK293T (A) y células HeLa (B) expresando construcciones EGFP-PRPF40B (verde) marcadas con anticuerpos contra SRSF2 (rojo). El diagrama estructural de cada construcción se representa a la izquierda de cada panel. Los números entre paréntesis indican la región de aminoácidos que contiene cada construcción. En los diagramas se indican la proteína EGFP, los dominios WW, la señal de localización nuclear (NLS) y los dominios FF. En cada panel se muestran las imágenes individuales y superpuestas

(merge). A la derecha están representados los análisis semi-cuantitativos de las intensidades de las señales, cuya región aparece indicada en las imágenes superpuestas. Barra = 3 μm .



Para determinar la importancia de los dominios FF en la localización a los *speckles* nucleares, probamos construcciones de PRPF40B donde se fueron excluyendo gradualmente dichos motivos. La eliminación de los dos primeros dominios FF mantuvo la distribución de PRPF40B, enriquecido en los *speckles* nucleares de forma significativa [Figura R-4A, EGFP-PRPF40B (394-871)]. En cambio, la delección del siguiente dominio FF provocó una pérdida parcial en la acumulación de la proteína expresada en los *speckles* nucleares [Figura R-4A, EGFP-PRPF40B (474-871)], sugiriendo que el dominio FF3 es importante en la correcta localización de PRPF40B en

dichos orgánulos nucleares. La exclusión del penúltimo motivo FF generó una proteína de fusión que se acumulaba de forma significativa en el nucléolo [Figura R-4A, EGFP-PRPF40B (618-871)]. Ante estos resultados, quisimos comprobar si realmente el dominio FF4 tenía capacidad por sí mismo para dirigir a la proteína a los *speckles* nucleares, o si por el contrario la dirigía preferentemente al nucléolo, como ocurría con la construcción FF5. Para ello creamos una nueva construcción donde el dominio FF4 quedaba fusionado a la región carboxilo-terminal que contenía la NLS y la EGFP. Las imágenes obtenidas de esos experimentos revelaron que la construcción de PRPF40B que contenía únicamente el FF4 es capaz de dirigirse a los *speckles* nucleares por sí misma [Figura R-4A, EGFP-PRPF40B (474-553/672-871)]. Para verificar los resultados y comprobar que estos patrones no eran específicos de la línea celular HEK293T, repetimos las inmunofluorescencias en células HeLa, donde se reprodujeron los mismos resultados observados para cada construcción de PRPF40B (Figura R-4B).

Los resultados obtenidos muestran que PRPF40B se distribuye en el nucleoplasma, exceptuando el espacio nucleolar. Sin embargo, a medida que se eliminan los dominios FF de la estructura de la proteína, aumenta la acumulación de la proteína en el nucléolo. Para profundizar en este comportamiento, estudiamos la localización de todas las construcciones de la proteína que se utilizaron previamente, marcando las células con anticuerpos dirigidos contra la forma fosforilada de la proteína nucleolar Nopp140 (Meier and Blobel, 1990). Las imágenes de inmunofluorescencia mostraron que la delección de los dos primeros dominios FF [Figura 5, EGFP-PRPF40B (394-871)] produce un aumento de su señal en el nucléolo, manteniendo también su localización en los *speckles* nucleares. A medida que se iban eliminando más dominios FF, la acumulación de PRPF40B en el nucléolo crecía, mientras su distribución en los *speckles* nucleares era afectada [Figura 5, EGFP-PRPF40B (474-871), [EGFP-PRPF40B (618-871)]. En cambio, cabe destacar que la proteína de fusión que contenía únicamente el FF4 de PRPF40B mostró una señal de intensidad nucleolar más tenue que estas últimas construcciones, y además en el análisis semi-cuantitativo se observa una mayor colocalización con SRSF2 [Figura 5, EGFP-PRPF40B (474-553/672-871)].

Todos estos datos sugieren que la NLS previamente identificada es esencial para la correcta localización nuclear de PRPF40B y sugieren asimismo que los dominios FF participan en el proceso de exclusión nucleolar. Además, las imágenes indican que el dominio FF4 es suficiente para reproducir el patrón subcelular característico de

PRPF40B, enriquecido en los *speckles* nucleares, aunque hay otros dominios FF, como el FF3, que parecen contribuir para que la localización de PRPF40B en esas regiones nucleares se produzca de forma eficiente.

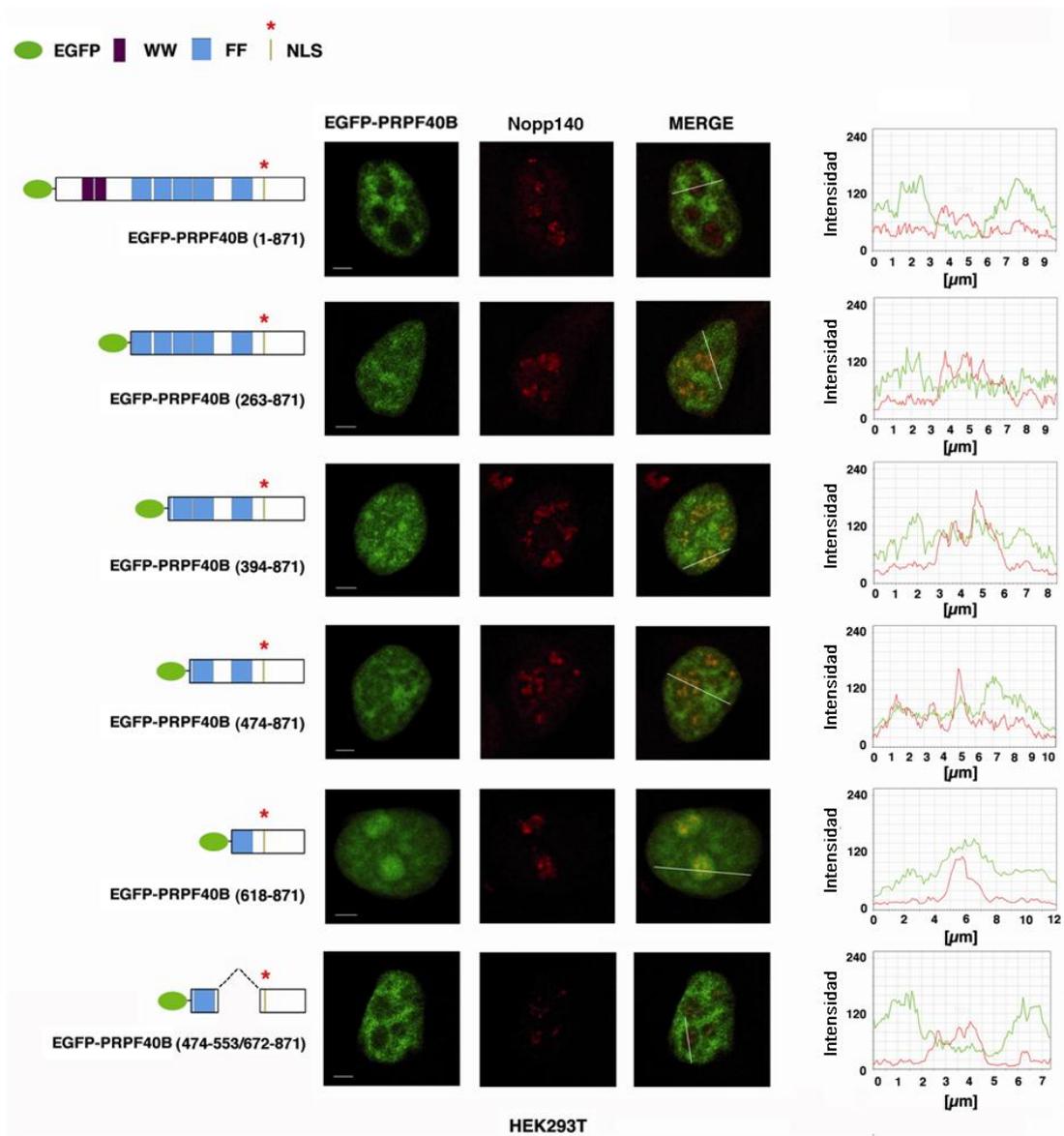


Figura R-5. Papel de los dominios FF en la exclusión nucleolar. Las imágenes muestran células HEK293T expresando construcciones EGFP-PRPF40B (verde) marcadas con anticuerpos contra la forma fosforilada de Nopp140 (rojo). El diagrama estructural de cada construcción se representa a la izquierda de cada panel. Los números entre paréntesis indican la región de aminoácidos que contiene cada construcción. En los diagramas se indican la proteína GFP, los dominios WW, la señal de localización nuclear (NLS) y los dominios FF. En cada panel se muestran las imágenes individuales y superpuestas (*merge*). A la derecha están representados los análisis semi-cuantitativos de las intensidades de las señales, cuya región aparece indicada en las imágenes superpuestas. Barra = 3 μ m.

2.3 La asociación de PRPF40B con los *speckles* nucleares no está afectada por el tratamiento con ARNasa o por la inhibición de la transcripción

Como ya se mencionó en la *Introducción* (Apartado 5.1.1), los *speckles* nucleares son estructuras dinámicas cuya función está relacionada con el almacenamiento, ensamblaje y modificación de componentes de *splicing*, así como proporcionar factores de procesamiento a los sitios activos de transcripción. La arquitectura de los *speckles* nucleares depende de interacciones que pueden ser resistentes al tratamiento con ARNasa (por ejemplo, SRSF2), o bien ser sensibles a la ARNasa (por ejemplo, ribonucleoproteínas Sm). Partiendo de esta base, decidimos estudiar la naturaleza de la asociación de PRPF40B con los *speckles* nucleares, para conocer si el mecanismo implicado en el aumento de la señal de PRPF40B en el compartimento de los *speckles* dependía de ARN. Para ello, transfectamos de forma transitoria la proteína de fusión EGFP-PRPF40B en células HEK293T y realizamos un tratamiento con ARNasa. Bajo estas condiciones, el patrón de PRPF40B y SRSF2 no cambió respecto a los controles sin tratar, mientras que el marcaje con anticuerpos anti-Sm se difuminó completamente en todo el nucleoplasma (Figura R-6). Estos resultados indican, que al igual que ocurre con el factor de *splicing* SRSF2, la localización de PRPF40B en los *speckles* nucleares no se produce por unión al ARN.

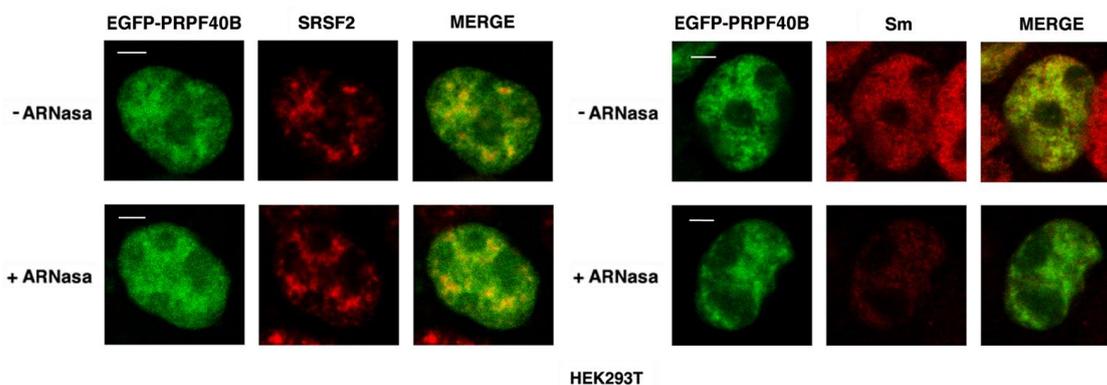


Figura R-6. La localización de PRPF40B en los *speckles* nucleares no depende de ARN. Tras la expresión de la proteína EGFP-PRPF40B, las células HEK293T fueron tratadas con ARNasa A (0,1 mg/ml, 2h a 37°C). Los marcajes con SRSF2 (izquierda) y Sm (derecha) fueron utilizados como controles negativo y positivo, respectivamente. Las imágenes muestran las tinciones individuales y la superposición de ambas (*merge*) sin el tratamiento (paneles superiores) y con el tratamiento con ARNasa A (paneles inferiores). Barras = 3 μ m.

Como ya se ha comentado, los *speckles* nucleares se encuentran próximos a sitios activos de transcripción. Para determinar si la actividad transcripcional está asociada con el enriquecimiento de PRPF40B en los *speckles*, transfectamos las células HeLa con la construcción EGFP-PRPF40B y tratamos las células con el inhibidor de la ARN polimerasa II, α -amanitina. Bajo estas condiciones, el número de *speckles* disminuye por núcleo, adoptando una forma más voluminosa y redondeada debido al reclutamiento y acumulación de la maquinaria de *splicing* encargada de procesar los pre-ARNm recién sintetizados (O'Keefe et al., 1994). Tras la inhibición de la transcripción dependiente de la RNAPII las imágenes tomadas de las inmunofluorescencias mostraron un marcaje de PRPF40B similar al de SRSF2, en el que los *speckles* sufrían un cambio en la morfología, característico de este tratamiento, volviéndose más grandes y redondeados (Figura R-7). Con estos resultados concluimos que la localización de PRPF40B en los *speckles* nucleares no depende de la transcripción activa en la célula, y su distribución es similar al comportamiento de otros factores de *splicing* en condiciones de inhibición transcripcional.

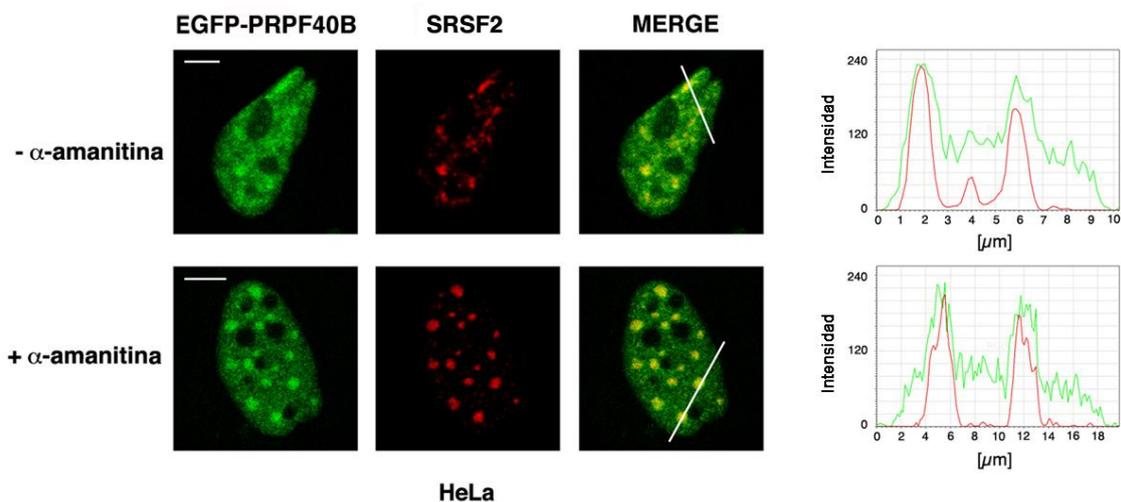


Figura R-7. La distribución de PRPF40B no es afectada con la inhibición de la transcripción. Tras la expresión de EGFP-PRPF40B, las células HeLa fueron tratadas con α -amanitina (25 μ g/ml, 6h a 37°C). Las imágenes muestran los marcajes individuales de PRPF40B (verde), SRSF2 (rojo) y la superposición de ambas (*merge*) en células sin el tratamiento (panel superior) y células tratadas (panel inferior). A la derecha se muestran los gráficos del análisis semi-cuantitativo de la distribución de las intensidades. Las líneas dibujadas en las imágenes superpuestas indican la región de análisis. Barra = 3 μ m.

3. ESTUDIO DE LA INTERACCIÓN DE PRPF40B CON LOS FACTORES DE *SPLICING* SF1 Y U2AF⁶⁵

3.1 PRPF40B colocaliza con los factores de *splicing* SF1 y U2AF⁶⁵ en el núcleo celular

Como ya se indicó en la *Introducción* (Apartado 5.2.2), Prp40, el ortólogo de PRPF40B en levaduras, participa en el reconocimiento del sitio 5'SS durante la primera etapa de ensamblaje del *spliceosoma*. Además, interacciona directamente con BBP y se asocia con Mud2p, los ortólogos en levaduras de los factores de *splicing* SF1 y U2AF⁶⁵, respectivamente (Abovich and Rosbash, 1997). SF1 y U2AF⁶⁵ son dos factores fundamentales del ensamblaje temprano del *spliceosoma*, y se unen de forma secuencial a la región 3'SS del intrón en el pre-ARNm. Por tanto, la interacción de estas tres proteínas situadas en los extremos del intrón podría ser importante para facilitar la definición del exón en levaduras. Con todos estos datos, quisimos investigar si PRPF40B podría estar estableciendo interacciones similares a Prp40 y asociarse con SF1 y U2AF⁶⁵, contribuyendo así en las primeras fases del ensamblaje del *spliceosoma*, desempeñando un papel similar al de su ortólogo en levaduras.

Para abordar esta hipótesis, comenzamos analizando la distribución espacial de PRPF40B, SF1 y U2AF⁶⁵ por microscopía confocal. Las imágenes obtenidas de SF1 y U2AF⁶⁵ mostraron un patrón difuso por todo el nucleoplasma con un incremento en la intensidad de la señal en los *speckles* nucleares (Figura R-8). Para determinar el nivel de proximidad de PRPF40B con estos dos factores de *splicing*, realizamos un doble marcaje utilizando anticuerpos específicos dirigidos contra PRPF40B y SF1, o bien PRPF40B y U2AF⁶⁵, tanto en células HEK293T (Figura R-8A) como HeLa (Figura R-8B). Los resultados observados mostraron un alto grado de solapamiento de la señal de PRPF40B con SF1 y U2AF⁶⁵. Por tanto, estos datos indican que PRPF40B se encuentra espacialmente cerca de estos dos factores de *splicing* en regiones del nucleoplasma celular.

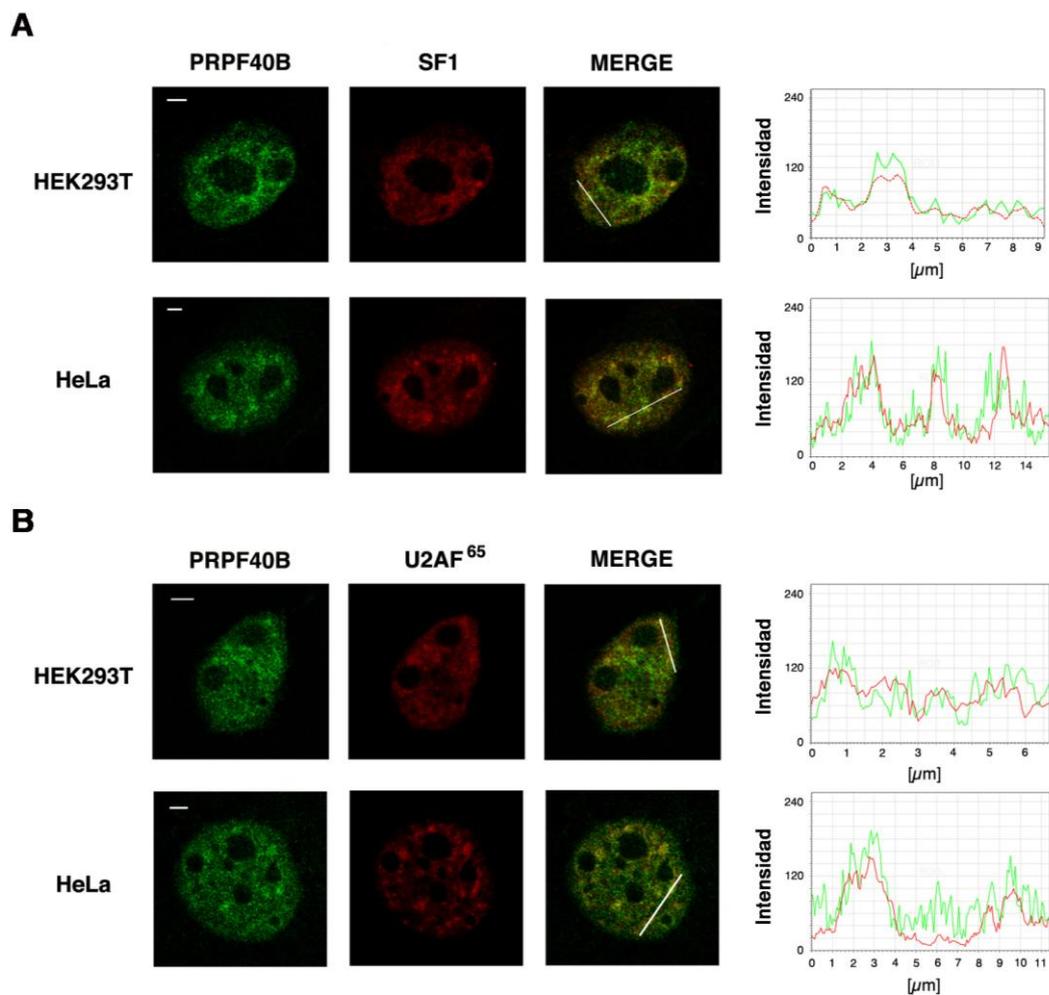


Figura R-8. Colocalización de PRPF40B con los factores de *splicing* SF1 y U2AF⁶⁵. Las imágenes muestran el análisis por inmunofluorescencia de células HEK293T y HeLa doblemente marcadas con anticuerpos que detectan PRPF40B (verde) y SF1 (rojo, A) o U2AF⁶⁵ (rojo, B), así como el solapamiento de ambas señales (*merge*). Las líneas dibujadas en las imágenes superpuestas indican la región utilizada en el análisis semi-cuantitativo de la distribución de las intensidades que se muestra a la derecha de cada panel. Barra = 3 μm .

3.2 Asociación de PRPF40B con el factor de *splicing* SF1 *in vitro*

Los resultados obtenidos en los análisis de inmunofluorescencia son el primer indicio de la relación de PRPF40B con factores implicados en el ensamblaje del *spliceosoma* temprano. Para continuar con este estudio, decidimos analizar las posibles interacciones establecidas entre PRPF40B con SF1 y U2AF⁶⁵.

Para ello, primero estudiamos la asociación de PRPF40B con SF1 mediante experimentos *in vitro* de unión a GST. Con este objetivo, purificamos proteínas recombinantes fusionadas a GST a partir de bacterias. Las construcciones utilizadas contenían la secuencia completa de PRPF40B (Figura R-9A, GST-PRPF40B), la región amino-terminal [Figura R-9A, GST-PRPF40B (1-333)] o la región carboxilo-terminal [Figura R-9A, GST-PRPF40B (446-871)]. Como control positivo del experimento, utilizamos proteínas recombinantes de la región amino de TCERG1 (Figura R-9B, GST-TCERG1 Nt) y U2AF⁶⁵ fusionadas a GST, cuyas interacciones con SF1 han sido descritas previamente (Goldstrohm et al., 2001). Como control negativo utilizamos la proteína GST y la región carboxilo de TCERG1 (Figura R-9B, GST-TCERG1 Ct). Estas proteínas fueron incubadas con glutation covalentemente unido a agarosa y posteriormente fueron incubadas con la proteína recombinante SF1 fusionada al epítipo T7, purificada a partir de células HEK293T. Tras varios pasos de lavado, las proteínas unidas a las bolas fueron eluidas y se separaron en geles de SDS-PAGE. Posteriormente se analizaron por *western blot* para detectar la unión de SF1.

Los resultados revelaron que PRPF40B y SF1 interaccionaban *in vitro*, a diferencia de los controles negativos utilizados. Además, cuando comparamos las regiones amino y carboxilo-terminal de la proteína PRPF40B, observamos que ambos fragmentos se unían a SF1, pero en la región amino existía una mayor fuerza de interacción que en la carboxilo [Figura R-9B, PRPF40B (1-333) y PRPF40B (446-871), respectivamente]. Estos resultados sugieren que PRPF40B contiene dominios de unión a SF1 en los dos extremos de la proteína, aunque la interacción más fuerte era mediada por el segmento amino-terminal. Para comprobar que la interacción establecida entre PRPF40B y SF1 no era consecuencia de una asociación indirecta mediada por ARN, realizamos un tratamiento con ARNasa A previo a los lavados finales en el ensayo de unión a GST. Los datos obtenidos muestran que el tratamiento con ARNasa A no

disminuía la señal de SF1 (Figura R-9C), indicando que en esas condiciones *in vitro* la interacción establecida entre PRPF40B y SF1 no es dependiente de ARN.

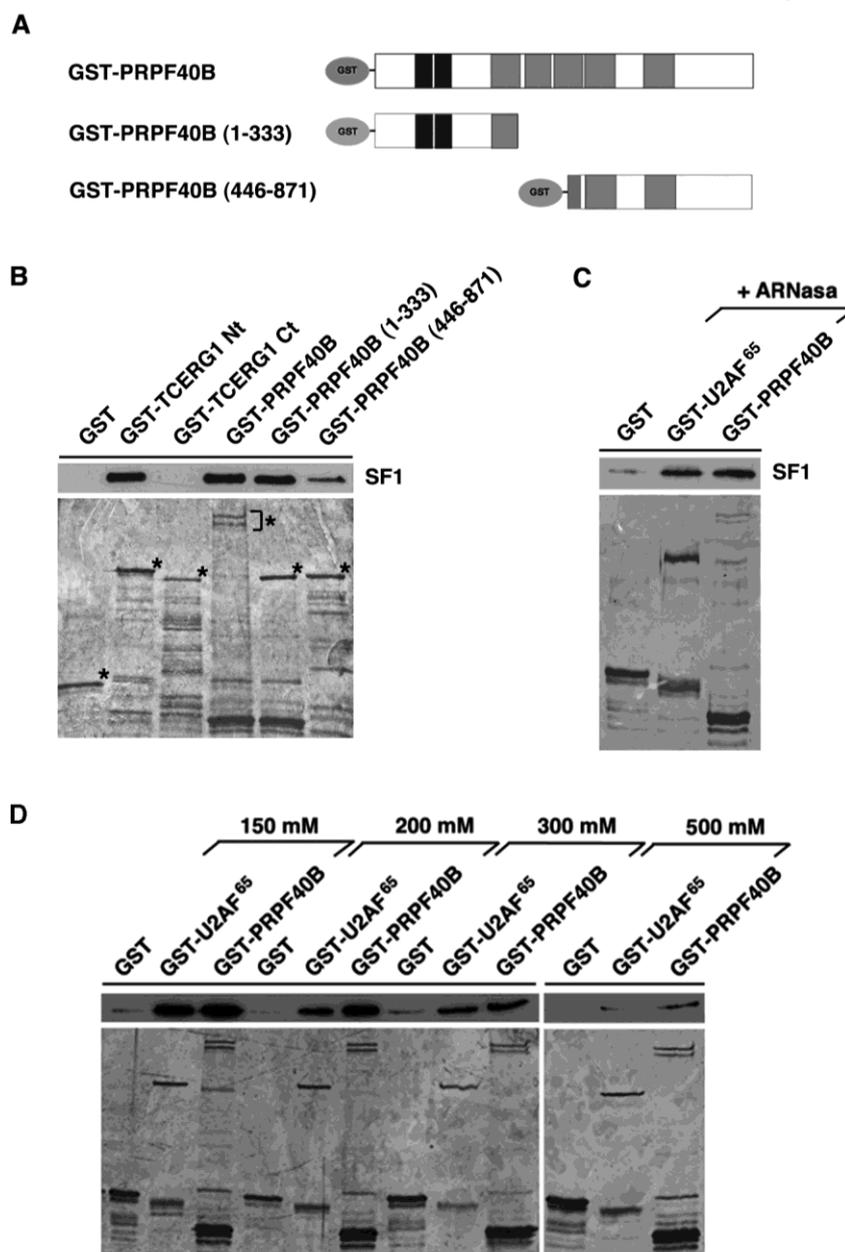


Figura R-9. PRPF40B se asocia *in vitro* al factor de *splicing* SF1 **A.** Representación esquemática de las proteínas de fusión GST-PRPF40B. Los números entre paréntesis indican los aminoácidos comprendidos en cada construcción. Se indican además los dos dominios WW y los cinco dominios FF en las regiones amino y carboxilo, respectivamente. **B.** Las regiones amino y carboxilo-terminal de PRPF40B son importantes para la interacción con SF1. La proteína recombinante GST-PRPF40B, así como las construcciones amino y carboxilo-terminal fueron unidas a resinas de glutatión-agarosa e incubadas con la proteína purificada T7-SF1. GST y GST-TCERG1 carboxilo-terminal (Ct), fueron utilizados como control negativo, y la proteína GST-TCERG1 amino-terminal (Nt) se usó como control positivo. Las proteínas eluidas fueron separadas por SFS-PAGE y analizadas por *western blot* (parte superior)

utilizando anticuerpos que detectaban SF1, o tinción con plata (parte inferior) para detectar las proteínas de fusión GST. **C.** Las proteínas recombinantes GST (control negativo), GST-U2AF⁶⁵ (control positivo) y GST-PRPF40B fueron utilizadas para unir la proteína purificada T7-SF1. Tras la unión, las muestras fueron tratadas con 0,01 mg/ml de ARNasa A, a temperatura ambiente durante 30 min. Posteriormente fueron lavadas y eluidas. Las proteínas unidas fueron separadas en un gel SDS-PAGE y analizadas por *western blot* utilizando anticuerpos para detectar SF1 o por tinción de plata para visualizar las proteínas de fusión GST. **D.** Las proteínas recombinantes GST, GST-U2AF⁶⁵ y GST-PRPF40B se unieron a la proteína recombinante SF1. Las proteínas unidas fueron eluidas utilizando un gradiente lineal de NaCl desde 150 a 500 mM. Tras la elución, las muestras fueron separadas por SDS-PAGE y analizadas por *western blot* o tinción de plata como se ha descrito anteriormente.

