



Contribución de prefoldina a la expresión génica en células humanas

Laura Payán Bravo

TESIS DOCOTORAL

Universidad de Sevilla

Marzo 2019

Índice

ABREVIATURAS.....	7
RESUMEN.....	13
INTRODUCCIÓN.....	17
1. La expresión génica.	19
1.1 La transcripción.	19
1.1.1 La RNA polimerasa II.	19
1.1.2 Etapas de la transcripción.	20
1.1.3 Regulación de la transcripción.	24
1.2 El <i>splicing</i> .	29
1.2.1 El espliceosoma.	30
1.2.2 Definición del exón y del intrón.	32
1.2.3 El <i>splicing</i> alternativo.	33
1.2.4 Tipos de <i>splicing</i> alternativos.	34
1.3 Acoplamiento de la transcripción y el <i>splicing</i> .	35
2. La prefoldina.	37
2.1 Tipos de prefoldinas y su estructura.	38
2.2 Funciones citoplasmáticas.	41
2.3 Funciones nucleares.	42
3. Antecedentes.	46
OBJETIVOS.....	49
RESULTADOS.....	53
1. PFDN5 se localiza mayoritariamente en el núcleo de células HCT116.	55
2. Estudio transcriptómico en células deficientes en PFDN2 y PFDN5.	58
2.1 Diseño de los experimentos del RNA-seq.	58
2.2 La deficiencia en PFDN5 reduce la viabilidad celular, no así la deficiencia en PFDN2.	59
2.3 La falta de suero modifica ligeramente el ciclo celular de la línea HCT116.	62
2.4 Obtención de muestras para los experimentos de RNA-seq.	63
2.5 La deficiencia de PFDN5, y en menor medida la de PFDN2, afecta a la expresión de un alto número de genes.	64
2.6 La deficiencia de PFDN5 y PFDN2 altera la expresión de categorías funcionales de genes.	68
2.7 La inducción por suero se ve más afectada por la deficiencia en PFDN5 que en PFDN2, aunque hay solapamiento.	75

3.	La ausencia y deficiencia en PFDN5 afecta al ciclo celular.	77
3.1	Tanto la ausencia como la deficiencia de PFDN5 provocan una disminución de los niveles del resto de subunidades, aunque no de forma homogénea.	77
3.2	La ausencia y deficiencia en PFDN5 baja los niveles de expresión de genes de histonas.	78
3.3	La ausencia y deficiencia en PFDN5 incrementa los niveles de expresión de <i>p21</i> .	80
3.4	La deficiencia y ausencia de PFDN5 provocan ligeros cambios en el ciclo celular.	81
4.	Prefoldina 5 se encuentra físicamente asociada al gen <i>p21</i> .	83
5.	La deficiencia en PFDN5 afecta negativamente a la inducción por suero de los genes más largos y con mayor número de intrones.	85
5.1	Tras 90 minutos de inducción por suero, genes más largos o con mayor número de intrones se expresan menos en las células deficientes en PFDN5.	85
5.2	El tamaño de los genes y el número de intrones ejercen un efecto sinérgico en el nivel de inducción tras la adición de suero en las células deficientes en PFDN5.	88
6.	La deficiencia en PFDN5 y PFDN2 aumenta los niveles de pre-mRNA.	89
6.1	La deficiencia en PFDN5 y PFDN2 provoca variaciones en la ratio exónica antes y después de la inducción con suero.	89
6.2	Los cambios en la ratio exónica ocasionados por la deficiencia en PFDN5 se asocian al nivel de expresión de los genes.	90
6.3	Validación de las alteraciones de splicing observadas tras la inducción por suero en células carentes de PFDN5.	93
7.	La deficiencia en PFDN5 afecta sutilmente a la maduración alternativa del pre-mRNA.	95
8.	La deficiencia en PFDN5 ocasiona defectos en el acoplamiento del <i>splicing</i> del pre-mRNA a la transcripción.	99
9.	La deficiencia y ausencia de PFDN5 disminuye los niveles de RNAPII Ser2-P.	101
10.	No logramos detectar interacciones físicas entre PFDN5, CDK9 y factores de <i>splicing</i> .	104
11.	PFDN5 se asocia físicamente a los genes <i>CTNNBL1</i> y <i>CD44</i> .	105

DISCUSIÓN.....	107
1. Algunos de los cambios observados en células deficientes en PFDN2 y PFDN5 podrían estar mediados por la función citoplásmica de la prefoldina.	109
2. Otros de los cambios observados en células deficientes en PFDN5 podrían deberse a su relación con c-Myc.	110
3. Cambios de expresión de genes relacionados con el control del ciclo celular.	111
4. PFDN5 contribuye de forma general al proceso de transcripción dependiente de RNAPII.	112
5. La deficiencia o carencia de PFDN5 causa defectos en el <i>splicing</i> del pre-mRNA.	114
CONCLUSIONES.....	119
MATERIALES Y MÉTODOS.....	123
1. Organismo y condiciones de cultivo.	125
1.1 Línea bacteriana y condiciones de cultivo.	125
1.2 Líneas celulares y condiciones de cultivo.	125
2. Oligonucleótidos y plásmidos.	126
2.1 Oligonucleótidos.	126
2.2 Pequeños RNAs de interferencia (siRNAs).	128
2.3 Plásmidos.	128
3. Manipulación y análisis del DNA.	130
3.1 Asilamiento de DNA plasmídico bacteriano.	130
3.2 Electroforesis de DNA en geles de agarosa.	131
3.3 Purificación, cuantificación y secuenciación de DNA.	131
3.4 Amplificación de fragmentos de DNA mediante reacción en cadena de la polimerasa.	131
3.5 PCR a tiempo real.	132
3.6 Transformación de bacterias por quimiocompetencia.	132
3.7 Transfección transitoria de líneas celulares animales con DNA plasmídico.	133
4. Manipulación y análisis del RNA.	133
4.1 Aislamiento y cuantificación del RNA de líneas celulares animales.	133
4.2 Tratamiento con DNAsa y reacción de retrotranscripción.	133
4.3 Transfección transitoria de líneas celulares animales con siRNAs.	134
5. Extracción y análisis de proteínas.	134
5.1 Obtención de extractos totales de proteínas de líneas celulares animales.	134
5.2 Electroforesis de proteínas en geles de SDS-poliacrilamida.	135
6. Métodos inmunológicos.	135
6.1 Anticuerpos	135

6.2 Ensayo de <i>Western Blot</i> (WB).	137
6.3 Co-inmunoprecitación.	138
6.4 Inmunoprecitación de la cromatina (ChIP).	139
6.5 Inmunofluorescencia (IF).	141
7. Microscopía de fluorescencia	142
8. Determinación de focos de γ -H2A.X: IF y conteo.	142
9. Ensayo de viabilidad celular.	142
10. Análisis del ciclo celular por citometría de flujo.	143
11. Análisis masivo de la expresión génica mediante RNA-seq.	144
11.1 Metodología.	144
11.2 Análisis bioinformático.	145
12. Ensayos de <i>splicing</i> .	146
12.1 Minigenes.	146
12.2 Análisis de la velocidad y procesividad de la RNAPII y el <i>splicing</i> co-transcripcional.	146
BIBLIOGRAFÍA.....	149

Abreviaturas

A

A	Adenina
aa	aminoácido
ATP	Adenosín trifosfato

B

BrdU	5-bromo-2-desoxiuridina
BSA	Albúmina de suero bovino

C

°C	Grado centígrado
C	Citosina
cDNA	Ácido desoxirribonucleico complementario
ChIP	Inmunoprecipitación de la cromatina
CMV	Citomegalovirus
CTD	Dominio carboxi terminal

D

DAPI	4',6-Diamino-2-fenilindol diclorhidrato
DNA	Ácido desoxirribonucleico
dNTPs	Desoxirribonucleótidos trifosfatos
DRB	5,6-Diclorobenzimidazol 1-β-D-ribofuranosido
DSP	Ditiobis (succinimidil-propionato)

E

EDTA	Ácido etilendiaminotetraacético
EGFP	Versión mejorada de GFP
EGTA	Ácido etilenglicol-bis [-aminoetil éter] N, N, N', N' tetraacético
Ex	Exón

F

FBS	Suero fetal bovino
FDR	Tasa de descubrimientos falsos
Fw	Sentido 5'-3'

G

g	Aceleración gravitatoria terrestre
G	Guanina
GAPDH	Gliceraldehido 3-fosfato deshidrogenasa
GFP	Proteína fluorescente verde
GO	Categorías funcionales o procesos ontológicos

H

H	Hora
HEPES	Ácido 2-(4-(2-hidroxietil)-1-piperacil-(1)) etanosulfónico

I

IEG	Gen de respuesta inmediata
IF	Inmunofluorescencia
In	Intrón

K

kb	Kilobase
kDa	Kilodalton
KO	Del inglés <i>knockout</i> . Hace referencias a mutantes a los que se le ha suprimido la expresión de un determinado gen

L

LB	Medio Luria-Bertani
Log ₂ FC	Logaritmo en base 2 del cambio

M

M	Molar
mM	miliMolar
mRNA	RNA mensajero

N

nt	Nucleótidos
----	-------------

P

pb	Pares de bases
PBS	Solución salina fosfato
PCR	Reacción en cadena de la polimerasa
PFD	Complejo prefoldina
Pfd1	Subunidad 1 del PFD en <i>S.cerevisiae</i>
Pfd2	Subunidad 2 del PFD en <i>S.cerevisiae</i>
Pfd3	Subunidad 3 del PFD en <i>S.cerevisiae</i>
Pfd4	Subunidad 4 del PFD en <i>S.cerevisiae</i>
Pfd5	Subunidad 5 del PFD en <i>S.cerevisiae</i>
Pfd6	Subunidad 6 del PFD en <i>S.cerevisiae</i>
PFDN1	Subunidad 1 del PFD en eucariotas superiores
PFDN2	Subunidad 2 del PFD en eucariotas superiores
PFDN3	Subunidad 3 del PFD en eucariotas superiores
PFDN4	Subunidad 4 del PFD en eucariotas superiores
PFDN5	Subunidad 5 del PFD en eucariotas superiores
PFDN6	Subunidad 6 del PFD en eucariotas superiores
PMSF	Fluoruro de fenilmetilsulfonilo
pre-mRNA	mRNA inmaduro
p/v	peso/volumen

Q

qPCR	Reacción en cadena de la polimerasa cuantitativa
------	--

R

RNA	Ácido ribonucleico
RNAPI	RNA polimerasa I
RNAPII	RNA polimerasa II
RNAPIII	RNA polimerasa III
RT	Temperatura ambiente
RT-qPCR	Reacción de la retrotranscriptasa seguida de amplificación de cDNA mediante PCR cuantitativa
Rv	Sentido 3'-5'

S

SD	Desviación estándar
SDS	Sodio dodecil sulfato
Ser	Serina
siRNA	RNA interferente pequeño
snRNA	RNA nucleolar pequeño
snRNP	Partículas ribonucleoproteicas nucleolares pequeñas
SS	Sitio de <i>splicing</i>

T

T	Timina
TE	Tampón Tris/EDTA
TSS	Sitio de comienzo de la transcripción
Tris	Tris-(hidroximetil) amino metano

V

v/v	Volumen/volumen
-----	-----------------

W

WT	Del inglés <i>wild type</i> . Hace referencia a la estirpe silvestre
----	--

X

X-gal	5-bromo-4-cloro-3-indolil-D-galactopiranosido
-------	---

Resumen

La prefoldina es una cochaperona que está presente en todos los organismos eucarióticos. Es conocida principalmente por su importante función en la dinámica del citoesqueleto, al favorecer el plegamiento de los monómeros de actina y tubulina durante su ensamblaje en filamentos de actina y microtúbulos, respectivamente. Adicionalmente, diferentes subunidades de este complejo heterohexamérico han sido encontradas en el núcleo y relacionadas con procesos nucleares en levaduras, plantas y metazoos.

En esta tesis nos hemos propuesto estudiar qué papel desempeña la prefoldina en la expresión génica en células humanas. Para ello, hemos realizado un ensayo de RNA-seq que nos permitiera analizar qué efectos ocasionaba la deficiencia de las subunidades PFDN2 y PFDN5 en el transcriptoma de las células, bajo diferentes condiciones nutricionales.

Los resultados mostraron que la deficiencia de estas dos subunidades afectó a la expresión de multitud de genes, siendo el efecto de PFDN5 más severo que el de PFDN2. Una parte de estos defectos puede ser explicado como consecuencia indirecta del papel de la prefoldina en la dinámica del citoesqueleto. Otros, pueden explicarse por la conocida función de PFDN5 como co-represor del protooncogén c-Myc.

Además de estos efectos predecibles según lo conocido previamente, en nuestros resultados encontramos evidencias de la participación de estas subunidades en otros fenómenos. Por ejemplo, hemos encontrado un efecto de PFDN5 sobre la expresión del gen *CDKN1A*, que codifica para p21, un represor de la progresión del ciclo celular. Además, hemos detectado que PFDN5 se localiza físicamente sobre dicho gen. Este resultado nos permite explicar su influencia sobre la expresión de genes que se expresan cíclicamente, como los que codifican histonas.

Asimismo, hemos detectado que las células deficientes en PFDN5 presentan deficiencias en la expresión de genes en función de sus características físicas, no funcionales, como su longitud o el número de intrones que presentan. También observamos que la deficiencia de PFDN5, y en menor medida la de PFDN2, empeoró de forma generalizada la eficiencia del procesamiento de intrones en el pre-mRNA, especialmente en los transcritos inmaduros de los genes que se expresan más.

Igualmente hemos comprobado que las células deficientes en PFDN5 mostraron defectos en el acoplamiento temporal entre *splicing* y elongación de la transcripción, así como una disminución drástica de los niveles de fosforilación de la RNA polimerasa II en los residuos Ser2 de su dominio CTD. Los genes que muestran estas deficiencias en ausencia de PFDN5 presentan ocupación de esta proteína a lo largo de su región transcrita.

En conjunto, nuestros resultados apoyan que la prefoldina contribuye de forma general al proceso de expresión génica en células humanas, al margen de los efectos regulatorios que sus otras funciones celulares puedan causar indirectamente.

Introducción

1. LA EXPRESIÓN GÉNICA.

La expresión génica es el proceso por el cual la información contenida en el genoma es utilizada para dirigir la síntesis de un producto génico específico según el requerimiento de cada organismo en un determinado momento. La molécula más importante en este proceso es el RNA (*ácido ribonucleico*), pues se encarga de transferir la información genética desde el núcleo, donde reside el DNA (*ácido desoxirribonucleico*), al citoplasma, donde se lleva a cabo la síntesis proteica por los ribosomas.

1.1. La transcripción.

Una de las principales etapas de la expresión génica es la transcripción, donde a partir del DNA se van a sintetizar los distintos tipos de RNA. Entre ellos, los RNA mensajeros (mRNA) van a ser procesados para formar mRNA maduros. Finalmente, serán exportados al citoplasma y traducidos a proteínas por los ribosomas (Svetlov and Nudler, 2013).

1.1.1. La RNA polimerasa II.

En los organismos eucariontes, la transcripción es llevada a cabo por tres polimerasas diferentes: la RNA polimerasa I, que transcribe únicamente genes que codifican para los RNA ribosómicos (rRNA) 5.8S, 18S y 28S; la RNA polimerasa II (RNAPII), que transcribe genes que codifican proteínas, así como la mayoría de RNAs no codificantes pequeños nucleares (snRNAs); y la RNA polimerasa III (RNAPIII), que transcribe genes cortos que dan lugar a RNA transferentes (tRNA) y al rRNA 5S (Roeder, 2005). Todas ellas comparten homología en muchas de sus subunidades (Allison et al., 1985; Cramer, 2002). Sin embargo, la subunidad mayor (RPB1) de la RNAPII presenta un extenso dominio carboxiterminal (CTD) sin una estructura definida y evolutivamente conservado, del que carecen el resto de las polimerasas (Chapman et al., 2008; Corden et al., 1985). El CTD está compuesto de varias repeticiones del heptapéptido Tyr₁-Ser₂-Pro₃-Thr₄-Ser₅-Pro₆-Ser₇. Entre los eucariontes, hay una relación directa entre la

complejidad del genoma y la longitud del CTD, desde las 26 repeticiones de levaduras hasta las 52 de mamíferos (Egloff and Murphy, 2008).

El CTD de la RNAPII es esencial para la viabilidad de los organismos y la síntesis del RNA, aunque se ha visto que deleciones parciales permiten la supervivencia de éstos (Nonet *et al.*, 1987; Litingtung *et al.*, 1998; Meininghaus *et al.*, 2000). Sin embargo, este dominio es prescindible en la iniciación de la transcripción *in vitro* (Fong, 2003; McCracken *et al.*, 1997; Zehring *et al.*, 1988).

Las fosforilaciones de algunos de los residuos del CTD están relacionadas con diferentes etapas de la transcripción, lo que da lugar a lo que se conoce como código CTD (Egloff and Murphy, 2008). De éstas, las más estudiadas son las fosforilaciones en serina 2 (Ser2) y serina 5 (Ser5) (ver apartado 1.1.2) (Egloff and Murphy, 2008; Patturajan *et al.*, 1998). También la serina 7 (Ser7) es fosforilada, jugando un papel fundamental en la transcripción de snRNAs (ver apartado 1.1.2) (Chapman *et al.*, 2007; Egloff *et al.*, 2007; Yurko and Manley, 2018), así como la treonina 4 (Thr4), requerida para un correcto procesamiento del extremo 3' de los mRNA de histonas (Hsin *et al.*, 2011). Por tanto, en el código CTD se requieren quinasas y fosfatasas capaces de discernir entre los distintos residuos del CTD que van a fosforilar o desfosforilar, respectivamente.

1.1.2. Etapas de la transcripción.

Clásicamente, el proceso de transcripción se divide en 3 etapas: iniciación, elongación y terminación. En cada una de ellas intervienen numerosos factores reguladores como la secuencia del DNA, la estructura de la cromatina (influenciada por las modificaciones de histonas y/o los factores remodeladores de la cromatina) o los factores de transcripción (Panning and Taatjes, 2008). Además, el patrón de fosforilación también regula el acoplamiento de cada fase de la transcripción con otros procesos necesarios para la expresión génica, respondiendo de este modo a diferentes condiciones celulares y extracelulares (Egloff and Murphy, 2008; Montes *et al.*, 2012).

Iniciación.

La iniciación de la transcripción comienza con la unión de activadores de la transcripción, en respuesta a un estímulo, a secuencias específicas del DNA situadas aguas arriba del sitio de inicio de la transcripción (TSS, *transcription start site*). Aunque esta unión es necesaria para inducir la expresión génica, hay genes que muestran niveles de expresión basal en su ausencia (Pugh, 2000). A su vez, estos activadores van a reclutar a otros factores llamados co-activadores, los cuales desempeñarán diversas funciones como regular la actividad del complejo transcripcional, modificar y remodelar la cromatina, comunicar con los activadores y reclutar a la maquinaria transcripcional (Taatjes et al., 2004). Entre los principales co-activadores se encuentra el complejo multiproteico *Mediator*, el cual es necesario para el ensamblaje del complejo de preiniciación (PIC, *pre-initiation complex*) (Conaway et al., 2005). El PIC está formado por la RNAPII y los factores generales de transcripción (GTFs), los cuales posicionan a la RNAPII sobre los elementos reguladores del DNA, determinando, de este modo, el sitio de inicio de la transcripción. Los GTFs son: TFIIA, TFIIB, TFIID (complejo multiproteico capaz de unirse a secuencias promotoras específicas), TFIIE, TFIIF y TFIIH (complejo encargado de fosforilar la Ser5 del CTD de la RNAPII), todos ellos muy conservados evolutivamente (Buratowski, 2009; Roeder, 1996). La RNAPII y los GTFs, una vez posicionados en el promotor, abrirán la doble cadena del DNA y comenzará la transcripción, formándose los primeros enlaces entre ribonucleótidos complementarios a los de la cadena molde (Grünberg and Hahn, 2013; Sainsbury et al., 2015).

Elongación.

Durante la elongación se produce la adición de ribonucleótidos a la cadena de RNA creciente. Esta etapa se divide a su vez en dos partes: elongación temprana y elongación tardía, diferenciadas principalmente en el residuo de serina del CTD que se encuentra fosforilado en cada una de ellas.

Una vez terminada la fase de iniciación (**Figura I-1A**), la serina 5 del CTD es fosforilada por la quinasa CDK7 (*cyclin dependent kinase 7*) que forma parte del TFIIH (**Figura I-1B**), provocando que el complejo *Mediator* se disocie de la RNAPII y desestabilice la unión entre la RNAPII y el PIC (Max et al., 2007) Esta fase transitoria

entre la iniciación y la elongación temprana se conoce como *promoter clearance* o escape del promotor, puesto que dicha modificación facilita que la RNAPII escape del promotor (Buratowski, 2009; Saunders et al., 2006). En la mayoría de los genes de mamíferos, la elongación temprana acaba en una pausa de la RNAPII a unos 20-100 pb del promotor (Jonkers et al., 2014). La fosforilación en Ser2 es necesaria para la salida de esta pausa (Ni et al., 2008). Durante la elongación tardía, la serina 2 del CTD es fosforilada por la quinasa CDK9, subunidad catalítica del factor de elongación positivo P-TEFb (*Positive Transcription Elongation Factor-b*), aunque también pueden hacerlo las quinasas CDK12 y CDK13 (**Figura I-1C**). Esta modificación convierte a la RNAPII en su forma procesiva (Buratowski, 2009; Liang et al., 2015). En esta fase también va a ser fosforilada la Ser7, lo cual es necesario para reclutar a la fosfatasa que va a desfosforilar a Ser5-P (Yurko and Manley, 2018). Las fosforilaciones en Ser2 y Ser5 permiten a la RNAPII reclutar a diferentes enzimas y factores que intervienen en los procesos de poliadenilación, *capping* y *splicing* (**Figura I-1C**) así como metiltransferasas y desacetilasas de histonas (Egloff and Murphy, 2008; Jeronimo et al., 2013). También la fosforilación en Ser7 facilita la interacción entre la maquinaria de transcripción e *Integrator*, un complejo que participa en el procesamiento del extremo 3' de los snRNAs (Yurko and Manley, 2018).

Terminación.

La terminación de la transcripción es esencial para evitar la transcripción de genes que estén corriente abajo y prevenir la formación de RNAs antisentidos que puedan interferir con la síntesis normal de pre-RNAs (Richard and Manley, 2009).

Esta etapa está íntimamente relacionada con el procesamiento del extremo 3' del RNA y la poliadenilación (Yonaha and Proudfoot, 2000) (**Figura I-1D**). De forma general, el reconocimiento del sitio polyA produce cambios conformacionales en el complejo de elongación, haciendo que la RNAPII se disocie de ciertos factores de elongación al mismo tiempo que otros necesarios para la terminación de la transcripción son reclutados (Richard and Manley, 2009).

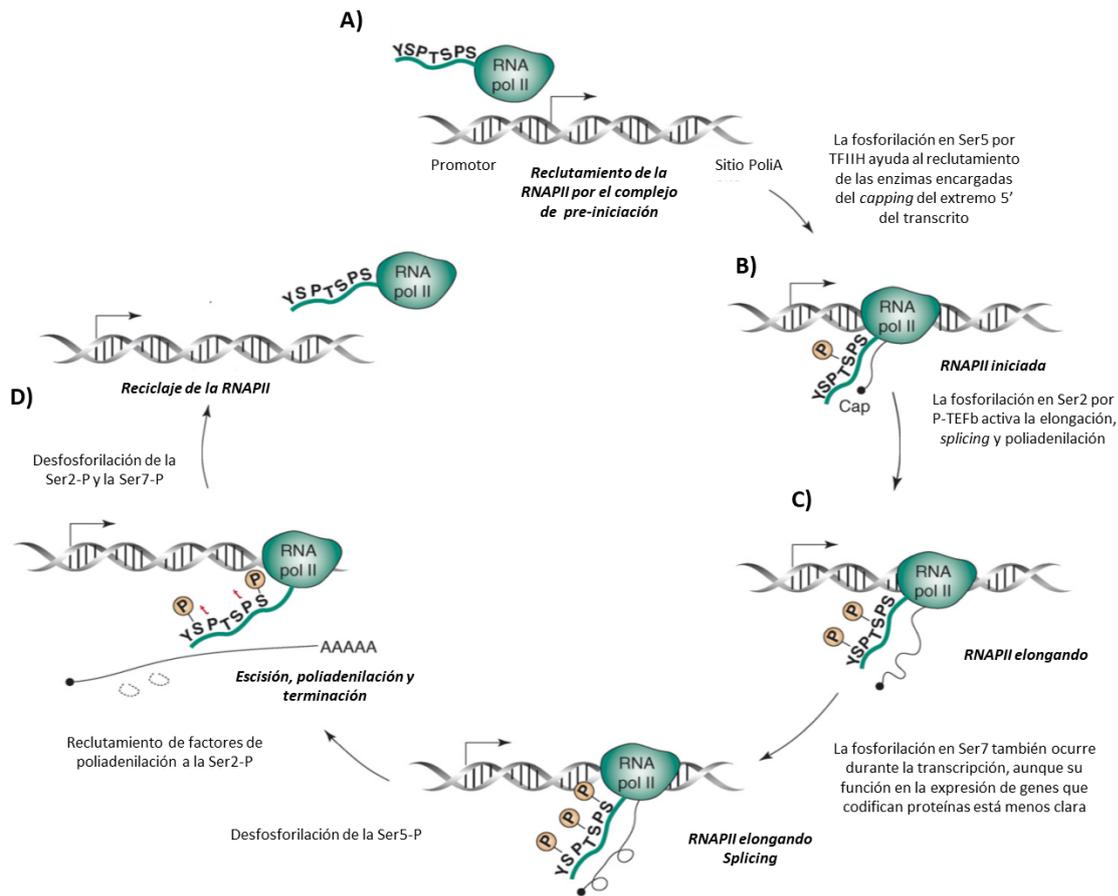


Figure I-1. Modificación del CTD de la RNAPII durante la transcripción de genes que codifican proteínas. **A)** El CTD de la RNAPII, la cual es reclutada al promotor por el complejo de pre-iniciación, está desfosforilado. **B)** El CTD se localiza cerca del canal de salida del RNA. La fosforilación de la Ser5 por la subunidad CDK7 del factor general de transcripción TFIIH, justo después de la iniciación, ayuda a reclutar y activar a las enzimas encargadas del *capping* del transcrito emergente. **C)** La subsecuente fosforilación en la Ser2 por la subunidad CDK9 del factor de elongación positivo b (P-TEFb), activa la elongación y el procesamiento del RNA. La fosforilación de la Ser7 también ocurre durante la transcripción, aunque su función en la expresión de genes que codifican proteínas está menos clara (Yurko and Manley, 2018). En levaduras y algunos genes de mamíferos, la Ser5 es desfosforilada a medida que la transcripción avanza hacia el extremo 3' (Phatnani and Greenleaf, 2008). **D)** Tras la escisión del pre-mRNA y la poliadenilación, se produce la desfosforilación del CTD, lo cual podría ayudar a que la RNAPII se desacople y esté lista para otra ronda de transcripción. Las fosforilaciones se muestran en círculos naranjas con una "P". Imagen adaptada de Egloff and Murphy, 2008.

En los genes de mamíferos que codifican proteínas, la terminación de la transcripción se desencadena por el factor específico de corte y poliadenilación (CPSF), el factor estimulador del corte (CstF), y los factores de corte CFI y CFII (Porrua and Libri, 2015). Cuando CPSF, junto con CstF, reconozcan secuencias específicas, como la señal de poliadenilación (PAS), en la región 3' que no se traduce del RNA, los factores de corte

escinden el RNA. A continuación, éste es poliadenilado en el extremo 3' para favorecer su exportación al citoplasma donde será traducido (Porrúa and Libri, 2015). El corte del extremo 3' del mRNA proporciona un sitio de entrada para la exonucleasa 5'-3' Xrn2, la cual va a degradar el mRNA naciente hasta alcanzar a la RNAPII que lo está transcribiendo (Porrúa and Libri, 2015). Esto hará que la RNAPII sea liberada de la cadena molde de DNA poniendo fin al proceso de transcripción (Richard and Manley, 2009).

Una vez liberada la RNAPII, ésta se recicla para sucesivos ciclos de transcripción, lo cual facilita nuevas rondas productivas de elongación (**Figura I-1D**) (Yao et al., 2007).

1.1.3. Regulación de la transcripción.

Hace algo más de una década, la iniciación de la transcripción era considerada como el paso limitante en la regulación de la transcripción. Sin embargo, numerosos estudios realizados en los últimos años establecen la fase de elongación como una etapa determinante en la regulación de este proceso en eucariontes (Price, 2008).

Factores reguladores de la elongación de la transcripción.

Los factores que regulan la elongación transcripcional se pueden clasificar en dos grupos: factores negativos de elongación (N-TEFs, *Negative transcription factor*) y factores positivos de elongación (P-TEFs, *Positive transcription factor*) (Marshall and Price, 1992). Entre los primeros, estarían los factores DSIF y NELF (que como se explicará en siguiente apartado, promueven la parada de la RNAPII). Y pertenecerían a los P-TEFs tanto los clásicos factores TFIIIS, TFIIIF y p-TEFb, como otros que también promueven la elongación de la transcripción, por ejemplo, las quinasas/fosfatasa encargadas de modificar el CTD de la RNAPII (Sims et al., 2004).

A medida que la RNAPII va elongando, va encontrando obstáculos como los nucleosomas o la tensión torsional derivada de la apertura necesaria de la doble cadena, conocida como burbuja transcripcional (Kouzine et al., 2013; Kulaeva et al., 2013). Algunos de esos obstáculos pueden resultar en la terminación prematura de la transcripción (Sheridan et al., 2019). Como consecuencia, los genes más largos, en los

que la probabilidad de terminación temprana aumenta, son más dependientes de factores de elongación (Chávez et al., 2001; King et al., 2013; Sheridan et al., 2019)

La velocidad de la RNAPII a lo largo de los genes no es constante, añadiendo una nueva variable que afecta a la producción del mRNA (Jonkers et al., 2014). Además, se ha visto que dicha velocidad puede afectar a la fidelidad del *splicing* (Jonkers and Lis, 2015).

Aparte de los clásicos factores de elongación mencionados anteriormente, otros factores van a regular la elongación indirectamente, como por ejemplo: las topoisomerasas, que resuelven las torsiones producidas en el DNA durante la transcripción o las proteínas implicadas en reparación del DNA (García-Muse and Aguilera, 2016); los factores implicados en el procesamiento del mRNA y su posterior exportación fuera del núcleo (Fong and Zhou, 2001; Reed, 2003); e incluso proteínas del proteasoma (Ferdous et al., 2001). Por todo ello, la elongación se propone como un proceso central que coordina múltiples fases de la biogénesis y maduración del mRNA (Sims et al., 2004).

Parada próxima al promotor.

El primer ejemplo donde se vio que esta fase debía estar regulada, fue en el gen *Hsp70* de *Drosophila melanogaster*, el cual codifica una proteína de choque térmico (*heat-shock protein*). En condiciones en las que no existía inducción de la expresión, se observó que este gen era incapaz de producir moléculas de mRNA completas, y, además, se detectó una acumulación de RNAPII en la zona próxima al inicio de la transcripción de este gen, a lo que se le llamó “parada próxima al promotor” (*promoter proximal pausing*) (Rougvie and Lis, 1988).

Este fenómeno, por el cual la maquinaria transcripcional se detiene a unos 20-100 nucleótidos del sitio de inicio de la transcripción y sólo progresa de manera eficiente si se producen las señales apropiadas, constituye un importante punto de control de la transcripción en la gran mayoría de genes (Jonkers et al., 2014). Los principales factores implicados en esta regulación son DSIF (heterodímero formado por Spt4 y Spt5) y NELF (complejo multiproteico constituido por 4 subunidades), los cuales inducen la parada transcripcional de la RNAPII al unirse a ella (y posiblemente, al transcrito que ya se ha

comenzado a sintetizar) e incrementar su afinidad al promotor (**Figura I-2**). La reactivación de la maquinaria transcripcional se produce por el reclutamiento de P-TEFb (quinasa compuesta por CDK9 y la ciclina T), que fosforila el CTD en la serina 2, haciendo reversible la inhibición causada por DSIF y NELF, y por tanto, reanudando la transcripción (**Figura I-2**) (Yamaguchi, Shibata and Handa, 2013).

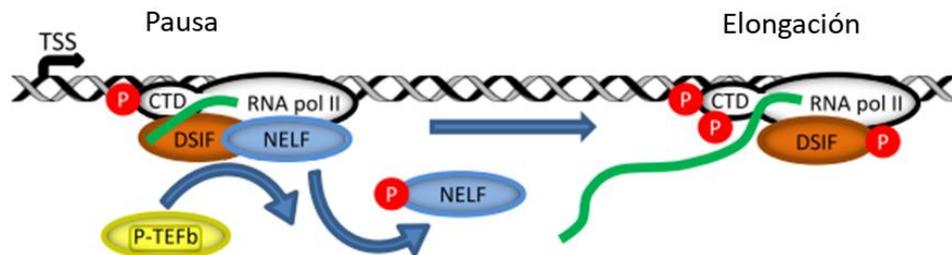


Figura I-2. Modelo de activación por estimulación específica de la transcripción de los IEGs. La transcripción comienza en el sitio de comienzo de la transcripción (TSS). El complejo DSIF-NELF detiene directamente a la RNAPII en la región próxima al promotor de los IEGs. Tras la estimulación, P-TEFb es reclutada a la maquinaria transcripcional, lo que provoca que NELF se desprenda del promotor y DSIF funcione ahora como factor de elongación acelerado, lo cual reactiva la transcripción. Imagen adaptada de Bahrami and Drabløs, 2016.

El factor transcripcional P-TEFb mencionado anteriormente, forma parte de una gran variedad de complejos, entre ellos, el súper complejo de elongación (SEC) (Lin et al., 2010). SEC es el principal regulador de los genes de respuesta inmediata (IEG), los cuales se caracterizan por presentar cinéticas de inducción rápida en respuesta a un estímulo y RNAPII pausadas en la región próxima al promotor (Bahrami and Drabløs, 2016) (**Figura I-2**). Estudios realizados en células HCT116, incubadas en medio sin suero durante 40 horas, demostraron que este complejo fue reclutado a los IEGs 30 minutos después de la inducción con suero, justo en el momento en el que estos genes presentaron su pico máximo de expresión (Lin et al., 2011).

El suero es, por tanto, un elemento que también va a regular la transcripción de muchos genes a través de factores de transcripción específicos conocidos como factores de respuesta al suero (SRF) (Miano, 2003). La expresión de los SRF conlleva a una reprogramación génica que implica tanto la reducción como el incremento de la expresión de diversos reguladores transcripcionales, como por ejemplo, los encargados

de mantener las células en la fase G₀ del ciclo celular (Iyer et al., 1999) o los que ejercen un papel en la diferenciación celular (Miano, 2003), respectivamente.

Influencia de la cromatina en la transcripción.

Otro de los elementos que influyen en la regulación de la transcripción es la cromatina. La cromatina es la estructura altamente organizada y dinámica que forma el DNA empaquetado en el núcleo eucariótico. Los nucleosomas son las unidades básicas que la constituyen, y están formados por un octámero de histonas, constituido por dos subunidades de cada una de las cuatro histonas nucleosómicas (H2A, H2B, H3 y H4), rodeadas de 147 pares de bases de DNA en forma de súper-hélice levógira (**Figura I-3**) (Kornberg and Thomas, 1974; Luger et al., 1997; Richmond et al., 1984).

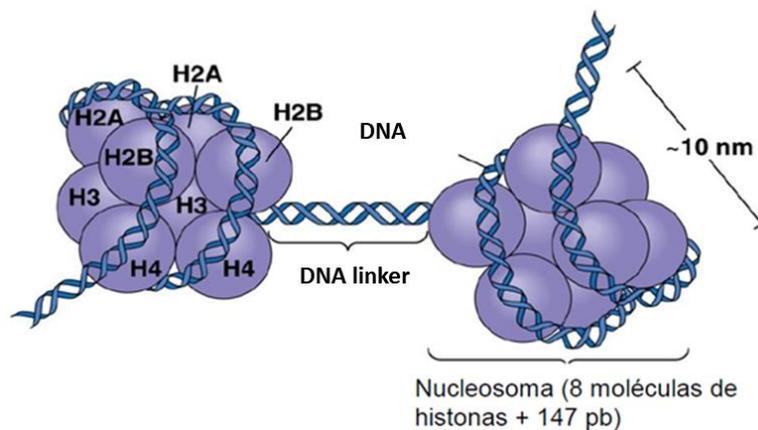


Figura I-3. Representación esquemática del nucleosoma de mamíferos. El nucleosoma es la unidad básica de la cromatina y está formado por un octámero de histonas (H2A-H2B) x2 y (H3-H4) x2. La doble hélice de DNA rodea al octámero de histona 1,7 veces. Imagen adaptada de http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-18/18_21.jpg

En la década de los 60, se observó por primera vez que las histonas podían ser modificadas, y se pensó que esto podría estar influyendo en la síntesis del RNA (Allfrey et al., 1964). Hoy en día se sabe que tanto sus colas como sus dominios globulares pueden sufrir modificaciones postraduccionales, algunas de las cuales influyen en la regulación transcripcional, como las metilaciones, ubiquitinaciones, fosforilaciones y acetilaciones (Gates et al., 2017; Li et al., 2007; Meas and Mao, 2015; Zhou et al., 2017). Éstas se dividen en dos grupos, según si están asociadas con la transcripción activa (modificaciones eucromatínicas) o con regiones o genes inactivos (modificaciones heterocromatínicas). Este fenómeno da lugar a lo que se conoce como “código de

histonas”, y va a regular el estado de la transcripción (activa, reprimida o silenciada) modificando el estado de apertura de la cromatina, y consecuentemente, la velocidad a la que la RNAPII transcribe (Li et al., 2007). También, ciertas marcas de histonas son capaces de reclutar factores transcripcionales y/o factores de *splicing* (Li et al., 2007). Además de esto, las marcas de histonas también controlan la replicación, recombinación y reparación del DNA (Li et al., 2007).

Pero la forma en que la cromatina regula la síntesis de RNA no reside únicamente en las modificaciones postraduccionales de las histonas que la componen, ya que existen otros mecanismos que también lo hacen: i) los nucleosomas pueden actuar como barreras de la RNAPII disminuyendo la tasa de elongación y provocando una acumulación de la RNAPII (Subtil-Rodríguez and Reyes, 2010); ii) la estructura de la cromatina también afecta a la velocidad de transcripción de la RNAPII. Por ejemplo, en el laboratorio del Dr. Reyes hemos demostrado que una cromatina con un bajo suministro de histonas produjo un aumento de la tasa de elongación de la RNAPII, que además afectó a la regulación del *splicing* alternativo del gen *CD44* (Jimeno-González et al., 2015); iii) el desalojo de histonas también es un proceso que influye en la transcripción (Li et al., 2007); iv) diversos factores remodeladores de la cromatina, como por ejemplo FACT o SWI/SNF, intervienen en la regulación transcripcional al modificar localmente la estructura de la cromatina y hacerla accesible al paso de la RNAPII (Li et al., 2007); v) y por último, la metilación del DNA en islas CpG está correlacionada con la represión transcripcional (Goldberg et al., 2007).

Regulación transcripcional durante el ciclo celular.

La transcripción de muchos genes también se encuentra regulada por el ciclo celular. Los cambios dinámicos en la expresión génica dependientes de la progresión del ciclo celular son llevados a cabo por las ciclinas dependiente de quinasas (CDKs), que son factores transcripcionales que regulan ciertos genes a lo largo del ciclo (Bertoli et al., 2013). En la mayoría de eucariontes, la transcripción regulada por el ciclo celular se agrupa en 3 grandes oleadas de expresión que coinciden con los diferentes puntos de transición de fases de este proceso, que son: G1/S, G2/M y M/G1 (Bertoli et al., 2013).

Entre los genes con expresión regulada por el ciclo celular se encuentran los de histonas, los cuales incrementan notablemente el nivel de sus transcritos durante la fase S (Ewen, 2000). En esta etapa del ciclo se va a llevar a cabo la replicación del genoma, por lo que el papel de las histonas empaquetando al DNA en el núcleo es esencial para que este proceso se lleve a cabo de manera eficiente (Ewen, 2000). La abundancia de los mRNA de histonas durante la fase S no solo se debe a una mayor transcripción de éstos, sino también al procesamiento de sus pre-RNAs que implica el corte endonucleolítico de sus extremos 3', y también a la estabilidad de los mensajeros mediada por las estructuras secundarias que se forman en el mismo extremo (Ewen, 2000). En fase G2, los niveles de mRNA de las histonas decaen, puesto que ya no se requiere un alto contenido de éstas en la cromatina (Ewen, 2000).

Las CDKs también regulan los *checkpoints* o puntos de control del ciclo de la fase en la que se encuentran expresadas (Bertoli et al., 2013). La actividad de éstas en dicho proceso viene determinada por los inhibidores de CDKs (CKIs), como por ejemplo, p21 (Bertoli et al., 2013). p21 es un inhibidor de CDK capaz de parar la progresión del ciclo celular en G1/S. La inducción de p21 por p53, en respuesta al daño en el DNA, va a inhibir a los complejos ciclinas-CDKs de fase G1 y a ICBP90 (un factor de transcripción implicado en la replicación del DNA) evitando la transición G1/S (Bertoli et al., 2013). También se ha visto que altos niveles de RAF, un efector de la vía de señalización RAS, induce la expresión de p21 y consecuentemente el arresto de las células en fase G1 (Bertoli et al., 2013). Por el contrario, la vía de señalización STAT induce la expresión de proteínas que antagonizan los efectos de p21, como por ejemplo c-Myc, lo cual hace que se reduzca el arresto de las células en fase G1 (Bertoli et al., 2013). Además de lo anterior, p21 también es capaz de inhibir el complejo ciclina-CDK de la fase G2, evitando la transición G2/M del ciclo celular (Bertoli et al., 2013).

1.2. El *splicing*.

En eucariontes, el RNA inmaduro producido por la RNAPII es generalmente procesado antes de dar lugar al RNA maduro. Los principales procesos de maduración de los transcritos primarios son: la adición de residuo de metilguanosa en el extremo 5' del RNA o *capping*; la poliadenilación y corte del extremo 3'; y la eliminación de

intrones (secuencias no codificantes) y posterior ligación de exones (secuencias codificantes) o *splicing*. Por interés de este trabajo, se hará a continuación una descripción más detallada sobre el proceso de *splicing*.

1.2.1. El espliceosoma.

Durante el *splicing*, las secuencias intrónicas son identificadas y eliminadas de los transcritos del precursor del mRNA (pre-mRNA) y las secuencias exónicas situadas a ambos lados de los intrones son unidas de forma secuencial, para dar lugar a un mRNA maduro, que posteriormente será traducido a proteína (Gilbert, 1978). Este proceso es llevado a cabo por el espliceosoma. En la mayoría de los organismos eucariotas, existen 2 tipos de espliceosomas: el espliceosoma mayor (que es el más frecuente) y el menor (Wahl et al., 2009).

Espliceosoma mayor.

El espliceosoma mayor es un complejo macromolecular formado por las cinco partículas ribonucleoproteicas ricas en uridina U1, U2, U4, U5 y U6 snRNPs (*Uridine-rich small nuclear ribonucleoprotein*) que constituyen el centro catalítico, y más de 150 proteínas auxiliares. A su vez, las partículas snRNP están compuestas por un snRNA y al menos siete proteínas auxiliares. Para que el *splicing* se desarrolle adecuadamente, el snRNA de las partículas ribonucleoproteicas debe reconocer las secuencias consenso del pre-mRNA, las cuales marcan las uniones exón-intrón. En metazoos, es necesario que intervengan también las proteínas auxiliares, las cuales ayudan a estabilizar a las snRNP (Ast, 2004).

Actualmente, no está claro cómo se lleva a cabo el ensamblaje del espliceosoma mayor, aunque diversas investigaciones realizadas tanto *in vitro* como *in vivo* apoyan la idea de que se trata de un proceso secuencial y ordenado, donde se pueden diferenciar varias etapas determinadas por la formación de complejos característicos (Hoskins et al., 2011; Huranová et al., 2010; Montes et al., 2012).

El ensamblaje del espliceosoma mayor (**Figura I-4**) comienza con la unión, por complementariedad de bases, de U1 snRNP al sitio 5' de *splicing* (*5' splice site*, *5'SS*) del intrón. En eucariontes superiores, esta unión RNA-RNA es débil, pero consigue ser

estabilizada con proteínas de la propia U1 snRNP y la presencia de miembros de la familia de proteínas ricas en serina y arginina (SR) (Wahl et al., 2009). Después de esto, se produce la unión del factor de *splicing* SF1/BBP a la secuencia intrónica conocida como BPS (*Branch-Point Sequence*) localizada cerca del extremo 3' del intrón. Al mismo tiempo, también tiene lugar la unión de U2AF65 al tracto de polipirimidinas (Py) y la unión de U2AF35 a la secuencia AG del sitio 3' de *splicing* (*3' splice site, 3'SS*). Juntas, todas estas interacciones tempranas conforman el complejo E, el cuál es esencial para el reconocimiento inicial de los sitios 5'SS y 3'SS de un intrón (Wahl et al., 2009).

A continuación, U2 snRNP se va a unir, por complementariedad de bases, al BPS del pre-mRNA, desplazando del mismo al factor SF1-BBP en una reacción dependiente de ATP que da lugar al complejo A o pre-espliceosoma. Esta unión RNA-RNA es estabilizada por proteínas de la propia U2 snRNP así como por U2AF65 (Wahl et al., 2009).

