



---

---

TESIS DE MAESTRÍA EN CIENCIAS BIOLÓGICAS  
PEDECIBA BIOLOGÍA  
SUB ÁREA BIOLOGÍA CELULAR Y MOLECULAR

# **CARACTERIZACIÓN ESTRUCTURAL DE LA PROTEÍNA CCDC28B, UN MODIFICADOR DEL SÍNDROME DE BARDET-BIEDL**

Lic. Matías Fabregat  
Orientador: Dra. Florencia Irigoín  
Co-orientadores: Dr. José Luis Badano y Dr. Alejandro Buschiazzo

**Laboratorio de Genética Molecular Humana - Institut Pasteur de Montevideo**

---

---

**MARZO 2017**

<b>1. RESUMEN</b> .....	5
<b>2. INTRODUCCIÓN</b>	
1.1. La cilia primaria.....	7
1.2. Las ciliopatías.....	11
1.3. El Síndrome de Bardet-Biedl.....	12
1.4. El modificador secundario CCDC28B.....	14
1.5. CCDC28B y su función.....	15
1.6. Aproximación a CCDC28B desde predicciones informáticas.....	18
1.7. De la estructura a la función.....	20
1.8. Determinación de la estructura tridimensional de una proteína.....	22
<b>3. OBJETIVO GENERAL Y ESPECÍFICOS</b> .....	25
<b>4. MATERIALES Y MÉTODOS</b>	
3.1. Clonado en vectores de expresión	
3.1.A. Clonado por enzimas de restricción.....	26
3.1.B. Clonado por amplificación del plásmido entero.....	27
3.2. Producción de proteínas recombinantes.	
3.2.A. Plásmidos utilizados.....	29
3.2.B. Cepas bacterianas utilizadas.....	30
3.2.C. Medios de cultivo bacteriano.....	31
3.2.D. Transformación de bacterias.....	31
3.2.E. Inducción de la expresión y producción de proteínas recombinantes.....	32
3.3. Purificación de proteínas recombinantes.	
3.3.A. Sistemas de purificación de proteínas.....	35
3.3.B. Cromatografía de afinidad a metales inmovilizados.....	35
3.3.C. Cromatografía de intercambio aniónico.....	38
3.3.D. Cromatografía de exclusión molecular.....	40
3.3.E. Cuantificación de proteínas.....	41
3.3.F. SDS-PAGE.....	42
3.3.G. Western Blot.....	45
3.4. Ensayos de cristalización	
3.4.A. Pruebas de solubilidad.....	48
3.4.B. Búsqueda de condiciones de cristalización.....	50
3.4.C. Optimización de las condiciones de cristalización.....	51
3.5. Ensayos de cultivo celular	
3.5.A. Líneas celulares utilizadas y condiciones de cultivo.....	52
3.5.B. Sobre-expresión de genes en células humanas.....	52
3.5.C. Co-inmunoprecipitación.....	53
3.5.D. Inmunofluorescencia.....	54
3.5.E. Microscopía confocal y cuantificación de imágenes.....	56

3.5.F. Diseño y obtención de una línea celular CCDC28B <i>knockdown</i> por la tecnología de CRISPR/Cas9.....	56
3.5.G. Obtención y caracterización de clones CCDC28B <i>kd</i> .....	59
<b>4. RESULTADOS</b>	
4.1. Producción de CCDC28B <i>wild type</i> recombinante	
4.1.A. Optimización de condiciones de inducción.....	61
4.2. Purificación y caracterización de CCDC28B <i>wild type</i>	
4.2.A Optimización de la cromatografía de afinidad por metales (IMAC).....	64
4.2.B Cromatografía de intercambio aniónico.....	66
4.2.C. Cromatografía de exclusión molecular de las distintas fracciones del intercambio aniónico.....	68
4.2.D. SEC del material purificado por IMAC.....	70
4.3. Caracterización de CCDC28B para la búsqueda de condiciones de cristalización	
4.3.A. Análisis por Dispersión de la Luz Dinámica.....	71
4.3.B. Estabilidad al congelado y descongelado a -80°C y la concentración.....	71
4.3.C. Pruebas de concentración en distintos buffer del material purificado por IMAC.....	73
4.3.D. Estudio de CCDC28B en condiciones no reductoras.....	75
4.4. Prueba de solubilidad y búsqueda de condiciones de cristalización.....	77
4.5. Estrategias para mejorar la cantidad/calidad de CCDC28B recombinante	
4.5.A. Expresión de CCDC28B fusionada a otras proteínas.....	80
4.4.B. Coexpresión de CCDC28B con interactores.....	81
4.4.C. Corte de la etiqueta 6-His de CCDC28B.....	83
4.6. Estudio del estado oligomérico de CCDC28B en células eucariotas.....	88
4.7. Identificación de dominios o regiones de CCDC28B que sean importantes para su función	
4.7.A. Estudio de funcionalidad de los mutantes de CCDC28B a través de su rol en la función de mTORC2.....	91
4.7.B. Diseño de una línea celular CCDC28B- <i>knockdown</i> por tecnología de CRISPR/Cas9 .....	93
4.7.C. Evaluación de la funcionalidad de diferentes mutantes de CCDC28B.....	96
<b>5. DISCUSIÓN</b> .....	99
<b>6. AGRADECIMIENTOS</b> .....	111
<b>7. BIBLIOGRAFÍA</b> .....	113
<b>8. ANEXO</b> .....	122

*“Almost all aspects of life are engineered at the molecular level, and without understanding molecules we can only have a sketchy understanding of life itself.”*

Francis Crick (1988)

## RESUMEN

---

La cilia primaria es un organelo presente en la mayoría de los tipos celulares en mamíferos, que se proyecta desde la superficie celular a modo de antena, participando en procesos de recepción y transducción de estímulos mecánicos y químicos. Para esto, participa de diferentes vías de señalización esenciales para el desarrollo y mantenimiento de los tejidos y órganos, como es la vía de Sonic-Hedgehog, Wnt, y PDGF $\alpha$ , entre otras. Esta importancia funcional la convierten en un organelo esencial, cuyo mal funcionamiento está asociado a un importante número de patologías complejas denominadas ciliopatías. Una ciliopatía modelo, y objeto de estudio del Laboratorio de Biología Molecular Humana del Instituto Pasteur de Montevideo, es el Síndrome de Bardet-Biedl (BBS). BBS es una enfermedad genéticamente heterogénea para la cual a la fecha se han identificado 21 genes causales. Si bien puede ser heredado con un patrón compatible con el de una enfermedad autosómica recesiva, en muchas familias se comporta como un rasgo oligogénico. De allí surge la importancia de modificadores secundarios de la enfermedad, como lo es CCDC28B (del inglés: *coiled coil domain containing protein 28B*). No hay evidencias de que mutaciones en este gen sean suficientes para causar el síndrome pero sí para modular su expresividad modificando la severidad de su manifestación en un contexto genéticamente sensible por la presencia de otras mutaciones en genes BBS. CCDC28B es una proteína sin homólogos de función conocida ni dominios funcionales identificables. Trabajos previos del grupo sugieren que CCDC28B funcionaría a través de modificar la actividad de proteínas con las cuales interactúa. Así, se ha mostrado que a través de la interacción con SIN1 juega un rol importante en el proceso de ciliogénesis, y en la modulación del complejo mTORC2, efectos observados tanto en células en cultivo como *in vivo* en el pez cebra. Pese a esto, no se sabe cómo CCDC28B interacciona con otras proteínas modificando su actividad y no se dispone de ninguna información estructural más allá de algunas predicciones *in silico*.

De esta forma, el objetivo central de este trabajo fue caracterizar desde el punto de vista estructural a CCDC28B y establecer relaciones entre su estructura y su función que nos permitan entender la forma de actuar de CCDC28B. Para alcanzar este objetivo se utilizaron dos enfoques diferentes. Por un lado se produjo y purificó la proteína recombinante en *E.coli* para intentar obtener cristales y realizar ensayos de cristalografía de rayos X buscando resolver la estructura tridimensional de la proteína. Si bien no se logró la cristalización de la proteína recombinante, se obtuvo información sobre la proteína en sí y sobre el proceso de su producción, que nos será de gran utilidad para continuar el trabajo a futuro. Sabemos que la proteína puede formar dímeros mediados por puentes disulfuro, algo que no ocurre *in vivo*, e indica la necesidad de utilizar reductores durante la producción y purificación de la misma. Se logró obtener una forma de CCDC28B a la que le faltan los primeros 16 aminoácidos que permite eliminar la cola de 6-Histidinas utilizada para el primer paso de purificación y carece de una serie de aminoácidos cargados que podrían dificultar la formación de cristales. La segunda aproximación a nuestro objetivo consistió en determinar la importancia de ciertas regiones/dominios en la actividad de

CCDC28B. Para ello se generó, por la tecnología de CRISPR/Cas9, una línea estable de células hTERT-RPE donde la actividad de CCDC28B está disminuida, lo cual se refleja en una clara disminución del porcentaje de células ciliadas. Utilizando este parámetro como reportero se analizó la capacidad de mutantes en regiones involucradas en determinar la localización subcelular de CCDC28B, de rescatar el fenotipo de ciliación. Se observó que, mientras la expresión de la proteína salvaje recupera el nivel de ciliación, la proteína mutante en la señal de localización nuclear, que no es capaz de entrar al núcleo, no lo hace. Esto sugiere la importancia de la localización nuclear de CCDC28B para su actividad en ciliogénesis. Este modelo nos permitirá seguir testeando otros mutantes, en regiones/residuos predichos *in silico* como potencialmente funcionales, en la actividad de CCDC28B en ciliogénesis.

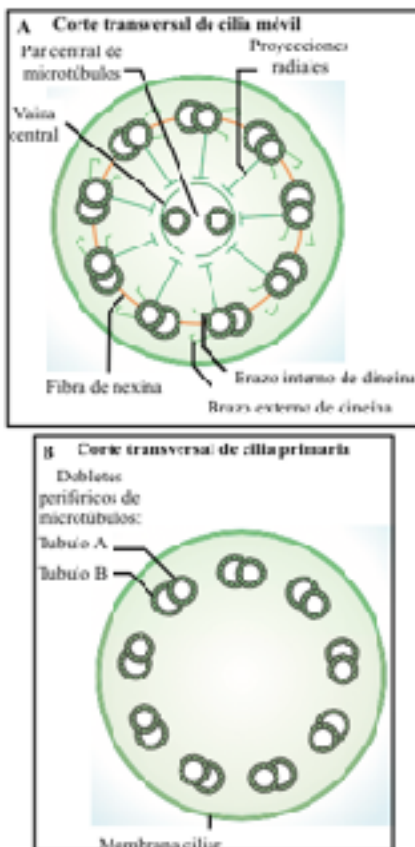
# INTRODUCCIÓN

## 1.1. La cilia primaria

La cilia primaria es un organelo celular conservado durante la evolución. Mientras que las plantas superiores y hongos, no poseen cilia, en invertebrados están restringidas a las neuronas sensoriales y en mamíferos están presentes en casi todos los tipos celulares del organismo ([www.bowserlab.org/primarycilia/cilialist](http://www.bowserlab.org/primarycilia/cilialist) y Hoyer-Fender, S., 2013). Se presenta como una estructura en forma de antena, generada por la proyección de microtúbulos desde la superficie celular, particularmente desde el cuerpo basal -centro organizador de microtúbulos derivado del centriolo-, hacia el medio extra-celular. El esqueleto de microtúbulos, denominado axonema, puede diferenciarse según su ultraestructura caracterizando dos grandes grupos distintos de cilias: cilias primarias o inmóviles, y cilias móviles (ver Figura 1).

Las cilias móviles presentan un axonema formado por nueve dobletes de microtúbulos periféricos (llamados A y B) y un par de microtúbulos centrales, este tipo de configuración se denomina por lo tanto “9+2” (Figura 1A). Cada doblete periférico está formado por un microtúbulo completo (A), formado por 13 protofilamentos y un microtúbulo incompleto (B) formado solamente por 10 protofilamentos. Al igual que los microtúbulos celulares, los que forman el axonema ciliar se forman gracias a la polimerización de heterodímeros de  $\alpha$  y  $\beta$  tubulina, con su extremo “más” (hacia el cual crecen) hacia el extremo distal de la cilia (Fisch, C. and Dupuis-Williams, P.,

2011). Además, el axonema presenta algunas estructuras accesorias, como los brazos externos e internos de dineína que colaboran al movimiento en este tipo de cilias. Normalmente se encuentran de a varias por célula en epitelios especializados, como en el tracto respiratorio o el aparato reproductor, y su función principal es la de generar movimiento para mover fluido extra-celular (Mitchison, H. M. and Valente, E. M., 2017). Por otro lado, las cilias primarias se encuentran de a una sola por célula y no son móviles. La configuración de su axonema es de “9+0”, ya que carece del par de microtúbulos centrales (Figura 1B). Tampoco posee muchas de las estructuras accesorias, como los brazos de dineína, y su función es principalmente mecánico y quimio-sensorial (Cardenas-Rodríguez, M. and Badano, J. L., 2009; Goetz, S. C. and Anderson, K. V., 2010). Pese a la caracterización simplista en estos dos grandes grupos, existen evidencias de la presencia de

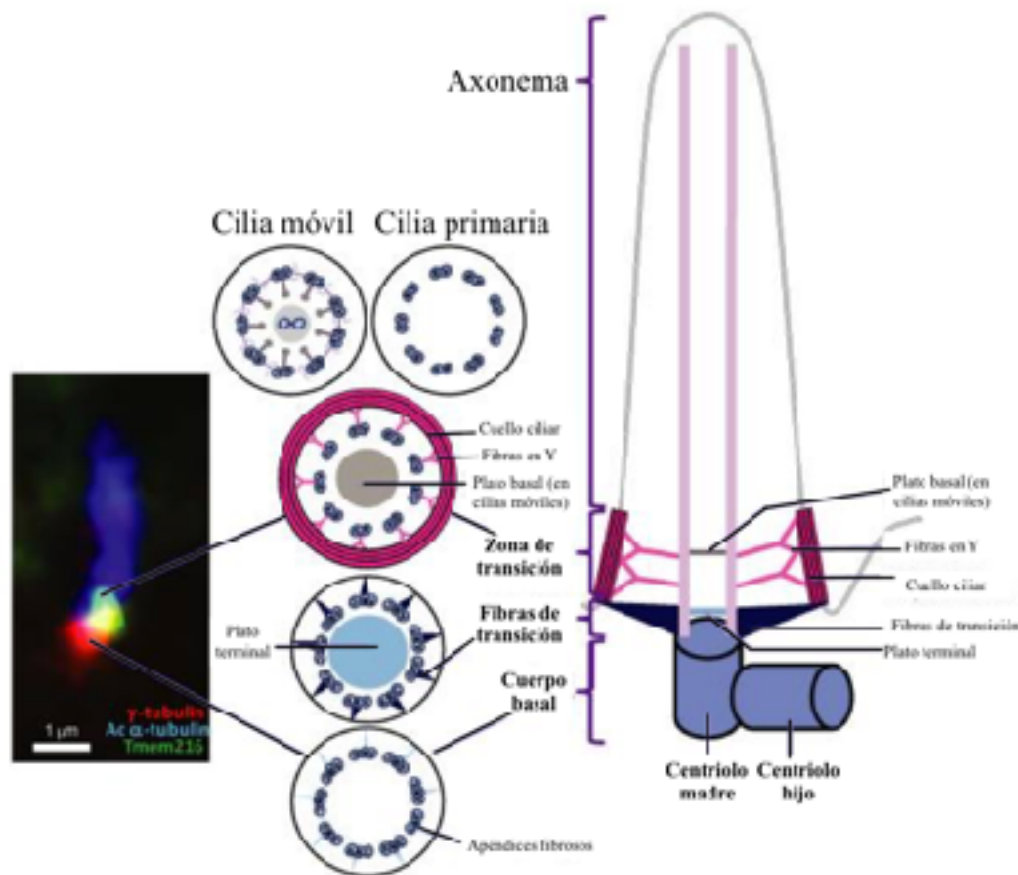


**Figura 1. Esquema mostrando un corte transversal de los dos tipos de cilias: móviles (A) y primarias (B).**

Figura adaptada de: Ishikawa, H. and Marshall, W. F., 2011.

cilias inmóviles con un axonema “9+2”, por ejemplo en el tracto olfatorio en humanos, y cilias móviles de axonema “9+0”, como en el nodo embrionario del ratón (Hoyer-Fender, S., 2013; Moran, D.T. et al, 1982; Sulik, K., et al, 1994).

La Figura 2 muestra una representación de las distintas regiones que encontramos en la cilia primaria. Allí se ve que el axonema nace desde el extremo distal del cuerpo basal, el cual es formado por el centriolo madre, es decir, el centriolo más antiguo. El cuerpo basal y los centriolos tienen la misma estructura en barril formada por nueve tripletes de microtúbulos, y es que, pese a que cumplen funciones diferentes, son el mismo organelo, que se transforma uno en otro en diferentes etapas del ciclo celular, asociando de esta manera el proceso de ciliogénesis al propio ciclo celular. Durante la fase G0 y G1, el centriolo se mueve hacia la superficie celular para actuar como cuerpo basal y promover la ciliogénesis. Mientras que durante la mitosis la cilia se reabsorbe y el cuerpo basal vuelve a actuar como centriolo formando el centrosoma (Irigoín, F. and Badano, J. L., 2011; Marshall, W. F., 2008). En el extremo distal del cuerpo basal se encuentran los apéndices fibrosos que al hacer contacto con la membrana celular pasan a formar las fibras de transición. Tanto los apéndices fibrosos, como las fibras de transición salen del túbulo central de los tripletes del cuerpo basal. Estas fibras no sólo hacen contacto con la membrana, sino que anclan el cuerpo



**Figura 2. Representación esquemática de la ultraestructura de la cilia primaria y cortes longitudinales de las secciones que la componen.**

A la izquierda se muestra una micrografía de una inmunofluorescencia de una cilia primaria. Con azul se marca la tubulina acetilada para ver el axonema, con rojo se marca la  $\gamma$ -tubulina para ver el cuerpo basal y con verde la proteína transmembrana (TMEM) 216, la cual se localiza en la zona de transición.

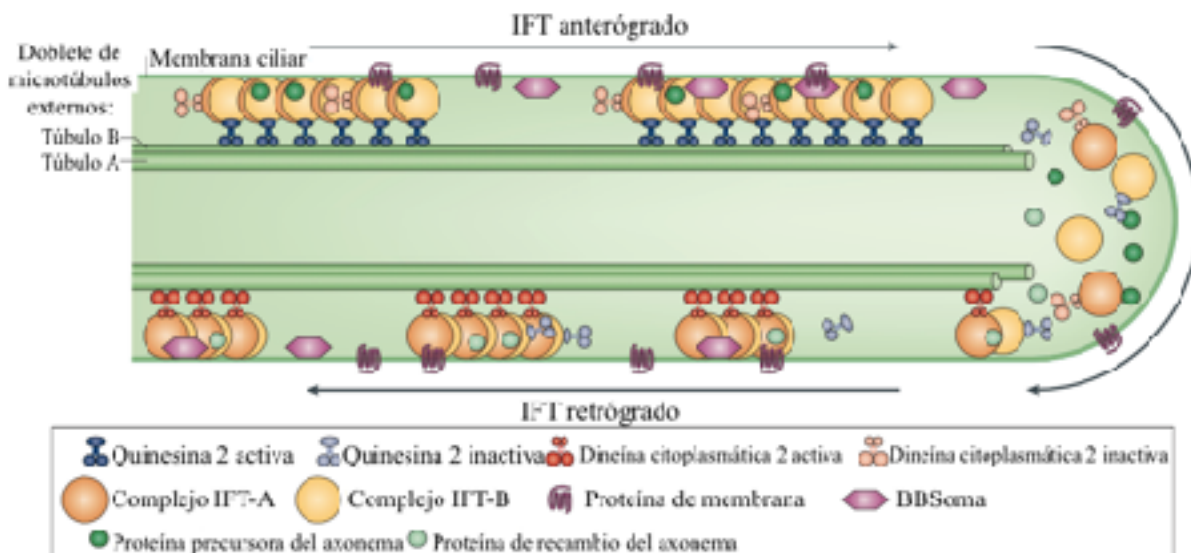
Figura adaptada de: Szymanska, K. and Johnson, C. A., 2012.



basal a la misma (Anderson, R. G. W., 1972). De hecho, Rosenbaum y Witman (2002) proponen que estas fibras de transición son parte de un complejo que funcionaría de manera similar al poro nuclear, regulando el tráfico de proteínas hacia y desde la cilia.

En la base de la cilia, por encima del cuerpo basal, existe una región llamada zona de transición (ZT), que separa el organelo del resto del citosol y de la membrana plasmática. Esta región actúa como una barrera que permite seleccionar los componentes solubles y de membrana del compartimiento intra-ciliar y la membrana ciliar, permitiendo que la cilia posea un conjunto característico de proteínas (Fisch, C. and Dupuis-Williams, P., 2011; Reiter, J. F. et al, 2012). En esta región se encuentran las fibras en Y, que unen el axonema con la membrana ciliar, y los platos terminal (sin función demostrada al momento) y basal (nuclea el par central de microtúbulos sólo en cilias móviles). Todas estas estructuras fueron identificadas por microscopía electrónica, pero aún resta la identificación exacta de todos sus componentes y funciones (Fisch, C. and Dupuis-Williams, P., 2011). Es interesante notar que muchas de las proteínas que se localizan en la ZT se han encontrado mutadas en diversas enfermedades (ciliopatías, se describirán más adelante), poniendo en evidencia la relevancia fisiológica de esta región (Czarnecky PG, Trends in Cell Biology, 2012, 22:201; Verhey, KJ, Current Opinion in Cell Biology, 2016, 41:109). La información existente hasta el momento indica que CEP164 (del inglés: *centrosomal protein of 164 kDa*) y ODF2 (del inglés: *outer dense fibre 2*) están unidas a las fibras de transición y ambas son necesarias para el anclaje a la membrana celular (Graser, S. et al, 2007; Ishikawa, H. et al, 2005). La membrana ciliar en la ZT se denomina cuello ciliar. Al momento, es desconocida la identidad de muchos componentes de este cuello ciliar pero se observó que no necesariamente son los mismos en diferentes organismos. Se postula que la función principal de este cuello ciliar es la de actuar como barrera selectiva para la difusión de lípidos y proteínas entre la membrana plasmática y ciliar (Fisch, C. and Dupuis-Williams, P., 2011). La membrana ciliar corresponde a la bicapa lipídica que cubre al axonema, que pese a ser continua con la membrana plasmática, es un compartimiento especializado hacia donde las moléculas tienen acceso restringido. Este compartimiento está enriquecido particularmente en receptores, canales iónicos específicos y cierto tipo de fosfoinosítidos (Rohatgi, R. and Snell, W. J., 2010).

La cilia primaria carece de la maquinaria necesaria para la síntesis proteica y por ende, todas las proteínas necesarias para la formación, mantenimiento y función de la cilia deben ser sintetizadas previamente en el cuerpo celular y luego transportadas hacia el organelo. Para ello existe un mecanismo especializado que se denomina transporte intraflagelar (IFT, del inglés: *intraflagellar transport*). Se trata de un proceso bidireccional que utiliza motores para mover complejos proteicos, denominados partículas IFT (o trenes IFT), hacia el extremo distal y basal de la cilia (movimiento anterógrado y retrógrado, respectivamente). Todos los motores basados en microtúbulos se mueven a lo largo de estos en una sola dirección, hacia el extremo “más” o hacia el “menos”, por lo tanto se necesitan motores diferentes para el movimiento anterógrado y retrógrado. Esta función la cumplen las dos grandes familias de motores basados en microtúbulos: las



**Figura 3. Maquinaria de transporte intraflagelar (IFT).**

El motor de IFT anterógrado, quinesina-2, transporta los complejos de IFT A y B, las proteínas precursoras del axonema y la dineína citoplasmática 2 hasta la punta de la cilia. Durante este movimiento anterógrado, la quinesina-2 está activa y el motor retrógrado, dineína citoplasmática 2, se mantiene de alguna manera inactivo para permitir un movimiento anterógrado progresivo. En la punta de la cilia, los trenes IFT anterógrados liberan las proteínas precursoras del axonema y reordenan su conformación para generar el IFT retrógrado. La dineína citoplasmática 2 se activa y transporta los trenes IFT retrógrados al cuerpo celular. Todos los trenes IFT están involucrados además en el transporte de proteínas de membrana hacia la membrana ciliar y del BBSoma (complejo proteico compuesto por al menos siete proteínas asociadas al síndrome de Bardet-Biedl) hacia la punta de la cilia.

Figura adaptada de: Ishikawa, H., Marshall, W. F., 2011.

quinesinas y las dineínas (Hirokawa, N., 1998). El motor de movimiento anterógrado, quinesina 2, es un complejo heterotrimérico formado por dos subunidades motoras, KIF3A y KIF3B (en mamíferos), y una subunidad no motora, KAP (Scholey, J. M., 1996). Por otro lado el motor de movimiento retrógrado, dineína citoplasmática 2, es de estructura bastante más compleja y está formado por una cadena pesada, una cadena intermedia-liviana, una cadena intermedia y tres cadenas livianas (Hou, Y. and Witman, G. B., 2015). La Figura 3, muestra esquemáticamente cómo funciona el IFT, allí se observa que las proteínas producidas en el cuerpo celular, son llevadas hacia la punta de la cilia por la quinesina 2 y el complejo IFT B. Entre las proteínas transportadas en esta dirección se encuentra la propia dineína citoplasmática 2, que una vez en la punta de la cilia, se hace cargo del transporte retrógrado junto con el complejo IFT A, pasando la quinesina 2 inactiva a ser carga. El cambio de movimiento anterógrado a retrógrado no está totalmente caracterizado, aunque sí se sabe que está regulado por proteínas ancladas a las fibras de transición (en la base de la cilia) y a algunos filamentos que unen el final de los túbulos A del axonema con la membrana ciliar en la punta de la cilia (Rosenbaum, J. L. and Witman, G. B., 2002).

Como se mencionó anteriormente, la función principal de la cilia primaria consiste en actuar como “antena celular”, sensando el medio extracelular y generando las respuestas necesarias. La relación superficie/volumen en la cilia es mayor que en el resto de la célula, por lo cual una gran densidad de receptores señalizan hacia un compartimiento donde las proteínas que participan en el proceso de transducción están concentradas, favoreciendo la interacción entre ellas y por lo tanto respondiendo en forma eficiente y rápida. Este tipo de argumentos se ha utilizado para explicar por qué son varias las vías de señalización que actúan a través de la cilia, como por ejemplo: *Wingless*

(Wnt), *Sonic Hedgehog* (Shh) o Factor de crecimiento derivado de plaquetas alfa (PDGF $\alpha$ ) (Cardenas-Rodríguez, M. and Badano, J. L., 2009; Gerdes, J. M. and Katsanis, N., 2008). Su participación en estas vías, hacen que la cilia primaria sea un organelo importante para procesos celulares fundamentales como la proliferación, establecimiento de la polaridad durante el desarrollo y la diferenciación celular.

Dada la distribución ubicua de la cilia, su conservación evolutiva y su participación en los mecanismos antes mencionados, no debe resultar sorprendente que la cilia primaria sea un organelo esencial. Efectivamente, la ausencia de cilia en mamíferos es incompatible con la vida y defectos en la formación o funcionamiento de este organelo ha sido ampliamente asociado a una serie de patologías complejas con un importante número de fenotipos, que conjuntamente se han denominado ciliopatías.

## 1.2. Las ciliopatías

Las ciliopatías han sido definidas como un grupo de patologías genéticas humanas que comparten un defecto celular común: problemas en la formación, estructura o función de las cilias. Estas enfermedades presentan un amplio rango de fenotipos, siendo algunos de ellos patologías altamente prevalentes en la población general, como la diabetes, obesidad, retardo mental y poliquistosis renal entre otras (Bergmann, C., 2012). Pero también algunas ciliopatías presentan otras características menos comunes, como degeneración de retina, anosmia y polidactilia, entre otras. Algunos ejemplos de ciliopatías con los fenotipos que las caracterizan se muestran en la Tabla 1 (Cardenas-Rodríguez, M. and Badano, J. L., 2009).

En general, las ciliopatías son caracterizadas por una gran variabilidad fenotípica inter e intra familiar. Por un lado, mutaciones en genes diferentes, necesarios para el ensamblado y función de la cilia, pueden dar lugar a fenotipos muy similares. A su vez, mutaciones en un mismo gen pueden dar lugar a presentaciones muy distintas de una determinada ciliopatía (Davis, E. E. and Katsanis, N., 2012). Por lo tanto, sumado a la variabilidad inherente a la posibilidad de que distintos genes, con funciones quizás similares pero no idénticas, se asocien a un mismo defecto celular/fenotipo, la variabilidad presentada dentro de una misma familia, segregando las mismas mutaciones, evidencia una significativa complejidad genética. Esto ha quedado evidenciado en distintos trabajos mostrando una compleja relación entre todos los alelos existentes para los genes asociados a estas enfermedades. Es así que muchas de las ciliopatías hoy día son consideradas enfermedades oligogénicas donde mutaciones/alelos en más de un gen interaccionan para determinar el fenotipo final. Por lo tanto, las ciliopatías son patologías donde generalmente se identifica un gen principal para el desarrollo de la patología, pero donde otros genes pueden actuar como moduladores de la presentación de la misma, llamados en algunos casos modificadores secundarios, necesarios pero no suficientes para el desarrollo de la enfermedad (Badano, J. L. and Katsanis, N., 2002).

	PKD	NPHP	LCA	SLSN	EVC	OFD	JS	MKS	JATD	ALMS	BBS
Malformaciones del SNC					X	X	X	X	X	X	X
Quistes renales	X	X		X		X	X	X	X	X	X
Diabetes										X	X
Malformaciones de las gónadas											X
Problemas cardíacos										X	X
Disfunción hepática	X	X	X	X		X	X	X	X	X	X
Retardo mental			X		X	X	X	X	X		X
Obesidad										X	X
Polidactilia					X	X	X	X	X		X
Disfunción pulmonar										X	
Degeneración de la retina			X	X			X	X	X	X	X
<i>Situs Inversus</i>		X		X			X	X	X		X
Malformaciones del esqueleto					X	X			X		

**Tabla 1. Fenotipos observados en las distintas ciliopatías.**

PKD: Poliquistosis renal; NPHP: Nefronoptosis; LCA: Amaurosis congénita de Leber; SLSN: Senior-Løken; EVC: Ellis van Creveld; OFD: *Oral-facial-digital* tipo 1; JS: Joubert; MKS: Meckel-Gruber; JATD: Displasia torácica asfixiante de Jeune; ALMS: Alström; BBS: Bardet-Biedl.

Tabla adaptada de: Cardenas-Rodríguez, M. and Badano, J. L., 2009.

Desde un punto de vista celular, los modelos de oligogenicidad se basan en la existencia de sistemas redundantes en los que distintas proteínas pueden suplantar la función de otras y por lo tanto se necesitan mutaciones en más de un gen para afectar la funcionalidad de la vía o complejo proteico en el que participan (Badano, J. L. and Katsanis, N., 2002). Esta complejidad hace que las ciliopatías hayan atraído un amplio interés en el campo de la genética. Una de las ciliopatías donde comenzó a disecarse este tipo de interacción es el Síndrome de Bardet-Biedl, objeto de estudio del Laboratorio de Genética Molecular Humana del Instituto Pasteur de Montevideo.

### 1.3. El Síndrome de Bardet-Biedl

El Síndrome de Bardet-Biedl (BBS, del inglés: *Bardet-Biedl syndrome*) fue descrito por primera vez en los años veinte por George Bardet y Arthur Biedl, quienes presentaron independientemente casos de pacientes con obesidad, discapacidad visual, retardo mental, polidactilia e hipogonadismo (Bardet, G., 1920; Biedl, A., 1922). Actualmente, casi 100 años después, la patología es definida por seis rasgos feotípicos principales, considerados como los



**Figura 4. Rasgos fenotípicos del Síndrome de Bardet-Biedl.**

**A-D-** Rasgos faciales típicos de BBS.

**E-** Braquidactilia y cicatrices debido a la escisión de dedos extra.

**F-** Apiñamiento dental.

**G-** Paladar arqueado alto.

**H-** Imagen oftalmoscópica que muestra una distrofia de conos rojos.

Figura tomada de: Forsythe, E. and Beales, P. L., 2013.

“sellos” de la enfermedad, de los cuales cinco son los observados en los primeros estudios de Bardet y Biedl (alguno de estos se observan en la Figura 4): degeneración de retina, obesidad, hipogonadismo, polidactilia, retardo mental y disfunciones renales (Zaghloul, N. A. and Katsanis, N., 2009). A su vez, los pacientes pueden también presentar otras patologías como ataxia, déficit en el habla, pérdida de audición, anosmia, anomalías dentales, defectos del corazón y diabetes entre otras (Forsythe, E. and Beales, P. L., 2013).

La prevalencia de esta enfermedad es baja pero variable según la población a analizar, yendo desde 1:160.000 en poblaciones europeas hasta 1:13.500 en poblaciones arábicas. De todas formas, la incidencia estimada en poblaciones caucásicas es de 1:100.000 (Forsythe, E. and Beales, P. L., 2013). Si bien no existen datos de la prevalencia de BBS en nuestro país es razonable asumir una incidencia en este rango. BBS puede ser heredada como una enfermedad autosómica recesiva, pero como ya se dijo, es clasificada como una enfermedad genéticamente compleja con herencia oligogénica (Katsanis, N. et al, 2001; Beales, J. P., et al, 2003).

Hasta el momento, se han identificado 21 genes causales de BBS y algunos modificadores secundarios, como CCDC28B (del inglés: *Coiled coil domain containing protein 28B*) (Novas, R. et al, 2015; Badano, J. L. et al, 2006; Schaefer, E. et al, 2016; Heon, E. et al, 2016). En cuanto a las proteínas codificadas por los genes BBS, al menos para las caracterizadas hasta la fecha, se localizan principalmente en el cuerpo basal, centrosoma o axonema ciliar y se ha demostrado que participan en la formación y funcionamiento de la cilia primaria (Kim, J. C. et al, 2004; Kim, J. C. et al, 2005; Nachury, M. V. et al, 2007). Para ocho de ellas (BBS 1, 2, 4, 5, 7, 8, 9 y 18) se sabe que su funcionamiento es a través de un complejo octamérico denominado BBSoma que actúa activamente en el proceso de ciliogénesis (Loktev, A. V. et al, 2008). En particular, el BBSoma actúa regulando el transporte de proteínas de membrana hacia y desde las cilias a través de su

asociación con las GTPasas Rabin8 y Rab8 (Nachury, M. V. et al, 2007). Pero no todas las BBS forman parte del BBSoma y últimamente se ha trabajado en búsqueda de disecar las funciones de cada una de ellas fuera del complejo como es el caso de BBS3, también llamada Arl6 (del inglés: *ADP-ribosylation factor like 6*) que se asocia al BBSoma, sin formar parte de este, para mediar su translocación hacia el interior de las cilias (Jin, H. et al, 2010). También se demostró que BBS6, BBS10 y BBS12 pertenecen a la clase de chaperoninas tipo II y colaboran con el ensamblaje del BBSoma (Kim, J. C. et al, 2004; Stoetzler, C. et al, 2006; Stoetzler, C. et al, 2007). Por otro lado, también se sabe que la ausencia de alguna de las proteínas BBS por separado, también genera problemas durante la transducción de señales reguladas por la cilia primaria, como la vía de Shh y Wnt (Gerdes, J. M. and Katsanis, N., 2008; Zhang, Q. et al, 2012).

Pese a que en los últimos años se ha avanzado en la caracterización de las diferentes BBS, para entender la base celular de las manifestaciones clínicas de la enfermedad se necesita conocer en detalle las funciones de estas proteínas, así como también la de los modificadores secundarios. En particular durante este trabajo nos centramos en el modificador secundario CCDC28B.

#### 1.4. El modificador secundario CCDC28B

La hipótesis de oligogenicidad para BBS alentó estudios apuntando a la identificación y caracterización de proteínas que interactuaran con las BBS conocidas como forma de ahondar en la función de estas proteínas y para identificar nuevos genes capaces, ya sea de causar BBS, como modificar la presentación de la enfermedad. La identificación de CCDC28B se realizó a partir de ensayos de doble híbrido en levaduras, utilizando a BBS4 como “carnada”, y una librería de cerebro fetal humano. En ese ensayo se encontraron 60 posibles interactores (Badano, J. L. et al, 2006). Debido a su posible actividad en el funcionamiento de la cilia primaria, esta lista de interactores se filtró con la base de datos del proteoma ciliar, una lista de proteínas enriquecida en proteínas asociadas a la cilia obtenida por análisis bioinformáticos, proteómicos y de expresión (Li, J. B. et al, 2004; Gherman, A. et al, 2006). Allí, de los 60 candidatos iniciales, la única secuencia que se encontraba en la base de datos fue la de CCDC28B, una proteína hasta ese momento totalmente desconocida (Badano, J. L. et al, 2006). Luego se demostró que la interacción de CCDC28B no está restringida a BBS4, sino que también interactúa con otras proteínas BBS, como BBS6 y BBS1 (Badano, J. L. et al, 2006). Resultó muy interesante detectar en pacientes BBS la presencia de una transversión C-T en la penúltima posición del exón 3 (C430T) de *CCDC28B* (el gen está compuesto por 6 exones). Se determinó que este cambio genera una secuencia de unión para la proteína



**Figura 5. Modificación del *splicing* en la variante C430T de CCDC28B.**

La mutación C430T genera una secuencia de unión para SC35 (proteína reguladora del *splice* (indicada en violeta). La unión de esta proteína favorece la delección de 5 pares de base (indicados en naranja) y la aparición de un codón de terminación prematuro (PTC).

Figura adaptada de: Badano, J. L. et al, 2006.

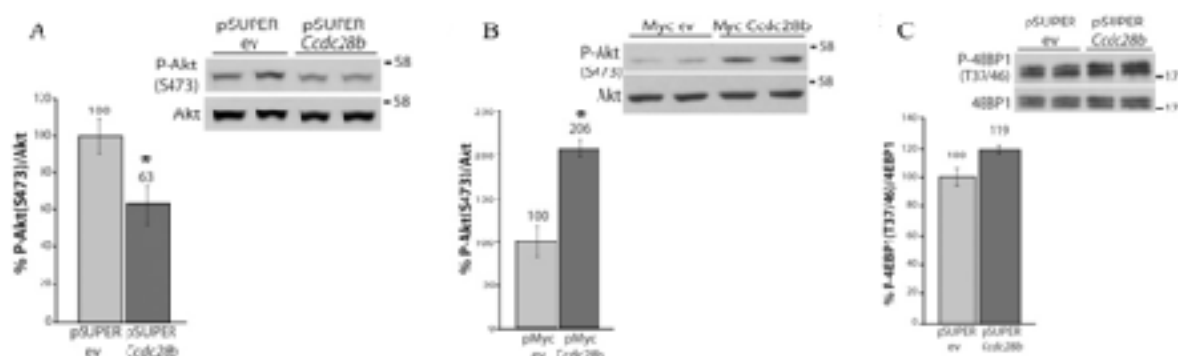
reguladora del *splice* SC35, que potencia el uso de un sitio aceptor de *splice* críptico, generando una isoforma con una pérdida de cinco pares de bases que resulta en un cambio en el marco abierto de lectura con la generación de un codón de terminación prematuro (ver Figura 5). Por lo tanto, el efecto generado sobre el procesamiento del ARNm determina que el cambio C430T se considere una mutación hipomórfica, ya que disminuye la cantidad de ARN mensajero de *CCDC28B* disponible para la traducción. Análisis poblacionales de esta mutación confirmaron el carácter de modificador secundario de *CCDC28B*. Se estableció que la mutación sola no es capaz de generar la enfermedad, ya que no se encontraron pacientes homocigotas C430T y sí individuos sanos (Badano, J. L. et al, 2006). Además, se observaron familias en las que, los individuos afectados de BBS eran aquellos que, además de ser homocigotos para una mutación en un gen causal BBS, eran heterocigotas para C430T en *CCDC28B*. En estas mismas familias, otros individuos homocigotas para la mutación en el gen BBS, pero sin mutaciones en *CCDC28B*, o los heterocigotas para los dos genes, no estaban afectados por la patología. Finalmente, también se encontraron familias con hermanos enfermos, homocigotas para la misma mutación en el gen BBS, pero con presentaciones diferentes de la enfermedad: aquellos que tenían una mutación en el *CCDC28B*, tenían una presentación más severa de la enfermedad. En conclusión, esta mutación hipomórfica en *CCDC28B* no sería suficiente para generar la enfermedad, pero en algunas familias con un contexto genéticamente sensibilizado por mutaciones en otros genes BBS, contribuyen para una manifestación más severa de la enfermedad (Badano, J. L. et al, 2006).

### 1.5. *CCDC28B* y su función

Se identificaron en un comienzo dos isoformas de *CCDC28B*: la NM\_024296 y la BC\_002462, las cuales difieren en el último exón. Aunque la importancia fisiológica de cada una de estas isoformas aún está por ser determinada, se sabe que ambas mantienen la porción del exón tres que se encontró mutada. Luego, con el avance de los estudios en genómica se estableció a la isoforma NM\_024296 como la forma canónica, mientras que la otra isoforma sería producto de la aparente retención de un segmento de intrón (ver <http://www.uniprot.org/uniprot/Q9BUN5>). Desde entonces se trabaja con la isoforma NM\_024296, y es a ella a la que nos referiremos a partir de ahora como *CCDC28B*.

*CCDC28B* codifica una proteína de 200 aminoácidos, conservada en todos los metazoarios ciliados (Cardenas-Rodríguez, M. et al, 2013 -B-) pero que no posee homólogos de función conocida y de la que no se posee información estructural. Una forma de aproximarse a la función de *CCDC28B* es identificar y caracterizar a las proteínas con las cuales interactúa. De esta forma en trabajos previos de nuestro laboratorio, se determinó que tanto en células en cultivo, como *in vivo* en pez cebra (*Dania rerio*), *CCDC28B* juega un rol importante en ciliogénesis (Cardenas-Rodríguez, M. et al, 2013-B-) y participa en la modulación de la vía mTORC2 a través de su interacción con la proteína SIN1 (Cardenas-Rodríguez, M. et al, 2013-A-).

La vía de mTOR (del inglés: *mechanistic or mammalian target of rapamycin*) es un regulador central de la homeostasis celular. Integrando señales intra y extracelulares juega un rol fundamental como regulador del metabolismo, crecimiento, proliferación y diferenciación celular (Laplante, M. and Sabatini, D. M., 2009). La vía de mTOR nuclea al menos dos complejos multiproteicos diferentes: el complejo 1 (mTORC1) y el complejo 2 (mTORC2). mTORC1 fue el primero en ser caracterizado, está formado por cinco componentes: mTOR, serina/treonina quinasa que nuclea ambos complejos, Raptor, mLST8, PRAS40 y DEPTOR, y por su función es definido como un regulador central del crecimiento celular y el metabolismo (Laplante, M. and Sabatini, D. M., 2012). Por otro lado, mTORC2 está formado por seis componentes diferentes: mTOR, Rictor, mSin1, Protor-1, mLST8 y DEPTOR. A diferencia de lo que sucede con mTORC1 se conoce menos sobre la biología del complejo mTORC2, en gran medida debido a que en ratones la delección de componentes de este complejo causa letalidad y tampoco se conocen inhibidores específicos del mismo, lo que dificulta su estudio (Laplante, M. and Sabatini, D. M., 2009). De todas maneras, se ha identificado que el complejo mTORC2 tiene un rol importante en la supervivencia, proliferación, metabolismo y fundamentalmente la reorganización del citoesqueleto de actina y la migración celular (Oh, W. J. and Jacinto, E., 2011). Nuestro grupo demostró que CCDC28B interactúa con SIN1 y a través de esta interacción afecta la actividad de mTORC2, actuando como agonista o regulador positivo de la vía (Cardenas-Rodríguez, M. et al, 2013 -A-; Cardenas-Rodríguez, M. et al, 2013 -B-). mTORC2 fosforila Akt en su dominio hidrofóbico, en la S473 (Oh, W. J. and Jacinto, E., 2011) y por ende los niveles de Akt S473 fosforilados pueden ser un buen reportero de la actividad del complejo. En la Figura 6 se muestran los resultados obtenidos por Cardenas-Rodríguez, et al. en 2013 (A) sobre los niveles de Akt S473 fosforilado al disminuir y sobre expresar CCDC28B. Allí se observó que la disminución de CCDC28B resulta en una reducción de la actividad de mTORC2 y la sobreexpresión de la proteína aumenta la actividad del mismo.



**Figura 6. CCDC28B modula la actividad del complejo mTORC2.**

**A-** La bajada de CCDC28B en células NIH3T3 resulta en una disminución de la fosforilación de AktS473, lo que indica una menor actividad de mTORC2. Se utilizó Akt total como control de carga.

**B-** La sobreexpresión de CCDC28B en células NIH3T3 resulta en un aumento de la actividad de mTORC2. Se utilizó Akt total como control de carga.

**C-** La bajada de CCDC28B en células NIH3T3 no afecta la actividad de mTORC1, medido a través de la fosforilación de 4EBP1 (T37/46) -sustrato de mTORC1-. Se utilizó 4EBP1 total como control de carga.

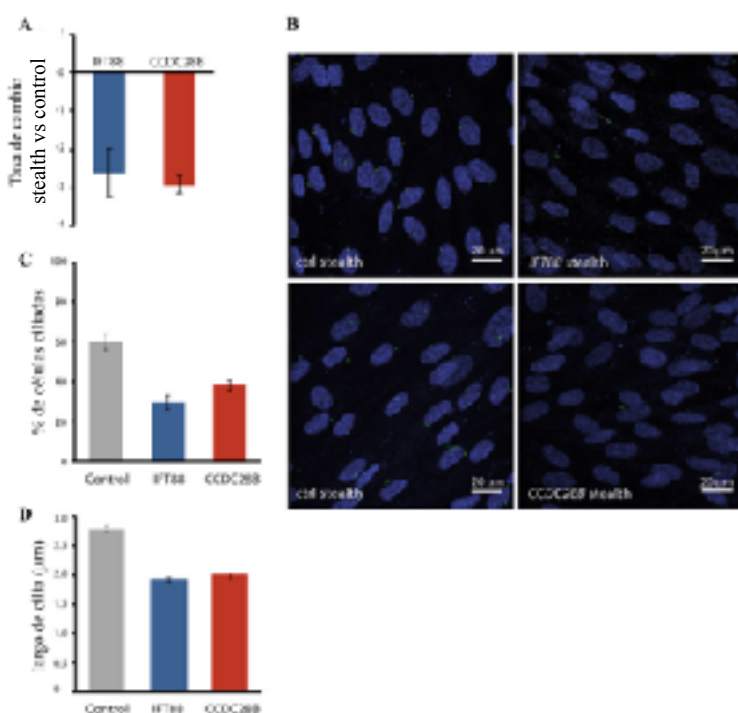
Figura adaptada de: Cardenas-Rodríguez, M. et al, 2013 -A-



Además se mostró que CCDC28B regula ciliogénesis, tanto en células en cultivo como *in vivo* en el pez cebra (Cardenas-Rodríguez, M. et al, 2013 -A-). En ambos casos se observó que la disminución de los niveles de CCDC28B reduce el porcentaje de ciliación y el largo de cilia (Cardenas-Rodríguez, M. et al, 2013 -A-). La Figura 7 muestra los resultados obtenidos en ese trabajo con células hTERT-RPE en cultivo, donde se observa la disminución de la ciliación y el largo de las cilia cuando disminuyen los niveles de CCDC28B, comparado con un control positivo como lo es IFT88 (del inglés: *intraflagellar transport gene 88*), indispensable para la ciliogénesis. Mediante ensayos de interacción se pudo ver que el efecto de CCDC28B sobre ciliogénesis depende, al menos en parte, de su interacción con SIN1 pero en forma independiente de mTORC2: mientras que la disminución de RICTOR (miembro del complejo mTORC2 e interactor de SIN1) y SIN1 disminuye la actividad mTORC2, solo esta última altera el largo de la cilia (Cardenas-Rodríguez, M. et al, 2013 -A-). Finalmente, también se demostró en ensayos *in vivo*, que la sobreexpresión de SIN1 es capaz de rescatar, al menos parcialmente, el fenotipo ciliar producido por la disminución de CCDC28B, mientras que la sobreexpresión de CCDC28B, no rescata el fenotipo ciliar inducido por la falta de SIN1. Estos resultados nos hablan de que SIN1 actúa corriente abajo de CCDC28B. A pesar de estos avances, aún no se conoce cómo CCDC28B interacciona con SIN1 y cómo esto se traduce en los efectos observados sobre ciliogénesis y la actividad de mTORC2.