El nivel de SF1 detectado en los análisis *in vitro* de interacción con PRPF40B era similar al observado para TCERG1 Nt e incluso algo más alto que para U2AF⁶⁵. Por ello nos pareció interesante medir la fuerza de interacción entre PRPF40B y SF1. Para este estudio nos basamos en el efecto que ejerce el incremento de la concentración salina en la desestabilización de las uniones proteína-proteína. Con este propósito, repetimos los ensayos de unión a GST, incrementando las concentraciones de NaCl a 150, 200, 300 y 500 mM en el tampón utilizado a lo largo del experimento. La agarosa con las proteínas unidas fueron lavadas repetidas veces también con la misma concentración de sales que durante la unión. La elución se hizo con calor en un tampón de carga de SDS-PAGE y las proteínas unidas fueron analizadas por *western blot*. Los resultados observados muestran que la interacción establecida entre PRPF40B y SF1 es estable incluso a concentraciones de NaCl de 500 mM, aunque esta unión se redujo de forma gradual conforme se aumentaba la cantidad de sal (Figura R-9D).

3.3 Asociación de PRPF40B con los factores de *splicing* SF1 y U2AF⁶⁵ *in vivo*

Los datos obtenidos en los ensayos *in vitro* nos parecieron una interesante aproximación sobre lo que podría estar ocurriendo *in vivo*. Por ello, decidimos continuar con el análisis de interacciones utilizando líneas celulares. A pesar de los intentos por detectar una interacción entre las proteínas endógenas, no pudimos obtener los resultados esperados, probablemente porque los anticuerpos utilizados no eran adecuados para realizar ensayos de inmunoprecipitación endógena. Por tanto, para realizar estos experimentos transfectamos células HEK293T con PRPF40B fusionado a un epítipo T7 o bien con un vector T7 vacío como control negativo. Tras

inmunoprecipitar con anticuerpos anti-T7, las muestras fueron analizadas por *western blot* con anticuerpos específicos dirigidos contra los factores de *splicing* SF1 y U2AF⁶⁵ endógenos. Los datos obtenidos revelaban la presencia de SF1 y U2AF⁶⁵ en los inmunoprecipitados, indicando una asociación entre PRPF40B con SF1 así como U2AF⁶⁵ (Figura R-10A). Para confirmar estos resultados, repetimos el ensayo de inmunoprecipitación, pero en este caso transfectando de forma transitoria un vector de expresión de SF1 fusionado a T7 o bien un vector T7 vacío. Los extractos inmunoprecipitados fueron tratados con anticuerpos específicos dirigidos contra PRPF40B y U2AF⁶⁵ endógenos. Al igual que en el experimento anterior, se detectó la asociación tanto de PRPF40B como U2AF⁶⁵ con SF1, confirmando las observaciones previas sobre la interacción de PRPF40B y SF1 (Figura R-10B). Sin embargo, cuando utilizamos anticuerpos dirigidos contra el factor ERK2, una proteína quinasa no relacionada con estos factores, no detectamos señal en los extractos procedentes de la inmunoprecipitación (Figuras R-10A y R-10B). Con estos resultados podemos concluir que PRPF40B se asocia a los factores de *splicing* SF1 y U2AF⁶⁵ *in vivo*.

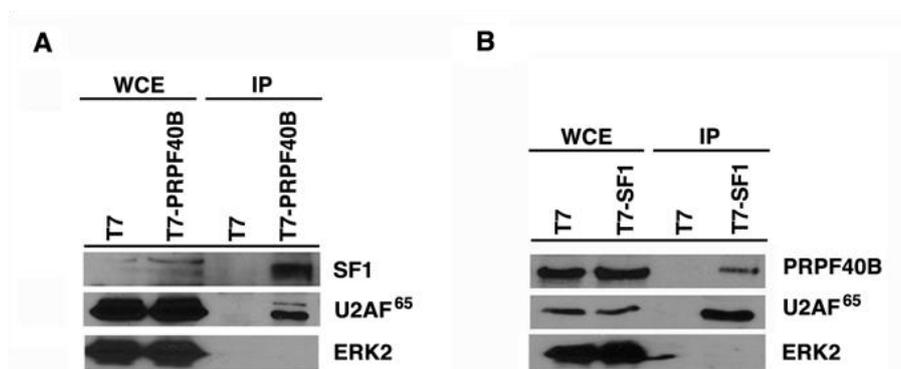


Figura R-10. PRPF40B se asocia con los factores de *splicing* SF1 y U2AF⁶⁵ *in vivo*. **A.** Células HEK293T fueron transfectadas transitoriamente con un plásmido que expresa PRPF40B unida al epítipo T7 o un vector vacío como control negativo. Fracciones de extracto celular total (WCE, del inglés *whole-cell extract*) fueron analizadas directamente por *western blot* o inmunoprecipitadas con anticuerpos específicos contra T7, seguida de una separación por SDS-PAGE y análisis por *western blot* utilizando anticuerpos específicos para SF1, U2AF⁶⁵ y ERK2. **B.** Se realizó el mismo procedimiento experimental descrito en A pero sobre-expresando T7-SF1 y, tras la inmunoprecipitación, los extractos fueron analizados con anticuerpos contra PRPF40B.

3.4 La asociación de PRPF40B con los factores de *splicing* SF1 y U2AF⁶⁵ depende de ARN *in vivo*

De nuevo quisimos estudiar si la interacción establecida entre estas proteínas estaba teniendo lugar de una forma directa proteína-proteína o era mediada por ARN. El tratamiento de los lisados celulares con ARNasa A antes de la inmunoprecipitación reveló que estas condiciones afectaban la asociación de PRPF40B con SF1 y U2AF⁶⁵ (Figuras R-11A y B). Estos resultados sugieren que en las células estas interacciones están siendo estabilizadas por la unión a ARN. Puesto que los factores de *splicing* SF1 y U2AF⁶⁵, así como el ortólogo de PRPF40B en levaduras (Prp40) participan en las etapas tempranas del ensamblaje, era de interés comprobar si algún ARN involucrado en este proceso podría estar mediando esta interacción. Por ello, decidimos analizar el posible papel mediador de U1 snRNA, componente de la partícula U1 snRNP, que media las interacciones en la formación del primer complejo del spliceosoma. Utilizamos un oligonucleótido antisentido Morpholino (AMO, *antisense morpholino oligonucleotide*), para bloquear la actividad funcional de U1 snRNA (Kaida et al., 2010). Se co-transfectaron las células HEK293T con los vectores T7-PRPF40B o T7 vacío como se indicó anteriormente, y además con el U1 AMO (Mo-U1) o el control negativo (Mo-CRTL). La capacidad de bloqueo del U1 AMO se confirmó analizando su efecto en el cambio del *splicing* alternativo del gen *Fas* (Figura R-11C). A continuación, se repitieron los experimentos de inmunoprecipitación como ya se describió previamente. De nuevo observamos la presencia de las proteínas endógenas SF1 y U2AF⁶⁵ en los extractos inmunoprecipitados (Figura R-11D), lo que sugiere que U1 snRNP no está mediando estas interacciones.

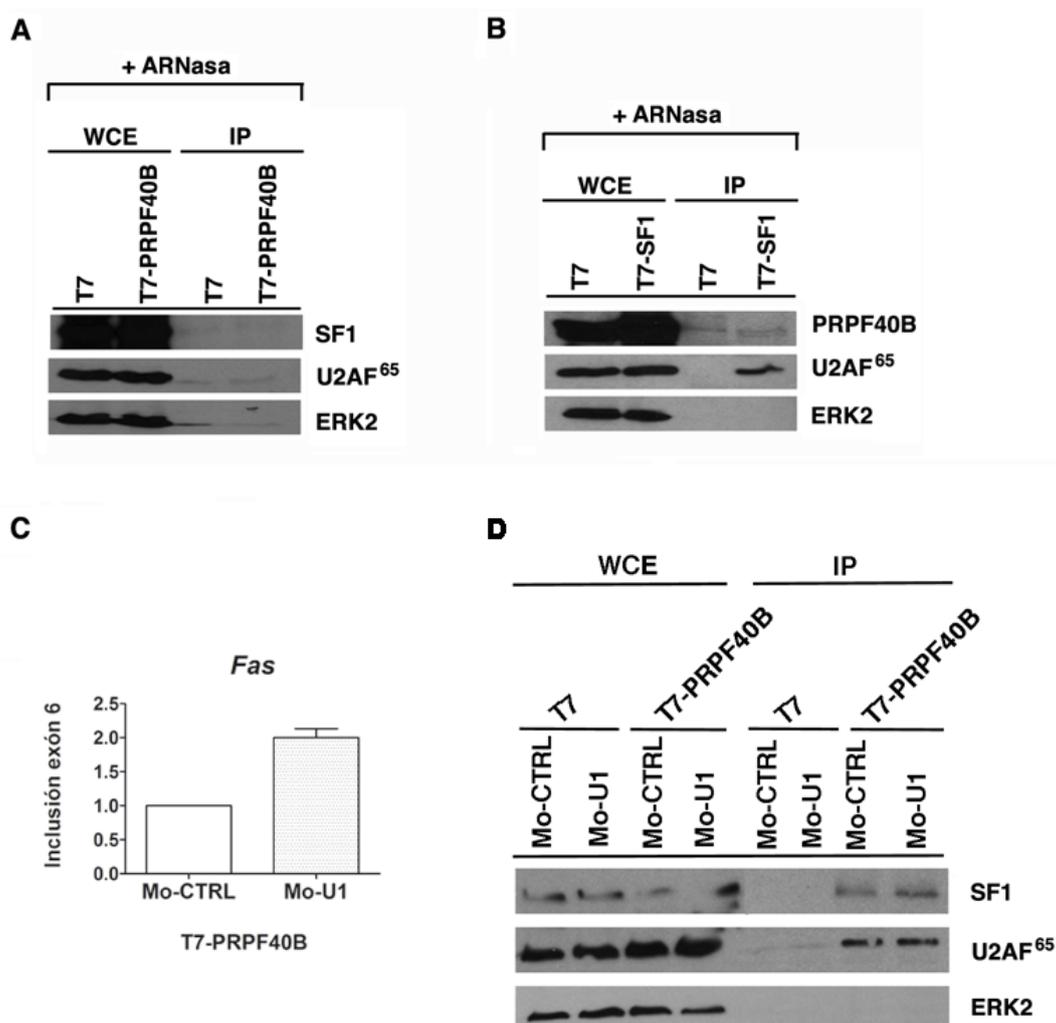


Figura R-11. La asociación de PRPF40B a los factores de *splicing* SF1 y U2AF⁶⁵ depende de ARN *in vivo*. **A.** Extractos celulares de HEK293T transfectadas transitoriamente con un plásmido de expresión PRPF40B o vacío fueron tratados con ARNasa A (0,1 mg/ml, 1h 4°C). Después una fracción era analizada por *western blot* (WCE) y el resto se inmunoprecipitaba con anticuerpos específicos de T7. A continuación se separaban las proteínas por SDS-PAGE y analizaban por *western blot*, utilizando anticuerpos que detectaban SF1, U2AF⁶⁵ y ERK2. **B.** El mismo procedimiento descrito en el apartado A se llevó a cabo transfectando SF1 y utilizando en el *western blot* anticuerpos específicos para PRPF40B. **C.** Representación gráfica de los niveles de *Fas* endógeno tras el bloqueo de U1 snRNA. Se extrajo ARN de una fracción de los extractos celulares de HEK293T donde se sobre-expresó T7-PRPF40B y tratados con un oligonucleótido antisentido morpholino (Mo-U1) y un oligonucleótido control (Mo-CTRL). La gráfica muestra los resultados de la RT-qPCR, como el ratio de inclusión del exón 6 relativo al control que es 1, de dos experimentos independientes. **D.** Células HEK293T fueron co-transfectadas con PRPF40B o un vector vacío y con Mo-U1 o Mo-CRTL. Los extractos celulares fueron inmunoprecipitados y analizados como se describe en la leyenda del apartado A.

4. ESTUDIO DEL EFECTO DE PRPF40B SOBRE EL *SPLICING* ALTERNATIVO

4.1 PRPF40B interviene en la selección del sitio de *splicing* alternativo

En los apartados anteriores, se ha localizado a PRPF40B en los *speckles* nucleares y se ha descrito su asociación con los factores de *splicing* SF1 y U2AF⁶⁵. Basándonos en estos datos, decidimos llevar a cabo ensayos funcionales de *splicing in vivo* para investigar la capacidad de PRPF40B de influir en el *splicing* alternativo. Para ello utilizamos construcciones reporteras que contienen exones alternativos flanqueados por intrones constitutivos clonados en vectores de expresión eucariota. Estos genes reporteros, conocidos como minigenes, son una buena herramienta para estudiar la regulación del *splicing* alternativo (Cooper, 2005).

Para estudiar el efecto de la sobre-expresión de PRPF40B, las células HEK293T fueron co-transfectadas con los minigenes reporteros junto con un vector que expresa PRPF40B fusionado al epítipo T7 o un vector vacío. Para estudiar el efecto del silenciamiento de PRPF40B, las células fueron co-transfectadas con los minigenes reporteros y siRNAs específicos. Debido a la alta homología de las proteínas PRPF40B y PRPF40A comprobamos mediante RT-pPCR que los siRNA utilizados para silenciar PRPF40B no afectaran a PRPF40A, como se indica en la Tabla R-1.

	Nivel ARN PRPF40B	Nivel ARN PRPF40A
siPRPF40B	0,437±0,058	1,258±0,107

Tabla R-1. Nivel de ARN de PRPF40A y PRPF40B al utilizar siPRPF40B. Representación de los datos obtenidos de la RT-qPCR de células HEK293T silenciadas con siPRPF40B. Se midió la cantidad relativa de ARN de PRPF40A y PRPF40B respecto al control GAPDH.

Los minigenes utilizados en este estudio han sido descritos previamente y contienen sitios alternativos 5'SS (*Fas* y *Bcl-x*) (Figura R-11). El receptor de muerte celular Fas/CD95, que codifica para el receptor de muerte transmembrana, puede ser procesado alternativamente para excluir el exón 6 de su ARNm dando lugar a una forma soluble cuya función es inhibir la apoptosis (Cascino et al., 1995; Cheng et al., 1994). El pre-ARNm del gen *Bcl-x* humano puede utilizar dos sitios 5'SS alternativos en el exón 2 para producir dos isoformas: una larga anti-apoptótica (*Bcl-x_L*) y una isoforma corta pro-apoptótica (*Bcl-x_S*) (Boise et al., 1993).

Con el fin de optimizar los resultados, se recogieron las células a diferentes tiempos según el minigen y la condición de silenciamiento o sobre-expresión, tal y como se indica en *Materiales y Métodos* (Apartado 3.2). Posteriormente el ARN fue extraído de las células transfectadas con los minigenes bajo condiciones de sobre-expresión o silenciamiento de PRPF40B, y se analizó el resultado del *splicing* de los minigenes por RT-PCR. Los resultados obtenidos mostraron que la sobre-expresión de PRPF40B incrementó la producción de la isoforma *Bcl-x_S* (Figura R-11A), mientras que la depleción de PRPF40B favoreció la expresión de la isoforma *Bcl-x_L*. En el caso del minigen *Fas*, la sobre-expresión de PRPF40B generó una mayor cantidad de la isoforma corta de *Fas* (Figura R-11B), mientras que la depleción de PRPF40B resultó en un incremento de la isoforma larga de *Fas* (Figura R-11B). Por tanto, estos datos muestran que PRPF40B es capaz de modular eventos de *splicing* alternativo.

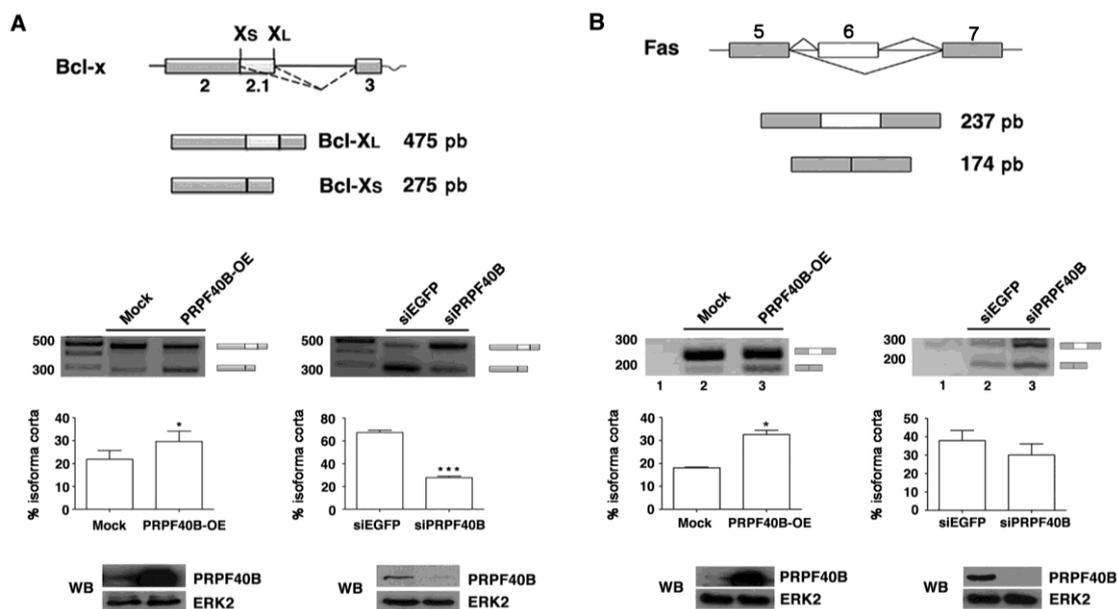


Figura R-11. PRPF40B afecta el *splicing* alternativo de minigenes de apoptosis. Efecto de la sobre-expresión y el silenciamiento de PRPF40B sobre los minigenes de *Bcl-x* y *Fas*. **A y B.** Representación esquemática de los minigenes *Bcl-x* y *Fas* y sus variantes de *splicing*. Se indican los exones (rectángulos), intrones (líneas horizontales) y los patrones de los eventos de *splicing* alternativo (líneas inclinadas). Se muestran los geles de agarosa donde se resuelven los productos de *splicing* generados mediante RT-PCR de células HEK293T co-transfectadas con los minigenes *Bcl-x* (**A**) o *Fas* (**B**) junto con un vector vacío (Mock) o el plásmido de expresión PRPF40B (PRPF40B-OE) en los experimentos de sobre-expresión; o bien, con siPRPF40B o el control siEGFP en los experimentos de depleción. Las gráficas muestran los resultados del análisis densitométrico como porcentaje de isoforma corta a partir de tres experimentos independientes (media \pm SEM). *, $P < 0,05$; ***, $P < 0,005$. Una fracción de los lisados celulares fueron analizados por *western blot* utilizando anticuerpos que detectan PRPF40B y ERK2.

A continuación decidimos profundizar en el papel de PRPF40B en la regulación del *splicing* alternativo tomando el minigen de *Fas* como modelo. Primero analizamos el patrón de *splicing* alternativo del gen *Fas* endógeno bajo condiciones de sobre-expresión o silenciamiento de PRPF40B mediante RT-PCR, utilizando oligonucleótidos que amplifican la isoforma corta anti-apoptótica (que codifica para una forma soluble) y la isoforma larga pro-apoptótica (que codifica para el receptor transmembrana). Los resultados obtenidos fueron consistentes con los datos previos, observando que la sobre-expresión de PRPF40B redujo los niveles de inclusión del exón 6, aumentando la cantidad de la isoforma corta anti-apoptótica. Además la disminución de la expresión de PRPF40B produjo un aumento en la isoforma larga pro-apoptótica (Figura R-12A). Para corroborar estos resultados repetimos los experimentos usando PCR cuantitativa con oligonucleótidos que amplifican exclusivamente la isoforma larga del receptor transmembrana, obteniendo los mismos resultados (Figura R-12B).

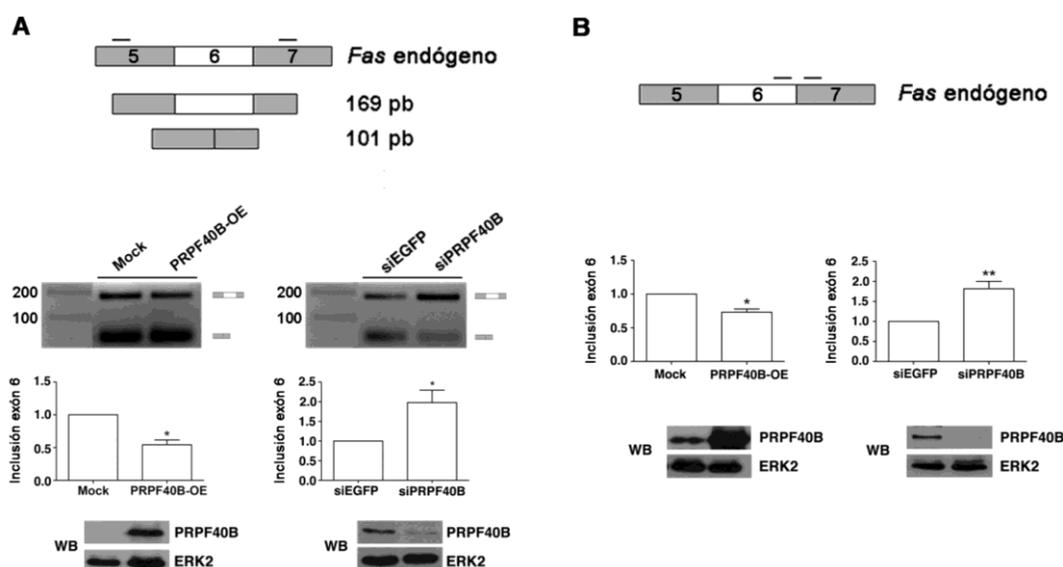


Figura R-12. PRPF40B regula el *splicing* alternativo del gen endógeno *Fas*. Células HEK293T fueron transfectadas con un vector vacío (Mock) o un plásmido de expresión PRPF40B (PRPF40B-OE) para experimentos de sobre-expresión, o bien con siPRPF40B o el control siEGFP para experimentos de silenciamiento. Tras la extracción del ARN, las muestras se analizaron por RT-PCR (**A**) o RT-qPCR (**B**) utilizando los oligonucleótidos que se indican en cada esquema de *Fas*. **A.** El gel muestra los productos derivados de la RT-PCR que amplifican la región comprendida entre los exones 5 y 7 de *Fas*, incluyendo las isoformas anti-apoptóticas (inclusión exón 6) y pro-apoptóticas (no inclusión exón 6). Los gráficos de barras muestran el análisis densitométrico como el ratio de inclusión del exón 6 relativo al control que es 1, a partir de tres experimentos independientes (media \pm SEM). *, $P < 0,05$. Una fracción de los lisados celulares fueron analizados por *western blot* utilizando anticuerpos que detectan PRPF40B y ERK2.

B. Los oligonucleótidos utilizados amplifican por RT-qPCR la isoforma pro-apoptótica que incluye el exón 6. Las gráficas representan el ratio de inclusión del exón 6 relativo al control que es 1, a partir de

cinco experimentos independientes (media \pm SEM). *, $P < 0,05$; **, $P < 0,001$. Una fracción de los lisados celulares se analizaron por *western blot* utilizando anticuerpos que detectan PRPF40B y ERK2.

4.2 El efecto de PRPF40B en el *splicing* alternativo de *Fas* necesita secuencias del exón 6 y los sitios 5'SS y 3'SS

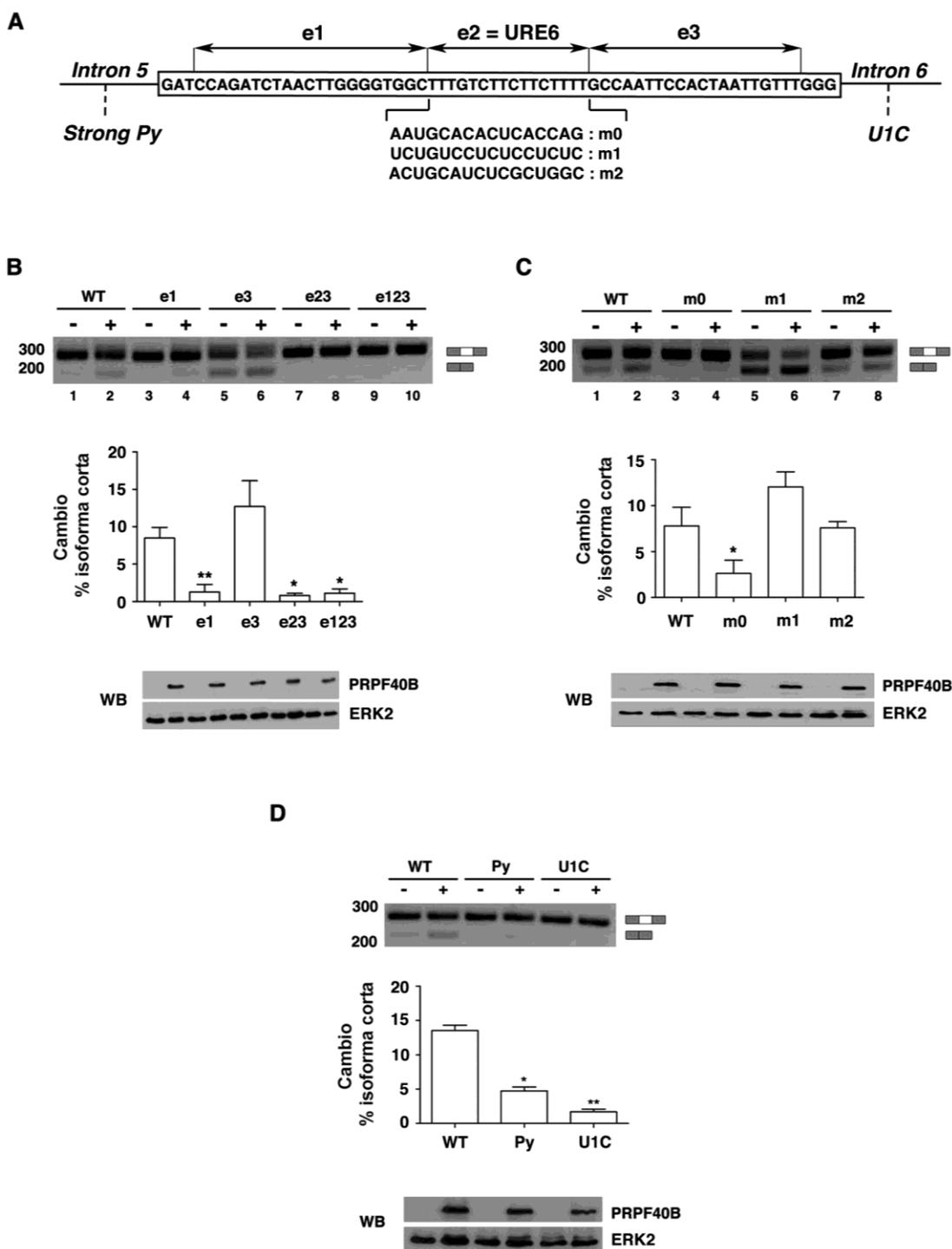
Para entender mejor la regulación del *splicing* alternativo de *Fas* por PRPF40B, comenzamos por determinar las secuencias reguladoras exónicas e intrónicas importantes para la función de PRPF40B. Aunque se asume que Prp40 en levaduras conecta los extremos 5' y 3' del intrón mediante la interacción con BBP y U1snRNP sin una unión directa con el pre-ARNm, el estudio de la secuencia de los elementos que afectan a la función de PRPF40B proporcionará información útil sobre el mecanismo molecular por el cual PRPF40B está modulando el *splicing* alternativo. Con este fin utilizamos minigenes reporteros que disponíamos en el laboratorio con la secuencia del exón 6 de *Fas* dividida en tres segmentos, aproximadamente de la misma longitud (e1, e2 y e3, Figura R-13A). Estos minigenes ya han sido utilizados previamente para caracterizar la función del gen supresor de tumores RBM5 en el exón 6 de *Fas* (Bonnal et al., 2008). Las células HEK293T fueron co-transfectadas con un vector de expresión de PRPF40B y cada uno de estos minigenes, utilizando la secuencia *Fas* WT como control, y posteriormente fueron procesadas para su análisis por RT-PCR. Los resultados obtenidos mostraron que en el caso del minigen con la secuencia e3 mutada no se produjo ningún cambio en la función de PRPF40B (Figura R-13B, carriles 5-6). Sin embargo, la sobre-expresión de PRPF40B produjo una bajada en los niveles de la isoforma corta anti-apoptótica respecto a los niveles obtenidos con *Fas* WT cuando se usaron los minigenes con las secuencias e1 y e2 mutadas (Figura R-13B, carriles 3-4 y 7-10). Estos datos indican que secuencias presentes en la región 5' del exón 6 de *Fas* son importantes para la función de PRPF40B.

El segmento e2 se corresponde con el elemento rico en uridinas del exón 6 de *Fas*, URE6. Datos previos han mostrado que el elemento URE6 contiene un silenciador exónico de *splicing*, a través del cual el factor represor de *splicing* PTB ejerce su efecto de silenciamiento y que depende del alto contenido en pirimidinas de la secuencia reguladora (Izquierdo et al., 2005). Para analizar si el efecto de PRPF40B es a través de la secuencia represora de URE6, realizamos experimentos de RT-PCR tras sobre-expresar PRPF40B en células HEK293T junto con el minigen *Fas* WT o construcciones

derivadas con mutaciones en el elemento URE6. Estas mutaciones cambian su composición y dan lugar a: una secuencia rica en citosinas y adenosinas, otra secuencia rica en uridinas que mantienen la actividad represora de PTB y una secuencia al azar de una composición balanceada (m0, m1 y m2 en la Figura R-13A, respectivamente) (Bonnal et al., 2008; Izquierdo et al., 2005). Los resultados de esos experimentos mostraron que sólo en el caso de la secuencia mutada m0 se reducía de forma significativa el efecto de sobre-expresión de PRPF40B observado con el minigen *Fas* WT (Figura R-13C, líneas 3-4). Estos datos coinciden con los resultados previos obtenidos con la mutación del elemento e2 y determinan que la composición del silenciador exónico de *splicing* URE6, rico en uridinas, es importante para la función de PRPF40B.