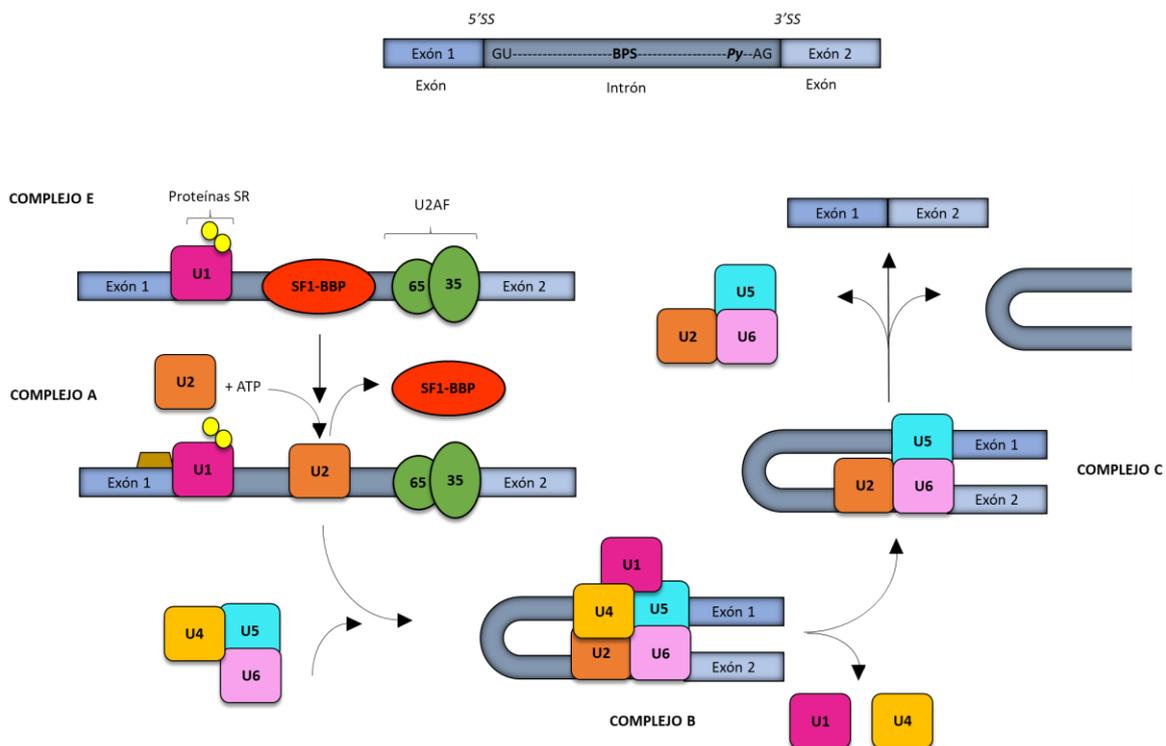


Figura I-4. Ensamblaje del espliceosoma mayor (modelo de definición de intrones) y reacción de *splicing*. Esquema de la formación de cada complejo (E, A, B, C) durante las etapas de ensamblaje del espliceosoma. Se representan las subunidades snRNPs U1, U2, U4, U5 y U6; así como los factores de *splicing* SF1 y U2AF. 3'SS, sitio de *splicing* 3' (AG); 5'SS, sitio de *splicing* 5' (GU); BPS, *branch point site*; Py, tracto de polipirimidinas.

Seguidamente, son reclutadas U4/U6 y U5 snRNP que se unen para formar el complejo B (Wahl et al., 2009). La U6 snRNP está formada por el U6 snRNA y el complejo LSM2-8 (Tharun, 2009). De entre las proteínas que forman U5, es destacable PRPF8, ya que forma parte del sitio activo del espliceosoma (Nguyen et al., 2013). Aunque todas las snRNPs están ya presentes en el complejo B, éste es catalíticamente inactivo y requiere de una serie de reordenamientos conformacionales para ser activado. Durante la activación del espliceosoma, U1 y U4 snRNP son liberados dando lugar finalmente al complejo C o complejo catalítico, por ser el responsable de llevar a cabo el proceso de *splicing* propiamente dicho, mediante una reacción de transesterificación dependiente de ATP. Finalmente, el espliceosoma libera el mRNA así como las U2, U5 y U6 snRNP, que son recicladas para posteriores ciclos de *splicing* (Wahl et al., 2009).

Espliceosoma menor.

El espliceosoma menor está formado por la U5 snRNP, y las ribonucleoproteínas específicas U11 (análogo funcional de U1), U12 (U2), U4atac (U4) y U6tac (U6). Los análogos funcionales presentan estructuras secundarias similares a las de su equivalente en el espliceosoma mayor. También, el ensamblaje del espliceosoma menor ocurre de forma parecida al del mayor, siendo la diferencia principal la ausencia del complejo E (Turunen et al., 2013).

El espliceosoma menor se encarga de procesar un tipo de intrones poco frecuentes conocidos como intrones U12. Estos intrones representan el 0,5% de todos los intrones de cualquier genoma y se diferencian del resto de intrones por las secuencias consenso de los sitios de *splicing* (AU y AC para los 3'SS y 5'SS, respectivamente) (Turunen et al., 2013).

1.2.2 Definición del exón y del intrón.

Numerosos estudios también han demostrado que la longitud de los exones e intrones afecta al reconocimiento de los sitios de *splicing*, y por tanto, a la eficiencia de este proceso (De Conti et al., 2013).

En vertebrados, los exones (con una media de 170 nucleótidos) son mucho más cortos que los intrones (miles de nucleótidos). En estos organismos, predomina el

mecanismo por el cual distintos componentes de la maquinaria de *splicing* se ensamblan en los sitios 5'SS y 3'SS de un exón. Este mecanismo se conoce como definición del exón (*exon definition*) (**Figura I-5B**). Por el contrario, en el resto de los eucariontes, la arquitectura del genoma se caracteriza por tener pequeños intrones y grandes exones. En este caso, el modelo de definición del intrón (*intron definition*) es el predominante, en el cual, los sitios de *splicing* se emparejan entre intrones, en lugar de exones (**Figura I-4 y I-5A**) (De Conti et al., 2013).

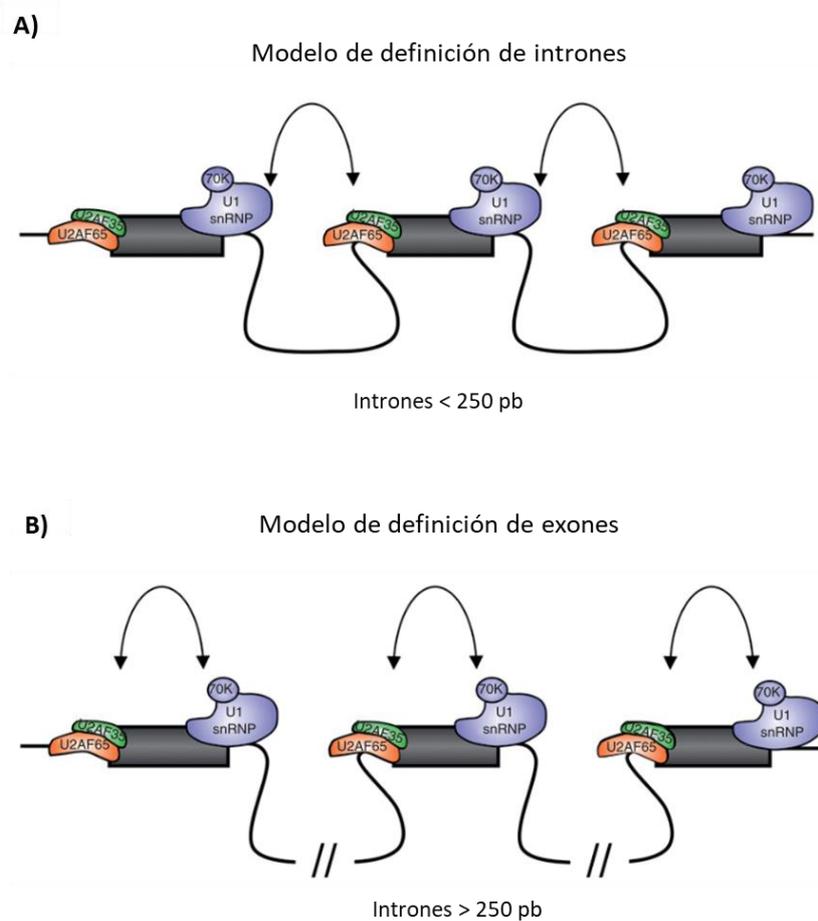


Figura I-5: Modelo de definición de intrones y exones. A) Modelo de definición de intrones acorde al emparejamiento entre los sitios de *splicing* a lo largo de un intrón, cuando los exones están separados por intrones cortos (< 250 pb). **B)** Modelo de definición de exones donde la comunicación entre los sitios de *splicing* ocurre a lo largo de un exón, cuando estos están separados por intrones largos (> 250 pb). Imagen adaptada de De Conti et al., 2013.

1.2.3 El *splicing* alternativo.

Según la eficiencia con la que se ensambla el espliceosoma y ocurre el *splicing*, se distinguen dos tipos: *splicing* constitutivo y *splicing* alternativo.

El *splicing* constitutivo es aquel que se produce cuando tanto el reconocimiento de un exón como el ensamblaje del espliceosoma se producen de forma correcta y del mismo modo en todos los pre-mRNAs de un determinado gen, dando lugar a una única molécula de mRNA.

Por el contrario, el *splicing* alternativo es aquel en el que los sitios 5'SS y 3'SS compiten con otros sitios 5'SS y 3'SS por la unión de los componentes del espliceosoma. Consecuentemente, el exón no será procesado del mismo modo en todos los pre-mRNAs de un gen, por lo que se originarán diferentes mRNAs a partir de un solo pre-mRNA. Estos mRNAs dará lugar a proteínas diferentes, con funciones que pueden ser incluso opuestas, según el tejido o el tipo de célula. Por ello, el *splicing* alternativo contribuye a la diversidad genómica y a la especificidad de tejidos (Chen and Manley, 2009).

Aunque lo normal sea que durante las primeras etapas del ensamblaje del espliceosoma se determine qué tipo de *splicing* tendrá lugar, estudios recientes han demostrado numerosos ejemplos en los que esta decisión se toma durante las etapas tardías de dicho proceso, e incluso durante la reacción de transesterificación (Chen and Manley, 2009).

1.2.4. Tipos de *splicing* alternativos.

Diversos análisis de secuenciación masiva estiman que más del 90% de los genes humanos son regulados por *splicing* alternativo. La existencia de este tipo de *splicing* implica el uso de 3'SS y 5'SS distintos a los convencionales. Por tanto, en un pre-mRNA con varios exones, el *splicing* alternativo puede darse de diversas formas (Black, 2003). Los diferentes tipos de *splicing* alternativo que existen son (Black, 2003):

- a. Exclusión de un exón: representa casi el 40% de los eventos de *splicing* alternativo en eucariontes superiores. Sin embargo, es bastante raro en eucariontes inferiores (**Figura I-6A**).
- b. Sitio donador alternativo: se produce cuando hay un sitio 3' de *splicing* alternativo, dando lugar a un acortamiento del exón. Representa el 18,4% de los eventos de *splicing* alternativo en eucariontes superiores (**Figura I-6B**).

- c. Sitio aceptor alternativo de *splicing*: se da cuando hay un sitio 5' de *splicing* alternativo, produciéndose un exón más largo. Representa casi el 7,9% de los eventos de *splicing* alternativo en eucariontes superiores (**Figura I-6C**).
- d. Retención de intrón: tiene lugar cuando un intrón es incluido en el transcrito final. Es un evento poco frecuente en vertebrados e invertebrados, pues ocurre con una frecuencia de 5% (**Figura I-6D**).
- e. Exones mutuamente excluyentes: genera transcritos en los que varios exones no coinciden en el mRNA maduro. Es poco frecuente (**Figura I-6E**).
- f. Promotor alternativo: tiene lugar cuando el transcrito puede empezar en exones diferentes. Es poco frecuente (**Figura I-6F**).
- g. polyA alternativa: se produce cuando el mRNA puede terminar en diferentes exones. Es poco frecuente (**Figura I-6G**).

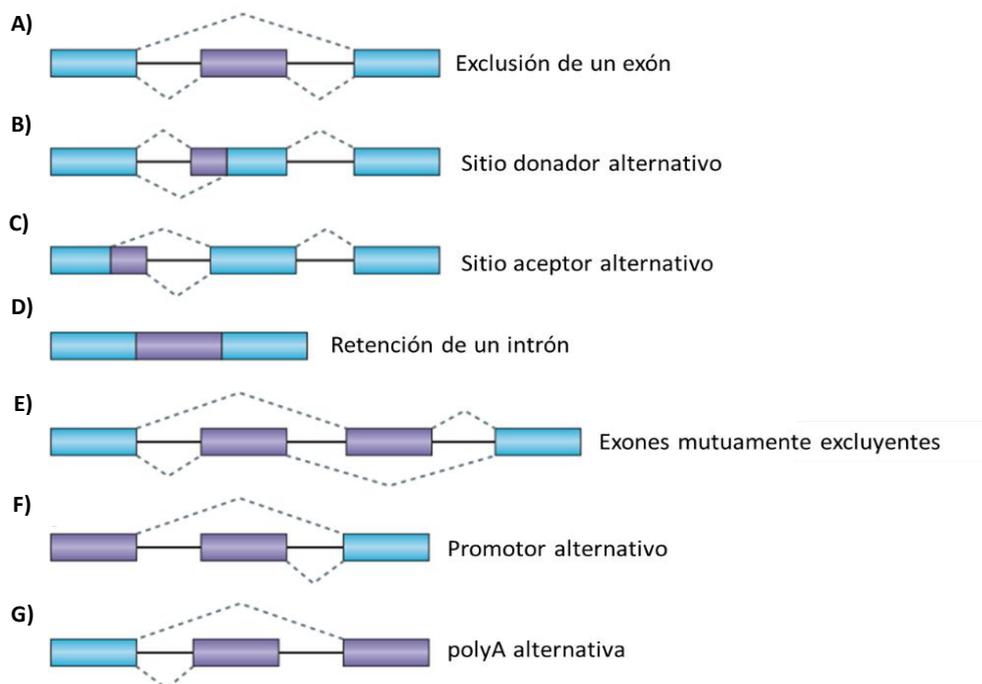


Figura I-6. Tipos de *splicing* alternativo. Los exones constitutivos son mostrados en azul y las regiones procesadas de forma alternativa en morado. Los intrones se representan con líneas continuas y las diferentes opciones de *splicing* con líneas discontinuas. Imagen adaptada de Black, 2003.

1.3. Acoplamiento de la transcripción y el *splicing*.

Durante las dos últimas décadas se ha demostrado que muchos eventos de *splicing* ocurren de manera cotranscripcional, es decir, a la vez que se produce la

transcripción del RNA, éste va siendo procesado por la maquinaria de *splicing* para dar lugar al transcrito maduro (Montes et al., 2012; Saldi et al., 2016). Un estudio reciente en levaduras ha demostrado que el *splicing* se produce cuando la RNAPII ha transcrito aproximadamente entre 26 y 129 nt aguas abajo del sitio 3'SS (Carrillo Oesterreich et al., 2016).

Los cambios en los patrones de fosforilación del CTD de la RNAPII juegan un papel esencial en el acoplamiento espacial de estos dos procesos, pues a través de ellos, en este dominio se van a reclutar o liberar factores de procesamiento (entre ellos, factores de *splicing*) de forma específica en cada etapa de transcripción (Buratowski, 2009; Egloff and Murphy, 2008). Por ejemplo, se ha comprobado que el CTD favorece la activación del *splicing* a través del reclutamiento del complejo U2AF65-Prp19 (David et al., 2011).

El reclutamiento al CTD de otros factores conocidos como acopladores de la transcripción y el *splicing*, podría modular la velocidad de elongación de la RNAPII a lo largo de un gen, o en una zona específica (Muñoz et al., 2010). Como se comentó en el apartado 1.1.3, la tasa de elongación de la RNAPII va a afectar a la transcripción, y por tanto, a los procesos cotranscripcionales como el *splicing* (Jonkers and Lis, 2015). Además, los cambios en la velocidad de elongación de la RNAPII también son capaces de regular el proceso de *splicing* alternativo cotranscripcional, puesto que va a determinar el periodo disponible para que el espliceosoma se ensamble en un sitio de *splicing* y elimine un intrón, antes de competir con el que se generará aguas abajo (Perales and Bentley, 2009).

El hecho de que el procesamiento de los transcritos inmaduros se produzca de manera cotranscripcional aporta numerosas ventajas: i) ayuda a prevenir que el pre-mRNA se degrade; ii) facilita el rápido reclutamiento de factores de *splicing* a la maquinaria transcripcional mientras está transcribiendo, de manera que el pre-mRNA pueda ser procesado inmediatamente después de haber sido transcrito, optimizando así la eficiencia de *splicing*; iii) y previene la hibridación entre el pre-mRNA y la hebra molde del DNA, un proceso que podría favorecer la inestabilidad genómica debido a la acumulación de rupturas en el DNA causadas por dicho estado de hibridación (Muñoz et al., 2010).

Por todo ello, los datos disponibles hasta el momento sugieren que la ventaja evolutiva de la cotranscripcionalidad reside en permitir el acoplamiento entre transcripción y *splicing* (Muñoz et al., 2010).

2. LA PREFOLDINA.

El correcto plegamiento de las proteínas es un proceso fundamental para permitir la función de la mayoría de los componentes celulares, y dado que la conexión entre genes y fenotipo no ocurre si la transcripción y traducción no producen proteínas funcionales, las chaperonas podrían ser consideradas como la última parte de la maquinaria de expresión génica.

Las chaperonas son un conjunto de factores presentes en todas las células, que actúan sobre proteínas no plegadas o parcialmente plegadas, impiden su agregación y modifican su estructura terciaria y/o cuaternaria. Normalmente, las chaperonas suelen estar ayudadas por una serie de cofactores conocidos como cochaperonas (Hartl et al., 2011). Ambas juegan un papel muy importante en el mantenimiento de la homeostasis de proteínas o proteostasis, es decir, en el equilibrio entre síntesis, plegamiento, ensamblaje y degradación de proteínas en respuesta a condiciones de estrés y medioambientales (Hartl et al., 2011).

La prefoldina (PFD) es una cochaperona que fue descrita hace dos décadas de forma independiente por Elman Schiebel en *Saccharomyces cerevisiae* y por Nicholas Cowan en células de mamíferos, quienes vieron que este nuevo complejo era capaz de capturar monómeros de actina y tubulina no plegados, y favorecer su ensamblaje para dar lugar a polímeros del citoesqueleto (Geissler et al., 1998; Vainberg et al., 1998).

Prefoldina, también conocida como GimC, no se encuentra presente en eubacterias, pero sí en arqueas, lo cual sugiere que éste podría ser el origen de su presencia en todos los organismos eucarióticos, desde protistas unicelulares hasta células humanas. Se trata de un complejo heterohexamérico que en arqueas está compuesto por dos subunidades α idénticas y cuatro subunidades β también iguales, y, sin embargo, en organismos eucarióticos el complejo canónico lo forman dos subunidades α y cuatro subunidades β distintas (Millán-Zambrano and Chávez, 2014).

Hasta el momento, ninguna de las subunidades de este complejo ha resultado esencial para los organismos estudiados, aunque sí que ha variado la gravedad de los fenotipos observados entre ellos. Por ejemplo, la deficiencia de prefoldina en *S.cerevisiae* empeoró el ensamblaje de filamentos de actina y tubulina. Por el contrario, en *Caenorhabditis elegans* y *Drosophila*, la deficiencia de prefoldina resultó en una letalidad embrionaria provocada por un citoesqueleto anormal. También los mutantes de *Arabidopsis* exhiben alteraciones del desarrollo relacionadas con una organización y dinámica de microtúbulos anormales (Payán-Bravo et al., 2018). La mutación de prefoldina en ratones (los cuales resultan viables hasta el nacimiento) también mostró fenotipos causados por defectos en el citoesqueleto, incluyendo degeneración de fotorreceptores, infertilidad masculina, disquinesia ciliar, pérdida neuronal y defectos en el desarrollo y función de las células B y T tubulina. Además de esto, la deficiencia de prefoldina también provocó acumulación de proteínas agregadas insolubles, lo cual se relaciona con el papel de prefoldina contra la agregación de proteínas (Chávez and Puerto-Camacho, 2016).

2.1. Tipos de prefoldinas y su estructura.

A lo largo de la evolución, han ido multiplicándose las subunidades de prefoldina en el genoma de los organismos eucarióticos, lo cual ha favorecido la diversificación de dichas subunidades y la aparición de complejos alternativos con funciones específicas (Chávez and Puerto-Camacho, 2016). A continuación, se describen los tres tipos de complejos de prefoldinas que existen.

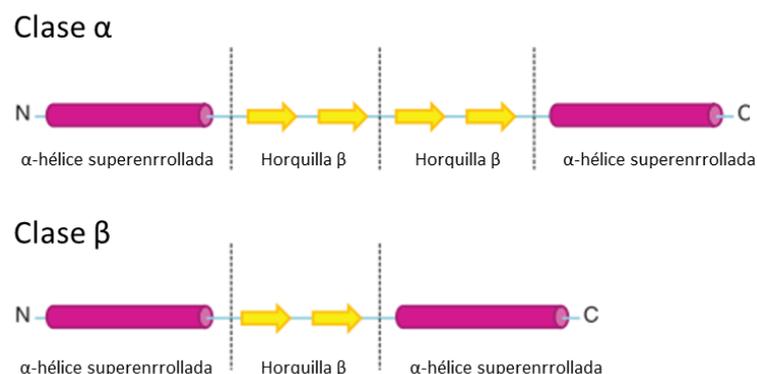


Figura I-7. Estructura secundaria de las subunidades α y β de prefoldina. El número de horquillas β determina la clase de subunidad: en las subunidades α hay dos, mientras que en las subunidades β hay solo una. Las horquillas β se localizan, en ambos casos, entre las α -hélices superenrolladas. Modificado de Chávez and Puerto-Camacho, 2016.

Complejo canónico de prefoldina.

Se trata de un complejo de 87 kDa que fue caracterizado por primera vez en *Methanobacterium thermoautotrophicum*. Todas sus subunidades presentan dos α -hélices que forman estructuras superenrolladas (coiled-coil) conectadas entre sí por una o dos horquillas β según si se tratan de subunidades β o subunidades α , respectivamente (**Figura I-7**) (Siegert et al., 2000).

En arqueas, la estructura del complejo recuerda a los tentáculos de una medusa, donde el cuerpo lo formarían dobles barriles β y los tentáculos 6 estructuras superenrolladas, lo que es consistente con el ensamblaje de cuatro subunidades β , dentro de las cuales van a dimerizar dos subunidades α (**Figura I-8A**). En eucariontes, los datos aportados por experimentos de microscopía electrónica de baja resolución sugieren que la PFD humana presenta una estructura similar a la de arqueas, donde las dos subunidades α (PFDN3 y PFDN5) ocuparían el centro del complejo, y rodeándolas se encontrarían las subunidades β (PFDN1, PFDN2, PFDN4 y PFDN6) (**Figura I-8B y I-8C**) (Arranz et al., 2018).

Las interacciones entre PFDN2 y PFDN3 y entre PFDN5 y PFDN6 son más estables que aquellas entre subunidades adyacentes (Simons et al., 2004) y este hecho confiere estabilidad al complejo (Chávez and Puerto-Camacho, 2016). Además, existe una regulación de los niveles de proteína de las distintas subunidades entre sí, lo cual ha sido demostrado *in vivo* mediante experimentos de deficiencia y sobreexpresión de subunidades, donde, por ejemplo, la deficiencia en una subunidad causa una bajada en la cantidad de las otras subunidades (Miyazawa et al., 2011).

Este canónico heterohexámero contribuye al proceso de plegamiento estabilizando a un amplio rango de proteínas no nativas a las que posteriormente va a liberar para su correcto plegamiento asistido por chaperoninas (Arranz et al., 2018). En arqueas las puntas de las estructuras superenrolladas de PFD α y PFD β interactúan con las proteínas no nativas e impiden su agregación (Okochi et al., 2004). En dicha interacción, van a intervenir al menos una pareja de PFD β adyacente y según el tamaño del sustrato, podrían ir acoplándose otras subunidades. En eucariontes, el mecanismo es diferente e involucra interacciones específicas con actina y tubulina basadas en la

carga de los residuos localizados en la región distal de las estructuras superenrolladas: mientras que las subunidades PFDN3 y PFDN4 son importantes para la interacción con actina, PFDN4 y PFDN5 lo son para la interacción con tubulina. Otra diferencia con respecto a arqueas, es que la prefoldina de eucariontes envuelve por completo a sus sustratos dentro de la cavidad (Arranz et al., 2018).

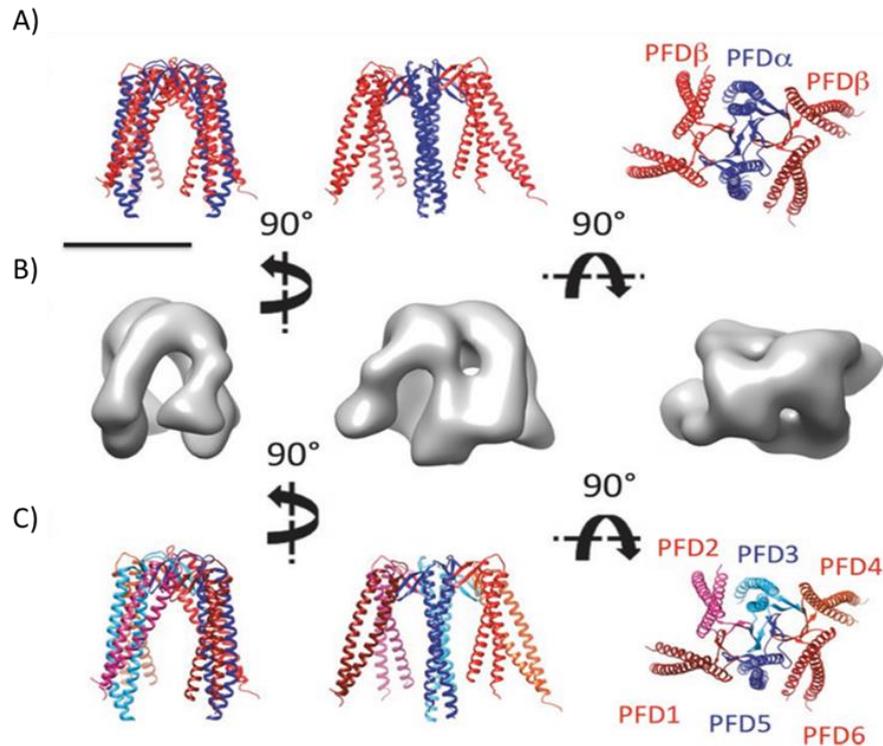


Figura I-8. Estructura de prefoldina en arqueas y humanos. (A) Diferentes vistas ortogonales de la estructura cristalográfica de prefoldina de *M. thermoautotrophicum*. Dos copias de PFD α (azul) ocupan la posición central y se encargan del mantenimiento del complejo, mientras que cuatro copias de PFD β (rojo) rodean a las PFD α . **(B)** Estructura a baja resolución de la prefoldina humana, donde se muestran las 3 mismas vistas ortogonales obtenidas por microscopía electrónica y procesamiento de imágenes. **(C)** Modelo de la prefoldina de organismos eucarióticos donde se muestran las mismas 3 vistas generadas usando la estructura de *M. thermoautotrophicum* con la secuencia de PFD humana y la disposición de las subunidades obtenida (Simons et al., 2004). Barra. = 50 Å. Imagen de Arranz, Martín-Benito and Valpuesta, 2018.

Complejo alternativo o *prefoldin-like complex*.

En eucariontes se ha visto que algunas de las subunidades β de prefoldina pueden participar en otros complejos estables en colaboración con otros polipéptidos diferentes, lo cual se conoce como complejos alternativos. En levaduras, solo la subunidad 6 forma parte de dichos complejos, mientras que en mamíferos podemos encontrar las subunidades 2 y 6. Otros de los componentes habituales de los complejos

alternativos en humanos son URI (polipéptido de tipo α), PDRG1, UXT, RPB5 y WDR92 (Payán-Bravo et al., 2018). Aunque la estructura del complejo alternativo no ha sido aún caracterizada, la presencia de 6 polipéptidos (de tipo α y β) sugiere que dicho complejo podría presentar una estructura parecida a la del complejo canónico (Chávez and Puerto-Camacho, 2016).

Prefoldina gamma.

Methanocaldococcus jannaschii es una arquea hipertermófila que presenta un tercer gen (γ) de prefoldina, el cual se expresa en respuesta al calor (Whitehead et al., 2007). Los monómeros proteicos que codifica este gen se ensamblan y dan lugar a largos filamentos que actúan como chaperonas *in vitro*. El ensamblaje de entre 6 y 8 monómeros de γ -prefoldina es esencial para que los filamentos tengan actividad funcional y prevengan la agregación de sus sustratos en respuesta a las altas temperaturas (Glover and Clark, 2015; Whitehead et al., 2007).

2.2. Funciones citoplasmáticas.

La prefoldina es un invento evolutivo de las arqueas (Chávez and Puerto-Camacho, 2016). En estos organismos, es capaz de reconocer a la gran mayoría de proteínas no plegadas y plegarlas. Sin embargo, en eucariontes este complejo heterohexamérico presenta un grado mayor de especificidad por sus sustratos desplegados, a los cuales también ayudará a plegarse (Chávez and Puerto-Camacho, 2016). Tanto en arqueas como en eucariontes, este papel suele ser desempeñado *in vivo* e *in vitro* con ayuda de chaperoninas de clase II dependiente de ATP (Chávez and Puerto-Camacho, 2016). Uno de los ejemplos mejor conocido en eucariontes es el de los monómeros de actina: después de la síntesis de los primeros 145 aminoácidos, el polipéptido naciente se asocia a la chaperonina CCT a la que permanece unida hasta que finaliza su traducción al completo (Hansen et al., 1999). De hecho, en eucariontes se ha propuesto que el sistema prefoldina/CCT podría hacer las veces de un compartimento funcional donde mantener de forma temporal a los polipéptidos recién sintetizados que son sensibles a la agregación, lo cual ayudaría a estabilizarlos y plegarlos de forma eficiente más tarde (Siegers et al., 1999). Además de esto, el sistema prefoldina/CCT también ayuda al ensamblaje de estructuras del citoesqueleto, función que está muy bien conservada en

todos los organismos eucariontes, tal y como ha sido demostrado al rescatar el fenotipo causado por la delección de genes de prefoldina en levaduras con ensayos de complementación en los que fueron utilizados genes de prefoldina de plantas y humanos (Payán-Bravo et al., 2018).

A parte de estas funciones, el complejo canónico también desarrolla un importante papel en contrarrestar la agregación de proteínas citoplasmáticas que nada tienen que ver con el citoesqueleto. Esta función es clave principalmente en condiciones de mal funcionamiento del proteasoma o de estrés en retículo endoplasmático, donde se ha demostrado que prefoldina colocaliza con proteínas poliubiquitinadas en células humanas (Abe et al., 2013).

Por su parte, el complejo alternativo contribuye al ensamblaje de otras proteínas no relacionadas con el citoesqueleto, como, por ejemplo, la RNAPII. Concretamente, URI interactúa con la subunidad Rpb5 de la RNA polimerasa I, II y III (Gstaiger et al., 2003). Esta función citoplasmática constituye un primer enlace de prefoldina con el proceso de expresión génica (Payán-Bravo et al., 2018).

Así como el complejo canónico colabora con chaperoninas de clase II, el complejo alternativo lo hace con otras chaperonas como HSP90 y R2TP. Esta última, junto al complejo alternativo, está implicada en el ensamblaje y la estabilización de proteínas relacionadas con la quinasa de fosfatidilinositol-3 en el citoplasma, incluida mTOR. Dado que mTOR regula la expresión génica en respuesta a la disponibilidad de nutrientes, esto supondría un segundo enlace indirecto de prefoldina con la regulación génica (Payán-Bravo et al., 2018).

2.3. Funciones nucleares.

Además de las funciones citoplasmáticas, tanto el complejo canónico como el complejo alternativo realizan diversas funciones en el núcleo de las células, lo cual ha sido demostrado en levaduras, plantas y metazoos (Millán-Zambrano and Chávez, 2014).

Un claro ejemplo de la función nuclear de prefoldina es su participación en la degradación de la integrasa del Virus de Inmunodeficiencia Humana (VIH). La integrasa se degrada en el núcleo mediante poliubiquitinación catalizada por VHL. La subunidad PFDN3 media la interacción entre la integrasa y VHL. La degradación de la integrasa es necesaria para la correcta expresión del virus. Las subunidades PFDN1, PFDN2 y PFDN6 también influyen en la expresión del VIH, lo cual sugiere que es el complejo canónico al completo el que actúa en dicha función (Mousnier et al., 2007).

En el caso del complejo alternativo, se ha descrito recientemente que la subunidad PDRG1 interactúa con enzimas que catalizan la síntesis de S-adenosil metionina, una coenzima necesaria para la metilación del DNA e histonas. Además, se ha determinado que el incremento de PDRG1 en ciertos tumores podría contribuir a la reprogramación epigenética y transformación celular (Pajares, 2017; Pérez et al., 2016). Otro ejemplo sería el protagonizado por la subunidad Bud27 (ortólogo de URI en levaduras), que es capaz de modular la actividad del complejo remodelador de nucleosomas *RSC* (*chromatin structure remodelling*), e influir indirectamente en el proceso de transcripción (Mirón-García et al., 2014).

Tabla I-1: Factores relacionados con la expresión de genes que física y funcionalmente interactúan con prefoldina. Modificado de Payán-Bravo, Peñate and Chávez, 2018.

FUNCIÓN	COMPLEJO CANÓNICO	COMPLEJO ALTERNATIVO
Interacción con factores de cromatina	HDAC1-mSin3	DMAP1-HDAC2-RSC
Interacción con promotores de genes	DELLA c-Myc Egr1 HBx	
Interacción con factores de transcripción	Gen de la ciclina A Genes de la espermatogénesis Gen Wnt4	EVI, NF-Kb Receptor de andrógenos VP16 HBX KAP1
Desalojo de histonas	Genes transcritos	
Localización		RNAPII-Thr4P Receptor aril hidrocarburo Receptor de estrógenos
Proteólisis	c-Myc	
Ensamblaje		RNAP RSC mTOR

A parte de estas funciones nucleares, tanto el complejo canónico como el alternativo interactúan física y funcionalmente con factores de transcripción (**Tabla I-1**) (Payán-Bravo et al., 2018).

Uno de los ejemplos mejor conocidos es el de la subunidad PFDN5, también conocida como MM-1, que actúa como correpressor de la actividad de transactivación de *c-Myc*, un proto-oncogén que regula el crecimiento celular, la apoptosis, el metabolismo y la diferenciación (a nivel celular, tisular o de organismo) mediante la regulación de numerosos genes dianas (Nie et al., 2012). Las isoformas nucleares PFDN5 α y PFDN5 γ se unen al dominio de transactivación de *c-Myc* localizado en el extremo amino terminal, y reprimen su actividad transcripcional (Hagio et al., 2006). La represión se produce mediante el reclutamiento de una desacetilasa de histonas que inhibe la remodelación de la cromatina necesaria para el inicio de la transcripción (**Figura I-9A**) (Satou et al., 2001). Se ha visto que en pacientes con leucemia o linfoma suelen aparecer mutaciones puntuales en PFDN5, lo cual sugiere que PFDN5 podría actuar como un supresor tumoral (Fujioka et al., 2001). Además de esto, PFDN5 es capaz de regular negativamente la acción de *c-Myc* de otras dos formas: una de ellas es favoreciendo su degradación reclutando a un complejo ligasa de ubiquitinas, y dirigiendo a *c-Myc* al proteasoma (**Figura I-9B**) (Kimura et al., 2007); la otra es mediante su propia mono-ubiquitinación a través de la proteína Rabring7, lo cual estimularía el papel de PFDN5 en el control de *c-Myc* (**Figura I-9B**) (Narita et al., 2012). También se ha descrito que PFDN5 puede inhibir indirectamente la expresión del gen *c-Myc*. La inhibición se produce, junto al represor Egr-1, reprimiendo la expresión de Wnt4, una proteína que regula positivamente la expresión de *c-Myc* (**Figura I-9C**) (Yoshida et al., 2008).

A parte de PFDN5, también se ha descrito que otras subunidades del complejo canónico ejercen una función represora en la transcripción. Por ejemplo, el incremento de los niveles de PFDN1 en células humanas tumorales de pulmón provoca que esta subunidad se localice en el núcleo, donde va a unirse al promotor del gen de la ciclina A, reprimiendo su expresión (Wang et al., 2017). Esta represión va a favorecer la transición epitelio-mesénquima, el crecimiento del tumor y la metástasis (Wang et al., 2017). Además, se ha visto la implicación de PFDN1 en el carcinoma de colon, donde induce la

proliferación celular y la motilidad, aunque se desconoce el mecanismo por el que lo hace (Wang et al., 2015). Estos ejemplos ponen de manifiesto que también existe conexión entre prefoldina y los fenómenos de tumorigénesis y metástasis.

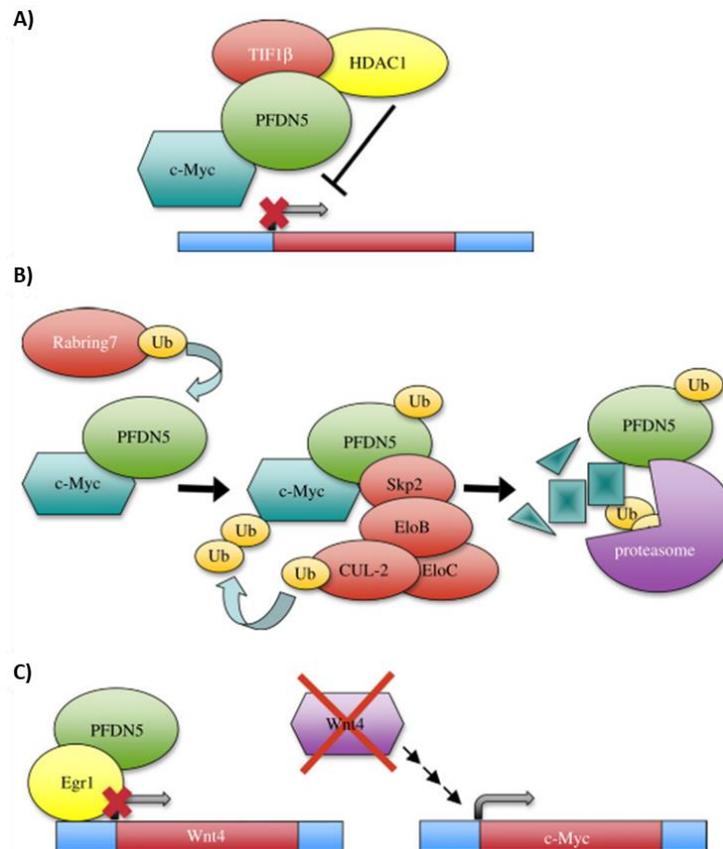


Figura I-9: La subunidad PFDN5 humana está involucrada en tres mecanismos de control diferentes de c-Myc. A) PFDN5 se une al extremo N-terminal de c-Myc y reprime su actividad transcripcional reclutando al correpresor TIF1 β y al complejo histona desacetilasa HDAC1-mSin3 (Hagio et al., 2006; Satou et al., 2001). **B)** PFDN5 conduce a c-Myc a degradarse en el proteasoma y para ello, recluta al complejo ligasa de ubiquitinas Skp2-ElonginC-ElonginB-Cullin2 (Kimura et al., 2007). Además, la mono-ubiquitinación de PFDN5 inducida por Rabring7, favorece dicha actividad (Narita et al., 2012). **C)** PFDN5 y Egr1 cooperan juntos en la represión transcripcional de Wnt4, un regulador positivo de la expresión del gen *c-Myc*. (Yoshida et al., 2008). Imagen de Millán-Zambrano and Chávez, 2014.

Prefoldina no solo ejerce un papel transcripcional en células humanas, sino que también se han descrito dichas funciones tanto en levaduras (ver apartado 3) como en plantas. En estas últimas, se ha demostrado que PFDN5 y PFDN6 intervienen en la regulación génica mediante la interacción física con el factor nuclear DELLA. Cuando las condiciones medioambientales son desfavorables para el crecimiento, DELLA retiene a prefoldina en el núcleo de la célula, lo cual compromete el papel de prefoldina en el citoplasma, previniendo de este modo la formación de microtúbulos y el crecimiento

celular (Locascio et al., 2013) . Sin embargo, cuando las condiciones medioambientales son favorables para el crecimiento, DELLA se degrada y prefoldina puede regresar al citoplasma (Locascio et al., 2013).

Todos estos ejemplos, y muchos más que no han sido mencionados en esta tesis, apoyan que tanto el complejo canónico como el alternativo ejercen un papel importante en la expresión génica, lo cual supera la visión inicial de la prefoldina eucariótica como factor de ensamblaje del citoesqueleto.

3. ANTECEDENTES.

En nuestro laboratorio se ha descrito que las prefoldinas Pfd1, Pfd2, Pfd5 y Pfd6 de levaduras son reclutadas juntas a los genes que se están transcribiendo (Millán-Zambrano et al., 2013). El perfil de reclutamiento que muestra una de estas subunidades (Pfd1) es paralelo al perfil de RNAPII fosforilada en la Ser2 de su CTD, por lo que parece que el reclutamiento es dependiente de transcripción (**Figura I-10**) (Millán-Zambrano et al., 2013). Además, se comprobó que Pfd1 contribuye a la elongación de la transcripción de los genes largos (Millán-Zambrano et al., 2013). La ausencia de Pfd1 provoca, además, un incremento en la densidad de histonas que permanece unida a los genes transcritos intensamente (Millán-Zambrano et al., 2013). Sin embargo, ensayos con nucleasa micrococcal demuestran que la ausencia de prefoldina no afecta al remodelamiento de los nucleosomas (Millán-Zambrano et al., 2013). Estos datos sugieren que la prefoldina participa en el desalojo de histonas después de la desestabilización cotranscripcional de los nucleosomas (Millán-Zambrano et al., 2013).

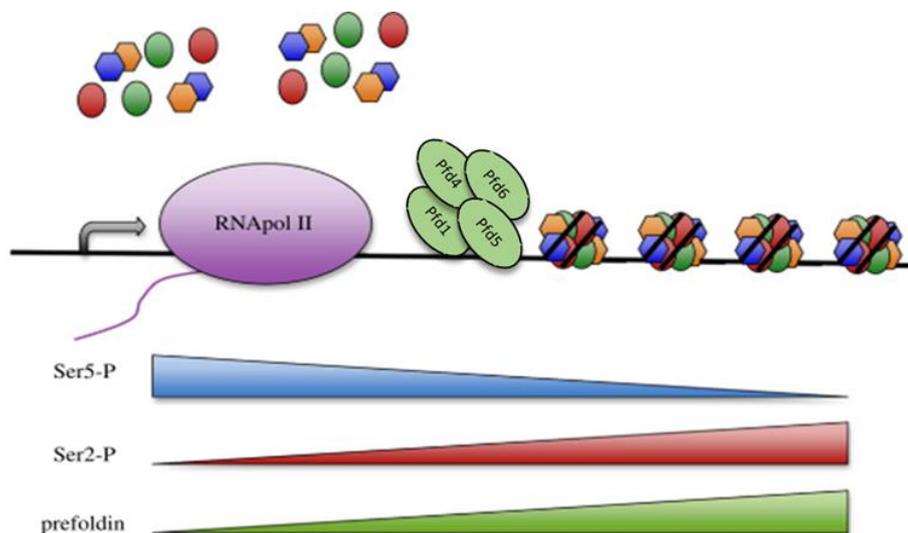


Figura I-10: En levaduras, la prefoldina estimula la elongación de la transcripción favoreciendo la dinámica de la cromatina. Las subunidades Pfd1, Pfd4, Pfd5 y Pfd6 se unen a genes que se están transcribiendo. Se ha visto que Pfd1 contribuye a la elongación de la transcripción. Este efecto se observa muy bien en genes largos, donde además Pfd1 estimula el desalajo de histonas. La unión de la prefoldina a los genes se localiza en el extremo 3' de estos, lo cual correlaciona con el perfil de la RNAPII fosforilada en Ser-2. Imagen adaptada de Millán-Zambrano and Chávez, 2014.

Más tarde, se realizaron experimentos de ChIP para analizar los niveles de RNAPII Ser2-P en mutantes de levaduras carentes de las diferentes subunidades del complejo canónico de la prefoldina (datos no publicados). En ellos se vio que la ausencia de cada una de las prefoldinas provocó un fuerte descenso de los niveles de esta marca de la polimerasa que está elongando (**Figura I-11**). Este resultado iba en concordancia con los defectos transcripcionales causados por la ausencia de prefoldina que fueron mencionados anteriormente.

