

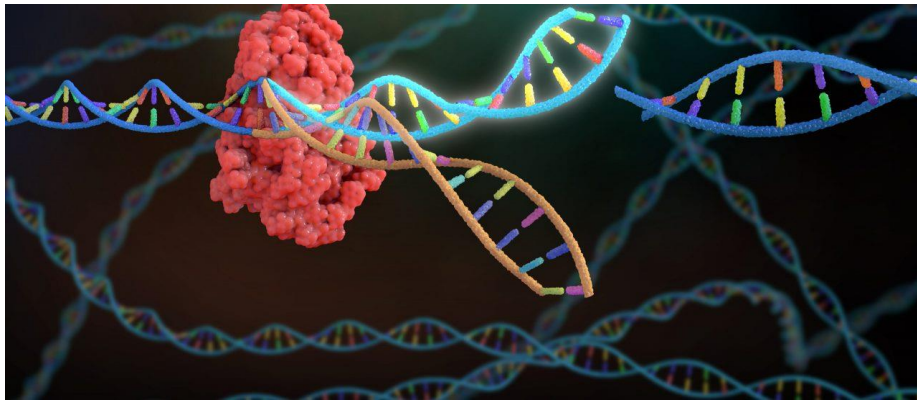


UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY



FACULTAD DE  
**CIENCIAS**  
UDELAR | [fcién.edu.uy](http://fcién.edu.uy)

“Construcción de un sistema *in vitro* para el etiquetado endógeno de AGO2 humana mediante edición genómica por CRISPR/Cas9 recombinante”



Tesis de grado - Licenciatura en Ciencias Biológicas

Laboratorio de Interacciones Moleculares

Facultad de Ciencias – UDELAR

2021

**Camilla Lucía Colantuono Seguí**

Tutora: Dra. María Ana Duhagón

Co-tutor: Lic. Juan Manuel Trinidad

# ÍNDICE

<b>Abreviaturas</b> .....	<b>3</b>
<b>Resumen</b> .....	<b>5</b>
<b>Introducción</b> .....	<b>7</b>
1.1 Cáncer .....	7
1.1.1 Cáncer de próstata .....	9
1.2 microARNs .....	11
1.2.1 Biogénesis y función .....	11
1.2.2 miARNs en el estudio y comprensión del cáncer de próstata .....	14
1.3 CRISPR/Cas9 .....	16
1.3.1 CRISPR/Cas y su rol en la inmunidad adquirida en procariotas .....	16
1.3.2 Clasificación de los sistemas CRISPR/Cas9 .....	18
1.3.3 Sistema CRISPR/Cas de tipo II .....	19
1.3.4 CRISPR/Cas9 como herramienta de edición genómica .....	21
1.3.5 Generación e introducción de los componentes a las células .....	23
1.3.6 Ventajas de CRISPR/Cas9 como herramienta de edición genómica .....	24
1.3.7 Dificultades y potenciales soluciones del sistema CRISPR/Cas9.....	26
<b>Objetivos</b> .....	<b>29</b>
2.1 Objetivo general .....	29
2.2 Objetivos específicos .....	29
<b>Materiales y métodos</b> .....	<b>30</b>
3.1 SaCas9 .....	30
3.1.1 Expresión .....	30
3.1.2 Purificación .....	32
3.2 sgRNA .....	33
3.2.1 Diseño .....	33
3.2.2 Síntesis .....	35
3.2.2.1 ADN molde para transcripción <i>in vitro</i> .....	35
3.2.2.2 Transcripción <i>in vitro</i> .....	36

3.3 ADN blanco .....	37
3.3.1 mCherry .....	37
3.3.2 AGO2 .....	38
3.4 Evaluación de la actividad de la SaCas9 .....	41
3.5 Purificación del ADN: fenol-cloroformo .....	42
3.6 Análisis de resultados mediante electroforesis .....	43
3.6.1 Electroforesis en gel de poliacrilamida .....	43
3.6.2 Electroforesis en gel de agarosa .....	44
3.7 Diseño ADN donante .....	45
<b>Resultados y discusión .....</b>	<b>46</b>
4.1 Diseño del experimento de edición genómica CRISPR/SaCas9 por HDR .....	46
4.1.1 sgRNA .....	46
4.1.2 ADN donante .....	49
4.2 Síntesis de los componentes CRISPR/SaCas9 .....	52
4.2.1 SaCas9 .....	52
4.2.2 sgRNA .....	55
4.2.3 ADN molde .....	56
4.3 Evaluación de la actividad de la SaCas9 .....	61
<b>Conclusiones y perspectivas .....</b>	<b>68</b>
<b>Referencias bibliográficas.....</b>	<b>70</b>
<b>Agradecimientos .....</b>	<b>78</b>
<b>Anexo .....</b>	<b>79</b>

## ABREVIATURAS

**A:** Adenina

**ADN:** Ácido Desoxirribonucleico

**AGO2:** Argonauta 2

**ARN:** Ácido Ribonucleico

**ARNasa A:** Ribonucleasa A

**ARNi:** ARN interferente

**ARNm:** ARN mensajero

**ATP:** Adenosín Trifosfato

**APS:** Ammonium Persulfate (Persulfato de Amonio)

**BrEt:** Bromuro de Etidio

**C:** Citosina

**CaP:** Cáncer de Próstata

**Cas:** CRISPR associated

**CMCaP:** Células Metastásicas de Cáncer de Próstata

**CRISPR:** Clustered Regularly Interspaced Short Palindromic Repeats (Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Espaciadas)

**crRNA:** CRISPR RNA (ARN asociado a CRISPR)

**Csp:** Cantidad suficiente para

**dCas9:** Dead Cas9 (Cas9 catalíticamente inactiva)

**DSB:** Double Strand Break (Rupture doble hebra)

**DO:** Densidad Óptica

***E. coli:*** *Escherichia coli*

**EDTA:** Ethylenediaminetetraacetic Acid (Ácido Etilenodiaminatetraacético)

**fCas9:** FokI-dCas9

**G:** Guanina

**g:** gramo

**HDR:** Homology Directed Repair (Reparación dirigida por homología)

**IPTG:** Isopropil- $\beta$ -D-1-tiogalactopiranosido

**IVT:** *In-Vitro* Transcription (Transcripción *in vitro*)

**kDa:** Kilodalton

**L:** Litro

**M:** Molar

**mg:** miligramo

**miARN:** microARN

**miRISC:** miRNA-Induced Silencing Complex (Complejo Silenciador Inducido por miARN)

**mL:** mililitro  
**mM:** milimolar  
**MPM:** Marcador de Peso Molecular  
**N:** Cualquier nucleótido  
**ng:** nanogramo  
**NHEJ:** Non-Homologous End-Joining (Unión de extremos no homólogos)  
**PAM:** Protospacer Adjacent Motif (Motivo adyacente a protoespaciador)  
**pb:** pares de bases  
**PCR:** Polymerase Chain Reaction (Reacción en cadena de la polimerasa)  
**R:** Purinas ( A o G)  
**RNP:** Ribonucleoprotein (Ribonucleoproteína)  
**rpm:** revoluciones por minuto  
**SaCas9:** *Staphylococcus aureus* CRISPR associated protein 9 (proteína 9 asociada a CRISPR de *Staphylococcus aureus*)  
**SDS-PAGE:** Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (Electroforesis en gel de poliacrilamida con dodecilsulfato sódico)  
**sgRNA:** Single Guide RNA (ARN guía único)  
**T:** Timina  
**TALEN:** Transcription Activator-Like Effector Nuclease (Nucleasa de actividad similar a activador de transcripción).  
**TEMED:** Tetramethylethylenediamin (Tetrametiletilendiamina)  
**TM:** Temperatura de Melting (Temperatura de fusión del ADN)  
**tracrRNA:** Trans-Activating CRISPR RNA (ARN transactivador asociado al CRISPR)  
**truRNA:** Truncated sgRNA (sgRNA truncado)  
**UA:** Unidades de Absorbancia  
**UTR:** Untranslated Region (Región no traducida)  
**ZFN:** Zinc Finger Nuclease (Nucleasa con dedos de zinc)  
**µg:** microgramo  
**µL:** microlitro  
**µM:** micromolar  
°C: grado Celsius