Dada la importancia de la presencia de secuencias reguladoras en el gen de *Fas*, continuamos el estudio analizando el efecto de los sitios de *splicing* en la función de PRPF40B. Para ello utilizamos mutantes del minigen *Fas* que han sido modificados bien para fortalecer el sitio 3'SS insertando un tracto de polipirimidinas (Py), bien para favorecer el reconocimiento del U1 snRNA con el sitio 5'SS (U1C). Los datos obtenidos revelaron que el efecto ejercido sobre *Fas* al sobre-expresar PRPF40B se disminuía al modificar los sitios de *splicing* asociados con el exón 6 (Figura R-13D). Estos resultados sugieren que los sitios 3'SS y 5'SS en el exón 6 modulan la respuesta de PRPF40B.

Figura R-13 (Página siguiente). Secuencias del exón 6 de *Fas* y los sitios 5'SS y 3'SS asociados, son necesarios para la regulación del *splicing* alternativo de *Fas* mediado por PRPF40B. A. Representación esquemática de la estructura del gen *Fas*, que incluye la secuencia genómica del exón 6 (rectángulo), y los intrones contiguos 5 y 6. Los tres segmentos e1, e2 y e3 están representados, de los cuales e2 se corresponde con el silenciador exónico rico en uridinas (URE6). Se indican las secuencias m0, m1 y m2, que se sustituyen en el ARN del silenciador URE6. El sitio fuerte de polipirimidinas en el intrón 5 que sustituye al sitio débil 3'SS está también indicado (*Strong Py*). La secuencia U1C en el intrón 6 se corresponde con la mutación que fortalece el sitio 5'SS, ayudando al reconocimiento de esta secuencia por el U1 snRNA. **B, C, D.** Análisis del efecto de la sobre-expresión de PRPF40B en la selección del sitio de *splicing*. Células HEK293T fueron co-transfectadas con un plásmido de expresión de PRPF40B junto con el minigen de *Fas* correspondiente. Las muestras se procesaron y analizaron por RT-PCR para visualizar los patrones de *splicing* alternativo. Las gráficas muestran los resultados del análisis densitométrico como cambio en el porcentaje de la banda corta a partir de tres experimentos independientes (media \pm SEM). *, $P < 0,05$; **, $P < 0,001$. Una fracción de los lisados celulares fueron analizados por *western blot* utilizando anticuerpos que detectan PRPF40B y ERK2.



Por último y para analizar si el reclutamiento de PRPF40B al elemento URE6 es suficiente para mediar los efectos observados en el *splicing* alternativo de *Fas*, utilizamos un ensayo basado en la proteína de unión a ARN del bacteriófago MS2. Datos ya publicados muestran que la sustitución de la secuencia URE6 por sitios de unión a la proteína MS2 (*MS2-Fas*) resulta en un aumento de la isoforma corta de *Fas* cuando se sobre-expresa la proteína de fusión MS2-PTB (Izquierdo et al., 2005). Nuestros experimentos mostraron que la co-transfección de un plásmido que expresa

PRPF40B (T7-PRPF40B) con el minigen *MS2-Fas* no produjo cambios en el *splicing* alternativo del mismo (Figura R-14B, carril 4), a diferencia de lo observado con la expresión de la proteína de fusión MS2-PTB (Figura R-14B, carril 2), indicando de nuevo que la secuencia del exón 6 es fundamental para la regulación del *splicing* alternativo de *Fas* por PRPF40B. Sin embargo, cuando sobre-expresamos la proteína de fusión MS2-PRPF40B junto con el minigen *MS2-Fas*, observamos que el *splicing* alternativo del minigen no se afectó por la expresión de la proteína de fusión MS2-PRPF40B (Figura R-14B, carril 3). Con estos resultados concluimos que el reclutamiento de PRPF40B al elemento URE6 no es suficiente para promover la isoforma anti-apoptótica de *Fas*, sino que la secuencia a través de la cual puede estar ejerciendo su efecto también es importante para su acción reguladora.

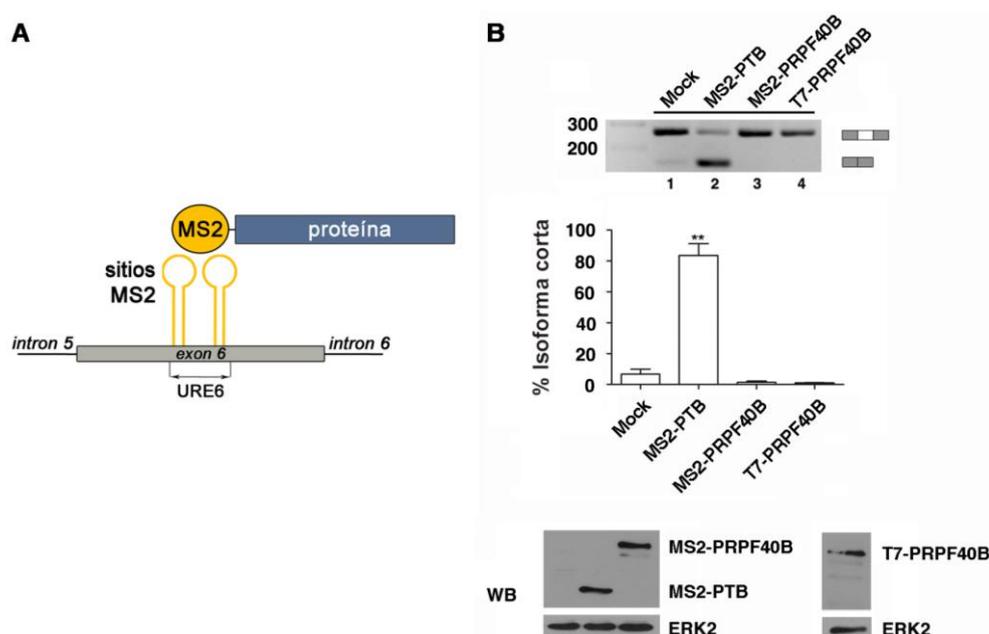


Figura R-14. El reclutamiento de PRPF40B al minigen de *Fas* a través del elemento MS2 no es suficiente para ejercer su regulación del *splicing*. **A.** Diagrama esquemático de la sustitución de la secuencia URE6 por dos sitios de unión MS2, que incluye el exón 6 (rectángulo gris) y los intrones 5 y 6 (líneas negras) del gen *Fas*. Se representan los dos dominios de unión MS2 en tándem (bucles amarillos) que sustituyen la secuencia URE6 y la proteína de fusión (rectángulo azul) unida al dominio de unión MS2 (círculo amarillo). **B.** El minigen *MS2-Fas* fue co-transfectado con plásmidos que expresaban las proteínas fusionadas a dominios de unión MS2 o un vector vacío (control negativo) o el plásmido T7-PRPF40B. El gel muestra los resultados de la RT-PCR para analizar los niveles de las isoformas generadas de *Fas*. Las gráficas muestran los resultados del análisis densitométrico como cambio en el porcentaje de la banda corta, a partir de tres experimentos independientes (media \pm SEM). **, $P < 0,01$. Una fracción de los lisados celulares fueron analizados por *western blot* utilizando anticuerpos que detectan PRPF40B y ERK2.

5. EFECTO DE PRPF40B EN APOPTOSIS

5.1 El silenciamiento de PRPF40B aumenta el número de receptores de Fas/CD95

Como se ha descrito en apartados anteriores, la depleción de PRPF40B promueve la inclusión del exón 6 de *Fas*, dando lugar a un incremento en la isoforma larga pro-apoptótica que codifica para el receptor transmembrana Fas/CD95 (Figura R-15).

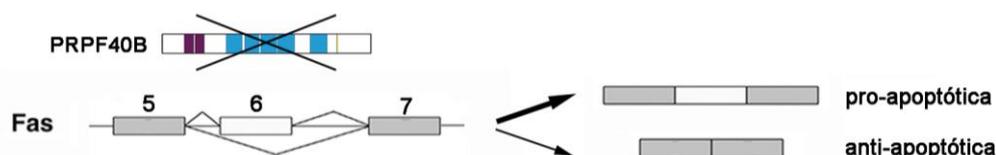


Figura R-15. Representación esquemática del efecto de PRPF40B en el *splicing* alternativo del exón 6 de *Fas*. Diagrama esquemático del gen *Fas* donde se representan los exones 5, 6 y 7 (rectángulos) separados por intrones (líneas) y el patrón de *splicing* alternativo (líneas oblicuas). Se indican las isoformas pro-apoptótica y anti-apoptótica generadas por *splicing* alternativo. Se muestra el silenciamiento de PRPF40B (tachado) y el incremento de la isoforma con la inclusión del exón 6 (flecha gruesa).

Para determinar si los cambios observados en el *splicing* alternativo de *Fas* se reflejan en una variación de la expresión de su receptor transmembrana, decidimos estudiar el número de receptores de *Fas* en células HEK293T control y células silenciadas transitoriamente con siPRPF40B mediante citometría de flujo utilizando anticuerpos específicos contra el receptor de Fas/CD95. Los datos de citometría se representan como histogramas de frecuencia de la inmunofluorescencia de Fas/CD95, producida a través de la señal emitida por el anticuerpo secundario conjugado con Alexa 647. Los resultados obtenidos mostraron un incremento en la expresión de receptores transmembrana de *Fas* cuando se depleciona PRPF40B (Figura R-16A). El análisis estadístico se efectuó utilizando el test no paramétrico de Kolmogorov-Smirnov para comparar dos muestras, que resultó en un número significativamente pequeño del valor P ($\leq 0,001$), lo que indica que esta distribución no se debe a cambios debidos al azar (Figura R-16B). Estos resultados sugieren que los efectos obtenidos en los ensayos de *splicing* del gen *Fas* tienen una relevancia fisiológica, puesto que el aumento de la isoforma larga pro-apoptótica se correlaciona con un aumento en el número de receptores de membrana Fas/CD95 que codifica.

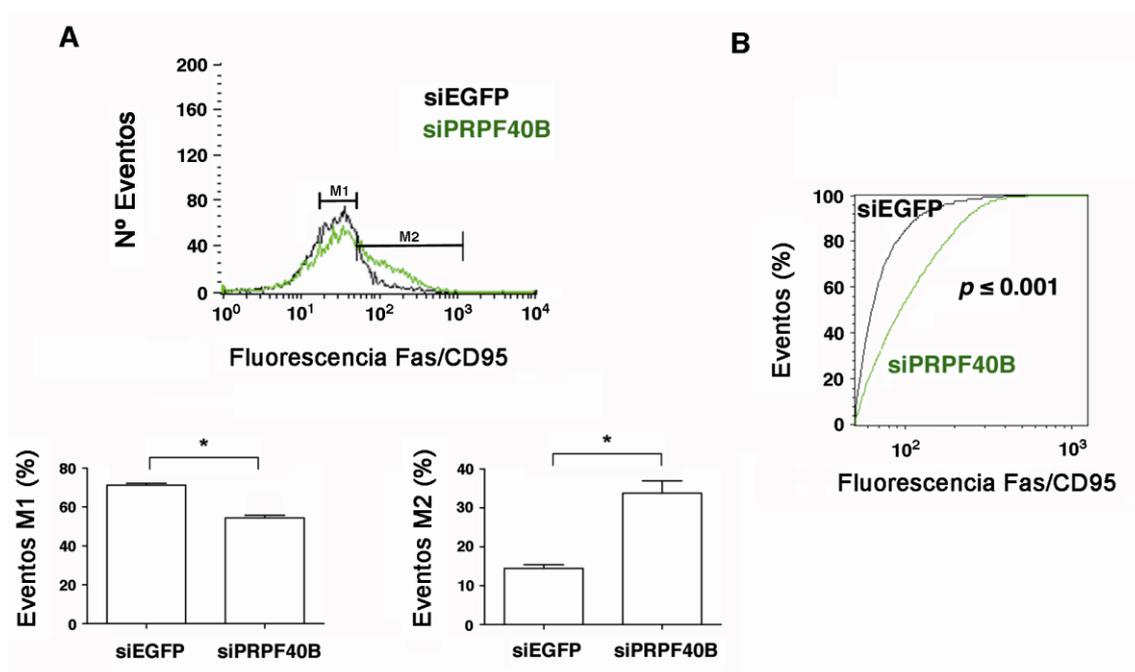


Figura R-16. El silenciamiento de PRPF40B incrementa el número de receptores de membrana de Fas/CD95. **A.** El gráfico superior representa el histograma de frecuencias generado por citometría de flujo de la fluorescencia asociada al receptor Fas/CD95 en células control (siEGFP, negro) y siPRPF40B (verde). Los resultados que se muestran proceden de un experimento representativo. Los gráficos de barra representan la cuantificación de la fluorescencia de las poblaciones M1 (izquierda) y M2 (derecha). Los datos proceden de tres experimentos independientes (media \pm SEM). *, $P < 0,05$. **B.** Gráfico que muestra los resultados obtenidos del test estadístico Kolmogorov-Smirnov al comparar la distribución de las poblaciones de fluorescencia asociadas a los receptores Fas/CD95 de células control siEGFP (negro) y siPRPF40B (verde). El valor P fue $\leq 0,001$, lo que indica que los dos grupos de muestras proceden de poblaciones con distribuciones diferentes y no generadas al azar.

5.2 La depleción de PRPF40B aumenta el porcentaje de muerte celular por apoptosis

El silenciamiento de PRPF40B da lugar a un incremento de la isoforma larga pro-apoptótica de *Fas* y a un aumento en el número de receptores de membrana Fas/CD95. Estos resultados sugieren un papel de PRPF40B en el proceso de muerte celular. Por ello, quisimos conocer la importancia fisiológica del efecto de PRPF40B realizando experimentos funcionales de apoptosis.

Puesto que los ensayos funcionales para estudiar del efecto de PRPF40B en el *splicing* alternativo de *Fas* se han llevado a cabo en células HEK293T, decidimos en una primera aproximación continuar utilizando esta línea celular, conociendo *a priori* su elevada resistencia a la muerte celular inducida y por ello los previsible bajos

resultados que se pudieran obtener en los experimentos de apoptosis. Comenzamos midiendo el porcentaje de células muertas totales por la incorporación de yoduro de propidio en células silenciadas transitoriamente con un siEGFP control o bien, con siPRPF40B. Observamos que la ausencia de PRPF40B aumentó la muerte celular (Figura R-17A). Sin embargo, esta técnica no permite distinguir entre células en apoptosis y necrosis. Por tanto para una mejor cuantificación de células en fase de apoptosis, llevamos a cabo experimentos de ciclo celular, donde se pueden determinar las células en fase sub-G1, que se corresponde con células en apoptosis. Los resultados de la citometría de flujo muestran un pequeño pero significativo incremento del porcentaje de células en sub-G1 en células siPRPF40B, respecto las células control (Figura R-17B). Estos resultados sugieren que la falta de PRPF40B aumenta la proporción de células apoptóticas.

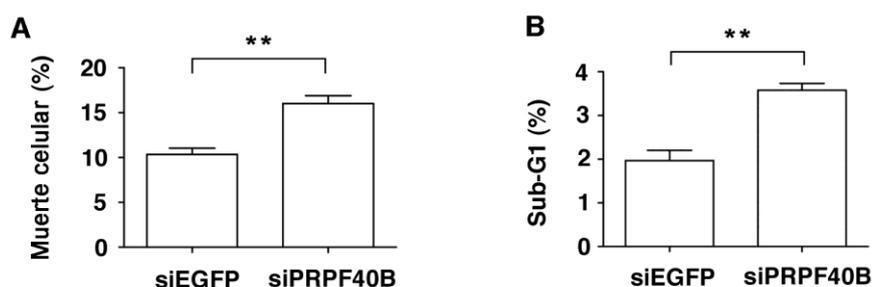


Figura R-17. El silenciamiento de PRPF40B incrementa la muerte celular por apoptosis. Representación gráfica del porcentaje de muerte celular total (A) o células apoptóticas (Sub-G1) (B) analizado por citometría de flujo de células silenciadas transitoriamente con siPRPF40B o siEGFP (control). Los gráficos muestran el porcentaje de células de tres experimentos independientes (media \pm SEM). **, $P < 0,001$.

Para corroborar estos resultados decidimos realizar ensayos de unión de anexina-V. La anexina-V se une de forma específica y con una alta afinidad al fosfolípido fosfatidilserina cuando esta molécula se transloca desde la cara interna de la membrana plasmática a la superficie celular bajo condiciones de apoptosis. Por tanto, las células anexina-V positivas se corresponden con células en fase de apoptosis. Tras el silenciamiento transitorio con un siEGFP control o siPRPF40B, las células HEK293T fueron incubadas con anexina-V para su análisis por citometría de flujo. Primero determinamos la viabilidad celular, es decir células negativas para anexina-V, con el fin de comparar el porcentaje de muerte celular inicial como control del experimento. Como ocurría en experimentos previos, la ausencia de PRPF40B disminuyó el número

de células viables (Figura R-18). Seguidamente se midieron los porcentajes de células positivas para anexina-V, observando un aumento pequeño pero significativo de la población de células silenciadas en PRPF40B respecto las células control (Figura R-18).

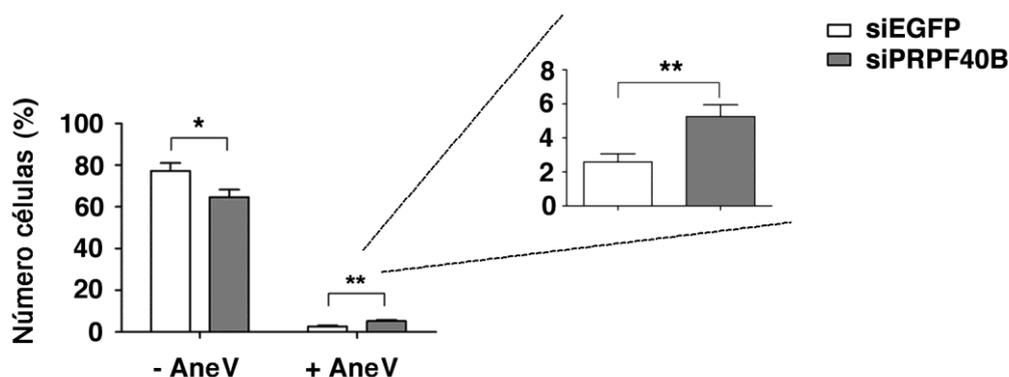


Figura R-18. El silenciamiento de PRPF40B incrementa el porcentaje de células anexina-V positivas. Células HEK293T silenciadas con siPRPF40B o siEGFP (control) fueron incubadas sin (izquierda) o con (derecha) anexina-V, y analizadas por citometría de flujo. Las gráficas muestran el porcentaje de células procedente de tres experimentos independientes (media \pm SEM). *, $P < 0,05$; **, $P < 0,001$.

5.3 El silenciamiento de PRPF40B aumenta la actividad de la caspasa-3

Otro parámetro indicador de apoptosis es la cantidad de caspasa-3 activa presente en la célula. Para medir la actividad de la caspasa-3, utilizamos un kit comercial basado en la adición de un sustrato de luciferasa con una secuencia de corte VEDG, que en presencia del enzima se corta emitiendo una señal de luminiscencia. El número de unidades de luciferasa obtenido es directamente proporcional a la cantidad de caspasa-3 activa presente en las células. Los resultados se corrigieron respecto a células a las que no se les adicionó el sustrato de luciferasa. El silenciamiento de PRPF40B dio lugar a un pequeño pero significativo incremento de la actividad de la caspasa-3 (casi dos veces de activación relativa) respecto las células control (Figura R-19A).

Para validar estos resultados a nivel bioquímico, analizamos la expresión de la caspasa-3 en ausencia de PRPF40B mediante *western blot*. La caspasa-3 se genera inicialmente como un precursor inactivo de 35 KDa (pro-caspasa 3). Durante el proceso de apoptosis, este precursor da lugar mediante una cascada de cortes proteolíticos a la caspasa-3 activa, un heterodímero compuesto de dos péptidos de 17 KDa (p17) y otros dos de 11 KDa (p11). Para ello, la pro-caspasa-3 primero es procesada para generar la

subunidad p11 y una subunidad de 20 KDa (p20) que a su vez generará la forma madura de 17 KDa. Utilizando anticuerpos específicos contra la caspasa-3 observamos una disminución significativa del precursor pro-caspasa 3, correspondiéndose con el aumento de los péptidos p20 y p17 bajo condiciones de silenciamiento de PRPF40B (Figura R-19B). Estos resultados están de acuerdo con los resultados anteriores sobre el aumento de la actividad enzimática de la caspasa-3 cuando bajan los niveles de PRPF40B en la célula. Además quisimos analizar los niveles de expresión de otras proteínas implicadas en el proceso de apoptosis, con un papel inhibitorio de la apoptosis como BCL-2. La detección mediante anticuerpos específicos contra BCL-2, mostraban la disminución de su expresión ante la ausencia de PRPF40B (Figura R-19B).

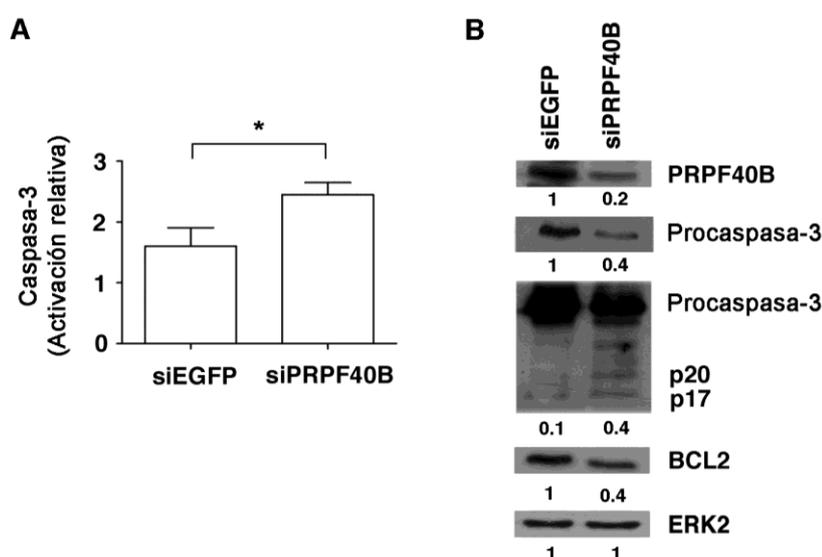


Figura R-19. El silenciamiento de PRPF40B aumenta la actividad de la caspasa-3. **A.** Medida de la quimioluminiscencia emitida por la activación de la caspasa-3 en células silenciadas para PRPF40B y control. La gráfica muestra la activación relativa de tres experimentos independientes (media \pm SEM). *, $P < 0,05$. **B.** Análisis de los niveles de expresión de PRPF40B, corte de la procaspasa-3 (precursor p35 y subunidades cortadas p20/p17) y BCL2. La expresión de las proteínas se detectó por *western blot*, utilizando anticuerpos específicos en los extractos celulares de células control y silenciadas para PRPF40B. ERK2 fue utilizado como control de carga. El ratio (p20+p17)/p35 se calculó para determinar los valores de las isoformas de la caspasa. La cuantificación de las bandas se muestra debajo de cada panel.

Estos datos apoyan la hipótesis de un papel regulador de PRPF40B en el proceso de apoptosis y sugieren una relación funcional entre este proceso y el control del *splicing* alternativo de genes relacionados con la muerte celular.

5.4 El silenciamiento de PTB aumenta el porcentaje de células en apoptosis

Con el fin de apoyar nuestros resultados con PRPF40B medimos el porcentaje de células sub-G1 y la unión de anexina-V al silenciar PTB. Como ya se ha comentado anteriormente, PTB es un represor del exón 6 de Fas, y por tanto su silenciamiento favorece la formación de la isoforma larga apoptótica. En ambos casos, observamos un incremento significativo del porcentaje de células apoptóticas cuando se deplecionó PTB, medido por incorporación de anexina-V (Figura R-20A) y cuantificación de la población celular en fase Sub-G1 (Figura R-20B). Estos resultados apoyan la influencia entre el cambio observado en el procesamiento del ARN de factores de apoptosis y el desarrollo de procesos de muerte celular.

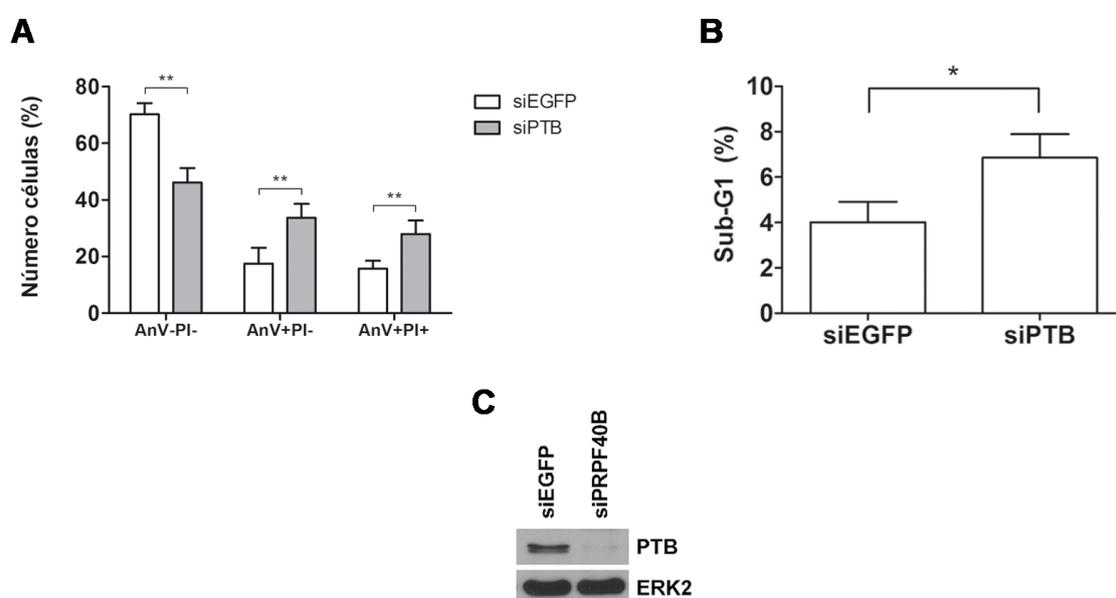


Figura R-20. El silenciamiento de PTB incrementa el porcentaje de células apoptóticas. **A.** Células HEK293T silenciadas con siPTB o siEGFP (control) fueron incubadas en ausencia (-) o presencia (+) de anexina-V y analizadas por citometría de flujo. Las gráficas muestran el porcentaje de células procedente de tres experimentos independientes (media \pm SEM). **, $P < 0,001$. **B.** Representación gráfica del porcentaje de células apoptóticas (Sub-G1) analizado por citometría de flujo de células silenciadas transitoriamente con siPTB o siEGFP (control). La gráfica muestra el porcentaje de células de tres experimentos independientes (media \pm SEM). *, $P < 0,05$. **C.** Una fracción de los lisados celulares fue analizada por *western blot* utilizando anticuerpos que detectan PTB o ERK2.

6. ANÁLISIS GLOBAL DEL SILENCIAMIENTO DE PRPF40B EN EL TRANSCRIPTOMA CELULAR

6.1 Análisis de los datos obtenidos usando el *Human Transcriptome Array* (HTA) en células SH-SY5Y

Con el objetivo de obtener una visión general del papel de PRPF40B en la célula como factor relacionado con el metabolismo del ARNm, decidimos realizar un análisis genómico del silenciamiento de esta proteína. Para elegir el contexto más apropiado para este análisis, investigamos los niveles de expresión por *western blot* de la proteína PRPF40B en líneas celulares de diferentes orígenes: HEK293T (células embrionarias renales), HepG2 (carcinoma hepatocelular), G3G1 (melanoma), MCF7GFPCJ (cáncer de mama), U87 (glioblastoma), SH-SY5Y (neuroblastoma), 3T3 (fibroblastos), HeLa (cáncer de cuello uterino) y Jurkat (linfocitos T).

Los resultados muestran una expresión diferencial de PRPF40B en líneas celulares relacionadas con el sistema nervioso, principalmente en células SH-SY5Y de neuroblastoma, respecto al resto de las líneas celulares ensayadas (Figura R-21). Nótese que a pesar de la alta expresión de PRPF40B en las líneas celulares HEK293T y HeLa, con las que se ha trabajado a lo largo de este trabajo, los niveles de expresión de PRPF40B son bastante menores en comparación con U87 o SH-SY5Y. Probablemente un alto nivel de expresión sea un indicativo de la importancia del papel de PRPF40B en el sistema nervioso. Los resultados obtenidos se correlacionan con los datos existentes en las bases de datos públicas referentes a la expresión de PRPF40B (<http://www.proteinatlas.org/ENSG00000110844-PRPF40B/cell/HPA038426>). En base a estas observaciones, y a los datos publicados donde se relaciona a este factor con disfunciones neurodegenerativas (ver *Introducción*), decidimos utilizar la línea celular SH-SY5Y para realizar el estudio genómico de PRPF40B.