Por otro lado, cómo ya mencionamos, CCDC28B fue originalmente identificada por su interacción con diferentes proteínas BBS (Badano, J. L. et al, 2006), aunque se desconocen las consecuencias funcionales de estas interacciones. En este contexto, pensamos que resolver la estructura de CCDC28B permitirá acercarse a la comprensión de cómo interacciona con estas proteínas y cómo podría regular sus funciones. También parece importante comenzar a obtener información estructural sobre proteínas mutadas en ciliopatías, ya que si bien hay muchas proteínas

descritas, no existe información estructural sobre casi ninguna de ellas.



**Figura 7. CCDC28B regula ciliogénesis.**

**A-** Eficiencia de los stealth RNA oligos, tanto de IFT88 (utilizado como control) como de CCDC28B, medidos por Real Time PCR. Se utilizó GAPDH para normalizar.

**B-** Imagen confocal de células hTERT-RPE transfectadas con Low GC Stealth RNAi control, o con los oligos RNAi para CCDC28B e IFT88. Para visualizar el núcleo y la cilia se utilizaron TOPRO (azul) y tubulina acetilada (verde), respectivamente.

**C-** Porcentaje de células ciliadas luego de la transfección con los Stealth: control, IFT88 y CCDC28B. Se contaron más de 100 núcleos y se calculó el porcentaje contento el número de cilia en esos núcleos.

**D-** Se midieron cilia de más de 300 células tanto en el control como en las depletadas para IFT88 y CCDC28B.

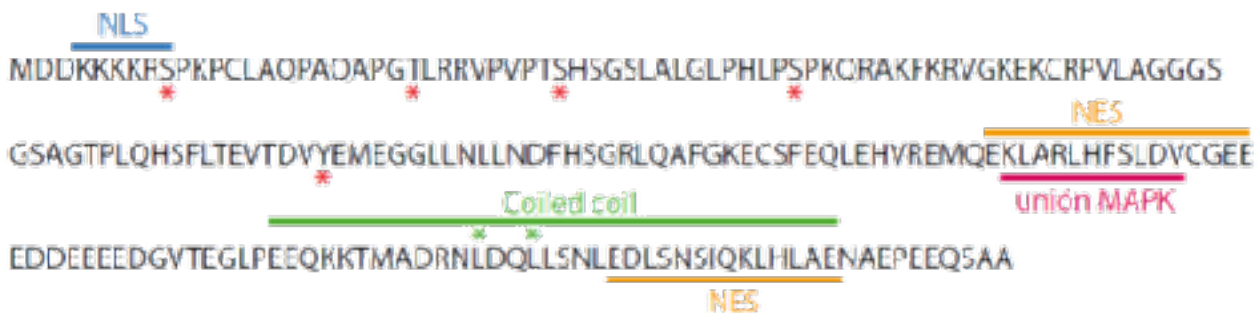
Figura adaptada de: Cardenas-Rodríguez, M. et al 2013 -B-

## 1.6. Aproximación a CCDC28B desde predicciones informáticas

Como se dijo previamente, no se tiene ninguna información estructural sobre CCDC28B. El análisis *in silico* de su secuencia aminoacídica, que se observa en la Figura 8, predice que en el extremo carboxilo terminal existe un dominio de tipo “*coiled-coil*” (Lupas, A., 1997), mientras que el extremo amino terminal podría comportarse como una región desestructurada (Ward, J. J. et al, 2004; Cardenas-Rodríguez, M. et al, 2013 -A-).

El dominio de tipo *coiled-coil* está caracterizado normalmente por dos a cinco hélices  $\alpha$  anfipáticas, enrolladas entre sí formando una super-hélice. Es un motivo típicamente involucrado en mediar interacciones proteína-proteína siendo también el principal motivo de oligomerización entre subunidades de proteínas (Burkhard, P. et al, 2001). En el caso de CCDC28B es el único dominio identificable, que además es el que le da nombre a la proteína y se encuentra entre los aminoácidos Glu158 y Glu193. Interesantemente, se ha determinado que tanto los centrosomas como las propias cilias están enriquecidas en proteínas con dominios de este tipo sugiriendo que la interacción entre proteínas es un aspecto fundamental en la biología de estos organelos (Nido, G. S. et al, 2012).

En el otro extremo de la proteína, los primeros 75-80 aminoácidos se predicen como desordenados (PSIPRED Protein Sequence Analysis Workbench, <http://bioinf.cs.ucl.ac.uk/disopred>). El hecho de que esta región pueda estar desorganizada no le resta importancia. A la fecha hay muchos estudios de asociación entre desorden intrínseco y función de las proteínas (Dunker, A. K. et al, 2002). Pese a que algunos insisten en asociar directamente la función a la estructura tridimensional, el número creciente de proteínas que se sabe son desestructuradas nos obliga a mirar esto desde otro punto de vista, ya que de lo contrario no se explicaría como este tipo de proteínas se han mantenido a lo largo de la evolución. Es más, incluso se ha establecido que muchas de estas están frecuentemente involucradas en las funciones regulatorias más importantes de la célula y su falta de estructura intrínseca deja de ser tal cuando se unen a sus ligandos, ya que esta flexibilidad le permitiría además unirse a más de un interactivo (Wright, P. E. and Dyson, H. J., 1999).



**Figura 8. Secuencia aminoacídica de CCDC28B humana (NM\_024296).**

Se observa la secuencia de aminoácidos de CCDC28B en la que se encuentran resaltados los sitios predichos *in silico* que pueden estar potencialmente involucrados en las funciones de la proteína. En azul se observa la secuencia de localización nuclear (NLS), en amarillo las secuencias de exportación nuclear (NES), en verde la región correspondiente al dominio coiled coil, en rosado el sitio de unión a MAPK y con asteriscos rojos se marcan los sitios de fosforilación.

CCDC28B posee a su vez otras regiones predichas *in silicio* que podrían ser importantes para determinar la localización sub-celular de la proteína, para regular su actividad o para interactuar con otras proteínas (ver Figura 8). En particular se predicen: un dominio de unión de la quinasa MAP (MAPK, del inglés: *mitogen-activated protein kinase*) entre los residuos K128 y V138; una señal de localización nuclear (NLS, del inglés: *nuclear localization signal*) en el extremo amino terminal (K4-S9) y dos señales de exportación nuclear (NES, del inglés: *nuclear export signal*) centradas en los aminoácidos F134 y D178. Los dominios NLS y NES sugieren que CCDC28B podría también tener roles extracelulares, o que su función en la regulación de la cilíogenesis dependiera del transporte hacia dentro y fuera del núcleo.

Nuestro grupo ha generado un panel de mutantes de CCDC28B en estas regiones y/o residuos potencialmente relacionados a la función de la proteína. Algunos de estos mutantes fueron evaluados en este trabajo para determinar la participación de dichas regiones en las actividades de CCDC28B, como una aproximación complementaria para obtener información funcional acerca de la proteína. Los mutantes generados son:

1. Mutantes en el dominio *coiled-coil*: El dominio *coiled-coil* fue predicho a través del servidor COILS ([http://www.ch.embnet.org/software/COILS\\_form.html](http://www.ch.embnet.org/software/COILS_form.html)). Para desestabilizar este dominio se mutagenizaron los residuos leucina 169 y 172 a ácido aspártico. Como control se generó un mutante (N168D y Q171D) en el que no se altera la predicción del dominio *coiled-coil*.
2. Mutantes en el dominio de interacción con proteínas MAPKs: a través del sitio ELM (<http://elm.eu.org/links.html>) se identificó este motivo, que podría estar implicado en interacciones específicas con miembros de la cascada MAPK. SIN1 también se conoce como MAPKAP1 (Mitogen Activated Protein Kinase Associated Protein1) por lo que este motivo resulta particularmente interesante a la hora de evaluar la interacción CCDC28B-SIN1 así como su capacidad de activar el complejo mTORC2.
3. Mutantes en residuos fosforilables: la fosforilación de proteínas es una modificación post-traduccional que puede regular no sólo actividad sino también la localización subcelular de una proteína. CCDC28B posee tres serinas (S9, S32, S46), una treonina (T23) y una tirosina (Y89) que podrían ser fosforiladas (<http://www.cbs.dtu.dk/services/NetPhos/>). Estos residuos fueron modificados por alaninas (mutantes fosfo-nulos) o ácido aspártico (fosfo-miméticos). En particular resulta interesante la T23 como posible blanco de PKC, una quinasa involucrada en la reorganización del citoesqueleto y la actividad de mTORC2.
4. Mutantes en las secuencias de localización y exportación nuclear: En particular la región NLS fue predicha utilizando el servidor: <http://www.moseslab.csb.utoronto.ca/NLStradamus/>. Además se predicen dos señales NES. Estas regiones parecen especialmente interesantes ya que podrían determinar la localización subcelular de CCDC28B. Se generaron mutantes donde se sustituyeron residuos claves de la NLS

(K5) y de la NES (L132 y L179) por alaninas de forma de eliminar la funcionalidad de estas regiones.

En todos los casos los cambios introducidos eliminan las predicciones *in silico* de los correspondientes sitios o secuencias funcionales.

### 1.7. De la estructura a la función

Hasta hace poco tiempo las proteínas a las que se le realizaba su determinación estructural eran seleccionadas según su función, previamente determinada, y de forma tal que la estructura tridimensional de la misma aportara detalles exactos del mecanismo que hay detrás de esa función. Esto hacía que los estudios estructurales fueran el punto final del proceso de descubrimiento. Ejemplos de este tipo sobran por ser una práctica muy común hasta hace pocos años. Uno particular es el de la tripsina, sobre la cual se sabía que era una proteasa, y su especificidad y cinética estaban muy bien descritas mucho tiempo antes de que se resolviera su estructura tridimensional (Stroud, R. M. et al, 1972). Este paradigma tradicional ha cambiado significativamente con el desarrollo de los diferentes proyectos de genómica estructural que plantean la posibilidad de utilizar los estudios estructurales más temprano en el proceso científico apuntando a preguntas como: ¿cuál es la función de esta proteína? (Shapiro, L. and Harris, T., 2000).

Las estructuras tridimensionales de las proteínas una vez obtenidas, plantean algunas ventajas por sobre el abordaje clásico de comparar secuencias realizando un BLAST (del inglés: *basic local alignment search tool*). En primer lugar la comparación de estructuras puede aportar información sobre el parentesco con proteínas de función conocida, que no necesariamente son comparables por secuencia. Y en segundo lugar, la información estructural de una proteína puede colaborar con la identificación de motivos de unión o sitios catalíticos, incluso en proteínas de función desconocida (Shapiro, L. and Harris, T., 2000).

Es sabido que la estructura de una proteína, en términos de su plegamiento general, es más conservada durante la evolución que la secuencia. A lo largo del tiempo dos secuencias pueden divergir mientras que las proteínas codificadas mantienen su plegamiento y posiblemente su función. El ejemplo más claro de esto quizás pueda ser el de los transportadores de oxígeno en la sangre y músculo, la hemoglobina y la mioglobina, respectivamente. En este caso, sus secuencias parecen no estar relacionadas al punto que si se realiza un BLAST con la secuencia de una la otra no aparece como resultado. Sin embargo, sus estructuras tridimensionales son extremadamente similares (Aronson, H. G. et al, 1994). Otro ejemplo se muestra en la Figura 9, y es el caso de la AdipoQ, una proteína secretada exclusivamente por adipocitos. La proteína se descubrió durante un estudio de secuenciación masiva de ADNc de adipocitos y la función de la misma era totalmente desconocida. La estructura resuelta por cristalografía de rayos X reveló una clara similitud estructural con la familia de citoquinas TNF (del inglés: *tumor necrosis factor*) y que también se



**Figura 9. Ejemplo de similitud inesperada entre proteínas sin ninguna homología de secuencia, el factor secretado por adipocitos AdipoQ (Acrp30) y la familia de citoquinas TNFs.**

**A-** Comparación entre diagrama de cintas de los trímeros de AdipoQ y TNF $\alpha$

**B-** Alineamiento de secuencia basado en estructura entre los TNFs: CD40L, TNF $\alpha$ , TNF $\beta$  y los dominios C1q-like de C1qA y AdipoQ. Los residuos idénticos en las cinco proteínas están sombreados de rojo, mientras que los residuos conservados en cuatro de las cinco proteínas están sombreados en azul. Las regiones de hojas  $\beta$  están indicadas con flechas verdes sobre la secuencia de AdipoQ y con flechas azules por debajo de la secuencia de CD40L. Las búsquedas basadas en secuencia como el BLAST no pueden identificar ningún tipo de relación entre el dominio C1q de AdipoQ y los miembros de la familia TNF.

Figura adaptada de: Shapiro, L. and Harris, T., 2000

ilustra en la Figura 9 (Shapiro, L. and Scherer, P. E., 1998). Esta similitud despertó la hipótesis de que AdipoQ puede funcionar como un factor de señalización celular, algo que al día de hoy ya se ha documentado (Oliveira, C.

S. et al, 2011). El hecho de que la similitud con la familia de las TNF no podía haber sido encontrada de otra manera que no fuera comparando las estructuras refuerza la importancia de este tipo de estudios.

Hoy por hoy, y debido al desarrollo de la genómica estructural, existen diferentes herramientas con el objetivo de acercarnos a la función mediante estudios estructurales. Para realizar una búsqueda informática de una estructura determinada con respecto a los patrones de plegamiento conocidos y depositados en la PDB (del inglés: *protein data bank*) existen distintos servidores basados en métodos como Dali, GRATH, VAST o SSM (todos ellos están revisados en: Laskowski, R. A. et al, 2002 y Najmanovich, R. J. et al 2005).

Creemos por lo tanto que la resolución de la estructura tridimensional de CCDC28B puede abrir la puerta a otro abanico de herramientas y posibilidades que nos ayuden a entender cómo la proteína afecta ciliogénesis, interactuando con otras proteínas, posiblemente regulando la función de las mismas y por lo tanto los procesos donde estas intervienen.

## 1.8. Determinación de la estructura tridimensional de una proteína

Si bien se pueden utilizar diferentes estrategias para la elucidación de la estructura tridimensional de proteínas, la cristalografía de rayos X y la resonancia magnética nuclear (NMR, del inglés: *nuclear magnetic resonance*) son, hasta el momento, los dos métodos principales que permiten conocer la estructura a resolución atómica. Las dos técnicas pueden ser de alguna manera complementarias y sus diferencias han sido tema de discusión a lo largo de los años (Garbuzynskiy, S. O. et al, 2005; Lee, A. A. et al, 2005; Snyder, D. A. et al, 2005). Actualmente se utiliza también la crio-microscopía electrónica, la cual emplea temperaturas extremadamente bajas y el congelado rápido de las muestras. Si bien la técnica está en pleno desarrollo aún no permite llegar a resoluciones tan buenas como las otras dos mencionadas anteriormente (entre 5 y 10 Å) y por ser una derivación de la microscopía es útil para estructuras de tamaño grande, lo suficiente al menos para ser detectadas por el instrumento (por lo menos 150 kDa). Pese a esto, esta técnica todavía está en desarrollo y siguen siendo las dos anteriores las más ampliamente utilizadas.

Las características más importantes de la cristalografía de rayos X y el NMR están descritas en la Tabla 2. Sin lugar a dudas, la gran diferencia entre ambas es que mientras que para la primera, como su nombre lo indica, es necesaria la formación de un cristal de proteína adecuado, para NMR la proteína se mantiene en estado soluble. Pese a que aparenta una diferencia muy importante, en realidad dentro del cristal las proteínas forman una red periódica que no es estrictamente inmóvil, debido a que es formada por interacciones débiles y no covalentes que contienen amplios canales de solvente. Estos canales, permiten una relativa difusión libre de pequeñas moléculas a través del cristal y a la vez permite una cierta libertad conformacional de cadenas laterales expuestas y bucles (Rupp, B., 2010). Incluso hay enzimas que mantienen su actividad en su forma cristalina. Estos datos determinan que el núcleo central de las estructuras resueltas por NMR y cristalografía de rayos X sea igual. El problema para la cristalografía comienza cuando se pretende resolver la estructura de proteínas con grandes movimientos conformacionales o regiones de gran libertad, problemas que podrían ser solucionados por NMR.

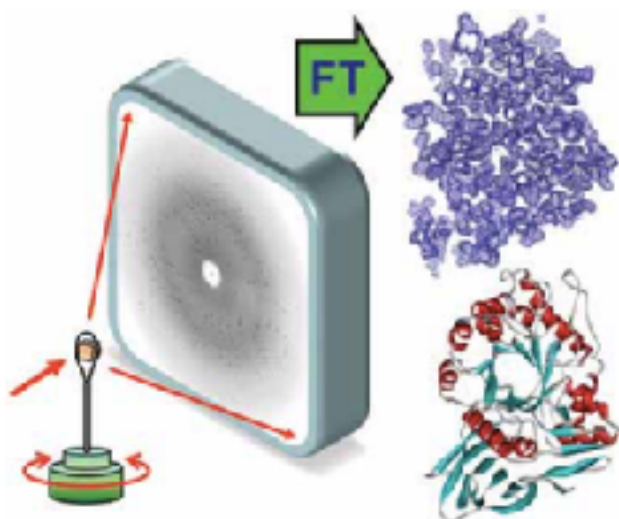
	<b>Cristalografía de rayos X</b>	<b>NMR</b>
<b>Requerimientos</b>	La macromolécula debe cristalizar y ese cristal debe difractar	La macromolécula debe ser soluble y marcada con isótopos pesados
<b>Estado de la muestra</b>	Estado solido-congelado	En solución buffer
<b>Tamaño</b>	Cualquier macromolecula	Pequeños polipéptidos (menores a 50kDa)
<b>Tiempo</b>	Largas pruebas de cristalización y optimización, pero poco tiempo de procesamiento de los datos	Poco tiempo de preparación pero mayor tiempo de análisis de los datos
<b>Ventajas principales</b>	Normalmente se obtiene alta resolución atómica para cristales que difractan bien	El estado solubles es más cercano al ambiente funcional y a su plegamiento en el mismo
<b>Desventajas principales</b>	La necesidad de resolver el problema de las fases, las pruebas de cristalización y optimización y en algunos casos la cristalización puede cambiar la conformación debido a interacciones debidas al empaquetamiento.	El marcaje con isótopos que contengan momento magnético puede ser caro; cuanto mayor sea el tamaño de la proteína puede acarrear un espectro más pobre e inestabilidad durante la colecta de datos.

**Tabla 2. Principales características comparativas de la cristalografía de rayos X y la resonancia magnética nuclear (NMR).**

En la cristalografía de rayos X, el principio básico es tomar un único cristal de una proteína, enfocalo bajo un haz de rayos X, y grabar las imágenes de difracción de los rayos X al atravesar el cristal. Cada uno de estos procesos es de por sí complejo y no es el objetivo profundizar en ellos, aunque en la Figura 10 se puede ver una representación esquemática del proceso. Luego de obtenidos los datos de difracción, la densidad electrónica que representa la estructura atómica de la molécula es reconstruida a través del método de Fourier (Figura 10). Finalmente se construye un modelo molecular que se ajuste a esa densidad electrónica. Una de las ventajas de la cristalografía sobre NMR es que normalmente se puede llegar a resoluciones mayores ( $0.6\text{-}4\text{\AA}$  vs  $2\text{-}3\text{\AA}$ ), pero para que esto suceda se necesita producir un cristal que difracte bien. Por lo tanto, el proceso de cristalización es uno de los cuellos de botella más grandes de esta técnica, siendo una disciplina casi artesanal. Pese a los cientos de estudios al respecto (McPherson, A. and Gavira, J. A., 2014) hay muy pocos parámetros generalizables en términos de la obtención de cristales con buena difracción, y por lo tanto el proceso se basa en gran medida en una cuestión de ensayo y error.

Por otro lado, el NMR se basa en la absorción de radiación electromagnética por los núcleos atómicos dispuestos en un campo magnético externo. El campo magnético, consigue excitar los spin nucleares hacia su estado de máxima energía. Cuando se cortan las radiofrecuencias, los spin nucleares vuelven a su estado de baja energía, liberando una pequeña cantidad de radiación. La radiación emitida induce una señal de radiofrecuencia en un detector que luego es amplificada y transformada por el método de Fourier para producir la señal de NMR. Las señales provocadas por átomos con distintos momentos magnéticos, se analizan en conjunto para finalmente llegar a una densidad electrónica de toda la molécula (Goto, N. K. and Kay, L. E., 2000). La velocidad con la que se toman las medidas, mucho menor a la que se necesita en la cristalografía, hacen que las proteínas con mayor movilidad también puedan ser detectadas, consiguiendo una imagen para cada una de sus posiciones en el tiempo.

Como se dijo antes, hoy se utilizan ambas técnicas como complementarias. Pese a esto, los números indican que la mayoría de los científicos se siguen volcando hacia la cristalografía: del



**Figura 10. Principios de determinación de una estructura por cristalografía de rayos X.**

Un cristal montado en un goniómetro, con al menos un eje rotatorio y allí es expuesto a un haz de rayos X (aproximadamente entre los  $2,3$  y  $0,6\text{\AA}$  de longitud de onda). Las distintas imágenes de difracción son guardadas en el detector mientras el cristal va sufriendo pequeñas rotaciones. El patrón de puntos obtenido en las imágenes de difracción corresponderían al espacio recíproco y tienen una relación con la posición de los átomos en el cristal (que representa el espacio real). El patrón de difracción de un cristal es la transformada de Fourier de su estructura. Por lo tanto a través de la inversa de Fourier, habiendo obtenido las fases para cada punto de difracción (proceso complejo de por sí) podemos obtener la densidad electrónica de las moléculas dentro del cristal (representada en azul). Finalmente a partir de la densidad electrónica se puede construir un modelo molecular que se corresponda con la densidad electrónica.

Figura tomada de: Rupp, B., 2010.

total de las estructuras depositadas en la PDB, el 90% corresponde a estructuras resueltas por cristalografía, mientras que solo el 10% corresponde a NMR. En nuestro caso particular, nos volcamos a la cristalografía como primera aproximación no solo por lo antes mencionado sino fundamentalmente por la colaboración con la Plataforma de Cristalografía de Proteínas del IPMon.



## **OBJETIVO GENERAL**

El objetivo general de este trabajo de tesis fue contribuir a entender la función de CCDC28B a partir de caracterizar la proteína desde el punto de vista estructural y establecer relaciones de estructura-función.

## **OBJETIVOS ESPECÍFICOS**

- I. Producción y purificación de CCDC28B recombinante.
- II. Rastreo de condiciones de cristalización de CCDC28B recombinante y determinación de su estructura tridimensional.
- III. Identificar dominios o regiones de la proteína CCDC28B que sean importantes para su función. Este objetivo se realizó en colaboración con otros investigadores del grupo que analizarán estos mutantes *in vivo* utilizando el modelo de pez cebra.

## MATERIALES Y MÉTODOS

---

### 3.1. Clonado en vectores de expresión

El clonado de los genes de interés se realizó en diferentes vectores de expresión dependiendo del objetivo buscado. En todos los casos, el primer paso consistió en amplificar el marco abierto de lectura (ORF, del inglés: *open reading frame*), ya sea partiendo desde ADN copia (obtenido a partir del ARNm) o a partir de otro plásmido en el que estuviera clonado previamente, utilizando la técnica de reacción en cadena de la polimerasa (PCR, del inglés: *polymerase chain reaction*). Luego se utilizaron dos técnicas de clonado para insertarlo en el vector destino: por enzimas de restricción o por amplificación del plásmido entero.

#### 3.1.A. Clonado por enzimas de restricción

Para clonar el marco abierto de lectura de *CCDC28B* salvaje (WT, del inglés: *wild type*) en el vector pET21d se utilizó esta metodología, que consiste en generar extremos cohesivos y compatibles en el fragmento amplificado y en el vector destino, digiriendo luego ambos con las mismas enzimas de restricción, para finalmente ligarlos entre sí mediante la incubación con la enzima ADN ligasa.

Se utilizaron las enzimas de restricción *XhoI* y *NcoI* y los sitios de reconocimiento de estas enzimas se adicionaron introduciéndolos en los oligonucleótidos cebadores que se utilizaron para la amplificación del ORF de *CCDC28B*. Por otro lado también se digirió con dichas enzimas de restricción el vector pET21d vacío. Luego de eliminar las enzimas de restricción, se incuban juntos inserto y plásmido digeridos, en presencia de la enzima ADN ligasa del fago T4, para ligar el fragmento de interés dentro del vector.

Protocolo de digestión:

- 1 µg plásmido.
- 15 U *XhoI*.
- 15 U *NcoI*.
- 3 µl Buffer 10x.
- Completar hasta 30 µl con H<sub>2</sub>O.
- Incubar toda la noche (ON, del inglés: *overnight*) a 37°C.

Protocolo de ligación:

- Plásmido: inserto en una relación 1:2.
- 15 µl de Buffer 2X.
- 5 U Enzima T4 ADN ligasa.
- Completar hasta 30 µl con H<sub>2</sub>O.

- Incubar ON a 4°C.

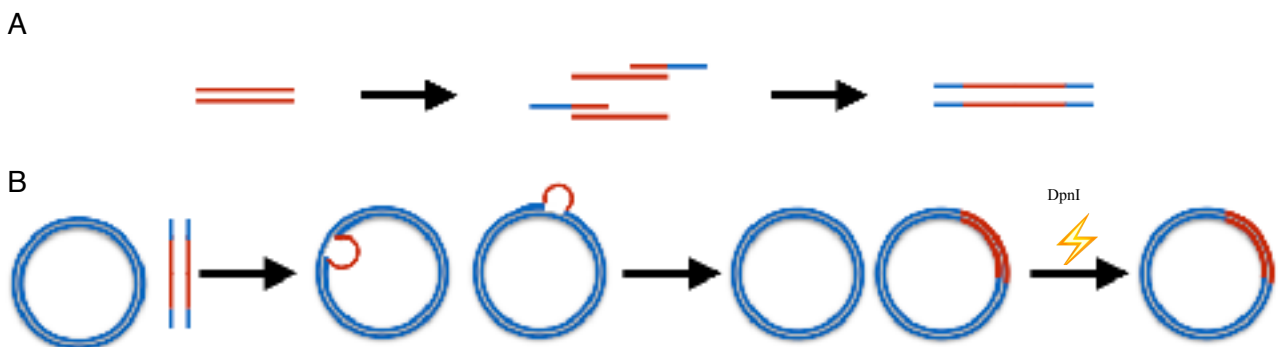
Luego de ligado el vector con el inserto, se transforman bacterias *Escherichia coli* DH5α para amplificarlo (protocolo de transformación descrito más adelante) y luego purificarlo para su uso mediante el kit PureLink™ Quick Plasmid Mini/Maxiprep Kit de Invitrogen. Este kit utiliza una lisis alcalina/SDS para romper las células, precipitación del ADN genómico con acetato de potasio y una columna de intercambio aniónico a la que se une selectivamente el ADN plasmídico.

### 3.1.B. Clonado por amplificación del plásmido entero

El otro protocolo utilizado para el clonado en diferentes plásmidos de expresión consistió en un método basado exclusivamente en la técnica de PCR, denominado amplificación del plásmido entero o RF *cloning* (RFC, del inglés: *restriction free cloning*) (Van den Ent F., Löwe J., 2006).

Este protocolo, que se ilustra en la Figura 11, consta de una primera PCR con la cual se genera un producto que funcionará de mega-cebador en una segunda reacción. Esta segunda reacción es operativamente semejante a una PCR, pero en verdad se trata de una extensión de cebador (*primer extension* en inglés). Para la primera reacción se utilizan cebadores más largos de lo convencional (aproximadamente 50 pares de bases, pb) ya que poseen una región homologa al ORF que se quiere clonar y otra región de homología con el vector de destino. Esta última servirá para hibridizar con el vector en la segunda PCR.

Luego de realizada la primera PCR se separó la mezcla de reacción por electroforesis en gel de agarosa, y se purificó el fragmento correspondiente al mega-cebador, utilizando el kit comercial PureLink™ Quick Gel Extraction Kit (Invitrogen). Una vez purificado, se procedió a la segunda reacción de extensión de cebador, donde el mega-cebador se une al vector de destino por sus



**Figura 11. Clonado por amplificación del plásmido entero (RFC).**

**A** - Primera reacción de PCR, donde se sintetiza el megacebador. Utilizando cebadores con una región de homología al ORF de la proteína de interés (rojo), y otra al vector destino (azul), que se utilizará como molde en la segunda reacción, se amplifica el ORF de la proteína formando el megacebador que tendrá en sus extremos las regiones de homología al vector destino.

**B** - Segunda reacción de PCR donde se amplifica el vector destino añadiendo el ORF de la proteína de interés. Aquí se utiliza como cebador el producto de la primera PCR y como molde el vector destino donde se quiere clonar la proteína de interés (azul). Ambas se unen por su región de homología (azul) y luego desde allí se amplifica el vector completo durante la PCR, obteniendo como producto final el vector con el inserto del ORF.

secuencias complementarias, y se extienden amplificando el plásmido en su totalidad, adicionando así el ORF que codifica para la proteína de interés (ver Figura 11).

Para degradar el ADN parental, la mezcla de reacción fue sometida a una digestión con la enzima *DpnI*. Esta enzima de restricción corta en el sitio de reconocimiento GATC, pero solamente cuando la adenina está metilada. Teniendo en cuenta que el ADN parental está metilado, pues proviene de la biosíntesis *in vivo* en la bacteria, la enzima lo digiere en muchos fragmentos; el producto de PCR en cambio, producto de una reacción *in vitro*, no está metilado y por ende no es digerido por *DpnI*.

El producto de la digestión por *DpnI* luego se purificó utilizando el kit comercial PureLink™ Quick PCR Purification Kit (Invitrogen), y con él se transformaron bacterias *E.coli* DH5α (protocolo de transformación descrito más adelante) para amplificarlo y luego purificarlo.

Los protocolos de cada una de las reacciones de PCR así como la lista de todos los cebadores utilizados se encuentran en el Anexo.

## 3.2. Producción de proteínas recombinantes

### 3.2.A. Plásmidos utilizados

A lo largo de todo este trabajo se utilizaron diferentes plásmidos con distintos objetivos (Tabla 3). En primer lugar, para clonar y expresar CCDC28B WT recombinante, se utilizó el pET21d (Novagen). Este es un vector con un promotor fuerte *T7/lac* y que adiciona una etiqueta de 6 histidinas (6-His) en el extremo C-terminal de la proteína de interés.

También se utilizó el vector pQE80 (Qiagen), que a diferencia del anterior posee un promotor *T5/lac* y adiciona la etiqueta de 6-His en el extremo N-terminal de la proteína. El pQE80 que utilizamos fue cedido gentilmente por la Unidad de Cristalografía de Proteínas del Instituto Pasteur de Montevideo, y había sido modificado por agregado de un sitio de corte para la proteasa TEV (*Tobacco Etch Virus*) entre el inicio del ORF y el 6-His, lo que permite eliminar la etiqueta en caso de ser necesario (ver construcción de este pQE80 en Trajtenberg, F. et al, 2014).

También se utilizó el vector pACYCDuet (Novagen) para co-expresar dos proteínas diferentes (por ejemplo, CCDC28B y alguno de sus interactores), ya que el mismo posee dos sitios de clonado simultáneo, cada uno de ellos con un promotor *T7/lac*. Además, posee el origen de replicación P15A (derivado del plásmido p15A, es de baja copia generando entre 10 y 12 copias por célula) y el gen de resistencia a cloramfenicol.

Finalmente, para algunas pruebas también utilizamos el pT5-*VectorSuite* construido en la Unidad de Proteínas Recombinantes (UPR) del Institut Pasteur de Montevideo (Correa A. et al 2014). Este *VectorSuite* fue construido a partir de pQE80, lo que hace que el mismo posea un promotor *T5/lac* y la etiqueta 6-His en el extremo N-terminal, pero además se le agregó un *Strep-tag* en el extremo C-terminal. El *Strep-tag* es una etiqueta formada por un péptido de ocho aminoácidos (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) que posee una alta afinidad intrínseca por la estreptavidina. Por otro lado esta construcción permite expresar la proteína de interés como fusión C-terminal de una proteína “potenciadora”, que procura favorecer la solubilidad de la fusión. Las proteínas potenciadoras incluidas en este set de plásmidos son: la proteína de unión a maltosa

PLÁSMIDO	PROMOTOR	RESISTENCIA	TAG	PROVEDOR	SITIO DE CORTE
pET21d	T7/lac	Amp/Cam	6His C-term	Novagen	-
pQE80	T5/lac	Amp	6His N-term	Qiagen	TEV
pACYCDuet	T7/lac	Cam	6His C-term StrepTag N-term	Novagen	-
pT5-VS	T5/lac	Amp	6-His N-term StrepTag C-term	UPR IPMont	TEV

Tabla 3. Diferentes vectores de expresión y clonado utilizados a lo largo del proyecto.

(MBP), la tiorredoxina A (Trx), la isomerasa C de puentes disulfuro (DsbC) o la endogluconasa CelD. El vector también posee un sitio de reconocimiento para la proteasa TEV, entre los ORFs de la proteína de interés y la proteína “potenciadora” para cortar la fusión de ser necesario.

### 3.2.B. Cepas bacterianas utilizadas

Se utilizaron diferentes cepas de *E. coli* en distintas etapas del trabajo según el objetivo perseguido (Tabla 4).

En la etapa de clonado y producción de plásmidos, las bacterias utilizadas fueron las DH5 $\alpha$  *max efficiency* (con muy altas eficiencias de transformación  $\sim 1 \times 10^9$  transformantes/ $\mu$ g de plásmido) o *library efficiency* ( $> 1 \times 10^8$  transformantes/ $\mu$ g de plásmido), para lograr buenas eficiencias de transformación. Esta cepa posee mutaciones tanto en la recombinasa *recA1* como en la endonucleasa *endA1*, que mejoran la estabilidad interna del material genético y evitan la degradación del plásmido a insertar, respectivamente. De esta forma podemos amplificar de forma más eficiente la construcción de interés.

Para la producción de CCDC28B recombinante, utilizamos la cepa Rosetta<sup>TM</sup>(DE3) pLysS. Esta cepa está diseñada para la expresión de proteínas de eucariotas utilizando codones raros que normalmente se utilizan poco en *E. coli*. Particularmente presenta sobre-expresados los ARNt: AGG, AGA, AUA, CUA, CCC y GGA. Por otro lado, también expresa la lisozima del bacteriófago T7, que posee la actividad natural de unirse e inhibir a la ARN polimerasa T7, reduciendo la expresión basal (*leaking*) de la proteína de interés en ausencia de inductor (típicamente la expresión

CEPA	CARACTERISTICA	OBJETIVO	RESITENCIA	PROVEDOR
DH5 $\alpha$	Gran eficiencia de incorporación de plásmidos. Generación de insertos estables.	Amplificación de diferentes construcciones.	-	Invitrogen
BL21 (DE3)	Cepa deficiente en proteasa OmpT.	Producción de proteína a partir del vector <i>pDuet</i>	-	Stratagene
Rosetta <sup>TM</sup> (DE3) pLysS	Expresa seis ARNt raros, facilitando la expresión de codones no comunes en <i>E. coli</i> . Reprime la expresión basan de proteínas recombinante por lisozima.	Producción de CCDC28B <i>wild type</i> recombinante	Cam	Novagen
SHuffle <sup>®</sup>	Promueve la formación de puentes disulfuro en el citoplasma, expresando constitutivamente la isomerasa DsbC, que también puede actuar como chaperona para el correcto plegado de la proteína a producir.	Producción de construcciones clonadas en el vector <i>pDuet</i>	-	New England BioLabs

Tabla 4. Cepas bacterianas utilizadas, todas ellas correspondientes a variantes de *E. coli*.

de la ARN polimerasa T7 está bajo control de un promotor inducible por isopropil- $\beta$ -tiogalactósido, que cuando se agrega al cultivo, resulta en un aumento en la cantidad de ARN polimerasa T7, por encima del nivel inhibible por la lisozima T7).

Finalmente, para la expresión de proteínas a partir de las construcciones generadas en el vector pACYCDuet se utilizaron las cepas BL21(DE3) o SHuffle® dado que la cepa Rosetta™ presenta resistencia a cloramfenicol, la misma que brinda el plásmido. La cepa SHuffle® tiene los genes codificantes de la glutaredoxina reductasa y de la tiorredoxina reductasa deletados ( $\Delta$ gor $\Delta$ trxB), lo que favorece la formación de puentes disulfuro en el citoplasma de estas bacterias. Además, expresa una versión de la isomerasa periplásmica formadora de puentes disulfuro (DsbC). Estos motivos hacen que SHuffle® sea una cepa muy utilizada para expresar proteínas recombinantes en las que se esperan puentes disulfuro y que pueden dar dificultades para su producción soluble. Por otro lado la cepa BL21(DE3) es derivada de la cepa B de *E.coli*. Una de las características que la distinguen es la delección de la proteasa OmpT, la cual corta en sitios dibásicos, afectando particularmente la degradación de la T7 ARN polimerasa. Es una de las cepas de *E.coli* más ampliamente utilizada para producir proteínas recombinantes.

### 3.2.C. Medios de cultivo bacteriano

Para todos los pasos que involucraron la transformación de bacterias o la amplificación de los vectores se crecieron las bacterias en medio Luria y Bertani (LB) (10 g/l de Triptona, 5 g/l de Extracto de levadura y 5 g/l de NaCl) y en medio S.O.C. (2 % Triptona, 0,5 % Extracto de levadura, 20 mM Glucosa, 10 mM NaCl, 2,5 mM KCl, 10 mM MgCl<sub>2</sub> y 10 mM MgSO<sub>4</sub>). Para plaquear las bacterias transformadas utilizamos medio LB-Agar (1,5 % de Agar en LB) con el antibiótico correspondiente.

Para la producción de proteínas recombinantes se utilizó como medio de cultivo el medio 2YT (1,6 % Triptona, 1 % Extracto de levadura y 0,5 % NaCl). También se realizó una prueba de producción en medio de autoinducción ZYP 5052 (1 % triptona, 0,5 % extracto de levadura, 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 50 mM KH<sub>2</sub>PO<sub>4</sub>, 25 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0,5 % Glicerol, 0,05 % Glucosa y 0,2 % Lactosa).

### 3.2.D. Transformación de bacterias

Para la transformación de las cepas de *E.coli* DH5 $\alpha$ , BL21(DE3) y Rosetta™ (DE3)pLysS se utilizó el método de choque térmico. Los protocolos para las tres cepas son prácticamente idénticos y consisten en:

- Alicuotar 50  $\mu$ l de células competentes.
- Agregar de 1 a 5  $\mu$ l (1-10 ng) de ADN.
- Mezclar suavemente e incubar en hielo por 30 minutos.

- Choque térmico a 42°C por 20 segundos (DH5 $\alpha$ ), 45 segundos (BL21), o 60 segundos (Rosetta).
- Incubar nuevamente en hielo por dos minutos.
- Agregar 950  $\mu$ l de medio S.O.C. e incubar a 37°C por una hora con agitación.
- Plaquear en LB-Agar con el antibiótico correspondiente.
- Incubar las placas ON a 37°C.

Por otro lado, se utilizaron bacterias electrocompetentes de la cepa SHuffle®T7, por lo que el método de transformación utilizado fue la electroporación. El protocolo utilizado fue el siguiente:

- Alicuotar 50  $\mu$ l de células competentes.
- Agregar de 1 a 5  $\mu$ l (1-10 ng) de ADN.
- Mezclar suavemente y pasar a la celda de electroporación.
- Realizar un pulso de 500 V y 8 mseg.
- Agregar 700  $\mu$ l de medio S.O.C. e incubar a 37°C por una hora con agitación.
- Plaquear en LB-Agar con el antibiótico correspondiente.
- Incubar las placas ON a 37°C.

### 3.2.E. Inducción de la expresión y producción de proteínas recombinantes

En su gran mayoría, la producción de proteína recombinante correspondió a CCDC28B a partir de pET21d utilizando la cepa Rosetta™(DE3)pLysS. A continuación se detalla el protocolo utilizado para dicha producción. Luego del protocolo general se detallan variantes utilizadas para mejorar la eficiencia de la producción o utilizadas para otras proteínas donde se cambió la cepa de *E.coli* y el vector de expresión.

El protocolo utilizado para la producción de CCDC28B WT recombinante a partir de pET21d en Rosetta™(DE3)pLysS fue el siguiente:

- Primero se realizó un precultivo ON a 37°C con agitación de 50 ml de LB + Ampicilina - Cloramfenicol, a partir de un stock de glicerol de las bacterias transformadas con pET21d - CCDC28B WT.
- Se inocularon 10 ml de ese precultivo por litro de medio 2YT + Ampicilina - Cloramfenicol.
- Se incubó el cultivo inoculado a 37°C con agitación hasta llegar a una densidad óptica (DO) a 600 nm de 1 (aproximadamente 4 horas).
- Luego de alcanzada esa DO, se incubó el cultivo 15 minutos a 20°C con agitación.
- Se indujo la expresión de la proteína con 0,1 mM o 1 mM de IPTG y se incubó el mismo a 20°C ON o a 37°C por 3 horas.

Como se observa en el protocolo, la inducción de la expresión de la proteína se realiza agregando IPTG (Isopropil- $\beta$ -D-1-tiogalactopiranosido) al medio. Este es un análogo no

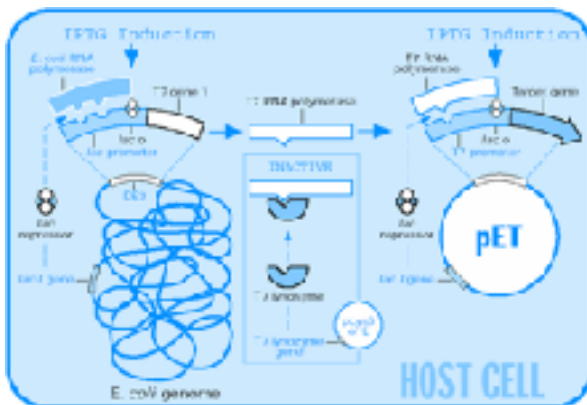


hidrolizable de la alolactosa, y por esto se utiliza para inducir de forma más eficiente la expresión de genes regulados por el operón *lac*. En ellos, la expresión del gen de interés esta reprimida por la unión del represor *lac* (*lacI*) al operador *lac* (evitando así que se una la polimerasa correspondiente). Para activar la expresión se necesita que el represor *lacI* se una a lactosa, permitiendo así que la polimerasa pueda unirse al promotor.

Particularmente en los sistemas pET, como se observa en la Figura 12, el gen de la proteína de interés está regulado por un promotor T7 (por lo que se necesita la transcripción a través de la ARN polimerasa del bacteriófago T7 -ARNPol T7-). Los pET, por lo tanto deben utilizarse con bacterias DE3, que reciben este nombre debido a que tienen insertado en su genoma el gen de la ARNPol T7, para expresarse bajo la regulación del promotor y operador *lac*. Allí el represor *lac* se une al operador *lac* y no permite expresión de ARNPol T7 (excepto una cantidad basal, que se inhibe por acción de la lisozima de T7 que se expresa desde el plásmido pLys en el caso de las Rosetta™). Cuando se agrega lactosa, o IPTG, este se une al represor y la ARNPol T7 se puede expresar. Esa misma ARNPol T7 es la que va a expresar el gen de interés clonado en el pET, siempre y cuando el IPTG siga uniéndose al represor *lacI* para que este no se una al operador. Vemos por lo tanto que un sistema pET/*E.coli*DE3 está doblemente regulado por el operón *lac*, lo que hace fundamental la acción del IPTG para liberar las dos vías de expresión.

Vale la pena destacar que para elegir las condiciones óptimas de inducción se realizó un control probando diferentes temperaturas, tiempos y concentraciones de IPTG. El mismo está descrito en la sección Resultados. Cuando expresamos construcciones clonadas en pET21d el protocolo utilizó como condiciones de inducción una concentración de 0,1 mM de IPTG incubando a 20°C ON, mientras que cuando la expresión se realizó a partir del vector pQE80, se utilizó 0,1 mM de IPTG incubando a 37°C por 3 horas.

Otra variante que eventualmente se introdujo en el protocolo de expresión fue el cambio de medio de cultivo, especialmente para probar la eficiencia del medio de autoinducción ZYP-5052. Este medio puede ser utilizado para inducir la expresión de proteínas cuya transcripción este regulada por el promotor T7/*lac*. Dicho medio activa la expresión de la proteína recombinante de interés a través del crecimiento de las bacterias en dos fases, que utilizan como fuentes de



**Figura 12. Sistema de regulación de la expresión de proteínas recombinantes en el sistema pET.**

En un sistema *E.coli*DE3/pET vemos que hay una doble regulación por parte del operón *lac*. Tanto el gen de la proteína de interés (clonado en pET), como el gen de la ARNPol T7 de las bacterias DE3, tienen en su promotor un operador *lac* que evita que dichos genes se expresen debido a la unión del represor *lacI*. Existe una pequeña expresión basal de la ARNPol T7 en las bacterias DE3 que puede saldarse con la inactivación de la ARNPol T7 que se exprese con la lisozima T7 expresada desde un plásmido pLys. Para liberar la expresión del sistema, se induce con IPTG (análogo de la alolactosa) que se une al represor y gatilla el sistema.

Imagen tomada de la edición 11 del manual del sistema pET, de Novagen.

carbono la glucosa y la lactosa respectivamente. En la primera fase, se genera un cultivo bacteriano de alta densidad, en el cual las bacterias crecen utilizando la glucosa como fuente de carbono. En esta etapa la lactosa no es internalizada por las bacterias y por ende el represor *lac* se une al operador impidiendo la transcripción del gen de interés. En la segunda fase de crecimiento, cuando la glucosa ya no es abundante, las bacterias comienzan a internalizar la lactosa del medio, la cual se une al represor activando de esta forma el promotor *lac* e induciendo la transcripción del gen de interés. Por lo tanto, el protocolo en sí no varía significativamente cambiando simplemente la forma en la cual se induce la expresión de la proteína. Al alcanzar una DO de 1 luego de la incubación a 37°C, en lugar de agregarle IPTG, simplemente se pasa a incubar a 20°C y allí una vez consumida la glucosa del medio, comienza a inducirse la expresión de la proteína de interés (ver Resultados).