## RESUMEN

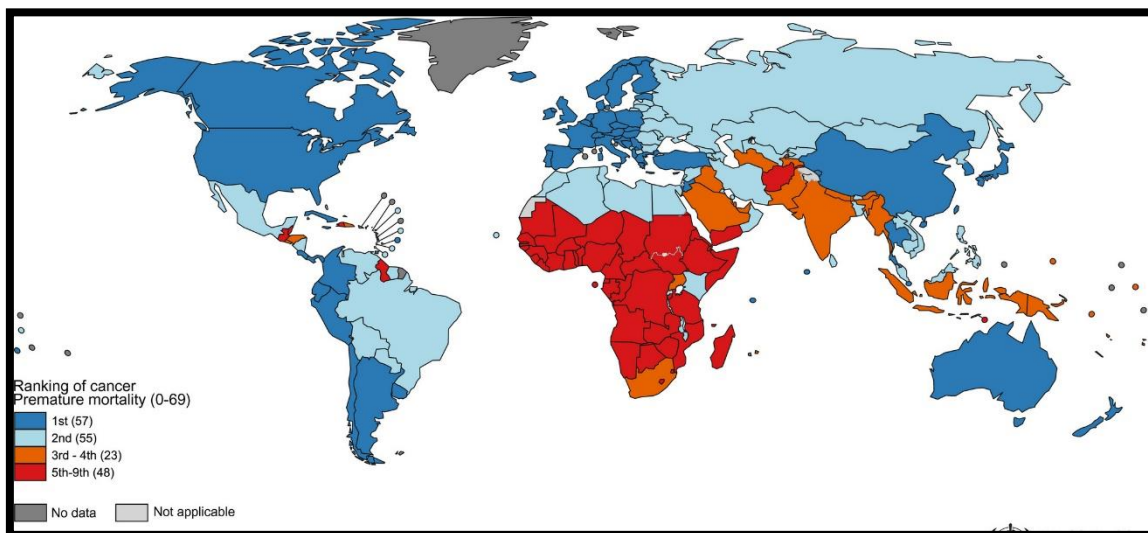
Los microARNs son pequeñas moléculas de ARN no codificante de aproximadamente 22 nucleótidos que actúan como importantes reguladores de la expresión génica a nivel post-transcripcional al asociarse con la proteína AGO2 formando el núcleo central del complejo miRISC. El complejo miRISC es el encargado de mediar el silenciamiento de los ARNm que contienen secuencias complementarias al microARN. Diversos estudios han demostrado un marcado vínculo entre la desregulación de expresión de microARNs y el desarrollo de varias enfermedades, entre las cuales se destaca el cáncer. El alto grado de complejidad de las interacciones microARN-ARNm representa un enorme desafío en estudios de identificación de ARNm candidatos a ser regulados por microARNs. La utilización de algoritmos informáticos como modelos predictivos de las interacciones suelen generar resultados positivos falsos, por lo que es necesario llevar a cabo estrategias experimentales complementarias, como la inmunoprecipitación de AGO2 (iCLIP, PAR-CLIP, CLASH) con la finalidad de obtener resultados más certeros. Este trabajo consistió en la construcción de un sistema *in vitro* para el etiquetado endógeno de la proteína AGO2 humana mediante edición genómica por CRISPR/SaCas9 recombinante. Esto haría posible la expresión de la proteína AGO2 desde su locus endógeno, asemejando su regulación y procesamiento natural de modo de reducir los artefactos de la expresión dirigida por promotores artificiales desde episomas o locus no naturales. En paralelo se construyó un sistema para editar el gen control mCherry que sirve de herramienta para evaluar sencillamente la eficiencia de la edición *in vivo*, asistiendo a la puesta a punto de protocolos de edición. El sistema de edición genómica CRISPR/Cas9 ha significado una gran innovación en el campo de la ingeniería genética como herramienta para la introducción de modificaciones secuencia específicas en el ADN. La demostración de que el gen AGO2 puede ser clivado *in vitro* por este sistema permitirá introducir el mismo en líneas celulares de cáncer de próstata y provocar la edición genómica por incorporación de un ADN molde con neo-epitope etiqueta, permitiendo así el análisis de ARNs mensajeros unidos a la proteína mediante ensayos de *pull-down*. Por otro lado, el corte *in vitro* del gen reportero mCherry otorgará una validación sencilla del sistema CRISPR/SaCas9 debido a la pérdida de fluorescencia en las células editadas. Para alcanzar el objetivo, se comenzó con la producción y purificación de la SaCas9. Esto fue realizado a partir de un plásmido comercial para expresión recombinante en *Escherichia coli* y luego se procedió a purificar la misma mediante cromatografía de afinidad. Posteriormente se evaluó la actividad de la SaCas9 *in vitro*. Para eso, se diseñaron los sgRNAs adecuados para introducir una etiqueta en el extremo N-terminal de la proteína AGO2 humana y para generar un knockout en mCherry mediante la incorporación de un péptido en posteriores ensayos *in vivo*. Además, se diseñaron los ADNs donantes correspondientes. Se demostró que los complejos ribonucleoproteicos SaCas9-sgRNA ejercieron *in vitro* un corte doble hebra específico en el gen control mCherry, evidenciando no sólo el correcto funcionamiento de la nucleasa sino también el diseño certero del sgRNA. Sin embargo, no se pudo cortar *in*

*in vitro* el gen AGO2, lo que podría indicar posibles limitaciones del sistema CRISPR/Cas9 frente a secuencias de ADN estructuralmente complejas. Este último resultado incita a continuar experimentando, optimizar estrategias y buscar alternativas para un gen en el que aún no se ha publicado un etiquetado endógeno.

# INTRODUCCIÓN

## 1.1 Cáncer

El cáncer se ubica como una de las principales causas de muerte y representa una barrera importante para aumentar la esperanza de vida en todos los países del mundo. Según estimaciones de la Organización Mundial de la Salud (OMS) en 2019, el cáncer es la primera o la segunda causa de muerte antes la edad de 70 años en 112 de 183 países y ocupa el tercer o cuarto lugar en otros 23 países (**Figura 1**) (Sung et al., 2021). La tasa de incidencia y mortalidad por cáncer está creciendo rápidamente; esto refleja tanto el envejecimiento y el crecimiento de la población como los cambios en la prevalencia y distribución de los principales factores de riesgo, varios de los cuales están asociados con el desarrollo socioeconómico (Sung et al., 2021).

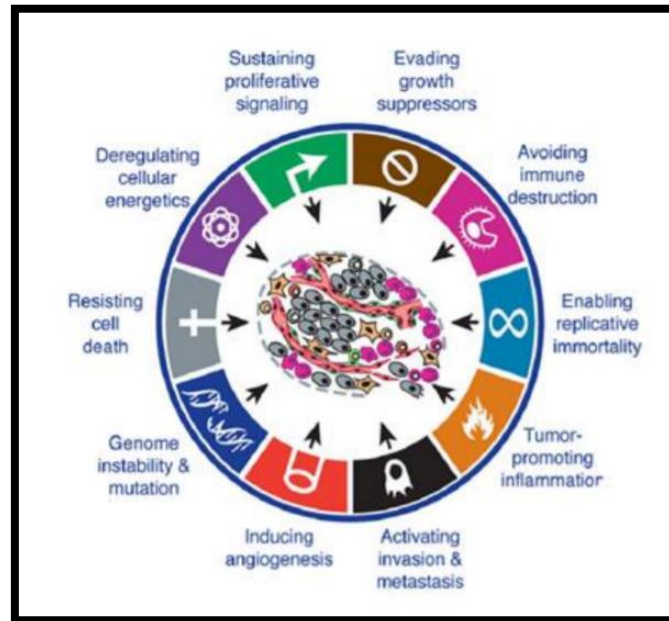


**Figura 1.** Ranking internacional del cáncer como causa de muerte en las edades <70 años en 2019. Fuente: Organización Mundial de la Salud (Sung et al., 2021).

El término cáncer hace referencia a un conjunto de enfermedades que se caracteriza por un crecimiento desregulado de las células que componen los diferentes tejidos del organismo, las cuales adquieren la capacidad de modificar el tejido de origen y también de propagarse invadiendo nuevos tejidos, tanto a nivel local como a zonas distantes (Hanahan & Weinberg, 2011). Las células afectadas presentan una acumulación de mutaciones en genes claves denominados “supresores de tumor” u “oncogenes” (Kent & Mendell, 2006). Durante el proceso de transformación maligna, las células adquieren una o más características que constituyen lo que se denominan las firmas o sellos distintivos del cáncer (en inglés, Hallmarks of Cancer) y proporcionan una base conceptual para la comprensión de la biología del cáncer. Estos pueden resumirse en: autosuficiencia en el desarrollo de señales de crecimiento y mantenimiento de señales de proliferación, evasión de las señales



antiproliferativas, activación y adquisición de capacidad invasiva (metástasis), desarrollo de una capacidad proliferativa inmortal, inducción de la angiogénesis y resistencia a la muerte celular (Hanahan & Weinberg, 2011). Estas firmas están siendo continuamente revisadas y ampliadas, en particular destacando la relevancia del ambiente tumoral en el desarrollo del cáncer (**Figura 2**) (Hanahan & Coussens, 2012).



**Figura 2.** Esquema de las principales características o “hallmarks” de las células cancerígenas. Fuente: Hanahan & Coussens, 2012.

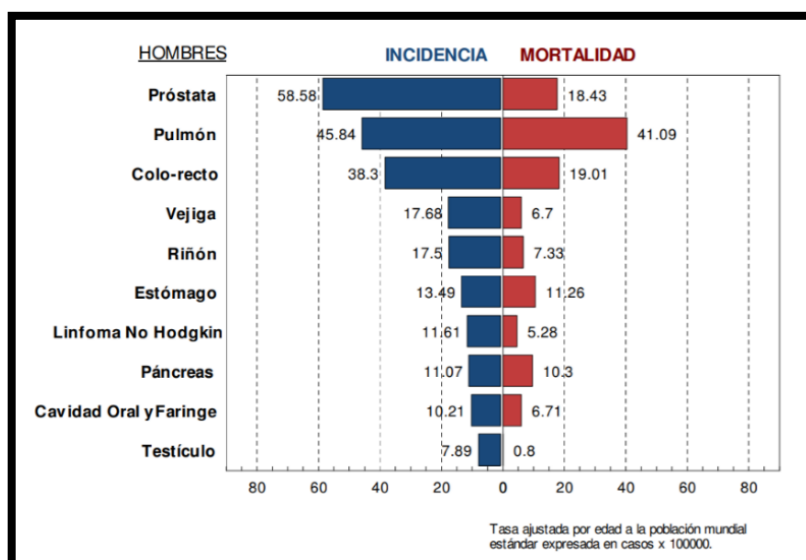
El desarrollo del cáncer es consecuencia de dos componentes: uno es hereditario, es decir, que existan mutaciones deletéreas en los parentales en la secuencia de genes que predisponen a la enfermedad, y el otro es esporádico donde se producen mutaciones espontáneas generadas por factores ambientales, siendo estos últimos los más comunes en el desarrollo de la enfermedad (Torre et al., 2015). Existen dos formas de clasificación de los tumores basadas en la agresividad de la expansión celular: si el tumor se encuentra localizado y no ha invadido otros tejidos es denominado “benigno”, mientras que el tumor que ha logrado invadir tejidos cercanos generando metástasis se lo denomina “maligno” (Weingberg et al., 2007). Un cáncer se vuelve metastásico cuando las células neoplásicas logran migrar (generalmente por vía sanguínea o linfática) a diferentes tejidos desencadenando tumores en aquellos que logran “colonizar”. El evento de colonización no es simple para la célula metastásica debido a que requiere un ambiente de señalización propicio para su desarrollo y multiplicación, entre otra gran cantidad de variables (Weingberg et al., 2007). Actualmente, se conocen los genes implicados en la metástasis y son clasificados en tres categorías según su contribución en el desarrollo. Los genes implicados en invasión e intravasación son denominados “Metastasis initiation genes”. En la categoría “Metastasis progression genes” se encuentran los de función dual que median la tumorigénesis primaria y la colonización de otros tejidos. Por

último, se encuentran los genes involucrados exclusivamente en la invasión de órganos distantes en la clasificación “Metastasis virulence genes” (Nguyen et al., 2007).