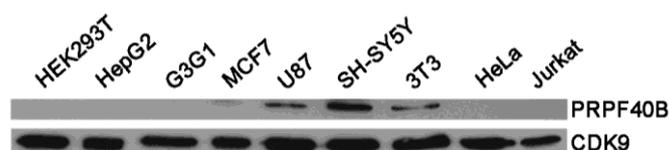


Figura R-21. PRPF40B muestra una expresión diferencial en líneas celulares de origen nervioso. Análisis de expresión de la proteína PRPF40B por *western blot* en extractos celulares procedentes de las líneas celulares: HEK293T (células embrionarias renales), HepG2 (carcinoma hepatocelular), G3G1 (melanoma), MCF7 (cáncer de mama), U87 (glioblastoma), SH-SY5Y (neuroblastoma), 3T3 (fibroblastos), HeLa (cáncer cuello uterino) y Jurkat (linfocitos T). CDK9 se utilizó como control de carga.

Dado que las células de neuroblastoma son difíciles de transfectar transitoriamente con siRNAs, generamos una línea celular silenciada establemente para PRPF40B mediante shRNAs (trabajo realizado por Juan Pablo Muñoz-Cobo Belart). Se obtuvieron dos líneas celulares establemente silenciadas para PRPF40B con dos shRNAs distintos. El grado de silenciamiento se comprobó analizando extractos celulares por técnicas de *western blot* (Figura R-22). Elegimos la línea silenciada con el shPRPF40B3 para experimentos posteriores puesto que fue en la que se observó consistentemente un mayor fenotipo del silenciamiento.

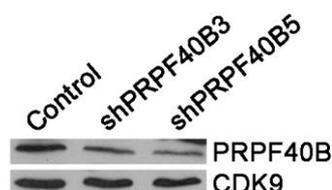


Figura R-22. Nivel de expresión de PRPF40B tras el silenciamiento en células SH-SY5Y con shRNAs. Representación del análisis de expresión por *western blot* de extractos celulares procedentes de células SH-SY5Y silenciadas con un shRNA control (pLKO-1), shPRPF40B3 o shPRPF40B5. Las proteínas fueron detectadas utilizando anticuerpos específicos para PRPF40B y CDK9.

A partir de extractos celulares procedentes de la línea SH-SY5Y control y silenciada (PRPF40B KD) se purificaron tres muestras de ARN para realizar el estudio global del transcriptoma humano usando el *Human Transcriptome Array* (HTA) de Affymetrix. El HTA pertenece a la última generación de *arrays* para el estudio de los perfiles de expresión génica. Se caracteriza porque utiliza un mayor número de sondas tanto en regiones codificantes como no codificantes, proporcionando una mayor cobertura y precisión necesarias para la detección de cada una de las isoformas

generadas por cada gen. De este modo, permite detectar cambios a nivel de transcripción, *splicing* alternativo, SNPs y ARNs no codificantes (ncRNA). El HTA tiene una sensibilidad equivalente a un RNA-seq de profundidad media.

Los cambios a nivel de transcrito y de exón se identificaron con el algoritmo EASANA (de la Grange et al., 2010). Los resultados obtenidos del análisis de los datos del HTA muestran que el silenciamiento de PRPF40B en células SH-SY5Y da lugar a 1197 cambios de expresión a nivel de transcripción (1022 se dan en genes conocidos) y a 102 cambios (102 genes) a nivel de *splicing*, de los cuales 42 genes son comunes a ambos grupos (Figura R-23A). Concretamente, de los 1197 cambios detectados a nivel de transcripción, 756 genes lo hacen disminuyendo su expresión (*down-regulated*) y 441 genes la aumentan (*up-regulated*). De los 102 genes que cambian a nivel de *splicing* alternativo, en 56 bajan su expresión y en 46 se ve incrementada (Figura R-23B). Estos datos fueron en principio sorprendentes, puesto que aunque no descartábamos el papel de PRPF40B en transcripción por tratarse de una proteína con dominios WW y FF, su estrecha relación con el factor de *splicing* en levaduras, Prp40, nos hizo centrarnos en la función de la proteína como reguladora de *splicing*. Todos estos datos, sin embargo, sugieren que PRPF40B podría estar actuando como un activador de la transcripción en células SH-SY5Y.

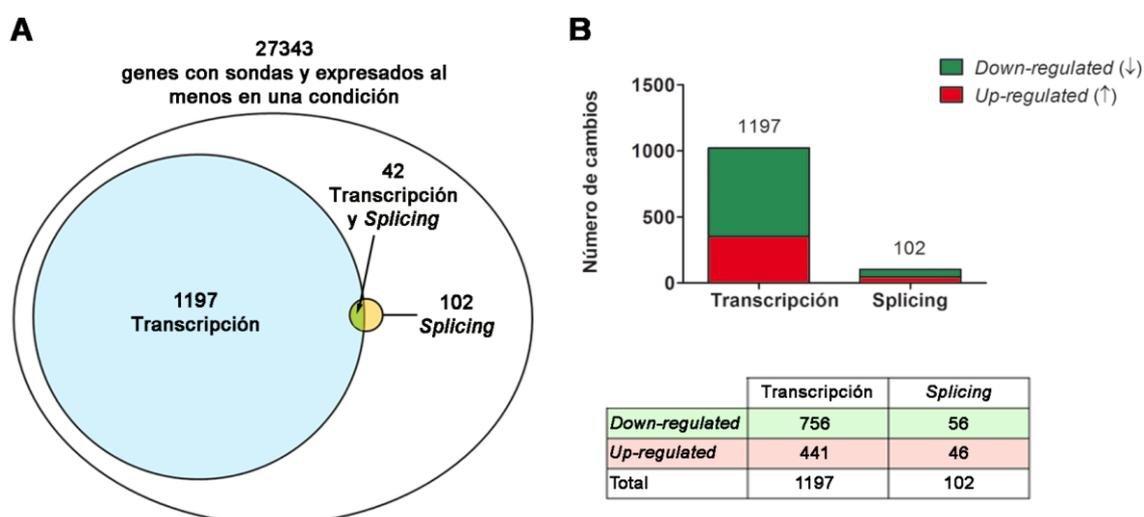


Figura R-23. Análisis global del silenciamiento de PRPF40B en células SH-SY5Y. A. Diagrama esquemático de los resultados generales obtenidos en el HTA. Se representan la proporción de cambios de regulación detectados a nivel de transcripción (azul), a nivel de *splicing* (amarillo) y los que coinciden en ambos grupos (verde). B. Representación gráfica del tipo de cambio de regulación detectado: disminución de la expresión (*down-regulated*, verde) o incremento de la expresión (*up-regulated*, rojo) tanto a nivel de transcripción como a nivel de *splicing*.

Por otra parte y para obtener una idea general de la función de PRPF40B en las células de neuroblastoma, quisimos estudiar las principales rutas afectadas. Para ello se utilizaron bases de datos del genoma, que organizan grupos de genes a gran escala en mapas pertenecientes a diversas rutas en que están implicados, como *KEGG* (<http://www.genome.jp/kegg/pathway.html>) y *Reactome* (<http://www.reactome.org/>). Teniendo en cuenta aquellos grupos con un mayor número de genes afectados tenemos que los procesos relacionados con cáncer y los procesos de señalización y transporte, en especial los relacionados con el sistema nervioso, son mayoritariamente afectados ante el silenciamiento de PRPF40B en células SH-SY5Y (Figura R-24).

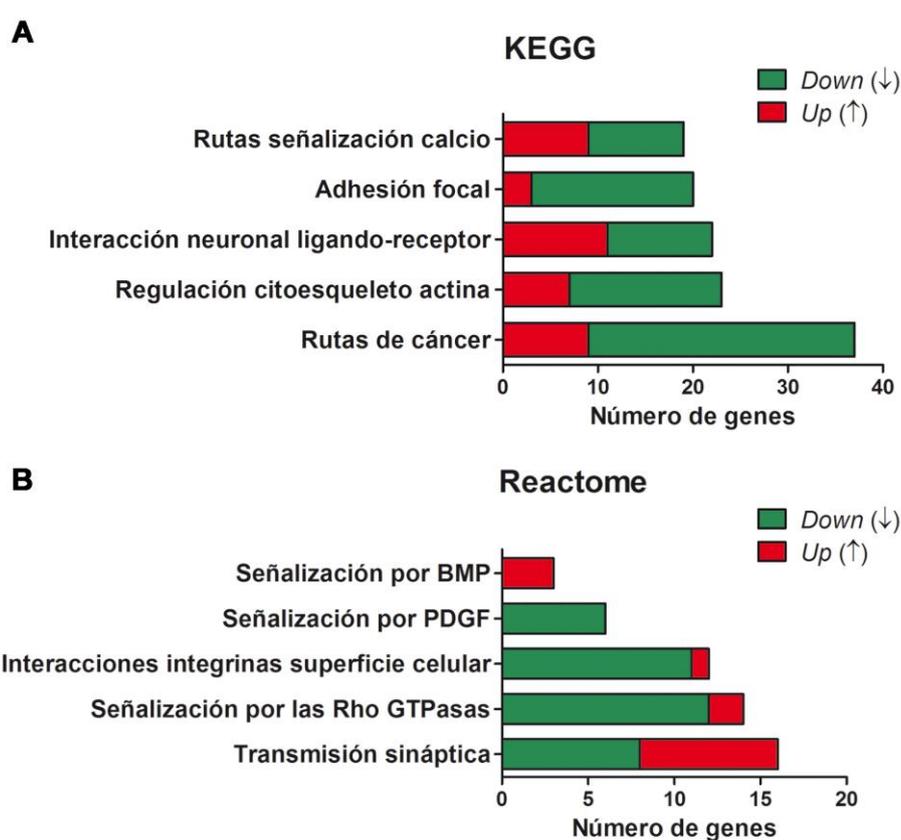


Figura R-24. El silenciamiento de PRPF40B en células SH-SY5Y altera rutas de cáncer, señalización y transporte. Representación gráfica de grupos de genes organizados en rutas atendiendo al criterio de las bases de datos *KEGG* (A) y *Reactome* (B). Cada categoría incluye genes que se encuentran activados (*up*) o inhibidos (*down*) ante el silenciamiento de PRPF40B en células SH-SY5Y.

6.2 Validación de genes regulados a nivel de transcripción

Para comprobar la veracidad de los datos obtenidos en el HTA, seleccionamos un conjunto de genes de cada grupo para conseguir una validación experimental representativa del total de datos. Para ello, nos centramos en genes con una mayor expresión diferencial, basándonos en el porcentaje de cambio (*fold-change*), así como algunos de los genes de mayor interés en nuestro estudio.

Entre los genes cuyos niveles de transcripción están considerablemente disminuidos ante el silenciamiento de PRPF40B en células SH-SY5Y, escogimos *NNAT*, *TNFAIP6*, *CACNG2*, *CCNA1*, *SV2C*, *CBLN4*, *LHX9* y *GBP4*. Además, como se describió en la *Introducción* (Apartados 4.2.2 y 8.3) y se amplía en apartados posteriores, la proteína PRPF40B ha sido implicada en síndromes mielodisplásicos, que se caracterizan por una predisposición a la leucemia mieloide aguda (AML) (Yoshida et al., 2011). Por tanto, nos pareció interesante validar genes implicados en AML para futuros experimentos. Para ello, nos basamos en el criterio establecido por la base de datos KEGG. A partir del grupo de genes incluidos en “Rutas de cáncer”, examinamos nuevamente en “KEGG Disease” si estos genes estaban categorizados en “Acute myeloid leukemia”. En esta búsqueda encontramos cinco genes: *LEF1*, *CCNA1*, *RUNX1T1*, *TCF7* y *PIK3R2*, implicados en diferentes puntos de la ruta que desencadena AML (Figura R-25). El conjunto de genes elegidos para realizar la validación de los resultados del HTA se muestra en la Tabla R-1.

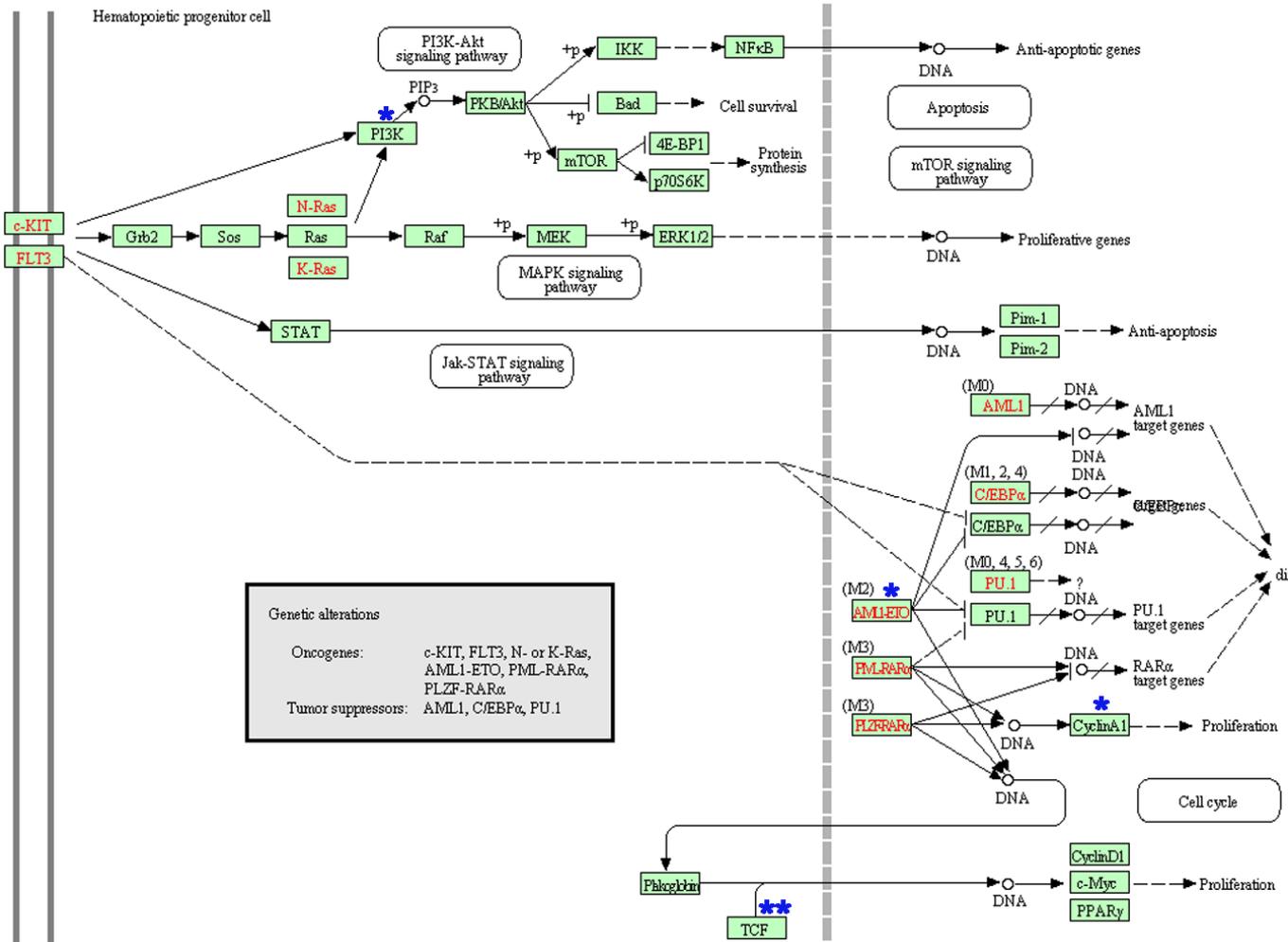


Figura R-25. Genes del HTA afectados en AML. Representación esquemática de las rutas afectadas en el desarrollo de AML durante la diferenciación hematopoyética. Se indican con un asterisco azul, los genes cuya expresión se encuentra alterada en el HTA del silenciamiento de PRPF40B o genes estrechamente relacionados con los mismos. Se señalan PI3K (*PIK3R2*), AML1-ETO (*RUNX1T1*), Cyclin A1 (*CCNA1*) y TCF (*TCF7*, *LEF1*).

Símbolo del gen	Regulación	Grado del cambio	Valor <i>P</i>	Nombre del gen
NNAT	Down	135,4	6,67e-04	Neuronatina
CACNG2	Down	23,36	3,42 e-04	Canal de calcio, dependiente de voltaje, subunidad gamma 2
TNFAIP6	Down	17,48	1,11 e-03	Factor de necrosis tumoral, proteína 6 alpha-inducida
SV2C	Down	13,94	8,28 e-04	Glicoproteína 2C vesícula sináptica
LEF1	Down	10,92	9,64 e-04	Factor linfoide de unión a enhancer 1
CCNA1	Down	2,84	1,09 e-02	Ciclina A1
RUNX1T1	Down	2,39	2,12 e-02	Factor de transcripción relacionado con Runt 1
TCF7	Down	1,59	2,48 e-03	Factor de transcripción 7 (específico células T)
CBLN4	Up	17,65	2,80 e-04	Precursor de cerebelina 4
LHX9	Up	8,46	6,38 e-03	LIM homebox 9
GBP4	Up	8,33	3,18 e-03	Proteína de unión a guanilato 4
PIK3R2	Up	1,5	9,45 e-03	Fosfoinositido-3-quinasa, subunidad reguladora 2 (β)

Tabla R-1. Lista de genes elegidos para validar los resultados del HTA a nivel de transcripción. Se muestran los datos de los doce genes seleccionados para evaluar su expresión diferencial a nivel de transcripción. Se indican el símbolo del gen, el tipo de cambio que experimentan, el grado del cambio (*fold-change*), la fiabilidad del valor obtenido (Valor *P*) y la descripción del gen. Los genes *NNAT*, *TNFAIP6*, *CACNG2*, *CCNA1*, *SV2C*, *CBLN4*, *LHX9* y *GBP4* fueron elegidos por su cambio significativo. Los cinco genes *LEF1*, *CCNA1*, *RUNX1T1*, *TCF7* y *PIK3R2* fueron seleccionados por estar implicados en las rutas de AML.

Para la validación de los genes se extrajo ARN a partir de las mismas líneas celulares utilizadas para obtener las muestras del HTA, células SH-SY5Y silenciadas establemente con shControl y shPRPF40B. A continuación, se analizaron los niveles de expresión mediante RT-qPCR, utilizando como control interno GAPDH. Los resultados corroboraron los datos obtenidos en el análisis del HTA, observando una disminución en la expresión de *NNAT*, *TNFAIP6*, *CACNG2*, *LEF1*, *CCNA1*, *SV2C*, *RUNX1T1* y *TCF7* (Figura R-26A), y un aumento en la expresión de *CBLN4*, *LHX9*, *GBP4* y *PIK3R2* (Figura R-26B). Para muchos de los genes, los resultados obtenidos a partir de la RT-qPCR presentan cambios mayores de lo esperado en el análisis del HTA. Este hecho puede ser debido a limitaciones relacionadas con los niveles de saturación en la lectura del *array*.

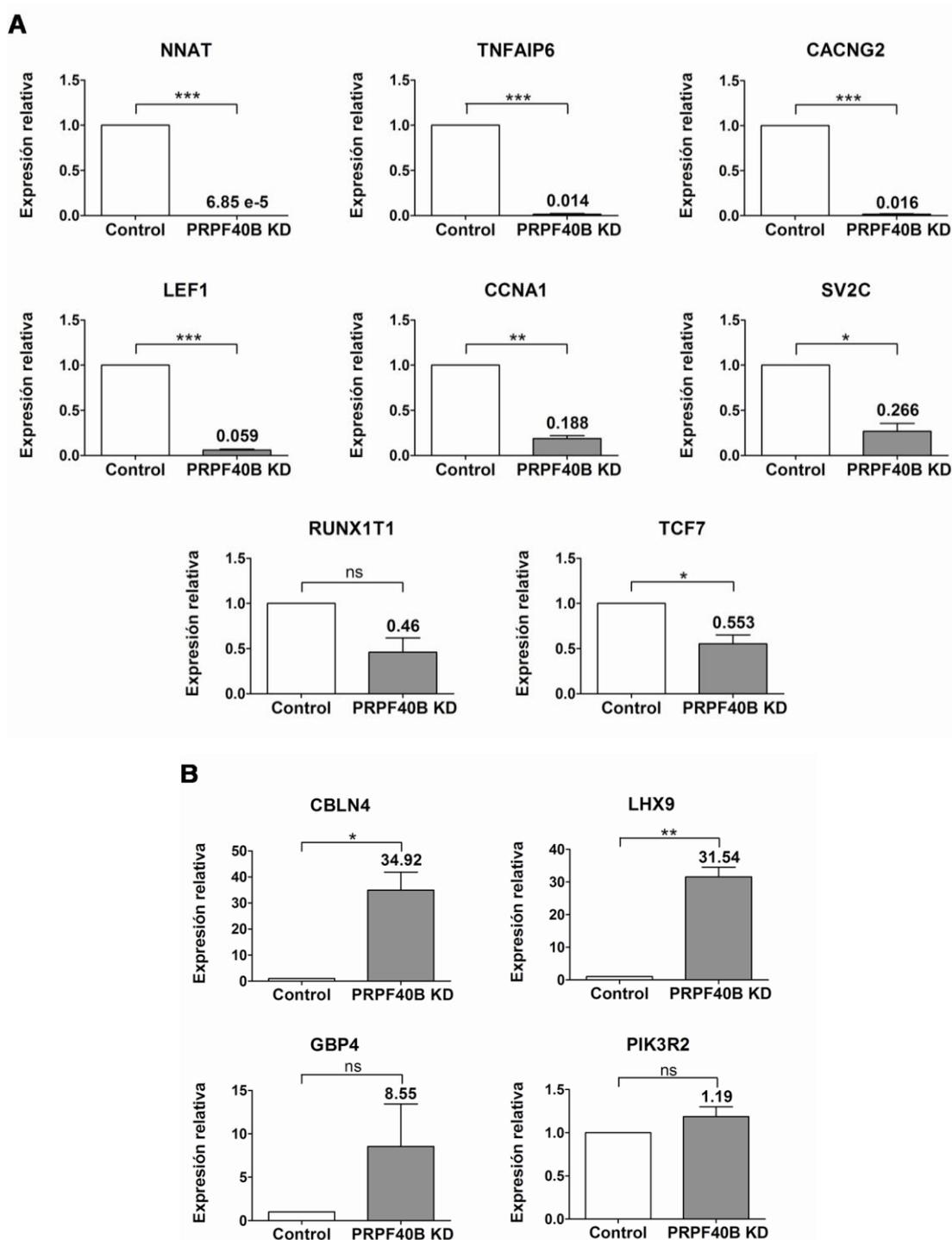


Figura R-26. La validación de los genes que cambian a nivel de transcripción es consistente con los datos obtenidos en el HTA. Representación gráfica de los resultados obtenidos de la RT-qPCR de genes con una disminución (*down-regulated*, **A**) o un incremento (*up-regulated*, **B**) en sus niveles de expresión. Las muestras proceden de las líneas celulares SH-SY5Y establemente transfectadas con el shControl (Control) o shPRPF40B (PRPF40B KD). Las gráficas muestran los niveles de expresión relativa al control que es 1. Los niveles de expresión fueron normalizados utilizando como control interno el gen GAPDH. Los datos proceden de tres experimentos independientes (media \pm SEM). Ns, no significativo; *, $P < 0,05$; **, $P < 0,001$; ***, $P < 0,005$.

6.3. Validación de genes regulados a nivel de *splicing*

Como ya se describió anteriormente, el método utilizado en el análisis del HTA no sólo permite observar cambios a nivel de expresión sino también variación en los eventos de *splicing* alternativos. Ya se ha comentado que 102 genes veían alterada su regulación en el *splicing* alternativo con el silenciamiento de PRPF40B. En total, hemos identificado 187 cambios en los eventos de *splicing*, de los cuales 91 veían disminuida su frecuencia (*down-regulation*) mientras que los 96 restantes eran incrementados (*up-regulation*). Dentro de los tipos de eventos conocidos los más alterados fueron la exclusión de exones (*exon skipping*), seguido del uso de primeros exones alternativos y últimos exones alternativos (Figura R-27).

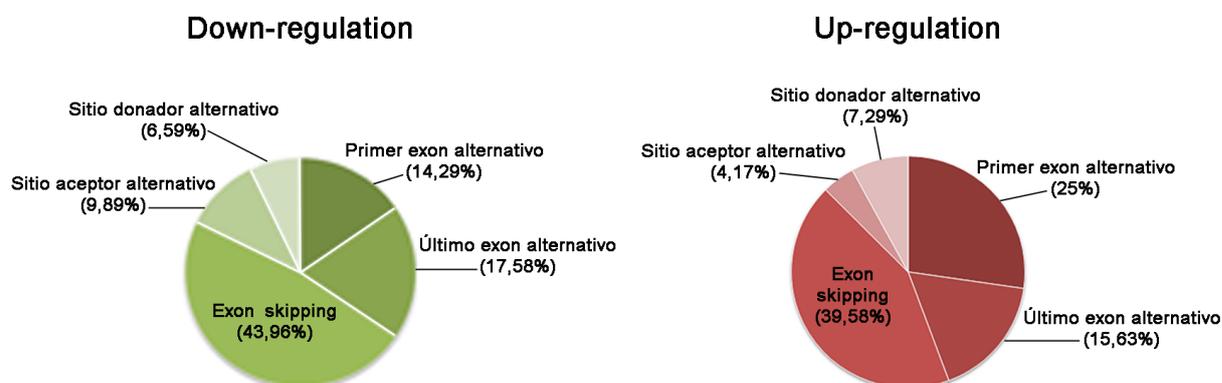


Figura R-27. La exclusión de exones es el evento más afectado por el silenciamiento de PRPF40B. Representación gráfica de la distribución de los eventos de *splicing* alternativo alterados cuando se silencia PRPF40B en células de neuroblastoma SH-SY5Y. Se clasifican en: exclusión de exón (*exon skipping*), último exón alternativo, primer exón alternativo, sitio donador alternativo de *splicing* y sitio aceptor alternativo de *splicing*. Los sectores indican el porcentaje del evento regulado respecto al número total de eventos que disminuyen su frecuencia (izquierda, *down-regulation*) o la aumentan (derecha, *up-regulation*).

Para evaluar los datos obtenidos de los genes afectados, al igual que se realizó en el apartado anterior con la regulación transcripcional, quisimos corroborar los resultados de los cambios de regulación de genes a nivel de *splicing*, en este caso con una muestra menor puesto que el grupo era más reducido. Para ellos elegimos algunos de los genes cuyos eventos sufrían un mayor nivel de cambio con el objetivo de evaluar genes muy relacionados con PRPF40B. Entre ellos escogimos a *BLCAP* (exón 4, exclusión de exón y retención de intrón), *SMARCA2* (exon 27, evento desconocido), *KCNMA1* (exón 39,

último exón alternativo, sitio donador alternativo), *NRG1* (exón 19, último exón alternativo) y *SLC6A11* (exón 1, primer exón alternativo) (Tabla R-2).

Símbolo del gen	Regulación	Exón	Grado del cambio	Valor <i>P</i>	Descripción del gen
BLCAP	Down	4	15,15	6,06 e-04	Proteína asociada al cáncer de vejiga
SMARCA2	Down	27	4,45	6,93 e-03	Relacionado con SWI/SNF, regulador de cromatina dependiente de actina, subfamilia a, miembro 2
KCNMA1	Down	39	2,27	6,72 e-03	Canal de potasio de gran conductancia, activado por calcio, subfamilia M, miembro α -1
NRG1	Up	19	5,35	2,26 e-03	Neurregulina 1
SCL6A11	Up	1	5,49	1,13 e-02	Neurotransmisor transportador, familia 6, miembro 11

Tabla R-2. Lista de genes a evaluar para validar los resultados del HTA a nivel de exón. Se muestran los datos procedentes del análisis del HTA de los cinco genes donde se analizó la expresión diferencial de los exones alterados. Se indican el símbolo del gen, el tipo de cambio que experimentan, el exón de estudio, el grado del cambio (*fold-change*), la fiabilidad del valor obtenido (Valor *P*) y la descripción del gen. Los cinco genes, *BLCAP*, *SMARCA2*, *KCNMA1*, *NRG1* y *SLC6A11*, fueron elegidos por su cambio significativo.

Las muestras de ARN utilizadas para validar los datos del HTA, al igual que en el caso anterior, fueron obtenidas de líneas celulares de neuroblastoma SH-SY5Y silenciadas establemente con shControl o shPRPF40B. Estas muestras fueron analizadas por RT-qPCR utilizando como control interno exones constitutivos de cada gen. Los resultados obtenidos reproducen la tendencia esperada en el cambio de regulación de los exones afectados (Figura R-28).