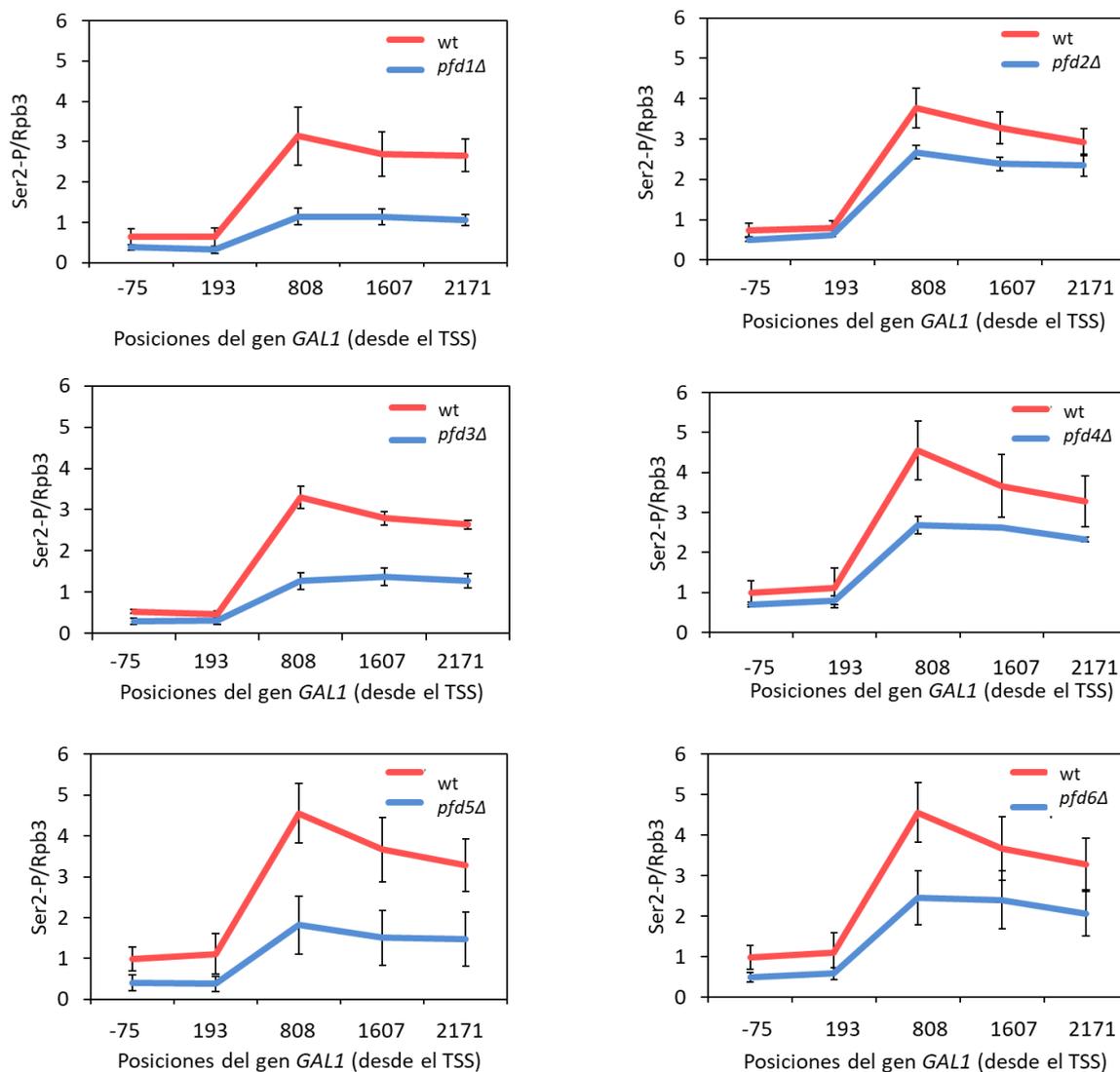


Figura I-11: La ausencia de las diferentes subunidades del complejo canónico de prefoldina los niveles de RNAPII Ser2-P. Se realizaron ensayos de CHIP a partir de extractos obtenidos de 6 mutantes diferentes de levaduras correspondientes a las 6 subunidades canónicas de prefoldina, y se analizó los niveles de Ser2-P relativizados a los niveles de RNAPII total (Rpb3). En todas las gráficas se representa valores promedios obtenidos de, al menos, 3 experimentos

Este conjunto de datos muestra que la prefoldina de levaduras ejerce un papel positivo en la elongación transcripcional y favorece la dinámica de la cromatina de forma cotranscripcional.

Objetivos

En este trabajo, nuestro propósito ha sido estudiar la contribución de la prefoldina a la expresión génica en el núcleo de las células humanas. Para lograr esto, nos planteamos los siguientes objetivos:

1. Estudiar los efectos globales que provoca la deficiencia de las subunidades PFDN2 y PFDN5 en el transcriptoma de células humanas y su regulación por suero.
2. Estudiar los efectos globales que provoca la deficiencia de PFDN2 y PFDN5 en el procesamiento del pre-mRNA de células humanas.
3. Explorar posibles mecanismos que expliquen los resultados de los objetivos anteriores.

Resultados

1. PFDN5 SE LOCALIZA MAYORITARIAMENTE EN EL NÚCLEO DE CÉLULAS HCT116.

La subunidad 5 del complejo prefoldina presenta 4 isoformas (α , β , γ y δ) como consecuencia del *splicing* alternativo de su transcrito (Hagio et al., 2006). Clásicamente, estas variantes se han clasificado en dos grupos según su localización en la célula: el grupo formado por las isoformas α y γ , que se localizan predominantemente en el núcleo; y el grupo de las isoformas β y δ , que lo hacen en el citoplasma (Hagio et al., 2006). La isoforma predominante en las células es la isoforma α , formada por 154 aminoácidos (Hagio et al., 2006; Mori et al., 1998). Su transcrito contiene 6 exones, de los cuales destaca el exón 4 porque es el que codifica para el dominio de localización nuclear (Hagio et al., 2006). Tanto su localización en el núcleo como su presencia en la mayoría de las células hacían que la variante α fuese la más idónea para llevar a cabo el estudio de esta tesis, por lo que decidimos enfocar este trabajo en PFDN5- α (a partir de ahora PFDN5).

El primer paso que dimos para investigar un posible papel de PFDN5 en expresión génica fue comprobar la localización de esta proteína en el núcleo de células HCT116, con las cuales se trabajó a lo largo de esta tesis. Se trata de células de carcinoma colorrectal de humanos, tejido dónde se expresan tanto PFDN5 como PFDN2 (subunidad común del complejo canónico y del complejo *prefoldin-like* con la que también hemos trabajado en esta tesis). Además, esta línea inmortal puede ser transfectada con gran facilidad, y no detiene su ciclo celular en ausencia de suero (véase apartado 2.3 de los resultados de esta tesis). Por todo ello, la línea HCT116 reunía las características idóneas para el desarrollo de este trabajo.

Para localizar PFDN5, se diseñó una línea estable que la expresara etiquetada en el extremo N-terminal con EGFP (una variante mejorada de la proteína de fluorescencia verde). GFP está formada por 238 aminoácidos y su principal característica se basa en su fotoactividad, ya que al ser fotoestimulada emite fluorescencia visible (Prasher et al., 1992). Es por ello que GFP se ha convertido en uno de los genes marcadores más usados a nivel experimental en distintos tipos celulares, ya que permite estimar los niveles de expresión de la proteína y su ubicación en la célula (Misteli and Spector, 1997).

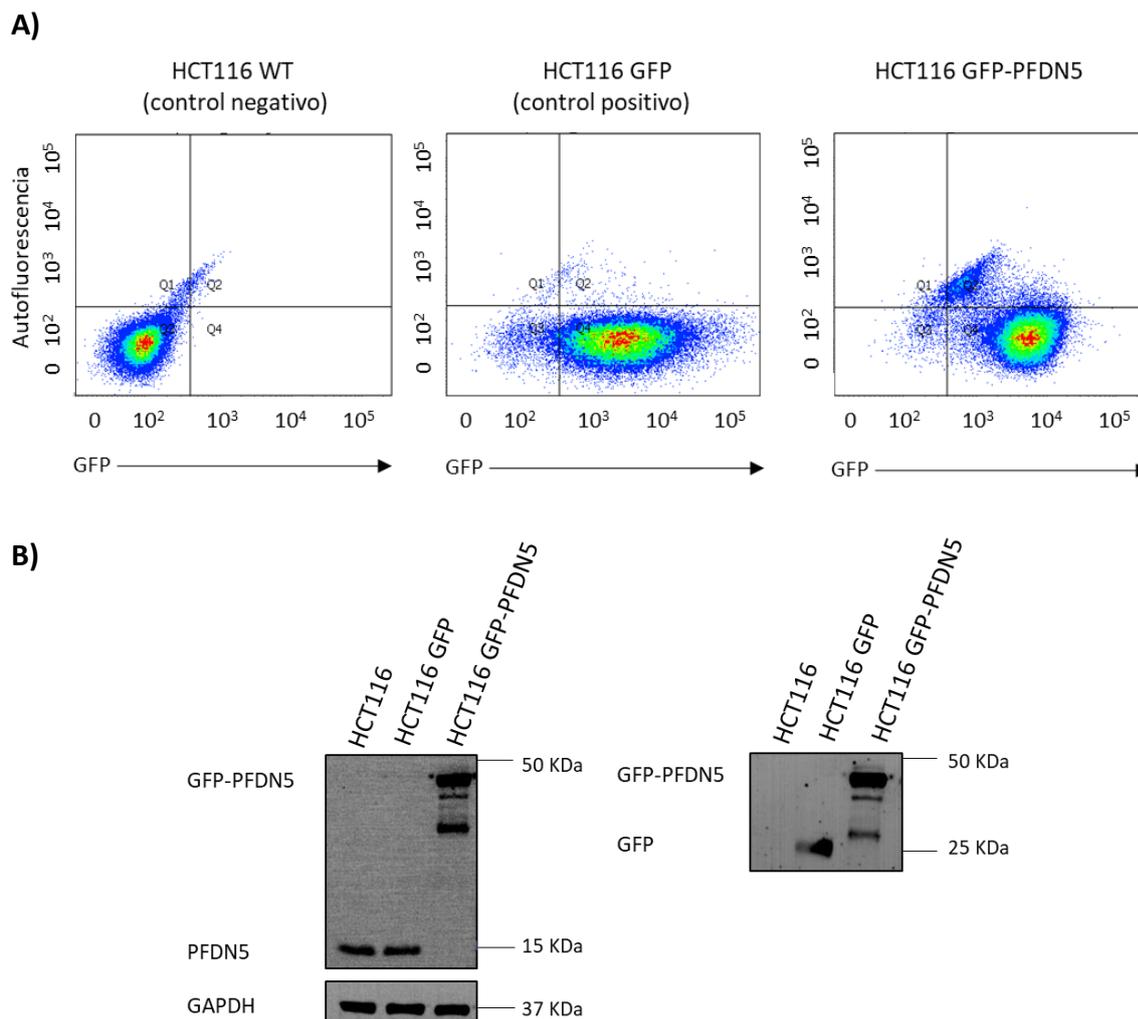


FIGURA R-1. Caracterización de las líneas estables HCT116 GFP y HCT116 GFP-PFDN5. A) Análisis por citometría de flujo de los clones elegidos. En el eje horizontal se representa la intensidad del parámetro fluorescente (GFP) y en el vertical la intensidad de la autofluorescencia. **B)** Análisis por *Western Blot* del contenido de PFDN5 en las poblaciones seleccionadas. El panel superior izquierdo corresponde a la incubación con anticuerpo anti-PFDN5; el panel inferior izquierdo a la incubación con anti-GAPDH; y el panel derecho a la incubación con anti-GFP.

La generación de la línea reportera comenzó con la transfección de células HCT116 con el vector de expresión de EGFP-PFDN5. Este plásmido contiene el gen que codifica la proteína EGFP-PFDN5 bajo control del promotor constitutivo CMV. Como línea control, otra población del mismo tipo celular fue transfectada con el plásmido vacío (pEGFP-C1). Más tarde, se seleccionaron los clones capaces de crecer en medio selectivo suplementado con G418 (marcador de resistencia de los plásmidos mencionados anteriormente). Por último, se analizó el carácter fluorescente de varios

de los clones seleccionados de cada línea por citometría de flujo, para así determinar cuáles eran los más apropiados basándonos en que tuviesen un nivel de fluorescencia parecido (**Figura R-1A**). Además, se comprobó mediante *Western blot* el tamaño de PFDN5 en dichas líneas. Tanto la WT (*wild type*, silvestre) como el control negativo mostraron el mismo tamaño y nivel de expresión de PFDN5. Por el contrario, la línea HCT116 GFP-PFDN5, mostró un nivel de expresión de PFDN5 mayor y su tamaño correspondía a su fusión con EGFP. También se observaron varias bandas más pequeñas y menos intensas que corresponden, posiblemente, a la degradación de EGFP-PFDN5. En esta línea celular, no se observó banda correspondiente a PFDN5 endógena (**Figura R-1B**).

Finalmente se realizaron ensayos de inmunofluorescencia con células de las líneas modificadas genéticamente, HCT116 GFP-PFDN5 y HCT116 GFP, y se comprobó que PFDN5 se expresa tanto en el núcleo como en el citoplasma y que, como era de esperar al tratarse de la isoforma α , su localización es preferentemente nuclear. Sin embargo, nuestra proteína de interés se encuentra excluida del nucléolo, el cual fue marcado con anticuerpo anti fibrilarina (**Figura R-2**).

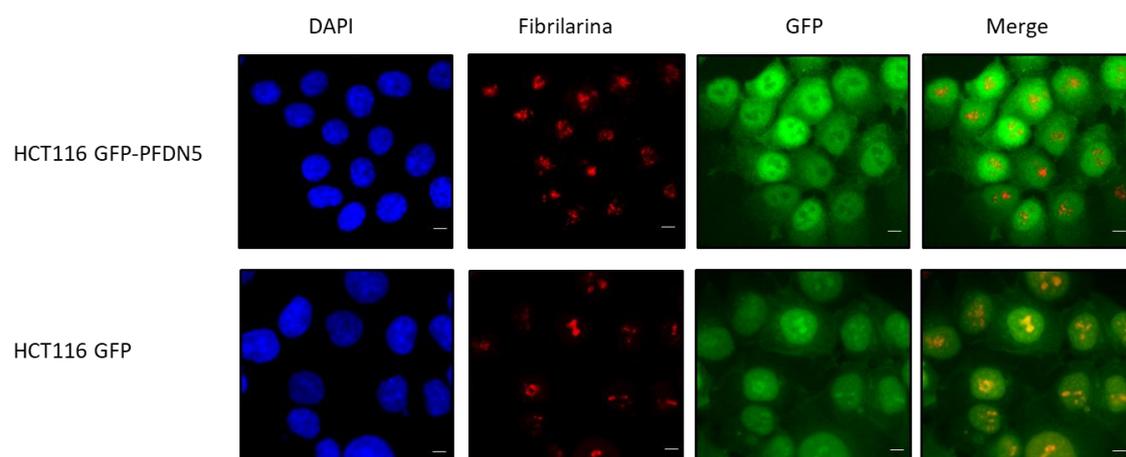


FIGURA R-2. En células HCT116, PFDN5 se localiza en el núcleo y en el citoplasma, pero no en el nucléolo. Células HCT116 que expresan EGFP o EGFP-PFDN5 (verde) teñidas con un anticuerpo que detectaba el marcador nucleolar fibrilarina (rojo). Los núcleos se teñieron con DAPI (azul). Se muestran las imágenes individuales de cada marcaje, así como la superposición del canal rojo y verde (*merge*). Barra = 5 μ m.

2. ESTUDIO TRANSCRIPTÓMICO EN CÉLULAS DEFICIENTES EN PFDN2 Y PFDN5.

2.1 Diseño de los experimentos de RNA-seq.

El siguiente paso que dimos con la intención de explorar las posibles funciones de las prefoldinas en la expresión génica en células humanas, fue diseñar un ensayo de RNA-seq que nos permitiera llevar a cabo un análisis masivo del transcriptoma de células HCT116 que expresaran bajos niveles de PFDN5, y compararlas con el de células HCT116 control. Tal y como se detalla en el apartado 2.1 de la introducción de esta tesis, PFDN5 forma parte del complejo prefoldina canónico, pero no del complejo *prefoldin-like*, por tanto, el análisis de nuestro experimento iba a quedar limitado, únicamente, al complejo canónico. Por ello, decidimos dar un paso más y realizar un estudio más completo, analizando también el transcriptoma de células con bajos niveles de PFDN2, subunidad presente en los dos complejos mencionados anteriormente. De este modo, tendríamos una aproximación más cercana a si, tanto el complejo prefoldina canónico como el complejo alternativo, ejercen o no un papel en la regulación de la expresión génica. Los genes cuya expresión esté influida por el complejo canónico deben responder a la depleción de ambas subunidades. Aquellos otros cuya expresión esté influida por el complejo *prefoldin-like* deben responder únicamente a la depleción de PFDN2. Si existieran otras combinaciones (por ejemplo, genes que respondan a la deficiencia en PFDN5, pero no a la de PFDN2), éstas no se podrían asociar a ninguno de los dos complejos hasta ahora conocidos.

Para la obtención de células que expresaran bajos niveles de las subunidades indicadas, se transfectaron células HCT116 con diferentes siRNAs que tuviesen como diana el transcripto del gen *PFDN2*, por un lado, y el de *PFDN5* por otro (**Figura R-3**).

En cuanto al ensayo de RNA-seq, se desarrolló en células privadas de suero a las que posteriormente se le añadió suero (**Figura R-3**). En esta situación se producen cambios en la expresión de una gran cantidad de genes, que responden a dicho estímulo externo (Hendrickson et al., 1984). Por ello, consideramos que este proceso de cambios generalizados de expresión era el momento óptimo para comprobar el posible papel de las prefoldinas en expresión génica.

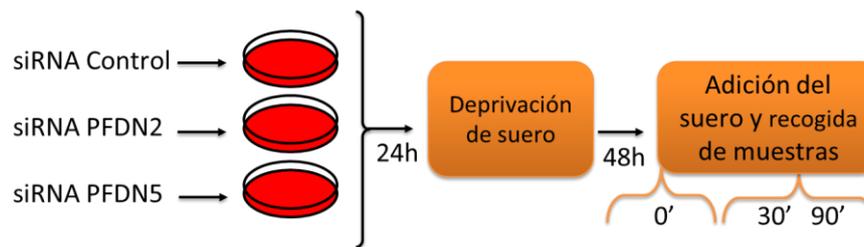


FIGURA R-3. Flujo de trabajo del ensayo de RNA-seq. El experimento comenzó con la transfección de las células, por separado, con los diferentes siRNA. Un día después, se reemplazó el medio completo por medio sin suero y las células fueron incubadas en estas condiciones durante 48 horas. Por último, el medio libre de suero fue sustituido por medio completo, y se procedió a recoger las muestras justo antes de la adición del suero ($t=0'$), y 30 y 90 min después de ésta ($t=30'$ y $t=90'$).

Los primeros genes en responder al suero son los denominados genes de respuesta temprana o IEGs (de sus siglas en inglés "*Immediate early genes*"), cuya transcripción se suele activar unos minutos después de la presencia del suero. A continuación, entre una y dos horas más tarde, se produce una segunda oleada transcriptómica en la que activan su transcripción los genes de respuesta tardía (Harvey, 1991; Lanahan et al., 1992). Teniendo en cuenta estos tiempos, se analizaron los cambios producidos en el transcriptoma de las células deficientes en PFDN2 y PFDN5, respecto a las células control, antes de la inducción con suero, así como 30 y 90 minutos después de ésta (**Figura R-3**).

2.2. La deficiencia en PFDN5 reduce la viabilidad celular, no así la deficiencia en PFDN2.

Diseñamos tres siRNAs que redujeran la expresión de *PFDN2* (siPFDN2 A, siPFDN2 B y siPFDN2 C). Estos, se localizaban en los nucleótidos con posiciones 212-230, 250-268 y 356-374 respectivamente, del ADNc de *PFDN2* (**Figura R-4A**). En la **Figura R-4B** podemos ver el efecto de estos tres siRNA en la expresión de PFDN2, 24 y 48 horas después de la transfección. Aunque el efecto de los 3 candidatos fue parecido, nos quedamos con el siPFDN2-C ya que fue el que mostró mayor eficiencia (niveles menores de mRNA PFDN2). El Dr. José Carlos Reyes nos cedió un RNA de interferencia para reducir la expresión de *PFDN5*, que previamente había sido escogido entre un grupo de tres candidatos. Este siPFDN5 afecta a las isoformas mayoritarias α y γ de PFDN5 (**Figura R-4C**).

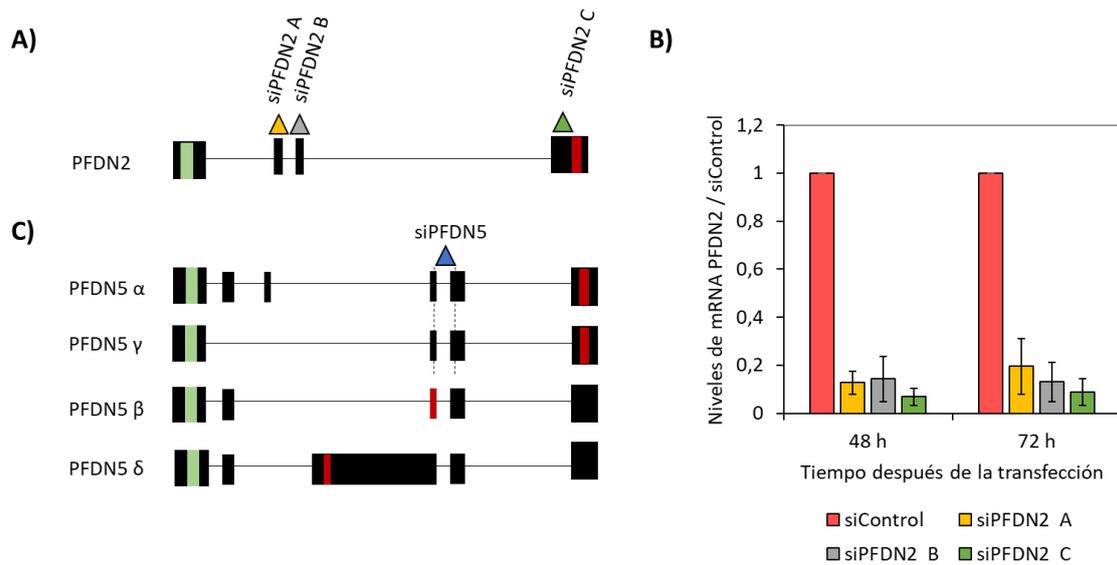


FIGURA R-4. RNAs de interferencia utilizados. **A)** Esquema que representa el gen *PFDN2* completo, con la localización de los tres siRNA utilizados (siPFDN2 A, siPFDN2 B y siPFDN2 C). Cada caja negra representa un exón, el codón de inicio se muestra en verde y el de STOP en rojo. **B)** Expresión de *PFDN2* en células HCT116, 24 y 48 horas después de ser transfectadas con tres siRNAs diferentes contra *PFDN2* o con un siRNA control (siControl). **C)** Esquema que representa las cuatro variantes de splicing de *PFDN5*, con la localización del siRNA utilizado, donde se puede observar que afecta a las isoformas nucleares α y γ de *PFDN5*. Cada caja negra representa un exón, el codón de inicio se muestra en verde y el de STOP en rojo.

Ya con los RNA de interferencia escogidos, se procedió a comprobar su eficiencia 24, 48 y 72 horas después de ser transfectados en las células. Para ello, se extrajo RNA y se prepararon extractos totales de proteínas tras la transfección de los siRNAs en los tiempos mencionados. En las **Figura R-5A** y **Figura R-5B** se muestran los niveles de mRNA de *PFDN2* y *PFDN5* respectivamente. Mientras que el mRNA de *PFDN2* sólo se expresó alrededor de un 20% en todos los tiempos analizados, el mRNA de *PFDN5* se expresó alrededor de un 60% a las 24 horas, un 50% a las 48 horas y un 70% a las 72 horas. Sin embargo, al analizar los extractos totales por *Western blot*, vimos que el contenido más bajo de ambas subunidades se dio 72 horas después de transfectar las células, tiempo en el que se había reducido el contenido de ambas subunidades al menos un 50% (**Figura R-5C** y **Figura R-5D**).

A continuación, se realizaron ensayos de viabilidad celular (ver apartado 9 de materiales y métodos). La deficiencia de *PFDN2* no causó defectos significativos en la viabilidad celular (**Figura R-5E**). Sin embargo, la deficiencia de *PFDN5* sí generó defectos en la viabilidad celular: tal y como puede verse en la **Figura R-5F**, el

porcentaje de células vivas transfectadas con el siPFDN5 disminuía a lo largo del tiempo.

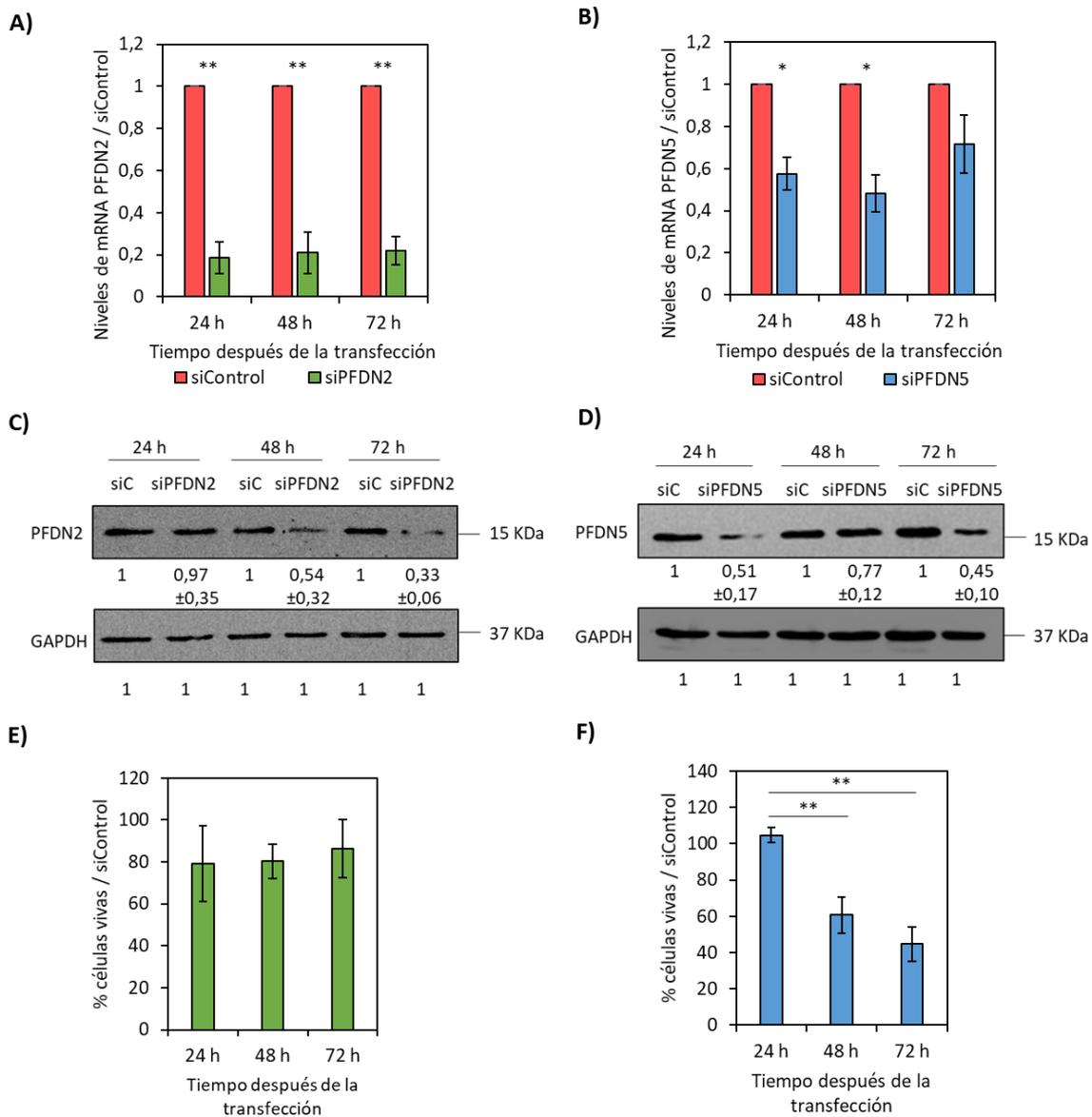


Figura R-5. La deficiencia de *PFDN5*, pero no de *PFDN2*, en la línea HCT116, ocasiona una pérdida de la viabilidad. **A y B)** Expresión de *PFDN2* (**A**) y de *PFDN5* (**B**) en células HCT116, 24, 48 y 72 horas después de ser transfectadas con el siRNA contra *PFDN2/PFDN5* respectivamente, o con un siRNA control (siControl). **C y D)** A partir de los extractos totales obtenidos de las mismas condiciones, se analizaron 40 µg de estos mediante *western blot*, utilizando los anticuerpos contra *PFDN2* (**C**), *PFDN5* (**D**) y GAPDH como control de carga. Los resultados fueron cuantificados con el software *Image Lab*. **E y F)** Ensayo de viabilidad celular en la misma línea celular transfectada con el siPFDN2 (**E**) y el siPFDN5 (**F**) y el siControl por separado. Se contaron las células cada 24 horas durante 3 días, y se representó el porcentaje de células vivas relativizado al control. En todas las figuras se representan valores promedios ± la desviación estándar obtenidos de tres experimentos diferentes. **t*- Student con valor $P < 0,05$ y ***t*- Student con valor $P < 0,005$.

2.3 La falta de suero modifica ligeramente el ciclo celular de la línea HCT116.

Tal y como se describe en el apartado 2.1 de los resultados de esta tesis, el ensayo de RNA-seq se desarrolló en células privadas de suero a las que posteriormente se les restituyó este componente del medio de cultivo. Generalmente, la ausencia de suero afecta al ciclo celular de las células (Chen et al., 2012). Sin embargo, dependiendo de la línea celular de la que se trate, el ciclo celular se puede ver más o menos afectado. Por ejemplo, la falta de suero durante 48 horas provoca que los fibroblastos dejen de proliferar y entren en fase G₀ (Vishwanath *et al.*, 1999). Por el contrario, el ciclo celular de las células H295R (procedentes de carcinoma de la glándula adrenal) no se ve significativamente alterado tras ser cultivadas 48 horas en medio libre de suero (Udhane et al., 2015). En nuestra línea celular, está descrito que la ausencia de suero durante 24 horas solo provoca una pequeña acumulación de células en fase G1 (Sui et al., 2015).

En nuestras condiciones experimentales, la ausencia de suero también ocasionó un aumento del porcentaje de células detenidas en la fase G1 del ciclo celular y una ligera disminución en las fases S y G2 (**Figura R-6**). Dado que los cambios detectados en la progresión de ciclo celular fueron sutiles, no esperamos que interfieran de forma apreciable en los perfiles de expresión analizados por RNA-seq.

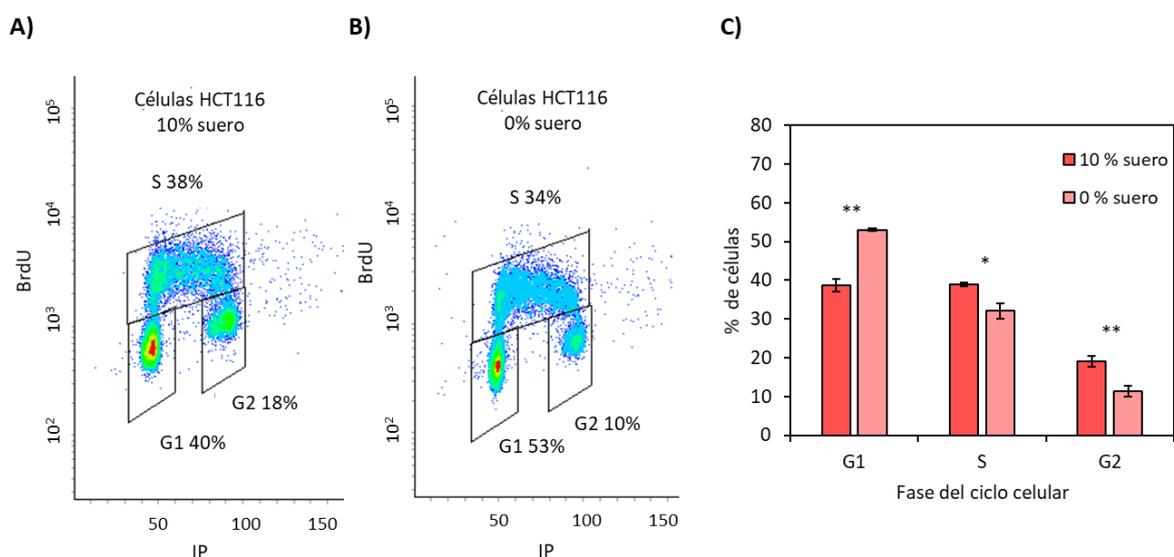


Figura R-6. La privación de suero afecta de forma leve a la progresión del ciclo celular de las células HCT116. Análisis de la progresión del ciclo celular de células HCT116 que 24 horas después de ser sembradas, fueron cultivadas durante 48 horas en medio completo **(A)** o medio libre de suero **(B)**, y, por último, marcadas con BrdU durante 15 minutos. El BrdU, que se incorpora a las células en fase S, fue visualizado marcándolo con anticuerpo anti-BrdU y usando un anticuerpo secundario unido a un fluoróforo. En el eje vertical se visualiza la intensidad emitida por el BrdU y en el horizontal la del Ioduro de Propidio (IP). **(C)** Histograma de barras del porcentaje de células en cada fase del ciclo celular. En la gráfica se representa valores promedios \pm la desviación estándar obtenidos de tres experimentos diferentes. **t*- Student con valor $P < 0,05$ y ***t*- Student con valor $P < 0,005$.

2.4. Obtención de muestras para los experimentos de RNA-seq.

El ensayo de RNA-seq quedó establecido finalmente del siguiente modo: antes de la extracción del RNA, las células se trataron con los diferentes siRNA durante 72 horas, de las cuales las últimas 48 horas estuvieron privadas de suero. La extracción del RNA se hizo antes de la adición del suero ($t = 0$ minutos), y 30 y 90 minutos después de ésta. El experimento se realizó por duplicado.

En una primera aproximación para confirmar que el ensayo se había ejecutado correctamente, y antes de enviar las muestras a secuenciar, se comprobó el contenido en PFDN2 en las muestras transfectadas con el siPFDN2, y el de PFDN5 en las muestras transfectadas con el siPFDN5 mediante *Western blot* realizado con extractos totales recogidos del mismo experimento **(Figura R-7A)**.

Además, se determinó que la activación de la transcripción con suero había funcionado analizando los niveles de mRNA de *FOS* **(Figura R-7B)**. *FOS* es un gen de respuesta temprana que presenta su pico máximo de expresión entre los 30 y los 60 minutos después de que se produzca el estímulo, y regresa a sus niveles de expresión basal a los 90 minutos (Healy et al., 2013). Nuestros resultados mostraron esta cinética de expresión sin cambios aparentes entre las células control y las deficientes en cada subunidad.

Una vez realizada esta serie de chequeos, el RNA extraído y tratado con DNAsa se envió a secuenciar a la Unidad de Genómica del Centro de Regulación Genómica de Barcelona. Posteriormente se realizó el análisis bioinformático con la ayuda del Dr. Ildefonso Cases.

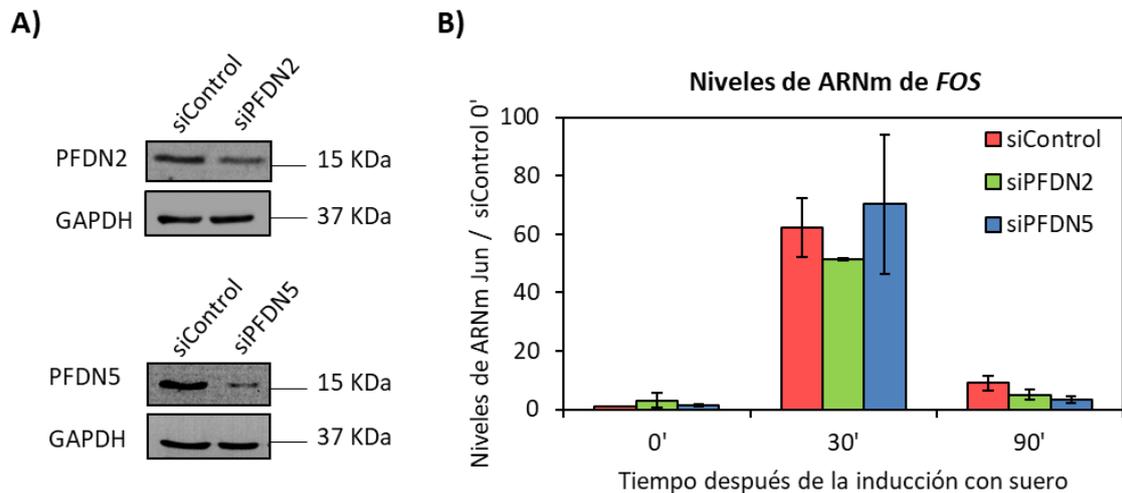


Figura R-7. Verificaciones realizadas en el ensayo de RNA-seq antes de enviar las muestras de cDNA a secuenciar. A) Se obtuvieron extractos totales de las diferentes muestras y condiciones, y se analizaron 40 μ g de estos mediante *Western Blot*, utilizando los anticuerpos contra PFDN2, PFDN5 y GAPDH. **B)** Se realizó un ensayo de RT-qPCR a partir del RNA extraído de las diferentes muestras y condiciones y se analizaron los niveles de mRNA de *FOS*. En la gráfica se representan los valores promedio \pm la desviación estándar de dos experimentos diferentes.

2.5 La deficiencia de PFDN5, y en menor medida la de PFDN2, altera la expresión de un alto número de genes.

Los primeros resultados obtenidos del análisis bioinformático mostraron que las células transfectadas con el siPFDN5 presentan multitud de cambios en su transcriptoma, antes y después de la inducción con suero. No obstante, se trata de cambios muy pequeños. Por el contrario, las células transfectadas con el siPFDN2 muestran pocos cambios en la expresión génica, principalmente en los tiempos 0' y 30', donde apenas se observan diferencias con respecto al control. En ambos casos, las muestras recogidas 90 minutos después de la inducción con suero son las que mayor número de cambios presentan (**Figura R-8A y 8B**). Estas diferencias entre las células deficientes en PFDN2 y PFDN5 son coherentes con el diferente impacto en la viabilidad celular mostrada anteriormente (**Figura R-5E**).

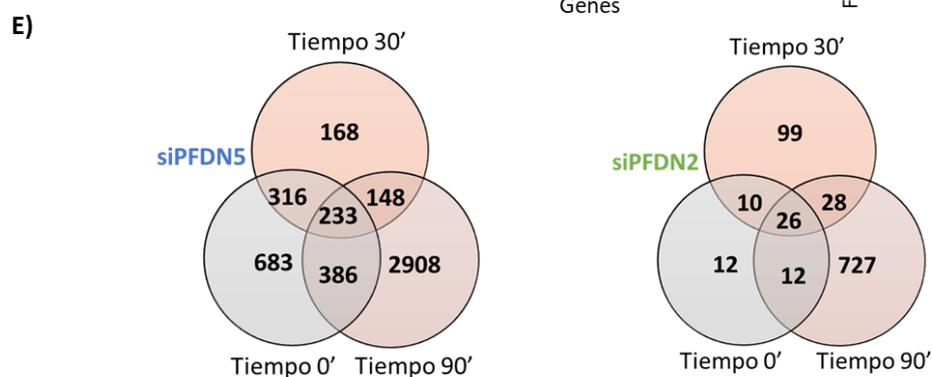
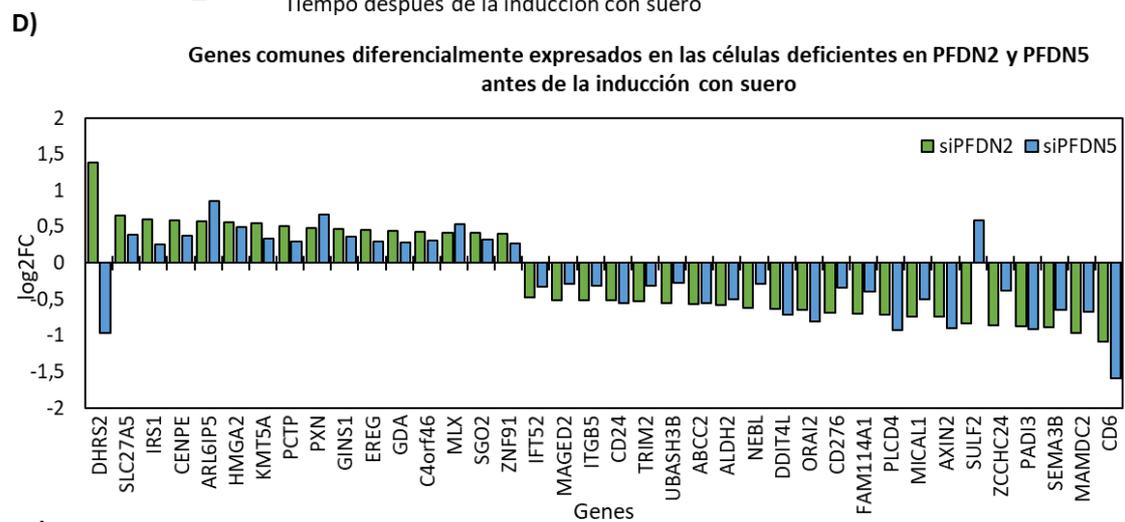
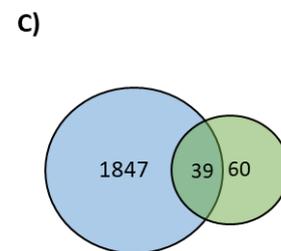
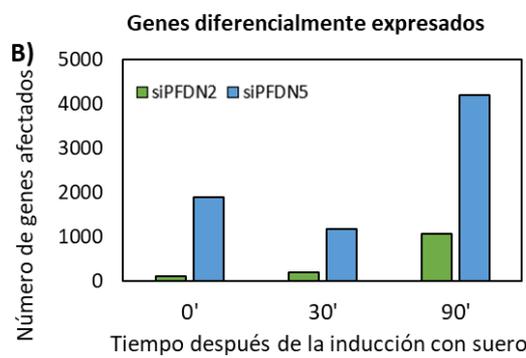
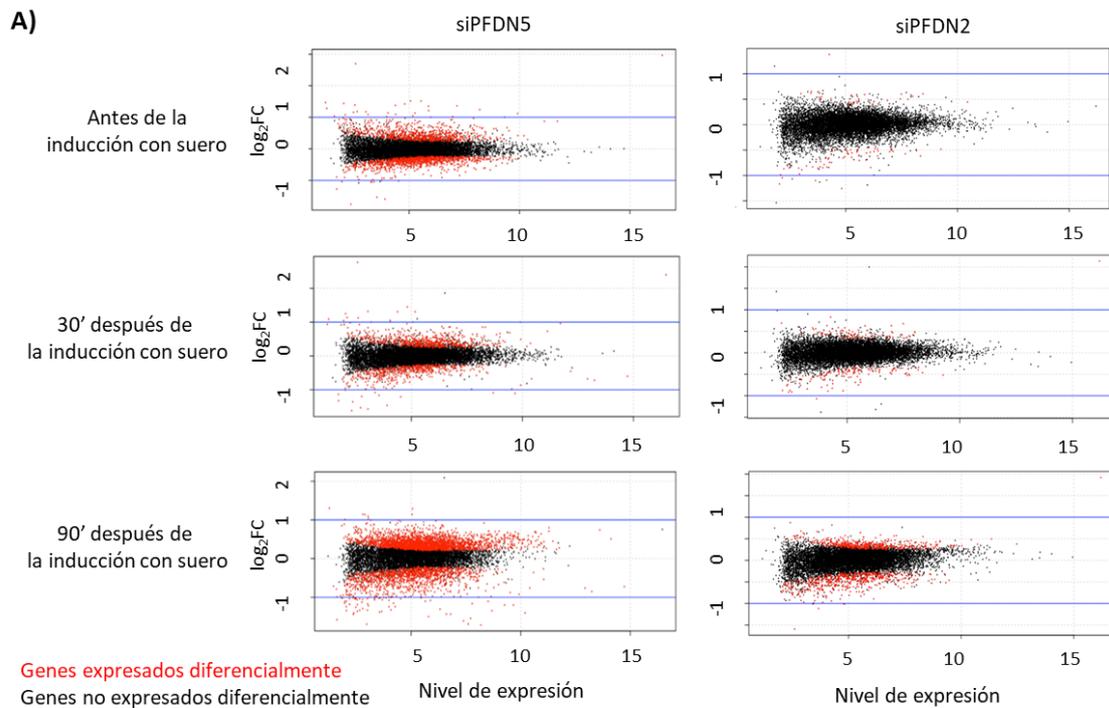


Figura R-8. La deficiencia en PFDN5, y en menor medida en PFDN2, causa cambios en el transcriptoma, tanto en ausencia de suero como al activar las células con suero. A) Representación de todos los genes secuenciados en el RNA-seq. Se muestra el logaritmo del cambio (\log_2FC) frente al nivel de expresión de los genes (definido como la media del logaritmo en base 2 de las cuentas por millón de las dos muestras de cada experimento). Las gráficas de la izquierda muestran los cambios transcriptómicos de las muestras transfectadas con el siPFDN5 con respecto al control, y las de la derecha, los de las muestras transfectadas con el siPFDN2. En negro se representan los genes no afectados, mientras que en rojo se muestran todos aquellos genes cuyo $\log_2FC > 0$, $FRD < 0.05$. **B)** Representación del número de genes afectados en cada muestra y tiempo. $\log_2FC > 0$, $FRD < 0.05$ **C)** Se muestra el número de genes expresados diferencialmente en las células transfectadas con el siPFDN2 (verde) y con el siPFDN5 (azul), y el solapamiento entre estos conjuntos, antes de la inducción con suero. **D)** Cambios de expresión de los genes comunes diferencialmente expresados en células deficientes en PFDN2 y PFDN5 antes de la inducción con suero. **E)** Representación del número de genes expresados diferencialmente en las células transfectadas con el siPFDN2 y con el siPFDN5 a cada tiempo y los genes que coinciden entre los distintos tiempos.