Los tumores presentan un perfil mutacional propio, constituido inicialmente por una población homogénea de células tumorales. Debido a la alta inestabilidad genómica y la alta proliferación, durante el desarrollo del tumor se generan nuevas subpoblaciones celulares que exhiben múltiples y variados fenotipos. A su vez, dentro de un mismo tumor se ha identificado una población denominada “Cancer Stem Cells” (CSC), la cual se distingue por presentar características de células pluripotentes y tener la capacidad de generar nuevos tumores cuando se inoculan en otros organismos (Hanahan et al., 2011). La inestabilidad genómica en estas poblaciones celulares modifica constantemente el perfil molecular de los tumores, lo cual constituye una de las razones en la generación de resistencia a tratamientos, estimulando al desarrollo de terapias más dirigidas con blancos específicos en cada tejido y para cada tumor (Ost et al., 2015).

### 1.1.1 Cáncer de próstata

Con un estimado de casi 1,4 millones de casos nuevos y 375.000 muertes en todo el mundo, el cáncer de próstata (CaP) es el segundo cáncer más frecuente y la quinta causa principal de muerte por cáncer entre los hombres en 2020. Además, es el cáncer diagnosticado con más frecuencia en hombres en más de la mitad (112 de 185) de los países del mundo (Sung et al., 2021). Particularmente en Uruguay, el CaP representa la tercera causa de mortalidad por cáncer en hombres luego del cáncer de pulmón, con casi 580 muertes anuales (**Figura 3**). Tanto las tasas de incidencia como de mortalidad de Uruguay se ubican en el segundo quintil superior a nivel global. Es conocida la fuerte dependencia del riesgo de padecer este cáncer en función de la edad, el 50% de los casos se diagnostican en mayores de 70 años (Ferlay et al., 2010).



**Figura 3.** Índice y tasa de incidencia/mortalidad del CaP en Uruguay. Fuente: Registro Nacional de Cáncer.

El origen tisular de este tipo de cáncer parece encontrarse en las células luminales y epiteliales, aunque aún es controversial. También se especula que puede ser una evolución de las células basales iniciales neoplásicas hacia un adenocarcinoma, mantenidas por células tipo luminal (Wang et al., 2009). Múltiples estudios han identificado mutaciones somáticas recurrentes, alteraciones en el número de copias y reordenamientos estructurales oncogénicos del ADN en el CaP primario. Estos incluyen mutaciones puntuales en SPOP, FOXA1 y TP53; alteraciones del número de copia que involucran a MYC, RB1, PTEN y CHD1; y fusiones específicas de transformación E26 (ETS), entre otros genes biológicamente relevantes (“The Molecular Taxonomy of Primary Prostate Cancer”).

En la metástasis del CaP los lugares más frecuentes de invasión son el tejido óseo (caracterizado por las lesiones osteoblásticas) y los ganglios linfáticos (Bubendorf et al., 2000). La combinación entre la pérdida de función del gen PTEN y la ganancia en número de copias de MYC son las variaciones genéticas que contribuyen a la inestabilidad genómica y conducen a metástasis que suele ser letal en los pacientes (Miyahira et al., 2016). Durante las primeras etapas las células metastásicas de CaP (CMCaP) se caracterizan por la gran expresión de genes embrionarios, anti-apoptóticos, de quiescencia y genes implicados en la transición epitelio-mesénquima (TEM). Adicionalmente, se evidenció un descenso en la expresión de genes marcadores de células luminales (Miyahira et al., 2016). En los estadios más terminales de las CMCaP la expresión se vuelve similar a las células parentales, es decir, a las neoplásicas localizadas que dieron lugar a estas metastásicas, y se caracteriza por altos niveles de expresión de genes luminales (Miyahira et al., 2016).

A pesar de los avances quedan desafíos por superar en el tratamiento del CaP. Aún no existen tratamientos que prometan una mayor efectividad en pacientes en situación metastásica. Por cierto, a lo largo de la última década el tratamiento utilizado por excelencia es la privación de andrógeno (PA), que, a su vez, tiene importantes efectos secundarios en la esperanza y calidad de vida de las personas tratadas. A los pacientes con oligo-metástasis (metástasis con poco desarrollo), además de tratarlos con PA, se los trata de forma directa con cirugías o radioterapia brindando mejores resultados que las personas con una metástasis ampliamente desarrollada (Crawford et al., 2018). Por otro lado, aunque la detección del antígeno específico de próstata (en inglés, Prostate Specific Antigen o PSA) ha facilitado la detección temprana del CaP en una etapa curable de la enfermedad, su baja especificidad y sensibilidad hacen necesario la identificación de nuevos marcadores para reducir la morbilidad y evitar tratamientos innecesarios (Vanacore et al., 2017). Recientemente, un ARN no codificante ha sido aprobado por la FDA como un biomarcador del CaP en orina. Se trata del antígeno de cáncer de próstata 3 (en inglés, Prostate Cancer Antigen 3 o PCA3), el cual se ha demostrado que se encuentra elevado en más del 90% de los hombres con CaP, no así en la glándula prostática normal o en la hipertrofia prostática benigna (Crawford et al., 2014).

## 1.2 microARNs

### 1.2.1 Biogénesis y función

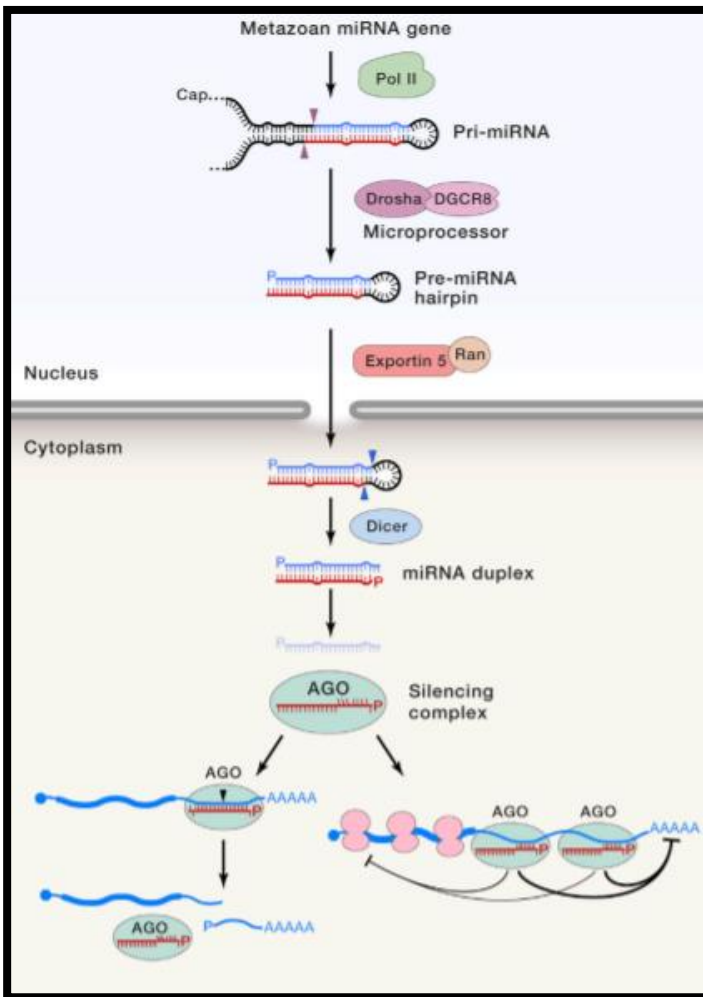
Los microARNs (miARNs) son pequeñas moléculas de ARN no codificante de aproximadamente 22 nucleótidos que actúan como importantes reguladores de la expresión génica a nivel post-transcripcional, siendo fundamentales para la homeostasis fisiológica y el desarrollo en los metazoarios (Sheu-Gruttadauria & MacRae, 2017). Lin-4 fue el primer miARN que se describió y se identificó como un elemento crucial en regular las distintas etapas de los estadios larvarios del nemátodo *Caenorhabditis elegans* (Bartel, 2004). Hoy día, se estima que el genoma humano contiene más de 1.000 genes que codifican para miARNs, representando aproximadamente el 3% de los genes conocidos (Gebert & MacRae, 2019). Los genes codificantes para miARNs se encuentran dispersos a lo largo de todo el genoma, pudiendo estar codificados en forma monocistronica o en clústeres con otros miARNs. Los miARNs presentes en clústeres representan el 37% de miARNs en humanos, y se cree que existe como mecanismo para coordinar de forma más eficiente la regulación de procesos celulares (Altuvia et al., 2005). En lo que refiere a su localización, los miARNs pueden estar en intrones, donde suelen presentar el mismo patrón de expresión de los genes en que residen, aunque también pueden encontrarse formando parte de exones de genes no codificantes. Por otro lado, existen miARNs que se ubican en zonas intergénicas conteniendo sus propios elementos reguladores, promotores y señales de terminación, generalmente para la ARN polimerasa II (Bartel, 2018).