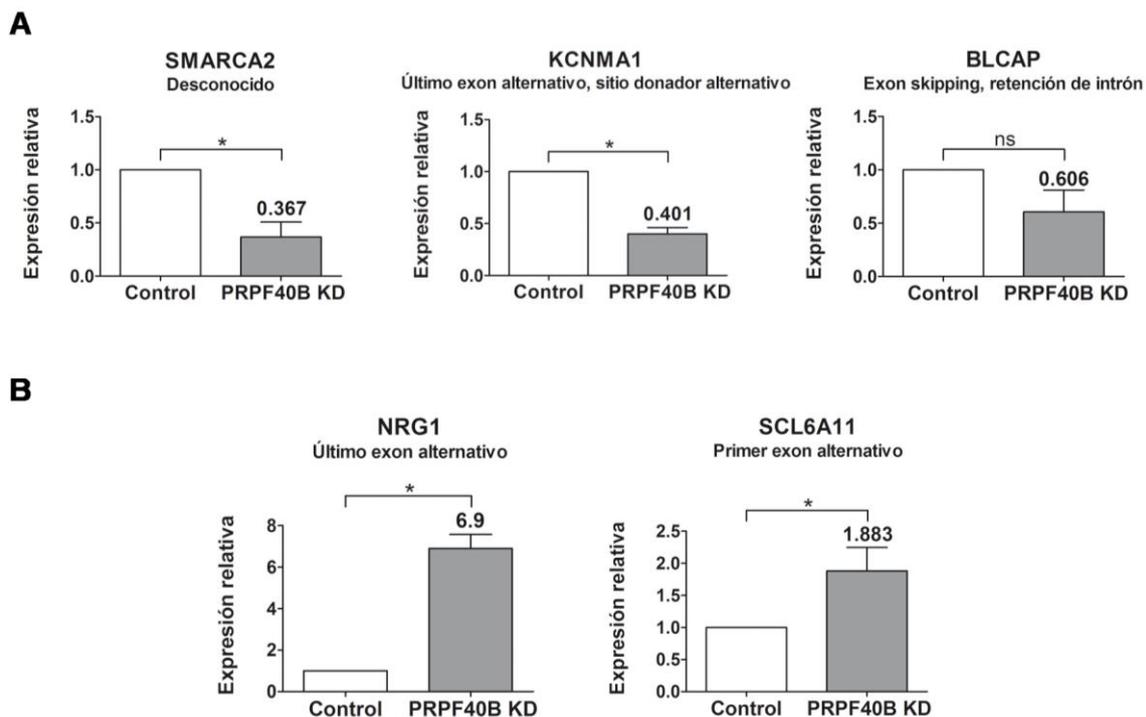


Figura R-28. La validación de los genes afectados a nivel de exón siguen la tendencia dictada por los datos del HTA. Representación gráfica de los resultados obtenidos por RT-qPCR de exones alterados cuando se silencia PRPF40B en células de neuroblastoma, disminuyendo (*down-regulation*, **A**) o incrementando su expresión (*up-regulation*, **B**). Las muestras proceden de las líneas celulares SH-SY5Y con el shControl (Control) y shPRPF40B (PRPF40B KD). Las gráficas muestran los niveles de expresión relativa al control que es 1. Los niveles de expresión fueron normalizados utilizando como control interno exones constitutivos de cada gen: exones 36/37 (SMARCA2), exones 21/22 (KCNMA1), exones 3/4 (BLCAP), exones 4/5 (NRG1) y exones 5/6 (SLC6A11). Los datos proceden de tres experimentos independientes (media \pm SEM). Ns, no significativo; *, $P < 0,05$.

7. ESTUDIO DE PRPF40B COMO ACTIVADOR DE LA TRANSCRIPCIÓN

7.1 PRPF40B activa la transcripción de promotores reporteros

Los datos obtenidos del HTA sugieren que PRPF40B podría actuar de forma general en las células SH-SY5Y promoviendo la activación de la transcripción, ya que al silenciar la proteína en esta línea celular la mayoría de los genes cambian a nivel de transcripción y aproximadamente dos tercios de los mismos lo hacen disminuyendo sus niveles de expresión.

Para abordar el posible papel de PRPF40B como activador de la transcripción utilizamos, como primera aproximación, un sistema reportero de luciferasa que contiene tres sitios de unión para el factor de transcripción NF- κ B (procedentes del gen *MHC-I*) seguidos de un elemento mínimo de iniciación conteniendo la caja TATA derivada del promotor *c-fos* murino. El número de unidades de luciferasa obtenido en los ensayos es directamente proporcional a la activación del promotor en las células. Para medir la capacidad de activación de la transcripción de PRPF40B, las células HEK293T se co-transfectaron con cantidades crecientes del vector que expresa PRPF40B o de un vector vacío junto con el plásmido reportero. Los resultados del ensayo de actividad luciferasa mostraron que PRPF40B es capaz de activar la transcripción dirigida desde este promotor reportero de una forma dependiente de dosis (Figura R-29B).

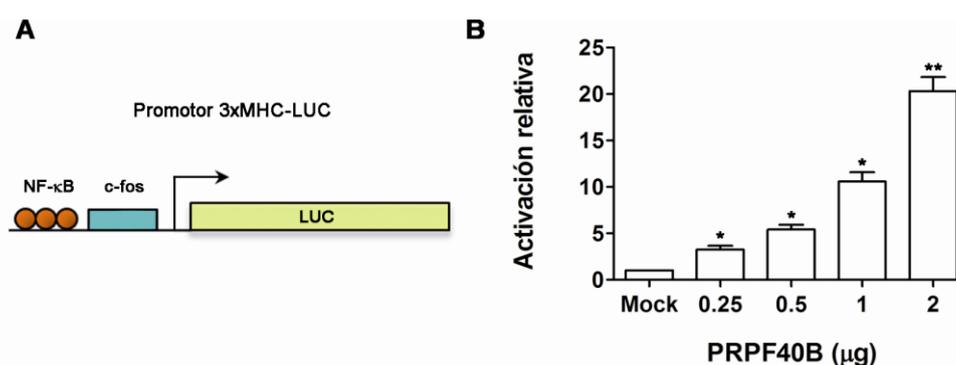


Figura R-29. PRPF40B activa la transcripción del promotor artificial p3xMHC-LUC. **A.** Representación esquemática del promotor artificial p3xMHC-LUC. Se indican los tres sitios de unión NF- κ B procedentes del gen *MHC-I* (círculos naranjas), el sitio mínimo de iniciación del promotor *c-fos* murino (rectángulo azul) y la secuencia del gen de luciferasa (rectángulo amarillo). **B.** Representación gráfica de la activación del promotor p3xMHC-LUC mediada por PRPF40B. Células HEK293T fueron co-transfectadas con cantidades crecientes del plásmido de expresión de PRPF40B, el plásmido p3xMHC-LUC y un plásmido de expresión de *Renilla* (pRL-TK-REN) como control interno del

experimento. La gráfica muestra los datos de tres experimentos independientes realizados en triplicado/duplicado (media \pm SEM). *, $P < 0,05$; **, $P < 0,001$.

Para determinar si PRPF40B es un activador general de la transcripción o su actividad depende del contexto o la presencia de determinados sitios de unión de factores de transcripción en el promotor, quisimos utilizar otros promotores basados en sistemas reporteros de luciferasa que disponíamos en el laboratorio. Para ello escogimos el plásmido pG5HIV-LUC que contiene cinco sitios de unión GAL4 y la secuencia mínima basal del promotor del gen *HIV* que contiene dos sitios NF- κ B, tres sitios Sp1 y una caja TATA, todo ello fusionado a la secuencia del gen de la luciferasa (Montanuy et al., 2008). Cuando co-transfectamos esta construcción junto con concentraciones crecientes del plásmido de expresión de PRPF40B observamos que la actividad transcripcional del promotor reportero aumentaba, aunque de una forma mucho más moderada que lo observado previamente con el promotor p3xMHCI-LUC (Figura R-30A).

Puesto que las dos construcciones utilizadas tenían sitios NF- κ B, quisimos descartar si esta actividad transcripcional podría ser causada por la presencia de estas secuencias de reconocimiento en el promotor. Para este experimento utilizamos el promotor pG5HIVNF κ Bm-LUC, similar al promotor del *HIV*, pero con los sitios NF- κ B mutados (Montanuy et al., 2008). Los resultados obtenidos revelaron que la actividad transcripcional mediada por PRPF40B se redujo considerablemente, reduciendo la capacidad de activación a menos de la mitad respecto al promotor sin mutar (Figura R-30B). Por tanto, aunque en este caso la actividad luciferasa es mucho menor a lo observado en los resultados previos, los datos obtenidos indican que los sitios NF- κ B podrían ser importantes para la capacidad de activación de la transcripción de PRPF40B.

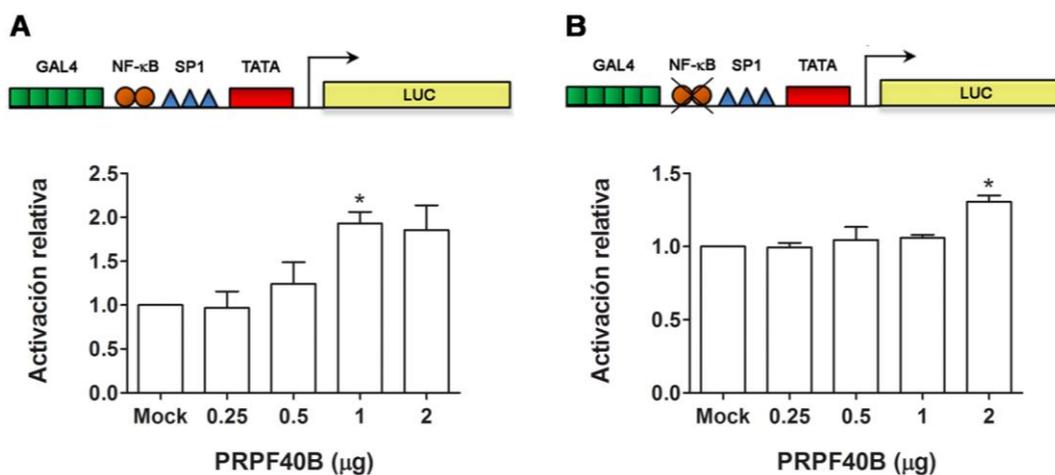


Figura R-30. Activación de la transcripción del promotor del HIV por PRPF40B. A. Representación esquemática de los promotores reporteros pG5HIV-LUC (A) o pG5HIVNF-κB (B) donde se indican los cinco sitios de unión GAL4 (cuadrados verdes), seguidos de dos sitios NF-κB (círculos naranjas), tres sitios Sp1 (triángulos azules), la caja TATA (rectángulo verde) y la luciferasa (rectángulo amarillo). En el caso del promotor pG5HIVNF-κB se muestran las mutaciones en los sitios NF-κB (tachados). Las células HEK293T fueron co-transfectadas con el promotor reportero correspondiente, concentraciones crecientes de PRPF40B y un plásmido pRL-TK-REN como control interno del experimento. Las gráficas muestran la activación relativa de los promotores reporteros de tres experimentos independientes realizados en triplicado/duplicado (media ± SEM). *, $P < 0,05$.

7.2. Importancia de los dominios WW y FF de PRPF40B en la activación de la transcripción

Los dominios WW y FF interaccionan con componentes de las maquinarias de transcripción y procesamiento del ARNm (ver *Introducción*). Por ello, quisimos investigar su papel en la activación de la transcripción mediada por PRPF40B. Con este objetivo generamos un conjunto de construcciones truncadas, donde se fueron eliminando gradualmente regiones del extremo 5' de la secuencia de PRPF40B pero conservando la zona carboxilo-terminal, lugar donde se encuentra la señal de localización nuclear (Figura R-31B). Las células HEK293T fueron co-transfectadas con cada una de estas construcciones o un vector vacío y el plásmido reportero p3xMHC-LUC, puesto que fue en el que se observaron los mayores niveles de actividad transcripcional. Los ensayos de actividad luciferasa obtenidos mostraron que la construcción de PRPF40B sin los dominios WW [Figura R-31C, PRPF40B (263-871)] fue capaz de activar al promotor reportero, incluso tres veces más que la proteína

completa, sugiriendo que estos dominios podrían estar ejerciendo una función represora, o bien que la conformación espacial de la proteína resultante hace que sea más activa que en condiciones nativas. En cambio, cuando además se eliminaron los dos primeros dominios FF de PRPF40B [Figura R-31C, PRPF40B (393-871)], la activación de la transcripción disminuyó aproximadamente a la mitad respecto al valor obtenido con la proteína completa. Finalmente, las construcciones que contenían la delección de los dominios FF3 [Figura R-31C, PRPF40B (474-871)] y FF4 [Figura R-31C, PRPF40B (618-871)], no fueron capaces de activar la transcripción de este promotor.

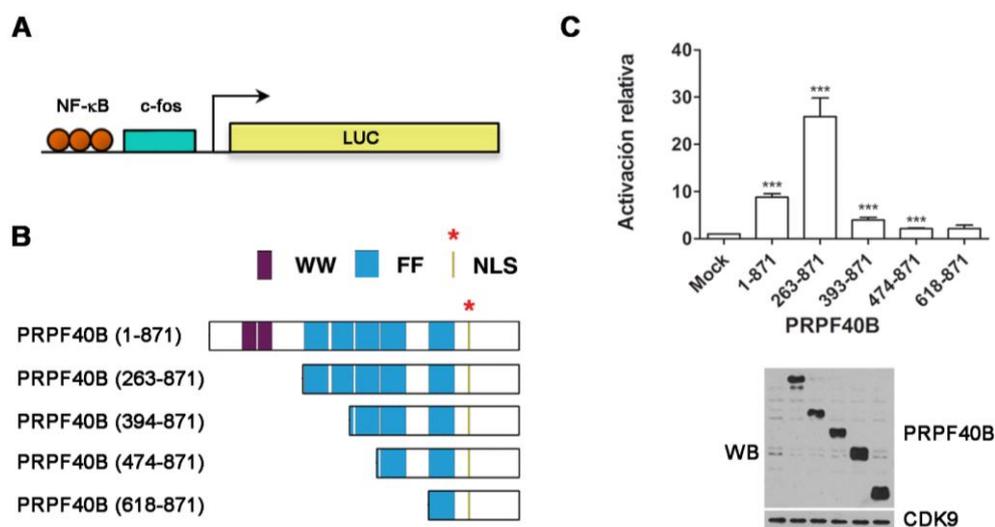


Figura R-31. Los dominios FF1 y FF2 son importantes para activar la transcripción. **A.** Diagrama esquemático del promotor reportero p3xMHC-LUC donde se indican los sitios NF-κB, el elemento mínimo del promotor *c-fos* y la secuencia de la luciferasa. **B.** Diagrama esquemático de las construcciones truncadas de PRPF40B. Entre paréntesis se indican la región que comprende la secuencia de aminoácidos. En cada estructura están trazados los dos dominios WW (violeta), los cinco dominios FF (azul) y la señal de localización nuclear, NLS (amarillo). **C.** Representación gráfica de la activación del promotor p3xMHC-LUC por la sobre-expresión de las distintas proteínas de PRPF40B. Células HEK293T fueron co-transfectadas con el plásmido reportero p3xMHC-LUC junto con cada una de las construcciones de PRPF40B y el plásmido pRL-TK-REN como control interno del experimento. La gráfica muestra los datos obtenidos de cuatro experimentos independientes realizados en triplicado/duplicado (media \pm SEM). ***, $P < 0,005$. Una fracción de los extractos celulares fue analizada por *western blot* con los anticuerpos indicados que reconocen PRPF40B y CDK9 (control).

Por tanto, todos estos resultados indican que PRPF40B es capaz de activar la transcripción, que puede estar mediada por el reconocimiento de sitios de unión de factores de transcripción como NF-κB y que los dominios FF1 y FF2 de la proteína pueden desempeñar un papel importante en esta actividad transcripcional.

8. IMPORTANCIA DE PRPF40B EN LA LEUCEMIA MIELOIDE AGUDA (AML)

Como ya se describió en la *Introducción* (Apartado 8.3), se han identificado mutaciones en componentes del *spliceosoma* implicados en el reconocimiento de los sitios de *splicing* en un estudio en pacientes con síndrome mielodisplásico (MDS), un grupo de anomalías que dan lugar a AML (Yoshida *et al.*, 2011). Uno de los factores identificados en ese estudio fue PRPF40B. Las mutaciones observadas en la secuencia de PRPF40B generan un cambio de aminoácido que se postula puede resultar en una falta de función de la proteína, alterando por ello el posible papel regulador de la proteína (Yoshida *et al.*, 2011).

Para comprobar si las mutaciones identificadas en PRPF40B afectan a la funcionalidad de la proteína modificamos, mediante mutagénesis dirigida, los nucleótidos alterados de PRPF40B identificados en ese estudio y que daban lugar a los cambios de aminoácidos: P383L, D442N y P540S (Figura R-32).



Figura R-32. Mutagénesis dirigida de los nucleótidos modificados en la secuencia de PRPF40B detectados en pacientes con MDS. Representación esquemática de los nucleótidos alterados de PRPF40B en pacientes con MDS. Se indican los codones donde tienen lugar la modificación de los nucleótidos (flecha), así como la base por la que es intercambiada (parte inferior). Se muestra el cambio de aminoácido que tiene lugar (rojo), tras mutar los nucleótidos indicados.

A continuación, llevamos a cabo ensayos funcionales para comparar la actividad de las construcciones mutadas con la proteína sin modificar (*wild type*). Basándonos en los datos de *splicing* alternativo del gen *Fas*, donde la sobre-expresión de PRPF40B favorece la isoforma más corta que excluye el exon 6, utilizamos los mutantes de PRPF40B para estudiar el efecto en el *splicing* alternativo de este minigen. Para ello comparamos cada una de las proteínas con las mutaciones P383L, D442N, P540S con la proteína *wild type*, sin embargo no apreciamos cambios significativos entre las distintas versiones de PRPF40B (Figura R-33).

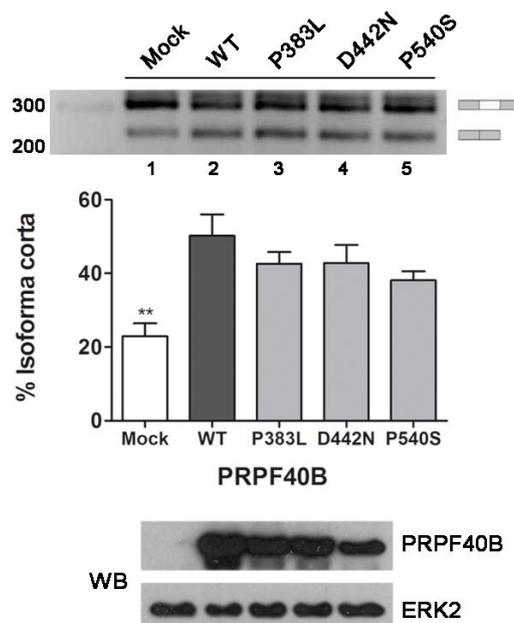


Figura R-33. Los mutantes de PRPF40B no ven alterado su efecto en el *splicing* alternativo del minigen de *Fas*. Se muestran los geles de agarosa donde se resuelven los productos de *splicing* generados mediante RT-PCR de células HEK293T co-transfectadas con el minigen de *Fas* junto con un vector vacío (Mock), PRPF40B *wild-type* (WT) o cada uno de los mutantes (P383L, D442N, P540S). La gráfica muestra los resultados del análisis densitométrico como porcentaje de isoforma corta a partir de tres experimentos independientes (media \pm SEM). **, $P < 0,001$. Una fracción de los lisados celulares fueron analizados por *western blot* utilizando anticuerpos que detectan PRPF40B y ERK2.

Debido al escaso efecto observado en los experimentos de *splicing*, quisimos investigar el comportamiento de las construcciones mutadas de PRPF40B en ensayos de activación transcripcional utilizando sistemas reporteros de luciferasa, de forma similar a experimentos previos. Para ello, co-transfectamos las células HEK293T con los promotores reporteros p3xMHC-LUC o pG5HIV-LUC junto con los vectores que expresan PRPF40B *wild-type*, las proteínas mutadas o un vector vacío de control. Los resultados obtenidos mostraron que la mutación P383L no afecta a la actividad transcripcional de la proteína en comparación con la proteína *wild type*. Sin embargo, las mutaciones D442N y P540S generaron proteínas que perdían completamente su capacidad de transactivación (Figura R-34).

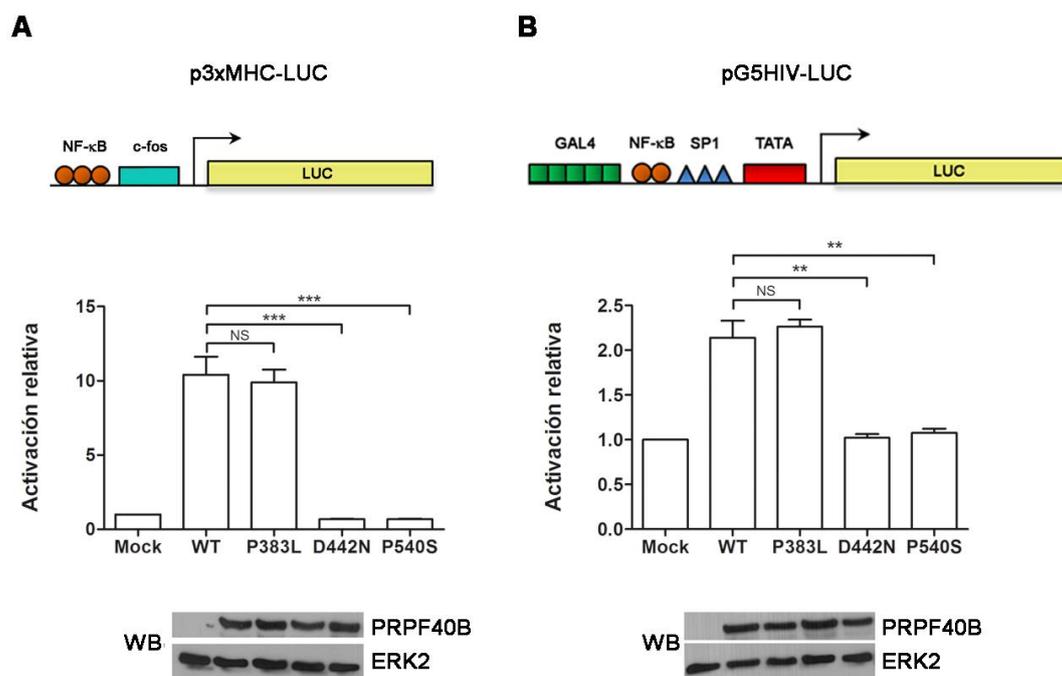


Figura R-34. Las mutaciones D442N y P540S de PRPF40B provocan una pérdida de la actividad transcripcional de PRPF40B. Ensayos de activación transcripcional con los sistemas reporteros de luciferasa p3xMHC-LUC (A) y pG5HIV-LUC (B). Se representan los esquemas de cada uno de los promotores artificiales donde se indican los sitios de unión en cada uno de ellos. La descripción de los esquemas es la misma que en figuras anteriores. Las células HEK293T se co-transfectaron con el correspondiente plásmido reportero, cada una de las construcciones de PRPF40B (WT, P383L, D442N, P540S) o un vector vacío (Mock) y el plásmido pRL-TK-REN como control interno del experimento. Las gráficas muestran los datos de la activación relativa de los promotores reporteros de tres experimentos independientes realizados en triplicado (media \pm SEM). **, $P < 0,001$; ***, $P < 0,005$. Una fracción de los extractos celulares fue analizada por *western blot* con anticuerpos específicos para PRPF40B y ERK2.

Estos resultados evidencian la pérdida de función de la proteína PRPF40B cuando se mutan los residuos D442N y P540S, sugiriendo un posible mecanismo molecular que puede estar afectado en pacientes con MDS.

Conclusiones

1. PRPF40B es una proteína estructuralmente conservada a lo largo de la evolución que contiene dos dominios WW y cinco dominios FF.
2. La distribución de PRPF40B es nucleoplasmática y presenta un enriquecimiento en los *speckles* nucleares.
3. La NLS de PRPF40B se encuentra en su extremo carboxilo terminal. Los dominios FF3 y FF4 junto con la NLS son importantes para la localización de PRPF40B en los *speckles* nucleares.
4. La localización de PRPF40B en los *speckles* nucleares no es afectada por el tratamiento con ARNasa A ni por la inhibición de la transcripción.
5. PRPF40B interacciona directamente con el factor de *splicing* SF1 *in vitro*. PRPF40B se asocia con los factores de *splicing* SF1 y U2AF⁶⁵ *in vivo* de manera dependiente de ARN.
6. PRPF40B promueve la selección de sitios de *splicing* alternativos en el ARN de los genes de apoptosis *Bcl-x* y *Fas*.
7. El reclutamiento de PRPF40B al sitio de *splicing* no es suficiente para su efecto sobre el *splicing* alternativo de *Fas*. PRPF40B necesita la presencia de secuencias en la región 5' del exón 6 de *Fas*, así como sitios 3'SS y 5'SS débiles para promover la exclusión de dicho exón.
8. El silenciamiento de PRPF40B aumenta el número de receptores de Fas/CD95 en la superficie celular y el porcentaje de muerte celular por apoptosis.
9. El estudio global del transcriptoma humano al silenciar PRPF40B en células de neuroblastoma SH-SY5Y, revela un papel relevante de PRPF40B sobre la transcripción y que PRPF40B podría estar implicado en procesos de cáncer, así como en señalización y transporte en el sistema nervioso.
10. PRPF40B activa la transcripción en genes reporteros. Los dominios FF1 y FF2 son importantes para esta activación.

11. Las mutaciones D442N y P540S de PRPF40B, asociadas al síndrome mielodisplásico, eliminan la capacidad de activación transcripcional de PRPF40B.

Anexo

1. ESTUDIO PRELIMINAR DE PRPF40A Y SU RELACIÓN CON PRPF40B

PRPF40A y PRPF40B son los homólogos humanos del factor esencial de *splicing* en levaduras Prp40. Estas proteínas comparten un porcentaje de identidad comprendido entre el 51-54% y de similitud del 65-68% al realizar alineamientos globales o locales con EMBOSS, respectivamente. Por tanto, nos pareció interesante estudiar de forma paralela el comportamiento de PRPF40A en relación con PRPF40B, en la medida de lo posible ya que no disponemos de construcciones plasmídicas de este gen, por su dificultad para clonarlo en un vector de expresión, y tuvimos que basar su estudio en experimentos de la proteína endógena.

Como se mostró en los Resultados (Apartado 2.1), el anticuerpo sintetizado para detectar PRPF40B también reconoce PRPF40A, por tanto, la tinción de la proteína endógena en los experimentos de microscopía podría mostrar la localización de ambos factores (Figura R-3A). Para comprobar la distribución de PRPF40B utilizamos el plásmido EGFP-PRPF40B, que tenía un patrón similar a la proteína endógena (Figura R-3B). Sin embargo, al no disponer de una construcción similar de PRPF40A, no podemos confirmar este dato y sólo podemos especular su enriquecimiento en los *speckles* nucleares.

Debido al gran parecido en las secuencias de los dos genes, tuvimos en cuenta la especificidad de los siRNAs. Por ello, de forma análoga a como hicimos con las células transfectadas con siPRPF40B (Resultados, Apartado 4.1, Tabla R-1), medimos por RT-qPCR los niveles de ARN de cada factor al transfectar las células HEK293T con siPRPF40A. Los datos indican que el nivel de ARN de PRPF40A se reduce aproximadamente un 40% sin afectar significativamente a PRPF40B (Tabla A-1).

	Nivel ARN PRPF40B	Nivel ARN PRPF40A
siPRPF40A	1,157±0,218	0,663±0,155

Tabla D-1. Nivel de ARN de PRPF40A y PRPF40B tras la transfección de siPRPF40A. Representación de los datos obtenidos de la RT-qPCR de células HEK293T silenciadas con siPRPF40A. Se midió la cantidad relativa de ARN de PRPF40A y PRPF40B respecto al control GAPDH.

Paralelamente al estudio del efecto de PRPF40B en el *splicing* de *Fas*, nos pareció interesante investigar el papel de PRPF40A en este proceso. Para ello

silenciamos células HEK293T con siPRPF40B, siPRPF40A o ambos siRNAs juntos y medimos los niveles de ARN de *Fas* endógeno. Sin embargo, a diferencia de los resultados observados con PRPF40B, no observamos un resultado concluyente sobre el efecto del silenciamiento de PRPF40A en el *splicing* del exón 6 de *Fas*. Además cuando deplecionamos ambos factores a la vez, no se observó una variación significativa en el *splicing* alternativo (Figura A-1).

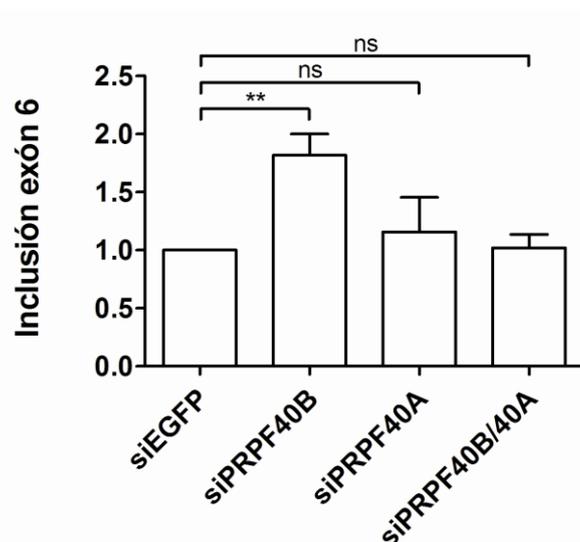


Figura A-1. Efecto del silenciamiento de PRPF40A en el *splicing* de *Fas*. Células HEK293T fueron transfectadas con siEGFP (control), siPRPF40B, siPRPF40A o siPRPF40B/siPRPF40A. Tras la extracción del ARN, las muestras se analizaron por RT-qPCR. La gráfica representa los niveles inclusión del exón 6 de *Fas* relativa al control que es 1. Los datos están normalizados respecto al control interno GAPDH y proceden de cinco experimentos independientes (media ± SEM). Ns, no significativo; **, $P < 0,001$.

A pesar de no observar un efecto de PRPF40A sobre el *splicing* alternativo del exón 6 de *Fas*, quisimos estudiar las consecuencias de su silenciamiento en la apoptosis. Para ello utilizamos el método de la doble tinción con anexina V y yoduro de propidio que marca las células en apoptosis tardía, cuando la membrana celular está afectada y el ADN se encuentra fragmentado. Los resultados observados revelan que tanto el silenciamiento de PRPF40B como PRPF40A aumentan el número de células en apoptosis respecto al control (Figura A-2). Sin embargo, el silenciamiento de las dos proteínas no mostró un aumento significativo en el número de células apoptóticas.