A pesar del limitado efecto de la deficiencia en PFDN2 sobre el transcriptoma de nuestras células en ausencia de suero, el reducido número de genes afectados en dichas condiciones mostró un considerable solapamiento con los afectados por la deficiencia en PFDN5 (**Figura R-8C**). Además, la gran mayoría de genes comúnmente afectados por la deficiencia en las dos subunidades mostraron cambios en el mismo sentido respecto a las células control, lo que sugiere que la alteración del complejo canónico de prefoldina explica, al menos, una parte de los cambios detectados (**Figura R-8D**).

Para identificar los genes afectados por la deficiencia en PFDN2 y PFDN5 independientemente de la presencia o ausencia de suero, buscamos aquellos genes que mostraron una alteración significativa ($FRD < 0.05$) de su expresión en todas las condiciones analizadas. Los resultados volvieron a mostrar que los efectos de la depleción de PFDN5 eran más severos que los efectos de la depleción de PFDN2, ya que mientras con siPFDN5 se encontraron 231 genes diferencialmente expresados en los 3 tiempos, con siPFDN2 fueron únicamente 25 (**Figura R-8E**).

El análisis de las categorías funcionales de genes (*GOs*) permite detectar cambios regulatorios en la expresión del genoma que pueden pasar inadvertidos con el análisis individual de genes. Los resultados obtenidos de dicho análisis no mostraron ninguna

categoría funcional significativamente enriquecida (valor $P < 0.05$) en los grupos de genes comunes entre los diferentes tiempos en las células deficientes en PFDN2, pero sí en las deficientes en PFDN5. Éstas fueron: “respuesta a estímulos externos”, “vía de señalización mediada por interleuquina-7” y “respuesta celular a estímulos químicos”.

Tabla R-1.

SUBUNIDAD	GENES COMUNES DIFERENCIALMENTE EXPRESADOS EN TODOS LOS TIEMPOS
PFDN2	<i>ABCC2, ABHD14B, ANKRD13C, DHRS2, DRAXIN, HMGA2, IRS1, ITGB5, KMT5A, LAMB1, MAGED2, MRC2, MXD4, NEBL, NEURL1B, NUCKS1, ORAI2, SAMD4B, SLC27A5, SLC2A3, SULF2, SYNJ2BP, TRIM2, TUBA1A</i> y <i>USP25</i>
PFDN5	<i>AAK1, ABCC2, ABCC3, ABCD1, ABHD14B, ACBD3, ACBD6, ADAMTS14, ADGRF1, AGRN, AHNAK, AHNAK2, ALS2CL, AMPD2, ANKHD1, AREG, ARHGEF3, ARL6IP5, ASCC3, ASNA1, ATF3, ATP13A3, ATP5IF1, BAZ2A, BCOR, BSCL2, C1RL, C6orf120, C9orf3, CA2, CFBF, CD55, CD6, CD99L2, CDC42EP4, CDC42SE2, CDK2AP2, CDKN1A, CDT1, CDYL, CEMIP, CENPX, CHD7, CHMP2A, CLDN12, CPA4, CPNE8, CPSF7, CST6, CXCL8, CXXC5, DDIT3, DDX11, DDX21, DDX39B, DGCR2, DHRS2, DNAJB11, DNAJC6, DNMT2, DNMBP, DPYSL2, DSP, DUSP6, E2F2, EFR3A, EGFR, EGR2, EID1, EIF3A, EIF3K, ELMO3, ELP5, EME2, EP400, ERFFI1, FAM86B1, FAS, FAT1, FGFR1, FHL, FLII, FNIP1, FRAS1, GALNT5, GDF15, GDI1, GEM, GMDS, GOT1, GPR161, HDAC10, HIPK3, HIST1H3A, HIST1H3B, HISTH3J, HIST2H3D, HLA-B, HMGA2, HR, HSP90B1, HSPA4L, HUWE1, IFITM2, IFNGR2, IGF2R, IGSF9, INF2, IRS2, TGA2, JMY, KANK2, KIF26A, KLHL9, KMT2A, KMT5A, KRT80, LAMB2, LAMC2, LATS2, LDLRAP1, LEPR, LPP, LRRK1, MAFF, MANF, MAPK9, MARVELD2, MCM6, MED24, MEGF6, METTL7A, MEX3A, MICALL2, MKI67, MKLN1, MLX, MLXIPL, MTA1, NBEAL2, NDUFS6, NEBL, NFAT5, NFATC2IP, NOTCH2NL, NPC2, OPLAH, RAI2, PARD6B, PDIA3, PDIA4, PEX26, PFDN2, PHLDA2, PHLDB1, PIEZO1, PIK3R1, PKD1, PLCG1, PLXNA1, PLXNB1, PMAIP1, POLH, POLM, PPM1D, PPME1, PRKDC, PROS1, PRRC2C, PSD4, PSMA4, RAB36, RAB3D, RAB8B, RCOR2, RGS2, RMI1, RMND5A, RNASEK, RPS17, RPS27, RPS27L, SAMHD1, SBDS, SBF1, SCAMP4, SCARA3, SCRIB, SDF2L1, SEM1, SERINC5, SERPINI1, SESN1, SESTD1, SH3TC2, SIVA1, SLC27A1, SMN1, SMN2, SOX4, SPPL2B, SRA1, SRC, REBF1, STAT3, STAT6, SULF, SULT1A4, SYBU, TAF1C, TBC1D9B, THRA, TIMM21, TMEM132A, TMEM173, TNFRSF6B, TNK1, TP53I3, TP53INP1, TSG101, TSPAN12, TTN, TUBGCP6, TULP4, USF1, USP53, TP3, VPS13C, ZFH3, ZMAT3, ZNF367 y ZSWIM6</i>

También detectamos que entre los 232 genes afectados por la deficiencia en PFDN5, antes y después de la inducción por suero, hubo varios relacionados con la organización y dinámica del citoesqueleto (*FLII, TUBGCP6, CDC42EP4*) lo que es esperable dada la demostrada vinculación funcional de las prefoldinas en ese proceso. Asimismo, la lista incluye al gen *PFDN2*, que se induce en condiciones de deficiencia de PFDN5, probablemente en respuesta a su propia desestabilización (véase más adelante). La lista de genes significativamente alterados por la deficiencia en PFDN5 en cualquier condición de suero también incluye genes funcionalmente implicados en diferentes fases del proceso de expresión génica, como factores de transcripción (*TAF1C, E2F2, STAT6, USF1, MED24*), factores de procesamiento del pre-mRNA (*SMN2*,

CPSF37, *DDX39B*) y factores modificadores de histonas (*HDAC10*, *KMT5A*) (**Tabla R-1**). Precisamente este último gen también está incluido en la lista de genes con expresión diferencial a tiempos 0', 30' y 90' de las células deficientes en *PFDN2* (**Tabla R-1**). Además, detectamos varios genes que codifican histonas (*HIST1H3A*, *HIST1H3C*, *HIST1H3D*, *HIST1H3J*) y reguladores del ciclo celular (*CDKN1A*, *MKI67*, *CDK2AP2*, *MCM6*) (**Tabla R-1**).

2.6. La deficiencia de *PFDN5* y *PFDN2* altera la expresión de categorías funcionales de genes.

Además de lo anterior, también se realizó un análisis de las *GOs* de los genes diferencialmente expresados por la deficiencia en *PFDN2* y *PFDN5* a tiempo 0', 30 y 90'.

Las células deficientes en *PFDN5* mostraron un total de 11 categorías funcionales afectadas a tiempo 0, de las cuales 7 mostraron expresión incrementada y 4 expresión disminuida (**Tabla R-2**). Entre las primeras se encontró la *GO* "organización de telómeros", y entre las segundas, las categorías funcionales de "ensamblaje de nucleosomas" y "cambios conformacionales del DNA" (**Tabla R-2**). Este resultado, junto a los cambios de expresión de genes que codifican reguladores del ciclo celular mencionados anteriormente, sugieren una conexión funcional de *PFDN5* con la replicación del genoma y el ciclo mitótico. La posible implicación de *PFDN5* en la progresión del ciclo celular se analizará en el apartado 3.4 de los resultados de esta tesis.

Asimismo, detectamos el incremento de expresión de varias categorías de genes relacionadas con señalización celular y respuesta a estímulos (**Tabla R-2**).

A tiempo 30 solo hubo 5 categorías funcionales afectadas por la deficiencia en *PFDN5*, 2 disminuyeron su expresión y 3 la incrementaron (**Tabla R-3**). Estas tres últimas estaban relacionadas con morfogénesis, desarrollo y diferenciación celular (**Tabla R-3**).

Tabla R-2.

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN5 EN AUSENCIA DE SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Disminución	Ensamblaje de nucleosomas	2,98E-15
	Cambios conformacionales del DNA	5,48E-13
	Regulación negativa de la diferenciación de megacariocitos	1,11E-09
	Desarrollo celular	1,2014E-06
Incremento	Comunicación celular	4,9237E-09
	Señalización	1,6012E-08
	Desarrollo del sistema	3,3783E-08
	Organización de los telómeros	1,31E-07
	Respuesta a estímulos externos	2,0133E-06
	Respuesta a hormonas esteroides	5,8453E-06
	Desarrollo del tejido	6,6571E-06

Tabla R-3.

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN5 30 MINUTOS DESPUÉS DE LA INDUCCIÓN CON SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Disminución	Regulación extracelular negativa de la transducción de señales	1,07902E-07
	Regulación negativa de la fase de ejecución de la apoptosis	7,89009E-07
Incremento	Desarrollo de organismos multicelulares	1,79E-07
	Diferenciación celular	5,06E-07
	Morfogénesis de la estructura anatómica	8,14E-06

Sin embargo, a tiempo 90 se detectaron muchas más categorías funcionales afectadas por la deficiencia en PFDN5 que en los tiempos anteriormente mencionados.

En total fueron 61 GOs afectadas, 28 de las cuales disminuyeron su expresión y 33 la incrementaron (**Tabla R-4**).

Curiosamente en este caso, y al contrario que a tiempo 30, se encontraron disminuidas en su expresión numerosas GOs relacionadas con la morfogénesis, el desarrollo y la diferenciación. También fue el caso de las categorías funcionales de “adhesión celular”, “movimiento de componentes celulares o subcelulares” y “organización del citoesqueleto” (**Tabla R-4**). Los cambios de expresión de estas GOs podrían estar relacionados con la función del complejo canónico de prefoldina en el plegamiento de monómeros de actina y tubulina durante el ensamblaje de estructuras citoesqueléticas.

Asimismo, entre las categorías funcionales con expresión incrementada encontramos un amplio número de GOs relacionadas con la biogénesis, funcionamiento y actividad metabólica de las mitocondrias (**Tabla R-4**). Este hecho puede derivar de la conexión funcional existente entre PFDN5 y c-Myc, un regulador que activa la expresión de genes mitocondriales en respuesta al suero (Kim et al., 2008).

También en las células deficientes en PFDN5 y a tiempo 90, detectamos disminuida la expresión de las categorías funcionales “reguladores positivos de la transcripción por la RNAPII” y “reguladores negativos de la transcripción por la RNAPII”, e incrementadas las GOs “terminación de la transcripción de la cadena molde de DNA”, “ensamblaje de las snRNP en el espliceosoma” y “*splicing* del pre-mRNA vía espliceosoma” (**Tabla R-4**). Estos cambios sugieren alteraciones en el proceso de expresión génica durante estadios posteriores a la iniciación de la transcripción y han de considerarse a la hora de interpretar resultados posteriores de esta tesis doctoral.

Tabla R-4.

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN5 90 MINUTOS DESPUÉS DE LA INDUCCIÓN CON SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Disminución	Desarrollo del sistema nervioso	0
	Desarrollo de órganos animales	8,88E-16

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN5 90 MINUTOS DESPUÉS DE LA INDUCCIÓN CON SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Disminución	Regulación negativa de la transcripción por la RNAPII	3,69E-12
	Regulación positiva de la transcripción por la RNAPII	2,84E-11
	Regulación de la transducción de señales de la proteína RAS	6,02E-11
	Regulación de procesos de organismos multicelulares	7,97E-11
	Desarrollo celular	1,91E-10
	Movimiento de componentes celulares o subcelulares	1,94E-09
	Adhesión celular	6,30E-09
	Morfogénesis de órganos animales	1,09E-07
	Desarrollo del sistema esquelético	2,16E-07
	Organización del citoesqueleto	2,32E-07
	Formación de estructuras anatómica implicadas en la morfogénesis	3,24E-07
	Transducción de señales intracelulares	3,64E-07
	Desarrollo de la vascularización	6,24E-07
	Regulación de la diferenciación celular	8,76E-07
	Desarrollo del corazón	1,94E-06
	Desarrollo del sistema renal	2,08E-06
	Respuesta a componentes del ciclo celular	3,74E-06
	Ensamblaje del complejo β -catenina-TCF	1,52E-05
	Desarrollo pulmonar	1,89E-05
	Regulación negativa del proceso de desarrollo	1,89E-05
	Regulación positiva de la actividad GTPasa	2,20E-05
	Desarrollo del tejido de músculo estriado	2,41E-05
	Organización sináptica	2,87E-05
	Regulación circadiana de la expresión génica	3,37E-05
	Regulación de estructuras o actividades sinápticas	4,61E-05
Desarrollo del tejido conectivo	5,72E-05	

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN5 90 MINUTOS DESPUÉS DE LA INDUCCIÓN CON SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Incremento	Iniciación de la traducción	0
	Organización mitocondrial	0
	Traducción mitocondrial	0
	Biogénesis mitocondrial	0
	Elongación de la traducción mitocondrial	0
	Establecimiento de la localización de proteínas de membrana	0
	Ensamblaje del complejo de la cadena respiratoria mitocondrial	8,88E-16
	Síntesis de ATP acoplado al transporte de electrones	1,22E-15
	Proceso metabólico ncRNA	4,44E-15
	Cadena de transporte de electrones	2,94E-14
	Plegamiento de proteínas	3,83E-10
	Ensamblaje del complejo DNA-proteína	1,09E-07
	Organización de la membrana mitocondrial	1,79E-07
	Establecimiento de la localización de proteínas de la mitocondria	2,33E-07
	Regulación de la actividad endopeptidasa	1,36E-06
	Regulación de la proteólisis	1,41E-06
	Respuesta inmune humoral antimicrobiana mediada por péptido antimicrobiano	1,50E-06
	Heterodimerización de proteínas	1,94E-06
	Detoxificación celular	4,76E-06
	Ensamblaje de las snRNP en el espliceosoma	4,84E-06
	Reparación de la escisión de nucleótido acoplada a la transcripción	5,46E-06
	Muerte celular	7,69E-06
	Proceso catabólico dependiente del complejo promotor de la anafase	9,39E-06
	Activación de neutrófilos	1,07E-05
	Regulación positiva de la localización de proteínas en la región telomérica del cromosoma	1,19E-05
	Procesamiento catabólico de proteínas independientes de ubiquitina	1,87E-05
	Transporte de protones acoplado a la síntesis de ATP mitocondrial	1,96E-05
	Terminación de la transcripción de la cadena molde del DNA	2,35E-05

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN5 90 MINUTOS DESPUÉS DE LA INDUCCIÓN CON SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Incremento	Desubiquitinación de proteínas	2,77E-05
	Regulación negativa de la actividad peptidasa	4,16E-05
	<i>Splicing</i> del pre-mRNA vía espliceosoma	4,20E-05
	Homeostasis redox celular	4,38E-05
	Muerte celular de otros organismos	5,45E-05

Las células deficientes en PFDN2 no mostraron categorías funcionales afectadas ni antes de la inducción con suero ni 30 minutos después de ésta. Sin embargo, a tiempo 90, coincidiendo con el tiempo en que estas células mostraron un número mayor de genes con expresión diferencial, hubo un total de 32 categorías funcionales alteradas, de las cuales 26 disminuyeron su expresión y 6 la incrementaron (**Tabla R-5**).

Entre las categorías funcionales con expresión disminuida encontramos, de nuevo, algunas relacionadas con morfogénesis, desarrollo, adhesión y diferenciación celular, así como con la organización del citoesqueleto. De hecho, nueve de estas GOs coincidieron con las halladas en las células deficientes en PFDN5 a tiempo 90 (compárense las **tablas R-4 y R-5**).

Asimismo, detectamos categorías funcionales con expresión incrementada relacionadas con la transición metafase/anafase durante la división celular y el ensamblaje del cinetocoro (**Tabla R-5**). Este hecho puede estar relacionado con el papel del complejo canónico en la estabilización de tubulinas durante la formación del huso (Delgehyr et al., 2012; Palumbo et al., 2015) o con la contribución del complejo alternativo en la correcta regulación de la citoquinesis desde los centrosomas (Hintermair et al., 2016).

Entre estas categorías apareció de nuevo “ensamblaje de nucleosomas” aunque en esta ocasión, y en contra de lo que ocurría en las células deficientes de PFDN5 privadas de suero, la expresión de sus genes estuvo incrementada (**Tabla R-5**).

Tabla R-5

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN2 90 MINUTOS DESPUÉS DE LA INDUCCIÓN CON SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Disminución	Adhesión celular	2,78E-14
	Morfogénesis celular involucrada en la diferenciación	1,48E-12
	Desarrollo del sistema nervioso	7,14E-12
	Regulación del desarrollo de organismos multicelulares	1,06E-11
	Desarrollo tubular	1,11E-11
	Regulación de la señalización	5,74E-11
	Migración celular	7,86E-11
	Regulación de la comunicación celular	1,05E-10
	Desarrollo tisular	1,13E-10
	Morfogénesis embrionaria	1,42E-10
	Morfogénesis de órganos animales	3,17E-10
	Regulación positiva de procesos en organismos multicelulares	3,98E-09
	Morfogénesis de la proyección de neuronas	9,48E-09
	Regulación de la diferenciación celular	1,26E-08
	Desarrollo de la vascularización	2,33E-08
	Regulación de la morfogénesis de estructuras anatómicas	3,13E-08
	Desarrollo del crecimiento	1,02E-07
	Respuesta celular al estímulo del factor de crecimiento	6,35E-07
	Regulación positiva del proceso de desarrollo	1,03E-06
	Desarrollo dendrítico	4,54E-06
	Regulación negativa del proceso de desarrollo	4,56E-06
	Procesos basados en filamentos de actina	5,85E-06
	Desarrollo de estructuras musculares	6,13E-06
	Organización sináptica	2,30E-05
	Organización del citoesqueleto	2,78E-05
	Desarrollo del corazón	4,70E-05

CATEGORÍAS FUNCIONALES CUYO NIVEL DE EXPRESIÓN CAMBIA EN CÉLULAS DEFICIENTES EN PFDN2 90 MINUTOS DESPUÉS DE LA INDUCCIÓN CON SUERO		
CAMBIO	CATEGORÍA FUNCIONAL (GO):	VALOR P
Incremento	División celular	2,8701E-10
	Segregación mitótica de las cromátidas hermanas	1,9616E-09
	Regulación de la transición metafase/anafase mitótica	1,15E-07
	Ensamblaje de nucleosomas	1,2658E-07
	Ensamblaje de complejos proteicos	5,9704E-07
	Localización de proteínas en la región centromérica del cromosoma	9,4478E-06

En resumen, el análisis del número de genes expresados diferencialmente y las categorías funcionales afectadas muestra que, en condiciones de inducción con suero, principalmente 90 minutos después de ésta, los bajos niveles de PFDN2 y PFDN5 causan cambios importantes y generalizados en la expresión del genoma. La comparación de los resultados obtenidos entre las células deficientes en estas dos subunidades sugiere que parte de los cambios están relacionados con funciones del complejo canónico de prefoldina. Por el contrario, otras de las alteraciones detectadas no son coincidentes y podrían deberse a funciones del complejo alternativo o de la subunidad PFDN5 fuera del contexto del complejo canónico.

2.7. La inducción por suero se ve más afectada por la deficiencia en PFDN5 que en PFDN2, aunque hay solapamiento.

A continuación, se analizó cómo afectaba la deficiencia en PFDN5 y PFDN2 a la inducción de genes por suero. Este análisis debía darnos información sobre un posible papel de las prefoldinas en el proceso de expresión génica. Cuantificamos el cambio de la expresión de cada gen en las células deficientes en cada subunidad después de la adición del suero y se comparó con el cambio producido en las células control. Consideramos como cambio significativo una alteración positiva o negativa de al menos dos veces lo expresado y con un FDR menor de 0.05. Los resultados del análisis nos mostraron que 90 minutos después de la inducción con suero hubo un mayor número de genes alterados en su regulación que a los 30 minutos. Y al igual que en los

análisis anteriores, las células transfectadas con el siPFDN5 presentaron mayor número de genes regulados diferencialmente que las transfectadas con el siPFDN2 (**Figura R-9**). En ambos casos hubo un solapamiento muy significativo ($FDR < 0.05$) de los genes alterados en células deficientes en PFDN2 y PFDN5. La mitad de los genes alterados en su inducción 30' después de la adición del suero en células deficientes en PFDN2 lo hicieron también en células deficientes en PFDN5 y, además, cambiaban en el mismo sentido respecto a las células control (**Figura R-9**). El solapamiento fue aún mayor en el caso de los genes alterados en su inducción 90' después de la adición del suero, donde dos tercios de los genes alterados por la deficiencia en PFDN2 también lo fueron por la deficiencia en PFDN5 (**Figura R-9**). En la mayoría de los genes comunes se produjo una disminución de la inducción (**Tabla R-6**). Estos resultados sugieren fuertemente que el complejo canónico juega un papel en la expresión de genes y que lo hace preferentemente de forma positiva.

Entre los genes con inducción alterada en ambos tipos de células deficientes en PFDN2 y PFDN5 se incluyen dos desmetilasas de lisinas (*KDM6B* y *KDM7A*) y varios reguladores transcripcionales (*ELF4*, *HES4*, *MN1*, *MNT* y *TP53INP2*) (**Tabla R-6**).

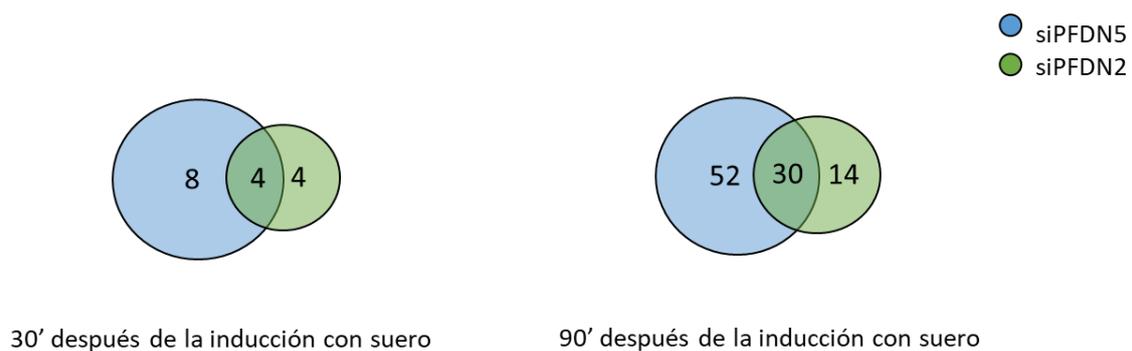


Figura R-9. Inducción diferencial por suero. Se muestra el número de genes regulados diferencialmente en las células transfectadas con el siPFDN2 (verde) y con el siPFDN5 (azul), y el solapamiento entre estos conjuntos, 30' y 90' después de la inducción con suero con $\log_2FC > 1$ y $FDR < 0.05$.

Tabla R-6

	GENES CUYA INDUCCIÓN POR SUERO CAMBIA TANTO EN RESPUESTA A LA DEFICIENCIA DE PFDN5 COMO A LA DE PFDN2	
	Genes cuya inducción disminuye	Genes cuya inducción aumenta
t = 30'	ARC, GPR183	ANKRD37, PTGER4
t = 90'	ANKRD33B, BCL2L12, DENND3, HES4, KDM6B, KDM7A, MN1, MNT, MYH9, PER2, RPLP1, RPS16, SERTAD2, STC2, TBX8, TP53INP2, TSC22D2, UBALD1, USP36, ZBTB5, ZYX	AADACP1, AATK, CLDN4, ELF3, RHOB, SERTAD1, TMEM158, TNFRSF12A, TNFSF18,

3. LA AUSENCIA Y DEFICIENCIA EN PFDN5 AFECTA AL CICLO CELULAR.

3.1. Tanto la ausencia como la deficiencia de PFDN5 provocan una disminución de los niveles del resto de subunidades, aunque no de forma homogénea.

El solapamiento entre las células deficientes en PFDN5 y PFDN2 en los resultados transcriptómicos obtenidos puede deberse a que la depleción de una subunidad provoque la desestabilización del conjunto de subunidades del complejo canónico (Miyazawa et al., 2011). Para comprobar el impacto de la deficiencia en PFDN5 sobre el resto de las subunidades que constituyen el complejo *prefoldina*, se realizaron ensayos de *western blot* en los que se utilizaron extractos totales obtenidos de tres réplicas biológicas de células transfectadas con siPFDN5. Los resultados de mostraron que la depleción de PFDN5 afectó negativamente a los niveles de las subunidades PFDN1, PFDN4 y PFDN6 (**Figura R-10A**). La subunidad PFDN3 no cambió sus niveles de expresión y la subunidad PFDN2 los aumentó sutilmente (**Figura R-10A**).

Paralelamente a los estudios descritos, la estudiante de máster Yerma Pareja Sánchez obtuvo una línea mutante estable que carecía de PFDN5 (HCT116 PFDN5 KO), utilizando para ello la técnica de mutagénesis CRISPR-CAS9. Esta tecnología permite generar mutantes nulos para un gen específicamente (Singh et al., 2017). Dicha línea fue muy útil para corroborar algunos de los efectos causados por la deficiencia en PFDN5. La ausencia de PFDN5 provocó la ausencia de la subunidad PFDN1, que no pudo ser detectada en la línea mutante. Por el contrario, las subunidades PFDN6 y en menor medida PFDN2, ambas componentes del complejo *prefoldin-like*, fueron menos

afectadas. Curiosamente el contenido de la subunidad PFDN4 también resultó poco alterado por la ausencia de PFDN5 (**Figura R-10B**).

Todo ello sugiere que la deficiencia o ausencia de PFDN5 afecta a los niveles del resto de subunidades del complejo prefoldina, probablemente por su desestabilización, aunque no de forma homogénea.

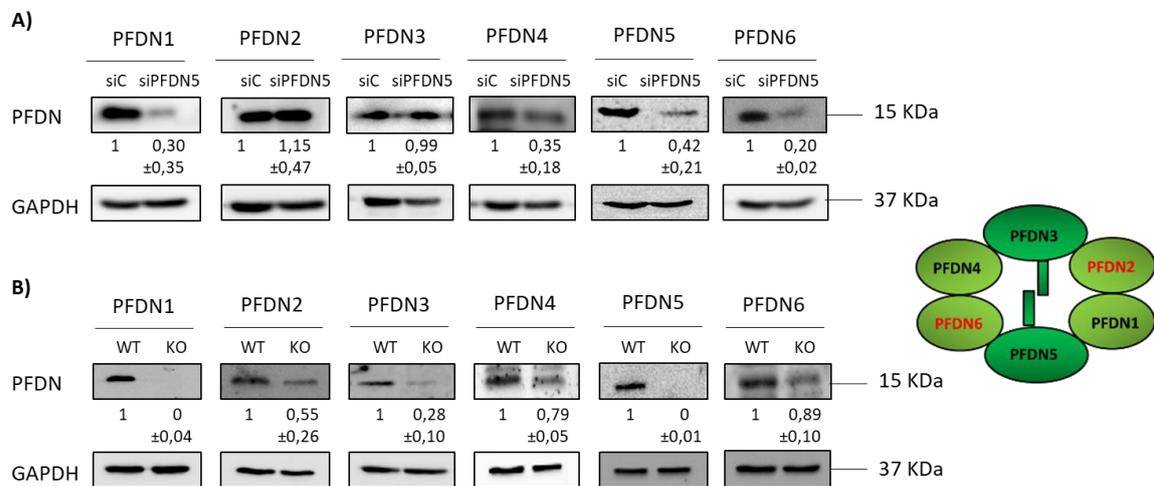


Figura R-10. Tanto la depleción de PFDN5, como la ausencia de esta, reducen los niveles de expresión del resto de subunidades, aunque no de forma homogénea. A partir de los extractos totales obtenidos de células transfectadas con el siPFDN5 o el siControl (**A**) y células carentes de PFDN5 (**B**) se analizaron 40 μ g de estos mediante *western blot*, utilizando los anticuerpos contra las diferentes subunidades de prefoldina y GAPDH como control de carga. A la derecha se muestra un dibujo esquemático de la posición de las diferentes subunidades, dentro del complejo prefoldina. Se muestran en rojo, las subunidades que también forman parte del complejo *prefoldin-like*.

3.2. La ausencia y deficiencia en PFDN5 baja los niveles de expresión de genes de histonas.

Como se mencionó en el apartado 2.5, la deficiencia en PFDN5 provocó una disminución de la expresión de los genes que forman parte de la categoría funcional de “ensamblaje de nucleosomas” a tiempo 0, es decir, antes de la inducción con suero. La mayoría de estos genes codifican histonas, cuya transcripción está regulada de manera dependiente del ciclo celular (Ewen, 2000).

Los niveles de mRNA de 4 de estos genes de histonas (*H2AB*, *H2BB*, *H3J* y *H4B*) fueron comprobados mediante ensayos de RT-qPCR, en los que se utilizó RNA extraído de células transfectadas con siPFDN5 y el mutante carente de PFDN5, que dos días después de ser sembradas, fueron incubadas en medio libre de suero. Los resultados demostraron que la deficiencia en PFDN5 ocasionó que los niveles de 3 de los 4 mRNA analizados disminuyeran ligeramente (**Figura R-11B**), de forma similar a lo detectado en el experimento de RNAseq (**Figura R-11A**). Sin embargo, la ausencia de PFDN5 provocó una caída más extrema en los niveles de todos los mRNA de histonas analizados (**Figura R-11C**).

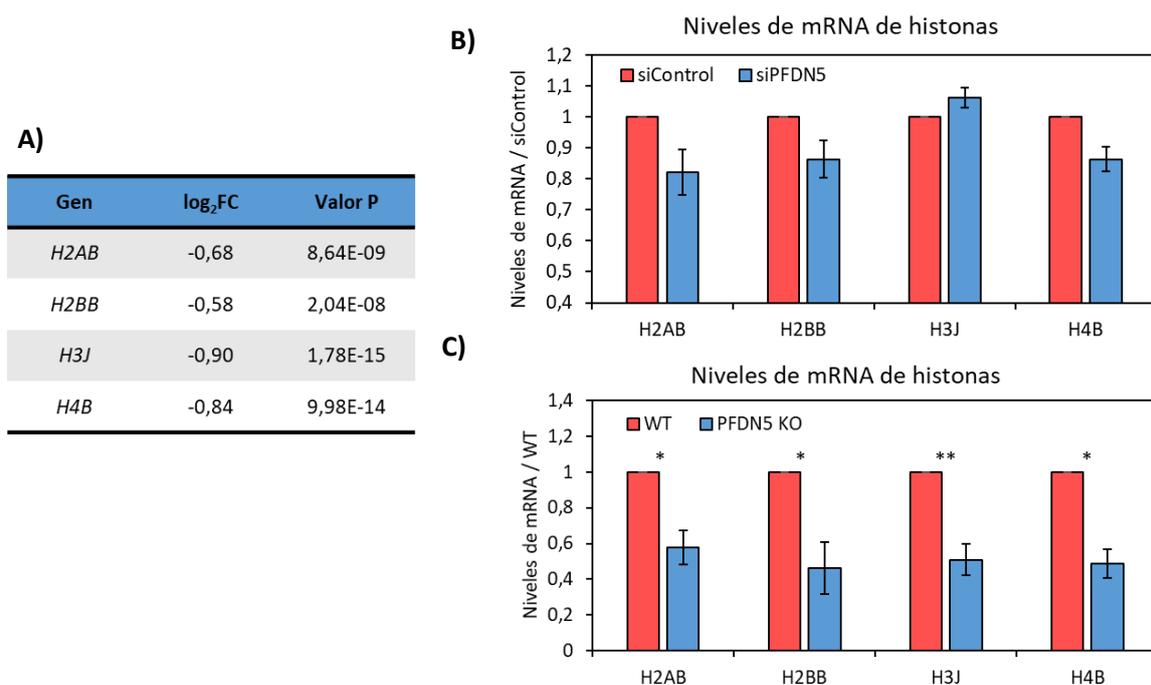


Figura R-11. La ausencia y deficiencia en PFDN5 disminuye los niveles de mRNA de genes de histonas. **A)** Genes de histonas que en los resultados obtenidos del RNA-seq mostraban expresión diferencial en las células transfectadas con siPFDN5 con respecto al control. **C y B)** Se realizaron ensayos de RT-qPCR a partir del RNA extraído de células deficientes en PFDN5 y células sin PFDN5, que crecieron en ausencia de suero durante 48 horas. Posteriormente, se analizaron los niveles de mRNA de los genes que aparecen en la tabla. En la gráfica se representan valores promedio \pm la desviación estándar obtenidos de 3 experimentos diferentes con células deficientes en PFDN5 (**B**) o carentes en PFDN5 (**C**). **t*- Student con valor $P < 0,05$ y ***t*- Student con valor $P < 0,005$.

Estos resultados validan los datos obtenidos en el experimento de RNA-seq y sugieren, por un lado, que PFDN5 podría estar regulando, directa o indirectamente, la expresión de los genes de histonas afectando de este modo a la dinámica del genoma;

y por otro, que el ciclo celular de las células carentes o deficientes en PFDN5 podría estar afectado.

3.3. La ausencia y deficiencia en PFDN5 incrementa los niveles de expresión del gen *p21*.

También nos llamó la atención que, en las mismas condiciones del ensayo (ausencia de suero), otros de los genes diferencialmente expresados fuese el gen *CDKN1A*, el cual incrementó su expresión en las células deficientes en PFDN5. *CDKN1A* codifica la proteína p21, un inhibidor de CDK que, entre otras funciones, detiene el progreso del ciclo celular ante una gran variedad de estímulos (Karimian et al., 2016). Dado que los datos transcriptómicos no solo mostraban que la deficiencia en PFDN5 provoca un descenso en la expresión de los genes de histonas, sino también una sobreexpresión de *p21* (**Figura R-12A**), predecimos que se podrían estar dando cambios en el ciclo celular.

Para corroborarlo, primero se realizaron ensayos de RT-qPCR en los que se utilizaron el mismo RNA extraído de células transfectadas con el siPFDN5 y el mutante carente de PFDN5, que dos días después de ser sembradas, fueron incubadas en medio libre de suero. Los resultados confirmaron que la deficiencia en PFDN5 provocó mayores niveles de mRNA de *p21* (**Figura R-12B**) y que la carencia de esta subunidad ocasionó una subida aún más drástica de este mensajero (entre 6 y 7 veces más que en la línea WT) (**Figura R-12C**).

Tanto los experimentos con siRNA como los realizados con la línea KO se realizaron en ausencia de suero. No sabíamos, por tanto, cómo afecta la ausencia o deficiencia en PFDN5 a la expresión de p21 en condiciones normales, es decir, en células incubadas en presencia de suero todo el tiempo. Para esclarecer si los efectos de la ausencia de PFDN5 sobre la expresión de *p21* estaban mediados por suero, se realizó un ensayo de *Western blot* con extractos totales extraídos de células WT y células PFDN5 KO incubadas en medio completo. Las células deficientes en PFDN5 no presentaba diferencias en el contenido en p21 (**Figura R-12D**), en cambio las células

mutantes carentes de PFDN5 mostraron un aumento del contenido de esta proteína (Figura R-12E).

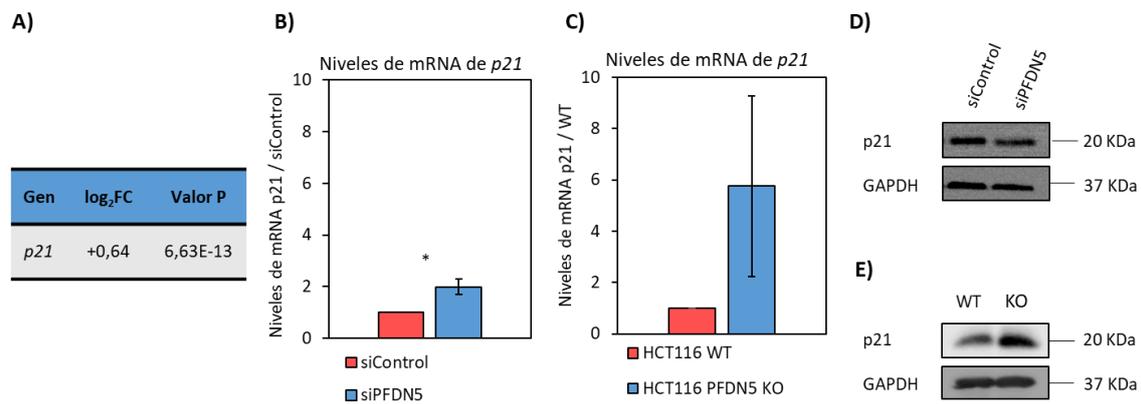


Figura R-12. La deficiencia en PFDN5 aumenta los de niveles del mRNA de *p21*, antes y después de la inducción con suero, y no afecta a sus niveles proteicos. A) Gen *p21* que en los resultados obtenidos del RNA-seq mostraba expresión diferencial en las células transfectadas con siPFDN5 con respecto al control. **B y C)** Se realizaron ensayos de RT-qPCR a partir del RNA extraído de células deficientes en PFDN5 (**B**) y células sin PFDN5 (**C**), que crecieron en ausencia de suero durante 48 horas. Posteriormente, se analizaron los niveles del mRNA de *p21*. En la gráfica se representan los valores promedio \pm la desviación estándar obtenidos de 3 experimentos diferentes. **t-Student* con valor $P < 0,05$ y ***t-Student* con valor $P < 0,005$. **D y E)** A partir de extractos totales obtenidos de células deficientes en PFDN5 (**D**) y carentes de PFDN5 (**E**) que crecieron en condiciones normales de suero se analizaron 20 μ g de estos mediante *western blot*, utilizando los anticuerpos contra *p21* y GAPDH como control de carga.

3.4. La deficiencia y ausencia de PFDN5 provocan ligeros cambios en el ciclo celular.

Hipotetizamos que quizás el origen de la alta expresión de *p21* podría ser que la falta de PFDN5 produjera daños en el DNA. De hecho, algunos de los genes con expresión alterada en ausencia de PFDN5, como *POLM*, están relacionados con reparación de daños en el DNA (**Tabla R-1**). Para averiguarlo, realizamos un ensayo de inmunofluorescencia que nos permitiera determinar el número de focos de γ -H2A.X (véase apartado 8 de materiales y métodos). Los focos de γ -H2A.X son marcadores del DNA dañado en su doble cadena, lo cual induce la expresión de *p21* y con ello la retención de las células en fase G1 (Fragkos et al., 2009). El experimento fue realizado en células transfectadas a las que se les privó de suero durante 48 horas. Los resultados

no mostraron un aumento en el número de focos en las células deficientes en PFDN5 con respecto a las células transfectadas con el siControl (**Figura R- 13**).

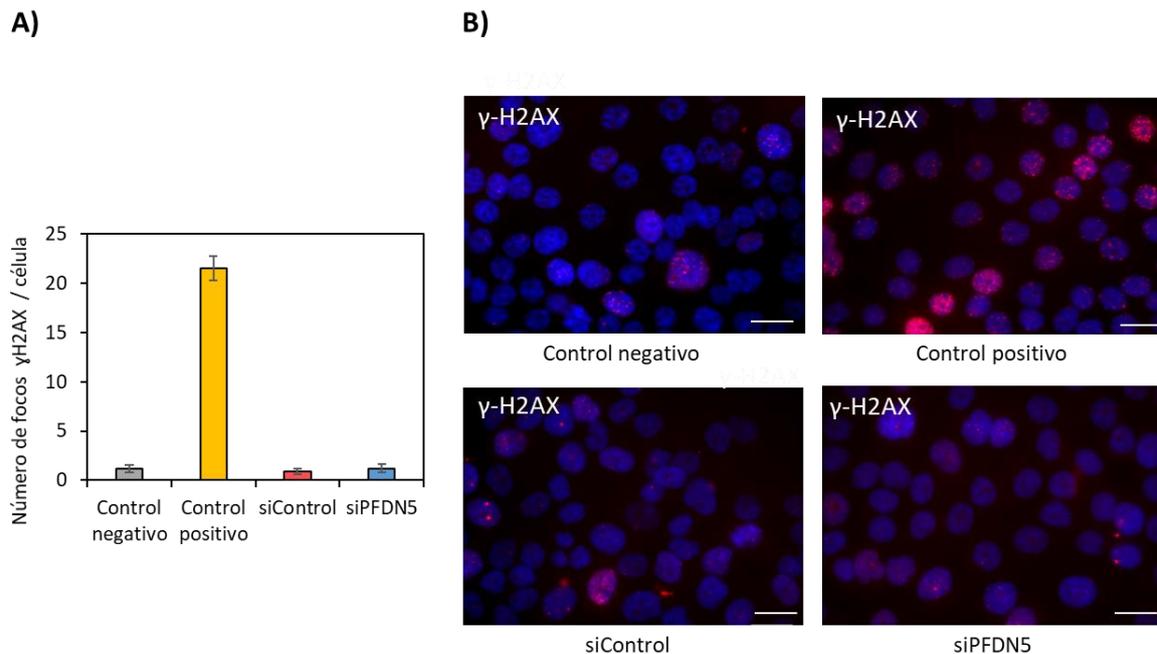


Figura R-13. Los efectos en el ciclo celular no se deben a que se genere daño en el ADN por la deficiencia en prefoldina. A) Histograma del número de focos de γ -H2A.X por célula, de células transfectadas con el siPFDN5 que crecieron sin suero durante 48 horas. El control positivo son células tratadas con etopóxido 20 μ M durante una hora, y el negativo, células tratadas con DMSO durante una hora. Para el análisis se usó el anticuerpo primario anti γ -H2A.X, y el secundario Alexa Fluor 488 (rojo). Para la cuantificación, se contó el número de focos en 30 células. **B)** Imágenes tomadas de las diferentes condiciones en ausencia de suero. Barra = 20 μ m.

Lo siguiente que hicimos fue realizar experimentos de citometría de flujo en células carentes y deficientes en PFDN5 que habían sido incubadas sin suero durante 48 horas (véase apartado 10 de materiales y métodos). La deficiencia en PFDN5 provocó cambios ligeros en el ciclo celular, con un porcentaje menor de células en fase G1, y mayor en fase G2 (**Figura R-13A**). Sin embargo, los resultados obtenidos en la línea mutante fueron distintos a los anteriores: la ausencia de PFDN5 ocasionó un aumento del porcentaje de células en fase G1, y una disminución de éstas en fase S (**Figura R-13B**). Mientras que los efectos en el ciclo celular causados por la deficiencia en PFDN5 no parecían ser debidos a p21, los efectos ocasionados por la ausencia de PFDN5 sí podrían estar mediados por una mayor expresión de p21, pues como se mencionó anteriormente, este factor retiene las células en fase G1 en respuesta a

multitud de estímulos externos (Karimian et al., 2016). Esto sugiere un papel proto-oncogénico para PFDN5, regulando la progresión del ciclo celular a través de p21.