La biogénesis de los miARNs requiere de varias etapas que tienen lugar en el núcleo y en el citoplasma (**Figura 4**) (Bartel, 2018). El proceso de transcripción es llevado a cabo por la enzima ARN polimerasa II para dar lugar a miARNs primarios (pri-miARNs) que forma estructuras características en forma de horquilla o stem-loop (Gebert & MacRae, 2019). Durante la biogénesis canónica estas estructuras primarias son procesadas por la ribonucleasa Drosha para producir precursores de miARNs de 70 nucleótidos de largo, conocidos como pre-miARNs, que conservan la forma de horquilla y se unirán a Exportina-V para ser exportados al citoplasma. Una vez en el citoplasma el pre-miARN es reconocido y procesado por la enzima Dicer, generando un miARN maduro de 22 nucleótidos (Bartel, 2018). En la vía no canónica, se generan pre-miARNs por la maquinaria de empalme de ARNm, independientemente de Drosha. Los pasos siguientes son idénticos en las vía canónica y no canónica (Oliveto et al., 2017). Para poder ejercer su función reguladora, una de las hebras del miARN se asocia con proteínas Argonauta (AGO) formando el núcleo central del complejo de silenciamiento miRISC (miRNA-Induced Silencing Complex) mientras que la otra hebra suele ser degradada (**Figura 4**) (Bartel, 2018). El genoma de los mamíferos codifica cuatro proteínas AGO (AGO1-4) siendo AGO2 la más expresada y la única de las cuatro capaz de escindir ARN (Liu et al., 2004). Esta proteína está compuesta por cuatro dominios característicos: el dominio amino-terminal (N), el dominio Piwi-Argonaute-Zwille (PAZ), el dominio del medio (MID) y el dominio PIWI. Dos dominios

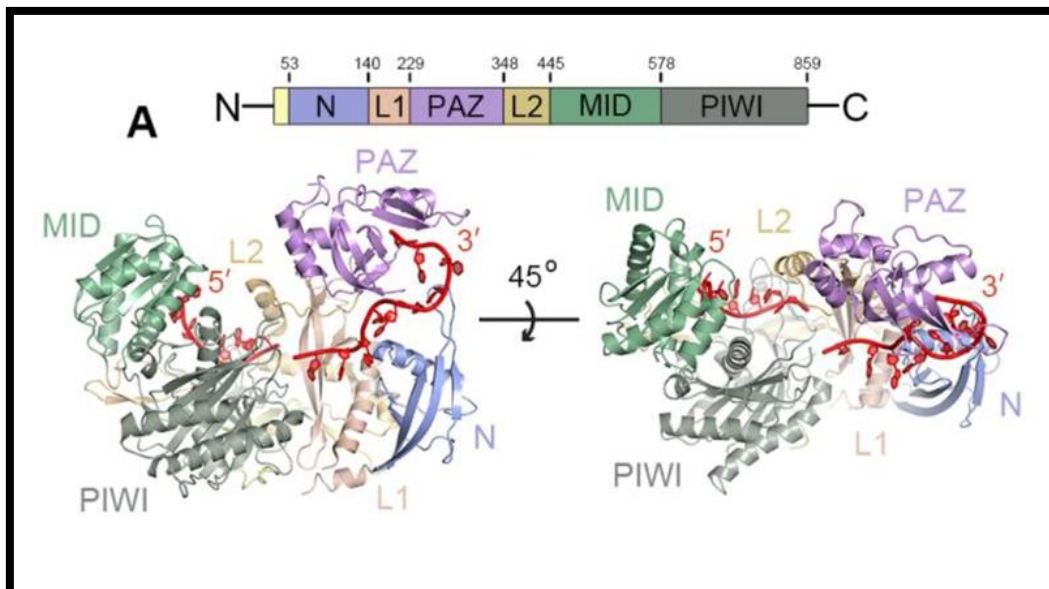
enlazadores (L1 y L2) conectan los dominios N y PAZ (L1) y los dominios PAZ y MID (L2). Los dominios N y PAZ forman un lóbulo de AGO y los dominios MID y PIWI forman otro segundo lóbulo. Los dominios MID y PIWI contienen el extremo 5' del miARN, mientras que el dominio PAZ se une al extremo 3' (**Figura 5**) (Schirle & MacRae, 2012). Los complejos miRISC son los encargados de mediar el silenciamiento de aquellos ARNm que contienen secuencias que son total o parcialmente complementarias al miARN (Jonas & Izaurralde, 2015). Los miARN se unen a la región 3' UTR (del inglés, UnTranslated Region) del ARNm diana a través de los nucleótidos 2-8 de su extremo 5', región denominada “semilla” (Bartel, 2009).

Los miARN silencian la expresión génica reprimiendo la traducción y promoviendo la degradación del ARNm blanco (**Figura 4**) (Jonas & Izaurralde, 2015). Generalmente, la regulación basada en la estabilidad del ARNm en humanos y otros animales bilaterales, comienza mediante el reclutamiento por parte de AGO de la proteína adaptadora GW182 (Bartel, 2018). Dicha proteína se caracteriza por contener múltiples repetidos de glicina-triptofano (GW) y su participación posee un rol fundamental ya que la formación de los complejos AGO-GW182 da inicio al mecanismo de regulación (Niaz & Hussain, 2018). GW182 se une a las proteínas AGO con su dominio N-terminal, mientras la parte C-terminal establece interacciones con la proteína de unión a poli A en el extremo 3' del ARNm, desencadenando la degradación de este por el reclutamiento de exoribonucleasas (Gebert & MacRae, 2019). Por otro lado, varios estudios han demostrado la participación de los miARNs como represores del inicio del proceso de traducción. Se han llevado a cabo trabajos experimentales en mamíferos que ponen en evidencia la existencia de una unión entre la proteína AGO2 y el CAP en el extremo 5' del ARNm sobre el que actúa sugiriendo una inhibición del inicio de la traducción generado por la competencia entre AGO2 y el ribosoma por el sitio de unión a CAP, entre otros mecanismos. Adicionalmente, se ha visto que los miARNs pueden inhibir la traducción en pasos posteriores a la iniciación, por ejemplo, provocando el desensamblaje del ribosoma durante la elongación (Oliveto et al., 2017).

Actualmente existe vasta evidencia que demuestra la participación de los miARNs en la mayoría de los procesos celulares como el desarrollo animal, diferenciación celular y la homeostasis. Además, las enzimas fundamentales en la biogénesis de los miARN, Dicer y Drosha, parecen estar involucradas en procesos nucleares esenciales como el splicing del pre-ARNm. La implicancia de los miARN en los mencionados procesos se pone de manifiesto cuando se generan deleciones de Dicer y/o Drosha en embriones de ratón ya que se obtiene un fenotipo letal (Gebert & MacRae, 2019). Los miARNs son agrupados en familias respondiendo a un modelo funcional donde se utiliza principalmente la identidad de la secuencia semilla. Al igual que con las proteínas parálogas, los miembros de las mismas familias de semillas a menudo tienen funciones al menos parcialmente redundantes, por ejemplo, regulación del sistema inmunológico, desarrollo o cáncer (Bartel, 2018).



**Figura 4.** Imagen ilustrativa de la biogénesis y función de un miARN típico. Una vez transcrito por la ARN polimerasa II, el pri-miARN se pliega sobre sí mismo para formar una estructura de horquilla distintiva que es escindida por Drosha (las puntas de flecha púrpura son sitios de escisión) para liberar un pre-miARN. La horquilla de pre-miARN se exporta al citoplasma a través de la acción de Exportina 5 y RAN-GTP. En el citoplasma, el pre-miARN es escindido por Dicer (las puntas de flecha azules son sitios de escisión) para producir un dúplex de miARN de 20 pb, el miARN maduro (rojo), se carga en una proteína AGO para formar el complejo silenciador miRISC, mientras que la hebra restante (azul) se degrada. Dentro del complejo de silenciamiento, los nucleótidos de miARN 2-5 (líneas verticales rojas hacia arriba) están preparados para interactuar inicialmente con el ARN diana. Este emparejamiento suele extenderse al nucleótido 7 u 8 del miARN y, en ocasiones, es más extenso. El ARNm objetivo puede cortarse (izquierda; punta de flecha negra es el sitio de corte), o puede sufrir otros tipos de represión (derecha). Fuente: Bartel, 2018.



**Figura 5.** Estructura proteína AGO2 humano. **A)** Esquema de la secuencia primaria de AGO2 **B)** Vista frontal y superior de AGO2 unido a un miARN (rojo). AGO2 contiene una gran hendidura central entre dos lóbulos (N-PAZ y MID-PIWI) conectados por dos dominios enlazadores (L1 Y L2) Fuente: Schirle et al., 2014.