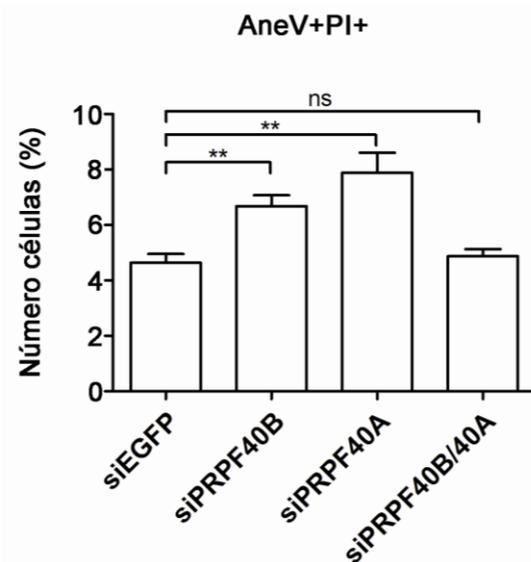


Figura A-2. Efecto del silenciamiento de PRPF40A y PRPF40B en la apoptosis. Células HEK293T silenciadas con siEGFP (control), si PRPF40B, si PRPF40A o siPRPF40B+siPRPF40A fueron incubadas con anexina-V y yoduro de propidio (PI), y posteriormente analizadas por citometría de flujo. Las gráficas muestran el porcentaje de células procedente de tres experimentos independientes (media \pm SEM). Ns, no significativo; **, $P < 0,001$.

Todos estos resultados indican que las proteínas homólogas humanas de Prp40, PRPF40A y PRPF40B podrían tener efectos distintos en la célula, incluso contrarios.

Discusión

El objetivo de este trabajo fue caracterizar el factor PRPF40B, uno de los homólogos humanos del factor esencial de *splicing* en levaduras Prp40, relacionado con el reconocimiento del sitio 5' de *splicing* e implicado en las interacciones que se establecen a ambos lados del intrón. Estas proteínas, caracterizadas estructuralmente por sus dominios de interacción proteína-proteína WW y FF, forman parte de un grupo de factores que se han relacionado con el metabolismo del ARNm, donde sus motivos estructurales tienen una función relevante.

Nuestros análisis bioinformáticos han identificado en la secuencia de aminoácidos de PRPF40B dos dominios WW y cinco dominios FF, muy conservados en ratón. El hecho de que el número de dominios FF ha aumentado uno más respecto a su homólogo en levaduras, así como la presencia de otra proteína humana procedente de Prp40, explica la complejidad de los organismos superiores y sus mecanismos de regulación.

Puesto que la información hasta ahora obtenida sobre PRPF40B era muy escasa y se especulaba sobre su posible papel en el *splicing*, nos pareció de interés estudiar sus características bioquímicas y funcionales para conocer la función de esta proteína en los mecanismos de regulación de la expresión génica celular.

1. PRPF40B ES UN FACTOR NUCLEAR ENRIQUECIDO EN LOS *SPECKLES* NUCLEARES

Experimentos de microscopía confocal mostraron la localización de PRPF40B en el nucleoplasma celular de una forma punteada, acumulándose en los *speckles* nucleares al igual que ocurre con otros factores de *splicing*. La distribución de PRPF40B en los *speckles* nucleares no se afectó tras el tratamiento con ARNasa y consecuente degradación del ARN, indicando que la localización a esas regiones no depende de la unión al ARN y puede estar sujeta a interacciones proteína-proteína.

El patrón de PRPF40B tampoco varió inhibiendo la transcripción, indicando que la localización de PRPF40B en los *speckles* nucleares no ocurre por su asociación a genes transcripcionalmente activos. Los *speckles* se encuentran cercanos a sitios de transcripción activa y el tratamiento de inhibición de la RNAPII provoca una disminución en el número de *speckles*, que se vuelven más grandes y redondeados

debido a la acumulación de factores de *splicing* (O'Keefe et al., 1994). Los resultados mostraron que PRPF40B tenía un comportamiento similar a otros factores de *splicing* al ser reclutada a los *speckles* bajo condiciones de inhibición de la transcripción.

Por el momento sólo podemos especular sobre las interacciones que establece PRPF40B para acumularse en los *speckles* nucleares y el motivo de su presencia. En principio, parece establecer uniones proteína-proteína que la dirigen a estos orgánulos. Los *speckles* nucleares están compuestos principalmente por factores de *splicing*, aunque también se localizan otras proteínas relacionadas con el metabolismo del ARNm, así como quinasas, fosfatasas, proteínas estructurales y lncRNA (Spector and Lamond, 2011). Aunque todavía desconocemos el significado funcional de la localización de PRPF40B en los *speckles* nucleares, estos datos proporcionan la primera evidencia de su relación con procesos relacionados con el metabolismo del ARNm.

2. CARACTERIZACIÓN DE ELEMENTOS ESTRUCTURALES NECESARIO PARA LA LOCALIZACIÓN DE PRPF40B

El análisis bioinformático de la señal de localización nuclear identificó una secuencia consenso KRRRR en la región carboxilo terminal. Este resultado es similar a otras proteínas relacionadas, como su homólogo en levaduras Prp40 o la proteína humana PRPF40A, que contienen la NLS en el extremo carboxilo terminal tras los dominios FF. Esto indica la importancia de conservación de secuencias. El análisis mediante microscopía confocal de una construcción truncada a la que faltaba el extremo carboxilo terminal, corroboró la predicción obtenida *in silico*.

Las imágenes de microscopía confocal del conjunto de construcciones de PRPF40B con deleciones seriadas en la región amino terminal, muestran que el dominio FF3 es importante para la correcta distribución de PRPF40B y que el dominio FF4 es capaz de dirigirse por sí mismo hacia los *speckles* nucleares. Estas observaciones indican que los dominios WW no intervienen en la localización de PRPF40B sino que son los dominios FF los importantes para la correcta distribución de la PRPF40B, al igual que ocurre con otras proteínas ricas en dominios WW/FF (Sánchez-Alvarez et al., 2006). Nuestros resultados muestran que la eliminación de los dominios FF3 y FF4 disminuye la localización de PRPF40B en los *speckles* y dirige a la proteína hacia el

nucléolo, sobre todo al eliminar FF4 en la construcción FF5-NLS, donde PRPF40B se encuentra casi exclusivamente en el nucléolo. Los datos sobre la construcción FF4-NLS, capaz de dirigirse por sí misma a los *speckles*, sugiere que aunque el dominio FF3 puede participar en la estabilización de la proteína para su correcta localización, la presencia de FF4-NLS es suficiente para la distribución de PRPF40B en los *speckles* nucleares. De todas formas sería necesario estudiar una construcción FF3-NLS y la delección de FF3 y FF4 en el contexto de la proteína entera, para comprobar si el dominio FF3 al igual que FF4 se dirige a los *speckles* nucleares de forma independiente, si es la combinación de ambos dominios FF3/FF4 o de estos dos junto con la participación de otros motivos lo que funciona como un elemento estructural de localización en PRPF40B. Se sabe que los dominios FF participan de forma coordinada e interacciones débiles de cada uno es lo que en conjunto generan una gran afinidad de unión por el ligando (Carty et al., 2000; Smith, Kulkarni, and Pawson, 2004). En cambio se ha descrito el papel de los dominios FF4/FF5 de TCERG1 como una unidad funcional única para dirigir a TCERG1 o proteínas heterólogas a los *speckles* nucleares (Sanchez-Hernandez et al., 2012).

La eliminación progresiva de los dominios FF tiene como consecuencia la pérdida de localización en los *speckles* y su acumulación en el nucléolo. Este dato sugiere que los dominios FF estarían evitando que PRPF40B se dirija hacia el nucléolo. Probablemente el acortamiento de la proteína genera estructuras poco funcionales que agregan y se dirigen hacia nucléolo para ser degradadas, ya que se ha relacionado este orgánulo nuclear con la degradación de proteínas ribosomales (Lam et al., 2007). Sin embargo, al comparar la distribución de las construcciones más pequeñas y de longitud similar observamos que FF4-NLS mantiene considerablemente su localización en los *speckles*, mientras que FF5-NLS se dirige casi en su totalidad al nucléolo. Este hecho indica que en efecto, existen dominios independientes y capaces de dirigir por sí mismos a una proteína hacia un orgánulo específico, en este caso, los *speckles* nucleares. En cambio, existen dominios que son dependientes del contexto de la proteína y no solo pierden su localización habitual, puesto que no contienen señal que lo dirijan, sino que también son funcionalmente dependientes de otros motivos que eviten que agreguen. Además de la hipótesis sobre la agregación de PRPF40B no funcional por defecto en el nucléolo, podemos especular sobre la existencia de una señal de localización nucleolar en el dominio FF5, que en condiciones normales sería

enmascarada por la señal de localización a los *speckles*. Este dato habría que corroborarlo generando una construcción FF1 o FF2, que parecen no dirigir a PRPF40B a los *speckles*, fusionados a FF5-NLS y comprobar su localización o no en el nucléolo. Aunque el nucléolo ha sido involucrado en diversas funciones celulares como el ciclo celular, estrés celular y biogénesis de los ribosomas, el motivo por el que PRPF40B estaría presente en el nucléolo estaría por determinar.

Aunque es necesario profundizar en la capacidad del dominio FF4 de dirigir a PRPF40B a los *speckles* y la importancia del dominio FF3 en la correcta distribución de la proteína, podríamos decir que las señales NLS y de localización en *speckles* parecen comportarse de manera independiente ya que al eliminar esta última, la proteína se mantiene en el núcleo. Estas señales seguirían comportamientos similares a los dominios FF4/FF5 de TCERG1, aunque en este caso la NLS se encuentra en el centro de la proteína separando los dominios WW y FF (Sanchez-Hernandez et al., 2012). Para corroborar esta hipótesis sobre la independencia de la NLS y la señal de localización de los dominios FF3/FF4 de PRPF40B sería necesario el estudio de una construcción de PRPF40B en la que sólo faltaran estos dos dominios.

3. PRPF40B INTERACCIONA CON LOS COMPONENTES DEL *SPLICEOSOMA* SF1 Y U2AF⁶⁵

Basándonos en los datos publicados sobre la asociación en levaduras de Prp40 con los homólogos de SF1 y U2AF⁶⁵, decidimos estudiar esta interacción en humanos. Observamos que PRPF40B se asocia con estos factores de *splicing* tanto *in vitro* como *in vivo*. Sin embargo, el tratamiento con ARNasa A disminuyó la asociación de PRPF40B con SF1 y U2AF⁶⁵ *in vivo* sin afectar la unión de estos dos últimos. A pesar de la interacción directa que parece establecerse entre PRPF40B y SF1 (datos *in vitro*), los experimentos de co-inmunoprecipitación indican que en células vivas esta unión es dependiente de la integridad del ARN, sugiriendo que PRPF40B interacciona con ARN, o bien, que esa asociación está establecida a través de otras proteínas unidas al ARN. El bloqueo funcional de U1 ARN mediante *morpholinos* (comprobado en nuestros experimentos de *splicing*) no cambió el resultado de las interacciones, indicando que probablemente la partícula U1 snRNP no media esta interacción. La naturaleza del ARN que está mediando esa interacción es desconocida.

Estos datos son relevantes porque son la primera evidencia directa que relacionan a PRPF40B con la formación del *spliceosoma*, puesto que SF1 y U2AF⁶⁵ son factores esenciales en el reconocimiento del sitio 3'SS.

4. PRPF40B MODULA EL *SPLICING* ALTERNATIVO

El ensamblaje de las primeras subunidades del *spliceosoma* en levaduras ocurre mediante la unión de la subunidad U1 snRNP al sitio 5'SS y Mud2p al tracto de polipirimidina, seguido de la unión de Bbp a la región BP (*branch point*). En este contexto Prp40, que forma parte de U1 snRNP, parece conectar ambos extremos del exón. Por tanto, se ha propuesto que las proteínas humanas homólogas a Prp40 puedan estar ejerciendo de mediadoras para conectar el sitio 5'SS con la región *BPS* durante la formación temprana del *spliceosoma* (Abovich and Rosbash, 1997; Padgett, 2012). Nuestros resultados muestran que PRPF40B es capaz de modular el *splicing* alternativo de los minigenes apoptóticos *Bcl-x* y *Fas*, promoviendo en ambos casos la isoforma corta, mientras su depleción provoca un aumento de la isoforma larga. En el caso de *Fas*, este efecto es visible también en el gen endógeno. Para el caso de *Bcl-x*, y a pesar de múltiples aproximaciones, no pudimos detectar el gen endógeno (datos no mostrados). Nuestros datos muestran el primer indicio funcional de PRPF40B como un factor regulador del *splicing* alternativo, al influir en el cambio del patrón de procesamiento del ARNm.

También mostramos que la presencia de sitios 3'SS y 5'SS débiles es necesaria para que PRPF40B ejerza su función de forma eficiente sobre el *splicing* alternativo del gen *Fas*, evitando que se produzca el reconocimiento del exón 6. Además, secuencias de este exón próximas al sitio 3'SS son importantes para mediar la función de PRPF40B. Estas secuencias podrían estar relacionadas con el reclutamiento de SF1 y U2AF⁶⁵ situados en el sitio 3'SS y con los que hemos visto que PRPF40B se asocia. También son importantes las secuencias exónicas que incluyen o se solapan con el elemento silenciador URE6, conocido por mediar el efecto represor de PTB. Sin embargo, a pesar de encontrar este conjunto de secuencias por el que PRPF40B ejerce su efecto sobre el gen de *Fas*, no conseguimos identificar un elemento consenso común a todos ellos que caracterice el efecto mediador de PRPF40B en el *splicing* (datos no mostrados).

Estos datos son coherentes con las interacciones previamente descritas de PRPF40B con SF1 y U2AF⁶⁵ durante la primera fase de formación del *spliceosoma* que conducen a la definición del exón durante el *splicing* de *Fas*. Sin embargo, nuestros datos muestran que PRPF40B participa en el *splicing* alternativo de *Fas* promoviendo la exclusión del exón (Figura D-1), cuando la idea que se tiene sobre las proteínas relacionadas con Prp40, es que funcionan como puentes de unión entre los sitio 5'SS y 3'SS del intrón para estabilizar el complejo E y definir el exón. Sin embargo, PRPF40B parece no actuar como un represor de *splicing* porque el reclutamiento directo de PRPF40B a través de los sitios MS2 no resultó en una exclusión del exón 6 de *Fas*. Por tanto, aunque los datos indican que PRPF40B participa en la selección del sitio de *splicing* durante la primera etapa de formación del *spliceosoma*, todavía queda por determinar el mecanismo mediante el cual está ejerciendo su efecto. En este contexto, señalar que intentamos comprobar si el comportamiento de PRPF40B era similar al de Prp40 mediante el rescatare de su función en un ensayo de complementación en levaduras; sin embargo, no obtuvimos resultados concluyentes debido a motivos técnicos.

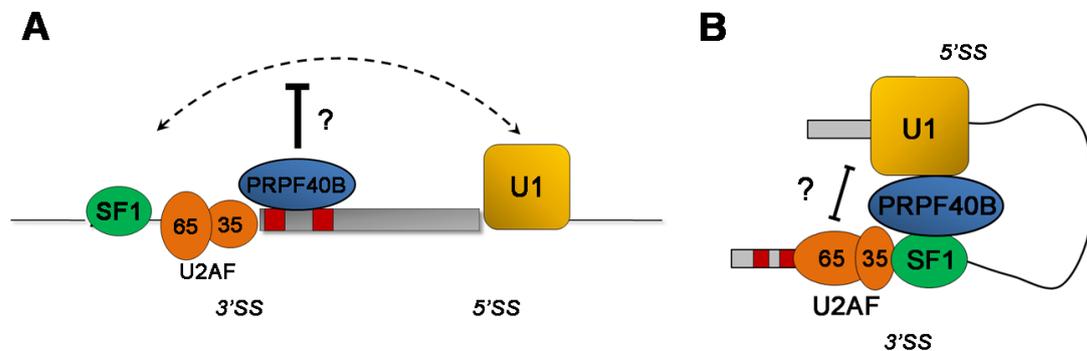


Figura D-1. Modelo del posible mecanismo de regulación del *splicing* alternativo del exón 6 de *Fas* por PRPF40B. Diagrama esquemático del mecanismo de inhibición de PRPF40B en la selección del sitio de *splicing* del exón 6 de *Fas* durante la definición del exón (A) o en la definición del intrón (B). Se representa la subunidad del *spliceosoma* U1 snRNP (U1), los factores de *splicing* SF1 y U2AF (subunidades 65 y 35 kDa), PRPF40B y los sitios 5'SS y 3'SS.

En resumen, los datos obtenidos sugieren que PRPF40B desempeña un papel en *splicing* alternativo del pre-ARNm durante los primeros estadios de formación del *spliceosoma*: PRPF40B se localiza en los *speckles* nucleares, enriquecidos en factores de *splicing*; PRPF40B interacciona con los factores esenciales de *splicing* SF1 y

U2AF⁶⁵, implicados en el reconocimiento del sitio 3'SS, y su asociación es dependiente de ARN; y PRPF40B regula el *splicing* alternativo de *Bcl-x* y *Fas*. Estos resultados son consistentes con lo esperado de un homólogo humano de Prp40, pudiendo estar implicado en la selección del sitio de *splicing* en eucariotas superiores, aunque es necesario continuar el estudio para determinar este punto.

5. PRPF40B INTERVIENE EN EL PROCESO DE APOPTOSIS

Los resultados muestran que PRPF40B modula el *splicing* alternativo de *Fas* y *Bcl-x*, genes implicados en el control de la apoptosis. El silenciamiento de PRPF40B aumenta la producción de la isoforma anti-apoptótica de *Bcl-x* e incrementa la formación de la isoforma pro-apoptótica de *Fas*. Estos resultados que en principio son contradictorios, también se han descrito para el modulador de *splicing* RBM10, cuya expresión favorece la isoforma anti-apoptótica de *Fas* y la isoforma pro-apoptótica de *Bcl-x* (Inoue et al., 2014). En cualquier caso, esos datos han servido para evaluar las consecuencias biológicas de la depleción de PRPF40B en el proceso de apoptosis. Además de los efectos observados en *Fas* y *Bcl-x*, quisimos investigar si otros genes apoptóticos tenían afectado su patrón de procesamiento. Sin embargo, en el caso de c-FLIP_L y c-FLIP_S no obtuvimos resultados significativos (datos no mostrados). De acuerdo con estos resultados, el número de receptores de *Fas* unidos a la superficie celular aumenta al silenciar PRPF40B (Figura R-16). Además hemos visto que al disminuir la expresión de PRPF40B se produce un incremento de inducción de apoptosis (Figuras R-17, R-18, R-19).

Aunque son necesarios más experimentos para explicar la relación entre el *splicing* de genes apoptóticos y la inducción de apoptosis, estos datos sugieren que PRPF40B altera el *splicing* alternativo de *Fas* para regular la supervivencia celular.

6. PRPF40B SE EXPRESA DIFERENCIALMENTE EN EL TEJIDO NERVIOSO Y FUNCIONA COMO ACTIVADOR DE LA TRANSCRIPCIÓN

PRPF40B se expresa de manera diferencial en el tejido nervioso y particularmente en la línea celular de neuroblastoma SH-SY5Y. El sistema nervioso es uno de los tejidos con mayor presencia de factores de *splicing* y donde más se expresan, por tanto, este dato apoya indirectamente un papel de PRPF40B en la regulación del

splicing y sugiere que es precisamente en este tipo celular donde desarrolla su función. Por eso, fue de gran interés analizar de forma general el papel de PRPF40B en el transcriptoma celular dentro de este contexto.

El análisis del global del transcriptoma humano mediante el HTA de Affymetrix reveló que el silenciamiento de PRPF40B en células SH-SY5Y genera 1197 cambios a nivel de transcripción y 102 a nivel de *splicing*, de los cuales, había 42 genes comunes a ambos grupos. De los 1197 cambios a nivel de transcripción, 756 lo hacía disminuyendo su expresión y 441 la aumentaban; mientras que de los 102 genes que cambiaban a nivel de exón, 56 bajaron su expresión y en 46 se incrementó. Estos datos sugieren que PRPF40B se comporta en el tejido nervioso como un activador de la transcripción. Este dato fue en parte sorprendente puesto que su ancestro en levaduras, Prp40, es un factor esencial de *splicing* y por ello los experimentos funcionales de esta tesis se centraron en el papel de PRPF40B en el procesamiento del ARNm. Sin embargo, se ha descrito que Prp40 interacciona con el CTD de la RNAPII y ha sido propuesto como posible acoplador entre procesos de transcripción y *splicing* (Morris and Greenleaf, 2000). De hecho si observamos en el HTA el grupo de genes en común entre los cambios detectados a nivel de transcripción y de exón, nos damos cuenta que casi la mitad de los genes alterados a nivel de *splicing* coinciden con genes que tienen alterada su expresión. Este número es un porcentaje bastante alto, lo que indica que en efecto, PRPF40B puede comportarse como un acoplador entre ambos procesos en genes específicos.

Por otra parte, el análisis de datos atendiendo a las rutas celulares alteradas, mostraron mayoritariamente genes implicados en cáncer y regulación del citoesqueleto. Estos datos concuerdan con los principales cambios que se detectan a nivel genómico individual. Entre los genes que más disminuyen su expresión bajo condiciones de silenciamiento de PRPF40B, encontramos el factor de necrosis tumoral *TNFAIP6*, así como diversos genes relacionado con la sinapsis y conexión intercelular, como el canal de calcio *CACNG2*, una glicoproteína de vesícula sináptica *SV2C* y cabe destacar la considerable disminución observada de $6,85 \times 10^{-5}$ veces en la neuronatina, *NNAT*, un proteolípido que relacionado con la regulación de los canales de iones durante el desarrollo cerebral (Sharma, Mukherjee et al., 2013). De forma similar, dentro de los genes que ven incrementados sus niveles de expresión se encuentran el precursor de la cerebelina 4, *CBLN4*, implicado en la formación de la sinapsis en diferentes puntos o el factor de transcripción LIM homeobox 4, *LHX9*. Datos del laboratorio han mostrado

que el citoesqueleto de actina se encuentra alterado en las células SH-SY5Y deplecionadas de PRPF40B (datos no mostrados). Estos datos están de acuerdo con los cambios observados en el transcriptoma de genes implicados en citoesqueleto cuando la expresión de PRPF40B se encuentra inhibida.

Si nos fijamos en los cambios producidos a nivel de exón, observamos que el evento más recurrente bajo el silenciamiento de PRPF40B es el *exon skipping*, como ocurre con el *splicing* alternativo del exón 6 de *Fas*, que supone casi la mitad de los casos anotados. Dentro de los mayores cambios observados a nivel de exón destacan BLCAP, una proteína asociada al cáncer de vejiga, SMARCA2, una subunidad del remodelador de cromatina SWI/SNF, un factor de un canal de potasio, KCNNA1, un neurotransmisor y la neuregulina 1. Ya que se ha descrito la estrecha relación existente entre alteraciones producidas en el procesamiento del ARNm y el cáncer, al generar patrones de *splicing* aberrantes que alteran el normal funcionamiento de procesos esenciales (Grosso, Martins and Carmo-Fonseca, 2008). Por otra parte, el citoesqueleto tiene un papel fundamental en la función neuronal, debido al tráfico de factores implicados en procesos como sinapsis o diferenciación neuronal. Por tanto, parece que PRPF40B está participando en procesos de transporte de moléculas relacionado con la sinapsis. Para corroborar estos resultados serán necesarios experimentos funcionales que permitan profundizar en el papel que está desarrollando PRPF40B en estas células.

Los resultados del HTA muestran que PRPF40B tiene una función bastante amplia como activador transcripcional en las células SH-SY5Y y nuestros resultados muestran que PRPF40B es capaz de activar la transcripción usando los vectores p3xMHCI-LUC y pG5HIV-LUC, de una forma dependiente de dosis. Cuando se mutaron los sitios de unión NF- κ B (comunes a ambos promotores) en la segunda construcción generando el sistema reportero pG5HIVNF κ Bm-LUC, PRPF40B perdía considerablemente su capacidad de activación. Sería interesante para futuros experimentos mutar o eliminar esos sitios NF- κ B en la construcción del MHC-I, puesto que es donde se observa una mayor activación de la transcripción, para estudiar la posible pérdida de actividad de PRPF40B. Aunque a priori estos sitios de unión NF- κ B parecen relevantes en la activación transcripcional de PRPF40B, al analizar los sitios de unión de los genes afectados a nivel de transcripción del HTA, utilizando la herramienta DiRE (*Distant Regulatory Elements of co-regulated genes*, <http://dire.dcode.org/>), sólo

detectamos cinco sitios NF- κ B en los promotores de los veintiuno detectados en sus genes dianas.

Los ensayos de actividad luciferasa muestran que los dominios FF de PRPF40B, concretamente los dominios FF1 y FF2 tienen una gran capacidad de activación de la transcripción, ya que al eliminar los dominios WW la activación es mucho mayor que la obtenida con la proteína entera. En cambio al quitar los dos primeros dominios FF y los siguientes de forma progresiva la activación disminuye considerablemente, por debajo del nivel de la proteína completa. Estos datos indican que los dominios WW no son importantes en la activación transcripcional e incluso podrían ejercer un efecto inhibitor, bien por sí mismos o por modificar la conformación de la proteína ocultando parcialmente los dominios de activación. En cambio los dominios FF sí que son relevantes para dicha función, principalmente los dos primeros motivos, aunque la presencia de los dominios FF3 y FF4 parecen incrementar la afinidad de unión por el ligando.

Los resultados obtenidos apoyan que PRPF40B es capaz de activar la transcripción. Aunque son necesarios más experimentos y estudios *in vivo* que corroboren esta función, estos datos son el primer indicio experimental de PRPF40B en transcripción.

7. IMPLICACIÓN DE PRPF40B EN AML

Estudios de secuenciación en pacientes con MDS identificaron mutaciones en un grupo de genes relacionados con las primeras fases del ensamblaje del *spliceosoma* entre los aparecieron tres mutaciones en PRPF40B que generaban un cambio de aminoácido, además de numerosos SNPs (Yoshida et al., 2011). Ensayos funcionales de *splicing* alternativo en el gen *Fas*, utilizando construcciones de PRPF40B que contenía cada una de las mutaciones asociadas a MDS, no mostraron cambios significativos en el procesamiento del ARN. En cambio, ensayos de activación de la transcripción con los promotores reporteros p3xMHCI-LUC y pG5HIV-LUC mostraron una drástica bajada de la activación cuando PRPF40B contenía las mutaciones D442 y P540S, y no se apreciaron cambios significativos con la mutación P383L.

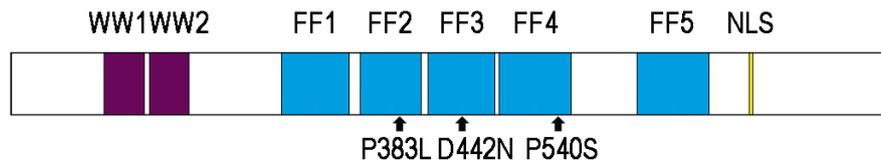


Figura D-5. Representación gráfica de las mutaciones de PRPF40B asociadas a MDS. Diagrama esquemático de PRPF40B donde se muestran las mutaciones P383L, D442N y P540S asociadas a MDS. Se indican los dominios WW (morado), los dominios FF (azul) y la NLS (amarillo).

Estos datos corroboran la importancia individual que desempeña cada dominio FF en la unión al ligando durante la activación transcripcional, puesto que, aunque anteriormente vimos que los principales dominios de activación eran FF1 y FF2, mutaciones puntuales en los dominios FF3 y FF4 son capaces de anular la funcionalidad de PRPF40B. Fue sorprendente que la mutación en el dominio FF2 no tuviera apenas efecto sobre la proteína, aunque puede ser que sea el dominio FF1 quien realmente ejerza su efecto en activación. De todas formas, atendiendo a la naturaleza del cambio de aminoácido, observamos que en el caso de la mutación P383L se da un cambio de aminoácidos no polares, a diferencia de la mutación D442N donde se sustituye un aminoácido ácido por otro polar sin carga y P540S en el que un aminoácido no polar o hidrofóbico da paso a uno polar sin carga. En el primer caso la estructura de la proteína podría no resultar muy afectada por esa mutación a diferencia de las otras dos, donde los aminoácidos sustituidos son de diferente carga o naturaleza pudiendo generar desestabilización en la estructura. Cambios en un único aminoácido pueden provocar profundas remodelaciones en la conformación de la proteína que conducen a cambios en la especificidad de las uniones a los ligandos y alteraciones en su función (Alexander et al. 2009).

Aprovechando los datos del HTA, decidimos realizar una búsqueda de genes implicados en la ruta de AML cuya expresión está alterada bajo condiciones de silenciamiento de PRPF40B. Encontramos cinco candidatos que cambiaban a nivel de transcripción, pero ninguno a nivel de exón: *LEF1*, *CCNA1*, *RUNX1T1*, *TCF7* cuyos niveles de ARNm estaban disminuidos considerablemente tras la depleción de PRPF40B en células SH-SY5Y y *PIK3R2* donde se observó un ligero aumento en su expresión. Estos datos indican que mutaciones en PRPF40B podrían estar bajando la activación de la transcripción de estos genes, tres de ellos factores de transcripción.

Aunque sería interesante estudiar estos cambios en líneas celulares procedentes de leucemias, estos resultados proporcionan la primera relación funcional entre mutaciones asociadas a esta enfermedad y alteraciones que podrían generarse como consecuencia de cambios, en este caso, causados por anomalías en PRPF40B.

8. SIMILITUDES Y DIFERENCIAS ENTRE PRPF40A Y PRPF40B

El hecho de que el factor esencial de *splicing* en levaduras Prp40 tenga dos homólogos humanos y de secuencia parecida despierta interés por la posible relación funcional que pueda haber entre ellas. Por eso realizamos algunos experimentos de forma paralela entre PRPF40A y PRPF40B. Aunque no tenemos datos que lo confirmen, *a priori* podemos sugerir que la localización de PRPF40A es principalmente nucleoplasmática con un enriquecimiento en los speckles nucleares de forma similar a lo que ocurre con PRPF40B y otras proteínas relacionadas con dominios WW/FF.