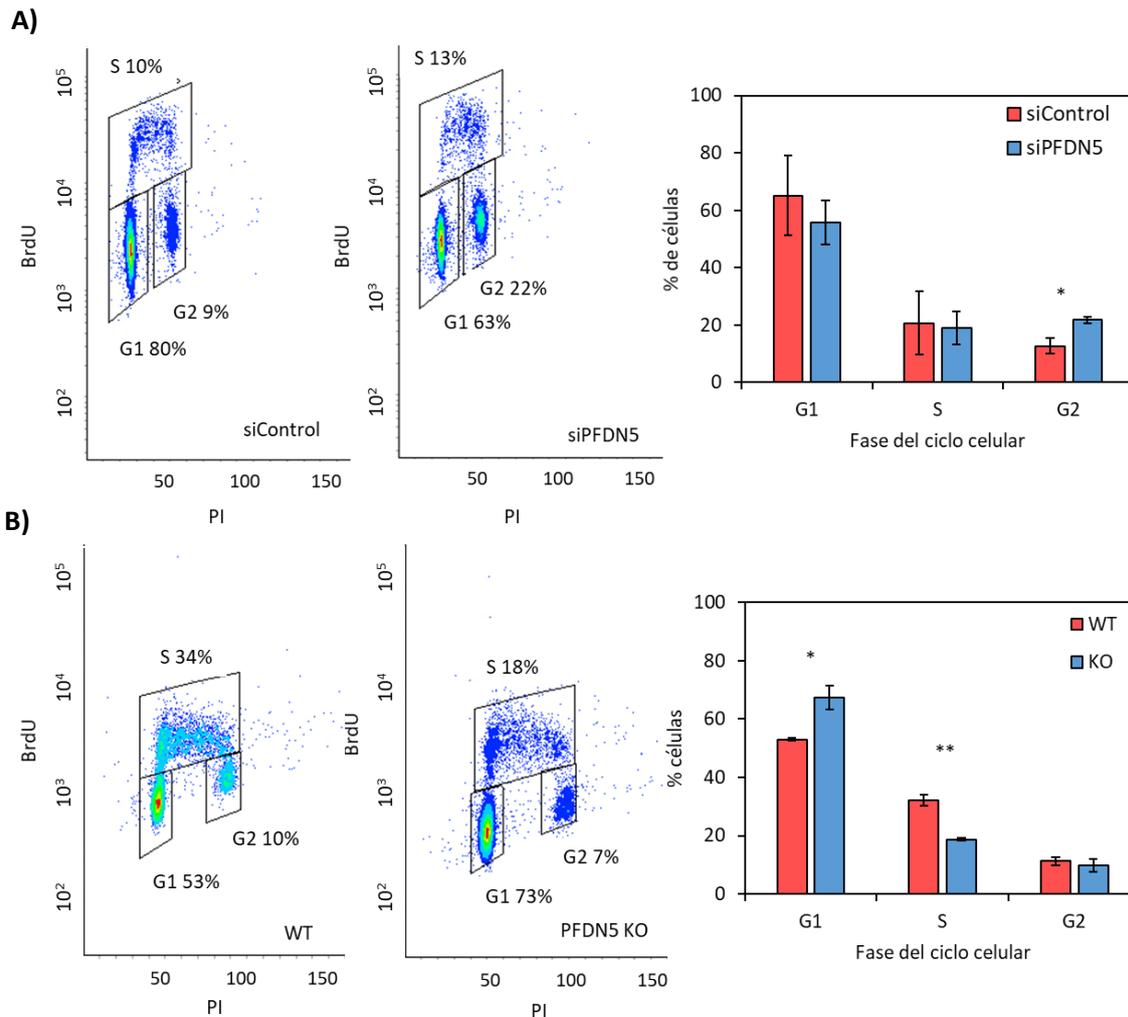


Figura R-14. La ausencia y deficiencia en PFDN5 produce efectos en el ciclo celular. A y B) Distribución en las distintas fases del ciclo celular de células HCT116 transfectadas con el siPFDN5 (A) o mutantes carentes de PFDN5 (B) incubadas en medio sin suero durante 48 horas. Por último, fueron marcadas con BrdU durante 15 minutos. El BrdU fue visualizado marcándolo con anticuerpo anti-BrdU y usando un anticuerpo secundario unido a un fluoróforo. En el eje vertical se visualiza la intensidad emitida por el BrdU y en el horizontal la del Ioduro de Propidio (IP). El experimento se repitió tres veces y los histogramas se representa la cuantificación y el análisis estadístico de los datos obtenidos. **t*- Student con valor $P < 0,05$ y ***t*- Student con valor $P < 0,005$.

4. PREFOLDINA 5 SE ENCUENTRA FÍSICAMENTE ASOCIADA AL GEN P21.

Con todo lo visto hasta ahora, decidimos estudiar si la expresión del gen *p21* podría estar regulada directamente por PFDN5. Dado que en levaduras varias subunidades de prefoldina son capaces de unirse al DNA (Millán-Zambrano et al.,

2013), un posible mecanismo de regulación podría conllevar la unión de PFDN5 al DNA del gen *p21*. Para averiguar si PFDN5 interacciona con el gen *p21*, se realizaron ensayos de inmunoprecipitación de la cromatina (ChIP).

Dado que el anticuerpo comercial anti-PFDN5 utilizado en esta tesis no funciona en ensayos de inmunoprecipitación, tuvimos que abordar el experimento creando dos nuevas líneas estables a partir de la línea HCT116 WT: una expresaba PFDN5 etiquetada a Flag en su extremo carboxi-terminal, y la otra, expresaba solo la etiqueta Flag (control negativo). Flag es un péptido de 8 aminoácidos comúnmente utilizado para ser añadido a proteínas de interés mediante técnicas de DNA recombinante (Einhauer and Jungbauer, 2001). Para la generación de estas líneas, las células fueron transfectadas con el vector de expresión de la proteína Flag fusionada a PFDN5 (pFlag-PFDN5), que contiene el gen que codifica para la proteína Flag bajo control del promotor constitutivo CMV. Como línea control de los experimentos, otra población del mismo tipo celular fue transfectada con el plásmido vacío (pCMV-Tag 2A). Más tarde, se seleccionaron los clones capaces de crecer en medio selectivo suplementado con G418 (marcador de resistencia de los plásmidos mencionados anteriormente). Por último, se comprobó mediante *Western blot* el tamaño de PFDN5 en dichas líneas usando el anticuerpo anti-PFDN5. Tanto el WT como el control negativo mostraron una banda del mismo tamaño, que correspondía a PFDN5. Sin embargo, la línea HCT116 Flag-PFDN5, mostró dos bandas: una correspondía a la proteína endógena y la otra a Flag-PFDN5 (**Figura R-15A**).

Una vez obtenidas las dos líneas, se realizaron cuatro experimentos independientes de ChIP a partir de extractos celulares recogidos de las líneas incubadas en medio completo. En el ensayo se analizaron dos regiones diferentes del gen *p21*: la primera situada en el primer exón y la segunda en el segundo exón (**Figura R-15B**). Los niveles de PFDN5 fueron normalizados primero al *imput*, y luego a una región intergénica del cromosoma V. Con todo ello se detectó una acumulación de PFDN5 asociada al primer exón del gen (**Figura R-15C**). Esto sugería que los cambios de expresión sobre este gen podrían estar mediados por una estrecha relación física entre prefoldina 5 y el primer exón de *p21*.

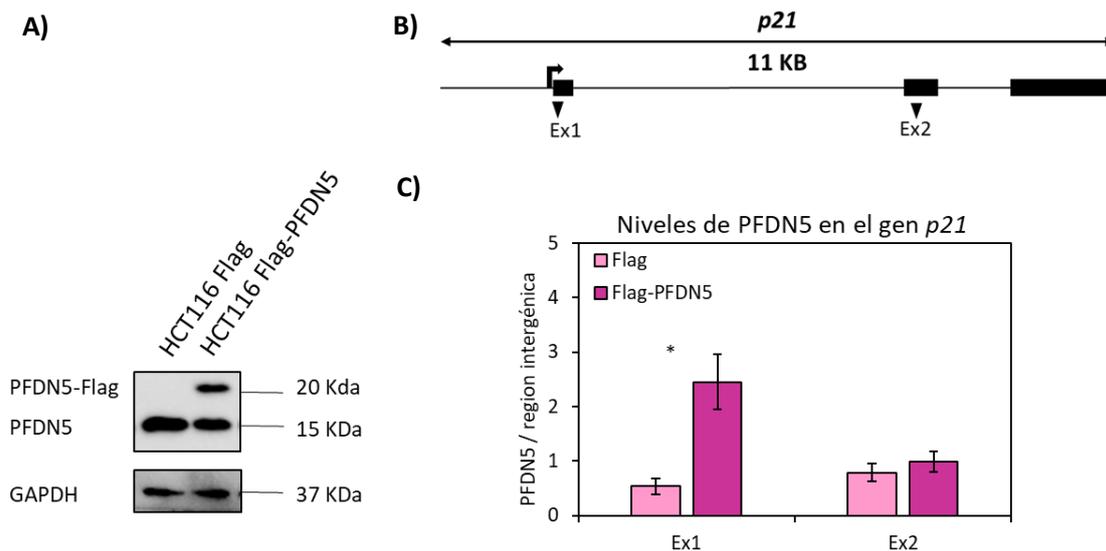


Figura R-15. Prefoldina 5 se encuentra físicamente asociada al exón 1 de p21. (A) *Western blot* realizado a partir de los extractos totales obtenidos de las líneas celulares HCT116 Flag y HCT116 Flag-PFDN5. Se analizaron 40 μ g mediante *western blot*, utilizando los anticuerpos contra PFDN5 y GAPDH como control de carga (B). Dibujo esquemático para representar la localización de los oligonucleótidos utilizados en el ensayo de ChIP. (C) Histograma de barras de los niveles de PFDN5 en el gen p21. En la gráfica se representan los valores promedio \pm la desviación estándar obtenidos de cuatro experimentos diferentes. **t-Student* con valor $P < 0,05$ y ***t-Student* con valor $P < 0,005$.

5. LA DEFICIENCIA EN PFDN5 AFECTA NEGATIVAMENTE A LA INDUCCIÓN POR SUERO DE LOS GENES MÁS LARGOS Y CON MAYOR NÚMERO DE INTRONES.

5.1. Tras 90 minutos de inducción por suero, genes más largos o con mayor número de intrones se expresan menos en las células deficientes en PFDN5.

Hasta el momento, nuestros datos han mostrado que la deficiencia en prefoldina 5, y en menor medida la deficiencia en prefoldina 2, afecta al transcriptoma de células humanas. Estos resultados pueden explicarse en términos regulatorios, dado que la deficiencia de prefoldinas puede originar alteraciones de la fisiología celular que den lugar a señales que activen o repriman la expresión de genes con capacidad de respuesta a la alteración producida. Sin embargo, los estudios realizados en *S. cerevisiae* han puesto de manifiesto que prefoldina contribuye de forma general a la elongación de la transcripción y la dinámica cotranscripcional de la cromatina (Millán-Zambrano et al., 2013). Este dato plantea la posibilidad de que parte de los efectos transcriptómicos observados en las células deficientes en PFDN2 y PFDN5 antes

descritos, sean consecuencia de una contribución similar de las subunidades de prefoldina humana a la elongación transcripcional y los fenómenos asociados a ella.

En nuestro laboratorio se ha demostrado que la ausencia de PFDN1 en *S. cerevisiae* provoca defectos transcripcionales en los genes largos (Millán-Zambrano et al., 2013). Además, numerosos estudios realizados a gran escala han relacionado la expresión génica con el tamaño de los genes (Castillo-Davis et al., 2002; Chiaromonte et al., 2003). Por tanto, decidimos estudiar cómo influye la longitud de los genes en el efecto que causa la deficiencia en PFDN5 en su expresión, antes y después de la inducción con suero. Para ello, se realizó un nuevo análisis bioinformático en el que los genes fueron divididos en 5 quintiles según su longitud. Se excluyeron los genes de expresión altísima, ya que estos genes suelen ser cortos (Castillo-Davis et al., 2002; Chiaromonte et al., 2003) y podrían interferir en el análisis. Los resultados de la **Figura R-16A** muestran que la longitud no tuvo un efecto importante sobre la diferencia de expresión a tiempos 0' y 30'. Sin embargo, a tiempo 90' la ratio de expresión de las células deficientes en PFDN5 respecto a las células control disminuía a medida que aumentaba el tamaño de los genes (**Figura R-16A**). En otras palabras, el efecto de la depleción de PFDN5 sobre la expresión génica 90 minutos después de la inducción por suero es ligeramente dependiente de la longitud del gen.

La longitud de los genes también está directamente relacionada con el número de exones que éstos tienen (Regulapati et al., 2008). Decidimos por tanto repetir el análisis anterior tras clasificar los genes en cinco quintiles en función de su número de intrones. Los resultados volvieron a no mostrar efectos relevantes del número de intrones en los tiempos 0' y 30', pero sí mostraron ratios de expresión menores en las células deficientes en PFDN5 respecto a las células control a medida que aumentaba el número de intrones (**Figura R-16B**). Es decir, a tiempo 90' los genes con más intrones se expresaron menos en las células deficientes en PFDN5, que los genes con menor número de intrones.

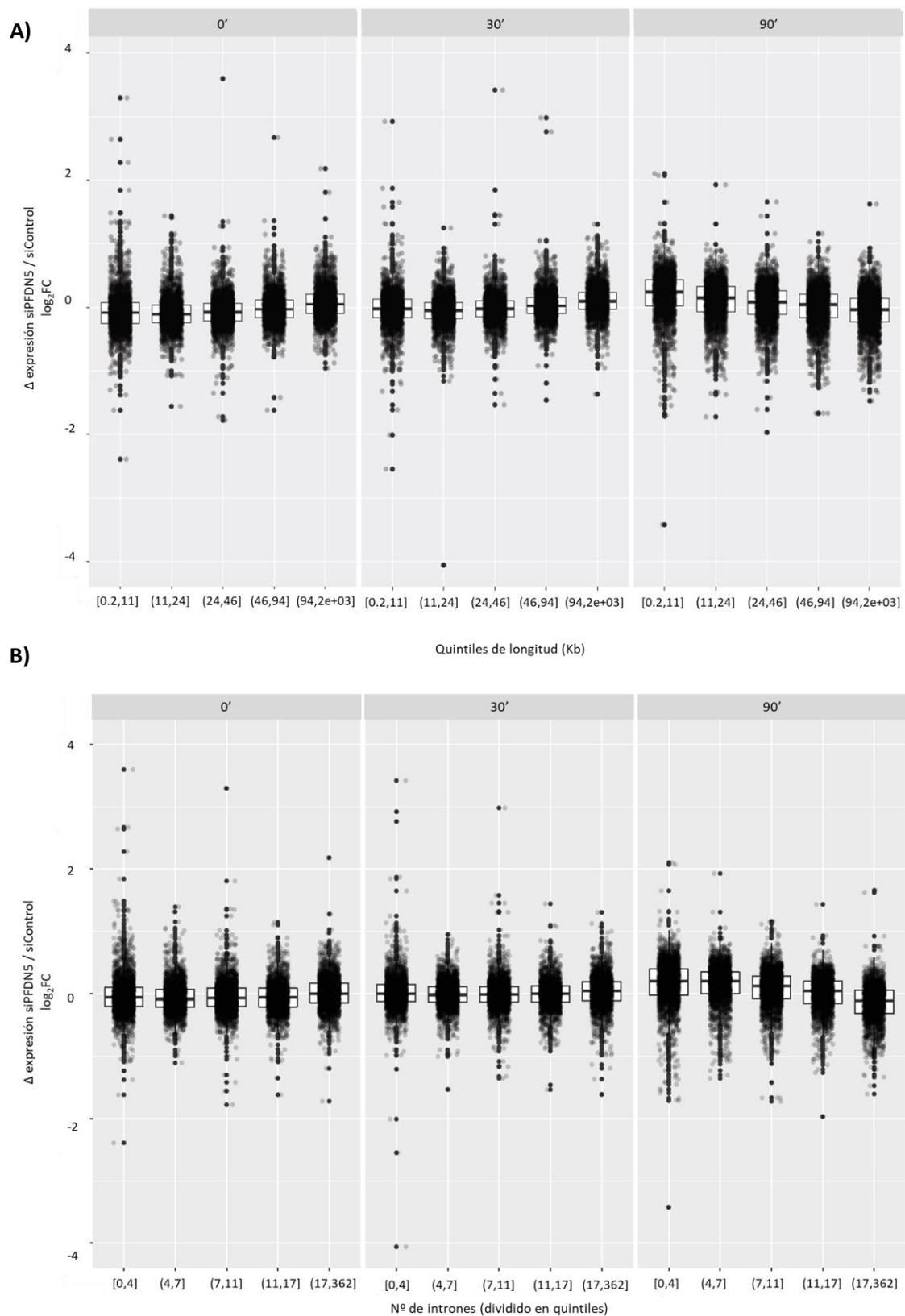


Figura R-16. La longitud de los genes y el número de intrones afectan a las diferencias de expresión causadas por la deficiencia en PFDN5. Se representa la variación de las diferencias de expresión siPFDN5 / siControl (eje vertical) con respecto a la longitud de los genes (**A**) y el número de intrones (**B**) a tiempos 0', 30' y 90' tras la inducción con suero.

5.2. El tamaño de los genes y el número de intrones ejercen un efecto sinérgico en el nivel de inducción tras la adición de suero en las células deficientes en PFDN5.

Puesto que la longitud de los genes correlaciona positivamente con el número de exones (Regulapati et al., 2008), y por tanto de intrones, ambos resultados podrían deberse únicamente a una de las dos variables. Alternativamente ambas variables podrían ser independientes, en cuyo caso esperaríamos un efecto sinérgico por longitud del gen y por número de intrones. Con el objetivo de dilucidar este aspecto, decidimos centrarnos en estudiar los genes que 90' después de la adición del suero se inducían en las células control y analizar si dicha inducción se veía afectada por la deficiencia de PFDN5. Los resultados nos mostraron que el número de intrones intensificaba de forma muy significativa el efecto negativo de la longitud sobre el nivel de inducción, pero esto ocurría exclusivamente en las células deficientes en PFDN5, no así en las células control. Estos resultados quedan reflejados gráficamente en la **figura R-17**, donde se representa el efecto de la longitud en los genes con un nivel medio de intrones frente a aquellos con un número de intrones inferior o superior a la media en una desviación estándar.

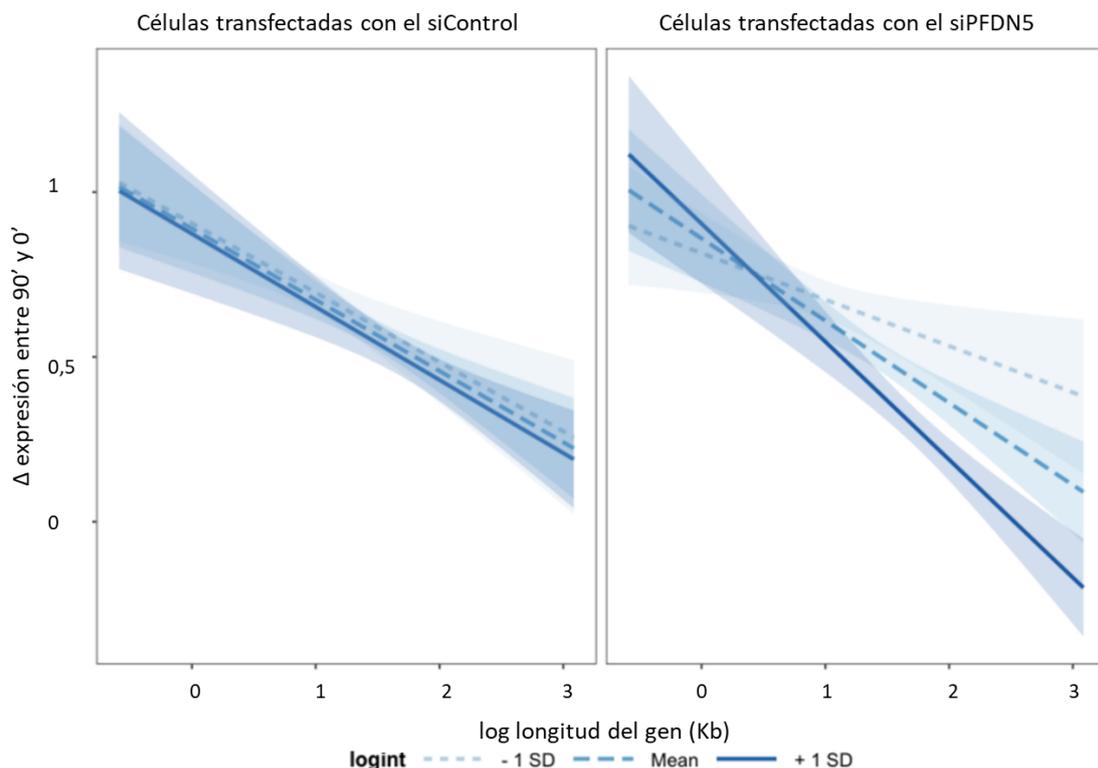


Figura R-17. La longitud de los genes y el número de intrones ejercen un efecto sinérgico sobre el nivel de inducción de los genes de las células transfectadas con el siPFDN5. A y B) Se seleccionaron los genes que se inducen 90' después de la adición del suero en la población control. Se dividieron en tres grupos según el número de intrones: El primer grupo contiene los genes con un número de intrones inferior a la media menos la desviación estándar (- 1 SD). El segundo grupo contiene los genes con un número de intrones dentro de un rango de una desviación estándar en torno a la media (Mean). El tercer grupo contiene los genes cuyo número de intrones es mayor que la media más la desviación estándar (+ 1 SD). Se representa el nivel de inducción de estos genes en la población transfectada con el siControl **(A)** y la población transfectada con el siPFDN5 **(B)** (eje vertical), frente al logaritmo de la longitud de los genes (eje horizontal).

Los resultados anteriores indican una relación entre PFDN5 y la expresión de genes según su longitud y número de intrones. Este hecho apoya la existencia de un efecto positivo de PFDN5 sobre la elongación transcripcional y un fenómeno mecánicamente asociado como es el procesamiento del pre-mRNA.

6. LA DEFICIENCIA EN PFDN5 Y PFDN2 AUMENTA LOS NIVELES DE PRE-mRNA.

6.1 La deficiencia en PFDN5 y PFDN2 provoca variaciones en la ratio exónica antes y después de la inducción con suero.

Ya que las células deficientes en PFDN5 parecían tener una mayor dificultad para expresar genes largos con gran número de intrones, nos planteamos si ello implicase que dichas células presentaran mayor dificultad para procesar los intrones de los transcritos inmaduros.

En una primera aproximación para dilucidar esta cuestión, se clasificaron las secuencias obtenidas en el análisis masivo del RNA total, según si mapeaban dentro de las coordenadas de los exones (lecturas exónicas), o lo hacían dentro de los genes completos (lecturas totales, que son la suma de lecturas intrónicas y exónicas). Esto se hizo para poder comparar las ratios exónicas (lecturas exónicas / lecturas totales) de los diferentes tratamientos con siRNA. En otras palabras, un incremento de la ratio exónica significa que el número de intrones está disminuyendo, es decir, el pre-mRNA está madurando de forma más efectiva; y una disminución de la ratio exónica indica que el número de intrones, con respecto al de exones, está aumentando, es decir, la maduración del pre-mRNA está siendo menos efectiva.

De este modo se detectó que la ratio exónica antes de la inducción con suero era mayor en las células deficientes en PFDN5 que en las células control y deficientes en PFDN2 (que presentaron ratios similares al control) (**Figura R-18**). Sin embargo, después de la inducción con suero el fenotipo se fue revirtiendo poco a poco: mientras que a tiempo 30' las ratios exónicas de las 3 condiciones fueron similares, a tiempo 90' disminuyó la ratio exónica en las células deficientes en PFDN5, y algo menos en las células deficientes en PFDN2, y, por el contrario, aumentó la del control (**Figura R-18**).

Estos resultados indican que, mientras en las células control la eficiencia de *splicing* aumenta tras la inducción con suero, en las células deficientes en PFDN5 disminuye.

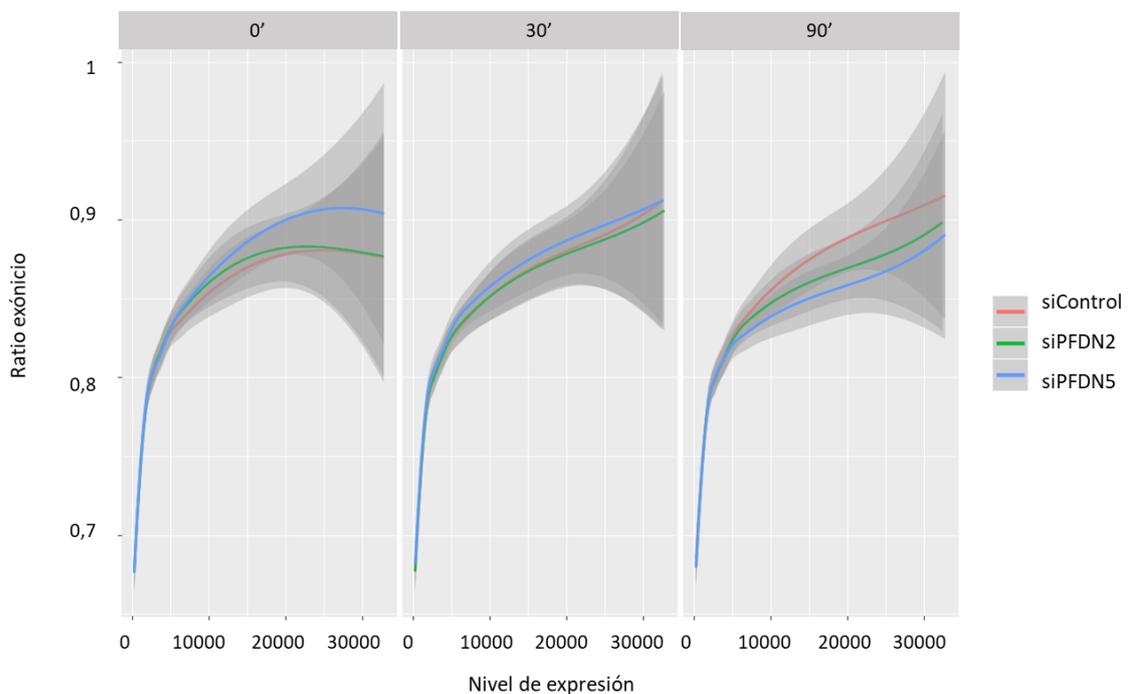


Figura R-18. La deficiencia PFDN5 provoca cambios en la ratio exónica antes y después de la inducción con suero. Se representa la ratio exónica (eje vertical) frente al nivel de expresión de los genes (eje horizontal) de las células transfectadas con el siControl (rojo), siPFDN2 (verde) y siPFDN5 (azul) a tiempos 0, 30 y 90 minutos tras la inducción con suero.

6.2. Los cambios en la ratio exónica ocasionados por la deficiencia en PFDN5 se asocian al nivel de expresión de los genes.

Además de detectar cambios en las ratios exónicas antes y después de la inducción con suero, observamos que dicho fenotipo era más prominente en genes

con mayor nivel de expresión, e imperceptible en aquellos que se expresan poco (**Figura R-18**). Para indagar más en este aspecto, y en cómo era el cambio de la ratio exónica tras activar las células con suero, se realizó un segundo análisis en el que se comparó cómo variaba la ratio exónica 30 y 90 minutos después de la inducción con suero ($(\text{lecturas exónicas} / \text{lecturas totales})^{30'} - (\text{lecturas exónicas} / \text{lecturas totales})^{0'}$) con respecto al nivel de expresión, esta vez dividido en tres categorías: genes de expresión baja (menos de 180 lecturas), media (entre 180 y 33000 lecturas) y alta (más de 33000 lecturas).

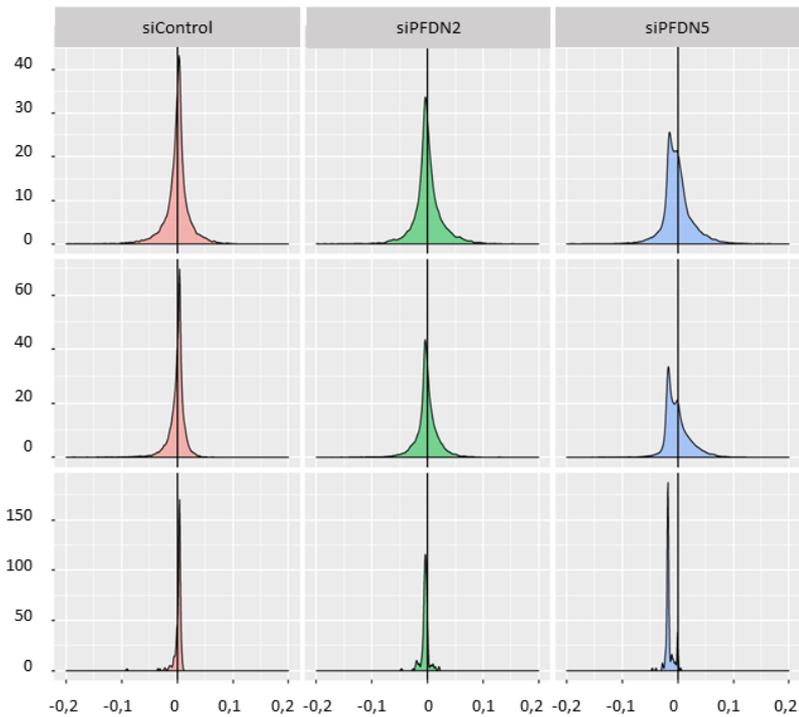
Los resultados mostraron que, 30 minutos después de la inducción con suero, la deficiencia en PFDN2 no ocasionó cambios en la ratio exónica respecto al tiempo 0', salvo una ligera disminución de esta en los genes de alta expresión (**Figura R-19A**). Sin embargo, la deficiencia en PFDN5 provocó una pequeña disminución de la ratio exónica mucho más frecuente entre los genes de expresión media y alta (**Figura R-19A**). En las células control únicamente se detectó un ligero aumento de la ratio exónica en los genes de alta expresión (**Figura R-19A**).

90 minutos después de la inducción con suero, la ratio exónica de las células deficientes en PFDN2 disminuyó ligeramente en la práctica totalidad de los genes de alta expresión y con menor frecuencia en el resto (**Figura R-19B**); la de las células deficientes en PFDN5 disminuyó notablemente en todos los genes, siendo más frecuente el cambio cuanto mayor era el nivel de expresión (**Figura R-19B**). Sin embargo, la ratio exónica de las células control se incrementó notablemente, sobre todo en los genes de alta expresión (**Figura R-19B**).

Por tanto, los datos de estos análisis nos aportaron las siguientes ideas: en primer lugar, que las células HCT116 mejoran la eficiencia de maduración del pre-mRNA tras la inducción con suero; en segundo lugar, que la deficiencia en PFDN5 empeora progresivamente la maduración del pre-mRNA tras la adición de suero en el medio de cultivo de células deprivadas de él; en tercer lugar, que la deficiencia en PFDN2 presenta, en rasgos generales, una situación intermedia entre las células control y las deficientes en PFDN5; y en cuarto lugar, que estas alteraciones en el procesamiento de intrones parece estar asociada al nivel de expresión de los genes.

A)

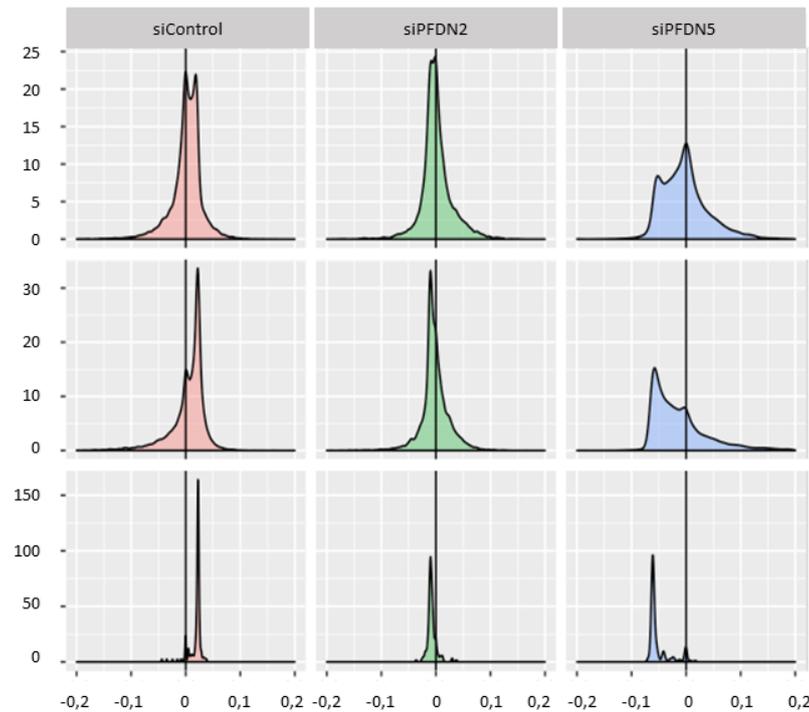
Nivel de expresión:

Bajo:
(<180 reads)Medio:
(180-33000 reads)Alto:
(>33000 reads)

Cambio de la ratio exónica 30' después de la inducción suero

B)

Nivel de expresión:

Bajo:
(<180 reads)Medio:
(180-33000 reads)Alto:
(>33000 reads)

Cambio de la ratio exónica 90' después de la inducción suero

Figura R-19. La deficiencia en PFDN5 empeora la maduración del pre-mRNA progresivamente tras la inducción con suero, mientras que en el control se produce el efecto contrario. A y B) Se representa la variación de la ratio exónica 30 minutos (A) y 90 minutos (B) después de la inducción con suero, respecto al tiempo cero. Los genes se dividieron en tres grupos según su nivel de expresión. El eje vertical representa el número de genes; el eje horizontal representa el cambio de la ratio exónica.

Siguiendo en la misma línea de investigación, decidimos estudiar si existía correlación entre la variación de la ratio exónica 90 minutos después de la inducción con suero de las células control y de las deficientes en PFDN5. En otras palabras, quisimos ver si los genes cuya eficiencia de *splicing* cambiaba tras la inducción con suero en la condición control eran los mismos que en la condición problema. Los genes fueron clasificados en 5 grupos, según su nivel de expresión en la población control. De este modo vimos que el índice de determinación (R^2) se hacía mayor conforme aumentaba el nivel transcripcional de los genes, es decir, cuanto mayor es el aumento de la ratio exónica tras la inducción por suero en las células control, mayor es también el decremento de ésta en las células deficientes en PFDN5, estando este fenómeno asociado al nivel de expresión de los genes (**Figura R-20**).

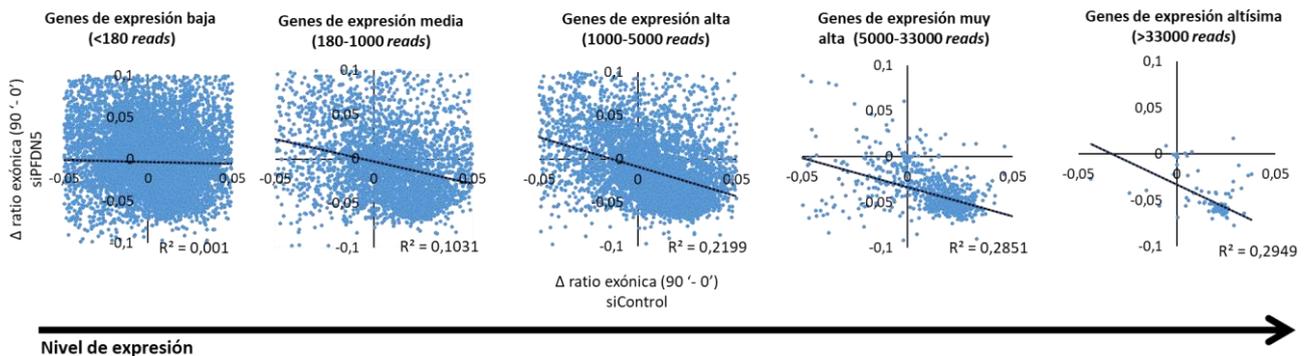


Figura R-20. El decremento de la ratio exónica en respuesta al suero provocado por la deficiencia en PFDN5 correlaciona con el incremento de esta ratio en las células control. Gráficas de dispersión donde se representan la variación de la ratio exónica 90 minutos después de la inducción con suero de las células control (eje horizontal) frente a la de las células deficientes en PFDN5 (eje vertical). Los datos se dividieron en 5 grupos según el nivel de expresión de los genes en la población control. Se representan, además, la línea de tendencia y el índice de determinación de cada gráfica.

6.3. Validación de las alteraciones de *splicing* observadas tras la inducción por suero en células carentes de PFDN5.

Para validar las alteraciones del *splicing* de los transcritos inmaduros observadas tras la inducción por suero en células deficientes en PFDN5 en los experimentos de RNA-seq, analizamos el fenómeno en la línea celular que carece de PFDN5. Para ello, se siguió el flujo de trabajo descrito en la **Figura R-3** (obviando el paso de las transfecciones de las células con los siRNAs) y se estudiaron los niveles de pre-mRNA

mediante ensayos de RT-qPCR en dos regiones diferentes de algunos de los genes que presentaban diferencias más prominentes en los experimentos de RNA-seq (**Figura R-21**). En este tipo de experimentos se usan dos parejas de oligonucleótidos: una mide la transcripción de un determinado intrón (pre-ARNm) (**Figura M-1A**, apartado 12.2 de materiales y métodos) y la otra, la eliminación de este (ARNm) (**Figura M-1B**, apartado 12.2 de materiales y métodos). Los datos del ARN inmaduro se normalizaron primero al maduro, y a continuación, al tiempo 0'.

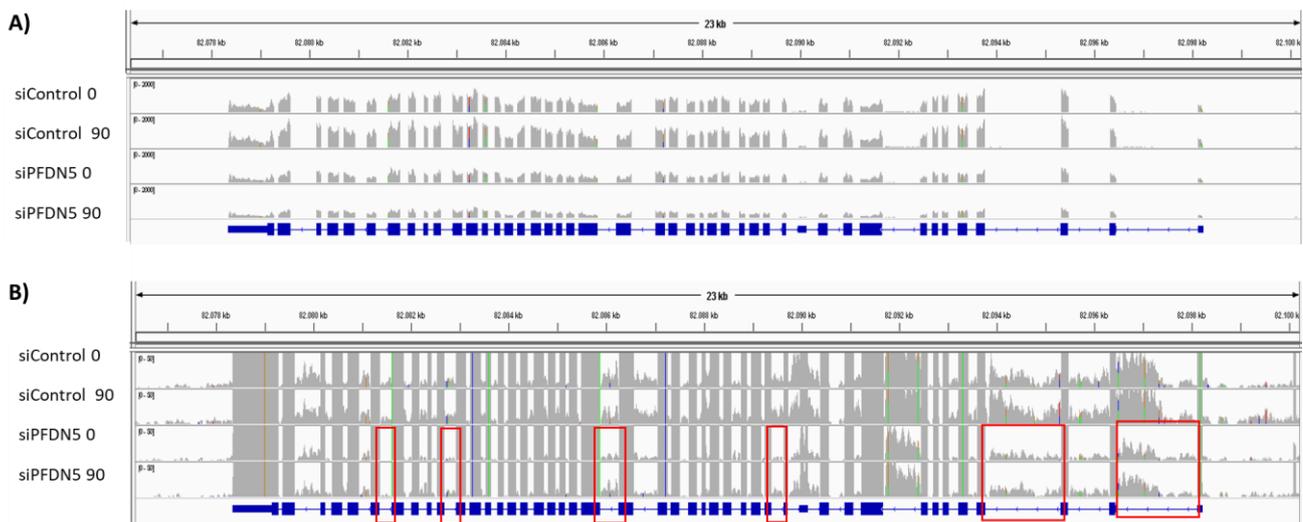


Figura R-21. Ejemplo de gen (*FASN*) con *splicing* del pre-mRNA afectado por la deficiencia de PFDN5 tras la inducción por suero. A y B) Se representa la distribución de las lecturas obtenidas en uno de las réplicas del ensayo de RNA-seq en células control y células deficientes en PFDN5 privadas de suero y 90' después de la adición de éste. Se muestran dos escalas. En **(A)** puede observarse que la inducción por suero no produce variaciones importantes en el número de lecturas exónicas. En **(B)** se aprecia un aumento de lecturas intrónicas tras añadir suero en las células deficientes en PFDN5, no así en las células control. Los recuadros rojos resaltan algunos intrones donde el cambio es especialmente claro.

Los resultados en las células control mostraron leves incrementos de los niveles de pre-mRNA 30 minutos después de la inducción por suero, que disminuían a los 90 minutos después de ésta (**Figura R-22**). Por el contrario, los niveles de pre-mRNA en las células carentes de PFDN5 aumentaron aún más 90 minutos después de la adición del suero (**Figura R-22**).

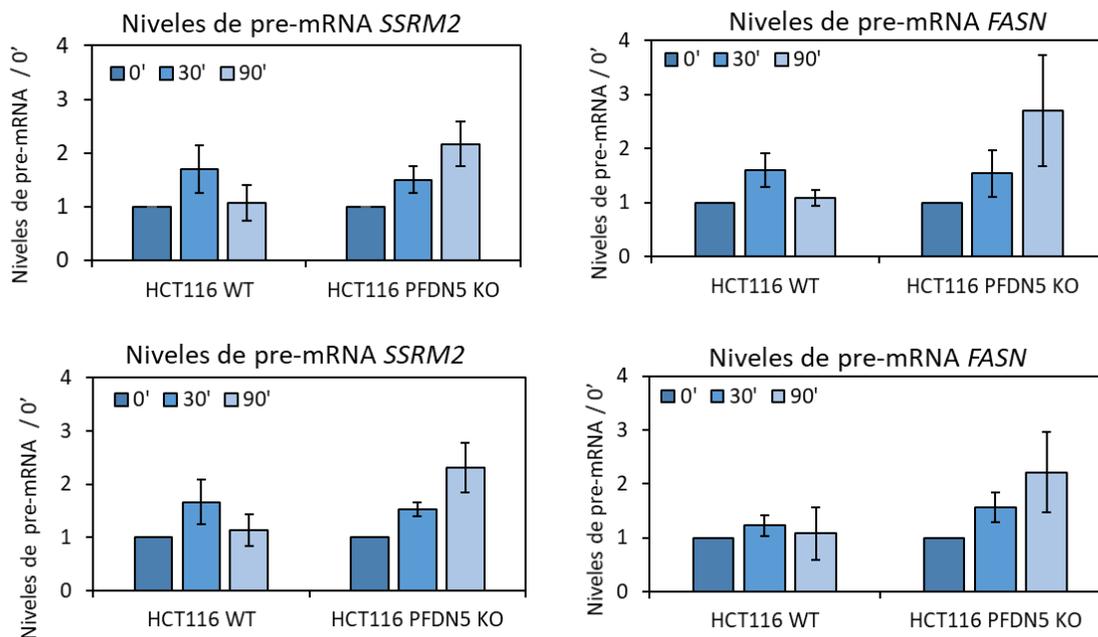


Figura R-22. La ausencia de PFDN5 provoca la acumulación de pre-mRNA tras inducir las células con suero. Se realizaron ensayos de RT-qPCR a partir del ARN extraído de células WT y célula PFDN5 KO de 3 réplicas biológicas ensayadas del mismo modo que el ARN-seq, y se determinó el nivel de pre-mRNA en dos regiones diferentes los genes *SSRM2* y *FASN*. En las gráficas se representan los valores promedio \pm la desviación estándar obtenidos de 3 experimentos diferentes. Los datos del ARN inmaduro se normalizaron primero al maduro, y a continuación, al tiempo 0'. En ninguno de los casos las diferencias medidas tuvieron significación con valor $P < 0,05$, pero en todos ellos se observó la misma tendencia de acumulación de pre-mRNA a tiempo 90' en las células carentes de PFDN5.

7. LA DEFICIENCIA EN PFDN5 AFECTA SUTILMENTE A LA MADURACIÓN ALTERNATIVA DEL PRE-mRNA.

Hasta el momento, nuestros datos mostraban que la deficiencia en prefoldina 5 afectaba a la eficiencia del pre-mRNA *splicing*. Nos preguntamos si ello estaría ocasionando también defectos en los procesos de *splicing* alternativo.

En una primera aproximación para responder a esta pregunta, se estudió si la deficiencia en PFDN5 afectó a la frecuencia con la que ocurren diferentes eventos de *splicing* alternativo utilizando para ello la herramienta bioinformática desarrollada por el Dr. Alamancos (Alamancos et al., 2015) (**Figura R-23A**). Este análisis bioinformático de los datos del RNA-seq nos permitió comparar la frecuencia de un determinado evento de *splicing* alternativo entre las células tratadas con el siRNA control y aquellas tratadas con el siRNA de PFDN5. Quisimos ver si la deficiencia en PFDN5 aumentaba o

disminuía la frecuencia de los distintos transcritos derivados de *splicing* alternativo. Para cada una de estas posibilidades se usó una aproximación distinta.

Para ver si disminuía la frecuencia de eventos de *splicing* alternativo, se tomaron en cuenta solo aquellos eventos que ocurrían en al menos un 10 % de los transcritos en el siRNA control. Sobre este total de eventos, el porcentaje de aquellos que disminuían su frecuencia en las células problema en al menos un 10 % está representado en la **Figura R-23B**. Por otro lado, para ver si aumentaba la frecuencia en las células problema, se tuvieron en cuenta todos los posibles eventos de *splicing* alternativo y se contaron aquellos cuya frecuencia sube al menos un 10% en las células transfectadas con el siPFDN5 (**Figura R-23C**).

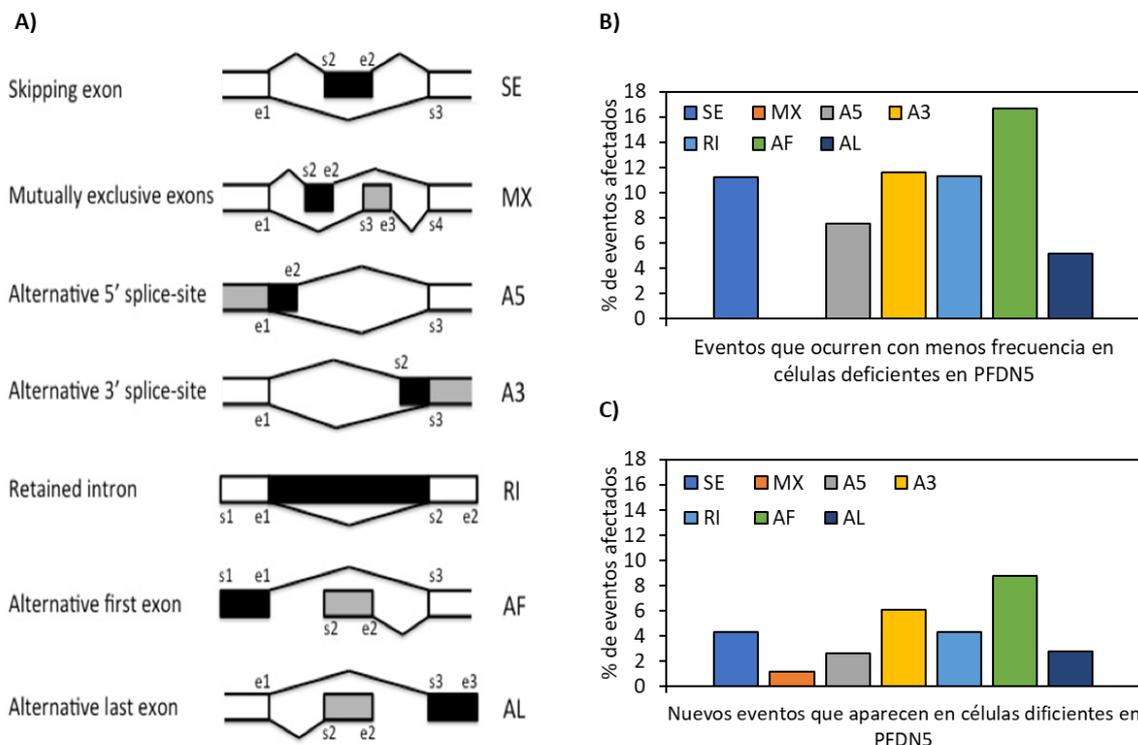


Imagen adaptada de Alamancos et al., 2015.