### 1.2.2 miARNs en el estudio y comprensión del cáncer de próstata

Los miARNs no solo influyen esencialmente en procesos celulares importantes sino también en diferentes enfermedades, incluido el cáncer (Bartel, 2018). La primera evidencia sobre la asociación de miARNs con el cáncer deriva de estudios sobre leucemia linfocítica crónica (LLC) siendo el grupo de Croce quien descubrió que dos miARNs denominados miR-15a y miR-16-1 los cuales derivan del mismo ARN policistrónico que se transcribe a partir de una región específica del cromosoma 13, con frecuencia se encuentra deletado en la LLC (Calin et al., 2002). Por otro lado, observaron que un porcentaje significativo de genes de miARN se localiza en sitios frágiles y/o en regiones del genoma que a menudo muestran alteraciones cromosómicas, incluidas amplificaciones o deleciones (Calin et al., 2002). Dicho hallazgo en conjunto con otras investigaciones posteriores, sugirieron que los miARN representan una clase de genes involucrados en procesos como la regulación de la patogénesis y el desarrollo del cáncer cuando su nivel de expresión se encuentra alterado (Oliveto et al., 2017). Estos miARNs desregulados pueden actuar como miARNs oncogénicos (oncomiRs) o como miARNs supresores de tumores basados en su capacidad para reprimir la expresión de genes supresores de tumores u oncogenes, respectivamente. La inhibición o estimulación de oncomiRs o miARNs supresores de tumores modula la proliferación de células cancerosas, el crecimiento tumoral, la formación de metástasis y la supervivencia celular (Svoronos et al., 2016). Experimentos de *knock-down* en oncomiRs regulados positivamente, han demostrado una reducción en la proliferación celular y el crecimiento tumoral tanto en modelos tumorales in vitro como in vivo (Iorio et al., 2005; Ventura et al., 2008). El desarrollo de tecnologías de secuenciación masiva ha permitido estudiar lo que se conoce como el miRNoma en tejidos normales y patológicos. El miRNoma es el estudio de todos los miARNs presentes en los distintos tejidos. Estos estudios han aportado datos que demuestran que los perfiles de expresión de miARNs pueden considerarse candidatos para el pronóstico, diagnóstico y tratamiento del cáncer (Yong & Dutta, 2009). Por ejemplo, se ha evidenciado que el grupo miR-130b/301b está regulado positivamente en pacientes con CaP con relación a personas sanas. Además, se observó una expresión mucho mayor de estos miARNs en sitios metastásicos que, en sitios primarios, pudiendo actuar entonces como posibles indicadores de malignidad de la enfermedad (Fort et al., 2018). Por otro lado, miR-148a también tiene un rol importante en el inicio de CaP; cuando se dirige al gen CDKN1B y cuando interactúa con CENPF podría inhibir la invasión (Hamilton et al., 2016). Asimismo, los supresores tumorales miR-15 y miR-16-1, cuyo objetivo es Bcl-2, están regulados negativamente no solo en CaP sino también en varios tipos de cánceres como mesotelioma y LCC (Reid, 2015; Aqeilan et al., 2010). Otro ejemplo relevante está representado por miR-125, que muestra propiedades supresoras de tumores en varios cánceres, como melanoma, osteosarcoma, cáncer de ovario y de mama, y funciones promotoras de tumores en cánceres de páncreas y próstata (Shaham et al., 2012; Sun et al., 2013). Diversos estudios han probado que varios miARNs están involucrados en la patogenia

del CaP ya que ejercen sus efectos al dirigirse a una variedad de vías moleculares (Aghdam et al., 2019). Con el propósito de obtener una mejor comprensión acerca de las vías moleculares que se encuentran alteradas en CaP, resulta pertinente la identificación de genes blanco que están siendo silenciados por miARNs desregulados en dicha enfermedad. Por ejemplo, es importante el estudio de las vías moleculares requeridas para el mantenimiento de la transformación maligna siendo crucial para el desarrollo de terapias específicas efectivas (Gandhi et al., 2018).

A pesar de varios años de esfuerzo para comprender mejor los factores subyacentes y los contextos involucrados en la inhibición eficiente de los ARNm de mamíferos por parte de los miARNs, todavía es un desafío identificar interacciones funcionalmente relevantes (Bartel, 2018). El alto grado de complejidad de las interacciones que ocurren entre los miARNs y los ARNm, representa un enorme desafío en estudios de identificación de genes candidatos a ser regulados por miARNs. Por ejemplo, recientemente se ha visto que el apareamiento insuficiente de la región semilla del miARN con el ARNm blanco puede ser compensada por una alta complementariedad entre ambos en la región 3' del miARN (Costa & De Lima, 2013). El apareamiento imperfecto y la secuencia semilla de 6-8 nucleótidos relativamente corta permiten una multitud de potenciales ARNm blancos por cada miARN. Adicionalmente, un único ARNm puede tener múltiples sitios blanco para múltiples miRNAs, creando una red molecular redundante para el control de la expresión génica (Hayes et al., 2014). En este sentido, la utilización de algoritmos informáticos como modelos predictivos de las interacciones miARNs-ARNm suelen generar resultados falsos positivos. Es por esto que, aunque el desarrollo de herramientas de predicción computacional implicó un gran paso en el descubrimiento de genes diana de miARNs, es necesario llevar a cabo estrategias experimentales complementarias como la inmunoprecipitación de AGO2 (iCLIP, PAR-CLIP, CLASH) con la finalidad de obtener resultados experimentales certeros (Clark et al., 2014). Pese a la disponibilidad del anticuerpo monoclonal anti AGO2 (#14-6519-82, Thermo Scientific™) único utilizado en la literatura, existe evidencia que demuestra la capacidad del mismo de establecer reactividad cruzada con otras proteínas, lo que conduce a hallazgos inespecíficos que requerirán estudios adicionales (Van Eijl et al., 2017). De esta forma, resulta conveniente el establecimiento de un etiquetado de AGO2 (fusión de un péptido inmunogénico) para garantizar la especificidad de los experimentos de inmunoprecipitación. El etiquetado de una proteína puede realizarse tanto en construcciones plasmídicas como por modificación de la secuencia endógena del gen AGO2, mediante el sistema CRISPR/Cas9. La primera de las opciones resulta ser metodológicamente más accesible, pero tiene la desventaja de producir cantidades de proteína distintas a las fisiológicas. Además, en el caso de integrarse en el genoma se ubicará al azar perdiendo entonces la geografía nuclear que es relevante para su regulación. La edición genómica presenta mayor dificultad técnica, pero permite mantener el locus lo que permite mantener mejor su procesamiento y regulación nativa. Esto es esencial



cuando se estudian proteínas y estas alteraciones provocan observaciones artefactuales.

## 1.3 CRISPR/Cas9

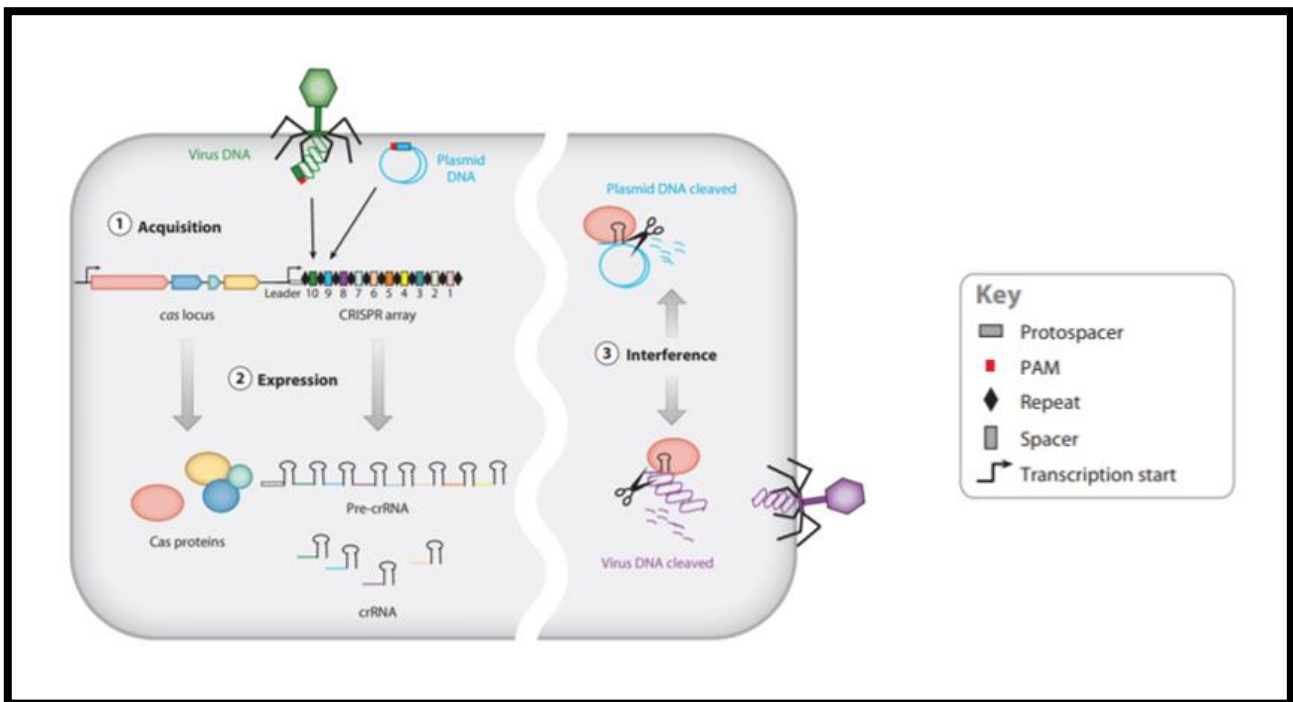
### 1.3.1 CRISPR/Cas y su rol en la inmunidad adquirida en procariontas

En el año 1987 se reportó la existencia de una secuencia de ADN repetitiva inusual de función desconocida en el genoma de *Escherichia coli* (Ishino et al., 1987). Tiempo después, Mojica notifica la presencia de estos elementos repetidos en diversos genomas procariontas de grupos filogenéticos ampliamente distribuidos incluyendo arqueas, bacterias gram positivas y negativas (Mojica et al., 2000). Estos repetidos cortos que consisten en motivos parcialmente palindrómicos de 24 a 40 nucleótidos, se encuentran regularmente espaciados por secuencias variables las cuales congruentemente reciben el nombre de espaciadores. Si bien en principio estos elementos repetidos fueron denominados SRSR (del inglés, Short Regularly Spaced Repeats), finalmente los investigadores acordaron hacer referencia a los mismos bajo la denominación CRISPR (del inglés, Clustered Regularly Interspaced Short Palindromic Repeats) (Mojica et al., 2000). Los repetidos CRISPR generalmente se encuentran cercanos a un operón con genes que codifican para un grupo de proteínas denominadas Cas (del inglés, CRISPR associated) (Jinek et al., 2012).