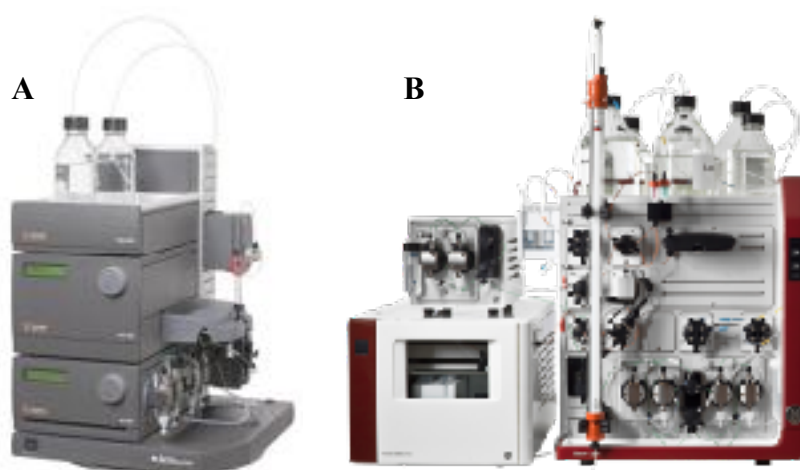
Durante el proceso de producción de la proteína de interés, se toman distintas fracciones para corroborar la expresión de la proteína. En primer lugar se obtiene la fracción sin inducir, que corresponde a una alícuota del cultivo una vez que alcanzó la DO de 1 pero antes de agregar el IPTG. Luego, se obtiene la fracción inducida, que corresponde a una alícuota del cultivo una vez culminada la inducción. Finalmente se obtiene un extracto total del cultivo inducido a partir del cual se separan las fracciones soluble e insoluble de acuerdo al siguiente protocolo:

- Se centrifugó el cultivo inducido a 4000 g por 15 minutos a 4°C.
- Se descartó el sobrenadante y se resuspendió el *pellet* en dH<sub>2</sub>O.
- Se centrifugó nuevamente a 4000 g por 15 minutos a 4°C.
- Se descartó el sobrenadante y se pesó el *pellet*.
- Se resuspendió el *pellet* en 5 ml de buffer A (IMAC) por gramo.
- Se le agregó inhibidores de proteasas (cocktail EDTA-free, Sigma) y lisozima 1mg/ml.
- Se congeló la suspensión a -80°C.
- Se descongeló la suspensión y se sonicó utilizando un sonicador Branson Ultrasonics Sonifier™ S-450 con el siguiente programa: un pulso intermitente de 1 minuto (2 segundos encendido y 1 segundo apagado) con una amplitud de intensidad del 40 %. Este pulso se repitió cuatro veces.
- Luego del sonicado se centrifugó por 45 minutos a 15000 g a 4°C.
- De esa centrifugación el sobrenadante corresponde a la fracción soluble del extracto total y el *pellet* a la fracción insoluble (para su análisis por SDS-PAGE una porción de la fracción insoluble se resuspendió en PBS y buffer de carga).

### 3.3. Purificación de proteínas recombinantes

#### 3.3.A. Sistemas de purificación de proteínas

Para todas las cromatografías utilizadas en la purificación y caracterización de la proteína recombinante se utilizaron dos sistemas: el ÄKTApurifier10 y el ÄKTApure, ambos de GE™ Healthcare Life Sciences (Figura 13). Los dos sistemas de purificación de proteínas están diseñados para ser utilizados a escala de laboratorio y ambos permiten la realización de todo tipo de cromatografías líquidas. Como todos los sistemas ÄKTA, utilizan el software de control UNICORN™ a través del cual se pueden automatizar muchos de los pasos de la cromatografía, entre ellos la inyección de la muestra y los gradientes de buffer, y también controlar algunos parámetros de la corrida como pH y conductividad. Ambos equipos presentan colectores diseñados para la utilización de casetes tanto de 24, 48 o 96 pocillos, como también tubos de tipo *Falcon* de 50 ml. Finalmente en cuanto a los detectores, el ÄKTApure posee tres detectores que se pueden utilizar simultáneamente dentro del rango de 190 a 700 nm. Por otro lado el ÄKTApurifier10 también tiene tres detectores pero estos pueden ser utilizados únicamente para las siguientes longitudes de onda fijas: 280 nm, 214 nm y 254 nm.



**Figura 13. Sistemas de purificación de proteínas utilizados durante este proyecto. Ambos fabricados por GE™ Healthcare Life Sciences.**

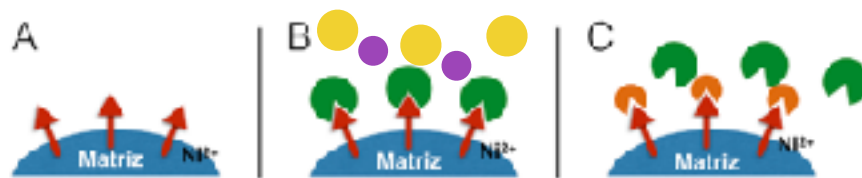
**A-** ÄKTApurifier10, ya discontinuado del mercado, pero aún en funcionamiento en el Laboratorio de la Unidad de Cristalografía de Proteínas del IPMon.

**B-** ÄKTApure, nuevo modelo diseñado para substituir el anterior. Adquirido hace dos años por el laboratorio de la Unidad de Proteínas Recombinantes.

#### 3.3.B. Cromatografía de afinidad a metales inmovilizados

Uno de los pasos utilizados para la purificación de proteínas recombinantes consistió en la realización de una cromatografía de afinidad a metales inmovilizados (IMAC, del inglés: *immobilized metal affinity chromatography*). La cromatografía de afinidad separa las proteínas en base a una interacción reversible entre la proteína de interés y un ligando específico unido a la matriz cromatográfica. En particular, la IMAC se basa en la interacción fuerte entre algunos aminoácidos de la superficie de las proteínas, especialmente repetidos cortos de histidinas en serie, que coordinan metales divalentes como  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$  o  $\text{Co}^{2+}$ , inmovilizados a la matriz.

En el marco de esta tesis, todas las proteínas recombinantes expresadas fueron producidas con una etiqueta de 6-His. Precisamente debido a la expresión de esta etiqueta es que utilizamos



**Figura 14. Etapas de una cromatografía de afinidad a iones metálicos inmobilizados (IMAC).**

**A-** Primera etapa, de equilibrado. La matriz de la columna IMAC, con los iones  $\text{Ni}^{2+}$  disponibles, es equilibrado con el buffer de carga.

**B-** Segunda y tercera etapa, de inyección de la muestra y posterior lavado. Aquí, el material de interés (verde) se une a los iones  $\text{Ni}^{2+}$  y posteriormente se realiza el lavado del material no unido (violeta y amarillo).

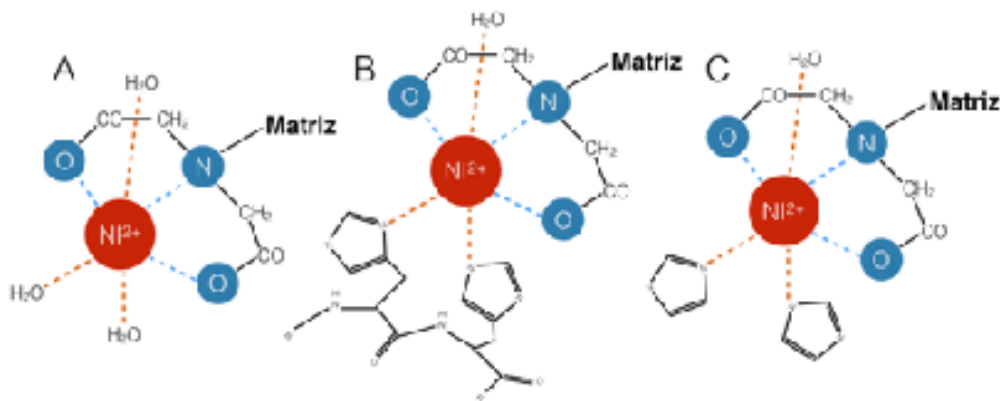
**C-** Etapa final, elución del material de interés (verde). A través del aumento de la concentración de imidazol (naranja) en el buffer y la competencia del mismo por el  $\text{Ni}^{2+}$ , se libera la muestra de interés.

este primer paso de purificación, ya que se sabe que una serie de seis histidinas en un bucle no estructurado ofrecerá una geometría óptima para establecer quelatos con cationes metálicos divalentes. Para la realización de este tipo de corridas cromatográficas utilizamos las columnas comerciales de GE<sup>TM</sup> Healthcare Life Sciences: HisTrap<sup>TM</sup> HP de 1ml y 5ml. Estas columnas pre-empaquetadas utilizan  $\text{Ni}^{2+}$  como ligando unido a una matriz de agarosa. Si bien la afinidad por la cadena lateral de las histidinas (un anillo imidazol) es alta, se sabe que también las cisteínas y los triptofanos pueden coordinar  $\text{Ni}^{2+}$ .

El proceso de purificación por IMAC puede ser dividido en tres grandes pasos (Figura 14): primero la columna debe ser equilibrada en el buffer de carga que se va a utilizar, para recién después inyectar la muestra en la columna (el buffer de carga genera condiciones que favorecen la unión específica al  $\text{Ni}^{2+}$ ); luego de inyectada la muestra, se lava la columna en esas mismas condiciones para eliminar todo el material no unido a la matriz; y finalmente en el último paso, el material unido a la matriz es recuperado/eluído cambiando las condiciones por otras que favorecen el despegado de la matriz. Normalmente, esta elución se puede desarrollar de dos formas: la específica, en la cual se usa un ligando que compite con nuestra proteína de interés por la unión a la matriz, o la no específica, en la que se cambian las condiciones en las que se da la interacción con la matriz, modificando por ejemplo el pH. En nuestro caso se utilizó una elución específica, aumentando la concentración de imidazol en el medio. El imidazol compite fuertemente con las cadenas laterales de las His por la unión a  $\text{Ni}^{2+}$  (Figura 15). Al aumentar la concentración de imidazol en el buffer, este compite por la unión a  $\text{Ni}^{2+}$  desplazando a las histidinas de la proteína y por ende liberándola. En todo el proceso, desde la pre-equilibración con buffer, es importante que la fuerza iónica sea alta (pues de otro modo la interacción con los cationes metálicos se podría dar por uniones inespecíficas puramente electrostáticas), y que el pH al momento de sembrar sea alcalino, para evitar que las histidinas tengan carga positiva (pues el grupo imidazol tiene un pK próximo a 6) lo que disminuiría su afinidad por el metal.

Buffers utilizados para la IMAC:

- Buffer A: 50 mM  $\text{Na}_2\text{HPO}_4$  pH 7,6; 0,5 M NaCl y 20 mM imidazol
- Buffer B: 50 mM  $\text{Na}_2\text{HPO}_4$  pH 7,6; 0,5 M NaCl y 0,5 M imidazol



**Figura 15. Estructura de la unión Matriz-Ni<sup>2+</sup>-ligando en la columna de IMAC.**

**A-** Matriz-Ni<sup>2+</sup> con los iones libres.

**B-** Matriz-Ni<sup>2+</sup> en su interacción con dos residuos de histidina consecutivos (recordar que la etiqueta utilizada en el trabajo es de seis residuos, 6-His).

**C-** Matriz-Ni<sup>2+</sup> en su interacción con dos moléculas de imidazol, donde se observa la similitud de secuencias con el grupo funcional de la histidina.

El protocolo que se utilizó para realizar la IMAC tuvo algunas variantes a lo largo del proyecto pero consta de los siguientes pasos:

- Equilibrado de la columna en 10 volúmenes de buffer A a 4°C.
- Cargado de la muestra en la columna a 4°C utilizando una bomba peristáltica, colectando la fracción no unida a la columna (N.U.).
- Conectar la columna ya cargada al equipo, y luego se realiza un primer paso de lavado con buffer A hasta que la DO a 280 nm sea menor a 10 mU.
- Luego de realizado el lavado comienza el protocolo propiamente dicho de la corrida, con un primer paso de lavado en que se aumenta la concentración de imidazol para eluir los contaminantes que interaccionan más débilmente con la matriz (4 % o 12 % de buffer B, 40 mM u 80 mM de imidazol, respectivamente).
- Para la elución de la proteína de interés, se utiliza un gradiente hasta 75 % de buffer B en 25 volúmenes de columna o un escalón directamente a 75 % de Buffer B (395 mM de imidazol).
- Finalmente, se usa un escalón a 100 % de B para asegurarse de que no quede nada unido a la columna.
- Antes de almacenar la columna se debe lavar con 10 volúmenes de H<sub>2</sub>O y 10 volúmenes EtOH 20 %.

Previo a la inyección en la columna de IMAC se realiza un filtrado de la muestra a través de un filtro de al menos 0,45 μm para evitar el tapado de la columna con impurezas o agregados de proteína.

### 3.3.C. Cromatografía de intercambio aniónico

Como segundo paso de la purificación se realizó una cromatografía de intercambio aniónico (AEC, del inglés: *anion exchange chromatography*) con el material purificado previamente a partir de la IMAC.

Las cromatografías de intercambio iónico separan biomoléculas en base a diferencias en su carga electrostática neta superficial. Las moléculas en general pueden interaccionar de diferentes formas con la matriz cromatográfica cargada según presenten diferencias en su carga total, densidad de carga o distribución de carga superficial. Las proteínas particularmente, por presentar propiedades anfóteras, es decir que se pueden comportar como un ácido o una base según cambie su carga superficial debido al pH del entorno, son un muy buen sustrato para la realización de esta técnica.

En las cromatografías de intercambio iónico, como muestra la Figura 16, se establecen interacciones reversibles entre las moléculas cargadas y la matriz cromatográfica cargada con polaridad opuesta. Estas interacciones pueden ser reguladas favoreciendo la unión o la elución de las moléculas de interés. Esta regulación se lleva a cabo teniendo en cuenta que una proteína determinada a un pH equivalente a su punto isoeléctrico (pI) no tiene carga neta, y por ende no interaccionará con la matriz cargada o lo hará débilmente (recordando que aún cuando la carga neta total sea nula, en superficie podría llegar a haber un potencial electrostático neto expuesto). Sin embargo, a un pH mayor que su pI, la proteína se cargará negativamente transformándose en un anión, por lo que se unirá a una matriz cargada positivamente (intercambiador aniónico). Mientras que a un pH menor a su pI, la carga de la proteína será positiva y se podrá unir a una matriz negativa (intercambiador catiónico) compitiendo con otros cationes del soluto. La elución en este tipo de corridas cromatográficas, se puede realizar cambiando la fuerza iónica del medio (aumentando la concentración de sales) o modificando el pH del mismo, siendo la primera la forma más utilizada. Al aumentar la fuerza iónica aumentará la concentración de los iones de la sal presente (por ejemplo  $\text{Na}^+$  y  $\text{Cl}^-$ ). Si la cromatografía es por intercambio de aniones, los iones  $\text{Cl}^-$  competirán con las proteínas aniónicas ya unidas a la columna, liberando a estas últimas en un orden que depende de sus afinidades por la matriz (las menos cargadas negativamente al pH



**Figura 16. Etapas de una cromatografía de intercambio aniónico.**

A- Primera etapa, de equilibrado. Aquí las cargas positivas superficiales de la matriz de la columna se equilibran con los aniones presentes en el buffer de inyección de la muestra.

B- Segunda etapa, de inyección de la muestra. Aquí, el material de interés (verde) cargado negativamente al pH del buffer de inyección se une a la matriz.

C- Etapa final, de elución del material de interés (verde). A través del aumento de la concentración de sal, aumentamos los iones  $\text{Cl}^-$  del medio, compitiendo con la unión de la matriz positiva por la proteína de interés, negativa, y liberando la misma.

utilizado y por ende unidas con menor afinidad, eluirán primero). También se puede lograr la elución diferencial cambiando progresivamente el pH del medio, pues ello cambiará la carga neta superficial de las proteínas de acuerdo a sus distintos puntos isoeléctricos.

En este trabajo, eluímos nuestra proteína de interés modificando la fuerza iónica del medio trabajando a pH fijo. CCDC28B tiene un pI teórico de 5,23 (ExPASy: [http://web.expasy.org/cgi-bin/compute\\_pi/pi\\_tool1?Q9BUN5@1-200@average](http://web.expasy.org/cgi-bin/compute_pi/pi_tool1?Q9BUN5@1-200@average)) por lo que nos resultó más práctico trabajar a un pH mayor a su pI, utilizando como intercambiador aniónico una columna MonoQ™ HR 5/5 (GE™ Healthcare Life Sciences) de 1 ml de lecho.

Previo a la realización de la AEC, las muestras deben ser equilibradas en el buffer adecuado, en nuestro caso esto se realizó por diálisis utilizando membranas de celulosa regenerada SnakeSkin® (ThermoFisher) de punto de corte 10 kDa. La diálisis se realizó contra un volumen 20 veces mayor al volumen de muestra, ON a 4°C con agitación. Luego de la diálisis, la muestra se centrifugó a 10000 g por 15 minutos a 4°C y se filtró por 0,22 µm previo a la inyección en la MonoQ.

Buffers utilizados para la MonoQ:

- Buffer A: 20 mM Tris/HCl pH 8, 20 mM NaCl
- Buffer B: 20 mM Tris/HCl pH 8, 1 M NaCl

El protocolo de una AEC, como ilustra la Figura 16, consiste en una primera etapa de equilibrado de la columna, otra de inyección de la muestra y lavado, y finalmente la etapa de elución de la muestra de interés. A continuación se describen los detalles del mismo:

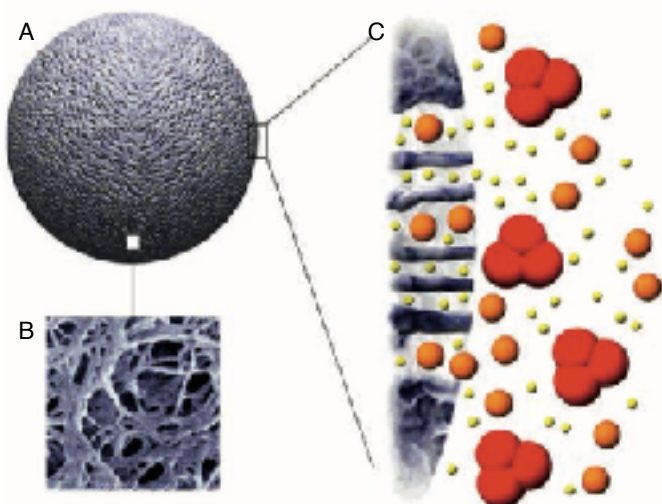
- Equilibrado de la columna en 5 volúmenes de columna de Buffer A, 10 volúmenes de Buffer B y 10 volúmenes de Buffer A, en ese orden.
- Inyección de la muestra y lavado en Buffer A hasta que se estabilice la medida de la DO a 280 nm.
- Finalmente se comienza la elución del material de interés. En nuestro caso, se realizó un primer gradiente hasta 50 % de Buffer B (520 mM NaCl) en 20 volúmenes de columna y luego un escalón a 100 % de Buffer B (1 M NaCl) para que termine de eluir todo lo que quedó unido a la columna.
- Antes de almacenarla, la columna se lava en 10 volúmenes de H<sub>2</sub>O y 10 volúmenes de EtOH 20 %.

### 3.3.D. Cromatografía de exclusión molecular

Como último paso de purificación se realizó una cromatografía de exclusión molecular (SEC, del inglés: *size exclusion chromatography*), o gel filtración.

La cromatografía de exclusión molecular es un método de separación de moléculas de acuerdo a su tamaño (o más precisamente, su radio hidrodinámico, en general directamente relacionado con el peso molecular para proteínas de arquitectura globular). Esta técnica, como se ilustra en la Figura 17, separa las moléculas de acuerdo a su tamaño mientras estas atraviesan la matriz empaquetada en el columna, que tiene poros de diferentes tamaños. Los componentes de la muestra inyectada atraviesan la matriz porosa tan profundamente como su tamaño lo permita. Es decir, las moléculas más grandes cuyo tamaño no permite entrar a los poros, van a ser excluidas corriendo por fuera de la matriz, y por lo tanto van a eluir más rápido (en el volumen de exclusión), en comparación con aquellas más pequeñas que sí puedan ingresar en los poros de la columna, incluyéndose en la misma, y recorriendo así una trayectoria mayor. Cuanto más pequeñas son las moléculas incluidas, más se retrasarán eluyendo después, hasta un máximo correspondiente al volumen de inclusión (determinado por el tamaño promedio de los poros de la matriz en cuestión).

Otra diferencia fundamental, con respecto a los otros tipos de cromatografía líquida, es que aquí las moléculas no interaccionan ni se unen con la matriz de la columna. Una consecuencia directa de esto es que el buffer en el que se realiza la gel filtración no altera la resolución de la misma (como si sucede en otros casos). En fin, se conoce que de manera inespecífica ciertas proteínas pueden establecer interacciones iónicas débiles con la matriz inerte, y es por ello que para reducir estos artefactos, las corridas SEC se realizan típicamente con una baja concentración de sal en el buffer de corrida. Por ende, este tipo de cromatografía se puede realizar en diferentes tipos de buffers sin que se altere la separación de las moléculas. Lo que sí es fundamental en SEC es el volumen de la muestra que se siembra, pues si dicho volumen es muy grande ( $> 5\%$ ) en relación al volumen de lecho, se compromete la resolución de la separación buscada, basada exclusivamente en un principio físico a modo de “colador” molecular. Así como los intercambiadores iónicos y la cromatografía de afinidad son principios cromatográficos concentradores (se puede partir de un volumen grande de muestra diluida, la molécula de interés se unirá a la



**Figura 17. Fundamentos de la cromatografía de exclusión molecular (Figura tomada del manual de GE Healthcare Life Sciences titulado Size Exclusion Chromatography).**

**A-** Imagen esquemática de una *bead* (componente básico de la matriz de la cromatografía de exclusión molecular)

**B-** Imagen de microscopía electrónica de los poros de las *beads*.

**C-** Difusión de moléculas de distinto tamaño a través de los poros de cada *bead*. donde se observa que las moléculas más grandes no entran a los poros, mientras que las más chicas lo hacen a los poros más profundos.



columna y será luego eluida concentrándose), SEC es una aproximación que diluye entre 5-8 veces desde el volumen de muestra inicial al del pico de elución de la proteína blanco.

Para las distintas corridas de SEC a lo largo del trabajo se utilizaron dos columnas diferentes: la Superdex 75 10/300, y la Superdex 75 16/600. Las Superdex 75 tienen la capacidad de separar proteínas en el rango de 3 a 70 kDa, y mientras que la 10/300 es una columna analítica, con un volumen de 25 ml en la que se pueden inyectar entre 25 y 500  $\mu$ l de muestra, la 16/600 es una columna preparativa, con un volumen de 120 ml y en la que puede inyectarse hasta 5 ml de muestra.

Buffer utilizado para la SEC:

- 5 mM  $\text{Na}_2\text{HPO}_4$  pH 7,6; 0,5 M NaCl

El protocolo de corrida de la cromatografía de exclusión molecular utilizando cualquiera de las columnas fue el mismo y consistió en:

- Equilibrado con dos volúmenes de columna de buffer.
- Inyectado de la muestra en la columna dejando que circule un volumen de columna por la misma.
- Lavado de la columna en 2 volúmenes de  $\text{H}_2\text{O}$  y EtOH 20 % antes de su almacenamiento.

Previo a la inyección en la columna de gel filtración, y para evitar el tapado de la misma, se realizó un centrifugado de la muestra a 15000 g por 10 minutos, para luego filtrarla utilizando un filtro de 0,22  $\mu$ m.

### 3.3.E. Cuantificación de proteínas

En muchos pasos de la producción y purificación de las proteínas recombinantes, se realizó una cuantificación del material con el objetivo de evaluar los rendimientos de los diferentes procedimientos. En todos los casos, las cuantificaciones se realizaron utilizando la misma técnica: BCA (del inglés: *bicinchoninic acid*, o ácido bicinconínico).

Este método se basa en la reacción de las proteínas con  $\text{Cu}^{2+}$  en medio alcalino para reducirlo a  $\text{Cu}^+$  que luego reaccionará específicamente con el ácido bicinconínico (Wiechelman K.J. et al, 1988). El producto de la reacción entre el BCA y el  $\text{Cu}^+$  es de color púrpura, soluble en agua y con un pico de absorbancia a 562 nm. De esta manera se puede realizar una cuantificación espectrofotométrica de las proteínas de una muestra en solución acuosa.

Para realizar este método utilizamos el kit comercial Pierce™ BCA Protein Assay Kit de ThermoFisher. Utilizamos albúmina de suero bovino (BSA) para realizar una curva de calibración en concentraciones desde 0,2 mg/ml a 1,4 mg/ml (rango de sensibilidad del ensayo). También

preparamos diluciones de la muestra a cuantificar, a fin de que la concentración quede dentro de ese mismo rango.

Luego de realizadas las diluciones de BSA y de la muestra a cuantificar, pipeteamos 10  $\mu$ l de cada una de las muestras en un pocillo de una placa de 96 pocillos. Realizamos una mezcla 50:1 de las soluciones A y B del kit comercial y agregamos 200  $\mu$ l de esta a cada pocillo. Finalmente incubamos 30 minutos a 37°C para después medir la absorbancia a 562 nm.

### 3.3.F. SDS-PAGE

También en distintos pasos del trabajo, se necesitó identificar las proteínas presentes en una mezcla, para lo cual nos apoyamos en el método de electroforesis en gel de poliacrilamida con dodecilsulfato sódico (SDS-PAGE, del inglés: *sodium dodecyl sulfate polyacrylamide gel electrophoresis*).

Este tipo de electroforesis, en una sola dimensión, da información sobre el tamaño de cada una de las proteínas que forman la muestra ya que aquí debido a la acción del SDS las proteínas se desnaturalizan, perdiendo su conformación tridimensional. Por ser un detergente aniónico, el SDS se une a las cargas positivas de las proteínas generando una cubierta de carga negativa a la muestra, la cual es proporcional al tamaño de la misma (cuanto más grande es la proteína se une más SDS, manteniendo constante la relación tamaño/carga). Y por ende al someter el gel a una corriente eléctrica, las proteínas migraran hacia el ánodo en función exclusivamente de su tamaño. De esta manera, las proteínas más pequeñas migran más y las más grandes migran menos. Otra variante que puede determinar el grado de migración de las proteínas es la concentración de acrilamida que tenga el gel, donde mayores concentraciones disminuyen el tamaño de los poros del gel.

Para aumentar la resolución de la separación de proteínas utilizamos un sistema discontinuo, donde la parte superior del gel, donde se siembra la muestra, corresponde a un gel concentrador (5% acrilamida en buffer Tris pH 6,8) y la parte inferior corresponde al gel separador (% de acrilamida variable en buffer Tris pH 8,8). El gel concentrador contiene iones Cl<sup>-</sup> que migran por delante de las proteínas de la muestra, mientras que el buffer en el que se realiza la corrida contiene iones glicina, que migran por detrás, dejando las proteínas atrapadas en una banda muy fina, concentrada las previo a su ingreso al gel separador. Al ingresar al gel separador la glicina se neutraliza (por el mayor pH) y las proteínas se separan en base a su tamaño a través de la red de poros más pequeños que en el gel concentrador. El tamaño de los poros, determinado por la concentración de acrilamida se debe elegir según el tamaño de la proteína de interés que se quiere separar. Generalmente para separar proteínas desnaturalizadas de 60 a 200 kDa se utiliza 8 % de acrilamida, mientras que para proteínas de 16 a 70 kDa se utiliza 10 % y para proteínas de 12 a 45 kDa 12 %.

Generalmente a lo largo de este trabajo utilizamos geles de acrilamida de 12% y se realizaron de acuerdo al siguiente protocolo:

- Preparar aproximadamente 10 ml de mezcla de gel separador por cada gel que se quiera realizar según la siguiente receta:

Preparación de Gel Separador 12% de acrilamida

Componentes	Volumen para una mezcla de 10 ml (en ml)
<b>H<sub>2</sub>O</b>	3,3
<b>30 % Acrilamida/ bisacrilamida</b>	4,0
<b>1,5 M Tris (pH 8,8)</b>	2,5
<b>10 % SDS</b>	0,1
<b>10 % Persulfato de amonio</b>	0,1
<b>TEMED</b>	0,004

- Una vez armado el soporte para los geles, cargar cada uno de ellos con la solución y dejar polimerizar sin moverlos por 15 minutos aproximadamente
- Luego de polimerizado el gel separador, preparar aproximadamente 5 ml de mezcla de gel concentrador por cada gel que se quiera realizar según la siguiente receta:

Preparación de Gel Concentrador 5% de acrilamida

Componentes	Volumen para una mezcla de 5 ml (en ml)
<b>H<sub>2</sub>O</b>	3,4
<b>30 % Acrilamida/ bisacrilamida</b>	0,83
<b>1,0 M Tris (pH 6,8)</b>	0,63
<b>10 % SDS</b>	0,05
<b>10 % Persulfato de amonio</b>	0,05
<b>TEMED</b>	0,005

- Cargar cada gel con la solución e inmediatamente agregar los peines para la formación de los pocillos en el gel y dejar polimerizar por 15 minutos.
- Preparar las muestras a correr agregando a las mismas buffer muestra para llegar a una concentración final 1x (62,5 mM Tris/HCl pH 6,8; 2,5 % SDS; 0,002 % Azul Bromofenol; 5 % β-mercaptoetanol y 10 % glicerol)
- Hervir las muestras a 100°C por 5 minutos
- Centrifugar a máxima velocidad por 5 minutos
- Colocar los geles ya polimerizados en la cuba de electroforesis y llenar la misma con buffer de corrida (25 mM Tris, 192 mM glicina, 0,1 % SDS, pH 8,3-8,6)
- Sembrar las muestras y el marcador de peso molecular en los pocillos

- Conectar los electrodos a la cuba de electroforesis y a la fuente de alimentación y aplicar corriente eléctrica (30 mA por gel).

Una vez terminada la electroforesis se procede a teñir los geles para observar los resultados. A lo largo del trabajo se utilizaron dos métodos para teñir geles de SDS-PAGE: la tinción con azul de Coomassie y la tinción con nitrato de plata.

El azul de Coomassie es un colorante de tipo aniónico y se une inespecíficamente a proteínas. Es no polar y por eso normalmente se usa en una solución 40 % metanol, 7 % ácido acético y 0,025 % azul de Coomassie. Al teñir, las proteínas son simultáneamente fijadas gracias a la acción del ácido acético y teñidas por el colorante. Su límite de detección se aproxima a los 50 ng de proteína. Para realizar la tinción con azul de Coomassie se realizó el siguiente protocolo:

- Transferir los geles de acrilamida a un recipiente con la solución de tinción de azul de Coomassie.
- Incubar con agitación suave a temperatura ambiente durante aproximadamente 20 minutos.
- Retirar la solución de tinción y agregar solución decolorante (40 % etanol y 7 % ácido acético) para eliminar el colorante en exceso.
- Incubar con agitación suave a temperatura ambiente hasta obtener el contraste deseado.
- Retirar la solución decolorante y dejar en solución de almacenamiento (5 % etanol y 7 % ácido acético).

El otro método utilizado para teñir los geles de SDS-PAGE es el nitrato de plata. Esta es una técnica más sensible que el azul de Coomassie, llegando a detectar bandas de hasta 1 ng de proteína. La mecánica de la tinción se basa en la unión de los iones de plata a las proteínas, y luego esos iones de plata pueden ser reducidos, generando un producto visible. Para realizar la tinción con nitrato de plata se realizó el siguiente protocolo:

- Transferir los geles de acrilamida a un recipiente con la solución de fijación (12 % ácido acético; 50 % etanol y 0,05 % formalina).
- Incubar con agitación suave a temperatura ambiente por al menos 1 hora.
- Realizar dos lavados con solución de lavado (50 % etanol) de 5 minutos.
- Realizar dos lavados con ddH<sub>2</sub>O de 1 minuto.
- Incubar con solución de sensibilización (0,02 % tiosulfato de sodio) por no más de 30 segundos.
- Realizar dos lavados con ddH<sub>2</sub>O de 1 minuto.
- Incubar con solución de teñido (0,2 % nitrato de plata y 0,075 % formalina) durante 20 minutos.
- Realizar dos lavados con ddH<sub>2</sub>O de 1 minuto.

- Revelar las proteínas, con solución de revelado (3 % carbonato de sodio; 0,02 % tiosulfato de sodio y 0,05 % formalina), hasta visualizar las bandas. No debería llevar más de 10 minutos.
- Terminar la reacción de revelado con solución de terminación (5 % ácido acético).
- Para almacenar, dejar los geles en solución de almacenamiento (1 % ácido acético).

### 3.3.G. Western Blot

El Western Blot (WB) permite la detección de una proteína particular en una mezcla de diferentes proteínas a través de su unión con anticuerpos específicos.

Durante esta técnica, las proteínas de una mezcla separadas mediante SDS-PAGE, son adsorbidas e inmovilizadas en la superficie de una membrana sintética de PVDF (*polyvinylidene difluoride*, Amersham-GE Healthcare), en las mismas posiciones que ocupaban en el gel. La adsorción se realiza por transferencia de las proteínas del gel a la membrana a través de un campo eléctrico perpendicular al gel y a la membrana. Bajo los mismos principios que en el SDS-PAGE, las proteínas, cargadas negativamente, migran hacia el ánodo (donde se encuentra la membrana). La velocidad de migración es nuevamente proporcional al tamaño de la muestra, por lo que el tiempo de transferencia dependerá del tamaño de la proteína a transferir.

La transferencia entonces se realizó bajo el siguiente protocolo:

- Una vez culminado el SDS-PAGE transferir el gel a un recipiente conteniendo buffer de transferencia (buffer de corrida + 20 % de metanol o etanol).
- Armar el cassette, dentro del recipiente con buffer de transferencia y siempre en inmersión, de la siguiente forma:
- Colocar el soporte plástico del cassette, que quedará sobre el lado del ánodo y la membrana abajo.
- Colocar encima una esponja.
- Colocar dos papeles de filtro.
- Colocar la membrana de PVDF y cortarle una esquina para recordar la orientación.
- Colocar el gel en el mismo sentido en el que fue sembrado para mantener el orden en la membrana y asegurarse que no queden burbujas entre la membrana y el gel.
- Colocar dos papeles de filtro.
- Colocar otra esponja, asegurarse de que no quede ninguna burbuja dentro del cassette y cerrar el mismo.

- Sumergir el cassette en la cuba de transferencia con buffer recordando orientar la parte en la que se encuentra la membrana hacia el ánodo.
- Correr la transferencia a 400 mA con refrigeración durante el tiempo que corresponda.

Una vez realizada la transferencia, las proteínas se encuentran inmovilizadas y expuestas en la superficie de la membrana. Allí se puede realizar una detección rápida de proteínas totales para controlar la eficiencia de la transferencia. Esto se puede realizar con una tinción con Rojo Ponceau, un colorante no permanente que puede por lo tanto ser compatible con la posterior detección a través de anticuerpos. La tinción con rojo Ponceau se realizó bajo el siguiente protocolo:

- Luego de la transferencia, incubar la membrana en una solución de rojo de Ponceau (0,1 % Ponceau S y 1 % ácido acético) durante cinco minutos con agitación.
- Transferir la membrana a la solución de desteñido (1 % ácido acético o agua) y agitar hasta que las bandas sean visibles.
- Descartar la solución de desteñido y lavar en agua varias veces.

Una vez realizada la detección de proteínas total se procede a bloquear la membrana. Ésta tiene gran afinidad por las proteínas y no todos sus sitios están ocupados por la muestra de interés, por lo tanto hay que bloquear los sitios que quedaron libres para evitar que los anticuerpos se unan a la membrana en forma inespecífica. Esto se realiza saturando el sistema con albumina bovina (BSA 5 % en PBS) o leche en polvo (5 % en PBS). Luego de realizar el bloqueo se procede a la incubación con los anticuerpos siguiendo este protocolo:

- Bloqueo la membrana incubando toda la noche en PBS 0,1 % Tween 20 y 5 % leche en polvo.
- Luego del bloqueo lavo varias veces con PBS 0,1 % Tween 20.
- Incubo las membranas con los anticuerpos primarios correspondientes y en las condiciones indicadas en cada caso.
- Lavo varias veces con PBS 0,1 % Tween 20.
- Incubo las membranas con los anticuerpos secundarios correspondientes por 1 hora a temperatura ambiente. Todos los anticuerpos secundarios están acoplados a la peroxidasa de rábano (HRP).
- Lavo varias veces con PBS 0,1 % Tween 20.
- Se procede a la detección de la proteína utilizando el sistema de quimioluminiscencia Pierce ECL Western Blotting Substrate siguiendo las indicaciones del fabricante.

El método de detección utilizado se basa en la reacción de oxidación del luminol, presente en el kit, por parte de la peroxidasa de rábano conjugada al anticuerpo secundario. Esta reacción emite luz, que luego es capturada en una lámina de rayos X.

En la Tabla 5 se detallan todos los anticuerpos utilizados durante este trabajo.

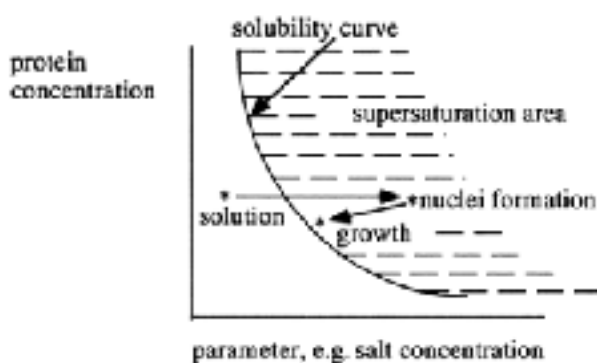
<b>Anticuerpos utilizados</b>	<b>Concentración</b>	<b>Tiempo</b>
Anti-CCDC28B (policlonal de conejo, Genscript)	1/1000	ON a 4°C o 2 hs a TA
Anti-Ran (policlonal de conejo, Bethyl)	1/1000	ON a 4°C
Anti-SIN1 (policlonal de conejo, Bethyl)	1/1000	ON a 4°C
Anti-Flag (monoclonal de ratón, Abcam)	1/800	ON a 4°C
Anti- $\alpha$ tubulina (monoclonal de ratón, Sigma)	1/2000	1 h a TA
Anti-histidina	1/1000	
<b>Conjugados utilizados</b>		
Anti-Ig de conejo conjugado a HRP (Sigma)	1/50000	1 h TA
Anti-Ig de ratón conjugado a HRP (Sigma)	1/20000	1 h TA

**Tabla 5. Anticuerpos utilizados durante este trabajo.**

### 3.4. Ensayos de cristalización

#### 3.4.A. Pruebas de solubilidad

Para que se pueda formar el cristal de una proteína en solución esta debe cambiar de fase, pasando de un estado estable e insaturado (donde la proteína está disuelta y desordenada en la solución) a un estado metaestable y supersaturado (donde las moléculas aún están solubles por una cuestión cinética, pero en equilibrio saldrán de solución y precipitarán, ya sea ordenadamente formando un cristal, o bien desordenadamente formando un precipitado amorfo). La concentración de proteína a la que ocurre esta transición de fases depende de cada proteína, así como de las condiciones en las que se haya disuelta. La Figura 18 muestra la curva de solubilidad típica de una proteína donde se ilustra el proceso de crecimiento de un cristal. Idealmente lo que se busca para generar un cristal es conseguir una condición dentro de la región de supersaturación, lo más cercano a la curva de solubilidad, porque si el material cristalino se genera a un alto grado de supersaturación, se obtendrían muchos cristales chicos. Pero esto ocurre en muy raras ocasiones, ya que normalmente cuando nos excedemos en la supersaturación, la muestra precipita. Por lo tanto, resulta práctico realizar pruebas de solubilidad que permitan estimar el punto de solubilidad de nuestra proteína recombinante: a qué concentración de proteína ocurre la precipitación en condiciones típicas de ensayos de cristalización (por ejemplo usando sulfato de amonio o PEG 4000 como agentes precipitantes). Luego podremos lanzar rastreos de condiciones de cristalización usando la proteína a una concentración por debajo pero cercana a su punto de solubilidad: nos aseguraremos de cruzar el punto de solubilidad en muchos casos (de no hacerlo nunca conseguiremos cristales), pero tampoco partir de una concentración tan alta que todas las condiciones generen precipitados amorfos rápidamente.



**Figura 18. Curva de solubilidad típica de una proteína (Figura tomada de *Principles of Protein X-Ray Crystallography* de Jan Drenth).**

En la imagen, se observa la curva típica de solubilidad de una proteína en función de algún parámetro, que puede ser la concentración de precipitante, por ejemplo de sal. Aquí podemos diferenciar las distintas fases. Como ser la fase soluble o la región supersaturada. La región óptima para la formación de cristales corresponde a la región dentro del área de supersaturación más cercana a la curva de solubilidad, para evitar la nucleación o precipitación de la muestra por exceso de supersaturación.

Una vez obtenida la proteína pura, se realizó una prueba de solubilidad que consistió en la incubación de la muestra a una determinada concentración en soluciones buffer con diferentes concentraciones de sulfato de amonio y polietilenglicol (PEG) 4000. Los buffers utilizados fueron los siguientes:

- Buffer A1: 0,1 M Tris/HCl pH 8,5; 2 M  $(\text{NH}_4)_2\text{SO}_4$
- Buffer B1: 0,1 M Tris/HCl pH 8,5; 1 M  $(\text{NH}_4)_2\text{SO}_4$
- Buffer A2: 0,1 M Tris/HCl pH 8,5; 0,2 M  $\text{MgCl}_2$  y 30 % PEG 4000
- Buffer B2: 0,1 M Tris/HCl pH 8,5; 0,2 M  $\text{MgCl}_2$  y 15 % PEG 4000



Se eligieron como precipitantes en la prueba de solubilidad el  $(\text{NH}_4)_2\text{SO}_4$  y el PEG 4000, por ser dos de los más utilizados dentro de las sales y los polímeros, respectivamente. Estos dos compuestos favorecen la supersaturación de la muestra mediante mecanismos diferentes y es por eso que se eligió uno de cada uno para probar como actuaban sobre CCDC28B. En primer lugar las sales, en nuestro caso el sulfato de amonio, modifican la interacción del solvente con la proteína. Como agente precipitante, las sales actúan de dos formas en paralelo, por un lado sus iones modifican las cargas eléctricas en la superficie de las proteínas, debilitando de esta forma las fuerzas repulsivas entre sus moléculas (efecto de *salting-in* o aumento de solubilidad a concentraciones de sal bajas). Pero además, alrededor de los iones salinos se genera una capa de hidratación que evita que el agua esté disponible para la proteína, deshidratándola (efecto de *salting-out* induciendo la precipitación de proteínas a altas concentraciones de sal). De ambas formas se favorece la supersaturación (McPherson y Gavira, 2013; Cheng et al, 2005). En particular dentro de las sales se eligió el sulfato de amonio ya que es muy soluble y no es dañino (caotrópico) para la mayoría de las proteínas, incluso en una alta concentración.

Por otro lado, el uso de polímeros como el PEG, disminuye el volumen de solvente disponible para que la proteína se mantenga solvatada en solución y por ende se favorece el acercamiento de las moléculas de proteína entre sí. El gran tamaño de los polímeros y su conformación variable y móvil, hacen que ocupen en solución un mayor espacio del que deberían, por lo que reducen la disponibilidad de solvente para la proteína. Como consecuencia, las moléculas de proteína se juntan pudiendo llegar a formar un estado sólido y ordenado como un cristal (McPherson y Gavira, 2013; McPherson A., 1976). El PEG en particular tiene como ventaja sobre otros polímeros que es hidrofílico, pudiendo actuar también deshidratando la muestra como una sal.

La prueba de solubilidad utiliza el método de la gota colgada (se explica en la sección Prueba Condiciones de Cristalización) y el protocolo que se siguió para la misma es el siguiente:

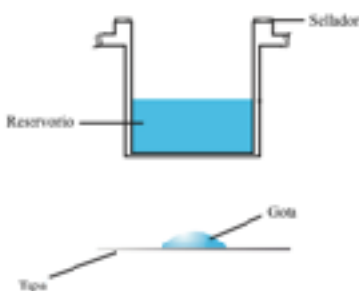
- Pipetear en cada uno de los reservorios (pocillos) 1 ml de cada uno de los buffers.
- En cuatro tapas diferentes pipetear 2  $\mu\text{l}$  de proteína.
- Agregar sobre las gotas de proteína 2  $\mu\text{l}$  de buffer (uno por cada una).
- Rápidamente invertir la tapa sobre el reservorio conteniendo el buffer correspondiente y sellar la tapa.
- Observar al otro día (o eventualmente a las 48 horas) el estado de cada una de las gotas en la lupa. Si hay formación de precipitado en todas las gotas, se baja la concentración de proteína del *stock* y se repite la prueba. Por el contrario, si no hay formación de precipitado con ninguna de las concentraciones de precipitante, es señal que se debe aumentar la concentración de proteína del *stock* madre.

### 3.4.B. Búsqueda de condiciones de cristalización

Luego de la prueba de solubilidad, se eligió la concentración de proteína óptima (en nuestro caso 5 mg/ml). El siguiente paso fue realizar una primera búsqueda de condiciones de cristalización. Allí nos apoyamos en la tecnología robótica con la que cuenta la Unidad de Cristalografía de Proteínas del IPMon, particularmente el robot Honeybee963® (Digilab) y la utilización de kits comerciales de condiciones de cristalización para ampliar lo más posible el rango de rastreo. Se utilizó como método de cristalización la difusión de vapor al equilibrio en gota sentada. Luego para las optimizaciones con gotas de mayor volumen (manuales o usando el robot Alchemist®, Rigaku) el mismo principio fue usado pero en gota colgante.

El método de difusión de vapor al equilibrio se ilustra en la Figura 19 y consiste en aproximarse a la supersaturación de la proteína contenida en la gota, por evaporación del agua. Las gotas se preparan mezclando proteína y buffer de reservorio (típicamente en proporción 1:1), con lo que la concentración de solutos en la gota es menor a la del reservorio, y por ende mayor la presión de vapor (una de las propiedades coligativas de soluciones). El sistema está herméticamente cerrado (la tapa donde se encuentra la gota está sellada sobre el volumen de reservorio), y la presión de vapor de la gota tenderá a buscar ser igual al del reservorio, para lo cual el agua contenida en la gota difunde hacia el reservorio hasta alcanzar el equilibrio de presiones de vapor. Durante este proceso, la concentración de la proteína también aumenta en la gota. Este proceso ocurre lentamente lo que favorece el proceso de cristalización.