Los datos de las inmunoprecipitaciones obtenidos con PRPF40B son similares con los experimentos previos que mostraban la asociación de los dominios WW del ortólogo PRPF40A con SF1 y U2AF⁶⁵ en ensayos de *pull down* (Lin, Lu, and Tarn, 2004). En nuestros experimentos también aparece una fuerte unión *in vitro* entre la región de PRPF40B que contienen los dominios WW con SF1. Este resultado es similar a los datos anteriormente observados sobre PRPF40A, aunque no es sorprendente debido a la alta homología entre sus dominios WW. Los porcentajes de identidad y similitud resultantes de los alineamientos locales y globales entre los dominios WW1 son 58% y 78% respectivamente, mientras que para los dominios WW2 esos porcentajes son del 81% y 96%. Aunque los ensayos de interacción se realizaron con construcciones que contenían ambos dominios WW, sería interesante estudiar la especificidad de cada uno de los dominios para determinar la especificidad de interacción con SF1.

Los resultados de los ensayos de *splicing* del gen *Fas*, no muestran un claro efecto de PRPF40A sobre el procesamiento del exón 6 de este gen. Sin embargo, estos datos no se corresponden con los resultados obtenidos en un trabajo reciente en el que identifican reguladores del *splicing* alternativo del exón 6 de *Fas* a gran escala, donde se ha encontrado a PRPF40A, indicando su participación en este proceso (Tejedor,

Papasaikas and Valcarcel, 2014). Aunque son necesarios más experimentos para confirmar estos resultados, nuestros datos muestran que el silenciamiento de ambas proteínas no genera un cambio en el procesamiento del exón 6 de *Fas*, sugiriendo que PRPF40A podría estar contrarrestando el efecto observado con PRPF40B.

Los ensayos de apoptosis muestran que el silenciamiento de PRPF40A genera un aumento significativo del porcentaje de células apoptóticas, al igual que ocurre cuando se depleciona PRPF40B. Sin embargo, al silenciar ambas proteínas, disminuye el porcentaje de muerte celular por apoptosis, prácticamente igual al nivel de las células control. Este dato es sorprendente, puesto que se esperaba en este último caso un aumento en la muerte celular como consecuencia de la suma de los efectos de ambas proteínas. Podemos interpretar este resultado por el efecto contrario que pueden estar ejerciendo estos dos factores relacionados, afectando a las rutas de apoptosis en distintos puntos de manera que por sí solos sí generan un aumento de muerte celular, pero juntos inhiben la acción del otro.

Estos datos sugieren que la divergencia evolutiva ha generado a partir de una proteína común, dos factores que podrían tener funciones relacionadas pero en principio contrarias. De momento desconocemos si PRPF40A y PRPF40B reconocen los mismos sustratos, aunque ya se han descrito ligandos comunes a ambas proteínas, como la huntingtina o como nuestros datos revelan, los factores de *splicing* SF1 y U2AF⁶⁵. En este caso, estas dos proteínas podrían actuar por un mecanismo de competición por un mismo sustrato, que explicaría la anulación de los efectos observados al silenciar ambas. Otra posibilidad es que exista una autorregulación entre PRPF40A y PRPF40B, como ocurre con otros factores de *splicing* como PTB/nPTB que controlan de forma recíproca su expresión (Boutz et al., 2007). En cualquier caso, aunque es evidente la relación entre ambas proteínas, son necesarios más experimentos que permitan dilucidar el papel que desempeña cada una de ellas juntas y por separado.

9. PRPF40B COMO ACOPLADOR DE LOS PROCESOS DE TRANSCRIPCIÓN Y *SPLICING*

PRPF40B ha sido relacionado con el acoplamiento físico entre los procesos de transcripción y *splicing*. Aunque no tenemos datos que confirmen el mecanismo funcional entre ambas maquinarias, nuestros resultados aportan evidencias sobre la

implicación de PRPF40B en ambos procesos y características comunes a otros factores donde sí se conoce mejor la influencia entre los dos procesos.

En primer lugar, PRPF40B contiene dos dominios WW y cinco dominios FF en su estructura. Estos motivos estructurales establecen importantes interacciones funcionales con factores que pertenecen a la maquinaria de transcripción y *splicing*, como el caso de TCERG1, que se une mediante sus dominios WW a factores de *splicing* como SF1 o U2AF⁶⁵ y a través de los dominios FF con factores de transcripción como Tat-SF1 o el CTD de la RNAPII (Carty et al., 2000; Goldstrohm et al., 2001; Lin, Lu, and Tarn, 2004; Sanchez-Alvarez et al., 2006; Smith, Kulkarni, and Pawson, 2004). En este trabajo, se ha descrito la asociación de PRPF40B con los factores de *splicing* SF1 y U2AF⁶⁵ a través de sus dominios WW. Estos factores de *splicing* están implicados en el reconocimiento del sitio 3'SS durante la primera etapa de formación del *spliceosoma*, situando por tanto, a esta fase del procesamiento del ARNm, de manera similar a su homólogo en levaduras Prp40. Los datos funcionales sobre el gen Fas corroboran este dato, ya que PRPF40B regula el *splicing* alternativo del exón 6 de este gen.

En segundo lugar, el hecho de que PRPF40B se exprese de forma diferencial en células nerviosas, como la línea de neuroblastoma SH-SY5Y, apoya indirectamente los datos de PRPF40B como factor de *splicing* puesto que este tejido se caracteriza por contener un alto número y expresión de factores de procesamiento del ARNm. Además atendiendo al comportamiento de la mayoría de factores de *splicing*, PRPF40B se encuentra enriquecido en los speckles nucleares, orgánulos nucleares donde más de la mitad de su composición consta de factores de *splicing* y el resto lo conforman otras proteínas relacionadas con el metabolismo del ARNm.

En tercer lugar, el análisis genómico del silenciamiento de PRPF40B en células de neuroblastoma SH-SY5Y revela que PRPF40B, al menos en esta línea celular, se comporta como un activador transcripcional. Ensayos de actividad luciferasa confirman su capacidad para activar la transcripción de genes reporteros. Además hemos mostrado que mutaciones de PRPF40B asociadas a MDS afectan a la actividad transcripcional pero no al *splicing* alternativo de las construcciones estudiadas, y varios genes implicados en la ruta de AML aparecen alterados a nivel de transcripción en el HTA. Por otro lado, datos no mostrados del laboratorio revelan la interacción de PRPF40B

con CDK9. Por tanto, todos estos datos sostienen el papel de PRPF40B en la transcripción.

Uno de los datos más relevantes obtenidos del HTA en relación con la posible función acopladora de PRPF40B, es que aproximadamente la mitad de los genes alterados a nivel de exón (42 de 102), también están afectados a nivel de transcripción. Este dato es bastante alto, en relación con otros factores como Tat-SF1, donde se ha descrito su relación con las maquinarias de transcripción y *splicing*, pero sólo 3 genes de los 86 que cambian a nivel de exón y 447, que están alterados a nivel de transcripción, son comunes a ambos procesos (Miller et al., 2011). Por tanto y aunque son previsible efectos independientes de PRPF40B sobre la transcripción y el *splicing*, al menos en un pequeño conjunto de genes PRPF40B parece estar mediando conexiones ente ambos procesos.

Destacar que los resultados en el HTA del silenciamiento de PRPF40B hay genes afectados que han sido directamente implicados en el acoplamiento entre la transcripción y el *splicing*. Es el caso de SMARCA2, subunidad catalítica del complejo remodelador de la cromatina SWI/SNF. Este factor, también conocido como Brm, es capaz de unirse a factores de *splicing* como Sam68 en el gen *CD44*, ralentizando a la tasa de elongación de la RNAPII y favoreciendo la inclusión de exones alternativos (Batsche, Yaniv and Muchardt, 2006). Además, este es un buen ejemplo de proteínas que aunque no están necesariamente regulando los procesos de la transcripción o el *splicing*, tienen capacidad para conectar elementos de ambas maquinarias, sirviendo de andamiaje y ayudando a la regulación recíproca.

Por tanto, aunque nuestros datos no proporcionan todavía un mecanismo de regulación y comunicación entre la transcripción y el procesamiento del ARNm, estos resultados presentan las primeras evidencias que justifican la figura de PRPF40B como acoplador entre los procesos de transcripción y *splicing*, al menos para un específico número de genes.

Bibliografía

- Abovich, N., and Rosbash, M. (1997). Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* **89**(3), 403-12.
- Aebi, M., Hornig, H., Padgett, R. A., Reiser, J., and Weissmann, C. (1986). Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. *Cell* **47**(4), 555-65.
- Alexander, P. A., He, Y., Chen, Y., Orban, J., Bryan, P. N., (2009). A minimal sequence code for switching protein structure and function. *Proc Natl Acad Sci U S A* **106**(50), 21149-54.
- Alexander, R. D., Innocente, S. A., Barrass, J. D., and Beggs, J. D. (2010). Splicing-dependent RNA polymerase pausing in yeast. *Mol Cell* **40**(4), 582-93.
- Alvarez, M., Estivill, X., and de la Luna, S. (2003). DYRK1A accumulates in splicing speckles through a novel targeting signal and induces speckle disassembly. *J Cell Sci* **116**(Pt 15), 3099-107.
- Allen, M., Friedler, A., Schon, O., and Bycroft, M. (2002). The structure of an FF domain from human HYPA/FBP11. *J Mol Biol* **323**(3), 411-6.
- Arango, M., Holbert, S., Zala, D., Brouillet, E., Pearson, J., Regulier, E., Thakur, A. K., Aebischer, P., Wetzel, R., Deglon, N., and Neri, C. (2006). CA150 expression delays striatal cell death in overexpression and knock-in conditions for mutant huntingtin neurotoxicity. *J Neurosci* **26**(17), 4649-59.
- Ast, G. (2004). How did alternative splicing evolve? *Nat Rev Genet* **5**(10), 773-82.
- Barboric, M., Lenasi, T., Chen, H., Johansen, E. B., Guo, S., and Peterlin, B. M. (2009). 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc Natl Acad Sci U S A* **106**(19), 7798-803.
- Barboric, M., Nissen, R. M., Kanazawa, S., Jabrane-Ferrat, N., and Peterlin, B. M. (2001). NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol Cell* **8**(2), 327-37.
- Bartkowiak, B., Liu, P., Phatnani, H. P., Fuda, N. J., Cooper, J. J., Price, D. H., Adelman, K., Lis, J. T., and Greenleaf, A. L. (2010). CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev* **24**(20), 2303-16.
- Bartolomei, M. S., Halden, N. F., Cullen, C. R., and Corden, J. L. (1988). Genetic analysis of the repetitive carboxyl-terminal domain of the largest subunit of mouse RNA polymerase II. *Mol Cell Biol* **8**(1), 330-9.
- Batsché, E., Yaniv, M., Muchardt, C. (2006). The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13**(1), 22-9.

- Bedford, M. T., Chan, D. C., and Leder, P. (1997). FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *EMBO J* **16**(9), 2376-83.
- Bedford, M. T., and Leder, P. (1999). The FF domain: a novel motif that often accompanies WW domains. *Trends Biochem Sci* **24**(7), 264-5.
- Bedford, M. T., Reed, R., and Leder, P. (1998). WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. *Proc Natl Acad Sci U S A* **95**(18), 10602-7.
- Berget, S. M. (1995). Exon recognition in vertebrate splicing. *J Biol Chem* **270**(6), 2411-4.
- Berk, A. J. (1999). Activation of RNA polymerase II transcription. *Curr Opin Cell Biol* **11**(3), 330-5.
- Bernard, D., Prasanth, K. V., Tripathi, V., Colasse, S., Nakamura, T., Xuan, Z., Zhang, M. Q., Sedel, F., Jourden, L., Couplier, F., Triller, A., Spector, D. L., and Bessis, A. (2010). A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J* **29**(18), 3082-93.
- Bernardi, R., and Pandolfi, P. P. (2007). Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nat Rev Mol Cell Biol* **8**(12), 1006-16.
- Beyer, A. L., and Osheim, Y. N. (1988). Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev* **2**(6), 754-65.
- Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72**, 291-336.
- Blencowe, B. J. (2006). Alternative splicing: new insights from global analyses. *Cell* **126**(1), 37-47.
- Boise, L. H., Gonzalez-Garcia, M., Postema, C. E., Ding, L., Lindsten, T., Turka, L. A., Mao, X., Nunez, G., and Thompson, C. B. (1993). bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell* **74**(4), 597-608.
- Bonnal, S., Martinez, C., Forch, P., Bachi, A., Wilm, M., and Valcarcel, J. (2008). RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. *Mol Cell* **32**(1), 81-95.
- Bonnal, S., Vigevani, L., and Valcarcel, J. (2012). The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discov* **11**(11), 847-59.
- Boronenkov, I. V., Loijens, J. C., Umeda, M., and Anderson, R. A. (1998). Phosphoinositide signaling pathways in nuclei are associated with nuclear

- speckles containing pre-mRNA processing factors. *Mol Biol Cell* **9**(12), 3547-60.
- Boutz, P. L., Stoilov, P., Li, Q., Lin, C. H., Chawla, G., Ostrow, K., Shiue, L., Ares, M., Jr., and Black, D. L. (2007). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev* **21**(13), 1636-52.
- Brede, G., Solheim, J., and Prydz, H. (2002). PSKH1, a novel splice factor compartment-associated serine kinase. *Nucleic Acids Res* **30**(23), 5301-9.
- Bres, V., Gomes, N., Pickle, L., and Jones, K. A. (2005). A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. *Genes Dev* **19**(10), 1211-26.
- Bres, V., Yoh, S. M., and Jones, K. A. (2008). The multi-tasking P-TEFb complex. *Curr Opin Cell Biol* **20**(3), 334-40.
- Brint, E., O'Callaghan, G., and Houston, A. (2013). Life in the Fas lane: differential outcomes of Fas signaling. *Cell Mol Life Sci* **70**(21), 4085-99.
- Brown, J. M., Green, J., das Neves, R. P., Wallace, H. A., Smith, A. J., Hughes, J., Gray, N., Taylor, S., Wood, W. G., Higgs, D. R., Iborra, F. J., and Buckle, V. J. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J Cell Biol* **182**(6), 1083-97.
- Buratowski, S. (2003). The CTD code. *Nat Struct Biol* **10**(9), 679-80.
- Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. *Mol Cell* **36**(4), 541-6.
- Buschdorf, J. P., and Stratling, W. H. (2004). A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. *J Mol Med (Berl)* **82**(2), 135-43.
- Caceres, J. F., and Kornblihtt, A. R. (2002). Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet* **18**(4), 186-93.
- Caceres, J. F., Misteli, T., Sreaton, G. R., Spector, D. L., and Krainer, A. R. (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J Cell Biol* **138**(2), 225-38.
- Carrillo Oesterreich, F., Preibisch, S., and Neugebauer, K. M. (2010). Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol Cell* **40**(4), 571-81.
- Carty, S. M., Goldstrohm, A. C., Sune, C., Garcia-Blanco, M. A., and Greenleaf, A. L. (2000). Protein-interaction modules that organize nuclear function: FF domains of CA150 bind the phosphoCTD of RNA polymerase II. *Proc Natl Acad Sci U S A* **97**(16), 9015-20.

- Cascino, I., Fiucci, G., Papoff, G., and Ruberti, G. (1995). Three functional soluble forms of the human apoptosis-inducing Fas molecule are produced by alternative splicing. *J Immunol* **154**(6), 2706-13.
- Cmarko, D., Verschure, P. J., Martin, T. E., Dahmus, M. E., Krause, S., Fu, X. D., van Driel, R., and Fakan, S. (1999). Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. *Mol Biol Cell* **10**(1), 211-23.
- Coiras, M., Montes, M., Montanuy, I., Lopez-Huertas, M. R., Mateos, E., Le Sommer, C., Garcia-Blanco, M. A., Hernandez-Munain, C., Alcami, J., and Sune, C. (2013). Transcription elongation regulator 1 (TCERG1) regulates competent RNA polymerase II-mediated elongation of HIV-1 transcription and facilitates efficient viral replication. *Retrovirology* **10**, 124.
- Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J. L., Bell, J. C., and Duncan, P. I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J* **15**(2), 265-75.
- Cooper, T. A., Wan, L., and Dreyfuss, G. (2009). RNA and disease. *Cell* **136**(4), 777-93.
- Corden, J. L., Cadena, D. L., Ahearn, J. M., Jr., and Dahmus, M. E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci U S A* **82**(23), 7934-8.
- Core, L. J., and Lis, J. T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319**(5871), 1791-2.
- Corey, S. J., Minden, M. D., Barber, D. L., Kantarjian, H., Wang, J. C., and Schimmer, A. D. (2007). Myelodysplastic syndromes: the complexity of stem-cell diseases. *Nat Rev Cancer* **7**(2), 118-29.
- Cotter, T. G., Lennon, S. V., Glynn, J. G., and Martin, S. J. (1990). Cell death via apoptosis and its relationship to growth, development and differentiation of both tumour and normal cells. *Anticancer Res* **10**(5A), 1153-9.
- Coutinho-Mansfield, G. C., Xue, Y., Zhang, Y., and Fu, X. D. (2007). PTB/nPTB switch: a post-transcriptional mechanism for programming neuronal differentiation. *Genes Dev* **21**(13), 1573-7.
- Cramer, P., Caceres, J. F., Cazalla, D., Kadener, S., Muro, A. F., Baralle, F. E., and Kornblihtt, A. R. (1999). Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol Cell* **4**(2), 251-8.
- Crawford, J. B., and Patton, J. G. (2006). Activation of alpha-tropomyosin exon 2 is regulated by the SR protein 9G8 and heterogeneous nuclear ribonucleoproteins H and F. *Mol Cell Biol* **26**(23), 8791-802.

- Chan, D. C., Bedford, M. T., and Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J* **15**(5), 1045-54.
- Chapman, R. D., Heidemann, M., Albert, T. K., Mailhammer, R., Flatley, A., Meisterernst, M., Kremmer, E., and Eick, D. (2007). Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* **318**(5857), 1780-2.
- Chen, M., and Manley, J. L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol* **10**(11), 741-54.
- Cheng, D., Cote, J., Shaaban, S., and Bedford, M. T. (2007). The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. *Mol Cell* **25**(1), 71-83.
- Cheng, E. H., Wei, M. C., Weiler, S., Flavell, R. A., Mak, T. W., Lindsten, T., and Korsmeyer, S. J. (2001). BCL-2, BCL-X(L) sequester BH3 domain-only molecules preventing BAX- and BAK-mediated mitochondrial apoptosis. *Mol Cell* **8**(3), 705-11.
- Cheng, J., Zhou, T., Liu, C., Shapiro, J. P., Brauer, M. J., Kiefer, M. C., Barr, P. J., and Mountz, J. D. (1994). Protection from Fas-mediated apoptosis by a soluble form of the Fas molecule. *Science* **263**(5154), 1759-62.
- D'Souza, I., Poorkaj, P., Hong, M., Nochlin, D., Lee, V. M., Bird, T. D., and Schellenberg, G. D. (1999). Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc Natl Acad Sci U S A* **96**(10), 5598-603.
- Danko, C. G., Hah, N., Luo, X., Martins, A. L., Core, L., Lis, J. T., Siepel, A., and Kraus, W. L. (2013). Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell* **50**(2), 212-22.
- Das, R., Yu, J., Zhang, Z., Gygi, M. P., Krainer, A. R., Gygi, S. P., and Reed, R. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell* **26**(6), 867-81.
- De Conti, L., Baralle, M., and Buratti, E. (2013). Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip Rev RNA* **4**(1), 49-60.
- de la Grange, P., Gratadou, L., Delord, M., Dutertre, M., and Auboeuf, D. (2010). Splicing factor and exon profiling across human tissues. *Nucleic Acids Res* **38**(9), 2825-38.
- de la Mata, M., and Kornblihtt, A. R. (2006). RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nat Struct Mol Biol* **13**(11), 973-80.

- Deveraux, Q. L., Roy, N., Stennicke, H. R., Van Arsdale, T., Zhou, Q., Srinivasula, S. M., Alnemri, E. S., Salvesen, G. S., and Reed, J. C. (1998). IAPs block apoptotic events induced by caspase-8 and cytochrome c by direct inhibition of distinct caspases. *EMBO J* **17**(8), 2215-23.
- Dong, B., Horowitz, D. S., Kobayashi, R., and Krainer, A. R. (1993). Purification and cDNA cloning of HeLa cell p54nrb, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and Drosophila NONA/BJ6. *Nucleic Acids Res* **21**(17), 4085-92.
- Dredge, B. K., Stefani, G., Engelhard, C. C., and Darnell, R. B. (2005). Nova autoregulation reveals dual functions in neuronal splicing. *EMBO J* **24**(8), 1608-20.
- Dye, B. T., and Patton, J. G. (2001). An RNA recognition motif (RRM) is required for the localization of PTB-associated splicing factor (PSF) to subnuclear speckles. *Exp Cell Res* **263**(1), 131-44.
- Earnshaw, W. C., Martins, L. M., and Kaufmann, S. H. (1999). Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annu Rev Biochem* **68**, 383-424.
- Egloff, S., Dienstbier, M., and Murphy, S. (2012). Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet* **28**(7), 333-41.
- Egloff, S., and Murphy, S. (2008). Cracking the RNA polymerase II CTD code. *Trends Genet* **24**(6), 280-8.
- Eilbracht, J., and Schmidt-Zachmann, M. S. (2001). Identification of a sequence element directing a protein to nuclear speckles. *Proc Natl Acad Sci U S A* **98**(7), 3849-54.
- Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P. W., Kobayashi, R., Blencowe, B. J., and Ingles, C. J. (2002). Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD. *RNA* **8**(9), 1102-11.
- Emmott, E., and Hiscox, J. A. (2009). Nucleolar targeting: the hub of the matter. *EMBO Rep* **10**(3), 231-8.
- Eperon, L. P., Graham, I. R., Griffiths, A. D., and Eperon, I. C. (1988). Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell* **54**(3), 393-401.
- Faber, P. W., Barnes, G. T., Srinidhi, J., Chen, J., Gusella, J. F., and MacDonald, M. E. (1998). Huntingtin interacts with a family of WW domain proteins. *Hum Mol Genet* **7**(9), 1463-74.
- Fakan, S. (1994). Perichromatin fibrils are in situ forms of nascent transcripts. *Trends Cell Biol* **4**(3), 86-90.

- Fischer, U., Janicke, R. U., and Schulze-Osthoff, K. (2003). Many cuts to ruin: a comprehensive update of caspase substrates. *Cell Death Differ* **10**(1), 76-100.
- Fong, N., Bird, G., Vigneron, M., and Bentley, D. L. (2003). A 10 residue motif at the C-terminus of the RNA pol II CTD is required for transcription, splicing and 3' end processing. *EMBO J* **22**(16), 4274-82.
- Fong, Y. W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**(6866), 929-33.
- Fox, A. H., and Lamond, A. I. (2010). Paraspeckles. *Cold Spring Harb Perspect Biol* **2**(7), a000687.
- Fricker, N., Beaudouin, J., Richter, P., Eils, R., Krammer, P. H., and Lavrik, I. N. (2010). Model-based dissection of CD95 signaling dynamics reveals both a pro- and antiapoptotic role of c-FLIPL. *J Cell Biol* **190**(3), 377-89.
- Gasch, A., Wiesner, S., Martin-Malpartida, P., Ramirez-Espain, X., Ruiz, L., and Macias, M. J. (2006). The structure of Prp40 FF1 domain and its interaction with the crn-TPR1 motif of Clf1 gives a new insight into the binding mode of FF domains. *J Biol Chem* **281**(1), 356-64.
- Gatchel, J. R., and Zoghbi, H. Y. (2005). Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet* **6**(10), 743-55.
- Goldstrohm, A. C., Albrecht, T. R., Sune, C., Bedford, M. T., and Garcia-Blanco, M. A. (2001). The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. *Mol Cell Biol* **21**(22), 7617-28.
- Gornemann, J., Barrandon, C., Hujer, K., Rutz, B., Rigaut, G., Kotovic, K. M., Faux, C., Neugebauer, K. M., and Seraphin, B. (2011). Cotranscriptional spliceosome assembly and splicing are independent of the Prp40p WW domain. *RNA* **17**(12), 2119-29.
- Graubert, T. A., Shen, D., Ding, L., Okeyo-Owuor, T., Lunn, C. L., Shao, J., Krysiak, K., Harris, C. C., Koboldt, D. C., Larson, D. E., McLellan, M. D., Dooling, D. J., Abbott, R. M., Fulton, R. S., Schmidt, H., Kalicki-Veizer, J., O'Laughlin, M., Grillo, M., Baty, J., Heath, S., Frater, J. L., Nasim, T., Link, D. C., Tomasson, M. H., Westervelt, P., DiPersio, J. F., Mardis, E. R., Ley, T. J., Wilson, R. K., and Walter, M. J. (2012). Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet* **44**(1), 53-7.
- Graveley, B. R. (2000). Sorting out the complexity of SR protein functions. *RNA* **6**(9), 1197-211.
- Graveley, B. R., Hertel, K. J., and Maniatis, T. (2001). The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA* **7**(6), 806-18.

- Grosso, A. R., Gomes, A. Q., Barbosa-Morais, N. L., Caldeira, S., Thorne, N. P., Grech, G., von Lindern, M., and Carmo-Fonseca, M. (2008). Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res* **36**(15), 4823-32.
- Grosso, A. R., Martins, S., and Carmo-Fonseca, M. (2008). The emerging role of splicing factors in cancer. *EMBO Rep* **9**(11), 1087-93.
- Gubitz, A. K., Feng, W., and Dreyfuss, G. (2004). The SMN complex. *Exp Cell Res* **296**(1), 51-6.
- Herrmann, C. H., and Mancini, M. A. (2001). The Cdk9 and cyclin T subunits of TAK/P-TEFb localize to splicing factor-rich nuclear speckle regions. *J Cell Sci* **114**(Pt 8), 1491-503.
- Hintermair, C., Heidemann, M., Koch, F., Descostes, N., Gut, M., Gut, I., Fenouil, R., Ferrier, P., Flatley, A., Kremmer, E., Chapman, R. D., Andrau, J. C., and Eick, D. (2012). Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *EMBO J* **31**(12), 2784-97.
- Hirose, Y., Tacke, R., and Manley, J. L. (1999). Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev* **13**(10), 1234-9.
- Holbert, S., Denghien, I., Kiechle, T., Rosenblatt, A., Wellington, C., Hayden, M. R., Margolis, R. L., Ross, C. A., Dausset, J., Ferrante, R. J., and Neri, C. (2001). The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis. *Proc Natl Acad Sci U S A* **98**(4), 1811-6.
- Hoskins, A. A., Friedman, L. J., Gallagher, S. S., Crawford, D. J., Anderson, E. G., Wombacher, R., Ramirez, N., Cornish, V. W., Gelles, J., and Moore, M. J. (2011). Ordered and dynamic assembly of single spliceosomes. *Science* **331**(6022), 1289-95.
- House, A. E., and Lynch, K. W. (2006). An exonic splicing silencer represses spliceosome assembly after ATP-dependent exon recognition. *Nat Struct Mol Biol* **13**(10), 937-44.
- Hsin, J. P., Sheth, A., and Manley, J. L. (2011). RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science* **334**(6056), 683-6.
- Huang, S., Deerinck, T. J., Ellisman, M. H., and Spector, D. L. (1994). In vivo analysis of the stability and transport of nuclear poly(A)⁺ RNA. *J Cell Biol* **126**(4), 877-99.
- Huang, X., Beullens, M., Zhang, J., Zhou, Y., Nicolaescu, E., Lesage, B., Hu, Q., Wu, J., Bollen, M., and Shi, Y. (2009). Structure and function of the two tandem WW domains of the pre-mRNA splicing factor FBP21 (formin-binding protein 21). *J Biol Chem* **284**(37), 25375-87.

- Huang, Y., Li, W., Yao, X., Lin, Q. J., Yin, J. W., Liang, Y., Heiner, M., Tian, B., Hui, J., and Wang, G. (2012). Mediator complex regulates alternative mRNA processing via the MED23 subunit. *Mol Cell* **45**(4), 459-69.
- Huranova, M., Ivani, I., Benda, A., Poser, I., Brody, Y., Hof, M., Shav-Tal, Y., Neugebauer, K. M., and Stanek, D. (2010). The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *J Cell Biol* **191**(1), 75-86.
- Hutchison, S., LeBel, C., Blanchette, M., and Chabot, B. (2002). Distinct sets of adjacent heterogeneous nuclear ribonucleoprotein (hnRNP) A1/A2 binding sites control 5' splice site selection in the hnRNP A1 mRNA precursor. *J Biol Chem* **277**(33), 29745-52.
- Ingham, R. J., Colwill, K., Howard, C., Dettwiler, S., Lim, C. S., Yu, J., Hersi, K., Raaijmakers, J., Gish, G., Mbamalu, G., Taylor, L., Yeung, B., Vassilovski, G., Amin, M., Chen, F., Matskova, L., Winberg, G., Ernberg, I., Linding, R., O'Donnell, P., Starostine, A., Keller, W., Metalnikov, P., Stark, C., and Pawson, T. (2005). WW domains provide a platform for the assembly of multiprotein networks. *Mol Cell Biol* **25**(16), 7092-106.
- Inoue, A., Yamamoto, N., Kimura, M., Nishio, K., Yamane, H., and Nakajima, K. (2014). RBM10 regulates alternative splicing. *FEBS Lett* **588**(6), 942-7.
- Izquierdo, J. M. (2008). Hu antigen R (HuR) functions as an alternative pre-mRNA splicing regulator of Fas apoptosis-promoting receptor on exon definition. *J Biol Chem* **283**(27), 19077-84.
- Izquierdo, J. M., Majos, N., Bonnal, S., Martinez, C., Castelo, R., Guigo, R., Bilbao, D., and Valcarcel, J. (2005). Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol Cell* **19**(4), 475-84.
- Izquierdo, J. M., and Valcarcel, J. (2007). Fas-activated serine/threonine kinase (FAST K) synergizes with TIA-1/TIAR proteins to regulate Fas alternative splicing. *J Biol Chem* **282**(3), 1539-43.
- Jagatheesan, G., Thanumalayan, S., Muralikrishna, B., Rangaraj, N., Karande, A. A., and Parnaik, V. K. (1999). Colocalization of intranuclear lamin foci with RNA splicing factors. *J Cell Sci* **112** (Pt 24), 4651-61.
- Jagiello, I., Van Eynde, A., Vulsteke, V., Beullens, M., Boudrez, A., Keppens, S., Stalmans, W., and Bollen, M. (2000). Nuclear and subnuclear targeting sequences of the protein phosphatase-1 regulator NIPP1. *J Cell Sci* **113** Pt 21, 3761-8.
- Jiang, Y. J., Che, M. X., Yuan, J. Q., Xie, Y. Y., Yan, X. Z., and Hu, H. Y. (2011). Interaction with polyglutamine-expanded huntingtin alters cellular distribution and RNA processing of huntingtin yeast two-hybrid protein A (HYPA). *J Biol Chem* **286**(28), 25236-45.