El hecho de que estos elementos sean la familia más ampliamente distribuida de repetidos entre los genomas procariontas sugería una función biológica de los mismos. Además, la ubicación de los genes Cas invariablemente adyacentes a un locus CRISPR, indicaban una relación funcional entre ambos (Jansen et al., 2002). Mediante análisis de secuencia del ADN espaciador ubicado entre los repetidos CRISPR se observó que la mayoría de éstos tenían homología con secuencias preexistentes, principalmente con elementos genéticos transmisibles tales como bacteriófagos y plásmidos conjugativos (Mojica et al., 2005). La mayoría de los genes con los que se observó homología estaban directamente involucrados en la transferencia plasmídica, el ensamblaje de los virus, la protección contra enzimas de restricción, síntesis del pili, transposición o integración del fago (Mojica et al., 2005; Jinek et al., 2012). Estos datos en conjunto con la descripción de dominios helicasa y endonucleasa en las proteínas Cas y teniendo en cuenta que elementos extracromosómicos como plásmidos o bacteriófagos no logran reinfectar la cepa portadora de una secuencia espaciadora específica, llevaron a determinar que el sistema CRISPR/Cas estaría involucrado en desempeñar un rol de “inmunidad adaptativa” en procariontas (Makarova et al., 2006).

La inmunidad mediada por CRISPR/Cas contra elementos genéticos invasores ocurre en tres etapas definidas (**Figura 6**). En la primera etapa, denominada ADQUISICIÓN, se adquieren nuevos espaciadores a partir de elementos invasivos y se integran en el locus CRISPR. Esta secuencia espaciadora cuando se encuentra presente en el ADN

invasor se denomina protoespaciador. En la siguiente etapa, fase de EXPRESIÓN, ocurre la transcripción de repetidos y espaciadores para dar lugar a moléculas precursoras denominadas pre-crRNA (pre-CRISPR RNAs). Seguido de un procesamiento enzimático se producen los crRNAs (CRISPR RNAs) maduros. Finalmente, en la etapa de INTERFERENCIA estos crRNAs que son ARNs antisentido hibridan con secuencias protoespaciadoras complementarias, dirigiendo el silenciamiento de estas secuencias exógenas a través de un corte doble hebra mediado por las proteínas Cas que funcionan en complejo con los crRNAs (Barrangou et al., 2013). Así es que el locus CRISPR/Cas está compuesto por un arreglo de repetidos idénticos intercalados por espaciadores homólogos a secuencias de ADN invasor que codifican los componentes del crRNA, y por un operón de genes que codifican los componentes de las proteínas nucleasas Cas (Doudna & Charpentier, 2014). Por lo tanto, los loci que albergan al sistema CRISPR son capaces de recopilar y almacenar fragmentos del ADN de bacteriófagos y plásmidos invasores, de esta forma estableciendo una memoria molecular de las invasiones pasadas (Makarova et al., 2006).



**Figura 6.** Ilustración del proceso de generación de inmunidad adaptativa a través de CRISPR/Cas en procariotas. **Etapa 1:** Se adquieren fragmentos específicos o protoespaciadores de ADN bicatenario de un virus o plásmido en el extremo líder del arreglo CRISPR en el ADN del huésped. El arreglo CRISPR consta de espaciadores únicos (cuadros de colores; los espaciadores se numeran secuencialmente con el espaciador adquirido más recientemente con el número más alto) intercalados entre repeticiones (diamantes negros). **Etapa 2:** El pre-crRNA se transcribe de la región líder por la ARN polimerasa y luego se escinde en crRNA más pequeño que contienen un solo espaciador y una repetición parcial (estructuras en horquilla con espaciadores de colores) por las proteínas Cas. **Etapa 3:** El crRNA que contiene un espaciador que tiene una fuerte correspondencia con el ácido nucleico extraño entrante (plásmido o virus) inicia un evento de escisión (mostrado con tijeras); Las proteínas Cas son necesarias para este

proceso. La escisión del ADN interfiere con la replicación del virus o la actividad del plásmido e imparte inmunidad al huésped. La interferencia se puede separar mecánicamente y temporalmente de la adquisición y expresión de CRISPR (representada por una barra blanca ondulada en la celda). Fuente: Bhaya et al., 2011.

### 1.3.2 Clasificación de los sistemas CRISPR/Cas

En un sistema de clasificación basado en múltiples criterios, incluidas las relaciones evolutivas de las proteínas y la organización del operón Cas, el sistema CRISPR/Cas puede dividirse en dos subsistemas parcialmente independientes. El primero consiste en un módulo de procesamiento de información y requiere de las proteínas Cas1 y Cas2, las cuales están involucradas en la adquisición de nuevas secuencias de ADN exógeno invasor o espaciadores en el locus CRISPR. Por otro lado, el segundo subsistema se basa en un módulo efector involucrado en el procesamiento de los transcritos CRISPR primarios (pre crRNAs), el reconocimiento y la consecuente degradación del ácido nucleico invasor. Aunque el primer módulo es universal, el módulo efector divide a los sistemas CRISPR/Cas en clase I y II (Bhaya et al., 2011).

La clase I comprende los tipos I y III, los cuales cuentan con complejos efectores formados por muchas subunidades de diversas proteínas Cas. Estos sistemas requieren de proteínas Cas especializadas que procesan los crRNAs, y una vez maduros, cada crRNA se ensambla con múltiples proteínas Cas en un gran complejo, capaz de reconocer y clivar el ADN foráneo complementario al crRNA (Jinek et al., 2012; Doudna & Charpentier, 2014). El sistema de tipo I es el más diverso y se caracteriza por la presencia de la proteína multidominio Cas3. Esta proteína posee actividades de ADN nucleasa y helicasa, y forma complejos tipo CASCADA (del inglés, CRISPR Associated Complex for Antiviral Defense) con múltiples proteínas Cas involucradas en el proceso de interferencia. Muchas de estas proteínas se encuentran en distintas superfamilias de RAMP (del inglés, Repeat-Associated Mysterious Protein) como la Cas5, Cas6, Cas7, entre otras (Makarova et al., 2011). El sistema CRISPR tipo III incluye la proteína RAMP Cas10, la cual está involucrada en la escisión del ADN diana siendo funcionalmente análoga a la CASCADA Tipo I. Este sistema también contiene la proteína Cas6, involucrada en el procesamiento de crRNA y en eventos adicionales (Anantharaman V et al., 2011). Hasta ahora, dos sistemas de tipo III se han distinguido (Tipo IIIA y IIIB). En *Pyrococcus furiosus*, un sistema de tipo IIIA, el ácido nucleico objetivo de la interferencia es el ARNm, mientras que en *Staphylococcus epidermidis*, un sistema de Tipo IIIB, el objetivo es el ADN. Esto resalta la naturaleza polimórfica de los sistemas CRISPR-Cas, incluso dentro de los sistemas Tipo III (Bhaya et al., 2011).

Por otro lado, la clase II incorpora al sistema CRISPR/Cas tipo II que se caracteriza por poseer complejos efectores con una única proteína Cas. Esta nucleasa llamada Cas9 se asocia a una molécula de ARN dual conformada por el crRNA y el ARN transactivador (del inglés, trans activating crRNA; tracrRNA). El tracrRNA se une al crRNA brindándole estabilidad al dúplex crRNA-tracrRNA, y permitiendo a su vez, la

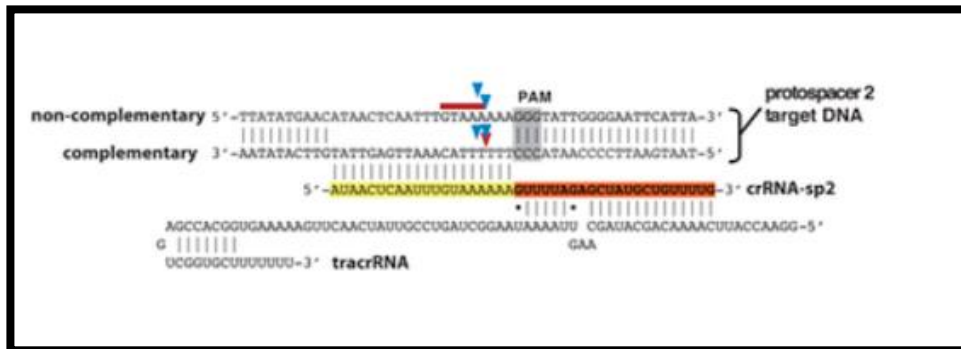
unión con la Cas9 para guiarla a su sitio blanco, donde debe llevar a cabo su función endonucleasa. El tracrRNA además participa en el procesamiento del pre-crRNA (Jinek et al., 2012). Como ejemplo de sistemas CRISPR/Cas de tipo II, se encuentra el sistema de *Streptococcus thermophilus* y *Streptococcus pyogenes* (Gasiunas et al., 2012).

Los sistemas CRISPR/Cas de tipo I se encuentran tanto en bacterias como en arqueas. En contraste, el tipo II está exclusivamente presente en bacterias, mientras que los sistemas de tipo III aparecen más comúnmente en arqueas, aunque también se encuentra en bacterias (Makarova et al., 2011; Terns et al., 2011). Parecería ser entonces que los procariotas han evolucionado generando un sistema inmunitario basado en ácidos nucleicos, donde la especificidad es determinada por los crRNA, mientras que la resistencia es provista por la maquinaria enzimática Cas. No todos los genes Cas están involucrados en la resistencia, algunos de ellos están encargados de la inserción de espaciadores y repetidos CRISPR (Barrangou et al., 2007).