Figura R-23. La deficiencia en PFDN5, ocasionó ligeros cambios en las frecuencias con la que ocurren diferentes eventos de *splicing* alternativo, sin que ninguna de éstas estuviese particularmente afectada. **A)** Dibujo esquemático de los diferentes eventos de maduración del pre-mRNA. **B)** Histograma de barras donde se representa el porcentaje de eventos de *splicing* alternativo que ocurren con menos frecuencia en células deficientes en PFDN5. **C)** Histograma de barras donde se representan el porcentaje de la frecuencia con que aparecen nuevos eventos de *splicing* alternativo en células deficientes en PFDN5.

La disminución de los niveles de PFDN5 ocasionó una variación en la frecuencia de los diferentes eventos de *splicing* alternativo de forma general, es decir, ninguna modalidad de maduración alternativa del pre-mRNA se encontraba particularmente afectada en las células deficientes en PFDN5 (**Figura R-23B**). Estos datos sugieren que la carencia de PFDN5 provoca disfunciones en el procesamiento del pre-mRNA que tienen un impacto general en los fenómenos de *splicing* alternativo.

Decidimos asimismo realizar un ensayo funcional que nos permitiera estudiar *in vitro* cómo impacta la ausencia de prefoldina 5 sobre la regulación de un determinado evento de *splicing* alternativo. Para ello, se decidió realizar experimentos de minigenes, los cuales se llevaron a cabo en el laboratorio del Dr. Carles Suñé (Instituto de Parasitología y Biomedicina López Neyra, Granada).

Los minigenes son vectores de expresión eucarióticos que contienen una región con maduración alternativa y secuencias genómicas flanqueantes, y por ello, constituyen una buena herramienta para estudiar la regulación del *splicing* alternativo (Cooper, 2005). El minigen reportero de *splicing* alternativo utilizado en este estudio fue el del gen *Fas* (**Figura R-24A**). El receptor de muerte celular Fas/CD95 para el que codifica este gen, puede ser procesado alternativamente para excluir el exón 6 de su mRNA dando lugar a una forma soluble cuya función es inhibir la apoptosis (Cascino et al., 1995; Cheng et al., 1994).

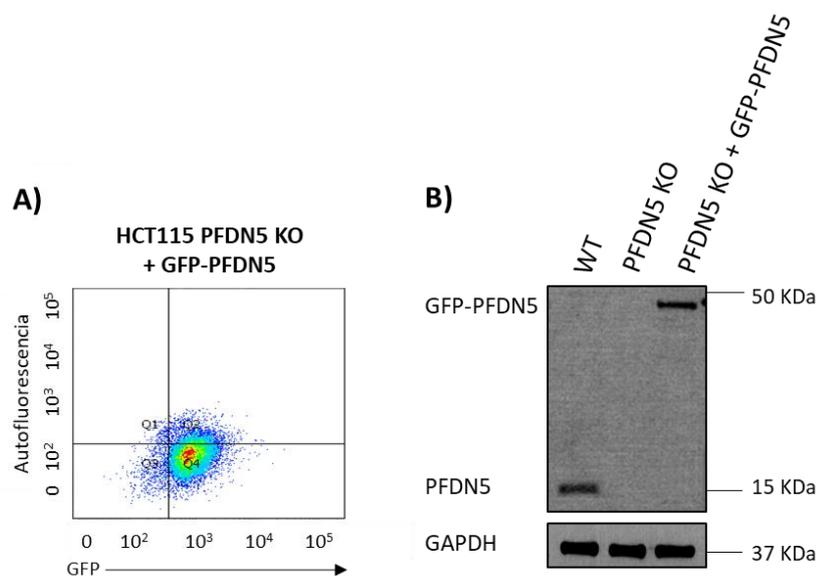


FIGURA R-24. Caracterización de la línea estable HCT116 PFDN5 KO + GFP-PFDN5. A) Análisis por citometría de flujo de clon seleccionado. **B)** Análisis por *Western Blot* del contenido de PFDN5 en el clon seleccionado.

Las líneas utilizadas en este ensayo fueron la WT, PFDN5 KO y una nueva línea “rescate” que se realizó a partir de la anterior. Para la generación de ésta, células HCT116 PFDN5 KO fueron transfectadas con el vector pEGFP-PFDN5 utilizado antes. Posteriormente, los clones capaces de crecer en medio selectivo suplementado con G418 fueron seleccionados. Por último, se analizó el carácter fluorescente de varios de los clones seleccionados por citometría de flujo (**Figura R-24A**), se eligió un candidato fluorescente y se comprobó mediante *Western blot* que expresaba PFDN5 (**Figura R-24B**).

Una vez extraído el RNA de las diferentes líneas transfectadas con la construcción reportera, se analizó el resultado del *splicing* de los minigenes por RT-PCR, cuyos productos fueron resueltos en geles de agarosa (véase apartado 3.8 de materiales y métodos). Los resultados obtenidos mostraron que la ausencia de PFDN5 generó algo más de cantidad de la isoforma corta de *Fas* (aquella que no contiene el exón 6). En cambio, cuando PFDN5 está presente, tanto en la línea WT como en la línea “rescate”, se observó un ligero incremento de la isoforma larga de *Fas* (**Figura R-25B**). Aunque las diferencias encontradas son sutiles y no resultaron estadísticamente significativas (t-sudent, $p < 0,05$), el rescate del fenotipo por parte de la línea que expresa EGFP-PFDN5 parece indicar que el efecto es específico de la falta de PFDN5.

En resumen, estos resultados indican que PFDN5 influye, aunque de forma sutil, en *splicing* alternativo.

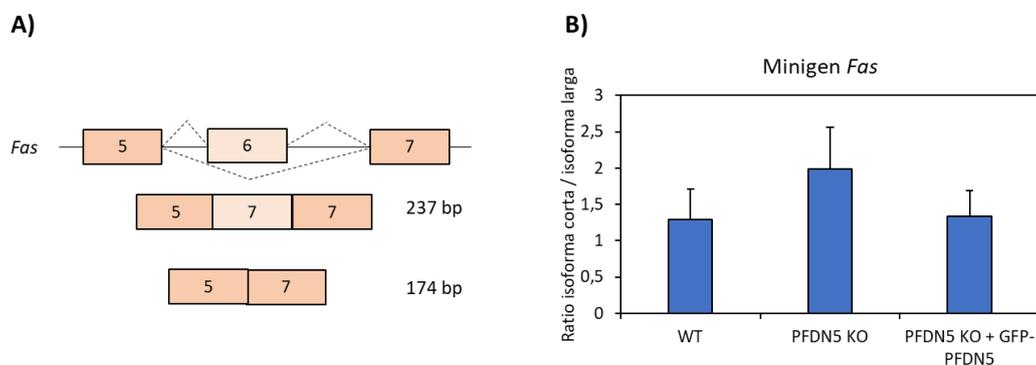


FIGURA R-25. Los efectos de la ausencia de prefoldina 5 sobre el *splicing* alternativo son muy leves. A) Dibujo esquemático de las isoformas alternativas del minigen *Fas*. **B)** Histograma de barras de los ratios isoforma corta / isoforma larga del minigen *Fas*. En la gráfica se representan los valores promedio \pm la desviación estándar obtenidos de cinco experimentos diferentes.

8. LA DEFICIENCIA EN PFDN5 OCASIONA DEFECTOS EN EL ACOPLAMIENTO DEL *SPLICING* DEL pre-mRNA A LA TRANSCRIPCIÓN.

Varios de los resultados ya citados apuntan a que la deficiencia en PFDN5 podría causar alteraciones tanto en la transcripción como en *splicing* del pre-mRNA. Es más, los resultados mostrados en la **figura R-17** indican que la influencia de PFDN5 sobre la inducción de la expresión génica en respuesta al suero depende tanto de la longitud de los genes como del número de intrones que tengan, lo que sugiere que las alteraciones en el *splicing* del pre-mRNA causadas por la deficiencia de PFDN5 podrían estar teniendo lugar en el contexto de la elongación transcripcional. Con todo lo visto hasta el momento, nos preguntamos si los defectos en *splicing* podrían deberse a una alteración del acoplamiento de éste a la transcripción, un fenómeno bien conocido (véase apartado 1.3 de la introducción de esta tesis). Un desacoplamiento de estos procesos podría provocar un retraso en el *splicing* con el consiguiente aumento de los niveles de RNA inmaduro y eventualmente el fallo en la formación de una molécula madura de RNA que impactaría negativamente en los niveles de expresión del gen.

Para averiguarlo, analizamos la aparición del pre-mRNA a lo largo de un gen modelo para este tipo de estudios, el gen *CTNNB1* (Singh and Padgett, 2010). Para conseguir tener una población sincronizada de RNAPIIs, las células se trataron previamente con DRB, un inhibidor reversible de la fosforilación dependiente de P-TEFb en la Ser2 del CTD que provoca que las RNAs polimerasas II se detengan al inicio de la fase de elongación (Price, 2000). El inhibidor es reversible, lo que permite a las células continuar la transcripción una vez eliminada la droga. Los niveles de pre-mRNA se analizaron a diferentes tiempos después del lavado del DRB en varias regiones del gen. Se utilizaron dos parejas de oligonucleótidos por cada región analizada, para así medir tanto la aparición de un intrón en el RNA naciente, como su eliminación por *splicing* (**Figura R-26A**) (véase apartado 12.2 de materiales y métodos). El experimento permite además determinar la velocidad de avance de la RNAPII a lo largo del gen. Los ensayos se realizaron tanto en células control (transfectadas con el siControl) como en células deficientes en PFDN5 (transfectadas con el siPFDN5).

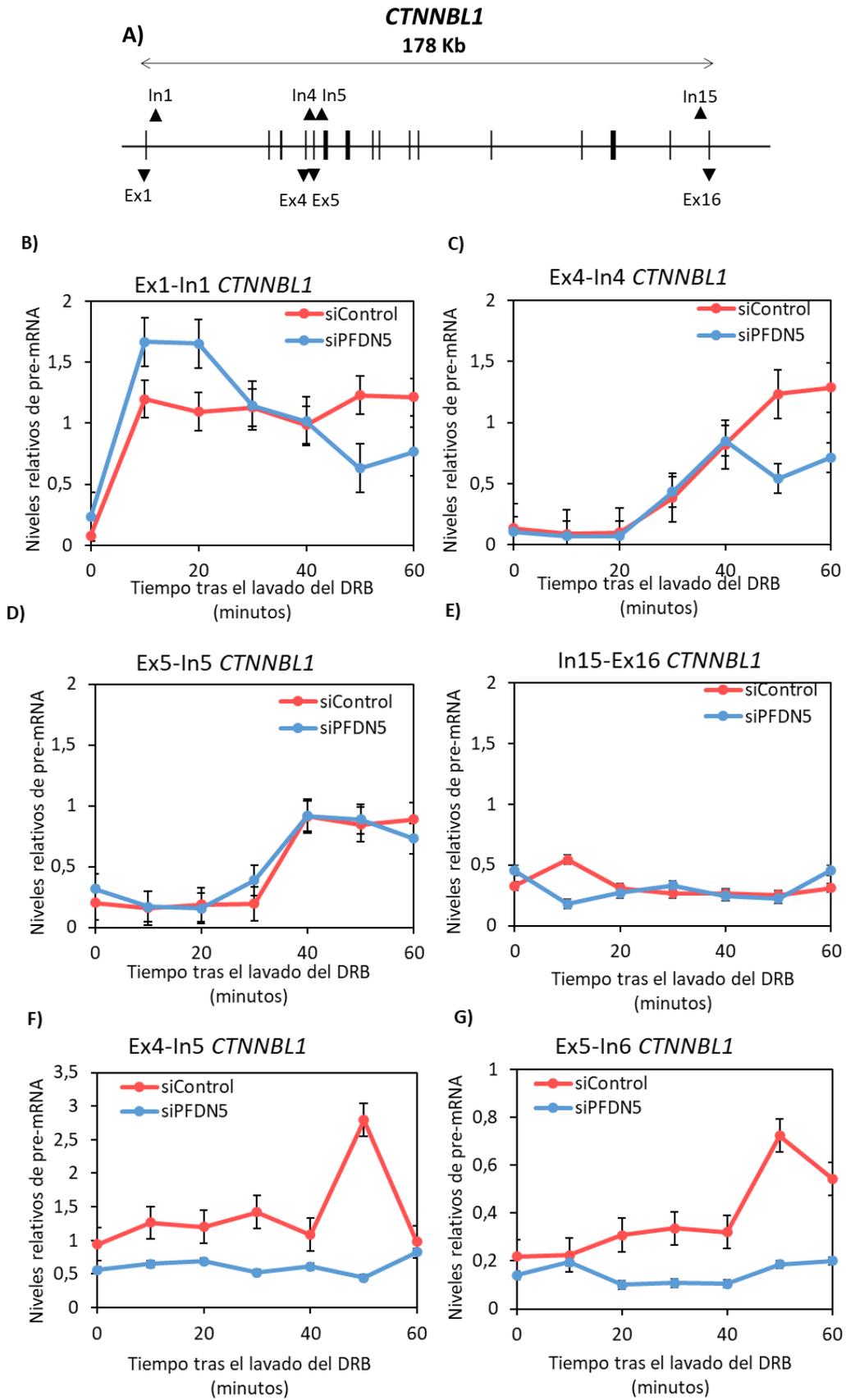


Figura R-26. La deficiencia en prefoldina 5 ocasiona defectos en el *splicing* cotranscripcional.
A) Esquema que representa el gen *CTNNB1* completo, con la localización de los oligonucleótidos utilizados. **B, C, D, E, F y G)** Se realizaron ensayos de RT-qPCR a partir del RNA extraído de células transfectadas con el siControl y el siPFDN5, y tratadas con DRB durante 3 horas. Se tomaron muestras cada 10 minutos a partir del lavado del DRB. Se analizaron diferentes regiones del gen *CTNNB1*. Los datos del RNA inmaduro se normalizaron primero al maduro, y a continuación, a la muestra sin tratar con DRB. **B, C, D y E)** Niveles de pre-mRNA en distintos puntos a lo largo del gen. **F y G)** Eventos de *splicing* cotranscripcional (medidos por la presencia del pre-mRNA sin el intrón x) en distintos puntos a lo largo del gen. Nomenclatura: Ex1 = Exón 1; In1 = Intrón 1.

Los resultados de este ensayo demostraron que la depleción de PFDN5 no afectó a la velocidad de la RNAPII, dado que la transcripción de los intrones 1, 4 y 5 ocurrió en el mismo momento después de lavar el DRB, tanto en el control como en las células deficientes en PFDN5 (**Figura R-26B, 26C y 26D**). La transcripción del intrón 15 no la llegamos a detectar en ninguno de los casos, ya que, seguramente, éste comience a transcribirse después del último punto tomado en nuestro experimento (**Figura R-26E**). Sin embargo, la falta de prefoldina 5 sí que ocasionó defectos en el procesamiento de los intrones analizados. En las células control tanto la eliminación del intrón 4 como la del intrón 5 ocurrió 50 minutos después del lavado del DRB, sin embargo, en las células transfectadas con el siPFDN5 no se detectó el procesamiento de estos intrones (**Figura R-26F y 26G**). Estos resultados indican que la deficiencia en PFDN5 desacopla el procesamiento de intrones de la elongación transcripcional. Como no se produce una alteración simultánea de la velocidad de las RNAs polimerasas II, la alteración del acoplamiento entre elongación y *splicing* no es consecuencia de una aceleración o desaceleración de la elongación transcripcional.

9. LA DEFICIENCIA Y AUSENCIA DE PFDN5 DISMINUYE LOS NIVELES DE RNAPII-Ser2-P.

Como ya mencionamos en la introducción de esta tesis, los cambios en los patrones de fosforilación del CTD juegan un papel fundamental en el *splicing* cotranscripcional, pues contribuyen al reclutamiento y liberación de factores de procesamiento del RNA en las inmediaciones del transcrito naciente al mismo tiempo que la polimerasa está elongando (Komarnitsky et al., 2000; Montes et al., 2012).

Dado que la depleción de prefoldina 5 ocasiona defectos en el *splicing* cotranscripcional, nos preguntamos si también estaría alterando los patrones de fosforilación del CTD, concretamente, la fosforilación en serina 2. Para averiguarlo, decidimos realizar ensayos de CHIP (véase apartado 6.4 de materiales y métodos) en los que estudiamos los niveles de RNAPII-Ser2-P en células deficientes en PFDN5. Además, corroboramos los resultados obtenidos realizando un segundo CHIP con células mutantes que carecían de PFDN5. Los niveles de Ser2 fosforilada fueron analizados en diferentes regiones del gen *CTNNB1*. Primero, todos los datos fueron normalizados al *input*, y luego, los valores de Ser2-P fueron normalizados a los niveles de RNAPII total. Tanto en las células deficientes en PFDN5 como en las células carentes de ésta, descendieron drásticamente los niveles de Ser2 fosforilada en todas las regiones analizadas (**Figura R-27A y 27B**). Resultados similares se obtuvieron en *CD44*, otro gen modelo para *splicing* alternativo (Saint-André et al., 2011) (**Figura 27C-D**). Merece destacarse que este fenotipo también se produce en mutantes de levaduras carentes de las diferentes subunidades del complejo prefoldina (**Figura I-13**), lo que indica que el efecto de prefoldina sobre los niveles de RNAPII fosforilada en Ser2 está conservado evolutivamente.

Para investigar si los defectos en la fosforilación de Ser2 podrían deberse a un efecto indirecto sobre la expresión de *CDK9*, un gen que codifica la principal quinasa que controla la fosforilación de la serina 2 del CTD (Liang et al., 2015), se estudiaron los niveles de mRNA de éste. En los resultados de las RT-qPCR se vio, por un lado, que la deficiencia en PFDN5 no modificó los niveles del mRNA de *CDK9* ni antes ni después de la inducción con suero (**Figura R-27E**), y, por otro lado, que la ausencia de PFDN5 aumentó ligeramente los niveles de éste (**Figura R-27F**). Esto no explicaría el descenso en los niveles de fosforilación en la serina 2 del CTD, por lo que descartamos que este fenotipo fuese debido a un efecto indirecto de la ausencia de PFDN5 sobre *CDK9*.

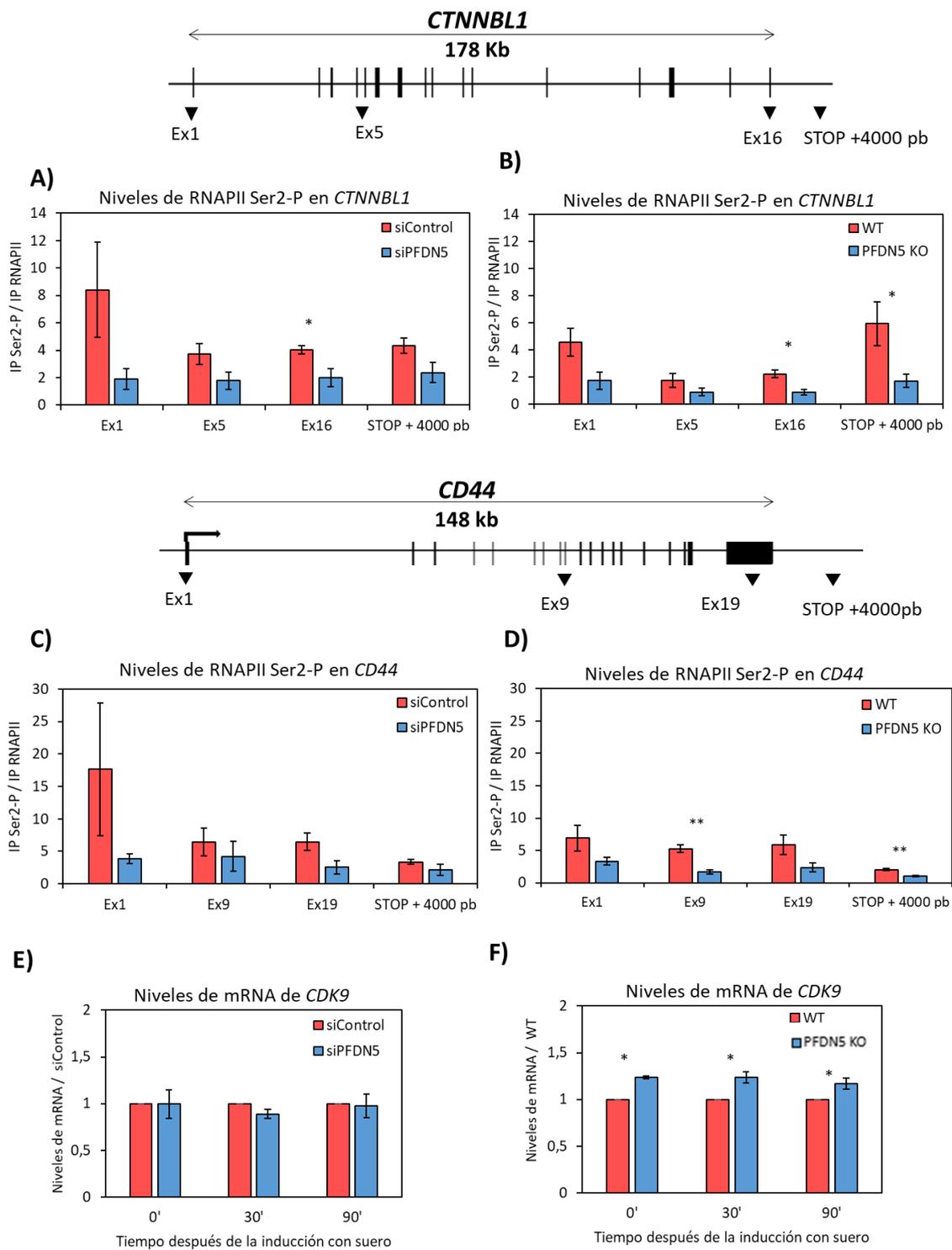


Figura R-27. La deficiencia y ausencia de prefoldina 5 disminuyen los niveles de RNAPII Ser2-P.

A, B, C y D) Se realizaron ensayos de ChIP a partir de extractos obtenidos de células transfectadas con los RNA interferentes (**A** y **C**) y células WT y PFDN5 KO (**B** y **D**) de 3 réplicas biológicas ensayadas en condiciones normales de suero. Todos los datos fueron normalizados primero al *input*, y luego los valores de Ser2-P fueron normalizados a los niveles de RNAPII total. También se realizaron ensayos de RT-qPCR a partir del RNA extraído de células transfectadas con los RNA interferentes (**E**) y células WT y PFDN5 KO (**F**) utilizadas en ensayos con las mismas condiciones que el RNA-seq, y se analizó los niveles de mRNA de *CDK9*. En todas las gráficas se representa valores promedios \pm la desviación estándar obtenidos de 3 experimentos diferentes. **t-Student* con valor $P < 0,05$ y ***t-Student* con valor $P < 0,005$.

10. NO LOGRAMOS DETECTAR INTERACCIONES FÍSICAS ENTRE PFDN5, CDK9 Y FACTORES DE *SPLICING*.

Para averiguar a qué se debían los defectos generales en la eficiencia de *splicing*, así como la bajada de los niveles de RNAPII Ser2-P, ocasionados por la deficiencia y ausencia de prefoldina 5, se planteó analizar si existían interacciones físicas directas entre PFDN5 y algún componente del espliceosoma y/o CDK9, por lo que decidimos realizar ensayos de co-inmunoprecipitación.

Teníamos dos factores candidatos a posibles interactores de prefoldina: el factor LSM8 y el factor PRPF8. El primero es un factor que forma parte del complejo nuclear LSM2-8, el cual se une y estabiliza al RNA nuclear pequeño (snRNA) U6, forma el cuerpo de la ribonucleoproteína (RNP) pequeña U6, e interviene en el *splicing* de los pre-mRNA (Tharun, 2009). Este factor fue sugerido por el Dr. David Alabadí (Instituto de Biología Molecular y Celular de Plantas, CSIC, Valencia) quién ha detectado interacciones físicas entre este factor y subunidades de prefoldina en *Arabidopsis* (datos aún no publicados). En cuanto a PRPF8, es un componente central del espliceosoma, que estabiliza al pre-mRNA en el sitio activo a través de interacciones con la subunidad U5 (Luo et al., 1999; MacRae et al., 2018). Los resultados de dos análisis masivos, que estudian las interacciones entre chaperonas y sus clientes en células humanas y levaduras, señalan a dicho factor de *splicing* como un posible interactor de prefoldina (Gong et al., 2009; Taipale et al., 2014).

En base a estos datos, se realizaron ensayos de co-inmuprecipitación de proteínas con las líneas HCT116 GFP-PFDN5 (que como se detalló en el apartado 1.2 es una línea que sobreexpresa PFDN5 etiquetado con GFP), y la línea HCT116 GFP (como control). Para favorecer la detección de interacciones débiles, este ensayo fue realizado con DSP, un entrecruzador que permite mantener unidos enlaces débiles (véase apartado 6.3 de materiales y métodos). Como control positivo se utilizó el anticuerpo anti-PFDN1 y como negativo el anticuerpo anti-GAPDH. No se detectaron interacciones físicas entre PFDN5 y los candidatos LSM8, PRPF8 y CDK9 (**Figura R-28**).

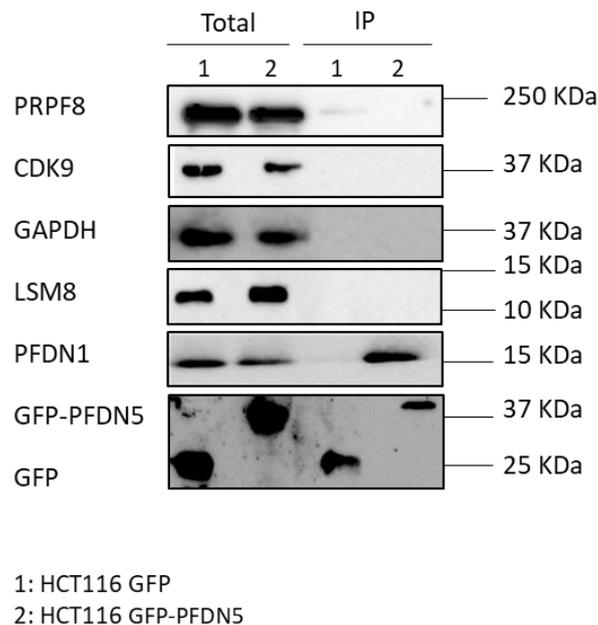


Figura R-28. No detectamos interacciones físicas entre prefoldina 5 y PRP8/CDK9/LSM8. Extractos de las líneas HCT116 GFP y HCT116 GFP-PFDN5 tratados con DSP 1 mM durante 20', se sometieron a inmunoprecipitación con GFP-trap, y posteriormente, el material inmunoprecipitado se analizó mediante *western blot*. Se utilizaron anticuerpos contra PRP8, LSM8 y CDK9, además de los anticuerpos contra PFDN1 (control positivo), GAPDH (control negativo), PFDN5 y GFP (controles de las líneas estables).

11. PFDN5 SE ASOCIA FÍSICAMENTE A LOS GENES *CTNBL1* Y *CD44*.

Puesto que no parece los efectos de PFDN5 sobre el *splicing* del pre-mRNA y la fosforilación de la RNAPII estén mediado por una interacción con la maquinaria de *splicing* o con la principal quinasa de Ser2, decidimos explorar la presencia física de PFDN5 a lo largo de los genes analizados anteriormente. Para averiguarlo, se realizaron 4 ensayos independientes de CHIP a partir de los extractos recogidos de cultivos celulares de las líneas HCT116 Flag y HCT116 Flag-PFDN5 que crecieron en medio completo. En ellos, se analizaron las mismas regiones que fueron estudiadas en el CHIP de RNAPII-Ser2-P. Los niveles de PFDN5 fueron normalizados primero al *impt*, y luego a la región intergénica del cromosoma V.

De esta forma, se detectó una pequeña acumulación de PFDN5 a lo largo de ambos genes, pero no aguas debajo de la región de terminación (**Figura R-29**). Este sugiere que los cambios en los niveles de RNAPII fosforilada en Ser 2 podrían estar

mediados directamente por la presencia de prefoldina 5 en el contexto de la elongación transcripcional.

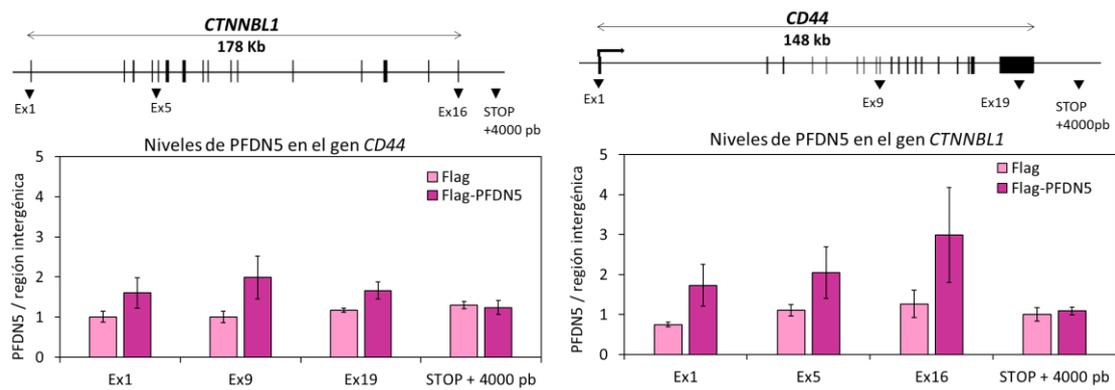


Figura R-29. Prefoldina 5 se encuentra físicamente asociada a los genes *CD44* y *CTNNB1*. Histograma de barras de los niveles de PFDN5 en los genes *CD44* y *CTNNB1*. En la gráfica se representan valores promedio \pm la desviación estándar obtenidos de cuatro experimentos de CHIP diferentes. Ex1 = Exón 1; STOP + 4000 pb: amplicón diseñado a 4000 pb aguas abajo del codón de STOP.

Discusión

1. Algunos de los cambios observados en células deficientes en PFDN2 y PFDN5 podrían estar mediados por la función citoplásmica de la prefoldina.

Para abordar el primer objetivo de esta tesis, que fue estudiar la función del complejo prefoldina en la expresión génica, realizamos ensayos de RNA-seq en células HCT116 deficientes en las subunidades PFDN2 y PFDN5, que fueron incubadas sin suero durante 48 horas y posteriormente inducidas con éste.

Los resultados mostraron que los bajos niveles de PFDN2, y sobre todo de PFDN5, afectaron a la expresión de multitud de genes, principalmente 90 minutos después de la adición del suero (**Figura R-8**), momento en el que se produce la segunda gran oleada de transcripción en respuesta al suero (Harvey, 1991; Lanahan et al., 1992). Aunque la mayoría de los cambios detectados en el transcriptoma de estas células no eran drásticos, el análisis de categorías funcionales permite relacionar estas dos subunidades de prefoldina con varios procesos y funciones celulares.

La función mejor conocida del complejo canónico de prefoldina es su participación en el plegamiento cotraduccional de los monómeros de tubulina y actina, contribuyendo a los procesos que implican cambios dinámicos del citoesqueleto (véase apartado 2.2 de la introducción). Algunos de los cambios transcriptómicos observados podrían por tanto estar mediados por esta función citoplásmica del complejo prefoldina. Si así fuera, esperaríamos que estos cambios se observaran tanto en las células deficientes en PFDN2 como en las deficientes en PFDN5, y que afectaran a genes que se agruparan en categorías funcionales relacionadas con el citoesqueleto o categorías funcionales en las que el citoesqueleto juegue un papel importante.

Esta predicción se corresponde muy bien con las categorías funcionales “organización del citoesqueleto”, “organización sináptica”, “morfogénesis de órganos animales”, “desarrollo del corazón”, “desarrollo de la vascularización”, “regulación de la diferenciación celular” y “regulación negativa del proceso de desarrollo”. Todas estas categorías funcionales disminuyeron su expresión, tanto en células deficientes en PFDN2 como en PFDN5, 90 minutos después de la inducción con suero (**Tablas R-4 y R-5**). En estas dos poblaciones y en este mismo punto, también encontramos inducido el

gen *RhoB* (**Tabla R-6**), que codifica para una GTPasa que interactúa con efectores que intervienen en diversos procesos celulares, entre ellos, la organización de la actina en el citoesqueleto, la migración celular, el tráfico de membranas, la proliferación celular y procesos del desarrollo (Vega and Ridley, 2016). Parece, por tanto, que estos cambios podrían explicarse por la deficiencia funcional del complejo canónico de prefoldina como cochaperona en el citoplasma, y especialmente, por su papel en el ensamblaje del citoesqueleto.

2. Algunos de los cambios observados en células deficientes en PFDN5 podrían deberse a su relación con c-Myc.

Algunos cambios detectados en el transcriptoma de las células deficientes en PFDN5 podrían estar mediados por *c-Myc*, puesto que esta subunidad es un regulador negativo de dicho protooncogén (véase apartado 2.3 de la introducción de esta tesis).

c-Myc se induce inmediatamente en respuesta a estímulos mitogénicos, como por ejemplo el suero, tras lo cual regresa a sus niveles basales de expresión (Nie et al., 2012). *c-Myc* se encuentra sobrepresado en diversas líneas celulares humanas de carcinoma de colon, incluida la línea HCT116 (Erisman et al., 1985; Wu et al., 2017; Zhao et al., 2017). Dado que *c-Myc* regula positivamente la proliferación celular (Nie et al., 2012), esto podría explicar que la línea HCT116 sea capaz de proliferar pese a la ausencia de suero en el medio del cultivo (**Figura R-6**).

PFDN5 es un co-represor de la actividad transcripcional de *c-Myc* (Hagio et al., 2006; Mori et al., 1998) y es capaz de favorecer su degradación reclutando a un complejo ligasa de ubiquitinas y dirigiéndolo al proteasoma (Kimura et al., 2007) (véase apartado 2.3 de la introducción de este trabajo). Por tanto, la deficiencia de PFDN5 debería aumentar la actividad transcripcional de *c-Myc* y consecuentemente, deberíamos detectar un aumento de los productos de los genes diana de este regulador. Ese es el caso, por ejemplo, de todas las categorías funcionales sobrepresadas que están relacionadas con la mitocondria (“organización mitocondrial”, “traducción mitocondrial”, “elongación de la traducción mitocondrial”, “ensamblaje del complejo de cadena respiratoria mitocondrial”, “organización de la membrana mitocondrial” y “establecimiento de la localización de proteínas de la

membrana de la mitocondria”) (**Tabla R-4**), puesto que los genes implicados en la biogénesis mitocondrial son claras dianas de c-Myc (Kim et al., 2008; Morrish and Hockenbery, 2014). Además, la expresión de los genes que constituyen la categoría funcional “regulación de la diferenciación celular” se encuentra disminuida en las células deficientes en PFDN5 (**Tabla R-4**), algo que podría relacionarse con el papel represor de c-Myc sobre los genes implicados en diferenciación celular (Nie et al., 2012). Ninguna de estas categorías funcionales se encontró alterada en las células deficientes en PFDN2, lo que refuerza que el cambio de estas GOs por deficiencia de PFDN5 sea mediado por c-Myc.

3. Cambios de expresión de genes relacionados con el control del ciclo celular

Entre los genes alterados en el transcriptoma de las células deficientes en PFDN5 antes y después de la inducción con suero, se encontraban varios relacionados con el control del ciclo celular y la replicación del genoma. Uno de los genes significativamente inducido por la carencia de PFDN5 fue *CDKN1A*, que codifica para la proteína p21.

Se ha descrito que c-Myc es capaz de antagonizar la actividad de p21 en el ciclo celular, reprimiendo la expresión de este inhibidor de ciclinas a través de diversos mecanismos. El mejor estudiado es el mecanismo por el cual c-Myc interrumpe la interacción de p53 con el promotor de *p21*, a través de la cual se induce la expresión de éste último (Bretones et al., 2015). p53 es capaz de inducir la expresión de p21 en respuesta a diferentes situaciones de estrés, como por ejemplo, el daño en el DNA (Karimian et al., 2016). En nuestros datos transcriptómicos obtenidos del análisis del RNA-seq, no detectamos un cambio significativo de la expresión de p53 en células deficientes en PFDN5 (\log_2FC 0.049 y valor P 0.61); y nuestros experimentos de inmunofluorescencia tampoco mostraron que la deficiencia de PFDN5 provocase daño en el DNA (**Figura R-13**). Por tanto, nuestros resultados señalan que la sobreexpresión de p21 causada por la deficiencia de PFDN5 es independiente de la vía de p53.

Alternativamente, c-Myc puede reprimir el promotor de *p21* de forma independiente a p53 (Gartel et al., 2001). Si esta función estuviese mediada por el papel represor de PFDN5 sobre la capacidad transactivadora de c-Myc, los bajos

niveles o la ausencia de PFDN5 podrían hacer que c-Myc ejerciese un papel activador sobre *p21*, desencadenando la inducción de este gen. El hecho de que hayamos detectado la presencia física de PFDN5 en la región 5' de *p21* es coherente con esta hipótesis.

En *Drosophila melanogaster* se ha descrito que la deficiencia de PFDN3 provoca alteraciones en los husos y centrosomas (Delgehr et al., 2012), lo cual podría provocar defectos en la segregación cromosómica. Si ocurriese algo similar por la deficiencia en PFDN5 en nuestras células, la inducción de *p21* podría ser una respuesta que permitiría detener la mitosis (Bertoli et al., 2013). Esta hipótesis se ajusta a la leve acumulación de células en fase G2 observada en células deficientes en PFDN5 (**Figura R-14A**), sin embargo, no explica el pequeño incremento de células en fase G1 en el mutante que carece de PFDN5 (**Figura R-14B**).

Acorde a la sobreexpresión de *p21* y la ligera acumulación de células en fase G1 en el *knockout*, o en G2 en el *knockdown*, detectamos una disminución de la expresión de la categoría funcional “ensamblaje de nucleosomas”, de la cual forman parte principalmente los genes de histonas, cuya expresión se encuentra estrictamente regulada por el ciclo celular (Ewen, 2000) (véase apartado 1.1.3 de la introducción de este manuscrito). La disminución de la expresión de los genes de histonas fue mucho más acusada en el mutante de PFDN5 que en las células tratadas con el siPFDN5, como así lo fue también la inducción de *p21*. Estos resultados sugieren que la inducción de *p21* en células carentes de PFDN5 retrasa su entrada en fase S, lo que hace disminuir la expresión de genes con expresión específica durante esta fase del ciclo, como las histonas.

4. PFDN5 contribuye de forma general al proceso de transcripción dependiente de RNAPII.

Nuestro laboratorio ha puesto de manifiesto un papel general de la prefoldina en la elongación de la transcripción en *Saccharomyces cerevisiae*, favoreciendo la expresión de genes largos y la dinámica de cromatina asociada a la transcripción (Millán-Zambrano et al., 2013). De existir una función similar en la prefoldina humana, esperaríamos que su deficiencia provocase alteraciones de expresión en función de las

características físicas de los genes, por ejemplo, de su longitud, al margen de su pertenencia a categorías funcionales. El análisis de los datos transcriptómicos de las células deficientes en PFDN5 confirma esta hipótesis, puesto que la expresión en condiciones de inducción por suero resultó significativamente disminuida en función de la longitud de los genes en las células deficientes en PFDN5 (**Figura R-24A**).

Obtuvimos resultados similares cuando analizamos el efecto de la deficiencia en PFDN5 sobre la expresión de los genes en función del número de intrones, otra característica independiente de las categorías funcionales (**Figura R-24B**). Puesto que el número de intrones correlaciona con la longitud de los genes (Regulapati et al., 2008) este resultado confirma una influencia de PFDN5 sobre la transcripción de genes largos y/o sobre la expresión de genes con alto número de intrones.

Dada la correlación antes apuntada, ambos fenómenos (efectos de longitud y de número de intrones) podrían ser dependientes. Sin embargo, nuestro análisis del efecto de la deficiencia de PFDN5 sobre la inducción de la expresión por suero indica que la longitud y el número de intrones contribuyen sinérgicamente al efecto negativo causado por la deficiencia de PFDN5 (**Figura R-25**). Ello sugiere que la deficiencia en PFDN5 provoca defectos en la elongación transcripcional, al igual que lo observado en levaduras con varias subunidades de prefoldina (Millán-Zambrano et al., 2013).

Los análisis del estado de fosforilación de la Ser-2 del CTD de la RNAPII en *CTNNBL1* y *CD44* muestran un descenso muy acusado en las células deficientes en PFDN5 o carentes de esta subunidad (**Figura R-27A-D**), tal y como se ha detectado en los mutantes de levaduras carentes de las diferentes subunidades del complejo prefoldina (**Figura I-11**). Este hecho apoya fuertemente que la deficiencia de PFDN5 cause defectos transcripcionales durante la fase de elongación también en células humanas. Sin embargo, no hemos detectado diferencias de velocidad de la elongación en el gen *CTNNBL1* de células deficientes en PFDN5 (**Figura R-26**) lo que descarta alteraciones drásticas del complejo de elongación.

5. La deficiencia o carencia de PFDN5 causa defectos en el *splicing* del pre-mRNA.

El efecto negativo del número de intrones sobre la expresión en células deficientes en PFDN5 es consistente con los resultados del análisis de la eficiencia de procesamiento del pre-mRNA. Hemos detectado que, mientras las células control mejoraban la eficiencia del *splicing* tras la adición del suero, las células deficientes en PFDN5, y en menor medida las deficientes en PFDN2, la empeoraban (**Figura R-19**). El hecho de que ambas subunidades provoquen efectos similares en el *splicing* del pre-mRNA indica que estamos ante un fenómeno mediado, probablemente, por el complejo canónico de prefoldina.

Puesto que el *splicing* es un fenómeno que ocurre en gran medida de forma cotranscripcional (Herzel et al., 2017), nuestros resultados pueden explicarse en el contexto del acoplamiento entre *splicing* y transcripción (**Figura D-1A**). Y así lo demuestran los resultados obtenidos en nuestros ensayos de *splicing* en el gen *CTNNBL1*, donde observamos un claro desacoplamiento del *splicing* del pre-mRNA respecto de la elongación de la transcripción (**Figura R-26**).

Los defectos en Ser2-P producidos por la deficiencia en PFDN5 pueden explicar este desacoplamiento, ya que la fosforilación del CTD en Ser2 juega un papel fundamental en el reclutamiento de factores de *splicing* a la RNAPII elongante (Gu et al., 2013). Por tanto, la deficiencia en PFDN5 causaría un estado defectuoso de la fosforilación del CTD que, a su vez, podría resultar en un reclutamiento anómalo de los factores de *splicing* durante la transcripción (**Figura D-1B**). Asimismo, los bajos niveles de Ser2 ocasionaría problemas en la síntesis de mRNA maduro de forma general, pero posiblemente estén afectando más a los genes largos y con alto número de intrones, tal y como nuestros resultados lo evidencian. De hecho, ha sido comprobado que la transcripción de los genes de mayor longitud y con muchos exones se ve más afectada por la disminución de los niveles de Ser2-P que la del resto de genes (Blazek et al., 2011).

No hemos detectado un descenso de la expresión de la quinasa CDK9 que fosforila la Ser2 del CTD (**Figura R-27E-F**), ni tampoco una interacción física entre

PFDN5 y CDK9 (**Figura R-29**). En nuestro estudio transcriptómico tampoco se detectó una disminución de la expresión de CDK12 y CDK13, las otras dos quinasas capaces de fosforilar a Ser2. No disponemos por tanto de evidencias que permitan explicar la disminución de la fosforilación de Ser2 por una acción de la prefoldina sobre las quinasas. Por tanto, es posible que los reducidos niveles de fosforilación en Ser2 sean consecuencia de un funcionamiento defectuoso del proceso de elongación que cause una desfosforilación prematura.