El robot Honeybee963® automatiza y miniaturiza el pipeteado de nanogotas de proteína. Para hacerlo usa una variante del método descrito, usando gotas sentadas en lugar de colgantes, aunque el principio de difusión de vapor es exactamente el mismo. Así se probaron 384 condiciones de cristalización, correspondientes a los kits comerciales: PEGs Suite (96 condiciones), AmSO<sub>4</sub> Suite (96 condiciones) y pHClear Suite (96 condiciones) de Qiagen, así como el CrystalScreen (96 condiciones) de Hampton Research. En todos los casos los kits ya vienen con las placas de 96 pocillos prontas, con cada una de las condiciones diferentes en los reservorios. Las gotas fueron producidas utilizando el robot Honeybee963®, el cual tiene un brazo de 96 agujas que permite el dispensado rápido de las soluciones que se encuentran en los reservorios y tres agujas synQUAD®



**Figura 19. Método de difusión de vapor por gota colgante (Figura modificada del manual *Crystallization Screening* de Hampton Research).**

En la imagen se observa un pocillo correspondiente a una condición de cristalización mediante el método de la gota colgante. A la izquierda se observa como se diseña el experimento, primero cubriendo los bordes del pocillo con sellador, para que una vez colocada la tapa, esta quede fija. Luego se le agrega el buffer a testear en el reservorio y finalmente se produce la gota en la tapa. Allí se ponen volúmenes iguales de proteína y buffer del reservorio, para luego dar vuelta la tapa sobre el pocillo y verificar que quede bien sellado.

que permiten dispensar volúmenes de hasta tres proteínas distintas llegando a los 100 nl. En todos los casos las gotas se forman con volúmenes iguales de proteína y buffer del reservorio.

Las placas se observaron cada dos días para buscar la presencia de material cristalino.

### 3.4.C. Optimización de las condiciones de cristalización

Luego de encontrar una condición en la que la proteína de interés forme sólidos cristalinos, se debe optimizar esa condición para conseguir el mejor cristal posible, es decir de buen tamaño y que difracte los rayos X.

Para esto se escaló el experimento de cristalización en la condición favorable encontrada previamente, pasando de placas de 96 pocillos a placas de 24 pocillos, nuevamente utilizando el método de difusión de vapor, ahora en gota colgante. Las placas se diseñaron manualmente, modificando sutilmente la condición en la que se vio material cristalino, por ejemplo, cambiando el pH de la solución o la concentración de agente precipitante. Para generar las placas se utilizó el robot Alchemist DT®, este es un robot de mesada, de manejo de líquidos que dispensa en diferentes tipos de placas volúmenes entre 1  $\mu$ l y 10 ml partiendo de 26 soluciones madre, previamente preparadas. El robot funciona acompañado del *software* CrystalTrak™, en el cual se ingresan cuales son las 26 soluciones madre con las que se cargó el robot y se diseña la placa, para que después se calcule automáticamente la receta para cada uno de los reservorios y el robot pueda proceder.

Una vez armada la placa se generaron las gotas manualmente para comenzar el experimento:

- Primero se repasó el contorno de todos los reservorios con sellador.
- Luego se tomó una tapa sobre la que se depositó una gota de 2  $\mu$ l de proteína.
- Sobre esa gota se agregó 2  $\mu$ l de buffer del reservorio.
- Se dio vuelta la tapa sobre el reservorio correspondiente y se selló.
- Se repitieron los tres pasos anteriores para todos los reservorios de la placa.
- La placa se dejó en un cuarto con temperatura controlada, a 20°C, observando la misma dos veces por semana para ver la evolución de las gotas.

### 3.5 Ensayos de cultivo celular

#### 3.5.A Líneas celulares utilizadas y condiciones de cultivo

Para los ensayos de cultivo celular se utilizaron dos líneas humanas: las hTERT-RPE y las Hek 293FT.

La hTERT-RPE (ATCC) es una línea celular derivada de células del epitelio pigmentado de retina humana, de allí su nombre en inglés *retina pigmented epithelium* (RPE). Esta línea fue immortalizada utilizando la telomerasa transcriptasa reversa humana (hTERT), ésta es una subunidad catalítica de la telomerasa que permite la síntesis de los telómeros. Este proceso es muy importante en el control de la senescencia celular ya que la enzima está reprimida normalmente en células somáticas postnatales, generando el acortamiento de los telómeros en el mecanismo de envejecimiento celular. La sobreexpresión de la hTERT en las células RPE evita que esto suceda logrando la inmortalidad de la línea.

Las hTERT-RPE se crecieron en medio de cultivo DMEM alta glucosa (3151 mg/l) y F12 (1:1) suplementado con 10 % suero fetal bovino (SFB), 1 % de penicilina/streptomina (50 µg/µl) y 0,01 mg/ml de Higromicina B (para mantener la expresión de la hTERT).

Por otro lado la Hek 293FT (ATCC) es una línea originalmente derivada de células de riñón de embrión humano, de allí su nombre en inglés *human embryonic kidney* (Hek), que fueron transformadas con Adenovirus 5, incorporando parte de su ADN.

Las Hek 293 se crecieron en medio de cultivo DMEM con glucosa suplementado con 10 % SFB, 1 % de penicilina/streptomina (50 µg/µl), geneticina G418 (500 µg/ml) -para mantener la expresión del antígeno T de SV40, encargado de conferir la estabilidad a la línea-, piruvato de sodio (1 mM) y aminoácidos no esenciales (0,1 mM).

Todas las células fueron crecidas en una estufa a 37°C y a una presión constante de CO<sub>2</sub> de 5 %.

#### 3.5.B Sobre-expresión de genes en células humanas

La sobre-expresión de los genes de interés se realizó a través de la transfección transitoria de plásmidos en las células humanas utilizando FuGENE®HD de Promega. El FuGENE®HD es un método no liposomal de transfección pero que utiliza una mezcla patentada de lípidos para lograr una gran eficiencia y baja toxicidad.

Para desarrollar el protocolo de transfección se utilizará como ejemplo el correspondiente a la transfección de una placa de seis pocillos, aunque para la utilización de otros sustratos el protocolo no cambia, simplemente se deben modificar las cantidades proporcionalmente a la superficies utilizadas. El protocolo utilizado para transfectar una placa de seis pocillos (8,87 cm<sup>2</sup> de superficie de crecimiento por pocillo) es el siguiente:

- Plaqueo  $2,5 \times 10^5$  células por pocillo y dejo creciendo 24 horas en medio de cultivo sin antibiótico para alcanzar una confluencia de aproximadamente el 60 % para transfectar al otro día.
- Preparo para transfectar cada pocillo una mezcla de 2 µg de ADN + 8 µl FuGENE®HD + 250 µl Opti-MEM™ y la incubo por 20 minutos a temperatura ambiente.
- Agrego la mezcla de transfección a cada pocillo.
- El análisis se realiza luego de 48 o 72 hs post-transfección según el ensayo.

### 3.5.C Co-inmunoprecipitación

Para realizar los experimentos de coimmunoprecipitación (CoIp) se transfectaron células Hek 293FT cultivadas en placas de 100 mm de diámetro (P100, 58 cm<sup>2</sup> de superficie de crecimiento por placa) utilizando el mismo protocolo descrito anteriormente con las cantidades ajustadas a la superficie de la placa. Las células fueron transfectadas con HA-CCDC28B + myc-CCDC28B o, como control, HA-EV (vector vacío) + myc-CCDC28B. Para la inmunoprecipitación se utilizó EZview™ Red Protein G Affinity Gel (Sigma Aldrich), una resina de 4 % agarosa acoplada covalentemente a proteína G. La proteína G que tiene la capacidad de interaccionar con la región constante de los anticuerpos, que de esta forma se inmovilizan a la matriz para luego, al incubarlos con una mezcla proteica, unir específicamente la proteína de interés.

Buffer de CoIp: 50 mM Tris pH7,5; 150 mM NaCl; 1 mM EDTA y 1 % NP-40.

Una vez transfectadas las células, se cultivan por 48 horas, y se obtiene un extracto para realizar el experimento de coimmunoprecipitación siguiendo el siguiente protocolo:

- Levanto las células de cada P100 con 1 ml. de tripsina 0,05 % e incubando a 37°C por 5 minutos.
- Luego de levantar las células inactivo la tripsina con 6 ml de medio de cultivo con SFB, resuspendiendo bien las células.
- Centrifugo la suspensión por 10 minutos a 90 g.
- Descarto el sobrenadante y resuspendo las células en 1 ml de PBS 1X.

**Tabla 6. Anticuerpos utilizados durante este trabajo.**

- Centrifugo nuevamente por 10 minutos a 90 g.
- Descarto el sobrenadante y resuspendo las células en 500 µl de buffer CoIp.
- Incubo por 30 minutos en hielo vortexeando cada 5 minutos.

- Centrifugo a 14000 g por 20 minutos y separo el sobrenadante.
- Cuantifico proteínas en el sobrenadante para tomar 1 mg para realizar la co-inmunoprecipitación.
- Equilibro 30  $\mu$ l de resina con dos lavados de 1 ml de buffer CoIp
- Realizo una limpieza previa de la muestra (para eliminar las proteínas que se unen en forma inespecífica a la resina), incubando 1 mg de extracto con los 30  $\mu$ l de resina por 1 hora con rotación a 4°C.
- Durante la incubación anterior, equilibro otros 30  $\mu$ l de resina con dos lavados de 1 ml de buffer CoIp.
- Luego de culminada la hora de la limpieza previa, hago una pequeña centrifugación de la mezcla anterior, y paso el sobrenadante al tubo que contiene la resina recién equilibrada..
- Agrego a cada tubo 4  $\mu$ l de anticuerpo anti-HA e incubo toda la noche con rotación a 4°C.
- Centrifugo las muestras guardando el sobrenadante por si el experimento no funcionó.
- Lavo cinco veces la resina con 1 ml de buffer CoIp.
- Luego del último lavado resuspendo la resina en 35  $\mu$ l de buffer de carga y queda listo para analizar por SDS-PAGE.

#### 3.5.D Inmunofluorescencia

La inmunofluorescencia (IF) se utilizó para observar la presencia de proteínas específicas o estructuras particulares en cultivos celulares a través de la utilización de anticuerpos específicos y conjugados a fluoróforos. Para realizar los ensayos de IF, en primera lugar se crecieron células sobre cubreobjetos dentro de placas de 6 pocillos. Luego se fijan las células para mantener las estructuras celulares, y para evitar la pérdida de contenido intracelular cuando las células se permeabilizan para permitir la entrada de los anticuerpos. La fijación se puede realizar por dos métodos diferentes: metanol o paraformaldehído (PFA). El metanol deshidrata las células, provocando además la desnaturalización de las proteínas y su precipitación *in situ*, sin modificar su ubicación. Por otro lado el PFA genera uniones covalentes reticulares entre las moléculas, literalmente pegándolas entre sí. Una vez fijadas, las células pueden ser permeabilizadas utilizando Triton X-100 que interacciona con los lípidos de membrana. Luego que se permeabilizan las células, estas se bloquean para ocupar los posibles sitios de unión inespecífica de los anticuerpos y finalmente son incubadas con los anticuerpos específicos correspondientes.

El protocolo utilizado para la IF fue el siguiente:

- Quitar el medio de cultivo.
- Lavar con PBS precalentado a 37°C.
- Fijación: Cubrir con metanol e incubar por 10 minutos a -20°C, o con PFA 4 % en PBS por 10 minutos a temperatura ambiente. En algunos casos se utilizaron

los dos métodos de fijación secuencialmente, primero con PFA y luego con metanol. Esto permite mantener epítopes que se desnaturalizan con metanol, y al mismo tiempo mantener la integridad del centrosoma (que se fija bien con metanol), que podrá ser detectado con anticuepos anti- $\gamma$  tubulina.

- Realizar dos lavados con PBS de 1 minuto cada uno con agitación.
- Permeabilización: Cubrir con Tritón X-100 0,1% en PBS e incubar 10 minutos a temperatura ambiente con agitación.
- Realizar dos lavados con PBS de 1 minuto cada uno con agitación.
- Bloqueo: Cubrir con suero fetal bovino 5,5 % en PBS e incubar una hora a temperatura ambiente con agitación .
- Realizar dos lavados con PBS de 1 minuto cada uno con agitación.
- Cubrir con el anticuerpo primario en las condiciones correspondientes e incubar 2 horas a temperatura ambiente.
- Realizar tres lavados con PBS de 10 minutos cada uno con agitación.
- Cubrir con el anticuerpo secundario en una dilución 1/1000 en PBS y con DAPI (marcador de núcleo) diluído 1/2500 en PBS e incubar 1 hora a temperatura ambiente. A partir de este punto, cubrir los preparados para evitar la luz.
- Realizar tres lavados con PBS de 10 minutos cada uno con agitación.
- Dejar escurrir vertical sobre papel.
- Montar sobre el portaobjetos con 20  $\mu$ l de Prolong antifade (Invitrogen).
- Dejar secar hasta el otro día.
- Sellar el cubreobjetos con esmalte de uñas.

En la Tabla 6 se detallan todos los anticuerpos utilizados durante este trabajo.

<b>Anticuerpo Primario</b>	<b>Objetivo</b>	<b>Concentración</b>	<b>Tiempo</b>
Anti-tubulina acetilada (monoclonal de ratón, IgG2, Sigma)	Detección de la cilia	1/1000	1 h TA
Anti- $\gamma$ tubulina (policlonal de conejo, Sigma o monoclonal de ratón, Sigma)	Detección del centrosoma	1/1000	1 h TA
Anti-HA (monoclonal de ratón, IgG1, Sigma)	Detección de proteínas fusionadas a HA	1/1000	1 h TA
<b>Conjugados</b>			
Anti-IgG2 de ratón-AF 568		1/1000	1 h TA
Anti-IgG1 de ratón-AF 488		1/1000	1 h TA
Anti-IgG de conejo-AF 633		1/1000	1 h TA

### 3.5.E. Microscopía confocal y cuantificación de imágenes

Para la captura de imágenes se utilizó el microscopio confocal Zeiss LSM 800, con objetivo 63x de inmersión en aceite. Se tomaron secciones cada 0,4  $\mu\text{m}$  y la imagen final se obtuvo por composición de los diferentes canales y la proyección en Z (algoritmo de suma).

Para el análisis de las imágenes se utilizó el software ImageJ Version: 2.0.0-rc-54/1.51h. Para realizar el conteo de células ciliadas, se contaron por un lado todos los núcleos presentes en un campo, y por otro lado las cilias presentes en el mismo. Los resultados se expresan como la proporción de células ciliadas con un intervalo de confianza de 95% calculado como

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

(Pagano M, Gauvreau K (2000) Principles of Biostatistics. USA: Duxbury. Thomson

Learning). El análisis estadístico se realizó aplicando un test de contraste de hipótesis para proporciones (Pagano M, Gauvreau K (2000) Principles of Biostatistics. USA: Duxbury. Thomson Learning).

### 3.5.F Diseño y obtención de una línea celular CCDC28B *knockdown* por la tecnología de CRISPR/Cas9

A fin de obtener una línea celular estable donde la expresión o funcionalidad de CCDC28B estuviera disminuida, utilizamos la tecnología de edición genómica CRISPR/Cas9 (Jinek y cols., 2012). Brevemente, se diseñan ARNs guía específicos para la secuencia de *CCDC28B* que guiarán a la nucleasa Cas9 a la zona del gen que se quiere modificar. La nucleasa corta la doble hebra de ADN y el error se resuelve por unión de extremos no homólogos o NHEJ (*non-homologous end joining*) que es un sistema que genera pequeñas pérdidas o inserciones en el sitio de ruptura. Guiando la nucleasa a un exón codificante del gen de interés se puede generar un corrimiento del marco abierto de lectura, la introducción de un codón de terminación prematuro y por tanto una mutación nula. Para ello utilizamos un sistema lentiviral pLENTI-CRISPR v2 (Sanjana N. E. et al, 2014), que permite expresar al mismo tiempo la enzima Cas9 y el ARN guía (unido al ARN al que se une la Cas9).

Lo primero que se realizó fue el diseño de varios ARNs específicos de CCDC28B, dirigidos al exón 2 y 3, a través del buscador del Laboratorio de Feng Zhang (<http://crispr.mit.edu>). El diseño de los ARNs se basa en una secuencia de homología de 20 pb con la región que se quiere modificar y la misma debe tener en su extremo 3' una secuencia PAM (del inglés: *protospacer adjacent motif*) de 3 pb de forma: NGG (donde N es cualquier nucleótido). Esta es la secuencia reconocida por la Cas9 una vez que el ARN la dirigió hasta allí, y recién en ese momento la nucleasa actúa. La estructura final de los ARNs se muestra en la Figura 20. Luego de que el buscador nos brindó una lista de muchos posibles ARNs, elegimos tres para cada exón y mandamos entonces a sintetizar los seis pares de ARNs que se detallan a continuación:



- Exón 2:
  1. hsCCDC28BgRNA\_1\_Fwd: **CACCG**TTGAACTTGGCCCGCTGCTT  
hsCCDC28BgRNA\_1\_Rev: **AAACA**AGCAGCGGGCCAAGTTCAAC
  2. hsCCDC28BgRNA\_2\_Fwd: **CACCG**TTTTGTCATCCATTGGGCGC  
hsCCDC28BgRNA\_2\_Rev: **AAACG**CGCCCAATGGATGACAAAAC
  3. hsCCDC28BgRNA\_3\_Fwd: **CACCG**GGCCCCAGGCACACTACGGA  
hsCCDC28BgRNA\_3\_Rev: **AAACT**CCGTAGTGTGCCTGGGGCCC
- Exón 3:
  4. hsCCDC28BgRNA\_4\_Fwd: **CACCG**AGACATCAGTCACCTCGGTC  
hsCCDC28BgRNA\_4\_Rev: **AAACG**ACCGAGGTGACTGATGTCTC
  5. hsCCDC28BgRNA\_5\_Fwd: **CACCG**TTTCCACTCTGGCCGGCTGC  
hsCCDC28BgRNA\_5\_Rev: **AAACG**CAGCCGGCCAGAGTGGAAC
  6. hsCCDC28BgRNA\_6\_Fwd: **CACCG**AGAGAAGTGCCGCCAGTCC  
hsCCDC28BgRNA\_6\_Rev: **AAACG**GACTGGGCGGCACTTCTCTC



**Figura 20. Estructura de ARNg para ser clonado en el vector lentiviral pLENTI CRISPR v2.**

El diseño de los ARNg incluye una secuencia de 20 pb homóloga a la región que se quiere editar (N en la secuencia) y con una secuencia PAM 3' (NGG). Además una secuencia específica detallada en la Figura en los extremos del ARNg para que el mismo sea reconocido por la Cas9.

Una vez obtenidos los ARNg se procedió a clonarlos en el vector lentiviral pLENTI CRISPR v2 de acuerdo al protocolo sugerido por los fabricantes (Sanjana N. E. et al, 2014):

- 5 µg de vector pLENTI CRISPR v2
- 3 µl FastDigest *BsmBI* (Fermentas)
- 3 µl Fast AP (Fermentas)
- 6 µl 10 X FastDigest Buffer
- 0,6 µl 100 mM DTT (preparado en el momento)
- Llevar a 60µl con ddH<sub>2</sub>O
- Incubar por 30 minutos a 37°C

El segundo paso consistió en correr en un gel de agarosa el producto de la digestión, para luego purificar la banda de alto peso molecular (aproximadamente 12 kb) utilizando el kit PureLink™ Quick Gel Extraction Kit de Invitrogen.

En paralelo, se fosforilaron e hibridaron cada par de oligonucleótidos según el siguiente protocolo:

- 1 µl Oligo #1 Fwd (100 µM)
- 1 µl Oligo #2 Rev (100 µM)
- 1 µl 10 X Buffer de ligación T4 (NEB)
- 6,5 µl ddH<sub>2</sub>O

- 0,5 µl T4 PNK (NEB M0201S)
- Incubar la mezcla en un termociclador con el siguiente programa: 30 minutos a 37°C, luego 5 minutos a 95°C y finalmente realizar una rampa descendente hasta 25°C a una velocidad de 5°C/minuto.
- Una vez culminada la reacción diluir la mezcla de oligos hibridados en una proporción de 1 en 200 en ddH<sub>2</sub>O

Finalmente se realiza la ligación de los oligonucleotidos hibridados y el plásmido pLENTI CRISPR v2:

- 50 ng de plásmido pLENTI CRISPR v2 digerido con *BsmBI*
- 1 µl dilución de oligonucleotidos hibridados
- 5 µl 2 X Buffer de Quick Ligasa (NEB)
- Llevar a 10 µl con ddH<sub>2</sub>O
- Agregar 1 µl de Quick Ligasa (NEB M2200S)
- Incubar a temperatura ambiente por 10 minutos

Una vez realizada la ligación, se transforman bacterias *Stbl3* (Invitrogen C7373-03) utilizando el protocolo recomendado por el proveedor. Se utilizn esta cepa ya que es deficiente en recombinación y por lo tanto se reduce la frecuencia de recombinación homóloga de repetidos terminales largos (LTR, del inglés: *long terminal repeats*) presentes en los plásmidos lentivirales.

Los lentivirus se obtuvieron transfectando células Hek293 en placas de 100 mm de diámetro con pLENTI CRISPR v2 (conteniendo ARNg, 5 µg) y los vectores pCMV-VSV-G (Addgene #8454, 2,5 µg) y psPAX2 (Addgene #12260, 3,8 µg), ambos de empaquetamiento. La transfección se realizó de la misma forma en la que se describió anteriormente (ver Sección 3.5.B) y 48 horas luego de transfectadas las células, se cosechó el sobrenadante de cultivo conteniendo los lentivirus. El sobrenadante se centrifugó a 4000 rpm por 15 minutos para eliminar restos celulares, se filtró por 0,45 µm y se congeló a a -80°C hasta su utilización.

Transducción de células hTERT RPE:

- Se plaquearon las células el día previo a la transducción para llegar en un 60 % de confluencia (en placas de seis pocillos se plaquearon 3 x 10<sup>5</sup> células por pocillo).
- Se preparó la siguiente mezcla para cada pocillo: 1 ml de medio de cultivo + 1 ml suspensión de lentivirus + Polybrene 8 µg/ml (concentración final).
- Se incubó por 24 horas.
- Luego de transcurridas las 24 horas se cambio el medio de cultivo a las células, por medio con puromicina (10 µg/ml), para comenzar la selección de las células que incorporaron el plásmido lentiviral a su genoma.

A diferentes tiempos obtuvimos extractos celulares y evaluamos la expresión de CCDC28B por WB, a fin de verificar cambios en la expresión de la proteína endógena.

### 3.5.G Obtención y caracterización de clones CCDC28B *kd*

Se transdujeron células hTERT-RPE cultivadas en una placa de 100 mm de diámetro de la misma forma que la descrita anteriormente (ver sección 3.5.D), pero utilizando lentivirus que codificaban solamente para el ARNg #1. Luego de 5 días de selección con puromicina (6 días post-transducción) se realizó un clonado por deposición de una única célula por pocillo en 3 placas de 96 pocillos utilizando el clasificador celular de alta velocidad BD FACSAria™ Fusion (BD Biosciences).

Las células fueron mantenidas en medio con 20% de SFB y a medida que iban llenando el pocillo se pasaron a placas con pocillos de mayor superficie, en forma progresiva: placas de 48, 24, 12 y 6 pocillos. No todas las células crecieron a la misma velocidad, por lo tanto, a medida que fue posible se obtuvieron extractos celulares para analizar la expresión de CCDC28B por WB y se plaquearon células sobre cubreobjetos para evaluar el porcentaje de ciliación por IF.

Seis clones que mostraron dificultades para ciliar fueron caracterizados más en profundidad, a fin de caracterizar el cambio que habían sufrido en el gen de *CCDC28B*. Para ello se obtuvo ADN genómico, se amplificó el exón 2 por RT-PCR utilizando primers específicos, se clonó el fragmento amplificado en el vector pGEM-T-easy, se transformaron bacterias DH5 $\alpha$ , se purificaron los plásmidos a partir de varias colonias bacterianas y se enviaron a secuenciar con el fin de conocer la secuencia de los dos alelos de CCDC28B.

Extracción de ADN genómico:

- Se resuspenden  $5 \times 10^5$  células en 300  $\mu$ l de buffer de lisis (10 mM Tris/HCl pH 8, 10 mM EDTA, 10 mM NaCl, 0,5 % SDS, 10  $\mu$ g/ml Proteinasa K (Sigma)) y se incuba 2 hs a 60°C.
- Se agrega 900  $\mu$ l de buffer de precipitación (150 mM NaCl en 100% de EtOH), se homogeniza y se incuba 30 min a TA.
- Se centrifuga a 15000 g por 15 min a 4°C.
- Se lava el precipitado con EtOH 70% 2 veces, centrifugando a 15000 g por 5 minutos.
- Se remueve el EtOH y el precipitado se deja secar por 10 min a TA y luego se disuelve en 50-100  $\mu$ l de agua.

En el clon que fue seleccionado para trabajar (#2.B1) se caracterizó también el ARN mensajero de CCDC28B. Para ello se extrajo ARN y se preparó ADNc utilizando oligo dT y el kit SuperScript™ First-Strand Synthesis System for RT-PCR (Invitrogen) de acuerdo a las instrucciones del fabricante. Se amplificó por RT-PCR el ADNc de CCDC28B utilizando primers específicos, se corrió la mezcla de PCR en un gel de agarosa 1.5 % y se observaron 2 transcriptos,

que por tamaño se corresponden con los alelos secuenciados en este clon. La mezcla de PCR se utilizó para clonar los ADNc en pGEM-T Easy, se transformaron bacterias y se purificaron plásmidos a partir de varias colonias para mandar secuenciar.

## RESULTADOS

---

### **OBJETIVOS 1 & 2: Producción y purificación de CCDC28B recombinante.**

#### **Rastreo de condiciones de cristalización de CCDC28B recombinante y determinación de su estructura tridimensional**

El objetivo principal de este trabajo de tesis fue obtener información acerca de la función de CCDC28B mediante la resolución de su estructura y la identificación y caracterización de dominios en la proteína. Para la caracterización estructural de CCDC28B, la técnica de elección fue la difracción de rayos X. Para lograr este objetivo, una parte significativa de este trabajo se centró en la producción y purificación de CCDC28B recombinante, un paso crítico para la posterior obtención de cristales y resolución de la estructura.

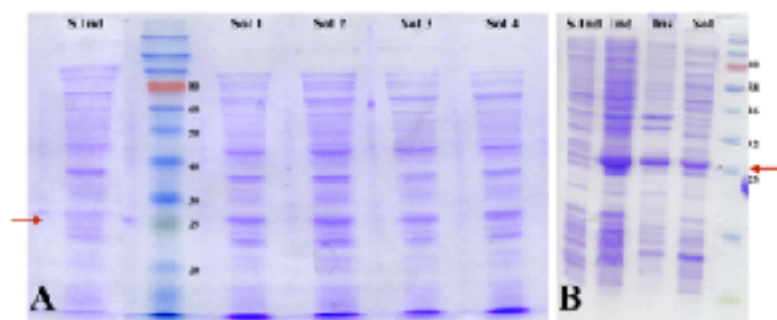
#### 4.1. Producción de CCDC28B *wild type* recombinante

En primer lugar, se clonó el marco abierto de lectura de CCDC28B humana en el vector de expresión pET21d, que permitió expresar nuestra proteína de interés fusionada a una etiqueta de 6 Histidinas (6-His) en su extremo C-terminal y posteriormente purificarla mediante una cromatografía de afinidad a iones metálicos (IMAC). Una vez clonado el ORF, se transformaron células *E. coli* Rosetta™(DE3)pLysS con dicho plásmido y a partir de allí se planificó una estrategia de expresión de CCDC28B recombinante estándar, para luego trabajar en su optimización.

##### 4.1.A Optimización de las condiciones de inducción

A la hora de realizar una producción de proteína recombinante, algunos de los parámetros que se pueden modificar para optimizar el proceso son: el vector en el que se clona el gen de interés, la cepa de *E. coli* donde se expresará (o incluso eventualmente el modelo, pasando de *E. coli* a células de mamífero o de *Drosophila*) y las condiciones de inducción. En nuestro trabajo, se eligió el vector pET21d ya que permite la expresión de la proteína con una etiqueta de 6-His que sirve para purificarla y la cepa Rosetta™(DE3)pLysS que está optimizada para expresar genes con codones eucariotas en bacterias y se trabajó en poner a punto las condiciones de inducción utilizando distintos medios de cultivo, modificando la temperatura a la cual se induce la expresión, o variando la concentración del agente inductor (en nuestro caso IPTG).

La producción de CCDC28B se realizó inicialmente creciendo las bacterias Rosetta™(DE3)pLysS en medio 2YT, que a diferencia del medio LB permite alcanzar mayor densidad celular y por lo tanto tener más biomasa. También se evaluaron cuatro condiciones diferentes de inducción: 1 mM IPTG incubando ON a 20°C, 0,1 mM IPTG incubando ON a 20°C, 1 mM IPTG incubando a 37°C por 3 horas y 0,1 mM IPTG incubando a 37°C por 3 horas. Se tomaron muestras de los cultivos antes del agregado de IPTG (extracto total sin inducir: S.Ind) y al final de



**Figura 21. Optimización de las condiciones de inducción de la producción de CCDC28B.**

**A-** SDS-PAGE 12% teñido con Azul de Coomassie donde se muestran los extractos solubles (Sol) obtenidos luego de la inducción en diferentes condiciones. S.Ind: extracto total sin inducir, 1: 1 mM IPTG a 20°C, 2: 0,1 mM IPTG a 20°C, 3: 1 mM IPTG a 37°C 3h., 4: 0,1 mM IPTG a 37°C 3h.

**B-** Protocolo seleccionado para producir CCDC28B recombinante. La inducción se realizó con 0,1 mM IPTG a 20°C. De izquierda a derecha: S.Ind: extracto total sin inducir, Ind: extracto total inducido, Ins: Fracción Insoluble, Sol: Fracción soluble. Las flechas rojas indican la banda correspondiente a CCDC28B.

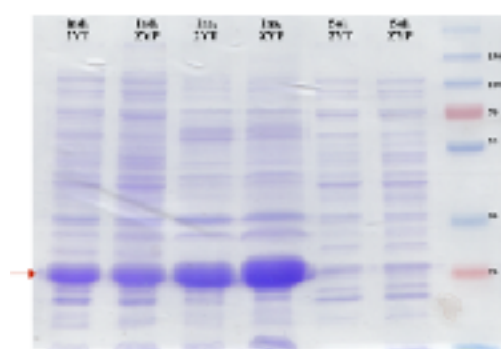
la inducción (extracto total inducido: Ind). Luego las células se sonicaron y del extracto total se obtuvo una fracción soluble (Sol) y otra insoluble (Ins) mediante centrifugación (ver Materiales y Métodos). Las mismas fueron analizadas por SDS-PAGE (Figura 21). Si bien no se observaron grandes diferencias de rendimiento entre las diferentes condiciones ensayadas se obtuvo una mayor cantidad de proteína soluble en las condiciones de inducción ON a 20°C. Dentro de esta también se ve una tendencia a una mayor inducción con 0,1 mM de IPTG. Basándonos en estos resultados se decidió utilizar esta última condición para la inducción de CCDC28B.



Como resultado general de la prueba de inducción observamos que el protocolo más eficiente es el que consiste en la inducción con 0,1mM de IPTG, O/N a 20°C.

Resulta importante destacar que tanto en el protocolo seleccionado (ver Figura 21B), así como en las otras condiciones evaluadas (datos no mostrados), hay una fracción importante de CCDC28B que queda insoluble, probablemente en cuerpos de inclusión. En este punto del trabajo se decidió seguir adelante con la purificación ya que el material obtenido de forma soluble, según nuestra evaluación, era suficiente. En paralelo sin embargo, evaluamos (y continuamos haciéndolo, ver Discusión) algunas alternativas para mejorar la solubilidad de nuestra proteína de interés.

Una estrategia utilizada fue producir CCDC28B WT utilizando el medio de autoinducción ZYP-5052 (Studier, F.W. 2005). Los resultados de ambas producciones (utilizando medio 2YT y ZYP-5052) se muestran a continuación. Luego de correr las fracciones solubles e insolubles



obtenidas con ambos medios en un SDS-PAGE (Figura 22), se observa que pese a no haber una diferencia importante en la cantidad de CCDC28B en las fracciones solubles (Sol.), sí

**Figura 22. Comparación de diferentes medios de inducción.**

SDS-PAGE 12% teñido con Azul de Coomassie donde se corrieron los extractos totales de cultivos inducidos (Ind) y las fracciones solubles (Sol) e insolubles (Ins) de cultivos en medio 2YT inducidos con IPTG o medio ZYP de autoinducción.

La flecha roja indica la banda correspondiente a CCDC28B.

la hay en las insolubles (Ins.), acumulándose más proteína en la producción del medio ZYP-5052. Teniendo en cuenta que la inducción (Ind.) parece haber sido similar en ambos casos, el hecho de que haya más cantidad de proteína en la fracción insoluble cuando la inducción total fue similar inclinó nuestra decisión por continuar nuestro trabajo utilizando el medio 2YT inducido por IPTG.



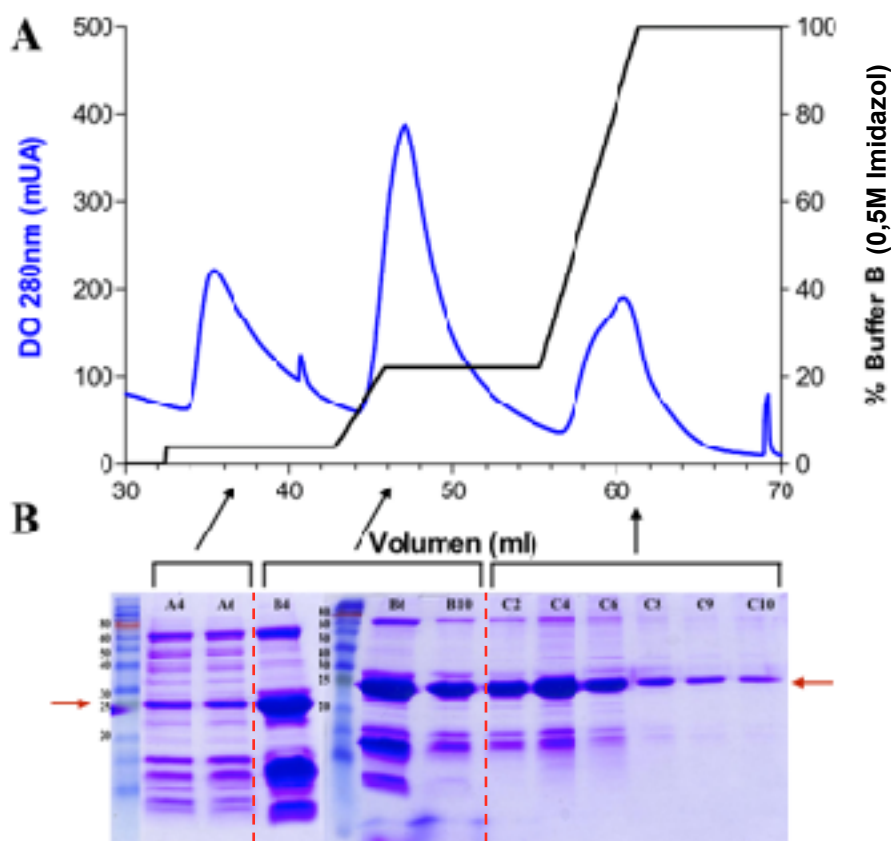
Otro resultado común a las cuatro condiciones evaluadas, fue que la gran mayoría del material inducido queda insoluble como cuerpos de inclusión. Pese a esto, se evaluó que la cantidad de CCDC28B soluble obtenida era suficiente para continuar trabajando.

## 4.2. Purificación y caracterización de CCDC28B *wild type*

### 4.2.A Optimización de la cromatografía de afinidad por metales (IMAC)

Como se mencionó anteriormente, la expresión de CCDC28B a partir del plásmido pET21d resulta en la producción de una proteína de fusión CCDC28B-6His. Por lo tanto, el primer paso de purificación utilizado fue una cromatografía de afinidad a Ni<sup>2+</sup> (IMAC), con columnas HisTrap™ HP. La etiqueta de 6-His se une con alta afinidad a iones metálicos (Ni<sup>2+</sup> en el caso de la HisTrap™) en la matriz de la columna. Posteriormente, se utilizó el imidazol como agente competitivo en la unión con el Ni<sup>2+</sup>, eluyendo de esta forma nuestra proteína de interés.

En primer lugar se utilizó una única columna HisTrap™ HP de 1 ml, a la cual se le inyectó el extracto soluble correspondiente a 0,5 l de cultivo bacteriano. Allí, se realizó un protocolo estándar de IMAC que consistió en un lavado con un escalón de 4 % buffer B (40 mM imidazol), un gradiente hasta 20 % de B (120 mM imidazol) en tres volúmenes de columna y finalmente otro gradiente a 100 % de B (0,5 M imidazol) en seis volúmenes. Utilizando este protocolo se observó que se conseguía un importante grado de purificación, obteniendo como mayoritaria la banda



**Figura 23. Purificación de CCDC28B por cromatografía de afinidad a metales (IMAC)**

A- Cromatograma obtenido al fraccionar por IMAC (HisTrap de 1 ml) el extracto soluble correspondiente a 0,5 l de cultivo bacteriano inducido para la producción de CCDC28B. La línea azul representa la absorbancia a 280 nm mientras la línea negra representa el porcentaje de buffer de elución.

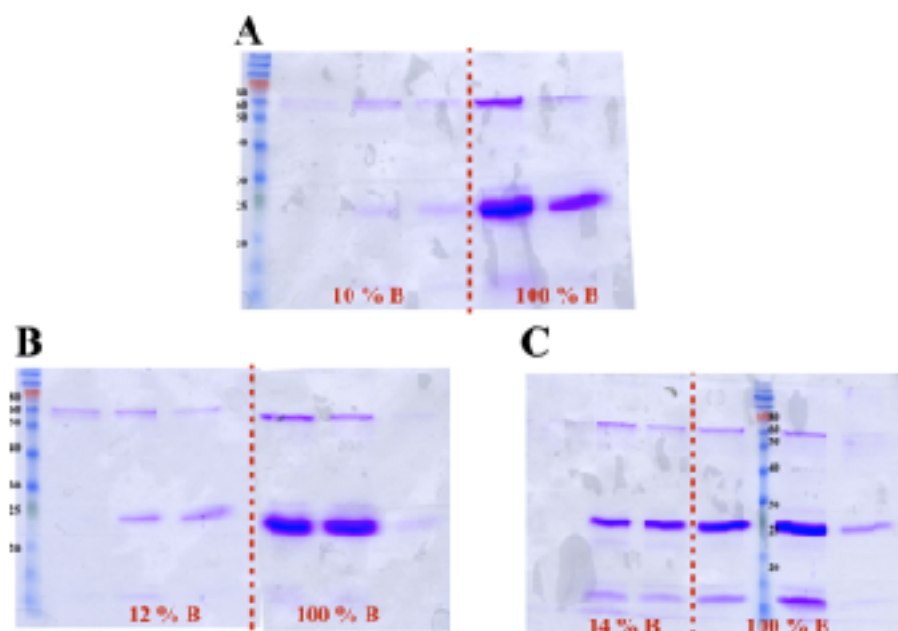
B- SDS-PAGE al 12% teñido con Azul de Coomassie donde se corrieron algunas fracciones de la purificación. Con paréntesis negros se señalan las fracciones correspondientes a cada uno de los picos mayoritarios de la IMAC, mientras que entre líneas punteadas rojo se marca el material que se colectó para continuar la purificación.



correspondiente a CCDC28B en las fracciones que eluyeron durante el gradiente de 20 % a 100 % de B (ver fracciones C2 a C6 en Figura 23). Sin embargo, una cantidad importante de CCDC28B se eluía previamente junto con muchos contaminantes, en el gradiente de 4 a 20 % de B, y además se eluía a lo largo de los gradientes y no como un pico discreto a una determinada concentración de imidazol (Figura 23). Por este motivo se pensó que una elución utilizando escalones podría ser más útil que un gradiente, ya que este último no aporta en términos de purificación y diluye la muestra purificada.

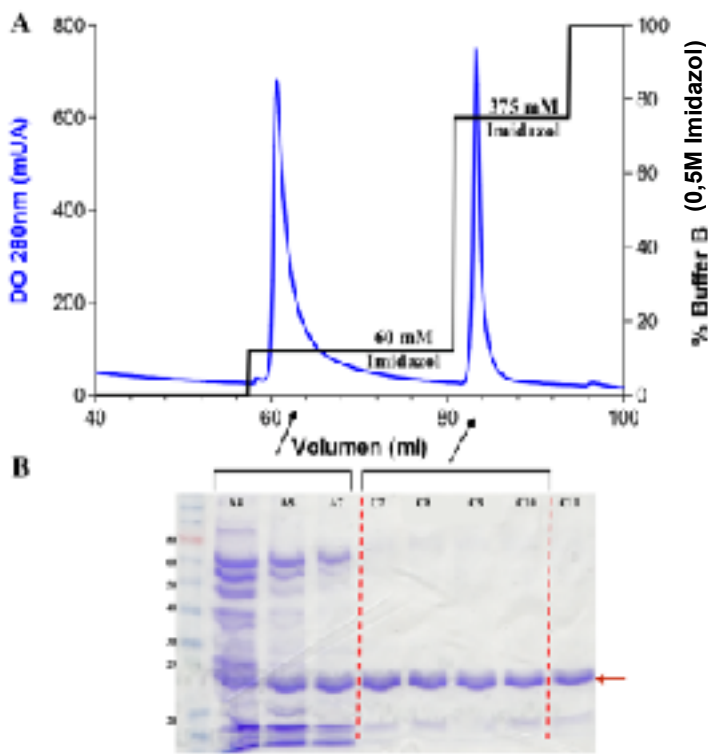
Por otro lado, se observó que la mayoría de los contaminantes eluían en el primer lavado (fracciones A4 y A6 en Figura 23) y en el primer gradiente (fracciones B4 a B10, Figura 23) y solo un contaminante mayoritario es el que se arrastra en las fracciones más puras de CCDC28B (la banda de aproximadamente 60 kDa; el contaminante de aproximadamente 15 kDa no se observó recurrentemente al reproducir este protocolo).

Se continuó optimizando el protocolo de purificación por IMAC modificando el lavado inicial intentando eluir la mayor cantidad de contaminantes sin que el material de interés se despegue de la columna. Los protocolos utilizados fueron los siguientes: 1) un escalón de lavado en 10 % (70 mM imidazol); 2) un escalón de lavado en 12 % (80 mM imidazol); 3) un escalón de lavado en 14 % (90 mM imidazol). En todos los casos, a posteriori se eluyó CCDC28B con un escalón de 100 % de B (0,5 M imidazol). Luego de analizar por SDS-PAGE las fracciones obtenidas con los tres protocolos, se observó que utilizando un escalón de 10 % de B no se consigue despegar el contaminante de mayor tamaño mientras que cuando el lavado se realizó con un



**Figura 24. Optimización del paso de lavado en la IMAC.**

A fin de optimizar la purificación de CCDC28B por IMAC se probó el efecto de cambiar la concentración de imidazol en el primer paso de lavado. Luego del lavado se eluyó con 100 % de buffer B (0,5 M imidazol). En cada cromatografía se inyectaron partes iguales de una muestra ya purificada previamente por IMAC y dializada para sacarle el imidazol. Se muestran los SDS-PAGE 12 % teñido con Azul de Coomassie de las fracciones obtenidas con los lavados y la elución con 100 % de B de las diferentes cromatografías. A- Lavado con 10 % de B; B- Lavado con 12 % de B; C- Lavado con 14% de B.



**Figura 25. Protocolo estándar del paso de purificación IMAC.**

**A** - Cromatograma correspondiente al fraccionamiento del extracto soluble obtenido a partir de 1 l de cultivo bacteriano sobre dos columnas HisTrap de 1 ml conectadas en serie. La línea azul indica la absorbancia a 280 nm, la línea negra indica el porcentaje de buffer de elución.

**B** - SDS-PAGE 12 % teñido con Azul de Coomassie de algunas de las fracciones de la cromatografía. Con paréntesis negros se señalan las fracciones correspondientes a cada uno de los picos mayoritarios de la IMAC. Para seguir trabajando se seleccionaron las fracciones indicadas entre líneas rojas punteadas (1 ml cada una) que eluyen con 375 mM de imidazol.

escalón de 14 % el contaminante de mayor tamaño se eluye en el lavado, pero a costo de una pérdida importante de la proteína de interés. Se estableció por lo tanto que la mejor relación pureza/cantidad se obtenía lavando con un

escalón de 12 % de buffer B (Figura 24), y es este el protocolo que se utilizó a partir de este momento para la purificación de CCDC28B recombinante.

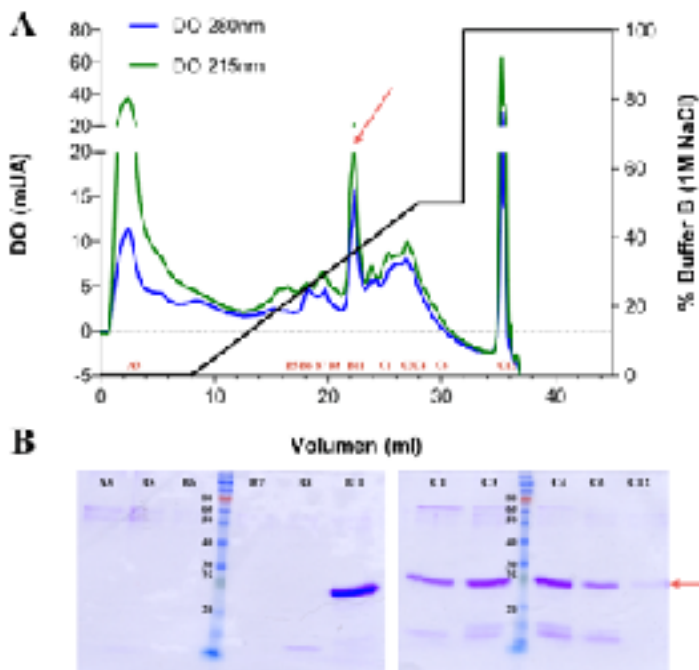
A partir de aquí entonces elegimos como protocolo estándar para la IMAC un primer lavado con 0 % de buffer B (100 % buffer A, 20 mM imidazol), luego un escalón con 12 % de buffer B como segundo lavado, luego elución con un escalón de 75 % de buffer B (donde esperábamos salgan las fracciones enriquecidas en CCDC28B), y finalmente un último escalón al 100 % de buffer B. Los resultados fueron los esperados (Figura 25): se obtuvo un conjunto de fracciones enriquecidas en CCDC28B, con un contaminante mayoritario correspondiente a la banda de aproximadamente 60 kDa (no observada claramente en esta figura), pero en una relación favorable hacia nuestra proteína de interés. Como resultado final, se logró alcanzar un rendimiento de hasta 3,2 mg de proteína por litro de cultivo.



La optimización de la IMAC arrojó como resultado que las condiciones óptimas consisten en la inyección del extracto soluble correspondiente a 0,5 l de cultivo en dos columnas HisTrap™ HP de 1 ml conectadas en serie y eluyendo con un protocolo de tres escalones de 12, 75 y 100 % de buffer B (0,5 M imidazol). El rendimiento obtenido es de 3,2 mg de proteína por litro de cultivo bacteriano.