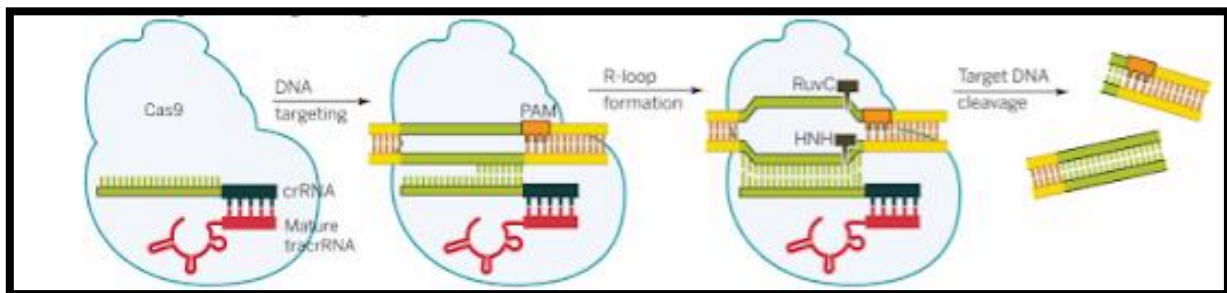
### 1.3.3 Sistema CRISPR/Cas de tipo II

El gen característico de los sistemas CRISPR-Cas de tipo II codifica para la proteína Cas9 multidominio, que combina todas las funciones de complejos efectores provocando la escisión del ADN diana y es también esencial para la maduración del crRNA (Doudna & Charpentier, 2014). Los fragmentos de ADN invasor se incorporan entre dos repetidos adyacentes próximos al final del locus CRISPR. El hecho de que estos espaciadores se integren al final del locus CRISPR otorga información posicional, representando una línea de tiempo sobre la adquisición de éstos. Los motivos periódicos de secuencias repetido-espaciador son transcritos en un largo pre-crRNA que luego es procesado en pequeños crRNA de interferencia de aproximadamente 40 nucleótidos de largo. Estos crRNA incluyen un espaciador, que es una secuencia complementaria al ADN invasor, flanqueado en ambos extremos por los repetidos CRISPR (Jinek et al 2013). La interacción entre los crRNAs y la proteína Cas9 depende de otro ARN transcripto en el locus denominado tracrRNA, el cual a través de su extremo 3' interacciona con la proteína Cas9 y al mismo tiempo interacciona con la región conservada 3' del crRNA dando lugar a una estructura de ADN dual denominada ARN guía (Jinek et al 2013). Este ARN guía es el encargado de dirigir a la proteína Cas9 a generar un corte doble hebra (DSB por double stranded break) en el ADN foráneo que presenta homología con la secuencia espaciadora del crRNA (**Figura 7 y 8**) (Jinek et al 2013). Es de vital importancia que próximo al protoespaciador en el ADN foráneo, se encuentre lo que se conoce como motivo PAM (del inglés, Protospacer Adjacent Motif). El sitio PAM es una pequeña secuencia de ADN ubicada corriente arriba de la región de complementariedad con el crRNA. La presencia de esta secuencia PAM es un requisito para que la Cas9 genere el clivaje específico 3-4 nucleótidos corriente arriba de la misma, luego de la formación del complejo DNA-Cas9 y de la separación de las hebras de ADN (Sternberg et al., 2014). El motivo conservado PAM varía según el organismo al cual pertenezca el sistema

CRISPR/Cas9, por ejemplo, el PAM que reconoce la Cas9 de *Streptococcus pyogenes* es un 5'-NGG-3' (Bortesi et al., 2015; Gasiunas et al., 2012).



**Figura 7.** Representación esquemática del ADN blanco, tracrRNA y crRNA. Se representan las regiones de complementariedad del crRNA con tracrRNA (naranja) y el DNA protoespaciador (amarillo). También se representa la secuencia PAM (gris). Los sitios de escisión están representados por flechas azules. Fuente: Jinek et al., 2012.



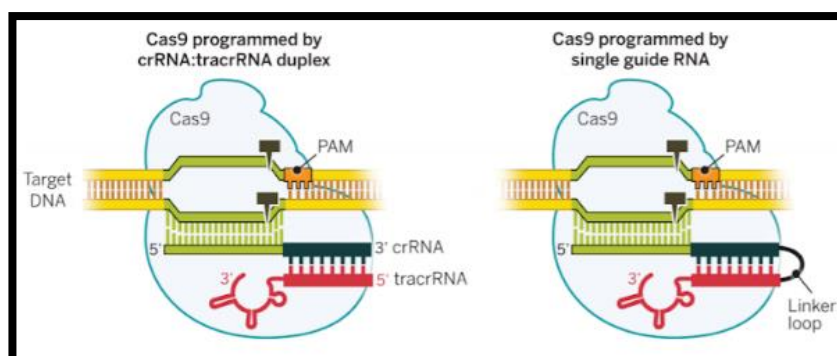
**Figura 8.** Biología del sistema CRISPR-Cas tipo II. Se ilustra el clivaje mediado por la Cas9 dirigido por el dúplex tracrRNA: crRNA. Fuente: (Doudna & Charpentier, 2014).

La enzima Cas9 de *Streptococcus pyogenes* (*SpCas9*), uno de los primeros grupos de bacterias en que se localizó el sistema, tiene aproximadamente 160 kDa y posee cuatro dominios principales. El dominio RuvC es el encargado de escindir la hebra no diana, mientras que el HNH escinde la hebra diana. Por otro lado, se encuentra el dominio responsable de la unión al ARN guía, específicamente al componente tracrRNA, que conduce a la unión específica de Cas9 en el ADN. Finalmente, el dominio PI (del inglés, PAM interaction) se une a la secuencia PAM descendente y es el componente más crítico para la unión del ADN- Cas9 (Nishimasu et al., 2014). Se logró caracterizar y aislar ortólogos de la *SpCas9* encontrándose proteínas Cas9 de tamaños variables. Esta diferencia recae en el lóbulo responsable del reconocimiento del ARN guía. Sin embargo, la arquitectura de sus dominios es similar y comparten regiones conservadas en los dominios nucleasa. De esta forma, se han encontrado Cas9 provenientes de otros organismos como *Staphylococcus aureus* (*SaCas9*) (Nishimasu et al., 2015; Ran et al., 2015). *SaCas9* cuenta con 1053 residuos de aminoácidos, mientras que la *SpCas9* posee 1368, esto se traduce en que la *SaCas9* es aproximadamente 40kDa más pequeña que la *SpCas9* (Nishimasu et al., 2015).

El hecho de que el sistema CRISPR/Cas9 de tipo II requiera únicamente de dos componentes esenciales, una proteína nucleasa Cas9 y un dúplex de ARN conformado por el crRNA y el tracrRNA lo hace muy propicio a ser explotado como herramienta de edición genómica (Doudna & Charpentier, 2014).

### 1.3.4 Sistema CRISPR/Cas9 como herramienta de edición genómica

En el año 2012 dos equipos independientes perfeccionaron el sistema CRISPR/Cas de tipo II de *Streptococcus pyogenes* transformándolo en una tecnología eficaz y aplicable en la ingeniería genética. Uno de los principales aportes para la simplificación del sistema fue el desarrollo de un único ARN guía de la proteína Cas9 denominado “single guide RNA” (sgRNA) (Figura 9). Esta molécula de ARN quimérica es capaz de conservar las funciones cruciales del dúplex crRNA:tracrRNA. En su extremo 5' contiene la secuencia de nucleótidos complementaria al sitio blanco al que se quiere dirigir el sistema, y en su extremo 3' se forman las estructuras necesarias para la asociación con la proteína Cas9. Debido a que la especificidad de corte de la nucleasa Cas9 está determinada por complementariedad de bases del extremo 5' de sgRNA, la edición se puede dirigir a prácticamente cualquier locus dentro del genoma (Jinek et al., 2012; Doudna & Charpentier, 2014). De este modo, el sistema CRISPR/Cas9 se ha adoptado como técnica de edición genómica y se ha demostrado la generación eficiente de modificaciones sitio específicas en células humanas y en una amplia variedad de modelos eucariotas (Cong et al., 2013; Jinek et al., 2013; Church et al., 2013). De esta forma, el diseño correcto del sgRNA y la secuencia PAM son dos factores importantes y esenciales para que ocurra el corte doble hebra en el sitio deseado y de forma eficiente (Gasiunas et al., 2012).

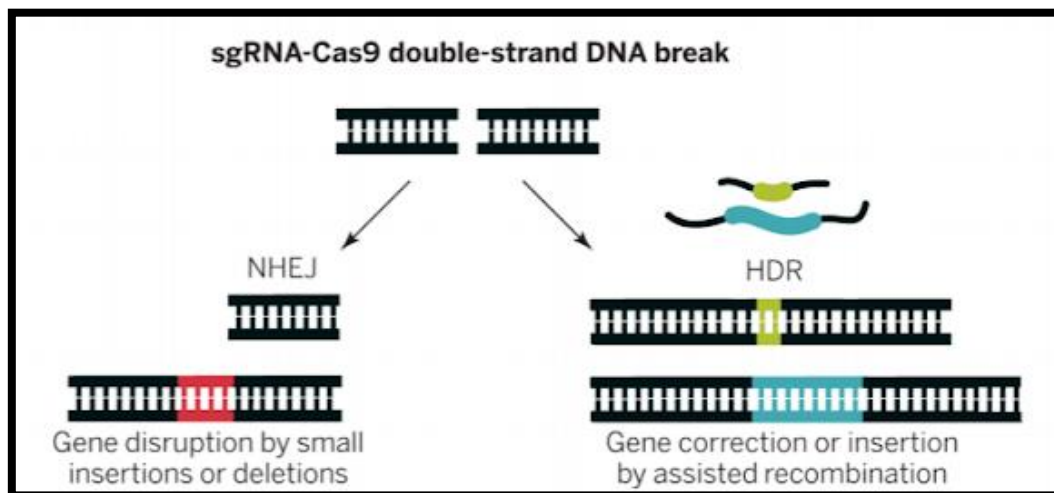


**Figura 9.** Ilustración que muestra el reemplazo al dúplex crRNA-tracrRNA por la molécula de ARN quimérica sgRNA que combina las funciones de estructura dual. Fuente: (Doudna & Charpentier, 2014).