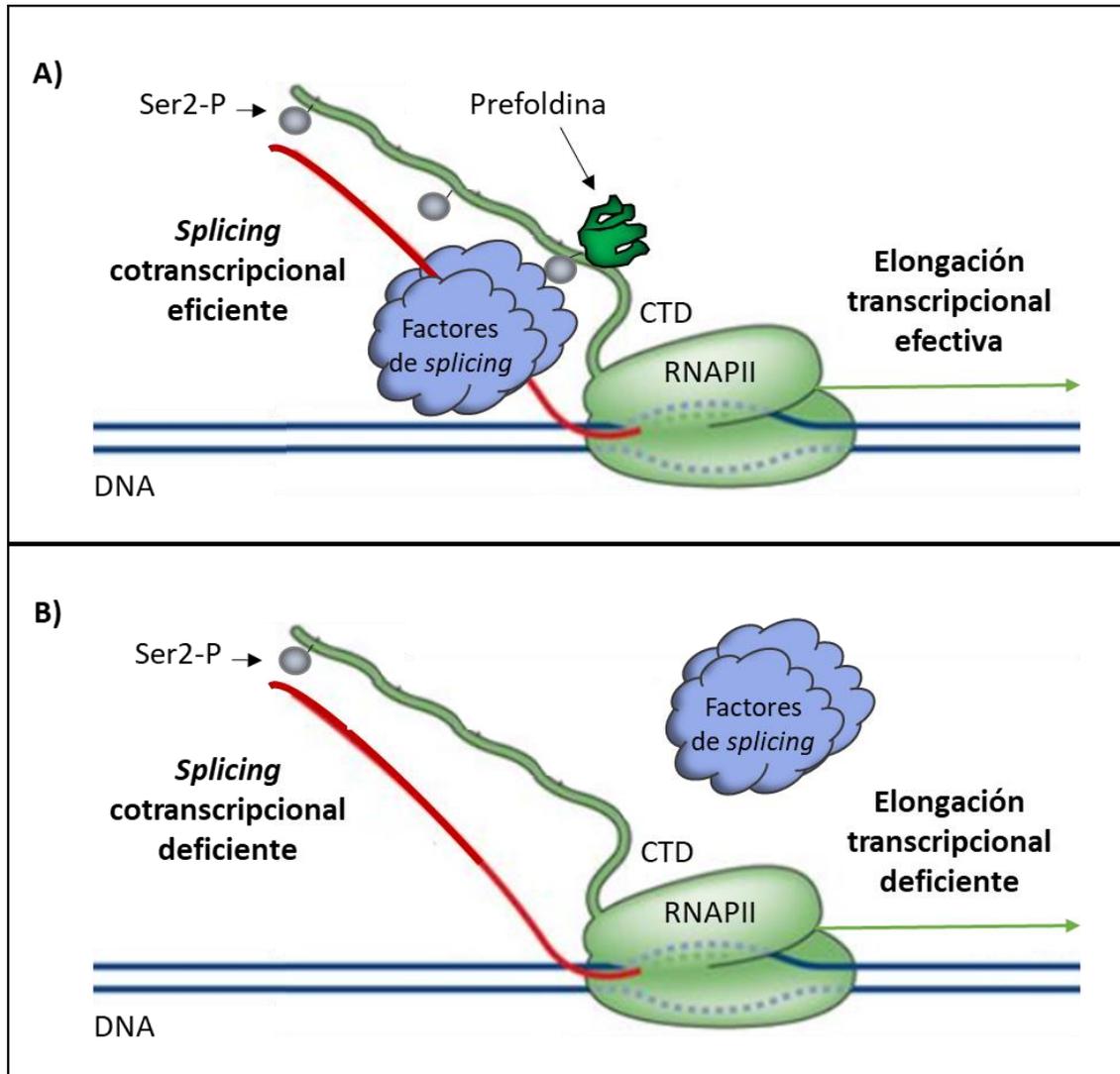


Figura D-1. Modelo por el cual se explica que prefoldina podría influir de forma general en la elongación de la transcripción, y el *splicing* cotranscripcional. A) Prefoldina contribuye positivamente al proceso de elongación de la transcripción, lo que facilita el *splicing* cotranscripcional. **B)** La falta de prefoldina ocasiona defectos transcripcionales y disminuye los niveles de Ser2-P. Este hecho dificulta el reclutamiento cotranscripcional de la maquinaria de *splicing* y provoca un desacoplamiento entre transcripción y *splicing*.

Alternativamente, los defectos transcripcionales detectados podrían ser consecuencia de un mal funcionamiento del pre-mRNA *splicing*. Un mal procesamiento de los transcritos inmaduros podría producir defectos en la fosforilación del CTD en Ser2, tal como ha sido propuesto (Koga et al., 2015). Según este modelo alternativo, los defectos en la fosforilación de Ser2 serían consecuencia de las alteraciones en el *splicing*, y no al revés, lo que implicaría que la prefoldina actuase funcionalmente sobre la maquinaria de *splicing*. Varias subunidades del espliceosoma han sido propuestas como posibles clientes de las diferentes subunidades de prefoldina a partir de estudios proteómicos realizados en células humanas y levaduras (Gong et al., 2009; Taipale et al., 2014). Entre ellas destaca PRPF8, ya que forma parte del sitio activo del espliceosoma (Nguyen et al., 2013). Un estudio reciente ha demostrado además que Hsp90 y R2TP, dos de las cochaperonas con las que colabora el complejo alternativo de prefoldina, estabilizan a PRPF8 en el citoplasma de células humanas y lo ayudan a ensamblarse a U5 snRNP en el núcleo (Malinová et al., 2017), lo cual refuerza la idea de que prefoldina pueda estar interaccionando físicamente con factores de *splicing*. Otro de los posibles clientes de PFDN5 propuestos fue el factor LSM8, que forma parte del complejo LSM2-8 que interacciona con la U6 snRNP (Tharun, 2009). Este factor ha sido detectado como interactor de prefoldina en plantas (comunicación personal de los laboratorios de Julio Salinas (CIB-CSIC) y Miguel Ángel Blázquez y David Alabadí (IBMCP-CSIC)).

A pesar de nuestros esfuerzos por detectar posibles interacciones de PFDN5 con PRP8 y LSM8 mediante experimentos de co-inmunoprecipitación, éstos no han dado resultados positivos (**Figura R-28**). Por tanto, no podemos confirmar esta hipótesis.

Nuestros resultados experimentales no nos permiten por tanto confirmar ninguno de los dos modelos propuestos. Sin embargo, las siguientes evidencias nos hacen inclinarnos por el primero de ellos (**Figura D-1**):

1. El hecho de que la deficiencia en PFDN5 no afecte en particular a ninguna modalidad de procesamiento alternativo del pre-mRNA (**Figura R-23**) sugiere que la prefoldina no estimula específicamente ningún paso del proceso de *splicing*, sino que ejerce un efecto general sobre el mismo, probablemente de forma indirecta.

2. Hemos detectado que en ausencia de suero la eficiencia del *splicing* del pre-mRNA fue más alta en las células deficientes en PFDN5 que en las células control (**Figura R-19**), lo que parece indicar que dicha deficiencia no produce un efecto primario sobre el *splicing*, sino que éste es el resultado indirecto de la función de PFDN5 en otros procesos celulares que se ven influidos por el estado nutricional de las células.

3. Por último, si la PFDN5 actuase directamente sobre la maquinaria de *splicing* no se requeriría la presencia física de ésta en el sitio de la transcripción. Sin embargo, nuestros resultados de CHIP muestran a PFDN5 asociada físicamente al gen *CTNNB1* (**Figura R-29**), en el que hemos detectado el efecto de prefoldina sobre el acoplamiento transcripcional del *splicing* (**Figura R-26**), así como al gen *CD44* (**Figura R-29**), donde la deficiencia en PFDN5 produce también un descenso claro de la fosforilación de Ser2.

Conclusiones

1. Dos subunidades del complejo prefoldina, PFDN2 y PFDN5, son necesarias para mantener los niveles de expresión de multitud de genes en las células de carcinoma humano, HCT116.
2. PFDN5 influye ligeramente en el ciclo celular de estas células, tal vez a través de la represión del gen p21.
3. PFDN5 favorece la fosforilación en Ser2 de la RNA polimerasa II en dichas células.
4. PFDN5 es necesario para el acoplamiento temporal entre transcripción y *splicing* en la línea celular HCT116.

Materiales y métodos

1. ORGANISMOS Y CONDICIONES DE CULTIVOS.

En esta tesis se han utilizado bacterias y líneas celulares de origen humano. La bacteria *Escherichia coli* (*E. coli*) se utilizó como herramienta para las técnicas de DNA recombinante. Para el resto de los ensayos de esta tesis, se usaron líneas celulares de origen humano (**Tabla M-1**).

1.1 Línea bacteriana y condiciones de cultivo.

La estirpe bacteriana DH5 α de *E. coli* se cultivó en medio Luria-Bertani (LB) esterilizado mediante autoclave y suplementado, en el caso en que fuera necesario, con ampicilina 100 $\mu\text{g/ml}$ o kanamicina 25 $\mu\text{g/ml}$. El cultivo sólido se llevó a cabo en estufas a 37 $^{\circ}$ C. El cultivo líquido, además, requirió agitación orbital continua.

- **LB:** NaCl 5 g/l; bacto triptona 10 g/l; extracto de levadura 5 g/l. Se añadieron 20 g/l de bacto agar para la obtención de medio sólido.

1.2. Líneas celulares y condiciones de cultivo.

La línea celular HCT116, cedida por el Dr. José Carlos Reyes, y las obtenidas a partir de ésta (**Tabla M-1**), se cultivaron en medio de cultivo Mckoy 5a Medium Modified suplementado con 10% (v/v) de suero fetal bovino (FBS) de la marca Biowest y 1% penicilina/estreptomicina 100 X de la marca Biowest. Además, al medio de las líneas celulares modificadas HCT116 GFP-PFDN5, HCT116 GFP, HCT116 FLAG-PFDN5, HCT116 FLAG y HCT116 PFDN5 KO + GFP-PF5, se le añadió G418 (Melford) a una concentración final de 2 $\mu\text{g}/\mu\text{l}$. Todos los cultivos se mantuvieron en una atmósfera húmeda, a 37 $^{\circ}$ C, y con CO $_2$ al 5%.

Tabla M-1: Línea celular.

LÍNEA CELULAR	ESPECIE	ORIGEN / CARACTERÍSTICA
HCT116	<i>Homo sapiens</i>	Carcinoma de colon
HCT116 Flag-PFDN5	<i>Homo sapiens</i>	Expresa PFDN5 etiquetada con Flag en su extremo C-terminal
HCT116 Flag	<i>Homo sapiens</i>	Expresa la etiqueta Flag

LÍNEA CELULAR	ESPECIE	ORIGEN / CARACTERÍSTICA
HCT116 GFP-PFDN5	<i>Homo sapiens</i>	Expresa PFDN5 etiquetada con GFP en su extremo C-terminal
HCT116 GFP	<i>Homo sapiens</i>	Expresa la etiqueta GFP
HCT116 PFDN5 KO	<i>Homo sapiens</i>	Línea CRISPR que no expresa PFDN5
HCT116 PFDN5 KO + GFP-PFDN5	<i>Homo sapiens</i>	Línea HCT116 PFDN5 KO que expresa PFDN5 etiquetada con GFP en su extremo C-terminal

2. OLIGONUCLEÓTIDOS Y PLÁSMIDOS.

Los oligonucleótidos utilizados para la biología molecular del DNA se muestran en la **Tabla M-2**, los oligonucleótidos usados en la biología molecular del RNA (siRNA) en la **Tabla M-3** y los plásmidos usados en la **Tabla M-4**.

2.1. Oligonucleótidos.

Tabla M-2: Oligonucleótidos utilizados.

NOMBRE	SENTIDO 5'-3'	ANTISENTIDO 5'-3'	USO	REFERENCIA
PFDN5	CGCAGCTAGAAATGCTCAAGAA	GCAATGGACGTGGACAAGAA	qPCR-RT	Esta tesis
PFDN2	GTTGGAGATGGAGTTGAATG	CTCCTTGTGTCTCAAAG	qPCR-RT	Esta tesis
GAPDH	GAGTCAACGATTTGGTCGT	AATGAAGGGTCATTGATGG	qPCR-RT	(Jimeno-González <i>et al.</i> , 2015)
JUN	TGACTGCAAAGATGGAAACG	CAGGGTCATGCTCTGTTCA	qPCR-RT	(Lopez-Bergami <i>et al.</i> , 2010)
CTNNB1 Ex1-In1	AGTGCAGGGAAGTGGAGTATTTGC	AGAGGAGGTGAGATGAAAGGGCT	qPCR-RT / ChIP	(Singh and Padgett, 2009)
CTNNB1 Ex4-In4	TACCACCTTCTGGTGGAGCTGAAT	GGAGAAAGAGTAACAGCACTTCCC	qPCR-RT	(Singh and Padgett, 2009)
CTNNB1 Ex5-In5	TGAAGAGGGAGCAGAAGTGCTCAT	TATTACTIONCAGCCTCACCACACT	qPCR-RT	(Singh and Padgett, 2009)
CTNNB1 In15- Ex16	TTCTGGGAAAGGTATGAAGCGACG	TAGAAGTTCTCCAGCAAGCCCA	qPCR-RT / ChIP	(Singh and Padgett, 2009)
CTNNB1 Ex4-In5	TACCACCTTCTGGTGGAGCTGAAT	TATTACTIONCAGCCTCACCACACT	qPCR-RT	(Singh and Padgett, 2009)

NOMBRE	SENTIDO 5'-3'	ANTISENTIDO 5'-3'	USO	REFERENCIA
CTNBL1 Ex5-In6	TGAAGAGGGAGCAGAAGTGCTCAT	CAAAGCAACTCAGGCAATGGCA	qPCR-RT	(Singh and Padgett, 2009)
CTNBL1 STOP + 4000 pb	TGGGTAGCTGAGAGGGAGGTG	ATTATTGGCCCCACTGTGCAG	qPCR-RT / ChIP	Esta tesis
CD44 Ex1	TCCCTCCGTCTTAGGTCAGTGT	CCTCGGAAGTTGGCTGCAGTTT	ChIP	(Saint-André <i>et al.</i> , 2011)
CD44-Ex9	GAAACTGGAACCCAGAAGCACA	TGATGCTCATGGTGAATGAGGG	ChIP	(Saint-André <i>et al.</i> , 2011)
CD44-Ex19	TGATCAACAGTGGCAATGGAGC	TCTGACGACTCCTTGTTACCA	ChIP	(Saint-André <i>et al.</i> , 2011)
CD44 STOP + 4000 pb	CCTGTGGTCCCAGCTACTC	CTCACTTTCACCCAGGCTG	ChIP	Esta tesis
NT Chr.V	CTGTACCTGGGGTTCATTCATT	CAGTAAGCCGTTCACTCTCAC	ChIP	(Gómez-Herreros <i>et al.</i> , 2014)
P21 Ex1	GGGGCGGTTGTATATCAGG	GGCTCCACAAGGAACTGACT	ChIP	Esta tesis
P21 Ex2	CTGCCAAGCTCTACCTTC	GGTCCACATGGTCTTCTCT	qPCR-RT / ChIP	(Saramaki, 2006)
HIST1H3J	AGCTGCTCATCCGAAACTG	AGCACAGAGGTTGGTGTCTTCAAAG	qPCR-RT	Esta tesis
HIST1H4B	CGGTAAAGGTTTGGGTAAGGGA	CAAACCGGAAATTCGCTTAACC	qPCR-RT	Esta tesis
HIST1H2AB	GCTTAATAAACTCTTGGGGCGT	TCCCTTGGCCTTATGATGG	qPCR-RT	Esta tesis
HIST1H2BJ	TATTCCATCTATGTGTACAAGTTCTG AAGC	ATGCGCTCGAAAATGTCGTTT	qPCR-RT	Esta tesis
HIST3H2BB	ATCACGTCCC GCGAAGTG	TTCAGGACGCCGAGGAAC	qPCR-RT	Esta tesis
DSP Ex13-In13	CATAGAGAAGATCAGGGCCATG	CACAAAAGCACTGCTACACAGC	qPCR-RT	Esta tesis
DSP Ex14		ATCATCTCCAAACATCTCTGAGC	qPCR-RT	Esta tesis
FASN Ex22-In22	AGAACATGCCAGCCTGAAG	TGTCTGTGCTGTCTCCTGCA	qPCR-RT	Esta tesis
FASN Ex23		GTGGCCGTGTAGCTCAGCT	qPCR-RT	Esta tesis
EVPL Ex21-In21	TTGCAAAGGCCTATACTGAGG	GGAGAGGAGGGAAGGAGACT	qPCR-RT	Esta tesis

NOMBRE	SENTIDO 5'-3'	ANTISENTIDO 5'-3'	USO	REFERENCIA
EVPL Ex22		CTTTGCATCATGGGTCCTTC	qPCR-RT	Esta tesis
HSP90B1 Ex4-In4	ACTGATGAAAATGCTCTTTCTGG	GCCACTGCAAATAATCTAGCC	qPCR-RT	Esta tesis
HSP90B1 Ex5		GCTTGTCCCAGATTTGGCTA	qPCR-RT	Esta tesis
csnPT1 y csnPT2	GTCGACACTTGCTCAAC	AAGCTTGCATCGAATCAGTAG	qPCR-RT	(Becerra <i>et al.</i> , 2015)
EcoRI-PFDN5 Fw PFDN5-Sall Rv	GGCACGAATTCATGGCGCAGTCTAT TAACATCA	TTCGTGACCTCAGGCCTTAGCAGT AGCCT	PCR	Esta tesis
XhoI-PFDN5 Fw PFDN5-EcoRI Rv	GGCACTCGAGACATGGCGCAGTCTAT TAACATCA	GGCACGAATTCCTCAGGCCTTAGCA GTAGCCT	PCR	Esta tesis

2.2. Pequeños RNAs de interferencia (siRNAs).

Tabla M-3: siRNAs utilizados.

GEN	SIRNA	SECUENCIA (5'-3')	REFERENCIA
Luciferasa	siControl	CGUACGCGAAUACUUCGA	(Tjeertes, Miller and Jackson, 2009)
Prefoldina 5	siPFDN5	AGAGAAGACAGCUGAGGAU	Esta tesis
Prefoldina 2	siPFDN2 A	GACACUGACACAGCAGCUU	Esta tesis
Prefoldina 2	siPFDN2 B	CUCGUAAGUGCUACCGCAU	Esta tesis
Prefoldina 2	siPFDN2 C	CAGCCUAGUGAUCGAUACA	Esta tesis

2.3. Plásmidos.

El vector de expresión en eucariota pFlag-PFDN5 (construido a partir del vector comercial pCMV-Tag 2A, que codifica para la proteína FLAG) fue obtenido mediante técnicas convencionales de PCR (ver apartado 3.4). En ellas, se usaron el par de oligonucleótidos EcoR1-PFDN5 Fw / PFDN5-Sall Rv para amplificar la secuencia de la isoforma α de PFDN5 a partir del cDNA extraído de células HCT116. Posteriormente, el producto de PCR se incubó a 72 °C en presencia de Taq y ATP

para añadirle un nucleótido A en cada extremo que facilitaría la ligación de estos fragmentos al plásmido pGEM-T. Este vector, en su forma lineal, posee un nucleótido de timina en sus extremos 3', los cuales actúan como extremos cohesivos con los fragmentos de DNA amplificados por PCR, ya que son generados con un nucleótido de adenina en los extremos 3'. La presencia de un nucleótido de timina en los extremos 3' del vector impide además la recircularización de éste, aumentando la eficiencia del proceso de ligación. Una vez ligado el inserto al vector, se transformó la cepa DH5 α de *Escherichia coli* (ver apartado 3.6). Para seleccionar los clones positivos, el sitio de clonamiento múltiple del vector interrumpe la secuencia codificante del péptido α de la enzima β -galactosidasa. Dicha enzima hidroliza el sustrato 5-bromo-4-cloro-3-indolil- β -D-galactósido, X-gal, generando un precipitado de color azul. Así pues, proporcionando este sustrato al medio de cultivo (LB), las colonias que adquieren este color se identifican como negativas, y las blancas como positivas. A continuación, se aisló el plásmido (ver apartado 3.1) y se digirió con la pareja de enzimas *EcoRI-Sall*. Luego, se corrió un gel de agarosa para separar los fragmentos del DNA digerido, y se purificó la banda correspondiente al inserto (ver apartado 3.1). Finalmente, este se clonó en el vector pCMV-Tag 2A, previamente digerido con el mismo par de enzimas mencionado anteriormente.

Del mismo modo fueron generados los plásmidos pLSm8-EGFP y pEGFP-PFDN5, basados en los vectores comerciales pEGFP-N1 y pEGFP-C1 respectivamente, los cuales codifican para la proteína fluorescente EGFP. Las parejas de oligonucleótidos utilizados en la PCR fueron *EcoRI-LSm8 Fw / LSm8-Sall Rv* para amplificar la secuencia de LSm8 y el par *XhoI-PFDN5 Fw / PFDN5-EcoRI Rv* para amplificar la secuencia de PFDN5.

Tabla M-4: Plásmidos utilizados.

PLÁSMIDO	ANTIBIÓTICO DE SELECCIÓN	PROMOTOR	APLICACIÓN EN LA TESIS	REFERENCIA
pEGFP-C1	Kanamicina G418	CMV	Línea HCT116 con el plásmido integrado que expresa EGFP.	Addgene

PLÁSMIDO	ANTIBIÓTICO DE SELECCIÓN	PROMOTOR	APLICACIÓN EN LA TESIS	REFERENCIA
pEGFP-PFDN5	Kanamicina G418	CMV	Línea HCT116 con el plásmido integrado que expresa EGFP-PFDN5.	Esta tesis
pCMV-Tag 2A	Kanamicina G418	CMV	Línea HCT116 con el plásmido integrado que expresa Flag.	Addgene
pFlag-PFDN5	Kanamicina G418	CMV	Línea HCT116 con el plásmido integrado que expresa Flag-PFDN5.	Esta tesis
pGEM-T	Ampicilina		Clonaje.	Promega
Minigen <i>Fas</i>	Ampicilina		Análisis del <i>splicing</i> del gen reportero <i>Fas</i> en las líneas HCT116 y HCT116 KO.	Cedido por el Dr. Carles Suñé

3. MANIPULACIÓN Y ANÁLISIS DEL DNA.

3.1. Aislamiento de DNA plasmídico bacteriano.

El aislamiento de DNA plasmídico a pequeña escala se llevó a cabo mediante lisis alcalina. Para ello, se centrifugaron 2 ml de cultivo bacteriano a máxima velocidad durante 1 minuto. Las células, contenidas en el precipitado, se resuspendieron en 100 μ l de la solución 1 y tras 5 minutos a temperatura ambiente (RT) se añadieron 200 μ l de la solución 2, se mezcló por inversión y se incubó posteriormente 5 minutos en hielo. Tras la adición de 150 μ l de la solución 3, la mezcla resultante se invirtió varias veces y posteriormente se incubó 5 min en hielo. Las proteínas, el DNA cromosómico y los restos celulares se eliminaron por centrifugación a máxima velocidad durante 5 minutos. La precipitación del DNA plasmídico se llevó a cabo añadiendo dos volúmenes de etanol puro, incubando 2 minutos a temperatura ambiente y centrifugando posteriormente a máxima velocidad durante 5 minutos. Al precipitado obtenido se añadió etanol al 70 % para eliminar las sales. Se volvió a centrifugar a máxima velocidad durante 5 minutos y finalmente el precipitado resultante se resuspendió en 50 μ l de TE 1X. Para la obtención de DNA plasmídico más puro a mediana escala, se utilizó el kit comercial de midiprep de Zymo Research, siguiendo las instrucciones del fabricante.

- **Solución 1:** Glucosa 50 mM; Tris-HCl pH 8 25 mM y EDTA pH 8 10 mM.
- **Solución 2:** NaOH 0,2 M; SDS 1 %.

- **Solución 3:** Acetato potásico 3 M; ácido acético glacial 11,5 %.
- **TE:** Tris-HCl pH 8 10 mM y EDTA pH 8 1 mM.

3.2. Electroforesis de DNA en geles de agarosa.

La separación de fragmentos de DNA se llevó a cabo según el método de Sambrook (Sambrook, 1989). Para la electroforesis de DNA en se utilizaron geles de agarosa al 0,8-2,5 % (p/v) en tampón TAE. El porcentaje de agarosa varió en función del tamaño de los fragmentos de DNA que se quiso analizar. A todos los geles se les añadió Red Safe (Intron Biotechnology).

Antes de realizar las electroforesis, a cada muestra se le añadió tampón de carga de DNA. Para determinar el tamaño de los fragmentos de DNA analizados se utilizó el patrón comercial 1 Kb plus DNA ladder (Invitrogen).

El DNA se visualizó con el transiluminador BioDoc-it (UVP), y se imprimió con la impresora Digital Graphic Printer UP-D895MD (Sony).

- **TAE:** Tris-acetato 40 mM; EDTA pH8 1 mM; ácido acético glacial 20 mM.

3.3. Purificación, cuantificación y secuenciación de DNA.

Para la purificación de DNA desde geles de agarosa se usó el kit comercial de Intron Biotechnology siguiendo las indicaciones del fabricante. La concentración de DNA en solución se estimó por espectrofotometría (Nanodrop One, de Thermo). La secuenciación de fragmentos de DNA se llevó a cabo en el servicio de secuenciación del IBiS, usando un secuenciador automático AB3500 de Applied Biosystems.

3.4. Amplificación de fragmentos de DNA mediante reacción en cadena de la polimerasa.

Para amplificación de fragmentos de DNA mediante reacción en cadena de la polimerasa (PCR) se utilizó el termociclador T100™ (BioRad). Los oligonucleótidos utilizados como cebadores en la reacción (**Tabla M-2**) fueron sintetizados por SIGMA.

En los ensayos de clonings se utilizaron las siguientes condiciones estándar para cada reacción: 10 μl de tampón comercial 5X Phusion, 1 μl de una mezcla de dNTPs 10 mM, 2,5 μl de cada oligonucleótido 10 μM , 0,5 μl de la polimerasa de DNA Phusion (2 U/ μl) y entre 1 y 6 ng de DNA molde en un volumen final de 50 μl . La reacción tipo constó de 1 ciclo de 30 segundos a 98 $^{\circ}\text{C}$; 35 ciclos de 10 segundos a 98 $^{\circ}\text{C}$, seguido de 30 segundos a la temperatura de apareamiento de los oligonucleótidos al DNA molde y posteriormente 30 segundos/Kb de DNA a amplificar a 72 $^{\circ}\text{C}$; para finalizar, 1 ciclo de 10 minutos a 72 $^{\circ}\text{C}$.

3.5. PCR a tiempo real.

Las PCRs cuantitativas (qPCR) se realizaron con un termociclador de PCR a tiempo real (Light Cycler 480 II, de Roche) que permite conocer la cantidad de DNA amplificado en cada momento de la reacción. Para la qPCR se mezclaron en frío 4 μl de la muestra de CDNA y 6 μl de una mezcla que contenía 0.5 μL de la mezcla (sentido y antisentido) de los oligonucleótidos a 10 μM , 0.5 μL de H_2O miliQ y 5 μL de SYBR Green Premix Ex Taq 2X (TAKARA). La reacción tipo constó de: 1 ciclo de 10 segundos a 95 $^{\circ}\text{C}$, 40 ciclos de [(a) 5 segundos a 95 $^{\circ}\text{C}$, (b) 20 segundos a 60 $^{\circ}\text{C}$], y 1 ciclo de [(a) 1 segundo a 95 $^{\circ}\text{C}$, (b) 11 segundos a 65 $^{\circ}\text{C}$, (c) 1 segundo a 95 $^{\circ}\text{C}$]. Para testar la eficiencia de los oligonucleótidos empleados (**Tabla M-2**) se realizó una recta patrón a partir de diluciones seriadas de una muestra de cDNA por cada par de oligonucleótidos y se comprobó si las pendientes de las rectas tenían un valor cercano a -3.3, lo que indica una eficiencia de amplificación del 100 %.

3.6. Transformación de bacterias por quimiocompetencia.

La transformación de *E.coli* DH5 α se llevó a cabo mediante transformación por choque térmico. En este proceso, 100 μl de solución de células competentes, previamente preparadas en el laboratorio utilizando el protocolo descrito por Hanahan (Hanahan 1983), se mezclaron con 50 ng de DNA. Luego, se incubó en hielo 10 minutos y a continuación se sometió a un choque térmico durante 1,5 minutos a 42 $^{\circ}\text{C}$. Después, se mantuvo en hielo durante 5 minutos. A continuación, se añadieron 800 μl de LB a la mezcla que se incubó una o dos horas a 37 $^{\circ}\text{C}$ en agitación. Transcurrido este tiempo se recogieron las células por centrifugación y se

sembraron en medio LB sólido con su correspondiente antibiótico para seleccionar aquellas bacterias portadoras del plásmido que contenga el gen de resistencia.

3.7. Transfección transitoria de líneas celulares animales con DNA plasmídico.

Las células se transfectaron utilizando TurboFect™ Transfection Reagent (ThermoFisher). Este reactivo es una mezcla de lípidos y otros componentes que permite transfectar DNA en una gran variedad de líneas celulares con una alta eficiencia y baja toxicidad. El día previo a la transfección, se sembraron entre 10^5 y 10^6 células. Para la transfección de células sembradas en placas de 6 pocillos, se mezclaron en un tubo 6 μ l del agente de transfección y 100 μ l de Opti-MEM™ (ThermoFisher). La mezcla se vorteo durante 5 minutos y luego, se añadió 2 μ g del plásmido y se incubó durante 15 minutos a temperatura ambiente. Por último, la mezcla se añadió gota a gota sobre las células en cultivo. Al día siguiente se añadió el antibiótico de resistencia G418 para seleccionar positivamente las células portadoras del plásmido. Aproximadamente 48 horas más tarde se procesaron las células para su análisis.

Para la generación de líneas estables modificadas, se siguió este protocolo. Los clones positivos fueron seleccionados tras varios días de cultivo en presencia de G418. Posteriormente fueron comprobados por citometría de flujo y/o ensayos de *western blot*.

4. MANIPULACIÓN Y ANÁLISIS DEL RNA.

4.1. Aislamiento y cuantificación del RNA de líneas celulares animales.

El aislamiento de RNA se llevó a cabo con el sistema RNAeasy mini kit (Qiagen) siguiendo las instrucciones del fabricante. La concentración de RNA en solución se estimó por espectrofotometría (Nanodrop One, de Thermo).

4.2. Tratamiento con DNAsa y reacción de retrotranscripción.

Para la obtención CDNA total a partir de RNA (1 μ g/reacción) se empleó hexanucleótidos y la enzima retrotranscriptasa (RT) utilizando el sistema Super-

Script™ III First Strand Synthesis for RT-PCR (Invitrogen) siguiendo las instrucciones del fabricante. Como control, una de las reacciones careció de la enzima.

Antes de proceder con la retrotranscripción, se eliminó el posible DNA contaminante. Para ello se utilizó la enzima *RQ1 RNase-Free DNase* (Promega) siguiendo las instrucciones del fabricante. Por cada μg de RNA a tratar se utilizó una unidad de DNasa en un volumen de reacción de 10 μl .

4.3. Transfección transitoria de líneas celulares animales con siRNAs.

La transfección transitoria de siRNAs fue realizada en placas de distintos tamaños de acuerdo con los requerimientos de cada experimento. Para placas de 60 mm, fueron sembradas 450000 células el día previo a la transfección. Entonces, el medio de cultivo fue eliminado y reemplazado por 2 ml de medio Opti-MEM™. En esta ocasión, el agente de transfección utilizado fue la Oligofectamina (Invitrogen). Para la transfección se mezclaron en un tubo (A) 15 μl de siRNA a 20 mM y 350 μl de Opti-MEM™ y en un tubo (B) 30 μl de Opti-MEM™ y 8 μl de Oligofectamina. Los tubos se incubaron a temperatura ambiente durante 7 minutos. A continuación, se mezcló el contenido de los tubos A y B y se incubaron durante 20 minutos. Por último, se añadió la mezcla a las células gota a gota y se incubaron durante 5 horas, tras las cuales se eliminó este medio y se añadió medio completo. Las células fueron cultivadas en estas condiciones otras 48-72 horas más.

5. EXTRACCIÓN Y ANÁLISIS DE PROTEÍNAS.

5.1. Obtención de extractos totales de proteínas de líneas celulares animales.

Para la obtención de un extracto proteico total, se recogieron en PBS y se centrifugaron a 900 x *g* durante 3 minutos. Los precipitados obtenidos se resuspendieron en tampón Laemmli 2X (10⁵ células/100 μl de tampón) y se hirvieron durante 5 minutos. A continuación, se sonicaron las muestras en el biorruptor fijado a la máxima potencia y con ciclos de encendido y apagado de 30 segundos durante 2 minutos.

- **PBS:** NaCl 138 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1,8 mM, KCl 3 mM, pH 7,4.
- **Tampón Laemmli 2X:** Tris-HCl pH 6,8 125 mM; glicerol 20 %; SDS 4 %; DTT 0,2M y azul de bromofenol 1 %.

5.2. Electroforesis de proteínas en geles de SDS-poliacrilamida.

La electroforesis de proteínas en condiciones desnaturalizantes se llevó a cabo según el protocolo descrito por Laemmli (Laemmli, 1970) para geles discontinuos de poliácridamida, utilizando el sistema Mini-Potearn II™ (BioRad). Los geles constaban de dos partes: la de empaquetamiento, que favorecía la correcta migración de las muestras; y la de separación, donde las proteínas migraban según su peso molecular. En este trabajo se usaron geles en los que su zona de separación contenía poliácridamida al 6 % o 12 % según el peso molecular de las proteínas a estudiar. Una vez cargados los geles, migraron a 180 V a temperatura ambiente durante 50 minutos. Como patrón de masa molecular se usó una mezcla de proteínas comerciales, Dual Color (BioRad). Tras finalizar la electroforesis, las proteínas presentes en los geles de poliácridamida fueron transferidas a membranas de nitrocelulosa para un posterior tratamiento de *western blot*.

- **Tampón de electroforesis:** Trizma base 25 mM, glicina 250 mM, SDS 0.1 %.

6. MÉTODOS INMUNOLÓGICOS.

6.1. Anticuerpos.

A continuación, se detallan los anticuerpos primarios (**Tabla M-5**) y secundarios (**Tabla M-6**) utilizados en esta tesis, indicando el antígeno que reconocen, su origen, procedencia y las diluciones usadas para inmunofluorescencia (IF), inmunoprecipitación de la cromatina (ChIP) y *western blot* (WB).

Tabla M-5: Anticuerpos primarios utilizados.

ANTICUERPO CONTRA:	ANTÍGENO	ORIGEN	PROCEDENCIA	IF	ChIP	WB
GAPDH	Proteína completa (humana)	Ratón monoclonal	Santa Cruz 47724	-	-	1:1000

ANTICUERPO CONTRA:	ANTÍGENO	ORIGEN	PROCEDENCIA	IF	ChIP	WB
PFDN1	aa 105-118 (humano)	Conejo monoclonal	Abcam 151708	-	-	1:500
PFDN2	aa 104-154 (humano)	Conejo policlonal	Bethyl A304-807A	-	-	1:100
PFDN3	aa 1-143 (humano)	Conejo policlonal	Abcam 96085	-	-	1:500
PFDN4	Proteína completa (humana)	Conejo policlonal	Abcam 168438	-	-	1:200
PFDN5	aa 1-154 (humano)	Ratón monoclonal	Santa Cruz 271150	-	-	1:100
PFDN6	aa 1-103 (humano)	Ratón monoclonal	Santa Cruz 376733	-	-	1:100
FLAG M2	DYKDDDDK	Ratón monoclonal	F3165	-	8 µg / ml	-
GFP	Completa	Ratón monoclonal	Roche	-	-	1:1000
Fibrilarina	aa 1-100 (humano)	Conejo policlonal	Abcam 5821	1 mg / ml	-	-
BrdU	Bromodeoxiuridina	Ratón monoclonal	Santa Cruz	1:1000	-	-
8WG16	Proteína completa (humana)	Ratón monoclonal	ThermoFisher MA1-10882	-	8 µg / ml	-
RNAPII CTD repeat (phospho S2)	P-Ser2 del dominio C-terminal	Conejo policlonal	Abcam 5095	-	5 µg / ml	-
p21	aa 1-100 (humano)	Conejo monoclonal	Abcam 109520	-	-	1:1000
H2A.X (Ser139)	aa 134-142 (humana)	Ratón monoclonal	Millipore 05-636	-	-	1:1000
LSM8	Dominio C-terminal de LSM8	Ratón monoclonal	Santa Cruz 81315	-	-	1:20
PRPF8 [2834C1a]	Proteína completa (humana)	Ratón monoclonal	Abcam 51366	-	-	1:100
CDK9	Dominio C-terminal	Conejo policlonal	Abcam 484	-	-	1:2000

Tabla M-6: Anticuerpos secundarios utilizados.

ANTICUERPO CONTRA:	ORIGEN	CONJUGADO	PROCEDENCIA	IF	WB
IgG ratón	Cabra	Peroxidasa	Jackson 115-035-146	-	1:10000
IgG conejo	Cabra	Peroxidasa	Jackson 111-035-144	-	1:10000
IgG ratón (H + L)	Cabra	Alexa Fluor 488	Invitrogen A11001	1:1000	-
IgG conejo (H + L)	Cabra	Alexa Fluor 546	Invitrogen A11010	1:1000	-

6.2. Ensayo de *western blot* (WB).

Los WBs realizados en este trabajo se detectaron por quimioluminiscencia. Para ello, primero las proteínas fueron separadas en geles de SDS-poliacrilamida (ver apartado 5.2) y transferidas a una membrana de nitrocelulosa (Amersham™ Protran™ 0,45 µm, de GE Healthcare). La electrotransferencia húmeda se realizó en tampón de transferencia durante 45 minutos o 17 horas, dependiendo del peso molecular de la proteína que se quería analizar, aplicando una diferencia de potencial de 80V o 35V respectivamente. Tras finalizar el proceso las membranas se bloquearon con 5 % de leche en TTBS (LTTBS) durante una hora en agitación. La incubación con el anticuerpo primario (**Tabla M-5**) se realizó o/n a 4 °C en la misma solución. A continuación, y usando de nuevo la misma solución, se procedió a lavar la membrana 3 veces (5 minutos cada lavado) en agitación. Tras los lavados, la incubación con el anticuerpo secundario conjugado con peroxidasa de rábano (**Tabla M-6**), se realizó en solución de bloqueo durante 1 hora a temperatura ambiente. Posteriormente se procedió a lavar las membranas de la forma mencionada anteriormente. A continuación, se llevó a cabo el revelado de la actividad peroxidasa con un kit de inmunodetección quimioluminiscente de gran sensibilidad (Clarity™ Western ECL Substrate, de BioRad) siguiendo las instrucciones

del fabricante. Para la detección de la señal luminosa se utilizó el ChemiDoc-Touch (Biorad).

- **Tampón de transferencia:** Trizma base 25 mM, glicina 250 mM, SDS 0.1 %, metanol 20%.
- **TTBS:** Tris-HCl 15 mM pH 7.5, NaCl 200 mM y Tween 20 0.1%,

6.3. Co-inmunoprecitación.

Para las co-inmunoprecipitaciones realizadas en esta tesis se precipitó PFDN5 etiquetada con Enhanced Green Fluorescent Protein (EGFP). Para ello, se usó GFP Trap A (Chromotek), que son pequeñas bolas de agarosa a las que se les ha añadido el dominio V_HH (zona variable de la cadena pesada) del anticuerpo anti-GFP, que son los fragmentos intactos más pequeños (con un peso molecular aproximado de ~ 13 KDa) capaces de unir el antígeno. Estas nanopartículas son muy estables, específicas y tienen gran afinidad por GFP y proteínas fusionadas a GFP.

Primero, se recogió un pellet de unas 30×10^6 células que se resuspendió en 400 μ l de tampón de lisis 1 y se incubó 30 minutos en hielo. Posteriormente se añadió 50 μ l de DSP 1 mM, un entrecruzador que mantiene unidos enlaces débiles, y se incubó 20 minutos a 20 °C. Las muestras se centrifugaron a alta velocidad durante 15 minutos a 4 °C. El pellet se resuspendió en tampón de lisis 2, y se sonicó en el biorruptor fijado a la máxima potencia y con ciclos de encendido y apagado de 30 segundos durante 2 minutos. Luego, se añadió benzonasa (25 U / 100 μ l) a las muestras, se incubaron en hielo durante 30 minutos y se centrifugaron a alta velocidad durante 15 minutos a 4 °C. Después, se apartó el sobrenadante a un nuevo tubo y se ajustó su volumen a 1 ml con tampón de dilución. El GFP Trap se equilibró con tampón de dilución y luego se adicionaron 40 μ l del mismo a cada tubo con el lisado total. A continuación, se dejaron en agitación a 4 °C durante dos horas y una vez pasado este tiempo, se lavaron las muestras y se realizaron unas centrifugaciones seriadas para separar el sobrenadante del pellet en el que se encontraban las nanopartículas unidas a la proteína de interés. Finalmente se

adicionó tampón Laemmli 2X a cada tubo. Todas las muestras fueron hervidas durante 5 minutos y congeladas a $-20\text{ }^{\circ}\text{C}$ hasta su uso.

- **Tampón de lisis 1:** Tris- HCl 10 mM pH 7.5, NaCl 10 mM, EDTA 0,5 mM, MgCl_2 2,5 mM e inhibidores de proteasas (1/100).
- **Tampón de lisis 2:** Tris- HCl 10 mM pH 7.5, NaCl 150 mM, EDTA 0,5 mM, MgCl_2 2,5 mM, NP-40 0,5% e inhibidores de proteasas (1/100).
- **Tampón de dilución:** Tris- HCl 10 mM pH 7.5, NaCl 150 mM, EDTA 0,5 mM e inhibidores de proteasas (1/100).

6.4. Inmunoprecipitación de la cromatina (ChIP).

En los experimentos de ChIP se usó el protocolo descrito (Orlando, Strutt and Paro, 1997) y las soluciones que se muestran al final del protocolo. Unos 30×10^6 de células sujetas a diferentes condiciones de crecimiento se fijaron añadiendo, directamente al medio de cultivo, una solución del entrecruzador formaldehído a una concentración final de 1% (p/v), y se incubaron 10 minutos a $37\text{ }^{\circ}\text{C}$. A continuación, la fijación se detuvo añadiendo glicina a una concentración final de 0,1 M e incubando la placa durante 5 minutos a temperatura ambiente. La glicina reacciona bloqueando el agente entrecruzador.

A continuación, las células se lavaron dos veces usando PBS frío, se recogieron utilizando un raspador en PBS frío complementado con inhibidores de proteasas Complete *protease inhibitor cocktail tablets* (Roche) y PMSF 1 mM (Sigma) y se centrifugaron 5 minutos a $3600 \times g$ y $4\text{ }^{\circ}\text{C}$. El sobrenadante se eliminó y el *pellet* se resuspendió en 2,5 ml de tampón de lisis. Tras incubar 5 minutos y centrifugar durante 5 minutos a $3600 \times g$ y $4\text{ }^{\circ}\text{C}$, el sobrenadante se eliminó y el *pellet* se resuspendió en 1 ml de tampón de lisis nuclear. El lisado se sonicó en hielo para obtener fragmentos de DNA de entre 400 y 500 pb usando un baño de ultrasonidos *Bioruptor* (Diogenode), fijado a la máxima potencia y con ciclos de encendido y apagado de 30 segundos durante 10 minutos. Después de la sonicación, el material se centrifugó 5 minutos a $3600 \times g$ y $4\text{ }^{\circ}\text{C}$ para tomar el sobrenadante (cromatina) y descartar el *pellet* (restos celulares). Una alícuota de la cromatina se trató con

Proteinasas K y el DNA se recuperó mediante extracción con fenol/cloroformo para cuantificar la concentración del DNA y comprobar el tamaño de los fragmentos de DNA en un gel de agarosa al 1,2 %.

Para llevar a cabo la inmunoprecipitación de la cromatina, se utilizó 100 μ l de cromatina diluida 10 veces en tampón IP complementado con inhibidores de proteasas *Complete protease inhibitor cocktail tablets* (Roche) y PMSF 1 mM (Sigma) y se añadió el anticuerpo. La mezcla se incubó durante 4 horas a 4 °C en agitación. A continuación, se añadió 30 μ l de Dynabeads Protein A o G (Life technology), previamente lavadas en tampón IP, y se incubó durante 2 horas a 4 °C en agitación. Tras la incubación, la cromatina se lavó con cada uno de los tres tampones de lavado y luego con el tampón TE, utilizando en todos los casos la gradilla magnética que retuvo a las Dynabeads con la cromatina de interés. Después, el DNA se eluyó incubándolo con tampón de elución durante 10 minutos a 65 °C y con ayuda de una gradilla magnética se retiró el DNA del vial con cuidado de no tomar las Dynabeads.

Para la purificación del DNA se tomó 60 μ l de este y se incubó con 90 μ l de tampón de elución durante toda la noche a 4 °C. Paralelamente se realizaron los inputs con 10 μ l de DNA incubado con 140 μ l de tampón de elución en las mismas condiciones. A continuación, las muestras se trataron con Proteinasas K (10 mg/ml) y el DNA se recuperó mediante extracción con fenol/cloroformo. El DNA se precipitó con etanol absoluto, 10% de acetato sódico (v/v) y 0,1% de glicógeno (v/v), y luego fue lavado con etanol al 70 % (v/v). Finalmente, se resuspendió el pellet de DNA en 200 μ l de agua libre de DNAsas y se mantuvo a -20 °C.

La cuantificación de las regiones de interés del DNA resultante se realizó con un termociclador de PCR a tiempo real (ver apartado 3.5).

- **Solución de fijación:** Hepes 50 mM pH 8,0; NaCl 0,1 M; EDTA 1 mM pH 8,0; EGTA 0,5 mM pH 8,0.
- **Tampón de lisis:** Pipes 5 mM pH 8,0; KCl 85 mM; NP-40 0,5% (v/v).

- **Tampón de lisis nuclear:** SDS 1% (p/v); EDTA 10 mM pH 8,0; Tris-HCl 50 mM pH 8,1.
- **Tampón de CHIP:** SDS 0,01% (p/v); Triton X-100 1.1% (p/v); EDTA 1.2 mM pH 8,0; Tris-HCl 16,7 mM pH 8,1; NaCl 167 mM.
- **Tampón de lavado 1:** SDS 0,1% (p/v); Triton-X100 1% (v/v); EDTA 2 mM pH 8,0; Tris-HCl 20 mM pH 8,1; NaCl 150 mM.
- **Tampón de lavado 2:** SDS 0,1% (p/v); Triton-X100 1% (v/v); EDTA 2 mM pH 8,0; Tris-HCl 20 mM pH 8,1; NaCl 500 mM.
- **Tampón de lavado 3:** LiCl 0,25 M; NP-40 1% (v/v); Sodium Deoxycholate 1% (p/v); EDTA 1 mM pH 8,0; Tris-HCl 10 mM pH 8,1.
- **Tampón de elución:** SDS 1% (p/v); NaHCO₃ 0,1 M.
- **Tampón TE 1X:** Tris-HCl 10mM pH 8,0; EDTA 1 mM pH 8,0.