#### 4.2.B Cromatografía de intercambio aniónico

Luego de puesta a punto la IMAC nos enfocamos en el segundo paso de purificación, la cromatografía de intercambio aniónico, para tratar de eliminar el contaminante de mayor peso molecular que eluyó junto a CCDC28B en la IMAC. Para esta cromatografía se utilizó una columna



**Figura 26. Segundo paso de purificación de CCDC28B: Cromatografía de Intercambio Aniónico.**

**A** - Cromatograma del intercambio aniónico sobre una columna Mono Q™ HR 5/5 del material obtenido en la IMAC. La línea azul indica la absorbancia a 280 nm y la verde a 215 nm. En negro se indica el porcentaje de buffer B (aumentando la concentración de NaCl).

**B** - SDS-PAGE 12 % teñido con Azul de Coomassie de algunas de las fracciones de la cromatografía. En la fracción B11 (flecha en el cromatograma), correspondiente a 360 mM NaCl es posible recuperar CCDC28B con un nivel de pureza aceptable.

MonoQ™ HR 5/5 (GE™ Healthcare Life Sciences). La columna tiene un volumen de 1 ml, y tiene la capacidad de unir hasta ~50 mg de proteína total y separar hasta ~5-10 mg de cada proteína individual. Previo a la utilización de dicha columna, debimos dializar el material eluido de la

IMAC contra Buffer A MonoQ (20 mM Tris/HCl pH 8, 20 mM NaCl), para eliminar el imidazol y el NaCl de la muestra, equilibrándola en el buffer apropiado para el intercambio aniónico. Luego de dializado el pool, se centrifugó la muestra a 15000 g, logrando separar una fracción de material que había precipitado durante la diálisis. En principio, no se intentó resuspender dicho material y fue descartado. El sobrenadante de dicha centrifugación se filtró por 0,45 µm y se utilizó para inyectar en la MonoQ.

El protocolo utilizado para intercambio aniónico consistió de un lavado con 0 % de buffer B, para limpiar todo lo que no interaccionó con la columna, un gradiente hasta 50 % de buffer B (0,5 M NaCl) en 20 ml, y luego un escalón hasta 100 % de buffer B (1 M NaCl). La corrida cromatográfica arrojó como resultado un pico mayoritario que eluyó en 28 % de buffer B (~300 mM de NaCl), que al ser analizado luego por SDS-PAGE, demostró corresponder a nuestra proteína de interés, y con un alto grado de pureza (Figura 26). Sin embargo, se observó que en las otras fracciones también eluye la proteína de interés, prácticamente a lo largo de toda la corrida y junto con los otros contaminantes, disminuyendo sensiblemente el rendimiento de la purificación si se seleccionara para continuar trabajando únicamente la fracción donde CCDC28B eluye pura.

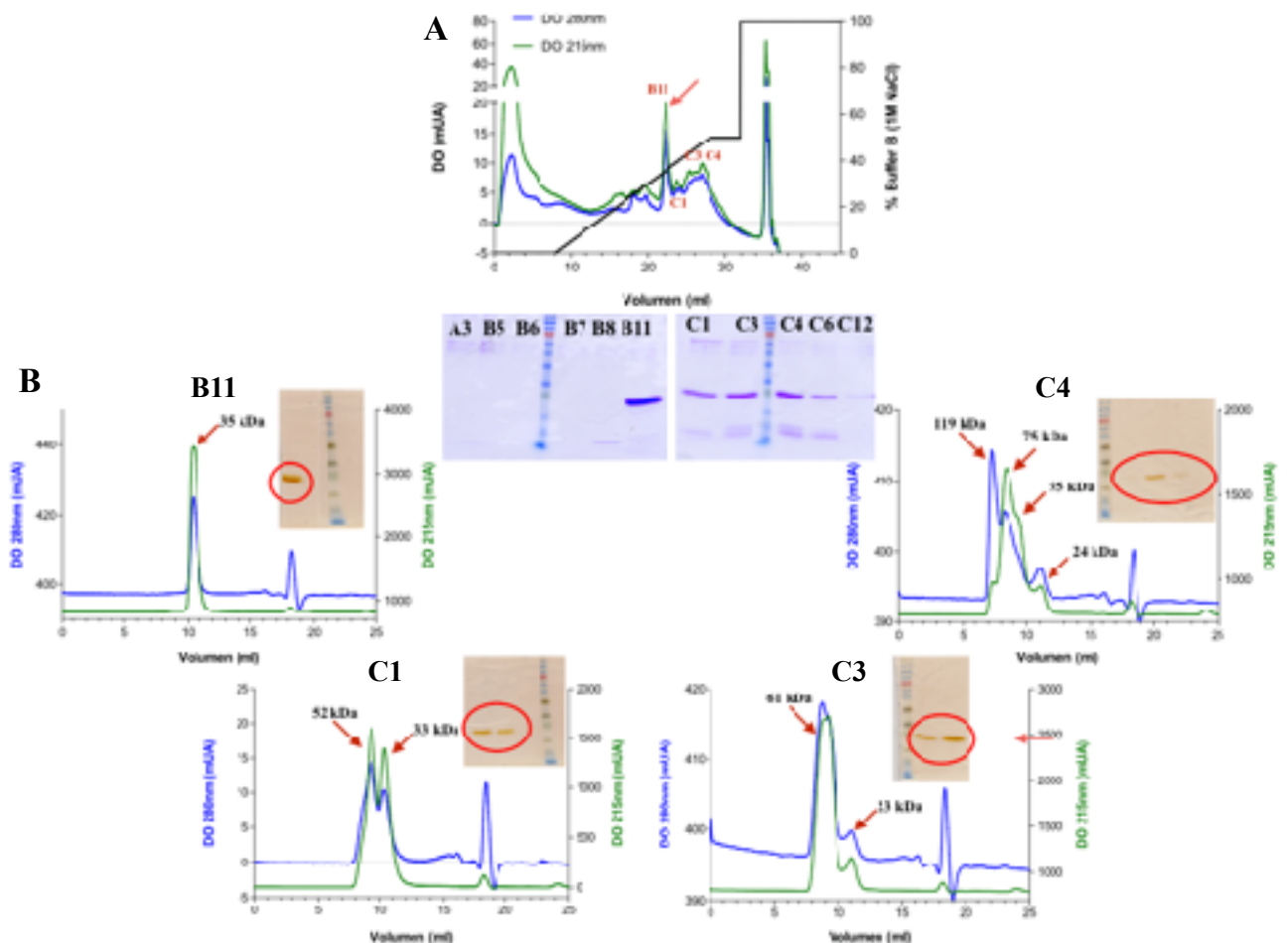


La cromatografía de intercambio aniónico arrojó como resultado un perfil donde CCDC28B eluye a lo largo de todo el gradiente salino, disminuyendo así la eficiencia de la corrida debido al poco aprovechamiento del material para continuar trabajando.

#### 4.2.C Cromatografía de exclusión molecular de las distintas fracciones del intercambio aniónico

Los resultados de la purificación por intercambio aniónico llamaron nuestra atención debido a que en principio no habría ninguna razón por la cual CCDC28B debiera eluir a lo largo de todo el gradiente. Pensamos que el hecho de que la proteína eluya de esta manera podría estar indicando que hay diferentes poblaciones de la muestra con cargas diferentes. Para evaluar esta posibilidad se analizaron algunas de las fracciones obtenidas en el intercambio aniónico por cromatografía de exclusión molecular (SEC, del inglés: *size exclusion chromatography*) de forma tal de observar si había diferencias en el estado oligomérico de la proteína.

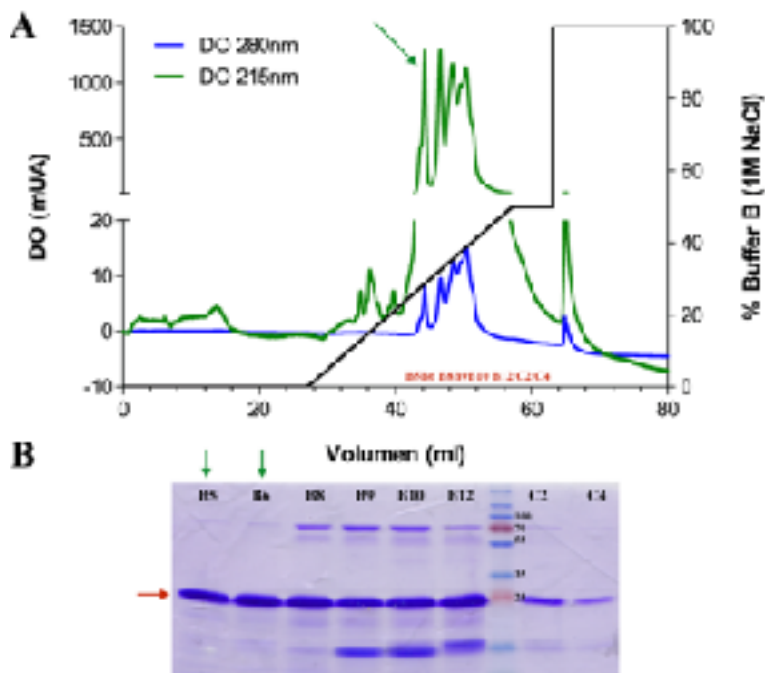
Para esto se corrieron las fracciones por SEC analítica utilizando la columna Superdex™ 75 10/300. Los resultados demostraron que en efecto, las diferentes fracciones obtenidas en el intercambio aniónico contienen distintos estados oligoméricos de CCDC28B (Figura 27). La SEC



**Figura 27. Análisis de las diferentes fracciones del intercambio aniónico por Gel Filtración.**

**A** - Adaptación de la Figura 23, donde se observa el cromatograma resultante del intercambio aniónico y el SDS-PAGE 12 % de las distintas fracciones del mismo. Las fracciones corresponden de izquierda a derecha a una elución con mayor concentración de sales.

**B** - Cromatogramas de las corridas de exclusión molecular de las fracciones de la MonoQ que se encuentran señaladas. Las líneas azules corresponden a la absorbancia a 280 nm, mientras que las verdes a 215 nm. Se señala para cada pico el peso molecular estimado según la calibración de la columna, observándose que las fracciones que eluyen del intercambio aniónico a mayor concentración de sal, corresponden a estados oligoméricos mayores de CCDC28B. En cada cromatograma se muestra además el SDS-PAGE 12 % con tinción por nitrato de plata de varias fracciones. Con un círculo rojo se marcó las fracciones correspondientes a cada pico, corroborando de esta manera que en todos los picos está presente CCDC28B.



**Figura 28. Intercambio aniónico del material conteniendo oligómeros de CCDC28B.**

Las fracciones de un intercambio aniónico previo, que contenían monómero y oligómeros de CCDC28B se juntaron y volvieron a cromatografiar sobre la MonoQ

**A** - Cromatograma del intercambio aniónico sobre una columna Mono Q™ HR 5/5. La línea azul representa la absorbancia a 280 nm y la línea verde a 215 nm. Comparar el perfil de elución con el de la primer MonoQ, en la figura 26.

**B** - SDS-PAGE 12% de algunas de las fracciones de la cromatografía. CCDC28B está presente en todas las fracciones nuevamente. Y la relación entre las cantidades de CCDC28B y el contaminante de mayor peso molecular es igual al encontrado anteriormente (figura 26).

de la fracción correspondiente al pico mayoritario (B11) muestra que CCDC28B eluye como un único pico a un peso molecular estimado, de acuerdo a la calibración de la columna,

de 35 kDa (Figura 27B), cercano al que corresponde a un monómero de 22 kDa. Por el contrario, al realizar las corridas de las fracciones siguientes el patrón de elución en SEC cambia, observándose varios picos (Figura 27B): junto al pico correspondiente al monómero (que continúa apareciendo en todas las fracciones) aparecen otros que corresponderían a oligómeros de mayor tamaño.

En resumen, el pico mayoritario obtenido del intercambio aniónico correspondía a una fracción de CCDC28B exclusivamente monomérica, mientras que en las fracciones que eluyen a mayor concentración de sal, se encuentran también oligómeros de mayor tamaño. Este patrón de elución se repitió incluso cuando realizamos la SEC partiendo de fracciones de un intercambio aniónico previo, que contenían tanto monómero como oligómeros de CCDC28B: se obtenía nuevamente una fracción exclusivamente monomérica y luego fracciones conteniendo el monómero junto a oligómeros de mayor tamaño (ver Figura 28).

A partir de este resultado, y como el objetivo era tener la mayor cantidad de proteína monomérica posible para comenzar a realizar las pruebas de cristalización, se diseñó un nuevo protocolo. De las siguientes corridas de intercambio aniónico, se juntaron todas las fracciones donde eluyen los oligómeros de CCDC28B, y este material se agregó al material purificado por IMAC que se iba a fraccionar en el intercambio aniónico. De esta manera se consiguió aumentar la cantidad de CCDC28B monomérico de las siguientes corridas, y permitió mejorar levemente el rendimiento de la purificación, de 0,3 a 0,8 mg de proteína pura monomérica por litro de cultivo.



El intercambio aniónico permitió obtener una fracción de CCDC28B monomérica con un elevado grado de pureza con un rendimiento de 0,8 mg de proteína pura por litro de cultivo.

#### 4.2.D SEC del material purificado por IMAC

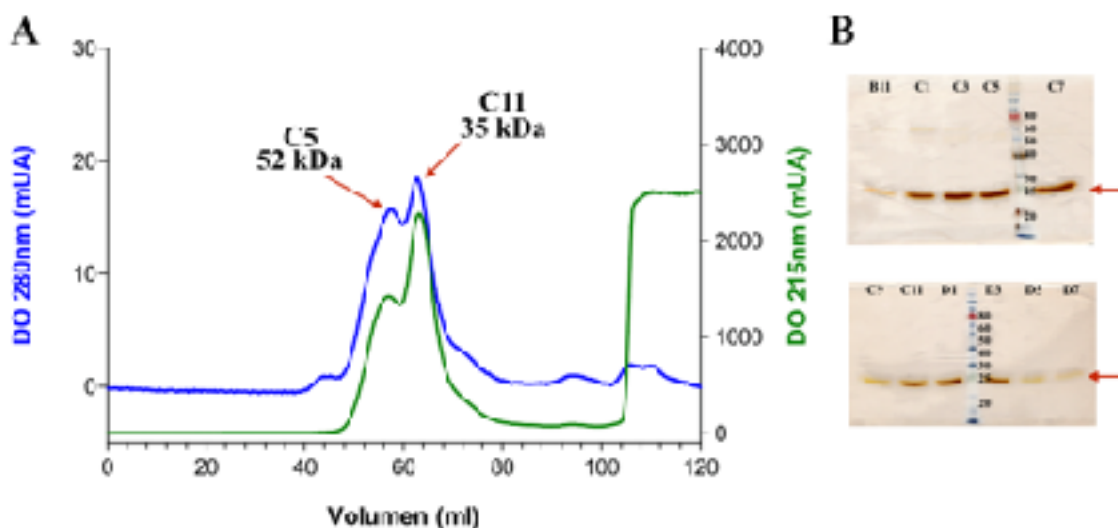
Teniendo en cuenta que en el intercambio aniónico CCDC28B mostraba una alta tendencia a formar agregados, se quiso investigar si los oligómeros estaban presentes en el material purificado por IMAC o se formaban durante la cromatografía sobre MonoQ como consecuencia del aumento de concentración que tiene lugar en el cargado de la columna. Para esto, se analizó por SEC el material purificado solamente por IMAC a partir de la fracción soluble correspondiente a un litro de cultivo bacteriano. El material obtenido (10 ml), enriquecido en CCDC28B, se concentró a la mitad para poder inyectarlo en una columna de SEC Superdex™ 75 16/600, en la que se puede inyectar hasta 5 ml de muestra.

El análisis de la corrida cromatográfica de exclusión molecular, que se observa en la Figura 29, evidencia un perfil con dos picos mayoritarios que eluyen a los 57 ml y 63 ml y que corresponderían a proteínas de peso molecular estimado en 52 kDa y 36 kDa, respectivamente. Recordando que el peso estimado de CCDC28B es de 22 kDa, estos volúmenes podrían corresponder a una fracción monomérica y una fracción dimérica.

Estos resultados indican que CCDC28B purificada luego de la IMAC está compuesta por dos poblaciones de proteína: una que correspondería a un monómero, y otra que posiblemente sea un dímero. Por lo tanto, los estados de agregación mayores observados en intercambio aniónico deben formarse durante la propia cromatografía.



El análisis por SEC de CCDC28B purificada por IMAC, demostró que la proteína se estaría comportando como una mezcla de dos poblaciones: una monomérica y una dimérica.



**Figura 29. Análisis por SEC del material purificado en IMAC.**

El material purificado por IMAC fue concentrado 2 veces e inyectado en una columna Superdex™ 75 16/600.

A- Cromatograma de la SEC. En azul se muestra la absorbancia a 280 nm, en verde a 215 nm. Las flechas indican el peso estimado de cada pico según la calibración de la columna.

B- SDS-PAGE 12% teñido con nitrato de plata de algunas fracciones de la SEC.

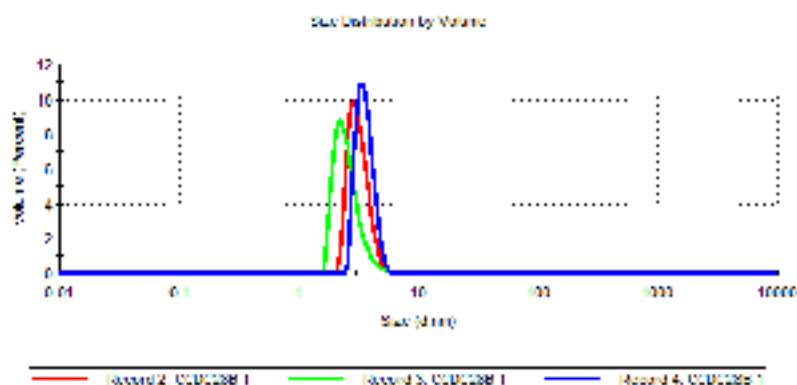
### 4.3. Caracterización de CCDC28B para la búsqueda de condiciones de cristalización

Luego de purificada la fracción monomérica de CCDC28B y como evaluación de calidad de la muestra en vistas a lanzar rastreos de condiciones de cristalogénesis, se analizó homogeneidad y estabilidad.

#### 4.3.A Análisis por Dispersión de la Luz Dinámica

Analizamos la fracción monomérica purificada, mediante DLS (del inglés: *Dynamic Light Scattering*). Esta aproximación permite estudiar el tamaño y homogeneidad de una muestra pura de proteína. Se basa en la determinación de la distribución de tamaños de las partículas de una solución: la muestra es iluminada por un láser y las fluctuaciones en la luz dispersada son detectadas a un ángulo fijo por un detector.

Los resultados de esta prueba no fueron concluyentes ya que se realizaron tres medidas diferentes que resultaron disímiles entre si. Una de las posibles causas de este resultado es que la muestra de CCDC28B sobre la que se realizó el análisis pudiera estar muy diluida. De todas formas como única conclusión del ensayo podemos ver que CCDC28B se comportó como una proteína homogénea ya que en las tres medidas aparece un solo pico, la cual sería de tamaño molecular aproximado al que corresponde al dímero (Figura 30), esto teniendo en cuenta las variantes que encontramos entre cada experimento.

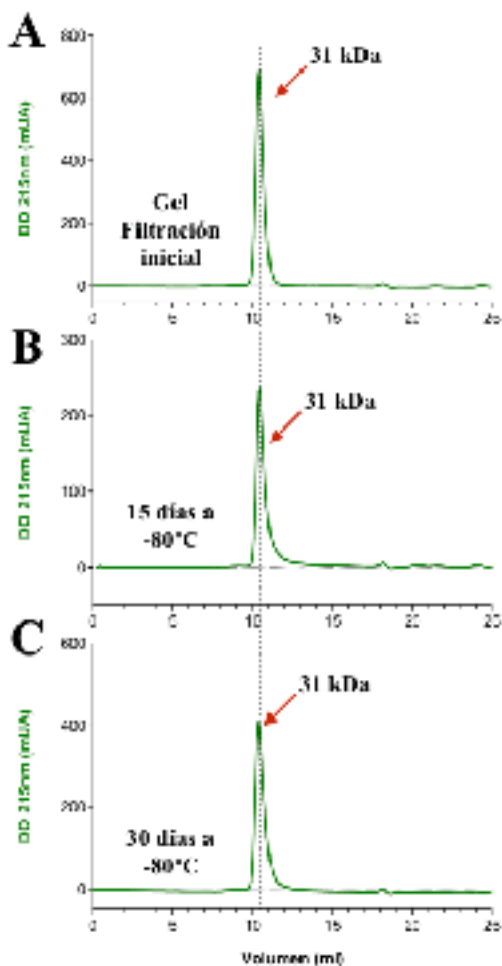


**Figura 30. Análisis por DLS de muestra purificada por IMAC e intercambio aniónico de CCDC28B.**

Se realizaron tres medidas diferentes de la misma muestra, observadas por diferentes colores. El hecho de que en los tres casos aparezca un solo pico puede ser síntoma de la homogeneidad de la muestra, pero no podemos ser contundentes en la conclusión ya que el hecho de que los tres picos sean diferentes habla de la variabilidad del experimento.

#### 4.3.B Estabilidad frente al congelado y descongelado a -80°C y la concentración

Estudiamos la estabilidad de la proteína en el tiempo y ante los procesos de congelado/descongelado y de concentración, ambos factores críticos a tener en cuenta dado los rendimientos de nuestra producción y los requerimientos del tratamiento para el proceso de cristalización. Las pruebas en búsqueda de condiciones de cristalización y luego el proceso de cristalización en sí mismo implicaban la necesidad de juntar distintas producciones, por lo que se debía almacenar



**Figura 31. Control de estabilidad de CCDC28B purificada luego de ser congelada a  $-80^{\circ}\text{C}$ .**

Cromatogramas de las diferentes cromatografías de exclusión molecular de las fracciones de CCDC28B utilizando una columna Superdex™ 75 10/300. Se observa únicamente la absorbancia a 215nm, marcada por la línea verde. Las flechas indican el peso estimado de cada uno de los picos, según la calibración de la columna. **A-** Corresponde a la Gel filtración en el día 0 previo al congelado. **B-** Luego de 15 días congelado a  $-80^{\circ}\text{C}$ . **C-** Luego de 30 días congelado a  $-80^{\circ}\text{C}$ .

parte del material a  $-80^{\circ}\text{C}$ . También se requería concentrar la muestra previo a los ensayos. De aquí nuestro interés en estudiar cómo se comporta la proteína en estos casos.

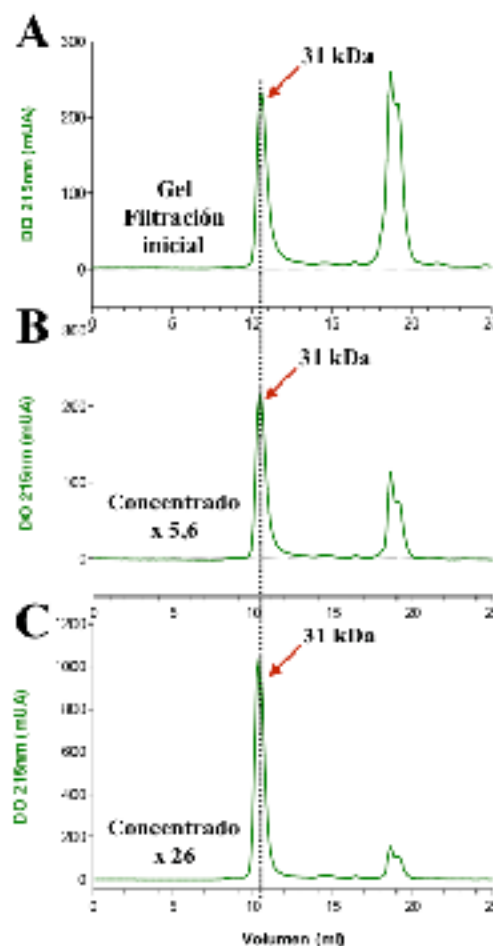
En primer lugar se generaron alícuotas de  $200\ \mu\text{l}$  de la proteína purificada para congelar a  $-80^{\circ}\text{C}$  y luego poder analizarlas por SEC cada 15 días. Es importante notar que luego de descongeladas, las fracciones se centrifugaron a velocidad máxima y no se evidenció la presencia de precipitado, lo que nos dio un primer indicio de que la proteína no se estaría agregando bajo estas condiciones. Luego, se realizó una SEC analítica inyectando los  $200\ \mu\text{l}$ . Esto se

realizó dos veces

con 15 días de diferencia entre cada SEC, para evaluar si el estado oligomérico de CCDC28B se veía alterado en ese período de un mes durante el cual la proteína se mantuvo congelada. Los resultados obtenidos, que se observan en la Figura 31, demuestran que el estado oligomérico de CCDC28B no sufrió modificaciones luego de un mes de congelado. El perfil de elución fue el mismo en los tres casos, con un único pico mayoritario correspondiente a un peso estimado de 31 kDa. De esta forma se pudo corroborar que CCDC28B se mantiene como una única población monomérica luego de un mes de almacenamiento a  $-80^{\circ}\text{C}$ .

**Figura 32. Control de estabilidad de CCDC28B purificada luego de ser concentrada.**

Cromatograma de las SEC utilizando una columna Superdex™ 75 10/300 de CCDC28B monomérica concentrada. Se observa sólo la absorbancia a 215 nm, marcada por la línea verde. Las flechas indican el peso estimado de cada uno de los picos, según la calibración de la columna. **A-** Corresponde a la Gel filtración previo al concentrado. **B-** Luego de concentrar 5,6 veces. **C-** Luego de concentrar 26 veces.





Por otro lado, un paso crítico para los estudios de solubilidad y cristalogénesis es la concentración de la muestra. Dado que teníamos evidencias de que CCDC28B tiene tendencias a agregarse, resultaba importante saber si la forma monomérica se mantenía como tal luego de concentrarla. Se concentró la proteína purificada 6 y 26 veces (correspondiendo a  $\sim 1$  y  $\sim 4,3$  mg/ml, respectivamente) por ultrafiltración utilizando los dispositivos Vivaspin™ 500 con una membrana de corte de 10 kDa. Luego las muestras concentradas se analizaron por SEC. Los resultados muestran que el estado de oligomerización de CCDC28B se mantiene incambiado luego de la concentración, comportándose en todos los casos como una única población monomérica (Figura 32).

Como conclusión a partir de los controles realizados se pudo establecer que la fracción monomérica de CCDC28B purificada por cromatografía de afinidad e intercambio aniónico se mantiene estable soportando el congelado y descongelado a  $-80^{\circ}\text{C}$  y su concentrado llegando hasta  $\sim 4,3$  mg/ml. Esto hace que sea un material óptimo para continuar trabajando de cara a las pruebas de cristalización.

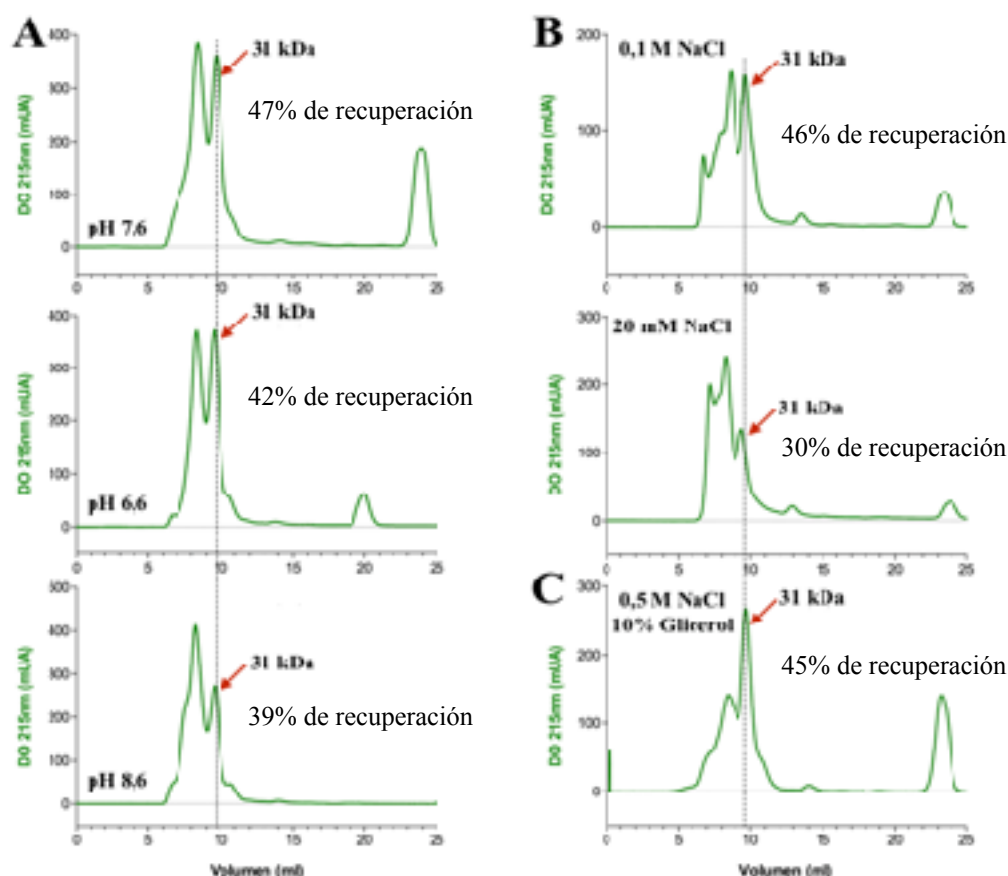


La fracción monomérica de CCDC28B purificada por IMAC y AEC demostró ser estable frente al congelado y descongelado a  $-80^{\circ}\text{C}$ , así como frente a su concentración hasta en 26 veces. En ambos casos la proteína mantiene su comportamiento como un monómero.

#### 4.3.C Pruebas de concentración en distintos buffers del material purificado por IMAC

A lo largo de la producción y purificación de CCDC28B recombinante hay varios pasos que implican la concentración de la muestra. Un resultado recurrente en este proceso fue la pérdida de parte sustancial del material de trabajo. Sobre el particular nos planteamos dos hipótesis: la primera consistió en que el material se perdía durante la concentración del mismo por pegado de la proteína a la membrana del dispositivo de ultrafiltración. Por lo tanto, se decidió utilizar otros dispositivos con filtros de diferente material: en particular se utilizaron los Vivaspin™ con filtro de polietersulfona y los Microsep™ Advance Centrifugal Device con una membrana Omega (polietersulfona modificada). En ambos casos sin embargo el resultado fue el mismo, perdiendo aproximadamente el 50% de la proteína durante la concentración.

La segunda posibilidad era que la proteína se estuviera precipitando cuando se la concentraba. En este sentido, se probó equilibrar y concentrar la muestra en diferentes buffers para evaluar si en alguno de ellos se evitaba la pérdida de proteína durante la concentración. Luego del procedimiento se estimó el rendimiento obtenido y el estado de agregación de la proteína mediante SEC. Con los diferentes buffers se evaluaron distintas posibilidades: en primer lugar, un cambio de pH manteniendo la concentración de sales; luego un cambio de fuerza iónica manteniendo el pH; y finalmente se probó el agregado de glicerol al buffer. Para estas pruebas se utilizó el material purificado luego de la IMAC que consistía en CCDC28B monomérica y dimérica. La muestra se



**Figura 33. Control de estabilidad de CCDC28B concentrada en distintos buffers.**

Se muestran los cromatogramas obtenidos al fraccionar por SEC la muestra de CCDC28B purificada por IMAC luego de ser equilibrada y concentrada en los buffer que se indican. En todos los cromatogramas se observa únicamente la absorbancia a 215 nm, marcada por la línea verde. Las flechas indican el peso estimado de cada uno de los picos, según la calibración de la columna. En cada cromatograma se indica además el porcentaje de recuperación de proteína luego de la concentración.

equilibró en el buffer a testear utilizando columnas HiTrap™ Desalting de 5ml, luego se concentró aproximadamente 20 veces, se determinó la concentración de proteína y se inyectó en la columna Superdex™ 75 10/300. La Figura 33 muestra los resultados de estas pruebas. Si bien en todos los casos se perdió material durante la concentración (el porcentaje de recuperación de proteína se indica junto a cada cromatograma en la Figura 33), se observaron algunos cambios en las proporciones monómero-dímero en los diferentes buffers. Se observó que para el caso de los distintos pHs, si bien no se aprecian cambios en la proporción monómero-dímero a pH 6,6 y 7,6, sí hay un cambio importante a pH 8,6 donde se observa que el equilibrio está volcado hacia la forma dimerica. Por otro lado, en la prueba donde se redujo la fuerza iónica del buffer (a 20 mM y 0,1 M de NaCl) el patrón de elución de la gel filtración cambió, apareciendo un pico correspondiente a un oligómero de mayor tamaño que el dímero sugiriendo una tendencia a la agregación. Finalmente, en la última prueba realizada se utilizó el buffer con el que se venía trabajando (20 mM Tris/HCl pH

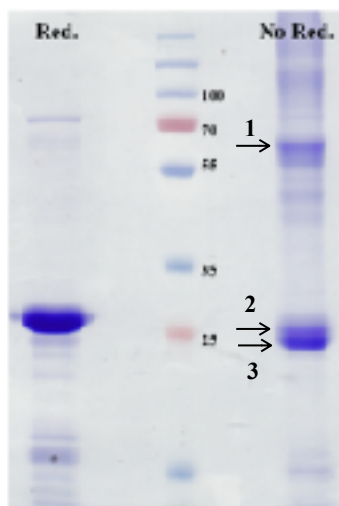


La prueba realizada con diferentes buffers arrojó como resultado que CCDC28B al ser concentrada en un buffer 20 mM Tris/HCl pH 7,6, 0,5 M NaCl y 10 % glicerol se comporta como una mezcla de dímero y monómero, con una mayor proporción del último.

7,6, 0,5 M NaCl) con el agregado de glicerol al 10 %. En este caso, se volvió a obtener el patrón inicial de dos poblaciones, una monomérica y otra dimérica, con la salvedad de que aquí se vio sensiblemente favorecida la población monomérica (segundo pico). Este resultado, mediante el cual se obtuvo una población de CCDC28B en un estado no agregado y particularmente enriquecida en su fracción monomérica, pareció ser una buena base para seguir adelante con las pruebas de cristalización y por lo tanto decidimos utilizar esta aproximación.

#### 4.3.D Estudio de CCDC28B recombinante en condiciones no reductoras

Para continuar caracterizando el material purificado por IMAC corrimos una muestra en SDS-PAGE en condiciones no-reductoras para evaluar la presencia de puentes disulfuro. En paralelo se corrió la misma muestra en condiciones reductoras, como control. Los resultados se observan en la Figura 34 y muestran que en condiciones no reductoras aparece una banda que no se observa en la muestra reducida, y que migra a la altura correspondiente al peso aproximado de un dímero de CCDC28B (55 kDa). Además, en condiciones no reductoras la banda que correspondería al monómero de CCDC28B migra como de menor tamaño que el monómero en condiciones reducidas, lo que podría sugerir la existencia de algún puente disulfuro intramolecular que le diera a la proteína una estructura más compacta.

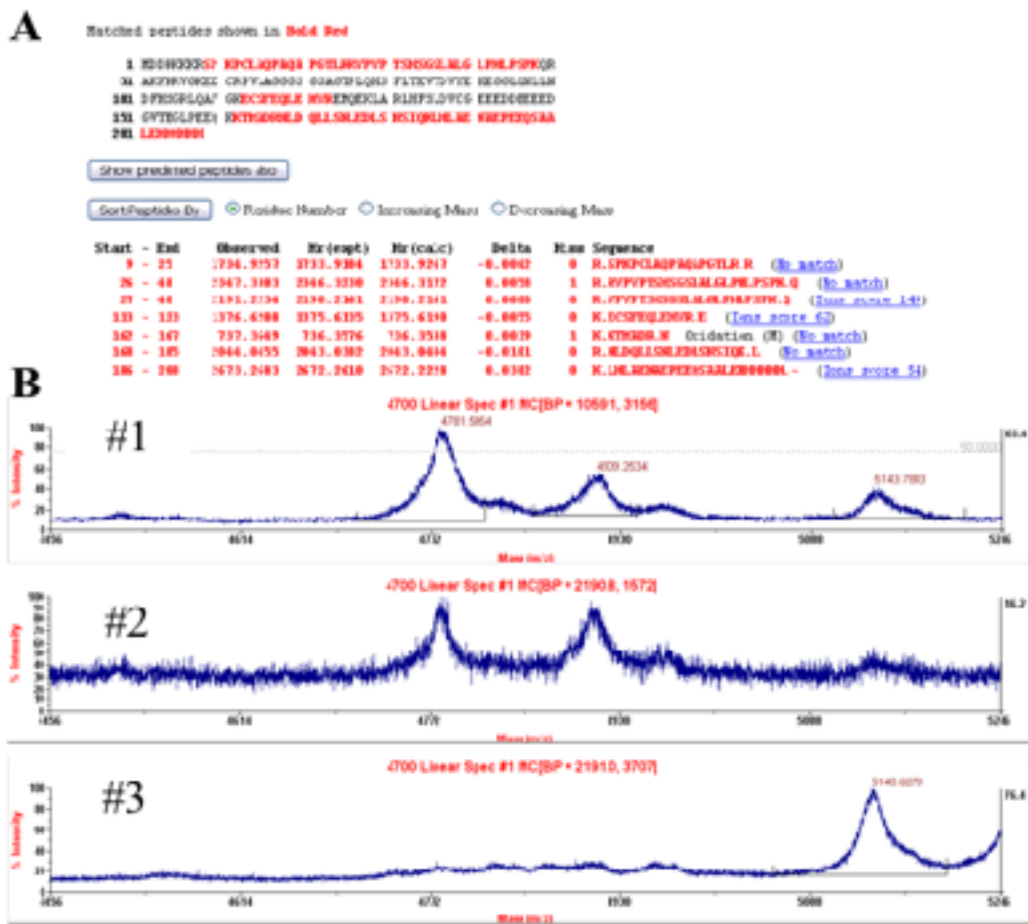


**Figura 34. Evaluación de CCDC28B en condiciones reductoras y no reductoras.**

SDS-PAGE 12% teñido con Azul de Coomassie donde se corrió la misma muestra de CCDC28B en condiciones reductoras (con  $\beta$ mercaptoetanol en el buffer de carga) y no reductoras (sin  $\beta$ mercaptoetanol en el buffer de carga). Con flechas negras se señalan las bandas cortadas para analizar por espectrometría de masas.

Para corroborar estos resultados, se cortaron las tres bandas presentes en el carril correspondiente a la muestra no reducida (a la altura correspondiente con la migración de CCDC28B monomérica, se pueden identificar dos bandas diferentes, ver flechas en la figura) y las mismas se analizaron por espectrometría de masa con la colaboración de la Unidad de Bioquímica y Proteómica Analíticas del IPMon. Los resultados mostraron que la banda de 55 kDa corresponde a CCDC28B, lo que estaría confirmando la formación de un dímero. Por otro lado en las otras dos bandas, que migran a la altura aproximada de 25 kDa se identificaron dos proteínas: CCDC28B y Peptidilpropil isomerasa de *E.coli*.

En la Figura 35 se observan parte de los resultados obtenidos por espectrometría de masas. En primer lugar la Sección A de la Figura 35 muestra la confirmación de la presencia de CCDC28B en la banda que migra a un peso de 55 kDa. Se muestran los péptidos trípticos identificados por MS y en algunos casos MS/MS, y a qué región de CCDC28B corresponden. En segundo lugar, se buscó la presencia de un puente disulfuro que pudiera explicar la dimerización de la proteína. En la Figura 35B se observan los espectros en modo lineal de las tres bandas recortadas del gel. Este modo permite identificar péptidos más grandes a los que normalmente se ven en el modo reflector, a costo



**Figura 35. Análisis por espectrometría de masa de las bandas observadas en SDS-PAGE en condiciones no reductoras de CCDC28B.**

**A-** Primer análisis de la fragmentación por tripsina de la banda de 55 kDa, confirma que la proteína es CCDC28B. Dentro de la secuencia aminoacídica de la proteína, se marcan en rojo los péptidos encontrados. En la parte inferior se observan los *scores* para cada péptido identificado teniendo en cuenta la masa esperada y la observada.

**B-** Espectros en modo lineal de las tres bandas (1, 2 y 3) observadas en el gel en condición no reductora. En las bandas #1 y #2 se observa un péptido de aproximadamente 4781 (m/z) que coincide con la masa esperada para dos péptidos unidos por un puente disulfuro. Esto fue confirmado por fragmentación MS/MS del correspondiente pico.

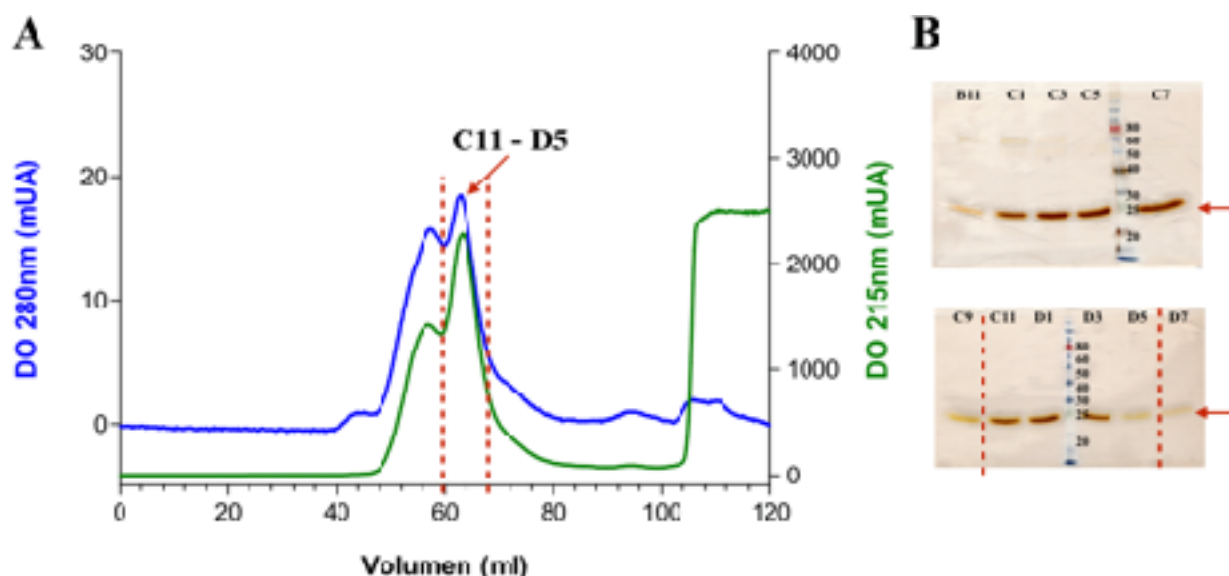
de una resolución un poco menor. Allí se observa una señal (m/z=4781) en las muestras #1 y #2, que correspondería a dos péptidos trípticos unidos por puentes disulfuro (113-EC<sup>S</sup>FEQLEJVR y 132-LHFSLDVCGEEEDDEEEEDGVTEGLPEEQK). La identidad de este péptido fue confirmada por fragmentación MS/MS. El hecho de que dicha señal también se observa en la muestra #2, que migra como monómero, puede indicar que dicho puente disulfuro se podría formar tanto intermolecularmente (#1) como intramolecularmente (#1 y #2).

#### 4.4 Prueba de solubilidad y búsqueda de condiciones de cristalización

A esta altura del trabajo se pensó que contábamos con material e información suficiente como para realizar una primera prueba de cristalización. La forma monomérica de CCDC28B parecía adecuada para este fin. Sin embargo, obtener el monómero puro por IMAC e intercambio aniónico era muy costoso debido al bajo rendimiento que se obtenía. Por lo tanto, se decidió fraccionar por SEC el material purificado por IMAC y utilizar las fracciones que correspondían a CCDC28B monomérica (Figura 36).

Esta muestra fue concentrada a 5 mg/ml y se sometió en primer lugar a una prueba de solubilidad para determinar si la concentración a la que se encontraba era compatible con la búsqueda de condiciones de cristalización. En la prueba de solubilidad se incubó la muestra en soluciones buffer con diferentes concentraciones de sulfato de amonio (1 a 2 M) y polietilenglicol (PEG) 4000 (15 a 30%), que se utilizaron como agentes precipitantes (ver Materiales y métodos), obteniéndose resultados positivos. Hubo algunas condiciones en las que la proteína precipitó (particularmente los buffers A1 -0,1 M Tris pH 8,5 y 2 M  $(\text{NH}_4)_2\text{SO}_4$ - y B1 -0,1 M Tris pH 8,5 y 1 M  $(\text{NH}_4)_2\text{SO}_4$ ), mientras que también se encontraron otras en las que la proteína se mantuvo soluble (A2 -0,1 M Tris pH 8,5; 0,2 M  $\text{MgCl}_2$  y 30 % PEG 4000- y B2 -0,1 M Tris pH 8,5; 0,2 M  $\text{MgCl}_2$  y 15 % PEG 4000-). Esto indicó que esa concentración de proteína era adecuada para realizar una prueba de cristalización con posibilidades de que alguna de las condiciones a testear funcionara.

A partir de este punto, lo que se realizó fue una primera prueba de cristalización utilizando kits comerciales, llegándose a probar 384 condiciones. Estas pruebas se realizaron utilizando la tecnología robótica de la Unidad de Cristalografía de Proteínas del IPMont en placas de 96 pocillos

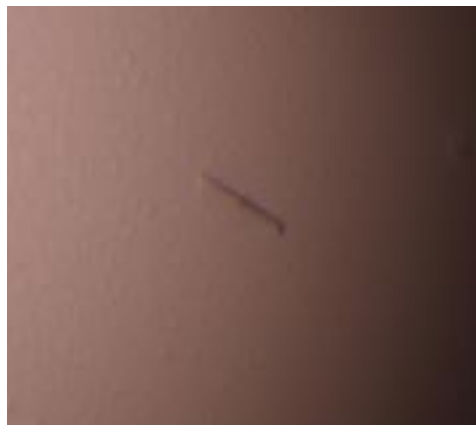


**Figura 36. Obtención de CCDC28B monomérica por SEC para utilizar en la prueba de cristalogénesis.**

A- Cromatograma obtenido utilizando una columna Superdex™ 75 16/600 donde se inyectó el material purificado de la IMAC. La línea azul corresponde a la absorbancia a 280 nm y la verde a 215 nm. Entre líneas puntadas rojas se marcan las fracciones que se seleccionaron para seguir trabajando, que pertenecen al pico al que le corresponde un peso molecular estimado de 31kDa.