- Johnson, C., Primorac, D., McKinstry, M., McNeil, J., Rowe, D., and Lawrence, J. B. (2000). Tracking COL1A1 RNA in osteogenesis imperfecta. splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. *J Cell Biol* **150**(3), 417-32.
- Kadener, S., Cramer, P., Nogues, G., Cazalla, D., de la Mata, M., Fededa, J. P., Verbajh, S. E., Srebrow, A., and Kornblihtt, A. R. (2001). Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing. *EMBO J* **20**(20), 5759-68.
- Kaida, D., Berg, M. G., Younis, I., Kasim, M., Singh, L. N., Wan, L., and Dreyfuss, G. (2010). U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**(7324), 664-8.
- Kameoka, S., Duque, P., and Konarska, M. M. (2004). p54(nrb) associates with the 5' splice site within large transcription/splicing complexes. *EMBO J* **23**(8), 1782-91.
- Kao, H. Y., and Siliciano, P. G. (1996). Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. *Mol Cell Biol* **16**(3), 960-7.
- Karlic, H., Herrmann, H., Varga, F., Thaler, R., Reitermaier, R., Spitzer, S., Ghanim, V., Blatt, K., Sperr, W. R., Valent, P., and Pfeilstocker, M. (2014). The role of epigenetics in the regulation of apoptosis in myelodysplastic syndromes and acute myeloid leukemia. *Crit Rev Oncol Hematol* **90**(1), 1-16.
- Kelly, W. G., Dahmus, M. E., and Hart, G. W. (1993). RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *J Biol Chem* **268**(14), 10416-24.
- Kerr, J. F., Wyllie, A. H., and Currie, A. R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* **26**(4), 239-57.
- Kim, H., Erickson, B., Luo, W., Seward, D., Graber, J. H., Pollock, D. D., Megee, P. C., and Bentley, D. L. (2010). Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nat Struct Mol Biol* **17**(10), 1279-86.
- Kim, J. B., Yamaguchi, Y., Wada, T., Handa, H., and Sharp, P. A. (1999). Tat-SF1 protein associates with RAP30 and human SPT5 proteins. *Mol Cell Biol* **19**(9), 5960-8.
- Klippel, S., Wiczorek, M., Schumann, M., Krause, E., Marg, B., Seidel, T., Meyer, T., Knapp, E. W., and Freund, C. (2011). Multivalent binding of formin-binding protein 21 (FBP21)-tandem-WW domains fosters protein recognition in the pre-spliceosome. *J Biol Chem* **286**(44), 38478-87.

- Ko, T. K., Kelly, E., and Pines, J. (2001). CrkRS: a novel conserved Cdc2-related protein kinase that colocalises with SC35 speckles. *J Cell Sci* **114**(Pt 14), 2591-603.
- Krawczak, M., Reiss, J., and Cooper, D. N. (1992). The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet* **90**(1-2), 41-54.
- Kuwana, T., Bouchier-Hayes, L., Chipuk, J. E., Bonzon, C., Sullivan, B. A., Green, D. R., and Newmeyer, D. D. (2005). BH3 domains of BH3-only proteins differentially regulate Bax-mediated mitochondrial membrane permeabilization both directly and indirectly. *Mol Cell* **17**(4), 525-35.
- Kuyumcu-Martinez, N. M., Wang, G. S., and Cooper, T. A. (2007). Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. *Mol Cell* **28**(1), 68-78.
- Lam, Y.W., Lamond, A. I., Mann, M., Andersen, J. S. (2007). Analysis of nucleolar protein dynamics reveals the nuclear degradation of ribosomal proteins. *Curr Biol* **17**(9), 749-60.
- Lamond, A. I., and Spector, D. L. (2003). Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol* **4**(8), 605-12.
- Lawrence, J. B., and Clemson, C. M. (2008). Gene associations: true romance or chance meeting in a nuclear neighborhood? *J Cell Biol* **182**(6), 1035-8.
- Lee, J. S., Smith, E., and Shilatifard, A. (2010). The language of histone crosstalk. *Cell* **142**(5), 682-5.
- Letai, A., Bassik, M. C., Walensky, L. D., Sorcinelli, M. D., Weiler, S., and Korsmeyer, S. J. (2002). Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. *Cancer Cell* **2**(3), 183-92.
- Levine, M. (2011). Paused RNA polymerase II as a developmental checkpoint. *Cell* **145**(4), 502-11.
- Li, B., Carey, M., and Workman, J. L. (2007). The role of chromatin during transcription. *Cell* **128**(4), 707-19.
- Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S. M., Ahmad, M., Alnemri, E. S., and Wang, X. (1997). Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* **91**(4), 479-89.
- Lin, C. H., and Patton, J. G. (1995). Regulation of alternative 3' splice site selection by constitutive splicing factors. *RNA* **1**(3), 234-45.
- Lin, K. T., Lu, R. M., and Tarn, W. Y. (2004). The WW domain-containing proteins interact with the early spliceosome and participate in pre-mRNA splicing in vivo. *Mol Cell Biol* **24**(20), 9176-85.

- Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X. D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol* **15**(8), 819-26.
- Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., and Ranum, L. P. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**(5531), 864-7.
- Listerman, I., Sapra, A. K., and Neugebauer, K. M. (2006). Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat Struct Mol Biol* **13**(9), 815-22.
- Liu, C., Cheng, J., and Mountz, J. D. (1995). Differential expression of human Fas mRNA species upon peripheral blood mononuclear cell activation. *Biochem J* **310** (Pt 3), 957-63.
- Lorenzo, H. K., Susin, S. A., Penninger, J., and Kroemer, G. (1999). Apoptosis inducing factor (AIF): a phylogenetically old, caspase-independent effector of cell death. *Cell Death Differ* **6**(6), 516-24.
- Lu, H., Zawel, L., Fisher, L., Egly, J. M., and Reinberg, D. (1992). Human general transcription factor IIIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358**(6388), 641-5.
- Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999). Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* **283**(5406), 1325-8.
- Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* **327**(5968), 996-1000.
- Macias, M. J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996). Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* **382**(6592), 646-9.
- Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., and et al. (1992). Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255**(5049), 1253-5.
- Makarov, E. M., Owen, N., Bottrill, A., and Makarova, O. V. (2012). Functional mammalian spliceosomal complex E contains SMN complex proteins in addition to U1 and U2 snRNPs. *Nucleic Acids Res* **40**(6), 2639-52.
- Malik, S., and Roeder, R. G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11**(11), 761-72.
- Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature* **416**(6880), 499-506.

- Marshall, N. F., and Price, D. H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem* **270**(21), 12335-8.
- McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D. L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* **11**(24), 3306-18.
- Meier, U. T., and Blobel, G. (1990). A nuclear localization signal binding protein in the nucleolus. *J Cell Biol* **111**(6 Pt 1), 2235-45.
- Miller, H. B., Robinson, T. J., Gordan, R., Hartemink, A. J., and Garcia-Blanco, M. A. (2011). Identification of Tat-SF1 cellular targets by exon array analysis reveals dual roles in transcription and splicing. *RNA* **17**(4), 665-74.
- Misteli, T., and Spector, D. L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol Cell* **3**(6), 697-705.
- Mizutani, K., Suetsugu, S., and Takenawa, T. (2004). FBP11 regulates nuclear localization of N-WASP and inhibits N-WASP-dependent microspike formation. *Biochem Biophys Res Commun* **313**(3), 468-74.
- Moen, P. T., Jr., Johnson, C. V., Byron, M., Shopland, L. S., de la Serna, I. L., Imbalzano, A. N., and Lawrence, J. B. (2004). Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. *Mol Biol Cell* **15**(1), 197-206.
- Monani, U. R., Lorson, C. L., Parsons, D. W., Prior, T. W., Androphy, E. J., Burghes, A. H., and McPherson, J. D. (1999). A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum Mol Genet* **8**(7), 1177-83.
- Montanuy, I., Torremocha, R., Hernandez-Munain, C., and Sune, C. (2008). Promoter influences transcription elongation: TATA-box element mediates the assembly of processive transcription complexes responsive to cyclin-dependent kinase 9. *J Biol Chem* **283**(12), 7368-78.
- Montes, M., Cloutier, A., Sanchez-Hernandez, N., Michelle, L., Lemieux, B., Blanchette, M., Hernandez-Munain, C., Chabot, B., and Sune, C. (2011). TCERG1 regulates alternative splicing of Bcl-x gene by modulating the rate of RNAPII transcription. *Mol Cell Biol*.
- Morris, D. P., and Greenleaf, A. L. (2000). The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **275**(51), 39935-43.
- Morris, G. E. (2008). The Cajal body. *Biochim Biophys Acta* **1783**(11), 2108-15.

- Mortillaro, M. J., Blencowe, B. J., Wei, X., Nakayasu, H., Du, L., Warren, S. L., Sharp, P. A., and Berezney, R. (1996). A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc Natl Acad Sci U S A* **93**(16), 8253-7.
- Myers, L. C., Gustafsson, C. M., Bushnell, D. A., Lui, M., Erdjument-Bromage, H., Tempst, P., and Kornberg, R. D. (1998). The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes Dev* **12**(1), 45-54.
- Nakayasu, H., and Ueda, K. (1984). Small nuclear RNA-protein complex anchors on the actin filaments in bovine lymphocyte nuclear matrix. *Cell Struct Funct* **9**(4), 317-25.
- Nasim, F. U., Hutchison, S., Cordeau, M., and Chabot, B. (2002). High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *RNA* **8**(8), 1078-89.
- Nechaev, S., Fargo, D. C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* **327**(5963), 335-8.
- Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., Bruce, J., Schuck, T., Grossman, M., Clark, C. M., McCluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., and Lee, V. M. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**(5796), 130-3.
- Nogues, G., Kadener, S., Cramer, P., Bentley, D., and Kornblihtt, A. R. (2002). Transcriptional activators differ in their abilities to control alternative splicing. *J Biol Chem* **277**(45), 43110-4.
- Nogues, G., Munoz, M. J., and Kornblihtt, A. R. (2003). Influence of polymerase II processivity on alternative splicing depends on splice site strength. *J Biol Chem* **278**(52), 52166-71.
- Norberg, E., Orrenius, S., and Zhivotovsky, B. (2010). Mitochondrial regulation of cell death: processing of apoptosis-inducing factor (AIF). *Biochem Biophys Res Commun* **396**(1), 95-100.
- Norris, A. D., and Calarco, J. A. (2012). Emerging Roles of Alternative Pre-mRNA Splicing Regulation in Neuronal Development and Function. *Front Neurosci* **6**, 122.
- O'Keefe, R. T., Mayeda, A., Sadowski, C. L., Krainer, A. R., and Spector, D. L. (1994). Disruption of pre-mRNA splicing in vivo results in reorganization of splicing factors. *J Cell Biol* **124**(3), 249-60.

- Okeyo-Owuor, T., White, B. S., Chatrikhi, R., Mohan, D. R., Kim, S., Griffith, M., Ding, L., Ketkar-Kulkarni, S., Hundal, J., Laird, K. M., Kielkopf, C. L., Ley, T. J., Walter, M. J., and Graubert, T. A. (2014). U2AF1 mutations alter sequence specificity of pre-mRNA binding and splicing. *Leukemia*.
- Padgett R. A. (2012) New connections between splicing and human disease. *Trends Genet* **28**(4). 147-54
- Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40**(12), 1413-5.
- Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G. W., Ares, M., Jr., and Fu, X. D. (2013). Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol Cell* **50**(2), 223-35.
- Panning, B., and Taatjes, D. J. (2008). Transcriptional regulation: it takes a village. *Mol Cell* **31**(5), 622-9.
- Papaemmanuil, E., Cazzola, M., Boulton, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J. S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., Godfrey, A. L., Rapado, I., Cvejic, A., Rance, R., McGee, C., Ellis, P., Mudie, L. J., Stephens, P. J., McLaren, S., Massie, C. E., Tarpey, P. S., Varela, I., Nik-Zainal, S., Davies, H. R., Shlien, A., Jones, D., Raine, K., Hinton, J., Butler, A. P., Teague, J. W., Baxter, E. J., Score, J., Galli, A., Della Porta, M. G., Travaglino, E., Groves, M., Tauro, S., Munshi, N. C., Anderson, K. C., El-Naggar, A., Fischer, A., Mustonen, V., Warren, A. J., Cross, N. C., Green, A. R., Futreal, P. A., Stratton, M. R., and Campbell, P. J. (2011). Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med* **365**(15), 1384-95.
- Papoff, G., Cascino, I., Eramo, A., Starace, G., Lynch, D. H., and Ruberti, G. (1996). An N-terminal domain shared by Fas/Apo-1 (CD95) soluble variants prevents cell death in vitro. *J Immunol* **156**(12), 4622-30.
- Paronetto, M. P., Achsel, T., Massiello, A., Chalfant, C. E., and Sette, C. (2007). The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *J Cell Biol* **176**(7), 929-39.
- Paronetto, M. P., Bernardis, I., Volpe, E., Bechara, E., Sebestyen, E., Eyraes, E., and Valcarcel, J. (2014). Regulation of FAS exon definition and apoptosis by the Ewing sarcoma protein. *Cell Rep* **7**(4), 1211-26.
- Passani, L. A., Bedford, M. T., Faber, P. W., McGinnis, K. M., Sharp, A. H., Gusella, J. F., Vonsattel, J. P., and MacDonald, M. E. (2000). Huntington's WW domain partners in Huntington's disease post-mortem brain fulfill genetic criteria for direct involvement in Huntington's disease pathogenesis. *Hum Mol Genet* **9**(14), 2175-82.

- Patturajan, M., Schulte, R. J., Sefton, B. M., Berezney, R., Vincent, M., Bensaude, O., Warren, S. L., and Corden, J. L. (1998). Growth-related changes in phosphorylation of yeast RNA polymerase II. *J Biol Chem* **273**(8), 4689-94.
- Paz, S., Krainer, A. R., and Caputi, M. (2014). HIV-1 transcription is regulated by splicing factor SRSF1. *Nucleic Acids Res.*
- Pearson, J. L., Robinson, T. J., Munoz, M. J., Kornblihtt, A. R., and Garcia-Blanco, M. A. (2008). Identification of the cellular targets of the transcription factor TCERG1 reveals a prevalent role in mRNA processing. *J Biol Chem* **283**(12), 7949-61.
- Pedrotti, S., Bielli, P., Paronetto, M. P., Ciccocanti, F., Fimia, G. M., Stamm, S., Manley, J. L., and Sette, C. (2010). The splicing regulator Sam68 binds to a novel exonic splicing silencer and functions in SMN2 alternative splicing in spinal muscular atrophy. *EMBO J* **29**(7), 1235-47.
- Peng, R., Hawkins, I., Link, A. J., and Patton, J. G. (2006). The splicing factor PSF is part of a large complex that assembles in the absence of pre-mRNA and contains all five snRNPs. *RNA Biol* **3**(2), 69-76.
- Perales, R., and Bentley, D. (2009). "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell* **36**(2), 178-91.
- Peter, M. E., and Krammer, P. H. (2003). The CD95(APO-1/Fas) DISC and beyond. *Cell Death Differ* **10**(1), 26-35.
- Phair, R. D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. *Nature* **404**(6778), 604-9.
- Phatnani, H. P., and Greenleaf, A. L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* **20**(21), 2922-36.
- Piacentini, L., Fanti, L., Negri, R., Del Vescovo, V., Fatica, A., Altieri, F., and Pimpinelli, S. (2009). Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in *Drosophila*. *PLoS Genet* **5**(10), e1000670.
- Rain, J. C., Rafi, Z., Rhani, Z., Legrain, P., and Kramer, A. (1998). Conservation of functional domains involved in RNA binding and protein-protein interactions in human and *Saccharomyces cerevisiae* pre-mRNA splicing factor SF1. *RNA* **4**(5), 551-65.
- Rino, J., Desterro, J. M., Pacheco, T. R., Gadella, T. W., Jr., and Carmo-Fonseca, M. (2008). Splicing factors SF1 and U2AF associate in extraspliceosomal complexes. *Mol Cell Biol* **28**(9), 3045-57.
- Rodriguez, J., and Lazebnik, Y. (1999). Caspase-9 and APAF-1 form an active holoenzyme. *Genes Dev* **13**(24), 3179-84.

- Roeder, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21**(9), 327-35.
- Roeder, R. G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett* **579**(4), 909-15.
- Roesler, J., Izquierdo, J. M., Ryser, M., Rosen-Wolff, A., Gahr, M., Valcarcel, J., Lenardo, M. J., and Zheng, L. (2005). Haploinsufficiency, rather than the effect of an excessive production of soluble CD95 (CD95 Δ TM), is the basis for ALPS Ia in a family with duplicated 3' splice site AG in CD95 intron 5 on one allele. *Blood* **106**(5), 1652-9.
- Rosonina, E., and Blencowe, B. J. (2004). Analysis of the requirement for RNA polymerase II CTD heptapeptide repeats in pre-mRNA splicing and 3'-end cleavage. *RNA* **10**(4), 581-9.
- Rosonina, E., Ip, J. Y., Calarco, J. A., Bakowski, M. A., Emili, A., McCracken, S., Tucker, P., Ingles, C. J., and Blencowe, B. J. (2005). Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. *Mol Cell Biol* **25**(15), 6734-46.
- Ruchaud, S., Korfali, N., Villa, P., Kottke, T. J., Dingwall, C., Kaufmann, S. H., and Earnshaw, W. C. (2002). Caspase-6 gene disruption reveals a requirement for lamin A cleavage in apoptotic chromatin condensation. *EMBO J* **21**(8), 1967-77.
- Salichs, E., Ledda, A., Mularoni, L., Alba, M. M., and de la Luna, S. (2009). Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment. *PLoS Genet* **5**(3), e1000397.
- Salton, M., Voss, T. C., and Misteli, T. (2014). Identification by high-throughput imaging of the histone methyltransferase EHMT2 as an epigenetic regulator of VEGFA alternative splicing. *Nucleic Acids Res.*
- Salvesen, G. S., and Duckett, C. S. (2002). IAP proteins: blocking the road to death's door. *Nat Rev Mol Cell Biol* **3**(6), 401-10.
- Sanchez-Alvarez, M., Goldstrohm, A. C., Garcia-Blanco, M. A., and Sune, C. (2006). Human transcription elongation factor CA150 localizes to splicing factor-rich nuclear speckles and assembles transcription and splicing components into complexes through its amino and carboxyl regions. *Mol Cell Biol* **26**(13), 4998-5014.
- Sanchez-Alvarez, M., Montes, M., Sanchez-Hernandez, N., Hernandez-Munain, C., and Sune, C. (2010). Differential effects of sumoylation on transcription and alternative splicing by transcription elongation regulator 1 (TCERG1). *J Biol Chem* **285**(20), 15220-33.
- Sanchez-Hernandez, N., Ruiz, L., Sanchez-Alvarez, M., Montes, M., Macias, M. J., Hernandez-Munain, C., and Sune, C. (2012). The FF4 and FF5 domains of

- transcription elongation regulator 1 (TCERG1) target proteins to the periphery of speckles. *J Biol Chem* **287**(21), 17789-800.
- Savkur, R. S., Philips, A. V., and Cooper, T. A. (2001). Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet* **29**(1), 40-7.
- Schmidt, U., Basyuk, E., Robert, M. C., Yoshida, M., Villemin, J. P., Auboeuf, D., Aitken, S., and Bertrand, E. (2011). Real-time imaging of cotranscriptional splicing reveals a kinetic model that reduces noise: implications for alternative splicing regulation. *J Cell Biol* **193**(5), 819-29.
- Schor, I. E., Rascovan, N., Pelisch, F., Allo, M., and Kornblihtt, A. R. (2009). Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci U S A* **106**(11), 4325-30.
- Schwartz, S., Meshorer, E., and Ast, G. (2009). Chromatin organization marks exon-intron structure. *Nat Struct Mol Biol* **16**(9), 990-5.
- Sharma, A., Takata, H., Shibahara, K., Bubulya, A., and Bubulya, P. A. (2010). Son is essential for nuclear speckle organization and cell cycle progression. *Mol Biol Cell* **21**(4), 650-63.
- Sharma J., Mukherjee D., Rao S. N., Iyengar S., Shankar S. K., Satischchandra P., Jana N. R. (2013). Neuronatin-mediated aberrant calcium signaling and endoplasmic reticulum stress underlie neuropathology in Lafora disease. *J Biol Chem* **288**(13), 9482-90.
- Shi, Y. (2002). Mechanisms of caspase activation and inhibition during apoptosis. *Mol Cell* **9**(3), 459-70.
- Shopland, L. S., Johnson, C. V., Byron, M., McNeil, J., and Lawrence, J. B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. *J Cell Biol* **162**(6), 981-90.
- Sims, R. J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18**(20), 2437-68.
- Sims, R. J., 3rd, Millhouse, S., Chen, C. F., Lewis, B. A., Erdjument-Bromage, H., Tempst, P., Manley, J. L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol Cell* **28**(4), 665-76.
- Sims, R. J., 3rd, Rojas, L. A., Beck, D., Bonasio, R., Schuller, R., Drury, W. J., 3rd, Eick, D., and Reinberg, D. (2011). The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science* **332**(6025), 99-103.

- Smith, M. J., Kulkarni, S., and Pawson, T. (2004). FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. *Mol Cell Biol* **24**(21), 9274-85.
- Spector, D. L., and Lamond, A. I. (2011). Nuclear speckles. *Cold Spring Harb Perspect Biol* **3**(2).
- Spector, D. L., Schrier, W. H., and Busch, H. (1983). Immunoelectron microscopic localization of snRNPs. *Biol Cell* **49**(1), 1-10.
- Spellman, R., and Smith, C. W. (2006). Novel modes of splicing repression by PTB. *Trends Biochem Sci* **31**(2), 73-6.
- Spies, N., Nielsen, C. B., Padgett, R. A., and Burge, C. B. (2009). Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* **36**(2), 245-54.
- Sudol, M. (1996). Structure and function of the WW domain. *Prog Biophys Mol Biol* **65**(1-2), 113-32.
- Sune, C., and Garcia-Blanco, M. A. (1999). Transcriptional cofactor CA150 regulates RNA polymerase II elongation in a TATA-box-dependent manner. *Mol Cell Biol* **19**(7), 4719-28.
- Sune, C., Hayashi, T., Liu, Y., Lane, W. S., Young, R. A., and Garcia-Blanco, M. A. (1997). CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription. *Mol Cell Biol* **17**(10), 6029-39.
- Swift, H. (1959). Studies on nuclear fine structure. *Brookhaven Symp Biol* **12**, 134-52.
- Tacke, R., and Manley, J. L. (1999). Determinants of SR protein specificity. *Curr Opin Cell Biol* **11**(3), 358-62.
- Tejedor, J. R., Papasaikas, P., and Valcarcel, J. (2014). Genome-Wide Identification of Fas/CD95 Alternative Splicing Regulators Reveals Links with Iron Homeostasis. *Mol Cell*.
- Tewari, M., Quan, L. T., O'Rourke, K., Desnoyers, S., Zeng, Z., Beidler, D. R., Poirier, G. G., Salvesen, G. S., and Dixit, V. M. (1995). Yama/ CPP32 beta, a mammalian homolog of CED-3, is a CrmA-inhibitable protease that cleaves the death substrate poly(ADP-ribose) polymerase. *Cell* **81**(5), 801-9.
- Thompson, C. B. (1995). Apoptosis in the pathogenesis and treatment of disease. *Science* **267**(5203), 1456-62.
- Tisserant, A., and Konig, H. (2008). Signal-regulated Pre-mRNA occupancy by the general splicing factor U2AF. *PLoS One* **3**(1), e1418.
- Trinkle-Mulcahy, L., Ajuh, P., Prescott, A., Claverie-Martin, F., Cohen, S., Lamond, A. I., and Cohen, P. (1999). Nuclear organisation of NIPP1, a regulatory subunit of

- protein phosphatase 1 that associates with pre-mRNA splicing factors. *J Cell Sci* **112** (Pt 2), 157-68.
- Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., Freier, S. M., Bennett, C. F., Sharma, A., Bubulya, P. A., Blencowe, B. J., Prasanth, S. G., and Prasanth, K. V. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **39**(6), 925-38.
- Tu, Y., Renner, S., Xu, F., Fleishman, A., Taylor, J., Weisz, J., Vescio, R., Rettig, M., Berenson, J., Krajewski, S., Reed, J. C., and Lichtenstein, A. (1998). BCL-X expression in multiple myeloma: possible indicator of chemoresistance. *Cancer Res* **58**(2), 256-62.
- Vaux, D. L. (2011). Apoptogenic factors released from mitochondria. *Biochim Biophys Acta* **1813**(4), 546-50.
- Vecerova, J., Koberna, K., Malinsky, J., Soutoglou, E., Sullivan, T., Stewart, C. L., Raska, I., and Misteli, T. (2004). Formation of nuclear splicing factor compartments is independent of lamins A/C. *Mol Biol Cell* **15**(11), 4904-10.
- Wagner, S., Chiosea, S., and Nickerson, J. A. (2003). The spatial targeting and nuclear matrix binding domains of SRm160. *Proc Natl Acad Sci U S A* **100**(6), 3269-74.
- Wajant, H. (2002). The Fas signaling pathway: more than a paradigm. *Science* **296**(5573), 1635-6.
- Wang, E. T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**(7221), 470-6.
- Wang, G. S., and Cooper, T. A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. *Nat Rev Genet* **8**(10), 749-61.
- Wansink, D. G., Schul, W., van der Kraan, I., van Steensel, B., van Driel, R., and de Jong, L. (1993). Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J Cell Biol* **122**(2), 283-93.
- Ward, A. J., and Cooper, T. A. (2010). The pathobiology of splicing. *J Pathol* **220**(2), 152-63.
- Wong, J. J., Lau, K. A., Pinello, N., and Rasko, J. E. (2014). Epigenetic modifications of splicing factor genes in myelodysplastic syndromes and acute myeloid leukemia. *Cancer Sci* **105**(11), 1457-63.
- Wyllie, A. H. (1980). Glucocorticoid-induced thymocyte apoptosis is associated with endogenous endonuclease activation. *Nature* **284**(5756), 555-6.

- Xing, Y., Johnson, C. V., Dobner, P. R., and Lawrence, J. B. (1993). Higher level organization of individual gene transcription and RNA splicing. *Science* **259**(5099), 1326-30.
- Yang, Z., Yik, J. H., Chen, R., He, N., Jang, M. K., Ozato, K., and Zhou, Q. (2005). Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19**(4), 535-45.
- Yao, J., Ardehali, M. B., Fecko, C. J., Webb, W. W., and Lis, J. T. (2007). Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. *Mol Cell* **28**(6), 978-90.
- Yonaha, M., and Proudfoot, N. J. (2000). Transcriptional termination and coupled polyadenylation in vitro. *EMBO J* **19**(14), 3770-7.
- Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., Chalkidis, G., Suzuki, Y., Shiosaka, M., Kawahata, R., Yamaguchi, T., Otsu, M., Obara, N., Sakata-Yanagimoto, M., Ishiyama, K., Mori, H., Nolte, F., Hofmann, W. K., Miyawaki, S., Sugano, S., Haferlach, C., Koeffler, H. P., Shih, L. Y., Haferlach, T., Chiba, S., Nakauchi, H., Miyano, S., and Ogawa, S. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**(7367), 64-9.
- Youle, R. J., and Strasser, A. (2008). The BCL-2 protein family: opposing activities that mediate cell death. *Nat Rev Mol Cell Biol* **9**(1), 47-59.
- Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R. V., Gentile, C., Gebara, M., and Corden, J. L. (1996). The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc Natl Acad Sci U S A* **93**(14), 6975-80.
- Zehring, W. A., Lee, J. M., Weeks, J. R., Jokerst, R. S., and Greenleaf, A. L. (1988). The C-terminal repeat domain of RNA polymerase II largest subunit is essential in vivo but is not required for accurate transcription initiation in vitro. *Proc Natl Acad Sci U S A* **85**(11), 3698-702.
- Zeng, C., and Berget, S. M. (2000). Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. *Mol Cell Biol* **20**(21), 8290-301.
- Zhang, Z., Lotti, F., Dittmar, K., Younis, I., Wan, L., Kasim, M., and Dreyfuss, G. (2008). SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* **133**(4), 585-600.
- Zhao, K., Wang, W., Rando, O. J., Xue, Y., Swiderek, K., Kuo, A., and Crabtree, G. R. (1998). Rapid and phosphoinositide-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell* **95**(5), 625-36.

- Zhong, X. Y., Ding, J. H., Adams, J. A., Ghosh, G., and Fu, X. D. (2009). Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. *Genes Dev* **23**(4), 482-95.
- Zhu, H., Hinman, M. N., Hasman, R. A., Mehta, P., and Lou, H. (2008). Regulation of neuron-specific alternative splicing of neurofibromatosis type 1 pre-mRNA. *Mol Cell Biol* **28**(4), 1240-51.