6.5. Inmunofluorescencia (IF).

En los ensayos de inmunofluorescencia, unas 150.000 células crecidas en monocapas y adheridas a lamelas de vidrio fueron, primero fijadas con PBS-PFA 4 % durante 10 minutos a temperatura ambiente, y después permeabilizadas con PBS-Tritón 0,2 % durante 2 minutos a 4°C. Tras esto, se realizaron 2 lavados con PBS y posteriormente se sumergieron las lamelas en solución de bloqueo durante 1 hora a temperatura ambiente. A continuación, las células se incubaron con el anticuerpo primario diluido en PBS-BSA 1% durante al menos 1 hora a temperatura ambiente. Transcurrido este tiempo, se hicieron 3 lavados de 10 minutos con PBS-Tween 0,1 % y se incubaron con los anticuerpos secundarios (Tabla 7) disueltos también en PBS-BSA 1% durante 1 hora a temperatura ambiente en oscuridad. Posteriormente se hicieron 3 lavados de 15 minutos con PBS-Tween 0,1 %. Después, se añadió sobre las muestras 4'-6-Diamino-2-fenilinol (DAPI) diluido en PBS a 0,1 µg/ml como fluorocromo marcador de DNA, y se incubaron durante 5 minutos. Luego, se hicieron dos lavados de 5 minutos, el primero con PBS y el segundo con agua miliQ. Finalmente, y una vez que las muestras se secaron, se montaron en portaobjetos sobre medio de montaje Vectashield Antifade Mounting Medium (H-1000 Vector Labs).

- **PBS-PFA 4 %:** PFA (Thermo Scientific) 4 % en PBS 1X.
- **PBS-Tritón 0,2 %:** Tritón™ (X-100 Sigma-Aldrich) 0,2 % en PBS 1X.
- **PBS-Tween 0,1 %:** Tween®20 (Sigma-Aldrich) 0,1 % en PBS 1X.
- **Solución de bloqueo:** BSA (Albumin Fraction V, PanReac) 5% en PBS 1X.

7. MICROSCOPIA DE FLUORESCENCIA.

Las preparaciones de esta tesis se observaron usando dos tipos diferentes de microscopio: el microscopio directo Olympus BX61, configurado con la lámpara BX-RFAA tanto para luz transmitida como para fluorescencia, o el microscopio confocal Nikon A1R+, con los filtros para DAPI, 488 y 546. Los softwares que controlan estos microscopios son el Cellprofiler, y el NIS-Element, respectivamente.

8. DETERMINACIÓN DE FOCOS DE γ -H2A.X: IF Y CONTEO.

Para realizar este experimento se sembraron, por duplicado, 10.000 células en placas de 24 pocillos con lamelas de vidrio. Al día siguiente se procedió a la transfección transitoria del siPFDN5 y el siControl con Oligofectamine™ Reagent (ver apartado 4.3). Pasadas 24 horas, a unas de las placas se le reemplazó el medio completo por medio sin suero y 48 horas después se hicieron las IF. Como control positivo se utilizaron células tratadas con etopóxido 20 μ M durante una hora, y como control negativo, células sin tratar. Para la IF se utilizó el anticuerpo primario anti γ -H2A.X (Ser139) y el anticuerpo secundario Alexa Flúor 488. Las muestras se analizaron en el microscopio directo Olympus BX61 y posteriormente fueron procesadas con el ImageJ. En el ensayo se cuantificaron 30 células.

9. ENSAYO DE VIABILIDAD CELULAR.

Para evaluar el efecto de la inhibición de la expresión de *PFDN2* y *PFDN5* en la evolución de un cultivo a lo largo del tiempo, se realizó un ensayo de viabilidad celular que consistió en cuantificar el número de células vivas en la placa de cultivo 24, 48 y 72 horas después de transfectar los RNA de interferencia siPFDN2 y siPFDN5 con Oligofectamine™ Reagent. Como control, se utilizaron cultivos transfectados con el siControl. Para las transfecciones se usaron placas P60

(Thermo). En cada pocillo se cultivaron 250.000 células que 24 h después se transfectaron según el protocolo habitual, y se recogieron para su conteo a las horas indicadas. Para contar las células se usó una cámara Neubauer. Cada conteo se realizó 3 veces y se calculó la media y la desviación estándar de los valores obtenidos.

10. ANÁLISIS DEL CICLO CELULAR POR CITOMETRÍA DE FLUJO.

Para realizar este experimento las células se trataron con BrdU 10 μ M durante 15 minutos a 37 $^{\circ}$ C. A continuación, se tripsinizaron, se recogieron y se centrifugaron a 900 x *g* durante 5 minutos a 23 $^{\circ}$ C. Posteriormente se hizo un lavado con PBS y se volvió a centrifugar en las mismas condiciones. Después, las muestras se fijaron con 1 ml de etanol frío al 70 %, gota a gota y en agitación constante con el vortex a mínima velocidad. Para realizar el marcaje con yoduro de propidio (PI), las células se centrifugaron a 900 x *g* durante 3 minutos, y el pellet obtenido fue resuspendido en 1 ml de 2 N HCl/Tritón x-100 en agitación constante con el vortex a mínima velocidad. Las muestras se incubaron 30 minutos a temperatura ambiente y luego, se centrifugaron 3 minutos a 900 x *g*. Para neutralizar las muestras, el pellet obtenido se resuspendió en 1 ml de 0,1 M Na₂Ba₄O₇, pH 8,5, y se centrifugó en las mismas condiciones. El pellet obtenido se resuspendió en 500 μ l de una mezcla madre que contenía PBS 0,5 % Tween 20/0,1 % BSA y el anticuerpo anti BrdU diluido. Las muestras se incubaron girando toda la noche a 4 $^{\circ}$ C. Al día siguiente, se centrifugaron las células 3 minutos a 900 x *g*, y el pellet obtenido se resuspendió en 500 μ l de una mezcla madre que contenía PBS 0,5 % Tween 20 / 0,1 % BSA y el anticuerpo Alexa Fluor 488 diluido. Las muestras se incubaron una hora a temperatura ambiente, y posteriormente se centrifugaron en las mismas condiciones. El pellet fue resuspendido en 500 μ l de una mezcla que llevaba PBS, RNase 10 mg / ml y 1.24 μ l de PI, y se incubaron durante 30 minutos en oscuridad a temperatura ambiente. Finalmente, las células se centrifugaron 3 minutos a 900 x *g*, y el pellet resuspendió en 500 μ l de PBS. Para evitar la agregación, las muestras fueron filtradas antes de pasarlas por el citómetro.

Este análisis se hizo en diferentes poblaciones de células, unas transfectadas con el siControl y otras transfectadas con el siPFDN5. Las transfecciones se hicieron con Oligofectamine™ Reagent (ver apartado 4.3). Además, los ensayos se llevaron a cabo con y sin suero, es decir, hubo poblaciones de células a las que se les privó de suero al día siguiente de ser transfectadas, y que crecieron en estas condiciones durante 48 horas y hubo poblaciones de células que crecieron en medio completo.

El experimento también fue ensayado en poblaciones de células WT y poblaciones de células que carecían de PFDN5. Del mismo modo los ensayos se llevaron a cabo con y sin suero.

11. ANÁLISIS MASIVO DE LA EXPRESIÓN GÉNICA MEDIANTE RNA-seq.

11.1. Metodología.

Para analizar el efecto de la depleción de dos de las subunidades del complejo Prefoldina, PFDN2 y PFDN5, en la expresión de genes a escala genómica, se realizó una secuenciación masiva del RNA Poly (A) (Poly (A) RNA-seq). Con el objetivo de estudiar mejor posibles defectos en la expresión de genes, el ensayo se desarrolló en condiciones de hambre de suero, para posteriormente añadir suero y estimular la expresión génica.

En este ensayo se partió de 375.000 células sembradas en placas P60 por condición, que 24 horas después, fueron transfectadas por separado con 3 RNA de interferencia diferentes: el siControl (condición control), y el siPFDN2 y siPFDN5 (para reducir los niveles de las subunidades de prefoldina 2 y 5 respectivamente). Las transfecciones se hicieron con Oligofectamine™ Reagent (ver apartado 4.3). Al día siguiente, a las células se les sustituyó el medio completo por medio sin suero y fueron cultivadas en estas condiciones durante 48 horas. Tras este periodo de tiempo, se reemplazó el medio sin suero por medio completo y se extrajo RNA total de las diferentes poblaciones celulares a tres tiempos distintos: antes de la inducción de la expresión con suero (tiempo 0), y 30 y 90 minutos después de la inducción. El RNA extraído fue tratado con DNAsa y se generó cDNA mediante retrotranscripción.

11.2. Análisis bioinformático.

La secuenciación masiva de las muestras de RNA se llevó a cabo en la Unidad de Genómica del Centro de Regulación Genómica de Barcelona. Para este trabajo se optó por una secuenciación total y pareada (*paired end*) del RNA. Las librerías correspondientes fueron preparadas tras eliminar la señal correspondiente a los RNAs ribosómicos de los datos.

Las secuencias obtenidas se sometieron a pruebas de control de calidad con el paquete FASTQC v.0.11.5. Tras esto, fueron alineadas al genoma de referencia (obtenido desde la web www.ensembl.org/info/data/ftp/index.html, versión Homo_sapiens_GRCh37.74_chr); usando el programa Kallisto (Bray *et al.*, 2016) para el estudio de la expresión diferencial entre las distintas muestras, y el programa HISAT2 (Pertea *et al.*, 2017) para calcular la ratio exónica. En este caso, las secuencias se mapearon al genoma completo y luego se cuantificaron aquellas que mapeaban dentro de las coordenadas de los genes completos (según la anotación de ENSEMBL versión 79) y las que mapeaban dentro de los exones.

Para el análisis de expresión diferencial entre las muestras transfectadas con el siControl, el siPFDN2 y el siPFDN5, en el tiempo 0', 30' y 90', y el análisis de los cambios de expresión de PFDN2 y PFDN5 con respecto al tiempo 0', se usó el paquete edgeR de Bioconductor, considerando expresados diferencialmente todos los genes con un FDR<0.05.

En el análisis de enriquecimiento funcional de las muestras transfectadas con el siPFDN2 y el siPFDN5 a tiempo 0', 30' y 90', se utilizó el paquete GOFfunction de Bioconductor seleccionando términos GO con un FDR<0.05.

El análisis de las ratios exónicas de las muestras transfectadas con el siControl, siPFDN2 y siPFDN5 a tiempo 0', 30' y 90', se llevó a cabo con el entorno estadístico R, mientras que las gráficas se generaron con el paquete ggplot2 de R.

Para analizar posibles cambios en el *splicing* alternativo entre el control y la muestra con PFDN5 deplecionada a tiempo 0', 30' y 90', se utilizó el paquete

DESeq2 de R, considerando diferencialmente expresados los exones con un $FDR < 0.05$.

Para determinar el efecto de la longitud o el número de intrones sobre la ratio de expresión siPFDN5/siControl y la ratio intrones/exones a tiempo 0', 30', y 90', se utilizó la anotación de genes de ENSEMBL versión 79., y el paquete estadístico R.

12. ENSAYOS DE *SPLICING*.

12.1. Minigenes.

Para analizar el efecto de PFDN5 en el *splicing*, células de las líneas HCT116 WT, HCT116 PFDN5 KO y HCT116 PFDN5 KO + GFP-PFDN5 fueron transfectadas con 0,5 μg del minigen de *Fas* usando turbofect (ver apartado 3.7). El RNA total era extraído 24 horas más tarde, y 800 ng del RNA obtenido se retro-transcribía con *MLV retrotranscriptase* (Invitrogen) durante 1 h a 37°C utilizando el oligo *csnPT2* para el minigen *Fas*. Un cuarto de la reacción se amplificó tras 35 ciclos de PCR (T_m 61,5°C) utilizando las siguientes condiciones estándar para cada reacción: 2,5 μl de tampón comercial 10X, 2 μl de MgCl_2 , 1 μl de una mezcla de dNTPs 10 mM, 2,5 μl de los oligonucleótidos *csnPT1/csnPT2* a 10 μM , 0,5 μl de la polimerasa de DNA. Los productos de la PCR fueron resueltos en geles de agarosa al 2,5%. Las intensidades de las bandas fueron cuantificadas con el programa *Image Lab 6.0* (Bio-Rad). El experimento se realizó cinco veces. Los valores están representados como la ratio entre la isoforma corta y la isoforma larga del gen.

12.2. Análisis de la velocidad y procesividad de la RNAPII y el *splicing* cotranscripcional.

En este ensayo se sembraron por duplicado 375.000 células en placas P60. Al día siguiente se procedió a la transfección del siPFDN5 y el siControl utilizando para ello Oligofectamine™ Reagent. 72 horas después de transfectar los RNA de interferencia, se añadió al medio de cultivo el inhibidor de la elongación transcripcional DRB (5,6-Dichloro-1- β -D-dibofuranosylbenzimidazole) a 1 μM , y se dejó actuar durante 3 horas en una atmósfera húmeda, a 37°C, y con CO_2 al 5%.

Como control, se dejó una placa por condición sin tratar con DRB. A continuación, las células fueron lavadas con PBS y se les puso de nuevo medio completo fresco. A partir de este momento, RNA total fue extraído cada 10 minutos de cada una de las poblaciones. El RNA obtenido fue tratado con DNAsa y retrotranscrito (ver apartado 4.2). Por último, a partir del CDNA obtenido se realizó una PCR a tiempo real (ver apartado 3.5) utilizando oligonucleótidos repartidos a lo largo del gen *CTNBL1*. Por un lado, para ver el efecto de la depleción de PFDN5 en la velocidad de la RNAPII, se utilizaron oligonucleótidos que amplificaban parte de un exón y su intrón adyacente. Por otro lado, para analizar el *splicing* cotranscripcional entre dichas condiciones, se utilizaron oligonucleótidos que amplificaban parte un exón y del intrón adyacente al exón posterior a este (**Figura M-1**).

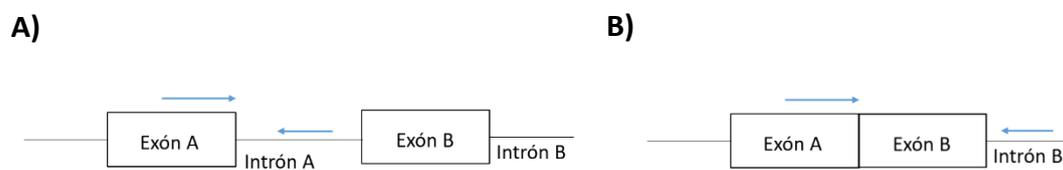


Figura M-1: Localización de los oligonucleótidos. A y B) Dibujo esquemático donde se muestra con flechas azules, el lugar donde se localizarían los oligonucleótidos utilizados según si se quiere analizar la transcripción de un intrón (**A**) o la eliminación de este (**B**).

Bibliografía

- Abe, A., Takahashi-Niki, K., Takekoshi, Y., Shimizu, T., Kitaura, H., Maita, H., Iguchi-Arigo, S.M.M., and Arigo, H. (2013). Prefoldin plays a role as a clearance factor in preventing proteasome inhibitor-induced protein aggregation. *J. Biol. Chem.* *288*, 27764–27776.
- Alamancos, G.P., Pagès, A., Trincado, J.L., Bellora, N., and Eyra, E. (2015). Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA* *21*, 1521–1531.
- Allfrey, G., Faulkner, R., and Mirsky, A.E. (1964). Possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci U.S.A.* *315*, 786–794.
- Allison, L.A., Moyle, M., Shales, M., and Ingles, C.J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* *42*, 599–610.
- Arranz, R., Martín-Benito, J., and Valpuesta, J.M. (2018). Structure and function of the cochaperone Prefoldin. In *Prefoldins: The New Chaperones*, pp. 119–131.
- Ast, G. (2004). How did alternative splicing evolve? *Nat. Rev. Genet.* *5*, 773–782.
- Bahrami, S., and Drabløs, F. (2016). Gene regulation in the immediate-early response process. *Adv. Biol. Regul.* *62*, 37–49.
- Becerra, S., Montes, M., Hernández-Munain, C., and Suñé, C. (2015). Prp40 pre-mRNA processing factor 40 homolog B (PRPF40B) associates with SF1 and U2AF 65 and modulates alternative pre-mRNA splicing in vivo. *RNA* *21*, 438–457.
- Bertoli, C., Skotheim, J.M., and de Bruin, R.A.M. (2013). Control of cell cycle transcription during G1 and S phases. *Nat. Rev. Mol. Cell Biol.* *14*, 518–528.
- Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* *72*, 291–336.
- Blazek, D., Kohoutek, J., Bartholomeeusen, K., Johansen, E., Hulinkova, P., Luo, Z., Cimermancic, P., Ule, J., and Peterlin, B.M. (2011). The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. *Genes Dev.* *25*, 2158–2172.
- Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal RNA-Seq quantification. *Nat. Biotechnol.* 1–21.
- Bretones, G., Delgado, M.D., and León, J. (2015). Myc and cell cycle control. *Biochim. Biophys. Acta* *1849*, 506–516.
- Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. *Mol. Cell* *36*, 541–546.
- Carrillo Oesterreich, F., Herzel, L., Straube, K., Hujer, K., Howard, J., and Neugebauer, K.M. (2016). Splicing of nascent RNA coincides with intron exit from RNA Polymerase II. *Cell* *165*, 372–381.
- 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*, 2706–2713.

- Castillo-Davis, C.I., Mekhedov, S.L., Hartl, D.L., Koonin, E. V., and Kondrashov, F.A. (2002). Selection for short introns in highly expressed genes. *Nat. Genet.* *31*, 415–418.
- 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*, 1780–1782.
- Chapman, R.D., Heidemann, M., Hintermair, C., and Eick, D. (2008). Molecular evolution of the RNA polymerase II CTD. *Trends Genet.* *24*, 289–296.
- Chávez, S., and Puerto-Camacho, P. (2016). Prefoldins. In *ELS*, (Chichester, UK: John Wiley & Sons, Ltd), pp. 1–8.
- Chávez, S., García-Rubio, M., Prado, F., and Aguilera, A. (2001). Hpr1 is preferentially required for transcription of either long or G+C-rich DNA sequences in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* *21*, 7054–7064.
- Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* *10*, 741–754.
- Chen, M., Huang, J., Yang, X., Liu, B., Zhang, W., Huang, L., Deng, F., Ma, J., Bai, Y., Lu, R., et al. (2012). Serum starvation induced cell cycle synchronization facilitates human somatic cells reprogramming. *PLoS One* *7*, e28203.
- 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*, 1759–1762.
- Chiaromonte, F., Miller, W., and Bouhassira, E.E. (2003). Gene Length and Proximity to Neighbors Affect Genome-Wide Expression Levels. *Genome Res.*
- Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* *30*, 250–255.
- De Conti, L., Baralle, M., and Buratti, E. (2013). Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip. Rev. RNA* *4*, 49–60.
- Cooper, T.A. (2005). Use of minigene systems to dissect alternative splicing elements. *Methods* *37*, 331–340.
- Corden, J.L., Cadena, D.L., Ahearn, J.M., 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*, 7934–7938.
- Cramer, P. (2002). Common structural features of nucleic acid polymerases. *BioEssays* *24*, 724–729.
- David, C.J., Boyne, A.R., Millhouse, S.R., and Manley, J.L. (2011). The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev.* *25*, 972–983.

- Delgehyr, N., Wieland, U., Rangone, H., Pinson, X., Mao, G., Dzhindzhev, N.S., McLean, D., Riparbelli, M.G., Llamazares, S., Callaini, G., et al. (2012). Drosophila Mgr, a Prefoldin subunit cooperating with von Hippel Lindau to regulate tubulin stability. *Proc. Natl. Acad. Sci.* *109*, 5729–5734.
- Egloff, S., and Murphy, S. (2008). Cracking the RNA polymerase II CTD code. *Trends Genet.* *24*, 280–288.
- Egloff, S., O'Reilly, D., Chapman, R.D., Taylor, A., Tanzhaus, K., Pitts, L., Eick, D., and Murphy, S. (2007). Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* *318*, 1777–1779.
- Einhauer, A., and Jungbauer, A. (2001). The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. *J. Biochem. Biophys. Methods* *49*, 455–465.
- Erismann, M.D., Rothberg, P.G., Diehl, R.E., Morse, C.C., Spandorfer, J.M., and Astrin, S.M. (1985). Deregulation of c-myc gene expression in human colon carcinoma is not accompanied by amplification or rearrangement of the gene. *Mol. Cell. Biol.* *5*, 1969–1976.
- Ewen, M.E. (2000). Where the cell cycle and histones meet. *Genes Dev.* *14*, 2265–2270.
- Ferdous, A., Gonzalez, F., Sun, L., Kodadek, T., and Johnston, S.A. (2001). The 19S regulatory particle of the proteasome is required for efficient transcription elongation by RNA polymerase II. *Mol. Cell* *7*, 981–991.
- Fong, N. (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*, 4274–4282.
- Fong, Y.W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. *Nature* *414*, 929–933.
- Fragkos, M., Jurvansuu, J., and Beard, P. (2009). H2AX is required for cell cycle arrest via the p53/p21 pathway. *Mol. Cell. Biol.* *29*, 2828–2840.
- Fujioka, Y., Taira, T., Maeda, Y., Tanaka, S., Nishihara, H., Iguchi-Ariga, S.M.M., Nagashima, K., and Ariga, H. (2001). MM-1, a c-Myc-binding protein, is a candidate for a tumor suppressor in leukemia/lymphoma and tongue cancer. *J. Biol. Chem.* *276*, 45137–45144.
- García-Muse, T., and Aguilera, A. (2016). Transcription–replication conflicts: how they occur and how they are resolved. *Nat. Rev. Mol. Cell Biol.* *17*, 553–563.
- Gartel, A.L., Ye, X., Goufman, E., Shianov, P., Hay, N., Najmabadi, F., and Tyner, A.L. (2001). Myc represses the p21(WAF1/CIP1) promoter and interacts with Sp1/Sp3. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 4510–4515.
- Gates, L.A., Shi, J., Rohira, A.D., Feng, Q., Zhu, B., Bedford, M.T., Sagum, C.A., Jung, S.Y., Qin, J., Tsai, M.-J., et al. (2017). Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation. *J. Biol. Chem.* *292*, 14456–14472.
- Geissler, S., Siegers, K., and Schiebel, E. (1998). A novel protein complex promoting formation of functional alpha- and gamma-tubulin. *EMBO J.* *17*, 952–966.

- Gilbert, W. (1978). Why genes in pieces? *Nature* 271, 501.
- Glover, D.J., and Clark, D.S. (2015). Oligomeric assembly is required for chaperone activity of the filamentous γ -prefoldin. *FEBS J.* 282, 2985–2997.
- Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: a landscape takes shape. *Cell* 128, 635–638.
- Gómez-Herreros, F., Schuurs-Hoeijmakers, J.H.M., McCormack, M., Grealley, M.T., Rulten, S., Romero-Granados, R., Counihan, T.J., Chaila, E., Conroy, J., Ennis, S., et al. (2014). TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nat. Genet.* 46, 516–521.
- Gong, Y., Kakahara, Y., Krogan, N., Greenblatt, J., Emili, A., Zhang, Z., and Houry, W.A. (2009). An atlas of chaperone–protein interactions in *Saccharomyces cerevisiae*: implications to protein folding pathways in the cell. *Mol. Syst. Biol.* 5.
- Grünberg, S., and Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem. Sci.* 38, 603–611.
- Gstaiger, M., Luke, B., Hess, D., Oakeley, E.J., Wirbelauer, C., Blondel, M., Vigneron, M., Peter, M., and Krek, W. (2003). Control of nutrient-sensitive transcription programs by the unconventional prefoldin URI. *Science* 302, 1208–1212.
- Gu, B., Eick, D., and Bensaude, O. (2013). CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Res.* 41, 1591–1603.
- Hagio, Y., Kimura, Y., Taira, T., Fujioka, Y., Iguchi-Ariga, S.M.M., and Ariga, H. (2006). Distinct localizations and repression activities of MM-1 isoforms toward c-Myc. *J. Cell. Biochem.* 97, 145–155.
- Hansen, W.J., Cowan, N.J., and Welch, W.J. (1999). Prefoldin-nascent chain complexes in the folding of cytoskeletal proteins. *J. Cell Biol.* 145, 265–277.
- Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature* 475, 324–332.
- Harvey, R.H. (1991). Primary response genes induced by growth factors and tumor promoters. *Annu. Rev. Biochem.* 60, 281–319.
- Healy, S., Khan, P., and Davie, J.R. (2013). Immediate early response genes and cell transformation. *Pharmacol. Ther.* 137, 64–77.
- Hendrickson, S., Cochran, B., Reffel, A., and Stiles, C. (1984). Regulation of gene expression by serum and serum growth factors. *Symp. Fundam. Cancer Res.* 37, 71–85.
- Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat. Rev. Mol. Cell Biol.* 18, 637–650.
- Hintermair, C., Voß, K., Forné, I., Heidemann, M., Flatley, A., Kremmer, E., Imhof, A., and Eick, D. (2016). Specific threonine-4 phosphorylation and function of RNA

- polymerase II CTD during M phase progression. *Sci. Rep.* **6**, 27401.
- 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**, 1289–1295.
- 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**, 683–686.
- Huranová, M., Ivani, I., Benda, A., Poser, I., Brody, Y., Hof, M., Shav-Tal, Y., Neugebauer, K.M., and Staněk, D. (2010). The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *J. Cell Biol.* **191**, 75–86.
- Iyer, V.R., Eisen, M.B., Ross, D.T., Schuler, G., Moore, T., Lee, J.C., Trent, J.M., Staudt, L.M., Hudson, J., Boguski, M.S., et al. (1999). The transcriptional program in the response of human fibroblasts to serum. *Science* **283**, 83–87.
- Jeronimo, C., Bataille, A.R., and Robert, F. (2013). The writers, readers, and functions of the RNA polymerase II C-terminal domain code. *Chem. Rev.* **113**, 8491–8522.
- Jimeno-González, S., Payán-Bravo, L., Muñoz-Cabello, A.M., Guijo, M., Gutierrez, G., Prado, F., and Reyes, J.C. (2015). Defective histone supply causes changes in RNA polymerase II elongation rate and cotranscriptional pre-mRNA splicing. *Proc. Natl. Acad. Sci.* **112**, 14840–14845.
- Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 167–177.
- Jonkers, I., Kwak, H., and Lis, J.T. (2014). Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3**, e02407.
- Karimian, A., Ahmadi, Y., and Yousefi, B. (2016). Multiple functions of p21 in cell cycle, apoptosis and transcriptional regulation after DNA damage. *DNA Repair (Amst.)* **42**, 63–71.
- Kim, J., Lee, J.H., and Iyer, V.R. (2008). Global identification of Myc target genes reveals its direct role in mitochondrial biogenesis and its E-box usage in vivo. *PLoS One* **3**.
- Kimura, Y., Nagao, A., Fujioka, Y., Satou, A., Taira, T., Iguchi-Arigo, S.M.M., and Ariga, H. (2007). MM-1 facilitates degradation of c-Myc by recruiting proteasome and a novel ubiquitin E3 ligase. *Int. J. Oncol.* **31**, 829–836.
- King, I.F., Yandava, C.N., Mabb, A.M., Hsiao, J.S., Huang, H.-S., Pearson, B.L., Calabrese, J.M., Starmer, J., Parker, J.S., Magnuson, T., et al. (2013). Topoisomerases facilitate transcription of long genes linked to autism. *Nature* **501**, 58–62.
- Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**, 2452–2460.
- Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. *Science* **184**, 865–868.

- Kouzine, F., Gupta, A., Baranello, L., Wojtowicz, D., Ben-Aissa, K., Liu, J., Przytycka, T.M., and Levens, D. (2013). Transcription-dependent dynamic supercoiling is a short-range genomic force. *Nat. Struct. Mol. Biol.* *20*, 396–403.
- Kulaeva, O.I., Hsieh, F.-K., Chang, H.-W., Luse, D.S., and Studitsky, V.M. (2013). Mechanism of transcription through a nucleosome by RNA polymerase II. *Biochim. Biophys. Acta - Gene Regul. Mech.* *1829*, 76–83.
- Lanahan, A., Williams, J.B., Sanders, L.K., and Nathans, D. (1992). Growth Factor-Induced Delayed Early Response Genes.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* *128*, 707–719.
- Liang, K., Gao, X., Gilmore, J.M., Florens, L., Washburn, M.P., Smith, E., and Shilatifard, A. (2015). Characterization of human cyclin-dependent kinase 12 (CDK12) and CDK13 complexes in C-terminal domain phosphorylation, gene transcription, and RNA processing. *Mol. Cell. Biol.* *35*, 928–938.
- Lin, C., Smith, E.R., Takahashi, H., Lai, K.C., Martin-Brown, S., Florens, L., Washburn, M.P., Conaway, J.W., Conaway, R.C., and Shilatifard, A. (2010). AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell* *37*, 429–437.
- Lin, C., Garrett, A.S., De Kumar, B., Smith, E.R., Gogol, M., Seidel, C., Krumlauf, R., and Shilatifard, A. (2011). Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC). *Genes Dev.* *25*, 1486–1498.
- Litingtung, Y., Lawler, A.M., Sebald, S.M., Lee, E., Gearhart, J.D., Westphal, H., and Corden, J.L. (1999). Growth retardation and neonatal lethality in mice with a homozygous deletion in the C-terminal domain of RNA polymerase II. *Mol. Gen. Genet.* *261*, 100–105.
- Locascio, A., Blázquez, M.A., and Alabadí, D. (2013). Dynamic regulation of cortical microtubule organization through prefoldin-DELLA interaction. *Curr. Biol.* *23*, 804–809.
- Lopez-Bergami, P., Kim, H., Dewing, A., Goydos, J., Aaronson, S., and Ronai, Z. (2010). c-Jun Regulates Phosphoinositide-dependent Kinase 1 Transcription. *J. Biol. Chem.* *285*, 903–913.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* *389*, 251–260.
- Luo, H.R., Moreau, G.A., Levin, N., and Moore, M.J. (1999). The human Prp8 protein is a component of both U2- and U12-dependent spliceosomes. *RNA* *5*, 893–908.
- MacRae, A.J., Mayerle, M., Hrabeta-Robinson, E., Chalkley, R.J., Guthrie, C., Burlingame, A.L., and Jurica, M.S. (2018). Prp8 positioning of U5 snRNA is linked to 5' splice site recognition. *RNA* *24*, 769–777.
- Malinová, A., Cvačková, Z., Matějů, D., Hořejší, Z., Abéza, C., Vandermoere, F., Bertrand,

- E., Staněk, D., and Verheggen, C. (2017). Assembly of the U5 snRNP component PRPF8 is controlled by the HSP90/R2TP chaperones. *J. Cell Biol.* *216*, 1579–1596.
- Marshall, N.F., and Price, D.H. (1992). Control of formation of two distinct classes of RNA polymerase II elongation complexes. *Mol. Cell. Biol.* *12*, 2078–2090.
- McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* *385*, 357–361.
- Meas, R., and Mao, P. (2015). Histone ubiquitylation and its roles in transcription and DNA damage response. *DNA Repair (Amst)*. *36*, 36–42.
- Meininghaus, M., Chapman, R.D., Horndasch, M., and Eick, D. (2000). Conditional expression of RNA polymerase II in mammalian cells. Deletion of the carboxyl-terminal domain of the large subunit affects early steps in transcription. *J. Biol. Chem.* *275*, 24375–24382.
- Miano, J.M. (2003). Serum response factor: toggling between disparate programs of gene expression. *J. Mol. Cell. Cardiol.* *35*, 577–593.
- Millán-Zambrano, G., and Chávez, S. (2014). Nuclear functions of prefoldin. *Open Biol.* *4*, 630–642.
- Millán-Zambrano, G., Rodríguez-Gil, A., Peñate, X., de Miguel-Jiménez, L., Morillo-Huesca, M., Krogan, N., and Chávez, S. (2013). The prefoldin complex regulates chromatin dynamics during transcription elongation. *PLoS Genet.* *9*, e1003776.
- Mirón-García, M.C., Garrido-Godino, A.I., Martínez-Fernández, V., Fernández-Pevida, A., Cuevas-Bermúdez, A., Martín-Expósito, M., Chávez, S., de la Cruz, J., and Navarro, F. (2014). The yeast prefoldin-like URI-orthologue Bud27 associates with the RSC nucleosome remodeler and modulates transcription. *Nucleic Acids Res.* *42*, 9666–9676.
- Misteli, T., and Spector, D.L. (1997). Applications of the green fluorescent protein in cell biology and biotechnology. *Nat. Biotechnol.* *15*, 961–964.
- Miyazawa, M., Tashiro, E., Kitaura, H., Maita, H., Suto, H., Iguchi-Ariga, S.M.M., and Ariga, H. (2011). Prefoldin subunits are protected from ubiquitin-proteasome system-mediated degradation by forming complex with other constituent subunits. *J. Biol. Chem.* *286*, 19191–19203.
- Montes, M., Becerra, S., Sánchez-Álvarez, M., and Suñé, C. (2012). Functional coupling of transcription and splicing. *Gene* *501*, 104–117.
- Mori, K., Maeda, Y., Kitaura, H., Taira, T., Iguchi-Ariga, S.M.M., and Ariga, H. (1998). MM-1, a novel c-Myc-associating protein that represses transcriptional activity of c-Myc. *J. Biol. Chem.* *273*, 29794–29800.
- Morrish, F., and Hockenbery, D. (2014). MYC and mitochondrial biogenesis. *Cold Spring Harb. Perspect. Med.* *4*, 24789872.
- Mousnier, A., Kubat, N., Massias-Simon, A., Ségéral, E., Rain, J.-C., Benarous, R., Emiliani, S., and Dargemont, C. (2007). von Hippel Lindau binding protein 1-mediated

- degradation of integrase affects HIV-1 gene expression at a postintegration step. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 13615–13620.
- Muñoz, M.J., de la Mata, M., and Kornblihtt, A.R. (2010). The carboxy terminal domain of RNA polymerase II and alternative splicing. *Trends Biochem. Sci.* *35*, 497–504.
- Narita, R., Kitaura, H., Torii, A., Tashiro, E., Miyazawa, M., Ariga, H., and Iguchi-Ariga, S.M.M. (2012). Rabring7 degrades c-Myc through complex formation with MM-1. *PLoS One* *7*.
- Nguyen, T.H.D., Li, J., Galej, W.P., Oshikane, H., Newman, A.J., and Nagai, K. (2013). Structural basis of Brr2-Prp8 interactions and implications for U5 snRNP biogenesis and the spliceosome active site. *Structure* *21*, 910–919.
- Ni, Z., Saunders, A., Fuda, N.J., Yao, J., Suarez, J.-R., Webb, W.W., and Lis, J.T. (2008). P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. *Mol. Cell. Biol.* *28*, 1161–1170.
- Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., et al. (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* *151*, 68–79.
- Nonet, M., Sweetser, D., and Young, R.A. (1987). Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* *50*, 909–915.
- Okochi, M., Nomura, T., Zako, T., Arakawa, T., Iizuka, R., Ueda, H., Funatsu, T., Leroux, M., and Yohda, M. (2004). Kinetics and binding sites for interaction of the prefoldin with a group II chaperonin: contiguous non-native substrate and chaperonin binding sites in the archaeal prefoldin. *J. Biol. Chem.* *279*, 31788–31795.
- Pajares, M.Á. (2017). PDRG1 at the interface between intermediary metabolism and oncogenesis. *World J. Biol. Chem.* *8*, 175–186.
- Palumbo, V., Pellacani, C., Heesom, K.J., Rogala, K.B., Deane, C.M., Mottier-Pavie, V., Gatti, M., Bonaccorsi, S., and Wakefield, J.G. (2015). Misato controls mitotic microtubule generation by stabilizing the TCP-1 tubulin chaperone complex. *Curr. Biol.* *25*, 1777–1783.
- Panning, B., and Taatjes, D.J. (2008). Transcriptional regulation: it takes a village. *Mol. Cell* *31*, 622–629.
- 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*, 4689–4694.
- Payán-Bravo, L., Peñate, X., and Chávez, S. (2018). Functional contributions of Prefoldin to gene expression. In *Advances in Experimental Medicine and Biology*, pp. 1–10.
- Perales, R., and Bentley, D. (2009). “Cotranscriptionality”: the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell* *36*, 178–191.
- Pérez, C., Pérez-Zúñiga, F.J., Garrido, F., Reytor, E., Portillo, F., and Pajares, M.A. (2016). The oncogene PDRG1 is an interaction target of methionine adenosyltransferases. *PLoS One* *11*, e0161672.

- Pertea, M., Kim, D., Pertea, G., Leek, J.T., and Steven, L. (2017). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* *11*, 1650–1667.
- Phatnani, H.P., and Greenleaf, A.L. (2008). Phosphorylation and functions of the RNA polymerase II CTD. *Cold Spring Harb. Lab. Press March* *10*.
- Porrua, O., and Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.* *16*, 190–202.
- Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* *111*, 229–233.
- Price, D.H. (2000). P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell. Biol.* *20*, 2629–2634.
- Price, D.H. (2008). Poised polymerases: on your mark...get set...go! *Mol. Cell* *30*, 7–10.
- Pugh, B.F. (2000). Control of gene expression through regulation of the TATA-binding protein. *Gene* *255*, 1–14.
- Reed, R. (2003). Coupling transcription, splicing and mRNA export. *Curr. Opin. Cell Biol.* *15*, 326–331.
- Regulapati, R., Bhasi, A., Singh, C.K., and Senapathy, P. (2008). Origination of the split structure of spliceosomal genes from random genetic sequences. *PLoS One* *3*, e3456.
- Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. *Genes Dev.* *23*, 1247–1269.
- Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D., and Klug, A. (1984). Structure of the nucleosome core particle at 7 Å resolution. *Nature* *311*, 532–537.
- Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* *21*, 327–335.
- Roeder, R.G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett.* *579*, 909–915.
- Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* *54*, 795–804.
- Sainsbury, S., Bernecky, C., and Cramer, P. (2015). Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* *16*, 129–143.
- Saint-André, V., Batsché, E., Rachez, C., and Muchardt, C. (2011). Histone H3 lysine 9 trimethylation and HP1 γ favor inclusion of alternative exons. *Nat. Struct. Mol. Biol.* *18*, 337–344.
- Saldi, T., Cortazar, M.A., Sheridan, R.M., and Bentley, D.L. (2016). Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *J. Mol. Biol.* *428*, 2623–2635.

- Saramaki, A. (2006). Regulation of the human p21(waf1/cip1) gene promoter via multiple binding sites for p53 and the vitamin D3 receptor. *Nucleic Acids Res.* *34*, 543–554.
- Satou, A., Taira, T., Iguchi-Arigo, S.M.M., and Ariga, H. (2001). A novel transrepression pathway of c-Myc: Recruitment of a transcriptional corepressor complex to c-Myc by MM-1, a c-Myc-binding protein. *J. Biol. Chem.* *276*, 46562–46567.
- Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* *7*, 557–567.
- Sheridan, R.M., Fong, N., D’Alessandro, A., and Bentley, D.L. (2019). Widespread backtracking by RNA Pol II is a major effector of gene activation, 5’ pause release, termination, and transcription elongation rate. *Mol. Cell* *73*, 107–118.e4.
- Siegers, K., Waldmann, T., Leroux, M.R., Grein, K., Shevchenko, A., Schiebel, E., and Hartl, F.U. (1999). Compartmentation of protein folding in vivo: sequestration of non-native polypeptide by the chaperonin-GimC system. *EMBO J.* *18*, 75–84.
- Siegert, R., Leroux, M.R., Scheufler, C., Hartl, F.U., and Moarefi, I. (2000). Structure of the Molecular Chaperone Prefoldin. *Cell* *103*, 621–632.
- Simons, C.T., Staes, A., Rommelaere, H., Ampe, C., Lewis, S.A., and Cowan, N.J. (2004). Selective contribution of eukaryotic prefoldin subunits to actin and tubulin binding. *J. Biol. Chem.* *279*, 4196–4203.
- Sims, R.J., Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* *18*, 2437–2468.
- Singh, J., and Padgett, R.A. (2009). Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.* *16*, 1128–1133.
- Singh, J., and Padgett, R.A. (2010). Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.* *16*, 1128–1133.
- Singh, V., Braddick, D., and Dhar, P.K. (2017). Exploring the potential of genome editing CRISPR-Cas9 technology. *Gene* *599*, 1–18.
- Søgaard, T.M.M., and Svejstrup, J.Q. (2007). Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J. Biol. Chem.* *282*, 14113–14120.
- Subtil-Rodríguez, A., and Reyes, J.C. (2010). BRG1 helps RNA polymerase II to overcome a nucleosomal barrier during elongation, in vivo. *EMBO Rep.* *11*, 751–757.
- Sui, X., Fang, Y., Lou, H., Wang, K., Zheng, Y., Lou, F., Jin, W., Xu, Y., Chen, W., Pan, H., et al. (2015). p53 suppresses stress-induced cellular senescence via regulation of autophagy under the deprivation of serum. *Mol. Med. Rep.* *11*, 1214–1220.
- Svetlov, V., and Nudler, E. (2013). Basic mechanism of transcription by RNA polymerase II. *Biochim. Biophys. Acta* *1829*, 20–28.
- Taatjes, D.J., Marr, M.T., and Tjian, R. (2004). Regulatory diversity among metazoan co-activator complexes. *Nat. Rev. Mol. Cell Biol.* *5*, 403–410.

- Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z.-Y., Larsen, B., Choi, H., Berger, B., Gingras, A.-C., and Lindquist, S. (2014). A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell* July 17, 434–44805.
- Tharun, S. (2009). Roles of eukaryotic Lsm proteins in the regulation of mRNA function. *Int. Rev. Cell Mol. Biol.* 272, 149–189.
- Tjeertes, J. V., Miller, K.M., and Jackson, S.P. (2009). Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO J.* 28, 1878–1889.
- Turunen, J.J., Niemelä, E.H., Verma, B., and Frilander, M.J. (2013). The significant other: splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* 4, 61–76.
- Udhane, S.S., Pandey, A. V, Hofer, G., Mullis, P.E., and Flück, C.E. (2015). Retinoic acid receptor beta and angiopoietin-like protein 1 are involved in the regulation of human androgen biosynthesis. *Sci. Rep.* 5, 10132.
- Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H.L., and Cowan, N.J. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell* 93, 863–873.
- Vega, F.M., and Ridley, A.J. (2016). The RhoB small GTPase in physiology and disease. *Small GTPases* 9, 384–393.
- Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701–718.
- Wang, D., Shi, W., Tang, Y., Liu, Y., He, K., Hu, Y., Li, J., Yang, Y., and Song, J. (2017). Prefoldin 1 promotes EMT and lung cancer progression by suppressing cyclin A expression. *Oncogene* 36, 885–898.
- Wang, P., Zhao, J., Yang, X., Guan, S., Feng, H., Han, D., Lu, J., Ou, B., Jin, R., Sun, J., et al. (2015). PFDN1, an indicator for colorectal cancer prognosis, enhances tumor cell proliferation and motility through cytoskeletal reorganization. *Med. Oncol.* 32, 264.
- Whitehead, T.A., Boonyaratanakornkit, B.B., Höllrigl, V., and Clark, D.S. (2007). A filamentous molecular chaperone of the prefoldin family from the deep-sea hyperthermophile *Methanocaldococcus jannaschii*. *Protein Sci.* 16, 626–634.
- Wu, G., Liu, J., Wu, Z., Wu, X., and Yao, X. (2017). MicroRNA-184 inhibits cell proliferation and metastasis in human colorectal cancer by directly targeting IGF-1R. *Oncol. Lett.* 14, 3215–3222.
- Yamaguchi, Y., Shibata, H., and Handa, H. (2013). Transcription elongation factors DSIF and NELF: Promoter-proximal pausing and beyond. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1829, 98–104.
- 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, 978–990.

- Yonaha, M., and Proudfoot, N.J. (2000). Transcriptional termination and coupled polyadenylation in vitro. *EMBO J.* *19*, 3770–3777.
- Yoshida, T., Kitaura, H., Hagio, Y., Sato, T., Iguchi-Ariga, S.M.M., and Ariga, H. (2008). Negative regulation of the Wnt signal by MM-1 through inhibiting expression of the *wnt4* gene. *Exp. Cell Res.* *314*, 1217–1228.
- Yurko, N.M., and Manley, J.L. (2018). The RNA polymerase II CTD “orphan” residues: Emerging insights into the functions of Tyr-1, Thr-4, and Ser-7. *Transcription* *9*, 30–40.
- 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*, 3698–3702.
- Zhao, X.-H., Wang, Y.-B., Qiu, W., Zheng, J.-H., and Li, G. (2017). MicroRNA-184 inhibits proliferation and promotes apoptosis of human colon cancer SW480 and HCT116 cells by downregulating C-MYC and BCL-2. *J. Cell. Biochem.* *119*, 1702–1715.
- Zhou, Z., Liu, Y.T., Ma, L., Gong, T., Hu, Y.N., Li, H.T., Cai, C., Zhang, L.L., Wei, G., and Zhou, J.Q. (2017). Independent manipulation of histone H3 modifications in individual nucleosomes reveals the contributions of sister histones to transcription. *Elife* *6*, 1–24.