B- SDS-PAGE al 12 % teñido con nitrato de plata en que se corrieron varias fracciones de la SEC y entre líneas punteadas rojas se indican algunas de las fracciones seleccionadas.

mediante el método de gota colgante que se detalla en Materiales y Métodos. Los kits utilizados en este caso fueron: PEGs Suite (96 condiciones), AmSO4 Suite (96 condiciones) y pHClear Suite (96 condiciones) de Qiagen y CrystalScreen (96 condiciones) de Hampton Research. El kit PEGs Suite utiliza PEG de distintos pesos moleculares dentro de sus 96 buffers de prueba, favoreciendo las interacciones proteína-proteína a través de la aglomeración molecular. Parecía una variante pertinente, ya que en la prueba de solubilidad, en los buffers que contenían PEG 4000 la muestra se mostró siempre clara y sin precipitar, por lo que una prueba utilizando PEG de otro peso molecular podía ser una aproximación interesante. Luego se utilizó el kit AmSO4 Suite, quizás el más prometedor, teniendo en cuenta los resultados de la prueba de cristalización, donde CCDC28B precipitó en los buffers con alta concentración de  $(\text{NH}_4)_2\text{SO}_4$ . Este kit nos provee por un lado de algunas soluciones de  $(\text{NH}_4)_2\text{SO}_4$  a distintos pH y por otro lado soluciones de  $(\text{NH}_4)_2\text{SO}_4$  en combinación con otras sales identificadas como buenos co-cristalizantes (McPherson A. y Cudney B., 2006). Por otro lado, el kit pHClear permite analizar un amplio espectro de cambios de pH en combinación con diferentes precipitantes (no solo  $(\text{NH}_4)_2\text{SO}_4$  y PEG, sino también MPD -2-Metil-2,4-pentanediol- y NaCl; todos en diferentes buffers). Finalmente el kit CrystalScreen, que al igual que el anterior también es de amplio espectro, prueba diferentes precipitantes en combinación con diferentes pH. Como ejemplo de las variantes entre estos dos últimos, el kit CrystalScreen



utiliza fosfato de amonio, acetato de sodio, sulfato de litio, 2-propanol, entre otros. En suma, la utilización de estos cuatro kits en su conjunto nos permitió probar 384 condiciones que abarcaron un amplio espectro dentro de las variables a modificar para encontrar una condición en la que cristalizara CCDC28B.

**Figura 37. Cristal obtenido en el primer screening de cristalización de CCDC28B.**

Cristal obtenido en el análisis con el material monomérico purificado por IMAC y SEC. La condición consiste en: 0,1 M Tris/HCl pH 8.5 y 2,0 M  $(\text{NH}_4)_2\text{SO}_4$ .

Los resultados obtenidos en este primer sondeo de cristalización arrojaron un único micro cristal. El mismo se muestra en la Figura 37 y la condición donde se encontró corresponde a 0,1 M Tris/HCl pH 8.5 y 2,0 M  $(\text{NH}_4)_2\text{SO}_4$ . El resto de los pocillos, mostraron o bien gotas claras, o algunas con material precipitado indicando un nivel excesivo de super-saturación o heterogeneidad en la muestra que impidió la formación de estructuras cristalinas. A partir de este punto se repitió la prueba de cristalización utilizando como punto de partida las condiciones en las que se obtuvo el micro cristal e intentando escalarla para la obtención de cristales de mayor tamaño. Para esto se utilizaron placas de 24 pocillos que fueron preparadas manualmente para hacer leves modificaciones con respecto a la condición de partida. Este ensayo también se basa en el método de gota colgante y el diseño de las placas se observa en la Figura 38. Como se ilustra allí, se intentó modificar sutilmente la condición en la que se encontró el primer micro cristal cambiando el pH (en las distintas filas) y la concentración de sulfato de amonio (en las diferentes columnas). Por último, la otra variable que introduce el cambio de escala es la cantidad de proteína que se necesita utilizar

	[ ] AmSO <sub>4</sub>						
		1,3	1,4	1,5	1,6	1,7	1,8
pH	7						
	7,5						
	8						
	8,5						

**Figura 38. Modelo de la placa de 24 pocillos en la que se intento optimizar la formación de cristales de CCDC28B.**

en cada gota, pasando de 300 nl en las pruebas de cristalización en placas de 96 pocillos a 2 µl en la optimización en placas de 24 pocillos.

Lamentablemente, este intento de mejorar las condiciones de cristalización no arrojó resultados positivos. Se observó la placa dos veces por semana

durante un mes, sin encontrar la presencia de ningún cristal en la placa.

#### 4.5. Estrategias para mejorar la cantidad/calidad de CCDC28B recombinante

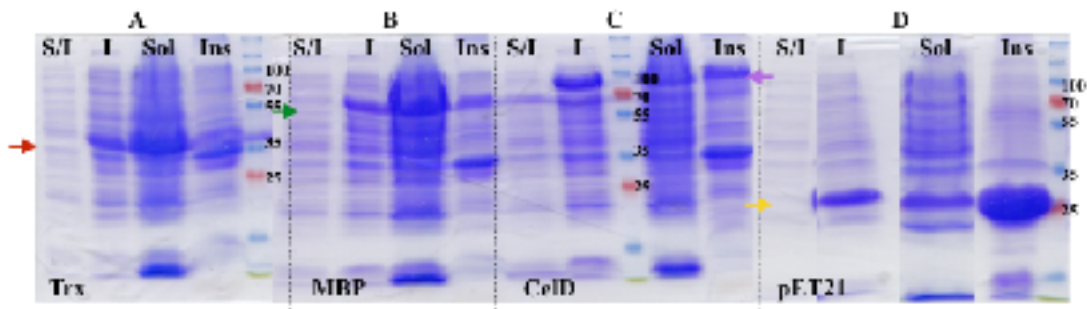
Uno de los cuellos de botella más importantes en proyectos de biología estructural y particularmente en la cristalografía de rayos X, es precisamente la formación de cristales para difractar. Esa dificultad puede ser intrínseca de la proteína o que la proteína recombinante no esté plegada correctamente causando una heterogeneidad conformacional en la muestra. Actualmente para tratar de solucionar este último problema, o al menos aumentar las probabilidades de encontrar un cristal, hay muchas variantes a realizar desde el punto de vista de ingeniería molecular (Derewenda Z., 2004; Correa A. y Oppezzo P., 2015). Luego de nuestros resultados adversos en la búsqueda y optimización de condiciones de cristalización, nos planteamos aplicar algunas de esas variantes en la producción de CCDC28B en un intento de lograr una mayor expresión soluble de la proteína así como de favorecer su plegamiento. Entre estas variantes nos planteamos: la expresión de CCDC28B en otro vector de clonado que permitiera cortar la etiqueta 6-His luego del primer paso de purificación, la co-expresión de CCDC28B junto a posibles interactores fisiológicos, y la expresión de CCDC28B como proteína de fusión a proteínas que colaboren en el plegamiento de la misma (Correa, A. et al 2014).

##### 4.5.A Expresión de CCDC28B fusionada a proteínas potenciadoras

Muchas proteínas, especialmente si son de origen heterólogo no logran plegarse de forma correcta cuando se expresan en grandes cantidades en *E.coli* (Balbás P., 2001) y una de las estrategias más utilizadas para evitar este problema es la expresión de la proteína de interés fusionada con proteínas accesorias que favorezcan el correcto plegado (Esposito D. y Chatterjee D., 2006). Con este objetivo, nos enfocamos en clonar CCDC28B en el *vector suite* diseñado en el laboratorio de la Unidad de Proteínas Recombinantes del IPMon (Correa, A. et al 2014). Esta metodología tiene como beneficio la posibilidad de clonar cualquier gen en 12 vectores diferentes utilizando los mismos dos cebadores. Los vectores están diseñados especialmente para favorecer la expresión soluble y el correcto plegamiento de la proteína de interés. Dentro de este amplio abanico de posibilidades que brinda el *vector suite* se eligieron cuatro de ellos, los que nos permitían expresar a CCDC28B como proteína de fusión a: Proteína de Unión a Maltosa (MBP), la Tiorredoxina A (Trx), la Isomerasa C Formadora de Puentes Disulfuro (DsbC) o la Endogluconasa CelD.

El clonado se realizó utilizando la técnica de RFC. Obtuvimos colonias para tres de las cuatro construcciones (Trx, CelD y MBP) y posteriormente se confirmó la correcta inserción y estructura de CCDC28B por secuenciación. Estos plásmidos se utilizaron para transformar bacterias *E.coli* Rosetta™(DE3)pLysS, y se realizó una prueba de producción siguiendo el mismo protocolo que veníamos utilizando para producir CCDC28B a partir de pET21d. En paralelo, para poder comparar niveles de producción y solubilidad y como control positivo, también se produjo CCDC28B a partir de la construcción en el vector pET21d. Se analizaron por SDS-PAGE las





**Figura 39. Control de inducción y solubilidad de las distintas construcciones de CCDC28B como proteína de fusión.**

SDS-PAGE al 12% con las fracciones sin inducir (S/I), inducido (I), soluble (Sol) e insoluble (Ins) de cada una de las construcciones. Con las flechas de colores se indican los tamaños esperados para cada construcción.

A - pT5-VS-CCDC28B-Trx (22 kDa + 11 kDa = 33 kDa)

B - pT5-VS-CCDC28B-MBP (22 kDa + 26 kDa = 48 kDa)

C - pT5-VS-CCDC28B-CelD (22 kDa + 70 kDa = 92 kDa)

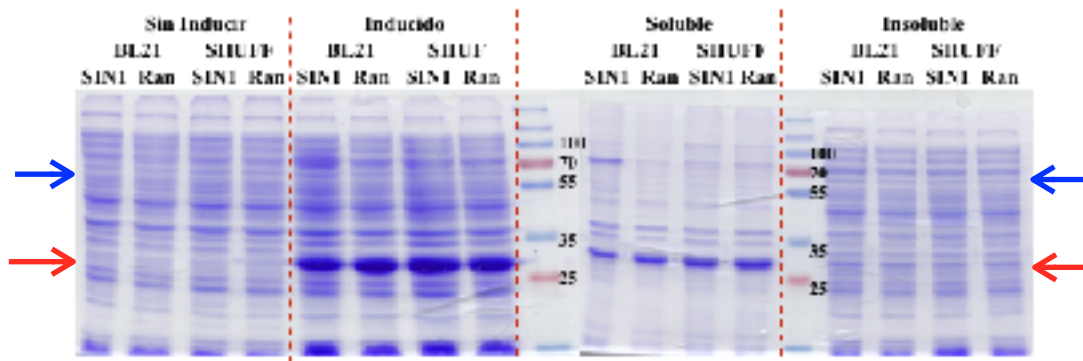
D - pET21d-CCDC28B (22 kDa)

fracciones totales sin inducir e inducidas, así como las fracciones solubles e insolubles obtenidas con los diferentes plásmidos. Los resultados mostraron que la proporción de material soluble para trabajar que se obtenía con las nuevas construcciones no era mayor a la que se obtenía en la producción de CCDC28B sola (ver Figura 39). Por lo tanto, la expresión de CCDC28B como proteína de fusión no resultó ser una alternativa de mejora en lo que respecta a aumentar la solubilidad de la proteína.

#### 4.5.B Co-expresión de CCDC28B con interactores

Otro de los abordajes realizados consistió en co-expresar CCDC28B junto a interactores fisiológicos. Esta es una estrategia muy utilizada a la hora de optimizar el proceso de cristalización (McPherson A. y Cudney B., 2006; McPherson A. y Cudney B., 2014). En nuestro caso, esta puede ser una ventaja importante ya que CCDC28B tiene su extremo N-terminal desestructurado según predicciones *in silico* (ver Introducción). Por ser una proteína pequeña y con una proporción importante sin una estructura 3D definida, no sería raro pensar que CCDC28B no cristalice sola, pues es difícil que moléculas móviles o sin estructura definida formen un sólido cristalino regular. Por lo tanto, el co-expresarla con un interactor fisiológico podría llevar a que la proteína se plegara correctamente reduciendo a su vez la movilidad de la región no estructurada de CCDC28B. En este sentido, decidimos co-expresarla con SIN1 y Ran ya que la interacción de ambas proteínas con CCDC28B fue previamente documentada en otros trabajos del laboratorio de Genética Molecular Humana (Cardenas-Rodríguez, 2013; y datos no publicados).

Para lograr la co-expresión se utilizó el plásmido pDUET donde se clonaron, nuevamente por RFC, los genes codificantes de las dos proteínas de interés. El clonado fue exitoso en todos los casos, verificándose las secuencias de los genes insertados por secuenciación. Luego se transformaron bacterias de dos cepas distintas de *E.coli*: BL21 y *Shuffle*. El cambio de cepa bacteriana se debió a que el plásmido pDUET tiene resistencia a cloramfenicol, y por lo tanto no



**Figura 40. Control de inducción y de solubilidad de las diferentes co-expresiones de CCDC28B con interactores fisiológicos.** SDS-PAGE al 12% con las fracciones sin inducir, inducida, insoluble y soluble de las coexpresiones de CCDC28B con los interactores fisiológicos: SIN1 y Ran. Se muestran los resultados de estas construcciones expresadas en dos cepas de *E.coli* diferentes: BL21 y Shuffle (SHUFF). Las flechas rojas marcan el tamaño esperado para CCDC28B y Ran, mientras que las flechas azules marcan el tamaño esperado para SIN1.

podíamos utilizar ninguna cepa con resistencia al mismo antibiótico, como es el caso de Rosetta™(DE3)pLysS.

Una vez transformadas las bacterias, se realizó una prueba de inducción utilizando el mismo protocolo que se utilizó con pET21d. Se corrieron las fracciones totales, solubles e insolubles de cada una de las co-expresiones en un SDS-PAGE (ver Figura 40). En el caso de la co-expresión con SIN1, no era claro que SIN1 se estuviera produciendo soluble (PM 59,1 kDa), ya que lo poco que parecía producirse se encontró en cuerpos de inclusión. Por otro lado, para la expresión junto a Ran, el SDS-PAGE teñido con Azul de Coomassie fue poco informativo, ya que CCDC28B y Ran poseen pesos moleculares similares (22 kDa y 24,4 kDa, respectivamente). Por lo tanto, si bien se observó una banda a la altura correspondiente, no se podía determinar a que proteína correspondía.

Por este motivo se realizó un *Western Blot* (WB) con anticuerpos contra SIN1, Ran y la etiqueta 6-His (todas las proteínas producidas estaban marcadas con 6-His). En primer lugar se intentó resolver si SIN1 se obtenía soluble. Para ello, se analizaron las fracciones solubles de las producciones en ambas cepas de *E.coli* por WB utilizando los anticuerpos anti-SIN1 y anti-His. La señal obtenida cuando revelamos con el anticuerpo anti-SIN1 mostró un patrón de bandas que parece específico, ya que no se observa en las muestras donde no se expresó SIN1, pero complejo, donde no es fácil identificar la banda correspondiente a SIN1 (Figura 41A). Pese a esto se observa que con este

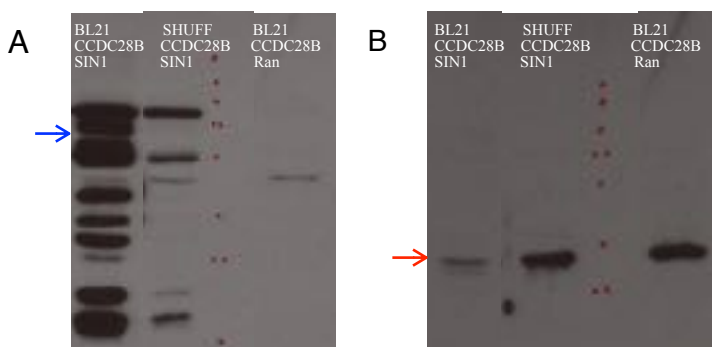
**Figura 41. Western Blot de coexpresiones de CCDC28B junto a los interactores fisiológicos SIN1 y Ran.**

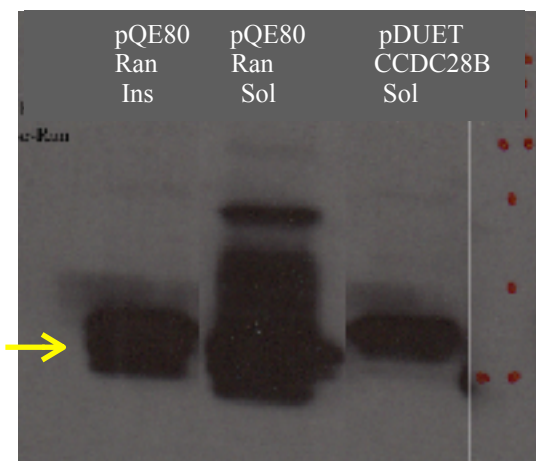
Se analizaron las coexpresiones de CCDC28B-SIN1 en BL21 y SHuffle; y se utilizó como control la coexpresión de CCDC28B con Ran en BL21.

**A** - Detección de Sin1 en los extractos solubles con un anticuerpo específico.

**B** - Detección de la etiqueta 6-His en las mismas muestras.

Con una flecha azul se señala el tamaño esperado para SIN1, que debería observarse en ambas figuras. Con una flecha roja se señala el tamaño esperado para CCDC28B y Ran, que debería observarse en la sección B, en los carriles que corresponde.





**Figura 42. Western Blot para la detección de Ran.**

Detección de Ran con anticuerpos específicos en fracciones insolubles y solubles obtenidas a partir de la expresión en pQE80 y fracción soluble de la co-expresión de Ran y CCDC28B a partir de pDUET. Con una flecha amarilla se señala el tamaño esperado para Ran

mismo anticuerpo, en los carriles correspondientes a los controles donde no se expresó SIN1, no hay ningún tipo de señal. Esto en principio nos dice que pese a no poder identificar la banda correspondiente a la proteína, la misma está allí ya que el anticuerpo la reconoce generando el patrón complejo. Por otro lado, en el WB

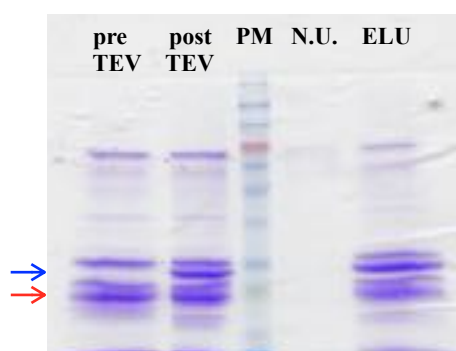
con el anticuerpo anti-His (Figura 41B) no se pudo observar ninguna banda a la altura esperada para el peso molecular de SIN1 (59,1 kDa), aunque si se observa la banda correspondiente a CCDC28B, indicando que el anticuerpo funciona correctamente. Es decir que observando los resultados del WB con el anticuerpo anti-His, no podemos afirmar que la proteína se exprese soluble.

Para evaluar la expresión de Ran se realizó un WB de las diferentes fracciones utilizando un anticuerpo específico (Figura 42). Los resultados indican que Ran se produce en forma soluble en las condiciones utilizadas. Pese a esto, si se observa en el gel teñido con Coomassie (Figura 40) la banda que correspondería a CCDC28B + Ran, no se ve más intensa que la banda donde se expresa solo CCDC28B, sugiriendo que la co-expresión con Ran no aumentaría sustancialmente la producción soluble de CCDC28B.

#### 4.5.C Corte de la etiqueta 6-His de CCDC28B

Finalmente, la otra prueba realizada consistió en eliminar la etiqueta 6-His previo a los ensayos de cristalización, si bien no hay una opinión consensuada sobre si la etiqueta puede afectar el plegamiento de la proteína y la cristalización de la misma (de hecho existen defensores de ambas posturas y ejemplos que respaldan ambas; por ejemplo ver Waugh D., 2005; Carson M. et al, 2007).

Se clonó el ORF de CCDC28B en el vector pQE80 (el vector pET21d no posee sitios de corte para proteasas entre el gen de la proteína y la etiqueta de 6-His), de forma de poder utilizar el sitio de corte para la proteasa TEV para separar la etiqueta 6-His una vez finalizada la IMAC. De esta manera, se produjo y purificó por IMAC CCDC28B a



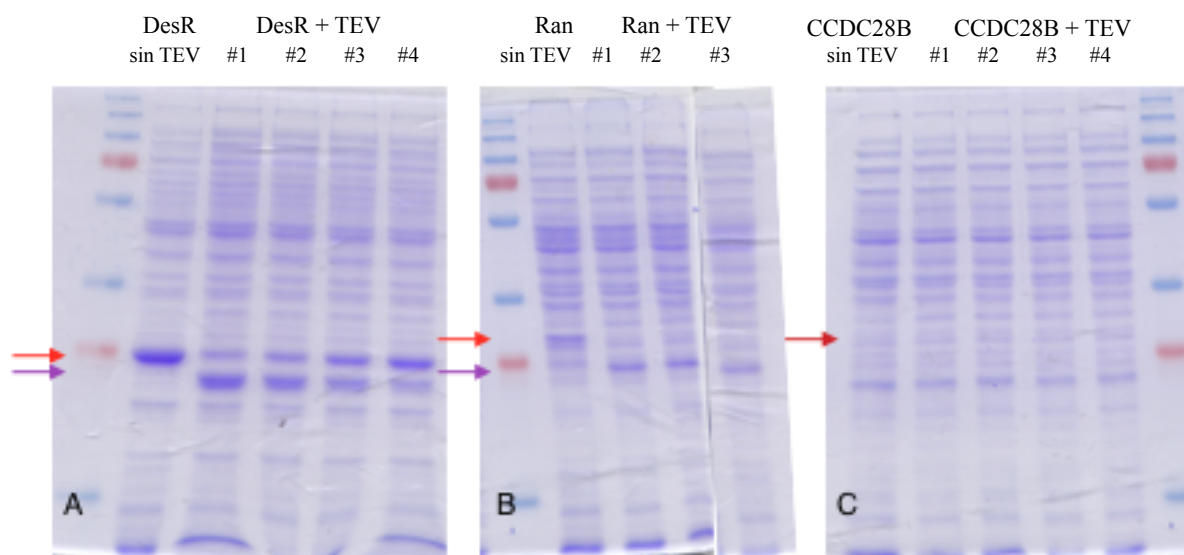
**Figura 43. Digestión de 6-His-CCDC28B con la proteasa TEV y posterior purificación por IMAC.**

CCDC28B purificada por IMAC fue incubada con TEV y luego de la digestión la muestra fue fraccionada por IMAC nuevamente para separar la proteína sin 6-His de la proteína sin cortar y de la propia TEV. SDS-PAGE 12% de las fracciones: pre (pre TEV) y post digestión (post TEV) y el material no unido a la IMAC post-digestión (N.U.) y eluido de la columna con 0,5 M Imidazol. Con una flecha azul se marca el peso molecular esperado para la proteasa TEV, y con una flecha roja se señala el peso molecular esperado para CCDC28B.

partir de pQE80 tal como se describió en Materiales y Métodos.

Luego se realizó la incubación del material purificado con la proteasa TEV, para después volver a purificar por IMAC el material digerido. En este caso, la proteína de interés debería salir en el material que no se une a la matriz, ya que no posee la etiqueta 6-His, mientras que TEV, que sí posee una etiqueta de 6-His, debería quedar retenida en la columna. Los resultados de esta prueba se muestran en la Figura 43 donde se observa que la proteasa no está actuando sobre CCDC28B. En el carril correspondiente al material no unido a la matriz (N.U.) no se observa proteína, lo que estaría hablando de que CCDC28B quedó unido a la columna, lo que es confirmado al observar el carril donde se sembró el material eluido. Allí se observa exactamente el mismo patrón previo a la digestión con TEV, además de la banda de la propia proteasa.

Dado este resultado, en primer lugar se verificó por secuenciación que la construcción pQE80-CCDC28B contuviera la secuencia de corte de TEV. Al confirmar que esto era así, se pensó que tal vez el sitio de corte quedara inaccesible a TEV en la proteína recombinante. Por lo tanto, realizamos una prueba utilizando diferentes concentraciones de un agente desnaturizante para desplegar parcialmente la estructura de CCDC28B y lograr que TEV acceda a su sitio de corte. Era importante utilizar en paralelo una proteína sobre la que se supiera que TEV funcionaba, para determinar la actividad de la proteasa en las diferentes condiciones. Para ello se utilizaron extractos solubles conteniendo las proteínas DesR (producido por Felipe Trajtenberg en la Unidad de Cristalografía de Proteínas) y Ran (producido por nosotros a partir de pQE80).



**Figura 44. Pruebas de digestión con la proteasa TEV en presencia de diferentes concentraciones de urea.**

Extractos solubles conteniendo las proteínas recombinantes DesR, Ran y CCDC28B fusionadas a 6-His fueron incubadas con la proteasa TEV en buffers conteniendo diferentes concentraciones de urea. Las flechas rojas marcan el tamaño esperado para cada proteína sin digerir y con flechas violetas el tamaño esperado para las proteínas digeridas.

**A** - SDS-PAGE 12 % de las muestras de DesR sin digerir y digerida en presencia de 4 concentraciones de urea. #1: 0,5 M urea; #2: 1 M urea; #3: 1,5 M urea; #4: 2 M urea.

**B** - SDS-PAGE 12 % de las muestras de Ran sin digerir y digeridas en presencia de las mismas 3 concentraciones de urea.

**C** - SDS-PAGE 12 % de las muestras de CCDC28B sin digerir y digeridas en presencia de las mismas 4 concentraciones de urea.

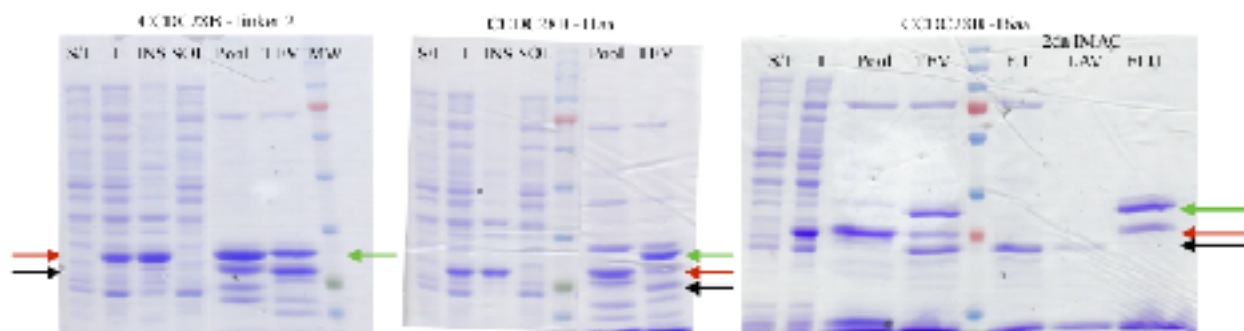
Como se indica en el protocolo, el buffer en el que normalmente incubamos la proteasa TEV fue: 50 mM Tris/HCl pH 8, 300 mM NaCl y 0,5  $\mu$ M DTT. Para estas pruebas agregamos a ese mismo buffer cuatro concentraciones diferentes de urea: 0,5; 1; 1,5 y 2 M. La Figura 44 muestra los resultados de esta nueva prueba de digestión con TEV. En todas las condiciones ensayadas se observan dos bandas de DesR, una superior coincidente con la proteína en ausencia de TEV y otra inferior que correspondería a la proteína sin la etiqueta 6-His (Figura 44A). Por otro lado, también se observa que a medida que aumenta la concentración de urea, la proteasa es menos eficiente para el corte aumentando por lo tanto la intensidad de la banda correspondiente a DesR-His y disminuyendo la de DesR cortada. En la sección B de la Figura 44 se observan los carriles que corresponden a la producción de Ran a partir de pQE80: la fracción soluble total y la incubada con TEV en los distintos buffers. Se observa el corte por parte de la proteasa, que se ve reflejado en el cambio del patrón de corrida de la banda correspondiente a la proteína. Por el contrario, no se observaron cambios en la migración de CCDC28B en ninguna de las condiciones analizadas (Figura 44C) indicando que TEV no cortó la etiqueta de 6-His en ninguna de las concentraciones de urea testeadas.



Los resultados de la digestión de CCDC28B con la proteasa TEV demuestran que en la proteína expresada a partir de pQE80 no se produce el corte de la etiqueta 6-His.

En conjunto los resultados mostraron que la proteasa TEV estaba activa y funcional pero que no era posible cortar la etiqueta de 6-His en CCDC28B posiblemente por alguna característica particular de la proteína que le está impidiendo reconocer o acceder a su sitio de corte.

Producto de los resultados adversos en los intentos de digerir 6-His-CCDC28B con TEV, y de nuestro interés en obtener la proteína sin la etiqueta 6-His, se propuso diseñar nuevas



**Figura 45. Control de expresión y digestión con proteasa TEV de las construcciones de CCDC28B diseñadas para cortar la etiqueta 6-His.**

SDS-PAGE 12 % teñido con azul de Coomassie de las fracciones correspondientes a la producción, purificación y digestión con TEV de las nuevas construcciones de CCDC28B. En todos los casos los carriles corresponden a: S/I- Fracción sin inducir; I- Fracción inducida; INS- Fracción insoluble; SOL- Fracción soluble; Pool- Fracción purificada por IMAC; TEV- Fracción purificada por IMAC e incubada posteriormente con TEV. La Sección A corresponde a la construcción CCDC28B-linker, la Sección B corresponde a la construcción CCDC28B-11aa y la Sección C corresponde a la construcción CCDC28B-16aa. En esta última se agregan los carriles correspondientes a la segunda purificación por IMAC luego de la incubación con TEV. En todos los casos la flecha roja corresponde al tamaño esperado para cada construcción sin digerir con TEV, la flecha negra el tamaño esperado para cada construcción digerida con TEV y la flecha verde el tamaño esperado para la proteasa TEV.

construcciones de CCDC28B que favorecieran la accesibilidad de la proteasa a su sitio de reconocimiento. Se diseñaron tres construcciones diferentes: una agregando un espaciador de tres aminoácidos (Ser-Gly-Ser) entre el sitio de reconocimiento de la proteasa TEV y el extremo N-terminal de CCDC28B (*linker*), y otras dos eliminando los primeros 11 y 16 aminoácidos de la proteína (-11aa y -16aa, respectivamente). Se utilizaron estas estrategias para mejorar la accesibilidad de la proteasa al sitio de corte de dos formas diferentes. En primer lugar el espaciador tuvo como objetivo alejar el sitio de corte de la proteína propiamente dicha. Esto es importante previendo que la estructura de la proteína esconda de alguna forma el sitio y la TEV no pueda acceder al mismo. Por otro lado, las deleciones de los primeros aminoácidos del extremo N-terminal, tuvieron por objetivo quitar una región de la proteína que se predice desestructurada. Al ser demasiado flexible esta región podría enmascarar el sitio de corte o impedir el acceso de TEV al mismo. Las construcciones se realizaron a través del diseño de nuevos cebadores. Para la eliminación de los primeros aminoácidos, se modificó la región de hibridación con el ORF de CCDC28B quitando la correspondiente a los primeros 11 o 16 aminoácidos según correspondiera. Y para obtener la construcción con espaciador, al cebador se agregaron nueve nucleótidos correspondiente a dos residuos de serina y uno de glicina que se insertaron entre el sitio de reconocimiento de TEV y el inicio del ORF de CCDC28B. Una vez realizados los clonados se confirmó la integridad de las construcciones por secuenciación.

Los protocolos de producción y purificación de las nuevas construcciones fueron los mismos que se detallaron anteriormente para CCDC28B en pET21d, teniendo en cuenta que solo se realizó la IMAC como paso de purificación. Luego de haber obtenido las fracciones enriquecidas en CCDC28B, las mismas se incubaron con la proteasa TEV y luego de la digestión se volvió a purificar por IMAC con el objetivo de separar la fracción cortada de la no cortada. Todas las fracciones de la producción, purificación y posterior digestión con TEV se analizaron por SDS-PAGE y los resultados se observan en la Figura 45.

Se observa que en dos de las tres construcciones (CCDC28B-*linker* y CCDC28B-11aa), los resultados no son claros, mientras que con la forma de CCDC28B carente de los primeros 16 aminoácidos (CCDC28B-16aa) la proteasa TEV si corta la etiqueta 6-His. En la sección A de la Figura 45 se observan las fracciones correspondientes a la construcción *linker*, allí en el carril correspondiente a la incubación con la proteasa TEV, parece difícil diferenciar la construcción en cuestión de la propia proteasa TEV, ya que ambas proteínas migran a un peso molecular similar (flechas negra y roja en la Figura 45A). Por otro lado en los carriles sembrados con las fracciones total (Pool) y digerida con TEV (TEV), se ve una banda por debajo de la correspondiente a CCDC28B, que puede ser un contaminante, pero que no permite diferenciar perfectamente las bandas correspondientes a CCDC28B con y sin 6-His. De todas formas, en la muestra incubada con TEV se ve un aumento de la intensidad de la banda que correspondería a la proteína digerida con TEV y una disminución en la intensidad de la banda superior (CCDC28B sin digerir y proteasa TEV). Este resultado nos hizo a pensar que en la construcción *linker* la TEV cortó la etiqueta. Para

corroborarlo purificamos la muestra nuevamente por IMAC, encontrando que la totalidad de la proteína se une a la columna y por ende mantiene su etiqueta 6-His (resultados no mostrados).

Por otro lado en la sección B de la Figura 45 se observan las fracciones correspondientes a la construcción -11aa. Los resultados allí son muy similares a los encontrados para *linker*, aunque en este caso parece ser más clara la acción de la proteasa. Allí se ve una disminución en la intensidad de la banda correspondiente a la proteína con etiqueta, a la vez que una nueva banda, de peso molecular inferior aparece en la muestra incubada con TEV, que podría corresponder a CCDC28B sin 6-His. De todas formas, cuando esta fracción se volvió a purificar por IMAC, se observó que toda la proteína queda unida a la columna, dejando dudas sobre si efectivamente se habría eliminado la etiqueta de 6-His. Finalmente, en la Figura 45C, se observan los resultados para la construcción -16aa donde si se observa la acción de TEV cortando la etiqueta 6-His. En el carril correspondiente a la incubación con la proteasa (TEV) se observa como la intensidad de la banda superior, correspondiente a la proteína con etiqueta disminuye y aparece una nueva banda de menor peso molecular que correspondería a la muestra cortada. Esta fracción digerida la purificamos nuevamente por IMAC y se vio como en la fracción que contiene el material no unido a la columna (N.U.) aparece la banda correspondiente a la proteína, por lo que TEV estaría cortando la etiqueta. Por otro lado en la fracción correspondiente a la elución final del material pegado a la IMAC (ELU) se ven dos bandas, una correspondiente a la propia proteasa, que se pega a la columna porque también se expresa con 6-His y otra correspondiente a una fracción de CCDC28B-16aa con la etiqueta, que TEV no cortó.

Con estos resultados podemos confirmar que encontramos una construcción de CCDC28B, sin sus 16 aminoácidos N-terminales, capaz de permitir el corte de su etiqueta 6-His por parte de la proteasa TEV. El hecho de poder generar una construcción de CCDC28B sin la etiqueta de 6-His y sin una porción de lo que se predice como su región desestructurada es positivo para continuar los estudios de cristalización e intentar obtener un cristal de CCDC28B.



La construcción de CCDC28B -16aa puede ser sustrato de la proteasa TEV para que la misma corte la etiqueta de 6-His. Siendo esta construcción por lo tanto interesante para repetir los ensayos de cristalización.

### **OBJETIVO 3: Identificar dominios o regiones de la proteína CCDC28B que sean importantes para su función.**

Hasta este momento, el trabajo estuvo centrado en la producción y caracterización de CCDC28B recombinante. En esta parte del trabajo, se pretendió aprender un poco más sobre aspectos fisiológicos de la proteína y comenzamos a trabajar en células humanas en cultivo.

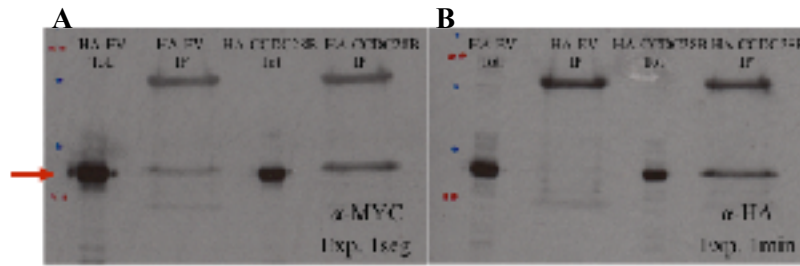
En primer lugar se buscó obtener información estructural sobre la proteína en células de mamífero, mientras que en la etapa final del trabajo, se intentó avanzar en la caracterización de regiones/motivos de CCDC28B que fueran importantes para su función. Para ello se trabajó en dos líneas celulares de humano: las HEK293FT (del inglés: *human embryonic kidney*) y las hTERT-RPE (del inglés: *retinal pigmented epithelial*). La funcionalidad de CCDC28B se evaluó a través de parámetros celulares que se sabe son afectados por CCDC28B: ciliogénesis (Cardenas-Rodríguez, M. et al 2013 -B) y la actividad del complejo mTORC2 (Cardenas-Rodríguez, M. et al 2013 -A), este último medido a través de la fosforilación de AKT en la Ser473. Luego de tener un buen sistema para evaluar el estado funcional de CCDC28B, se probaron mutantes de CCDC28B en regiones funcionalmente relevantes según predicciones *in silico*, con el objetivo de ver la relevancia de dichas regiones para la función de CCDC28B.

#### 4.6. Estudio del estado oligomérico de CCDC28B en células eucariotas

En las cromatografías de exclusión molecular realizadas con la proteína recombinante producida en *E.coli* se observó que la muestra eluía en dos poblaciones, una que correspondería a un monómero y otra a un dímero (ver sección 4.2.C). Posteriormente esto fue también observado por SDS-PAGE en condiciones no reductoras, donde también se observó la presencia de una fracción dimérica de CCDC28B (ver sección 4.3.B). En esta parte del trabajo se planteó obtener información sobre el estado oligomérico de CCDC28B en células humanas. En primer lugar se realizó una co-inmunoprecipitación (CoIp) a partir de un lisado de células HEK293FT transfectadas con HA-CCDC28B y myc-CCDC28B. Como control se utilizó un lisado de células transfectadas con el vector pCMV-HA vacío (EV, del inglés: *empty vector*) y myc-CCDC28B. La inmunoprecipitación se realizó utilizando un anticuerpo anti-HA comercial y una resina de agarosa acoplada a proteína G. El material precipitado se analizó por WB utilizando un anticuerpo anti-myc. A través de esta estrategia se buscó visualizar si en células eucariotas al menos dos moléculas de CCDC28B son capaces de interactuar entre sí.

En la Figura 46 se observa el resultado de la CoIp. La sección A, que muestra el WB revelado con anti-myc, indica que si bien se ve una banda tenue correspondiente a myc-CCDC28B en la CoIp control, probablemente debido al pegado inespecífico de la proteína a la matriz, hay una mayor presencia de myc-CCDC28B en la muestra en que se inmunoprecipitó HA-CCDC28B. Este resultado indica que habría interacción entre HA-CCDC28B con myc-CCDC28B y que por lo tanto





**Figura 46. Co-inmunoprecipitación de CCDC28B marcada con dos etiquetas diferentes.**

Se realizó la precipitación con HA y se evaluó por WB la presencia de myc-CCDC2B (en A) y luego como control la presencia de HA-CCDC28B (en B). La flecha roja indica el tamaño esperado para CCDC28B. Tot: extracto total, IP: fracción inmunoprecipitada.

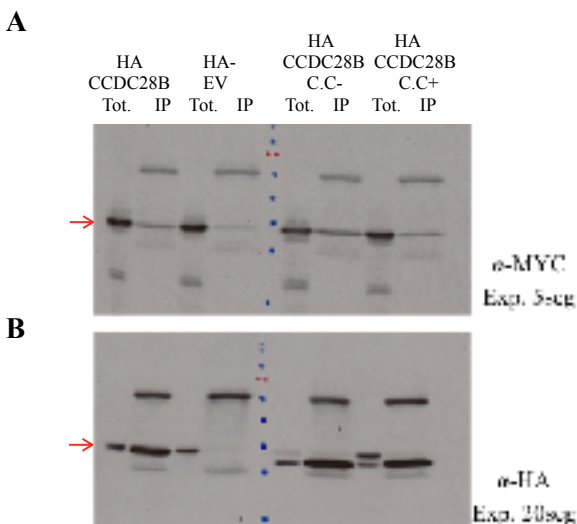
la proteína puede interactuar entre sí en células eucariotas. En la sección B de la Figura 46 se observa el WB revelado con anti-HA, que como control nos muestra la presencia de HA-CCDC28B solo en la muestra transfectada con dicho plásmido.

De esta forma, volvimos a obtener información que indica que CCDC28B estaría formando oligómeros también en células eucariotas. Sin embargo, es importante tomar en cuenta que en este caso estamos sobreexpresando la proteína, y la situación podría ser diferente con los niveles endógenos de CCDC28B.



El resultado de la CoIp de CCDC28B con dos etiquetas diferentes, junto a los obtenidos previamente en la cromatografía de exclusión molecular nos permiten afirmar que CCDC28B forma dímeros.

La información estructural que hasta el momento se tenía de la proteína (basada exclusivamente en predicciones *in silico*) hizo que desde un principio se hipotetizara que el dominio *coiled-coil* (correspondiente al extremo C-terminal de la proteína) fuera responsable de la dimerización o la oligomerización de la proteína. Para evaluar esta hipótesis repetimos la CoIp, pero esta vez analizando la interacción de myc-CCDC28B con dos mutantes de HA-CCDC28B en el dominio *coiled-coil*: una mutación que elimina la



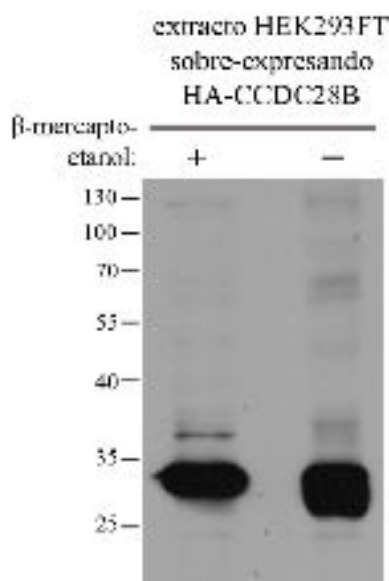
**Figura 47. Coimmunoprecipitación de CCDC28B y sus mutantes para el dominio *coiled-coil* marcadas con dos etiquetas diferentes.**

Se realizó la precipitación con HA y se evaluó por WB la presencia de MYC-CCDC2B o sus mutantes (en A) y luego como control la presencia de HA-CCDC28B o sus mutantes (en B). Las flechas rojas indican el tamaño esperado para CCDC28B.

predicción del mismo (C.C.-) y otra que no (C.C.+), utilizada como control. En la Figura 47 se observan los resultados de esta nueva CoIp, donde en primer lugar se confirma la interacción myc-CCDC28B con HA-CCDC28B. En el caso de los mutantes, y al contrario de lo que se esperaba, el mutante C.C.-, parece interactuar con CCDC28Bwt de la misma forma que el mutante C.C.+ o que la proteína WT.

Con estos resultados no podemos ser contundentes en nuestra conclusión, pero si nos da información particular sobre estos mutantes (C.C.+ y C.C.-): ninguno de los dos parece interferir en la interacción entre moléculas de CCDC28B para oligomerizar.

Por último, se tenía evidencia de que la proteína recombinante podía formar dímeros a través de puentes disulfuro (Figuras 34 y 35). Esto sería raro en células tratándose de una proteína citosólica, que por lo tanto se sintetiza y pliega en un ambiente altamente reductor. Por lo tanto, nos interesaba saber si los puentes disulfuro existían en condiciones fisiológicas o se estaban formando durante la producción/purificación de la proteína recombinante. Para ello se decidió realizar un SDS-PAGE en condiciones no reductoras seguido de WB. Si bien habría sido ideal analizar células que contuvieran CCDC28B a niveles endógenos, los anticuerpos disponibles contra CCDC28B reconocen otras proteínas en WB y se pensó que sería difícil identificar la o las bandas correspondientes a CCDC28B en condiciones no reductoras. Por lo tanto, se analizaron extractos celulares donde se había sobre-expresado HA-CCDC28B y se utilizó un anticuerpo anti-HA para identificar a CCDC28B.



El resultado obtenido mostró que en condiciones desnaturalizantes pero no reductoras CCDC28B se comporta como una única proteína cercana al peso molecular estimado para el monómero (Figura 48). Esto sugiere que en células eucariotas CCDC28B no formaría dímeros mediante puentes disulfuro.

**Figura 48. Evaluación del comportamiento de HA-CCDC28B sobreexpresada en HEK293FT en condiciones reductoras y no reductoras.**

Se realizó un experimento similar al mostrado en la Figura 34, pero esta vez utilizando la proteína expresada en células de mamífero y no en bacterias. Se detectó por WB utilizando un anticuerpo anti-HA la presencia de CCDC28B en condiciones reductoras (con β-mercaptoetanol) y no reductoras (sin β-mercaptoetanol).

#### 4.7 Identificación de dominios o regiones de CCDC28B que sean importantes para su función

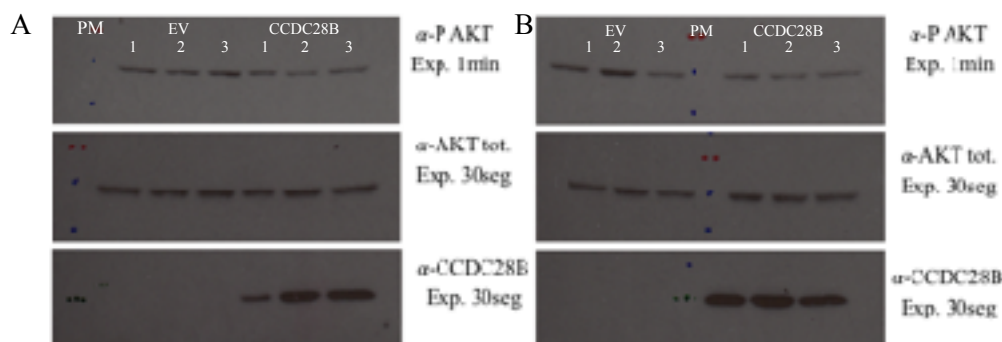
Dadas las dificultades que podían plantearse en la obtención de la estructura de CCDC28B, desde un principio se había propuesto un tercer objetivo que consistía en acercarnos a relaciones estructura-función en CCDC28B a través de analizar la relevancia de diferentes regiones de la proteína en sus funciones conocidas: modulación de la actividad de mTORC2 (Cardenas-Rodríguez, M. et al, 2013-A) y ciliogénesis (Cardenas-Rodríguez, M. et al, 2013-B). Como se mencionó en la introducción, CCDC28B presenta, de acuerdo a predicciones *in silico*, regiones o residuos que podrían estar relacionados a su función y/o regulación. De todos ellos, en esta parte de la tesis el foco se puso en regiones que podrían estar involucradas en determinar la localización subcelular de la proteína: el dominio de localización nuclear (NLS) y el de exportación nuclear (NES). Esta elección se debió a que trabajos del grupo habían mostrado que la entrada de CCDC28B al núcleo es importante, al menos, para su función en ciliogénesis (no se evaluó la actividad de mTORC2, datos no publicados). Los mutantes para eliminar estas regiones fueron construidos por Magdalena Cárdenas en su trabajo de posdoc en el laboratorio. En el mutante NLS se sustituyó un residuo básico característico y clave (KKKKRS) por alanina (KAKKRS). En el caso de las secuencias NES (EKLARLHFSLDVCGEE y EDLSNSIQKLHLAE) se sustituyeron los residuos L132 y L179 por alaninas (EKLARAHFSLDVCGEE y EDASNSIQKLHLAE). En ambos casos, las mutaciones introducidas eliminaban las predicciones de NLS y NES. Además, mediante fraccionamiento subcelular de células que sobre-expresaban los mutantes se observó que efectivamente las regiones eran funcionales, ya que el mutante NLS entraba menos a núcleo que la proteína WT, mientras que lo contrario sucedía con el mutante NES, tal como era de esperar (Novas, R., Cárdenas-Rodríguez, M. Fabregat, M. y cols, manuscrito en preparación). Los mutantes habían sido construidos en el vector pCS2+ y se sub-clonaron en pCMV-HA para poder expresarlos fusionados a una etiqueta contra la cuál se tenían anticuerpos.

##### 4.7.A. Estudio de funcionalidad de los mutantes de CCDC28B a través de su rol en la función de mTORC2

Una de las formas que utilizamos para analizar la funcionalidad de los mutantes de CCDC28B fue evaluando la actividad del complejo mTORC2, a través de medir la fosforilación de Akt en el residuo serina 473. La fosforilación de este residuo ocurre únicamente por la quinasa mTOR cuando forma parte del complejo mTORC2 (Hresko, R. C. and Mueckler, M., 2005). La actividad de CCDC28B sobre el complejo mTORC2 fue reportada por nuestro grupo (Cardenas-Rodriguez, M. et al 2013 -A) y sabemos que al disminuir los niveles de CCDC28B, disminuye el ensamblaje/estabilidad del complejo llevando a una disminución de su actividad. Este efecto es debido a la interacción de CCDC28B con SIN1, miembro del complejo mTORC2 (Cardenas-Rodriguez, M. et al 2013 -A). Así, la estrategia más sencilla era intentar aumentar la actividad del complejo al sobre-expresar CCDC28B WT, y ver si se obtenía el mismo efecto con los mutantes. En este tipo de ensayo es importante tener un buen porcentaje de transfección, para que el resultado

refleje lo que está sucediendo en las células transfectadas. Por esta razón se decidió utilizar células HEK 293FT, que son las que, en la experiencia del laboratorio, transfectan con mayor eficiencia.

Para evaluar si el sistema funcionaba en este tipo celular, se transfectaron por triplicado células HEK293FT con el vector pCMV-CCDC28B WT y pCMV-EV como control (protocolo detallado en Materiales y Métodos). Luego las células se cultivaron por 48 h en dos condiciones: medio de cultivo con suero fetal bovino (SFB) y medio sin SFB, para disminuir la actividad basal de mTORC2. En ausencia de SFB la actividad basal de mTORC2 debería ser menor, y tal vez esta condición permitiera un sistema más sensible para evaluar el aumento de la actividad por la sobreexpresión de CCDC28B. Luego se obtuvo un extracto celular en el que se analizaron los niveles de P-Ser437-Akt (sustrato de la actividad quinasa del complejo mTORC2) por WB utilizando anticuerpos específicos. En la Figura 49 se observan los resultados de los WB, donde la sección A corresponde a las células cultivadas sin SFB y la sección B corresponde a las cultivadas con SFB. Se verificó un claro aumento en la expresión de CCDC28B en las células transfectadas con pCMV-CCDC28B respecto a las transfectadas con pCMV-EV (ver el blot inferior, revelado con el anticuerpo anti-CCDC28B, donde la proteína endógena no es detectable). Sin embargo, al observar los niveles de P-Ser437-Akt, en ninguna condición se observa una diferencia entre las células control y las que sobre-expresan CCDC28B. El análisis de los niveles de Akt total en las muestras mostró que estos eran iguales en los dos tipos de células. Por lo tanto, en este tipo celular no podemos concluir que aumente la cantidad de P-Ser437-Akt, y por ende haya una mayor actividad del complejo mTORC2 en las células que sobre-expresan CCDC28B, lo que nos impidió utilizar este ensayo como reportero de la funcionalidad de CCDC28B.



**Figura 49. Evaluación del funcionamiento del complejo mTORC2, a través de la fosforilación de AKT en la Ser437.**

*Western Blot* de extractos de células HEK293FT transfectadas con pCMV-EV (control) o pCMV-CCDC28B para sobreexpresar la proteína. Se reveló con anticuerpos específicos contra: AKT fosforilada en la Ser437 (parte superior de la figura), AKT total como control (parte media de la figura), y CCDC28B como control de la sobreexpresión de la proteína (parte inferior de la figura). La sección A corresponde a las células incubadas sin SFB y la sección B a las células incubadas con SFB.

Una estrategia alternativa a la antes ensayada era disminuir los niveles de CCDC28B por ARN de interferencia (ARNi), lo que lleva a una disminución de la actividad mTORC2, y rescatar la actividad transfectando con pCMV-CCDC28B. Sin embargo, esta estrategia es técnicamente difícil ya que implica transfectar las células dos veces: en primera instancia con el ARN interferente y luego con el plásmido. La eficiencia de transfección con el plásmido en estos casos es muy baja y muchas veces las células no toleran bien el procedimiento. Dado que la técnica para seguir la

actividad de mTORC2 es WB, y por lo tanto no podemos separar las células transfectadas de las no-transfectadas, el sistema resultaría extremadamente poco sensible. La estrategia de la doble transfección tampoco funcionó cuando se evaluó ciliogénesis como reportero de la actividad de CCDC28B, ya que si las células no estaban saludables no ciliaban (datos no mostrados).



Al sobre-expresar CCDC28B en células HEK293, no logramos ver una sobreactivación del complejo mTORC2 mediante el aumento de P-Ser437-Akt, lo que nos impide utilizar este método para analizar los mutantes de la proteína.

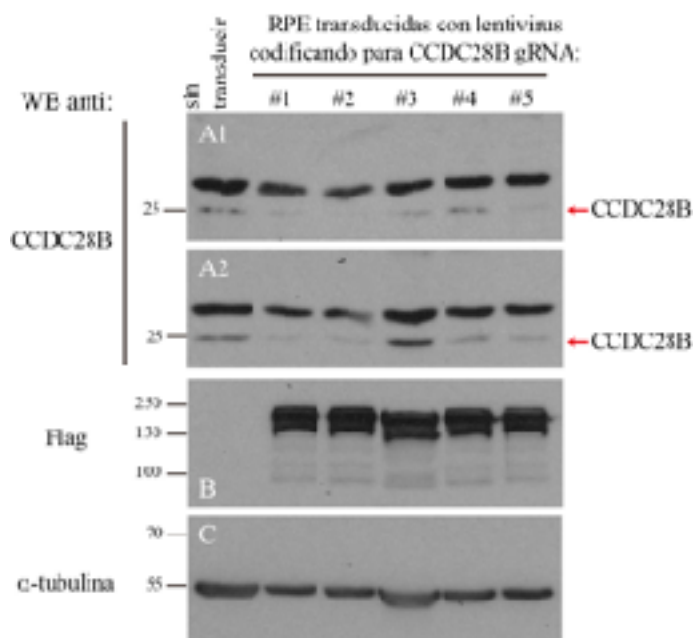
#### 4.7.B. Diseño de una línea celular CCDC28B-knockdown por tecnología de CRISPR/Cas9

Dadas las dificultades técnicas que existían en los ensayos planteados inicialmente para evaluar la funcionalidad de los mutantes, se decidió intentar obtener una línea celular estable carente (KO, del inglés: *knockout*) de CCDC28B o con niveles de proteína funcional claramente disminuidos (KD, del inglés: *knockdown*). En estas líneas se esperaba tener alterado en forma constitutiva alguno o ambos de los procesos donde se sabe CCDC28B participa: ciliogénesis y mTORC2. Por lo tanto, estas células podrían transfectarse con plásmidos para la expresión de CCDC28B WT o mutante y evaluar que sucedía con los fenotipos antes mencionados.

Para construir esta línea CCDC28B KO/KD se utilizó la tecnología de CRISPR/Cas9 (Shalem, Sanjana, *et al*, 2014). En una primera etapa se diseñaron varios ARN guía (ARNg) para dirigir a la enzima Cas9 a distintos lugares del gen *CCDC28B*, particularmente en los exones 2 y 3 (ver Materiales y Métodos). Estos ARNg fueron clonados en un vector de transferencia lentiviral que ya posee el gen de la enzima Cas9 (pLenti-CRISPRv2, Sanjana, N. E. *et al* 2014, Addgene). Esto permite introducir los genes y expresarlos en células hTERT-RPE utilizando lentivirus, lo cual tiene varias ventajas: i) puede lograrse un alto porcentaje de células infectadas y ii) los plásmidos lentivirales son integrativos, o sea que las secuencias codificantes del ARNg y la Cas9 quedan integrados al genoma de las células, lo que permite que ambos se expresen en forma permanente facilitando la alteración de los dos alelos de CCDC28B. Se eligieron las células hTERT-RPE porque son un modelo muy utilizado para estudiar el proceso de ciliogénesis.

Luego de obtener los lentivirus (ver Materiales y Métodos) que poseen los diferentes ARNg, se transdujeron células hTERT-RPE, y se cultivaron con el antibiótico correspondiente para seleccionar las células que integraron el plásmido al genoma. A diferentes tiempos se tomaron muestras de células para obtener extractos que permitieran analizar la expresión de CCDC28B y Cas9 por WB. La Figura 50 muestra los resultados obtenidos a los 3 (Figura 50A1) y 10 días (Figura 50A2) post-selección. Los resultados muestran que a los 3 días de selección, hay una disminución en los niveles de CCDC28B en las células que expresan los ARNg #1, #2 y #3, y las que fueron transducidas con los tres lentivirus al mismo tiempo. A los 10 días de selección, la disminución en la expresión de CCDC28B parece mantenerse con los ARNg #1 y #2, mientras que

con el #3 se recuperó la expresión de la proteína. Las células que expresan la mezcla de todos los ARNg, parecen también ir recuperando los niveles de CCDC28B. Por último, el ARNg #5 resultó ser el menos eficiente ya que no se observa una disminución en CCDC28B a ninguno de los tiempos analizados. De esta misma forma se analizaron las células a tiempos más largos post-transducción, y con todos los ARNg se observó la recuperación de los niveles de CCDC28B a



aquellos observados en las células control (datos no mostrados). Esta recuperación de CCDC28B, al menos a nivel de WB, ocurre a pesar de que la enzima Cas9 se sigue expresando a los tiempos analizados (Figura 50B).

**Figura 50. Evaluación de células hTERT-RPE CRISPR a expresando distintos ARNgüa durante varios días.**

*Western Blot* realizados sobre extractos de las células hTERT-RPE transducidas con lentivirus codificando para Flag-Cas9 y diferentes ARNg específicos para CCDC28B. A) Detección del niveles de CCDC28B a los 3 días (A1) y 10 días (A2) post-transducción. B) Detección de la expresión de Cas9 utilizando un anticuerpo anti-Flag y C) detección de alfa-tubulina como control de carga. En la Sección A se marcan con flechas rojas la banda correspondientes a CCDC28B.

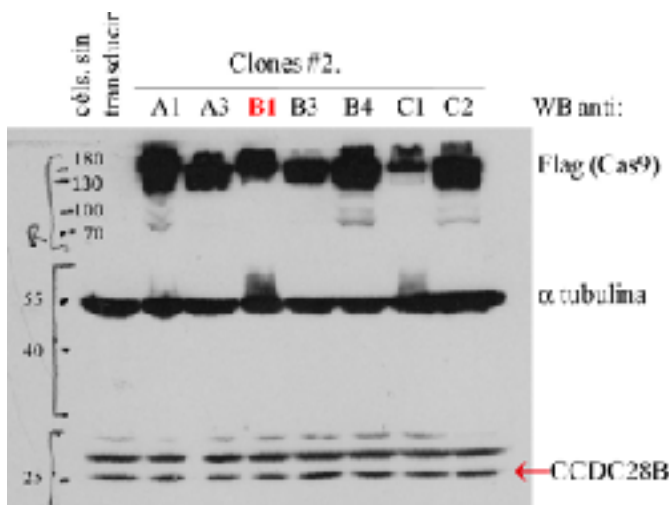
Esta recuperación de los niveles de CCDC28B podría ser explicado si la ausencia de la proteína fuera incompatible, o dificultara la proliferación celular. En este contexto, aquellas célula que tuvieran ausencia completa de la proteína podrían no ser viables. De esta manera, con el tiempo, el cultivo podría ser tomado por células en las que la modificación en el gen de CCDC28B diera lugar a proteínas ya sea funcionales o parcialmente funcionales. Debemos recordar en este punto que en esta primera instancia no se realizó clonado de las células luego de la transducción, por lo cual el cultivo es en realidad una mezcla de células donde la Cas9 puede haber introducido cambios muy diferentes en *CCDC28B* (deleciones, inserciones de distinto largo).

Para evitar este problema se comenzó el proceso nuevamente con la infección lentiviral, en este caso usando solo el ARNg #1. Las células fueron analizadas por citometría de flujo y utilizando el clasificador/separador celular, se realizó una deposición de una célula por pocillo en placas de 96 pocillos, para estudiar célula a célula y evitar la presencia de diferentes genotipos en un mismo cultivo. Las células resistentes al antibiótico de selección fueron amplificadas pasando a placas de 48, 24 y 6 pocillos sucesivamente y luego a placas de mayor área. Cuando fue posible se tomaron muestras para analizar la presencia de CCDC28B por WB y de cilia por inmunofluorescencia (IF). Es interesante notar que se observaron diferentes comportamientos durante el cultivo. Muchas células murieron, otras se dividieron algunas veces, luego dejaron de crecer y murieron. Otras, a pesar de no dividirse se mantuvieron viables durante un tiempo de 2-3 semanas, cambiando radicalmente su morfología, aumentado su tamaño a unas 10 veces el de una célula normal, en algunos casos presentando más de un núcleo y prolongaciones celulares, sugiriendo problemas de

citocinesis. Estos clones no pudieron ser analizados en profundidad ya que fue imposible generar una línea, por lo cual no se pudo saber que defectos portaban en *CCDC28B*.

De los 26 clones que se pudieron analizar, la gran mayoría presentó defectos marcados en ciliación, ya sea presentando un menor porcentaje de células ciliadas o cilias cortas en relación a células hTERT-RPE control. Sin embargo, cuando se evaluaron los niveles de *CCDC28B* por WB se observó que la intensidad y el tamaño de la banda correspondiente a *CCDC28B* era igual que en una célula control (Figura 51). Esto llamó especialmente la atención, ya que se esperaba una disminución del nivel de proteína o la generación de una forma de menor tamaño. Dado que observar una banda en WB no era indicativo de la existencia de una proteína funcional, se

seleccionaron clones que presentaban claros defectos en ciliación para evaluar la estructura génica de *CCDC28B*. Para ello se extrajo ADN genómico, se amplificó por PCR el exón 2 del gen de *CCDC28B* (donde se esperaba que Cas9 hubiera introducido

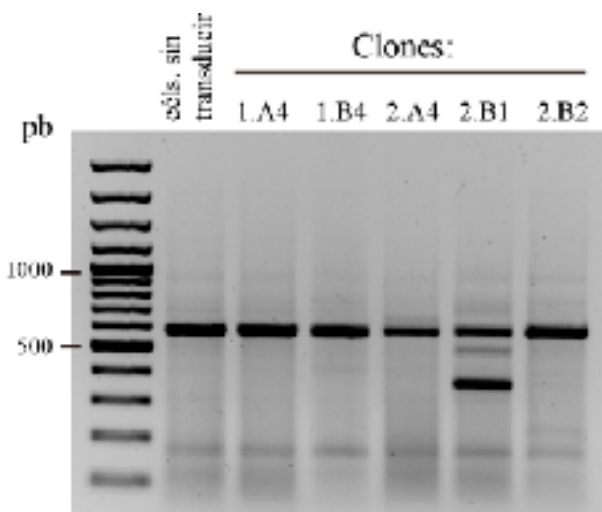


**Figura 51. Evaluación de la expresión de *CCDC28B* y Cas9 en diferentes clones.**

Los niveles de expresión de *CCDC28B*, Cas9 y tubulina (control de carga) fueron analizados por *Western Blot* utilizando anticuerpos específicos, tal como se realizó en la Figura 50. Con una flecha roja se señala la banda correspondiente a *CCDC28B*.

cambios) y el producto de PCR se clonó en el vector pGEM (Figura 52). Luego se seleccionaron varias colonias para secuenciar e intentar encontrar los dos alelos de *CCDC28B*. Los cinco clones analizados presentaban cambios en *CCDC28B*. En uno de clones, se observaron inserciones de dos y una base en cada uno de los alelos respectivamente (Clon 2.B2), lo que originaba un codón de terminación prematuro en ambos casos. Este clon fue uno de los que presentaba pocas células de gran tamaño y que no se dividían, por lo cual solo se pudo obtener ADN genómico antes de sacrificar el cultivo. En otro de los clones analizados se determinó la adición de tres pares de bases

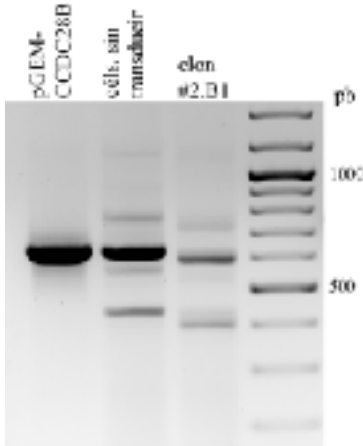
en cada uno de los alelos (Clon 2.A4), lo que origina una inserción de un aminoácido, sin cambio en el corrimiento del marco de lectura. En los tres últimos clones analizados se observó que en un alelo se adicionaba 1 o 2 pares de bases, dando lugar a un cambio en marco de lectura y un codón de terminación prematuro. En el otro alelo



**Figura 52. Amplificación por PCR del exón 2 de *CCDC28B* en diferentes clones.**

A partir de ADN genómico se amplificó por PCR el exón 2 de *CCDC28B* utilizando cebadores específicos y los productos de la reacción se corrieron en un gel de agarosa 1,5%.

se produjeron deleciones de 3 (Clon 1.A1), 12 (1.B4) y 200 (2.B1) nucleótidos que llevaban a una pérdida de 1, 4 o 31 aminoácidos, respectivamente. El caso del clon 2.B1, llamó la atención, que la pérdida de 31 aminoácidos no se reflejaba en una disminución del tamaño de CCDC28B por WB.



Por lo tanto, se purificó ARNm, se preparó cDNA y se amplificó el transcripto de CCDC28B. Cuando se corrió en un gel de agarosa se observó que la banda tenía un tamaño levemente menor que el del transcripto WT pero mayor al esperado para un alelo que codificara 31 aminoácidos menos (Figura 53). Cuando se obtuvo la secuencia

**Figura 53. Amplificación por PCR del ADN copia de CCDC28B en el clon 2.B1.**

Se preparó ADN copia a partir del ARN mensajero de las células correspondientes al clon 2B1 y células sin transducir como control. Luego se amplificó por PCR en ADNc de CCDC28B utilizando cebadores específicos. Como control positivo de la reacción se utilizó un plásmido pGEM que contiene el gen de CCDC28B. Todas las muestras se corrieron en un gel de agarosa 1,5%.

del transcripto mutante se observó que si bien se habían perdido 93 pares de bases del exón 2, se habían añadido 64 pares de bases provenientes del intrón 2, cambiando por tanto el patrón de

*splicing* y generando una proteína que tenía una secuencia diferente a la WT pero que era solo 11 aminoácidos más pequeña (Figura 54).



**Figura 54. Alineamiento de secuencia de CCDC28 WT y la proteína producida por el Clon 2.B1.**

En la Sección A de la figura se observan las secuencias aminoácidas de la proteína WT y la modificada en el clon 2.B1, deducidas a partir del secuenciado de ADN copia. En verde se marca la región modificada.

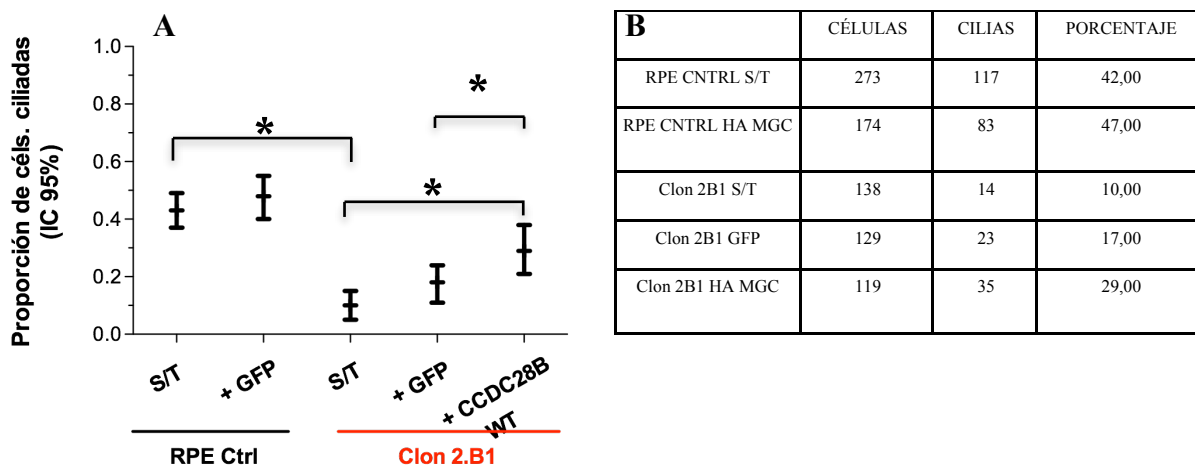
En la Sección B se observa el alineamiento de dichas secuencias, donde se observa que dentro de la región modificada aún existen algunos aminoácidos conservados.

Finalmente se terminó eligiendo este clon (#2.B1) para los estudios funcionales ya que se sabía que producía únicamente una proteína diferente a la WT y esto se reflejaba en un menor porcentaje de células ciliadas en el cultivo (Figura 55). Esto permitía tener un sistema reportero (porcentaje de ciliación) donde poder evaluar los mutantes de CCDC28B.

#### 4.7.C. Evaluación de la funcionalidad de diferentes mutantes de CCDC28B

En primer lugar se verificó que los problemas de ciliación que se observaron en el clon #2.B1 eran debidos a los cambios en *CCDC28B*. Para ello se transfeció el clon 2.B1 y células control con un plásmido para la expresión de CCDC28B WT o GFP como proteína control. Luego de 24 horas post-transfección se quitó el SFB al cultivo para estimular la ciliación, las células se fijaron 48 horas después, y se analizaron por IF.





**Figura 55. Cuantificación del porcentaje de ciliación en células control y el clon 2.B1.**

A. Proporción de células ciliadas expresadas como el intervalo de confianza 95%.

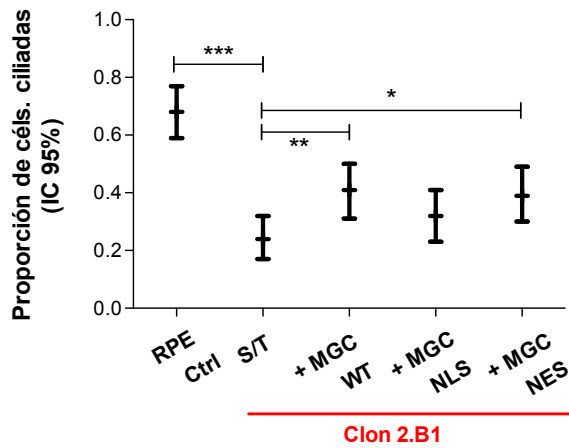
B. Tabla con los valores promedio de ciliación expresado en porcentaje.

En todos los casos  $*p \leq 0.01$

Los resultados de la cuantificación del porcentaje de ciliación se muestran en la Figura 55. Allí se observa en primer lugar que el clon #2.B1 posee un porcentaje de ciliación significativamente menor que las células hTERT-RPE control. Por otro lado, al comparar el clon #2.B.1 transfectado con los diferentes plásmidos se observó una clara diferencia: las células transfectadas con CCDC28B WT tienen estadísticamente más cilias que las células no transfectadas, o transfectadas con GFP, sugiriendo que la expresión de CCDC28 WT es capaz de recuperar la ciliogénesis en esta línea celular. Si bien el porcentaje de ciliogénesis con el clon #2.B1 transfectado con CCDC28B WT no alcanza el nivel de las células control, es posible que esto sea resultado de la forma en que se realiza el análisis, evaluando todas las células del cultivo y no únicamente las transfectadas. Esto podría resultar en que la recuperación en la ciliación inducida por la expresión de CCDC28B se viera enmascarada por una baja eficiencia de transfección. Si bien se intentó identificar las células transfectadas con HA-CCDC28B WT utilizando un anticuerpo anti-HA, esto no fue posible ya que no se pudo diferenciar con certeza las células transfectadas de las no-transfectadas, posiblemente debido a una baja expresión de la proteína.

En conclusión, el clon#2.B.1 tiene un defecto en ciliogénesis y dicho fenotipo se rescata al transfectar las células con CCDC28B WT. Por lo tanto, pudimos utilizar este sistema para evaluar la funcionalidad de los diferentes mutantes de CCDC28B.

Para evaluar los mutantes se plaquearon células hTERT-RPE y hTERT-RPE clon#2.B.1 en una placa de 6 pocillos para transfectar con plásmidos para la expresión de: GFP (control), HA-CCDC28B WT, HA-CCDC28B NLS (mutante en la señal de localización nuclear), HA-CCDC28B NES (mutante en la señal de exportación nuclear). El tratamiento de las células fue el mismo que en el experimento anterior, incubando por 24 horas luego de la transfección y luego 48 horas más sin SFB en el medio. La Figura 56 muestra los resultados del experimento donde se observa en primer lugar que los resultados anteriores se repiten, mostrando que las células WT tienen estadísticamente mayor ciliación que las del clon#2.B.1. Asimismo, en estas últimas aumenta el porcentaje de



**Figura 56. Cuantificación del porcentaje de ciliación de las células hTERT-RPE control y Clon #2.B1 transfectadas con los diferentes mutantes de CCDC28B.**

Proporción de células ciliadas en los diferentes cultivos, expresado por el intervalo de confianza (IC) 95%. Los asteriscos señalan las diferencias estadísticamente significativas. \* $p \leq 0.05$ , \*\* $p \leq 0.01$  y \*\*\* $p \leq 0.001$

ciliación al ser transfectadas con HA-CCDC28B WT. Por otro lado, la comparación de las células #2.B.1 al ser transfectadas con los distintos plásmidos muestra un aumento estadísticamente significativo de la ciliación en las células transfectadas con HA-CCDC28B WT y HA-CCDC28B NES, en relación a las células sin

transfectar. Por el contrario, las células transfectadas con HA-CCDC28B NLS no muestran un aumento significativo de la ciliación respecto de las células control.

Estos estudios preliminares parecen indicar que el tránsito de CCDC28B por el compartimento nuclear podría ser importante para su función en ciliogénesis. Será importante poder confirmar y expandir estos resultados preliminares. Es importante resaltar que esta línea celular KD para CCDC28B generada en este trabajo es una herramienta que nos servirá no solo para estudiar los mutantes NLS y NES sino un número de otras predicciones de dominios y residuos potencialmente importantes, como por ejemplo de fosforilación, que nos permita seguir entendiendo el rol de CCDC28B en ciliogénesis y los mecanismos que regulan su actividad.

## DISCUSIÓN

---

CCDC28B es una proteína que a nivel de secuencia no tiene homólogos de función conocida ni dominios funcionales predichos o posibles sitios activos reconocibles. Sin embargo, a través de ensayos de doble híbrido en levadura y posterior confirmación en células de mamífero, sí se han identificado un número de proteínas con las cuales interactúa, de alguna manera afectando su funcionalidad. Por ejemplo, sabemos que la interacción con SIN-1 modifica la estabilidad o ensamblado del complejo mTORC2 del cual SIN-1 forma parte (Cárdenas-Rodríguez, M. et al, 2013-A). También descubrimos que la interacción con SIN-1 media, al menos en parte, el rol de CCDC28B en ciliogénesis. Esto sucede independientemente de la actividad de SIN-1 en el complejo mTORC2. Respecto a este último punto sabemos que CCDC28B actúa río arriba a SIN-1 en el proceso de ciliogénesis (Cárdenas-Rodríguez, M. et al, 2013-A), pero aún no sabemos cómo SIN-1 participa en el proceso en sí. Por otro lado, también determinamos que CCDC28B interacciona con varias proteínas BBS, algunas de las cuales forman parte del BBSoma (Badano, J.L. et al, 2006), pero aún no conocemos las consecuencias funcionales de esa interacción. No sabemos si CCDC28B afecta el funcionamiento del BBSoma en células, a nivel de ciliogénesis o del transporte de proteínas de membrana a la cilia (una de las funciones principales del BBSoma), o si se afectan otras funciones del BBSoma que podrían ser independientes de la cilia (ref, review Rochi). Por último también sabemos por otros trabajos de nuestro grupo que CCDC28B interacciona con KLC-1, una de las cadenas livianas de la kinesina-1, pero en este caso, KLC-1 actuaría como intermediario entre CCDC28B y el motor de la kinesina para modificar la localización subcelular de CCDC28B (Novas, R, Cárdenas-Rodríguez, M., Fabregat, M. y cols, manuscrito en preparación). Debido a todo esto, la principal hipótesis manejada por nuestro grupo ha sido que la función de CCDC28B, más allá de poder tener una actividad directa, debe estar relacionada a modificar la actividad de otras proteínas con las cuales interactúa.

Sin embargo, en ninguno de los casos mencionados conocemos cómo CCDC28B interacciona con sus interactores. Por lo tanto, en esta tesis se propuso resolver la estructura tridimensional de la proteína como forma de aportar al entendimiento de su función. Ya se mencionó en la Introducción que al no existir proteínas homologas a nivel de secuencia, una alternativa posible puede ser evaluar la estructura, la cual sí puede mostrar homología con proteínas de función conocida que no se visualizan a nivel de secuencia primaria y por lo tanto sugerir qué regiones de la misma podrían estar interactuando con otras proteínas o ser funcionales en otro sentido.

Como estrategia para cumplir nuestro objetivo elegimos la producción de la proteína humana recombinante en bacterias seguida de la cristalización y difracción de rayos X. Sabíamos que esta estrategia podía presentar dificultades, ya que el análisis de la secuencia primaria de CCDC28B utilizando el PSIPRED (*Protein Sequence Analysis Workbench*) indicaba que

aproximadamente el 40 % de su secuencia puede tener una conformación desordenada, lo cual, como explicamos en la Introducción, puede dificultar la obtención de cristales. Sin embargo, se consideró que era importante intentar esta aproximación, máxime teniendo en cuenta la disponibilidad en nuestro instituto tanto de la tecnología necesaria como la experiencia de nuestros colaboradores. Lamentablemente el trabajo resultó más problemático de lo esperado, presentando dificultades en varios puntos de la estrategia experimental: baja producción de proteína recombinante soluble, tendencia de la proteína a agregarse y perderse en los pasos de concentración, y finalmente el fracaso en la obtención de cristales. Cabe destacar sin embargo que durante el proceso se aprendieron y descubrieron distintos aspectos de la proteína que no solo brindan pistas sobre como mejorar algunas etapas del proceso que podrían en el futuro permitir llegar a la obtención de cristales, sino que a su vez son importantes para continuar entendiendo su función. Estos puntos serán discutidos a continuación.

Por último, previendo que la elucidación estructural podía ser difícil, se planteó una estrategia complementaria para estudiar la relación estructura-función de CCDC28B analizando el papel de diferentes regiones de la proteína, con predicciones funcionales a partir de análisis *in silico*, en las actividades conocidas de CCDC28B, especialmente en su rol en ciliogénesis. Aunque aún queda mucho por explorar también con esta estrategia, en este trabajo de tesis logramos avances significativos poniendo a punto un sistema experimental que resultó adecuado para el propósito. Esto en parte involucró la obtención de una línea celular estable donde el gen de CCDC28B fue modificado por la tecnología CRISPR/Cas9 dando lugar a una proteína modificada de función disminuida, lo cual se refleja en un bajo porcentaje de células ciliadas. Creemos que esta línea celular nos permitirá, no solo testear mutantes de CCDC28B para determinar la relevancia de las regiones modificadas en la actividad de la proteína, sino estudiar el rol de CCDC28B en otros procesos, como en el ensamblado u actividad del BBSoma.

### **Lo que aprendimos sobre CCDC28B recombinante**

Durante la producción de la proteína recombinante se utilizaron diferentes vectores de expresión, cepas de bacterias, formas de inducir la expresión de CCDC28B y se intentó la expresión como fusión a diferentes proteínas potenciadoras. En todos los casos se observó que un gran porcentaje de la proteína que se produce va a cuerpos de inclusión, lo que sugiere un plegamiento incorrecto o tendencia a la agregación. También en este sentido, observamos que tanto en la purificación por IMAC así como en muchas fracciones del intercambio aniónico, co-purificamos una chaperona de *E. coli* de aproximadamente 70 kDa (identificada por espectrometría de masa como *molecular chaperone DnaK*, resultados no mostrados). Esto podría sugerir que al menos parte de la proteína aún no se ha terminado de plegar correctamente y por eso permanece unida a la chaperona.

Se observó que la proteína soluble luego del primer paso de purificación se comporta como una mezcla de monómero y dímero. La proporción de estas dos formas oligoméricas cambia ligeramente al disminuir la fuerza iónica, favoreciéndose la formación del dímero e incluso alguna forma de mayor tamaño. Ya se ha reportado para diversas proteínas que se sabe son funcionales en estados oligoméricos, que el aumento de concentración de sal favorece la despolimerización a estados monoméricos (Marcum, J. M. and Borisy, G., 1978; Wagner, R. et al, 1987; Kadima, W. et al, 1993). Este último dato no es menor, ya que tanto la IMAC como el estudio de estabilidad de la forma monomérica se realizaron a una alta concentración de sal (0.5 M NaCl). Esto podría indicar un efecto encubierto de la alta fuerza iónica del medio, que estuviera estabilizando a la proteína como monómero, cuando en realidad esta no sea la forma fisiológica de la misma.

La proteína tiene cuatro cisteínas y si bien no esperamos la formación de puentes disulfuro en condiciones intracelulares (Saaranen, M.J & Ruddock, L.W, 2013), es importante destacar que no se usaron reductores durante los pasos de obtención de los extractos solubles y la purificación, por lo que podrían haberse formado puentes disulfuro *ex-vivo*. Para determinar si el dímero observado en la gel filtración luego de la IMAC estaba mediado por puentes disulfuro, deberíamos haber reducido y alquilado la proteína, para luego repetir la exclusión molecular y ver si se seguía observando el pico correspondiente al dímero. No lo hicimos, y como alternativa realizamos un SDS-PAGE en condiciones no reductoras. Así, se confirmó que el material purificado de la IMAC en ausencia de reductores contenía dos bandas, a 55 y 25 kDa, y se confirmó por espectrometría de masa que CCDC28B estaba presente en ambas bandas, indicando la presencia de dímeros por puentes disulfuro en la banda de 55 kDa. Por lo tanto, parte de la proteína que estamos produciendo se está agregando a través de puentes disulfuro, algo que no parece ser fisiológico, por la localización predicha de la proteína y por resultados que obtuvimos analizando la proteína expresada en células eucariotas, como se discutirá más adelante. Por lo tanto, sería importante para las próximas producciones de proteína recombinante incluir un reductor para impedir la formación de puentes disulfuro. Esto podría permitir obtener una población más homogénea de proteína, lo cual sería beneficioso a la hora de cristalizar.

Otra cosa llamativa fue el comportamiento de la proteína durante el intercambio aniónico. Esto se puede observar en los resultados inusuales de este segundo paso de purificación, en el cual vemos la elución de CCDC28B a lo largo de todo el gradiente salino. Luego determinamos que este comportamiento es ocasionado porque CCDC28B se agrega durante la cromatografía probablemente debido al proceso de concentración que sufre al cargarse en la columna, formando oligómeros de diferente tamaño que se eluyen a diferentes concentraciones de sal. Lo llamativo fue determinar luego que la forma monomérica es sumamente estable, incluso cuando la proteína se concentra casi 30 veces. ¿Cómo podría explicarse entonces la formación de agregados mayores? Pensamos que en el material que se obtiene luego de la IMAC debe existir una fracción de la proteína mal plegada, que se agrega al concentrarse en la columna de intercambio aniónico, mientras que la forma monomérica se pliega de una forma estable.

Este comportamiento de la proteína recombinante, con tendencia a formar oligómeros probablemente por no estar del todo bien plegada no presagiaba buenos resultados a la hora de obtener cristales. De allí a que se intentaran estrategias alternativas para producir la proteína recombinante. Se buscó en primer lugar aumentar la cantidad de proteína soluble que se producía, pensando que esto podía indicar una mayor proporción de proteína bien plegada. Asimismo, analizar la proporción de proteína soluble era un método de evaluación sencillo y rápido a la hora de ensayar diferentes estrategias.. A pesar de que ninguna de las estrategias (expresión como fusión a una proteína potenciadora, co-expresión con interactores) aumentó significativamente la proporción de proteína que se producía soluble y fue descartada en el trabajo de esta tesis, creemos que varias de ellas plantean opciones interesantes para continuar el trabajo a futuro.

1) La expresión como proteína de fusión a las diferentes proteínas potenciadoras a pesar de que no aumentaba la cantidad de proteína soluble, podría igualmente favorecer el correcto plegamiento de CCDC28B, lo que podría hacer valer la pena un estudio más detallado de las proteínas producidas. Como perspectiva usando esta aproximación nos planteamos enfocarnos en una de las fusiones, ensayadas hacia el final de la tesis y que no fue mostrada en la sección de resultados: la fusión CelD-CCDC28B- con un espaciador rígido de 10 aminoácidos (GAEAAAKAGG) entre ambas proteínas. Ya probamos una construcción similar con un espaciador flexible entre ambas proteínas y observamos que luego de la purificación por IMAC la fusión se cortaba, obteniendo un bajo rendimiento de CCDC28B. Por el contrario, tenemos resultados preliminares donde encontramos que la fusión con el espaciador rígido se expresa muy bien soluble (en cuanto a cantidad total, ya que la relación soluble/insoluble fue muy similar a la encontrada en todos los demás casos). Esto podría ser muy beneficioso ya que una alternativa podría ser la cristalización de la fusión CelD-CCDC28B, donde no solo nos apoyaríamos en el facilidad de cristalización de CelD, sino que también al ya estar resuelta su estructura tridimensional (Chauvaux, S. et al, 1995), nos serviría para facilitar la interpretación de los datos para resolver la estructura de CCDC28B, realizando el remplazo molecular en la parte de la fusión correspondiente a CelD.

2) La co-expresión junto con interactores partía de la base de que al producirse las dos proteínas al mismo momento, la unión entre sí podría contribuir a estabilizar a CCDC28B. Sin embargo, es de esperar que CCDC28B se una a sus interactores una vez plegada en forma correcta, o sea que la co-expresión no contribuiría al plegamiento en sí sino a la estabilidad de la proteína una vez plegada. impidiendo que se agregue y precipite. De todas formas, la formación de un complejo podría estabilizar a la proteína impidiendo que se agregue y precipite. Pero aún más importante sería el hecho de: a) poder obtener cristales de un complejo, lo cual sería sumamente informativo para entender la interacción de CCDC28B con otras proteínas; b) la unión a un interacto podría inmovilizar la región N-terminal de CCDC28B, predicha como desestructurada y que podría ser flexible, dificultando la obtención de cristales y c) al igual que se discutió para la fusión CCDC28B-CelD, el hecho de poder cristalizar la proteína como fusión o en complejo, facilitaría la resolución de la estructura, siempre y cuando la estructura del interacto ya esté resuelta, realizando un

reemplazo molecular en esa sección de la estructura. En el caso de SIN1 y Ran, dos de los interactores conocidos, ambas estructuras están resueltas (Vetter, I. R. et al 1999; Pan, D. and Matsuura, Y., 2012) por lo que estaríamos ante una situación favorable. Sin embargo, la co-expresión con SIN 1 tuvo la dificultad de que esta última se produce en forma soluble en muy baja cantidad. Debido al interés de nuestro laboratorio en la interacción funcional CCDC28B-SIN1 se intentó por diversos métodos aumentar la expresión de SIN1 soluble, incluso expresándola sola, y nunca se logró aumentar el rendimiento (datos no mostrados). En cuanto a la co-expresión con Ran, o la idea de cristalizar el complejo entre ambas proteínas, creemos que aún quedan muchas cosas por explorar. En primer lugar, corroborar que realmente CCDC28B y Ran formen un complejo que pueda purificarse. En nuestros ensayos tanto Ran como CCDC28B se produjeron con la etiqueta 6-His y por lo tanto co-purificaban en IMAC cuando se hizo elución con escalones de concentración creciente de imidazol. Una variante podría haber sido la elución de la IMAC utilizando un gradiente de imidazol y ver si ambas proteínas co-eluían, lo que podría sugerir, aunque no confirmar, que estarían formando un complejo. Si el rendimiento fuera bueno también podría purificarse ese material para los ensayos de cristalogénesis. Otra posibilidad podría ser el clonado en un vector donde solo una de las proteínas tuviera la etiqueta 6-His, y así favorecer que solo una se ligue a la matriz y estudiar si la otra co-purifica. Finalmente, otra opción es la producción de las dos proteínas por separado para luego intentar formar el complejo *in vitro*, el cual podría ser purificarlo por gel filtración para ir luego al paso de cristalogénesis. El hecho de poder formar el complejo *in vitro* también tendría la ventaja de poder estudiar la interacción mediante otras metodologías, como resonancia plasmónica de superficie (Biacore) o microcalorimetría y así obtener más información sobre la interacción CCDC28B-Ran, la cual fue descubierta a través de un ensayo de co-inmunoprecipitación utilizando Tap-tag (R. Roepman, comunicación personal), pero sobre la cual no se tiene más información.

Finalmente, es bueno destacar, que con la fracción monomérica de CCDC28B purificada por IMAC e intercambio aniónico y en colaboración con Lucía Alfaya y Gualberto González (Cátedra de Inmunología, Facultad de Química) se inmunizaron llamas y se obtuvieron *nanobodies* específicos para CCDC28B. Utilizando el conocimiento adquirido durante este trabajo en cuanto a producción y purificación de proteínas recombinantes, se produjeron y purificaron los *nanobodies* que luego fueron caracterizados por Magdalena Cárdenas. Aunque los mismos finalmente no fueron útiles para los objetivos iniciales para los que fueron producidos, que eran poder mejorar la detección de la proteína endógena en inmunofluorescencia y *Western Blot*, sí permitieron inmunoprecipitar la proteína endógena, y confirmar la interacción con la kinesina-1 (Novas, R., Cárdenas-Rodríguez, M., Fabregat, M. y cols, manuscrito en preparación).

### **Sobre los ensayos de cristalización de CCDC28B**

Luego de varios ensayos, y en paralelo a las pruebas destinadas a la obtención de mayor cantidad de CCDC28B soluble, se optó por realizar pruebas de cristalización utilizando CCDC28B

monomérica. Esta decisión se basó en que el monómero mostró ser muy estable, lo cual parecía una buena señal a la hora de pensar en la obtención de cristales. Para ello se trabajó con la muestra purificada por IMAC y exclusión molecular, seleccionando en esta última cromatografía la fracción que por su volumen de elución correspondía a la forma monomérica..

Una de las cosas que deberíamos haber hecho es verificar que el material utilizado correspondiera en su totalidad al monómero. Es decir, una vez realizada la SEC y separada la fracción monomérica según el cromatograma, volver a correrla por exclusión molecular para asegurarnos de que el material seleccionado realmente está en el estado oligomérico deseado, y que no arrastramos parte de la población dimerica. A priori es difícil prever cómo se comportaría una mezcla monómero-dímero en un ensayo de cristalogénesis. Por un lado la cristalización en sí misma podría funcionar como un paso de purificación, cristalizando una sola de las formas de la proteína. Por otro lado podría suceder que la interacción entre ambas poblaciones afecte la disposición de las moléculas y por ende atente contra la formación del cristal. Como perspectiva en este sentido, si es que decidimos continuar trabajando exclusivamente con una forma de la proteína, deberíamos asegurarnos de obtener exclusivamente esa forma y disminuir lo máximo posible la presencia de otros estados oligoméricos en la solución. Asimismo, a la luz de resultados obtenidos con la proteína expresada en células eucariotas, donde observamos que CCDC28B podría formar oligómeros, sería conveniente re considerar cuál sería la forma oligomérica que deberíamos cristalizar para conocer la estructura fisiológicamente relevante de la proteína.

Los resultados negativos en los ensayos de cristalogénesis no nos sorprendieron demasiado. Sabíamos a partir de estudios *in silico* que aproximadamente el 40% de la secuencia en el extremo amino-terminal se predecía como desordenada, lo cual podría dar una gran flexibilidad a esa región de CCDC28B. Si bien no es una certeza que una proteína con una porción tan grande desestructurada se vea impedida de formar cristales ordenados, es cierto que esto va en sentido opuesto a la lógica que hay detrás de la formación del cristal (ver Materiales y Métodos). Existen en la literatura ejemplos de proteínas intrínsecamente desordenadas, con las cuales se ha logrado la formación de cristales e incluso a la resolución de su estructura (Meinke, G. et al, 2008), pero de todas formas sería inteligente tomar algunos recaudos al respecto para tratar de evitar dejar una porción tan grande de la proteína libre o desestructurada. Para esto surgen varias ideas como perspectiva, entre ellas las ya descritas sobre los interactores, cuya unión a CCDC28B podría estabilizar la región desestructurada. En el caso de otras proteínas puede probarse la eliminación de la región desordenada. En el caso de CCDC28B evaluamos que la eliminación de toda esta región nos dejaría una proteína muy pequeña con un dominio *coiled-coil*, lo cual no parecía muy interesante a la hora de determinar la estructura tridimensional. Por ello, optamos por la eliminación de una parte de la región, que incluyera los aminoácidos básicos del inicio de la proteína que podrían interferir con el correcto empaquetamiento de la proteína y la obtención de cristales.



Esta proteína, carente de los primeros 16 aminoácidos del extremo N-terminal (MDDKKKKRSPKPCLAQ) mostró ser la única sobre la cual la proteasa TEV pudo actuar, eliminando la etiqueta 6-His. Las otras construcciones analizadas, la proteína WT y otra carente de los primeros 11 aminoácidos, no pudieron ser cortadas por TEV, y nuestra hipótesis fue que el sitio de corte de la proteasa estaba inaccesible debido al plegamiento (o falta del mismo) de la proteína. El hecho de que la proteasa corte la proteína carente de los primeros 16 aminoácidos, ya nos está hablando de un cambio en el estado conformacional de la proteína que ameritaría un nuevo ensayo de cristalización. Además, existen casos descritos donde la etiqueta 6-His puede interferir en la formación de cristales (Waugh D., 2005; Carson M. et al, 2007). Es por esto que esta nueva construcción nos permitiría también, en caso de creerlo necesario, remover la etiqueta previo a los ensayos de cristalogénesis.

Por último, pero no menos importante, de persistir el problema con la formación de un cristal adecuado de CCDC28B, otra perspectiva que surge es la resolución de su estructura por NMR en lugar de cristalografía de rayos X. Además, puesto que para emplear esta técnica se necesita solamente la proteína en solución, tampoco nos deberíamos preocupar de sobremanera por las regiones móviles de la proteína o las predichas desorganizadas ya que estas también pueden ser resueltas por NMR. La gran diferencia radica en que para los ensayos de cristalografía de rayos X, podemos utilizar todos los instrumentos y la colaboración de la Unidad de Cristalografía de Proteínas del IPMon, mientras que no contamos con la infraestructura necesaria para realizar los estudios de NMR. Si decidimos realizar esta aproximación, la misma se llevará a cabo a través de las plataformas disponibles por la red CeBEM (<http://www.cebem.org.ar>), de la cual el IPMon es miembro activo. De proceder, se podría marcar CCDC28B con  $^{15}\text{N}$  para luego intentar obtener espectros HSQC (del inglés: *Heteronuclear Simple Quantum Correlation*) en los imanes disponibles en CeBEM. Luego, se rastrearían condiciones que resulten en espectros HSQC de buena calidad, y que permitan resolver la estructura usando protocolos estándar de NMR multidimensional, marcando sucesivamente la proteína recombinante con  $^{15}\text{N}$  y con  $^{13}\text{C}$ .

### **CCDC28B en su contexto fisiológico**

La otra aproximación de este trabajo consistió en recaudar información funcional sobre CCDC28B estudiando la proteína en su contexto fisiológico, particularmente en células de mamíferos. Trabajos previos del laboratorio permitieron identificar proteínas interactoras así como establecer un número de actividades de CCDC28B, deducidos a partir de la disminución o aumento de la expresión de la proteína en células en cultivo. Sin embargo resta mucho por conocer acerca del rol biológico de CCDC28B por lo que la información funcional que pudiéramos obtener en este trabajo resultaba novedosa.

Uno de los puntos sobre los que no se tenía información y era de sumo interés para nosotros debido a la información obtenida de la proteína recombinante, era conocer el estado oligomérico de la proteína en un ambiente fisiológico. En este sentido, durante este trabajo se mostró, mediante

ensayos de co-inmunoprecipitación, que CCDC28B puede interaccionar entre sí formando oligómeros. Cabe destacar que este estudio no nos permitió determinar la estequiometría del complejo por lo que se necesitarán nuevos estudios para determinar si se encuentra en forma de dímero u oligómeros de mayor tamaño. Pudimos determinar a su vez que, a diferencia de la proteína recombinante purificada, la asociación entre moléculas de CCDC28B no tendría lugar a través de puentes disulfuro, algo esperable para una proteína citosólica. Analizando la secuencia primaria de CCDC28B, la hipótesis principal entonces es que la asociación entre las distintas moléculas pueda ocurrir a través del dominio *coiled-coil*, ya que el mismo es uno de los principales motivos de oligomerización reportado en proteínas (Burkhard, P. et al, 2001). Este dominio está caracterizado por repeticiones periódicas de siete residuos (denominadas héptadas) las cuales están formadas preferencialmente por residuos apolares en las posiciones 1 y 4. Pese a su aparente simplicidad de secuencia, se ha demostrado que estos dominios pueden generar un gran número de variaciones estructurales: dímeros, trímeros, tetrámeros, pentámeros y hasta un heptámero, todos formados a través del dominio *coiled-coil*, con algunas variaciones en la orientación de las hélices o su alineamiento, tanto en homo-complejos como en hetero-complejos (Grigoryan, G. and Keating, A. E., 2008). Sin embargo, no pudimos confirmar esta posibilidad ya que continuamos detectando una interacción en los experimentos de co-inmunoprecipitación entre la proteína salvaje y un mutante para el dominio *coiled-coil*. Vale la pena aclarar que este experimento tiene algunas limitaciones, dada la falta de certeza sobre el grado de alteración que se genera en el *coiled-coil* por las mutaciones introducidas. En nuestro mutante *coiled-coil* se re-emplazaron por aspartatos las leucinas 169 y 172, correspondientes al residuo 1 y 4 de una misma héptada. Si bien estos cambios hacían perder la predicción *in silico* del dominio, debido a que el mismo en CCDC28B consta de 4 héptadas, el hecho de solo haber mutagenizado dos residuos de la misma héptada (la segunda), hace pensar que puede no haber sido suficiente para desestabilizar completamente el *coiled-coil*. Como perspectiva pensamos que sería importante repetir este experimento utilizando un nuevo mutante de CCDC28B donde las cuatro héptadas estén afectadas, maximizando así la posibilidad de que el *coiled-coil* se desestabilice.

Como se dijo anteriormente, el ensayo de co-inmunoprecipitación, si bien nos confirmó la existencia de interacción entre distintas moléculas de CCDC28B, no nos permite saber la estequiometría del complejo. Por ende, una importante perspectiva a futuro es caracterizar el estado oligomérico fisiológico de la proteína. Para esto nos planteamos la expresión de CCDC28B recombinante en células eucariotas, lo que permitirá que la proteína se exprese con las chaperonas, interactores fisiológicos y modificaciones post-traduccionales que correspondan para plegarse de manera correcta. La expresión en células de mamíferos y purificación de la proteína recombinante es sustancialmente diferente a lo realizado en *E.coli* (por una revisión del tema ver Dyson, M. R., 2016), utilizando plásmidos diferentes y una estrategia de purificación más compleja. Dado que en eucariotas hay mayor abundancia de proteínas ricas en histidina, la IMAC no es tan eficiente como método de purificación, por lo que generalmente se suma un paso de cromatografía de unión a estreptavidina, marcando la proteína con una etiqueta de Strep-tag (Cass, B. et al, 2004). Luego de

purificar la proteína recombinante podríamos efectuar una cromatografía de exclusión por tamaño para estimar su peso molecular o utilizar los ensayos de dispersión de luz dinámica.

La posibilidad de que CCDC28B forme oligómeros es importante en el contexto de intentar entender su función. Por ejemplo, trabajos previos de nuestro grupo han venido mostrando la interacción de CCDC28B con un gran número de proteínas, lo que *a priori* llamó la atención por el pequeño tamaño de la misma (200 aminoácidos y 22 kDa de peso molecular estimado). En este contexto, la formación de un oligómero seguramente aumente las posibilidades de regulación e interacción con más de una proteína en un mismo momento.

Dado que no logramos obtener la estructura tridimensional de CCDC28B quisimos acercarnos a establecer relaciones "región/dominio-función" a través de predecir *in silico* regiones o residuos potencialmente funcionales y evaluar si son importantes en la actividad de CCDC28B. Como se mencionó en la introducción, son varios los motivos que aparecen en la secuencia de CCDC28B que podrían tener relación con su función. De ellos, en este trabajo nos centramos en analizar únicamente el dominio *colied-coil* y regiones que podrían determinar la localización subcelular de CCDC28B: una secuencia de localización nuclear y una secuencia de exportación nuclear. Se mutagenizó la secuencia de CCDC28B en estas regiones introduciendo cambios que eliminaron las predicciones y luego se estudió si la expresión de estas formas mutantes rescataba el fenotipo que se obtenía al disminuir la expresión o funcionalidad de CCDC28B endógena. Para llevar a cabo este objetivo se tuvo que generar una línea celular estable en la que la funcionalidad de CCDC28B endógena estuviera disminuida. Para ello se utilizó la técnica de CRISPR/Cas9, y vale la pena reflexionar sobre lo que aprendimos en el propio proceso de generación de esta línea. Evaluamos cinco ARNg diferentes diseñados para generar un corte de doble hebra en los primeros exones de *CCDC28B*. Utilizamos un sistema lentiviral que permitía la integración de la construcción expresando el ARNg así como la Cas9 en el genoma. La primera observación interesante surgió al evaluar los niveles de *knockdown* de CCDC28B en estas células a distintos tiempos. Los cultivos en los que en algún momento se evidenció una disminución en la proteína endógena, luego recobraron los niveles normales. Dado que el proceso de mutagénesis basado en CRISPR/Cas9 depende del mecanismo de reparación con error NHEJ (del inglés: *Non-homologous end joining*) era de esperar que se generasen distintos tipos de mutaciones en las distintas células. Por lo tanto, en nuestro cultivo seguramente estuvieran co-existiendo mutaciones que determinaban un cambio en el marco abierto de lectura y la generación de proteínas truncas junto a pequeñas deleciones/inserciones que solo resultarían en el cambio de unos pocos aminoácidos. Por lo tanto, pensamos que esta recuperación muy probablemente fuera producto del rol funcional de la proteína, actuando seguramente como una proteína vital, cuya ausencia total hace inviable la línea celular en sí misma. Resulta interesante que *in vivo*, en pez cebra, también se demostró que la inyección de morfolinós (ARN modificados para inducir ARNi) por encima de 3 ng resulta en la muerte de todos los embriones, inclusive cuando todavía se puede detectar ARNm salvaje de *ccdc28b* (Cardenas-Rodríguez, M. et al 2013-B). A su vez, esta explicación sería consistente con lo observado en

pacientes BBS, donde la mutación reportada para CCDC28B corresponde a una mutación hipomórfica, producto de un cambio sinónimo que afecta el *splicing* del ARNm de CCDC28B y por lo tanto genera una disminución en los niveles, pero no la falta completa, del mensajero *wild type* (Badano, J. L. et al, 2006). Esto a su vez podría explicar por qué hasta ahora se ha encontrado solo una mutación hipomórfica en CCDC28B. En otras palabras empezar a responder las siguientes preguntas: si CCDC28B afecta ciliogénesis, ¿por qué no se han encontrado otras mutaciones en CCDC28B asociadas a BBS u otras ciliopatías? ¿Por qué CCDC28B es “solo” un modificador secundario de la enfermedad? En suma, todas las evidencias disponibles estarían indicando que la ausencia total de CCDC28B sería incompatible con la vida. Pensando en las ciliopatías, una posibilidad sería por tanto que mutaciones más fuertes en CCDC28B (nulas) contribuyan a síndromes más severos pero que se solapan genéticamente con BBS como por ejemplo Meckel-Gruber (MKS, del inglés: *Meckel-Gruber syndrome*). Ya se ha visto que mutaciones hipomórficas de genes MKS están presentes en pacientes BBS (Leitch, C. C. et al, 2008). Más experimentación será necesaria para testear si la situación opuesta, mutaciones fuertes en CCDC28B asociadas a MKS, también suceden.

Para sobrellevar este problema de viabilidad en la generación de la línea celular CCDC28B *knockout* decidimos entonces aislar y crecer células individuales luego de transducir las células con lentivirus codificantes para uno de los ARNg para CCDC28B. En una colaboración fuerte con la Dra. Irigoín quien llevó adelante este proceso, utilizamos *cell sorting* para obtener células individuales. Analizamos diferentes clones de células y en ningún caso logramos obtener *knockouts* completos de CCDC28B: en todos los clones analizados uno de los alelos presentaba un cambio en el marco abierto de lectura pero en el otro alelo solo un cambio de uno o pocos aminoácidos. Finalmente, se logró generar una línea estable (Clon #2.B1) donde la actividad de CCDC28B está reducida, reflejándose esto en un menor porcentaje de ciliación en estas células respecto de células control. En esta línea celular uno de los alelos de CCDC28B presenta una mutación nula y en el otro alelo se eliminaron 93 nucleótidos del exón 2 y se agregaron 60 nucleótidos provenientes del intrón adyacente. Esto hace que este clon exprese un solo ARNm para CCDC28B que origina una proteína donde se sustituyeron 29 aminoácidos (del 26 al 54) por 18 con una secuencia diferente a la WT. Esta región no contenía ningún motivo o residuo funcional predicho, pero tal vez el cambio altere la estructura de CCDC28B disminuyendo su actividad, al menos respecto a ciliogénesis. En primer lugar se verificó el rescate con CCDC28B WT en 3 experimentos independientes, lo que permitía afirmar que el defecto en la ciliación era debido a cambios en CCDC28B. El grado de rescate que se obtuvo fue variable, en general no alcanzando el nivel de ciliación observado en un cultivo de células normales. Esto podría explicarse por diferencias en el porcentaje de transfección en los distintos experimentos y en este sentido vale la pena tomar en cuenta que en estos experimentos podríamos estar subvalorando la eficiencia del rescate, ya que al no poder identificar las células transfectadas, también se contaron en el análisis células que no expresaban CCDC28B WT. Si bien intentamos subsanar este problema clonando tanto los mutantes como el WT como fusión a HA para luego identificar las células transfectadas con un anticuerpo anti-HA, esto no funcionó en IF, no pudiendo determinar si el problema era el anticuerpo o la baja expresión de

CCDC28B. Otra dificultad que podría disminuir el porcentaje de rescate de CCDC28B WT es que la línea 2.B1 sigue expresando la nucleasa Cas9 y el ARNg, y por lo tanto podrían estar también afectando el gen de CCDC28B clonado en el plásmido de transfección. Para tratar de evitar que ello suceda, se mutagenizó la secuencia de CCDC28B WT y mutantes de forma de no cambiar la secuencia de aminoácidos pero sí las secuencias de reconocimiento de Cas9 (PAM), evitando así la actividad de la nucleasa en el material transfectado. Estas últimas mutagénesis fueron realizadas recientemente y aún resta su evaluación en las células en cultivo.

En este punto del trabajo, comenzamos el análisis priorizando los mutantes que podrían modificar la localización subcelular de CCDC28B. Como ya se mencionó, se sabe que CCDC28B afecta ciliación, al menos en parte a través de su interacción con SIN1 (Cardenas-Rodríguez, M. et al, 2013-A). También se ha demostrado que al disminuir la expresión de CCDC28B, disminuye la ciliación y SIN-1 se concentra simultáneamente en el núcleo (Cardenas-Rodríguez, M. et al, 2013-A y Tesis Doctoral de Magdalena Cárdenas). Finalmente, también se demostró que otro interactor de CCDC28B, KLC-1 (del inglés: *kinesin light chain 1*) modifica la localización de CCDC28B: al disminuir KLC-1, CCDC28B se acumula en el núcleo. Por último, con una estrategia similar a la que usamos en este trabajo, pero utilizando el pez cebra, se observó que mientras CCDC28B WT es capaz de rescatar el fenotipo ciliar en larvas donde CCDC28B endógena fue disminuida, el mutante NLS, no rescata el fenotipo (Novas, R. Cárdenas-Rodríguez, M.; Fabregat, M. y cols, manuscrito en preparación). En suma, se demostró que la localización subcelular de CCDC28B juega, al menos en parte, un rol importante en la funcionalidad de la proteína, específicamente en lo que respecta a ciliogénesis. Es por eso que nos volcamos por intentar obtener algo más de información sobre la importancia de las regiones NLS y NES, en la función de CCDC28B.

Vale la pena recordar que mientras las proteínas pequeñas pueden entrar al núcleo por difusión a través del poro nuclear, las proteínas mayores a 45 kDa requieren de transportadores específicos, importinas y exportinas, para entrar y salir del núcleo respectivamente (Cautain, B. et al, 2014). Estos transportadores reconocen y se unen a las proteínas "carga" en secuencias específicas, llamadas de localización nuclear (NLS) para importinas y de exportación nuclear (NES) para exportinas. Algunas de estas secuencias pueden predecirse con herramientas de análisis de secuencia primaria de proteínas. Por lo tanto, si bien CCDC28B monomérica tiene un tamaño que le permitiría entrar al núcleo por difusión, alcanzaría solo con la formación de dímeros para que la entrada y salida requiriera de transportadores. De hecho, la secuencia de CCDC28B posee tanto una señal NLS clásica como una NES, y ensayos previos en nuestro laboratorio mostraron que esas secuencias son funcionales. Cuando analizamos la capacidad de los mutantes NLS y NES de rescatar el porcentaje de ciliación en las células 2.B1 observamos resultados coherentes con los obtenidos en el pez cebra: mientras que la proteína WT y el mutante NES rescatan el porcentaje de ciliación, el mutante NLS, incapaz de entrar al núcleo, no lo hace. Estos son resultados preliminares que esperamos corroborar pronto con la realización de nuevos experimentos.

Por último, también hay que decir que esta línea hTERT-RPE clon 2.B1 nos permitirá testear también la funcionalidad de otras regiones o residuos eventualmente importantes, respecto a la

actividad de CCDC28B en ciliación. Además, hasta el momento no hemos evaluado en esta línea la actividad de mTORC2, que como ya se discutió puede ser otra manera de observar la funcionalidad de CCDC28B.

En el comienzo de este trabajo nos planteamos como objetivo el contribuir a entender la función de CCDC28B a partir de caracterizar la proteína desde el punto de vista estructural, y establecer relaciones de estructura-función. Pese a que el resultado más esperado, la resolución de la estructura tridimensional de CCDC28B por cristalografía de rayos X, no se logró, el trabajo aportó información y herramientas importantes que nos permitirá avanzar en el diseño de estrategias para la resolución de la estructura y continuar disecando el rol biológico de CCDC28B, paso fundamental para comprender la base celular de BBS y posiblemente otras ciliopatías.

## AGRADECIMIENTOS

---

En primer lugar a Florencia Irigoín, mi orientadora, por su confianza y su paciencia. Por haberse tomado el trabajo de aceptarme bajo su tutela prácticamente sin experiencia en muchas de las metodologías desarrolladas a lo largo de este trabajo, confiar en mi y ayudarme a crecer tan exponencialmente como nunca lo hubiera imaginado. Por abrirme puertas que ni siquiera sabía que estaban allí y ayudarme siempre en el camino de descubrirlas.

A José Badano, por abrirme las puertas de su laboratorio, por haberme dado el privilegio de formar parte de este grupo de trabajo tan lindo que él formó. Por haber aparecido junto a Florencia y Alejandro Buschiazzo con ese llamado a cargo en un momento clave de mi vida. Por recordarme desde aquel momento cuanto me gusta lo que hago y darme la oportunidad de disfrutar haciéndolo. A José y a Florencia por permitirme y apoyarme siempre en esa locura de convivir con mis dos pasiones que son la ciencia y el deporte.

A mis compañeras de laboratorio: Rossina, Belén, Magdalena, Victoria y Paola. Por abrirme las puertas del grupo de la forma en la que lo hicieron. Por tener un grupo tan disfrutaba de trabajo, y preocuparse siempre por esas pequeñas cosas que te hacen sentir ganas de ir a trabajar. Y fundamentalmente por aceptar el fútbol como tema principal de los almuerzos. Llena de orgullo tener un grupo de trabajo así.

Al resto de los compañeros del Instituto, que terminan de conformar un lugar espectacular para trabajar. A todos los que molesté en mayor o menor medida durante mi trabajo y siempre estuvieron ahí para colaborar con la paciencia necesaria para alguien que venía de otra área. A Alejandro Buschiazzo, Felipe Trajtenberg, Nicole Larrieux, Juan Imelio, Agustín Correa, Federico Carrión, Madelón Portela y todos los que siempre están. A los del fútbol de los miércoles, el cable a tierra que necesita la semana.

A mi familia. A mi padres uno acá y la otra allá, en Canadá, pero que están siempre para apoyarme, en todas y cada una de las locuras que he hecho y sin los cuales no sería quien soy, pero más importante, no sería lo feliz que soy. A mi abuela, mi hermana y mis amigos, pocos pero buenos, esos que son un miembro más de la familia, ustedes me trajeron hasta los 29 en una sola pieza. Y a Ale, por enseñarme que todo esto, lo lindo y lo feo, lo bueno y lo malo, tiene mucho más gusto si se comparte con alguien. Me completás.

Finalmente, también agradecer a las agencias financiadoras. En primer lugar a la Agencia Nacional de Investigación e Innovación (ANII) y a la Comisión Académica de Posgrados (CAP) por apoyarme económicamente durante el desarrollo de este trabajo mediante la Beca de Maestría y la Beca de Finalización de Posgrados, respectivamente. También agradecer a la Comisión Sectorial

de Investigación Científica (CSIC) por financiar el proyecto en el cual estuvo enmarcada mi tesis de Maestría.

A todos ustedes, muchas gracias.



## BIBLIOGRAFÍA

---

- Anderson, R. G. W. **The three-dimensional structure of the basal body from the rhesus monkey oviduct.** *J Cell Biol*, 54(2): 246-65, (1972).
- Badano, J. L. and Katsanis, N. **Beyond Mendel: an evolving view of human genetic disease transmission.** *Nat Rev Genet*, 3(10): 779-89, (2002).
- Badano, J. L., Leitch, C. C., Ansley, S. J., May-Simera, H., Lawson, S., Lewis, R. A., Beales, P. L., Dietz, H. C., Fisher, S., Katsanis, N. **Dissection of epistasis in oligogenic Bardet-Biedl syndrome.** *Nature*, 439(7074): 326-30, (2006).
- Balbás, P. **Understanding the art of producing protein and nonprotein molecules in escherichia coli.** *Mol Biotechnol*, 19(3): 251-67, (2001).
- Bardet, G. **Sur un syndrome d'obestie infantile avec polydactylie et rétinite pigmentaire (contribution a l'étude des formes cliniques de l'obésité hypophysaire).** *Doctoral Thesis*, Sorbonne University of Paris, Paris, (1920).
- Beales, P. L., Badano, J. L., Ross, A. J., Ansley, S. J., Hoskin, B. E., Kirsten, B., Mein, C. A., Froguel, P., Scambler, P. J., Lewis, R. A., Lupski, J. R., Katsanis, N. **Genetic interaction of BBS1 mutations with alleles at other BBS loci can result in non-mendelian Bardet-Biedl syndrome.** *Am J Hum Genet*, 72: 1187-1199, (2003).
- Bergmann, C. **Ciliopathies.** *Eur J Pediatr*, 171: 1285-1300, (2012).
- Biedl, A. **Ein geschwisterpaar mit adiposo-genitaler dystrophie.** *Deutsch Med Wochenschr*, 48:1630, (1922).
- Burkhard, P., Stetefeld, J., Strelkov, S. V. **Coiled coils: a highly versatile protein folding motif.** *Trends Cell Biol*, 11(2): 82-8, (2001).
- Cardenas-Rodríguez, M. and Badano, J. L. **Ciliary biology: understanding the cellular and genetic basis of human ciliopathies.** *Am J Med Genet C Semin Med Genet*, 151C(4): 263-80, (2009).
- Cardenas-Rodríguez, M., Irigoín, F., Osborn, D. P. S., Gascue, C., Katsanis, N., Beales, P. L., Badano, J. L. **The Bardet-Biedl syndrome-related protein CCDC28B modulates mTORC2 function and interacts with SIN1 to control cilia length independently of the mTOR complex.** *Hum Mol Genet*, 22(20): 4031-42, (2013) -A-.
- Cardenas-Rodríguez, M., Osborn, D. P. S., Irigoín, F., Graña, M., Romero, H., Beales, P. L., Badano, J. L. **Characterization of CCDC28B reveals its role in ciliogenesis and provides**

- insight to understand its modifier effect on Bardet-Biedl syndrome.** *Hum Genet*, 132(1): 91-105, (2013) -B-
- Carson, M., Johnson, D. H., McDonald, H., Brouillette, C., DeLuca, L. J. **His-tag impact on structure.** *Acta Cryst*, D63: 295-301, (2007).
  - Cass, B., Lan Pham, P., Kamen, A., Durocher, Y. **Purification of recombinant proteins from mammalian cell culture using a generic double-affinity chromatography scheme.** *Protein Expr Purif*, 40(1): 77-85, (2005).
  - Cautain, B., Hill, R., de Pedro, N., Link, W. **Components and regulation of nuclear transport processes.** *FEBS J*, 282(3): 445-62, (2015).
  - Chauvaux, S., Souchon, H., Alzari, P. M., Chariot, P., Beguin, P. **Structural and functional analysis of the metal-binding sites of *Clostridium thermocellum* Endoglucanase CelD.** *J Biol Chem*, 270(1): 9757-62, (1995).
  - Cheng, Y., Lobo, R. F., Sandler, S. I., Lenhoff, A. M. **Kinetics and equilibria of lysozyme precipitation and crystallization in concentrated ammonium sulfate solutions.** *Biotechnol Bioeng*, 94(1): 177-88, (2006).
  - Correa, A., Ortega, C., Obal, G., Alzari P., Vincentelli, R., Oppezzo, P. **Generation of a vector suite for protein solubility screening.** *Front Microbiol*, 5:67, (2014).
  - Correa, A., Oppezzo, P. **Overcoming the solubility problem in E.coli: available approaches for recombinant protein production.** *Methods Mol Biol*, 1258: 27-44, (2015).
  - Davis, E. E. and Katsanis, N. **The ciliopathies: A transitional model into systems biology of human genetic disease.** *Curr Opin Genet Dev*, 22(3): 290-303, (2012).
  - Derewenda, Z. S. **The use of recombinant methods and molecular engineering in protein crystallization.** *Methods* 34(3): 354-63, (2004).
  - Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., Obradovic, Z. **Intrinsic disorder and protein function.** *Biochemistry*, 41(21): 6573-82, (2002).
  - Dyson, M. R. **Fundamentals of expression in mammalian cells.** *Adv Exp Med Biol*, 896: 217-24, (2016).
  - Esposito, D., Chatterjee, D. K. **Enhancement of soluble protein expression through the use of fusion tags.** *Curr Opin Biotechnol*, 17(4): 353-8, (2006).
  - Fisch, C. and Dupuis-Williams, P. **Ultrastructure of cilia and flagella - back to the future!** *Biol Cell*, 103: 249-270, (2011).
  - Fosythe, E. and Beales, P. L. **Bardet-Biedl syndrome.** *Eur J Hum Genet*, 21(1): 8-13, (2013).

- Garbuzynski, S. O., Melnik, B. S., Lobanov, M. Y., Finkelstein, A. V., Galzitskaya, O. V. **Comparison of X-ray and NMR structures: is there a systematic difference in residue contacts between X-ray and NMR-resolved protein structures?** *Proteins*, 60(1): 139-47, (2005).
- Gerdes, J. M. and Katsanis, N. **Ciliary function and Wnt signal modulation.** *Curr Top Dev Biol*, 85: 175-95, (2008).
- Gherman, A., Davis, E. E., Katsanis, N. **The ciliary proteome database: an integrated community resource for the genetic and functional dissection of cilia.** *Nat Genet*, 38(9): 961-2, (2006).
- Goetz, S. C. and Anderson, K. V. **The primary cilium: a signaling center during vertebrate development.** *Nat Rev Genet*, 11(5): 331-344, (2010).
- Goto, N. K. and Kay, L. E. **New developments in isotope labeling strategies for protein solution NMR spectroscopy.** *Curr Opin Struct Biol*, 10(5): 585-92, (2000).
- Graser, S., Stierhof, Y. D., Lavoie, S. B., Gassner, O. S., Lamla, S., Le Clech, M., Nigg, E. A. **Cep164, a novel centriole appendage protein required for primary cilium formation.** *J Cell Biol*, 179: 321-330, (2007).
- Grigoryan, G. and Keating, A. E. **Structural specificity in coiled-coil interactions.** *Curr Opin Struct Biol*, 18(4): 477-83, (2008).
- Guzik-Lendrum, S., Rank, K. C., Bensel, B. M., Taylor, K. C., Rayment, I., Gilbert, S. P. **Kinesin-2 KIF3AC and KIF3AB Can Drive Long-Range Transport along Microtubules.** *Biophys J*, 109(7): 1472-82, (2015).
- Heon, E., Kim, G., Qin, S., Garrison, J. E., Tavares, E., Vincent, A., Nuangchamng, N., Scott, C. A., Slusarski, D. C., Sheffield, V. C. **Mutations in C8ORF37 cause Bardet Biedl syndrome (BBS21).** *Hum Mol Genet*, 25(11): 2283-2294, (2016).
- Hirokawa, N. **Kinesin and dynein superfamily proteins and the mechanism of organelle transport.** *Science*, 279(5350): 519-26, (1998).
- Hou, Y. and Witman, G. B. **Dynein and intraflagellar transport.** *Exp Cell Res*, 334(1): 26-34, (2015).
- Hoyer-Fender, S. **Primary and Motile Cilia: Their Ultrastructure and Ciliogenesis.** *Cilia and Nervous System Development and Function*, Chapter 1: 1-53, (2013).
- Hresko, R. C. and Mueckler, M. **mTOR.RICTOR is the Ser473 kinase for Akt/protein kinase B in 3T3-L1 adipocytes.** *J. Biol. Chem.*, 280: 40406-40416, (2005).

- Irigoín, F. and Badano, J. L. **Keeping the balance between proliferation and differentiation: the primary cilium.** *Curr Genomics*, 12(4): 285-97, (2011).
- Ishikawa, H., Kubo, A., Tsukita, S., Tsukita, S. **Odf2-deficient mother centrioles lack distal/subdistal appendages and the ability to generate primary cilia.** *Nat Cell Biol*, 7(5): 517-24, (2005).
- Jabs, A., Weiss, M. S., Hilgenfeld, R. **Non-proline cis peptide bonds in proteins.** *J Mol Biol*, 286(1): 291-304, (1999).
- Jin, H., White, S. R., Shida, T., Schultz, S., Aguiar, M., Gygi, S. P., Bazan, J. F., Nachury, M. V. **The conserved Bardet-Biedl syndrome proteins assemble a coat that traffics membrane proteins to cilia.** *Cell*, 141: 1208-1219, (2010).
- Jinek, M., Chylinski, K. et al. **A programmable dual-RNA guided DNA endonuclease in adaptive bacterial immunity.** *Science*, 337: 816-21, (2012).
- Kadima, W., Øgendal, L., Bauer, R., Kaarsholm, N., Brodersen, K., Hansen, J. F., Porting, P. **The influence of ionic strength and pH on the aggregation properties of Zinc-free insulin studied by static and dynamic laser light scattering.** *Biopolymers*, 33: 1643-1657, (1993).
- Katsanis, N., Ansley, S. J., Badano, J. L., Eichers, E. R., Lewis, R. A., Hoskins, B. E., Scambler, P. J., Davidson, W. S., Beales, P. L., Lupski, J. R. **Triallelic inheritance in Bardet-Biedl syndrome, a mendelian recessive disorder.** *Science*, 293(5538): 2256-9, (2001).
- Kim, J. C., Badano, J. L., Sibold, S., Esmail, M. A., Hill, J., Hoskin, B. E., Leitch, C. C., Venner, K., Ansley, S. J., Ross, A. J. **The Bardet-Biedl protein BBS4 targets cargo to the pericentriolar region and is required for microtubule anchoring and cell cycle progression.** *Nat Genet*, 36: 462-470, (2004).
- Kim, J. C., Ou, Y. Y., Badano, J. L., Esmail, M. A., Leitch, C. C., Friedrich, E., Beales, P. L., Archibald, J. M., Katsanis, N., Rattner, J. B. **MKKS/BBS6 a divergent chaperonin-like protein linked to the obesity disorder Bardet-Biedl syndrome, is a novel centrosomal component required for cytokinesis.** *J Cell Sci*, 118: 1007-1020, (2005).
- Laplante, M. and Sabatini, D. M. **mTOR signaling at a glance.** *J Cell Sci*, 122(Pt 20): 3589-94, (2009).
- Laplante, M. and Sabatini, D. M. **mTOR signaling in growth control and disease.** *Cell*, 149(2): 274-93, (2012).
- Laskowski, R. A., Watson, J. D., Thornton, J. M. **From proteins structure to biochemical function?** *J Struct Funct Genomics*, 4(2-3): 167-77, (2003).

- Leitch, C. C., Zaghloul, N. A., Davis, E. E., Stoetzel, C., Diaz-Font, A., Rix, S., Al-Fadhel, M., Lewis, R. A., Eyaid, W., Banin, E., Dollfus, H., Beales, P. L., Badano, J. L., Katsanis, N. **Hypomorphic mutations in syndromic encephalocele genes are associated with Bardet-Biedl syndrome.** *Nat Genet*, 40:443-448, (2008).
- Li, J. B., Gerdes, J. M., Haycraft, C. J., Fan, Y., Teslovich, T. M., May-Simera, H., Li, H., Blacque, O. E., Li, L., Leitch, C. C., Lewis, R. A., Green, J. S., Parfrey, P. S., Leroux, M. R., Davidson, W. S., Beales, P. L., Guay-Woodford, L. M., Yoder, B. K., Stormo, G. D., Katsanis, N., Dutcher, S. K. **Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene.** *Cell*, 117: 541-552, (2004).
- Loktev, A. V., Zhang, Q., Beck, J. S., Searby, C. C., Scheetz, T. E., Bazan, J. F., Slusarski, D. C., Sheffield, V. C., Jackson, P. K., Nachury, M. V. **A BBSome subunit links ciliogenesis, microtubule stability, and acetylation.** *Dev Cell*, 15: 854-865, (2008).
- Lupas, A. **Predicting coiled-coil regions in proteins.** *Curr Opin Struct Biol*, 7(3): 388-93, (1997).
- Marcum, J. M. and Borisy, G. G. **Characterization of microtubule protein oligomers by analytical ultracentrifugation.** *J Biol Chem*, 253(8): 2825-33, (1978).
- Marshall, W. F. **Basal bodies platforms for building cilia.** *Curr Top Dev Biol*, 85: 1-22, (2008).
- McPherson, A. **Crystallization of proteins from polyethylene glycol.** *J Biol Chem*, 251: 6300-6303, (1976).
- McPherson, A., Cudney, B. **Searching for silver bullets: an alternative strategy for crystallizing macromolecules.** *J Struct Biol*, 156(3): 387-406, (2006).
- McPherson, A., Cudney, B. **Optimization of crystallization conditions for biological macromolecules.** *Acta Cryst*, F70: 1445-1467 (2014).
- McPherson, A., Gavira, J. A. **Introduction to protein crystallization.** *Acta Cryst*, F70, 2-20, (2014).
- Meinke, G., Ezeokonkwo, C., Balbo, P., Stafford, W., Moore, C., Bohm, A. **Structure of yeast poly(A) polymerase in complex with a peptide from Fip1, an intrinsically-disordered protein.** *Biochemistry*, 47(26): 6859-69, (2008).
- Mitchison, H. M. and Valente, E. M. **Motile and non-motile cilia in human pathology: from function to phenotypes.** *J Pathol*, 241(2): 294-309, (2017).
- Moran, D.T., Rowley, J.C., Jafek, B.W., Lovell, M.A. **The fine structure of the olfactory mucosa in man.** *J Neurocytol*, 11(5): 721-46, (1982).

- Nachury, M. V., Loktev, A. V., Zhang, Q., Westlake, C. J., Peränen, J., Merdes, A., Slusarski, D. C., Scheller, R. H., Bazan, J. F., Sheffield, V. C. **A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis.** *Cell*, 129: 1201-1213, (2007).
- Najmanovich, R. J., Torrance, J. W., Thornton, J. M. **Prediction of protein function from structure: insights from methods for the detection of local structural similarities.** *Biotechniques*, 38(6): 847-851, (2005).
- Nido, G. S., Méndez, R., Pascual-García, A., Abia, D., Bastolla, U. **Protein disorder in the centrosome correlates with complexity in cell types number.** *Mol Biosyst*, 8(1): 353-67, (2012).
- Novas, R., Cardenas-Rodríguez, M., Irigoín, F., Badano, J. L. **Bardet-Biedl syndrome: Is it only cilia dysfunction?** *FEBS Lett*, 589(22): 3479-91, (2015).
- Novas, R., Cardenas-Rodríguez, M., Rodao, M., Fabregat, M., Davison, C., Casanova, G., Alfaya, L., González-Sapienza, G., Irigoín, F., Badano, J. L. **Members of the kinesin 1 motor interact with the Bardet-Biedl syndrome related protein CCDC28B and participate in cilia length regulation.** *Manuscrito en preparación.*
- Oh, W. J. and Jacinto, E. **mTOR complex 2 signaling and functions.** *Cell Cycle*, 10(14): 2305-16, (2011).
- Oliveira, C. S., Giuffrida, F. M., Crispim, F., Saddi-Rosa, P., Reis, A.F. **ADIPOQ and adiponectin: the common ground of hyperglycemia and coronary artery disease?** *Arg Bras Endocrinol Metabol*, 55(7): 446-54, (2011).
- Pagano, M. and Guvreau, K. **Principles of biostatistics.** *Duxbury Thomson Learning, USA*, (2000).
- Pan, D. and Matsuura, Y. **Structures of the pleckstrin homology domain of *Saccharomyces cerevisiae* Avo1 and its human orthologue Sin1, an essential subunit of TOR complex 2.** *Acta Crystallogr Sect F Struct Bill Cryst Commun*, 68(Pt 4): 386-92, (2012).
- Portolano, N., Watson, P. J., Fairall, L., Millard, C. J., Milano, C. P., Song, Y., Cowley, S. M., Schwabe, J. W. **Recombinant protein expression for structural biology in HEK 293F suspension cells: a novel and accessible approach.** *J Vis Exp*, (92): e51897, (2014).
- Reiter, J. F., Blacque, O. E., Leroux, M. R. **The base of the cilium: roles for transition fibers and the transition zone in ciliary formation, maintenance and compartmentalization.** *EMBO Rep*, 13(7): 608-18, (2012).

- Robichon, C., Luo, J. , Causey, T. B., Benner, J. S., Samuelson, J. C. **Engineering *Escherichia coli* BL21(DE3) derivative strains to minimize *E.coli* protein contamination after purification by immobilized metal affinity chromatography.** *Appl Environ Microbiol*, 77(13): 4634-4646, (2011).
- Rohatg, R. and Snell, W. J. **The ciliary membrane.** *Curr Opin Cell Biol*, 22(4): 541-546, (2010).
- Rosenbaum, J. L. and Witman, G. B. **Intraflagellar transport.** *Nat Rev Mol Cell Biol*, 3(11): 813-25, (2002).
- Rupp, B. **Biomolecular crystallography: Principles, practice, and application to structural biology.** *Garland Science*, ISBN: 978-0-8153-4081-2, (2010).
- Saaranen, M. J. and Ruddock, L. W. **Disulfide bond formation in the cytoplasm.** *Antioxidant & Redox Signaling*, 19: 46-53, (2013).
- Sanjana, N. E., Shalem, O., Zhang, F. **Improved vectors and genome-wide libraries for CRISPR screening.** *Nat Methods*, 11(8): 783-784, (2014).
- Schaefer, E., Stoetzel, C., Scheidecker, S., Geoffroy, V., Prasad, M. K., Redin, C., Missotte, I., Lacombe, D., Mandel, J. L., Muller, J., Dolfus, H. **Identification of a novel mutation confirms the implication of IFT172 (BBS20) in Bardet-Biedl syndrome.** *J Hum Genet*, 61(5): 447-50, (2016).
- Scholey, J. M. **Kinesin-II, a membrane traffic motor in axons, axonemes, and spindles.** *J Cell Biol*, 133(1): 1-4, (1996).
- Shapiro, L. and Scherer, P. E. **The crystal structure of a complement-1q family protein suggests an evolutionary link to tumor necrosis factor.** *Curr Biol*, 8: 335-338, (1998).
- Shapiro, L. and Harris, T. **Finding function through structural genomics.** *Curr Opin Biotechnol*, 11(1): 31-5, (2000).
- Snyder, D.A, Chen, Y., Denissova, N. G., Acton, T., Aramini, J. M., Ciano, M., Karlin, R., Liu, J., Manor, P., Rajan, P. A., Rossi, P., Swapna, G. V., Xiao, R., Rost, B., Hunt, J., Montelione, G. T. **Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are complementary methods for small protein structure determination.** *J Am Chem Soc*, 127(47): 16505-11, (2005).
- Stoetzel, C., Laurier, V., Davis, E. E., Muller, J., Rix, S., Badano, J. L., Leitch, C. C., Salem, N., Chouery, E., Corbani, S., Jalk, N., Vicaire, S., Sarda, P., Hamel, C., Lacombe, D., Holder, M., Odent, S., Holder, S., Brooks, A. S., Elcioglu, N. H., Silva, E. D., Rossillion, B., Sigaudy, S., de Ravel, T. J., Lewis, R. A., Leheup, B., Verloes, A., Amati-Bonneau, P., Mégarbané, A., Poch, O.,

- Bonneau, D., Beales, P. L., Mandel, J. L., Katsanis, N., Dollfus, H. **BBS10 encodes a vertebrate-specific chaperoning-like protein and is a major BBS locus.** *Nat Genet*, 38(5): 521-4, (2006).
- Stoetzel, C., Muller, J., Laurier, V., Davis, E. E., Zaghoul, N. A., Vicaire, S., Jacquelin, C., Plewniak, F., Leitch, C. C., Sarda, P., Hamel, C., de Ravel, T. J., Lewis, R. A., Friederich, E., Thibault, C., Danse, J. M., Verloes, A., Bonneau, D., Katsanis, N., Poch, O., Mandel, J. L., Dollfus, H. **Identification of a novel BBS gene (BBS12) highlights the major role of a vertebrate-specific branch of chaperoning-related proteins in Bardet-Biedl syndrome.** *Am J Hum Genet*, 80(1): 1-11, (2007).
  - Stroud, R. M., Kay, L. M., Dickerson, R. E. **The crystal and molecular structure of DIP-inhibited bovine trypsin at 2.7 Angstrom resolution.** *Cold Spring Hard Symp Quant Biol*, 36: 125-40, (1972).
  - Studier, F.W. **Protein production by auto-induction in high density shaking cultures.** *Protein Expr Purif*, 41(1): 207-34, (2005).
  - Sulik, K., Dehart, D.B., Inagaki, T., Carson, J.L., Vrablic, T., Gesteland, K., Schoenwolf, G.C. **Morphogenesis of the murine node and notochordal plate.** *Dev Dyn*, 201(3): 260-78, (1994).
  - Trajtenberg, F., Albanesi, D., Ruétalo, N., Botti, H., Mechaly, A. E., Nieves, M., Aguilar, P. S., Cybulski, L., Larrieux, N., de Mendoza, D., Buschiazzo, A. **Allosteric activation of bacterial response regulators: the role of the cognate histidine kinase beyond phosphorylation.** *MBio*, 5(6): e02105, (2014).
  - Van den Ent, F., Löwe, J. **RF cloning: A restriction-free method for inserting target genes into plasmids.** *J Biochem Biophys Methods*, 67 (1): 67-74, (2006).
  - Vetter, I. R., Arndt, A., Kutay, U., Görlich, D., Wittinghofer, A. **Structural view of the Ran-importin  $\beta$  interaction at 2.3 Å resolution.** *Cell*, 97: 635-646, (1999).
  - Wagner, R., Gonzalez, D. H., Podesta, F. E., Andreo, C. S. **Changes in the quaternary structure of phosphoenolpyruvate carboxylase induced by ionic strength affects its catalytic activity.** *Eur J Biochem*, 164: 661-666, (1987).
  - Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., Jones, D. T. **Prediction and functional analysis of native disorder in proteins from the three kingdoms of life.** *J Mol Biol*, 337: 635-645, (2004).
  - Waugh, D. S. **Making the most of affinity tags.** *Trends Biotechnol*, 23(6): 316-20, (2005).



- Wiechelman, K. J., Braun, R. D., Fitzpatrick, J. D. **Investigation of the Bicinchoninic Acid Protein Assay: Identification of the Groups Responsible for Color Formation.** *Anal Biochem*, 175(1), 231-7, (1988).
- Wright, P. E. and Dyson, H. J. **Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm.** *J Mol Biol*, 293(2): 321-31, (1999).
- Yee, A. A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X., Skarina, T., Evdokimova, E., Liu, C. U., Semesi, A., Guido, V., Edwards, A. M., Arrowsmith, C. H. **NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins.** *J Am Chem Soc*, 127(47): 16512-7, (2005).
- Zaghoul, N. A. and Katsanis, N. **Mechanistic insights into Bardet-Biedl syndrome, a model ciliopathy.** *J Clin Invest*, 119(3): 428-37, (2009).
- Zhang, Q., Seo, S., Bugge, K., Stone, E. M., Sheffield, V. C. **BBS proteins interact genetically with the IFT pathway to influence SHH-related phenotypes.** *Hum Mol Genet*, 21(9): 1945-53, (2012).

# ANEXO

## Lista de primers utilizados

PRIMER	SECUENCIA 5' - 3'	FUNCIÓN
CCDC28B-NcoI Fwd	ACGTCCATGGATGGATGACAAAAAGAAGAAACG	Clonado de CCDC28Bwt en pET21d
CCDC28B-XhoI Rev	ACGTCTCGAGCTACGCAGCGGACTGCTC	Clonado de CCDC28Bwt en pET21d
CCDC28B-NM Fwd pColo	GGATCGGAAAACCTGTATTTTCAGGGATCCATGGATGACAAAAAGAAGAAACGGAG	Clonado de CCDC28Bwt en <i>vector suite</i> para clonar con proteínas de fusión
CCDC28B-NM Rev pColo	GAACTGCGGGTGGCTCCAGTGCCTCCGATCCCGCAGCGGACTGCTCCTCAGGCTCG	Clonado de CCDC28Bwt en <i>vector suite</i> para clonar con proteínas de fusión
pQE80-CCDC28B-Fwd	CCTGTATTTTCAGGGATCCGGTTCAGATGACAAAAAGAAGAAACGGAGTCCAAGCCCTGCC	Clonado de CCDC28Bwt en pQE80
pQE80-CCDC28B-Rev	GCTGCAGGTCGACGCCAAGATCCTTTTTTACGCAGCGGACTGCTCCTCAGGCTCGGCG	Clonado de CCDC28Bwt en pQE80
pACYCDuet S1 Fwd	CCATGGGCAGCAGCCATCACCATCATCACCCAGGATCAGGATCGGAAAACC	Primer genérico para clonar en el sitio 1 de pACYCDuet desde el pQE80
pACYCDuet S1 CCDC28B Rev	TCGACTTAAGCATTATGCGGCCCAAGCTTCGACTTACGCAGCGGACTGCTCCTCAGGCTCGGCG	Clonado de CCDC28Bwt en el sitio 1 de pACYCDuet desde el pQE80
pQE80-SIN1a Fwd	CCTGTATTTTCAGGGATCCGGTTCAGCCTTCTTGGACAATCCAATATCATTCTAGC	Clonado de SIN1a en pQE80
pQE80-SIN1a Rev	GCTGCAGGTCGACGCCAAGATCCTTTTTTACTGCTGCCGGATTCTTCTCCTTC	Clonado de SIN1a en pQE80
pQE80-RAN Fwd	CCTGTATTTTCAGGGATCCGGTTCAGTGCCTGCGCAGGGAGAGCCCAAGGTCAG	Clonado de RAN en pQE80
pQE80-RAN Rev	GCTGCAGGTCGACGCCAAGATCCTTTTTTACAGGTCATCATCCTCATCCGGGAG	Clonado de RAN en pQE80
pQE80-BBS2 Fwd	CCTGTATTTTCAGGGATCCGGTTCAGTGCCTGCTGTGTTACCCCTG	Clonado de BBS2 en pQE80
pQE80-BBS2 Rev	GCTGCAGGTCGACGCCAAGATCCTTTTTTAGGAAGAAGCTGTCCCACTC	Clonado de BBS2 en pQE80
pACYCDuet S2 Fwd	ATTAGTTAAGTATAAGAAGGAGATATACATAGAGAGGATCGCATCACCATCACC	Primer genérico para clonar en el sitio 2 de pACYCDuet desde el pQE80
pACYCDuet S2 SIN1a Rev	CGCAGCAGCGGTTTCTTTACCAGACTCGAGTTACTGCTGCCCGGATTCTTC	Clonado de SIN1a en el sitio 2 de pACYCDuet desde el pQE80
pACYCDuet S2 RAN Rev	CGCAGCAGCGGTTTCTTTACCAGACTCGAGTTACAGGTCATCATCCTCATCCGG	Clonado de RAN en el sitio 2 de pACYCDuet desde el pQE80
pACYCDuet S2 BBS2 Rev	CGCAGCAGCGGTTTCTTTACCAGACTCGAGTTAGGAAGAAGCTGTCCCACTC	Clonado de BBS2 en el sitio 2 de pACYCDuet desde el pQE80