La idea subyacente de utilizar este sistema como herramienta de edición genómica es la facilidad con la que se puede programar al sgRNA para que dirija a la nucleasa Cas9

a zonas específicas del ADN y generar una ruptura doble hebra. Dicha escisión desencadena los mecanismos de reparación del ADN (**Figura 10**) (Ceccaldi et al., 2016). En células eucariotas suele actuar el mecanismo conocido como unión de extremos no homólogos (NHEJ). Su objetivo es restablecer el enlace fosfodiéster, pero a menudo inserta o elimina algunos nucleótidos próximos al sitio de corte. Esta imprecisión del sistema resulta útil para generar knockouts génicos. Por ejemplo, es posible que se altere el marco abierto de lectura (Open Reading Frame, ORF) de la proteína codificada, lo que producirá un cambio en su secuencia aminoacídica, o la generación de un codón stop prematuro, que dará lugar a una proteína trunca y por ende a la pérdida de función del gen. Es recomendable que el sgRNA se diseñe lo más cerca posible del extremo N-terminal de la proteína que se verá afectada, asegurando la pérdida de función del gen blanco. A su vez, también es posible alterar la expresión de un gen disruptiendo secuencias de regiones reguladoras (Bortesi & Fischer, 2015; Cong et al., 2013). Para comprobar si ocurrió edición a través de NHEJ, se puede secuenciar el ADN o realizar un experimento denominado “Surveyor nuclease”, entre otros (Qiu et al., 2004).

Alternativamente, se puede forzar el sistema hacia la reparación por recombinación homóloga (HDR) presentándole a la célula un ADN donante con regiones de homología flanqueantes al sitio de corte. De este modo, se puede insertar una secuencia de ADN de interés en el sitio del corte, estrategia conocida como *knock in*. Por lo general, si el fragmento que se desea introducir es menor a 50 pb el molde es diseñado en forma de oligonucleótido de simple hebra junto con los brazos de homología flanqueantes. Para que pueda darse la recombinación homóloga con el ADN blanco estos últimos deben tener una longitud de entre 50 y 80 pb cada uno. Por el contrario, si la secuencia a introducir posee un tamaño mayor a 100 pb, se suelen utilizar plásmidos que contienen una región flanqueante de 800pb de homología (Bortesi & Fischer, 2015; Cong et al., 2013). Es muy importante que en el ADN donante no exista la secuencia PAM, ya que es probable que la nucleasa lo reconozca como blanco y genere un corte en él. Esta opción brinda la posibilidad de realizar modificaciones en la secuencia codificante para una proteína que permiten la incorporación de un péptido tipo etiqueta (en inglés, tagging) o modificar aminoácidos específicos. También permite generar knockouts mediante el reemplazo génico, incorporando por ejemplo un gen de resistencia a antibiótico que permita la posterior selección de las células editadas o una secuencia con codones stop en los tres marcos de lectura posibles (Salsman & Dellaire, 2017).



**Figura 10.** Posibles mecanismos de reparación luego de un DSB en una célula eucariota. La reparación por NHEJ usualmente genera inserciones o deleciones de pares de bases al azar, ocasionando un knockout génico. La HDR junto con un ADN donante puede ser utilizado para modificar un gen introduciendo una sustitución de nucleótidos precisa o para generar una inserción génica. Fuente: (Doudna & Charpentier, 2014).

### 1.3.5 Generación e introducción de los componentes a las células

Al momento de diseñar el experimento de edición genómica, existen diversas estrategias respecto a la introducción de la maquinaria CRISPR/Cas9 en la célula diana. En lo que refiere a la proteína Cas9, se puede transfectar con plásmidos comerciales que contengan un cassette de expresión de la nucleasa para la expresión ectópica en la célula blanco (ya sea como un sistema inducible o constitutivo) (Ran et al., 2013). Además, también se puede transfectar como ARN mensajero para ser traducido *in vivo*, o directamente como proteína producida de forma recombinante (Burle-Caldas et al., 2018). En el último caso utilizar SaCas9 confiere la ventaja de incrementar la eficiencia de la transfección debido a su bajo peso molecular en comparación con otras Cas9 (Burle-Caldas et al., 2018). Por otro lado, para el sgRNA se puede también transfectar como un plásmido para la transcripción *in vivo* del mismo. Es posible introducir en las células dos plásmidos, uno que exprese la Cas9 deseada y otro el sgRNA, sin embargo, se puede optar por utilizar un solo vector de expresión que contenga simultáneamente la nucleasa y el ARN (Ran et al., 2013). A su vez, el sgRNA puede ser transfectado luego de ser transcrito *in vitro*. Finalmente, considerando un ensayo de edición por HDR en células eucariotas, el ADN donante se puede transfectar como una molécula de ADN lineal doble hebra, o se puede clonar en un plásmido y transfectar al mismo, la opción más conveniente dependerá del tamaño del donador. Las diferentes combinaciones para cada componente del sistema CRISPR/Cas9 son posibles, sin embargo, entre las opciones más comunes se encuentran las líneas celulares que mediante un plásmido expresan de forma inducible o constitutiva la Cas9, y que son co-transfectadas con el sgRNA y el ADN



donante producidos *in vitro*. Otra alternativa frecuente consiste en la co-transfección de los tres componentes en conjunto (Kim et al., 2014). Para mejorar la eficiencia del corte como parte de esta estrategia, previo a la transfección se incubaba a la Cas9 junto al sgRNA facilitando que se forme el complejo ribonucleoproteico (RNP) (Burle-Caldas et al., 2018). Existen variados mecanismos de transfección para la incorporación del sistema CRISPR/Cas9 en las células deseadas. Por un lado, la transfección química se basa en la formación de complejos que las células sean capaces de adquirir e incorporar, ya sea directamente mediante la ruta endocítica o a través de la membrana plasmática. Un método muy frecuente se basa en el uso de moléculas catiónicas que forman complejos con la molécula a transfectar. Estos se adhieren a la membrana celular mediante interacciones electrostáticas provocando el posterior ingreso de la molécula en cuestión al interior celular. Generalmente, los métodos químicos resultan muy útiles para transfectar células en cultivo o embriones (Chandrasekaran et al., 2018). Por otro lado, los métodos físicos de transfección se basan en la introducción mecánica. Un ejemplo de esta categoría es la electroporación, proceso en el cual el ácido nucleico o proteína pasa a través de poros temporarios en la membrana celular consecuencia de someter las células a pulsos eléctricos. Una variante es la nucleofección, se trata de una electroporación optimizada que facilita el ingreso de moléculas al núcleo celular (Chandrasekaran et al., 2018). Adicionalmente, es posible la introducción de moléculas en la célula mediante una microaguja de 10  $\mu\text{m}$  de diámetro en un proceso conocido como microinyección. Finalmente, la transfección biológica comprende procesos como los que llevan a cabo los virus. En este último caso se utilizan lentivirus o adenovirus ya que son capaces de infectar células de mamíferos con alta eficiencia de transfección. Además, el tropismo de los distintos virus permite dirigir la edición a grupos celulares o tejidos específicos (Abrahimi et al., 2015; Chandrasekaran et al., 2018).

### **1.3.6 Ventajas del sistema CRISPR/Cas9 como herramienta de edición genómica**

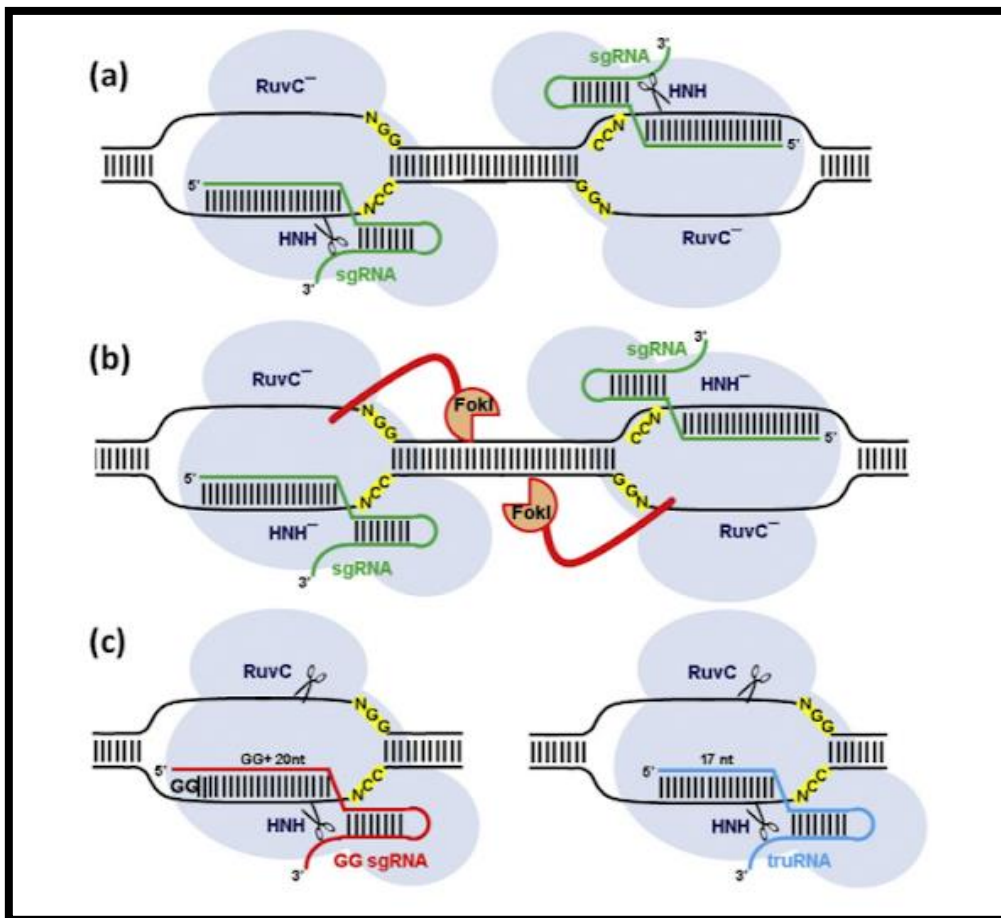
El principal mérito que se destaca del sistema CRISPR-Cas9 en relación con otras tecnologías como ZFN y TALEN radica en el diseño de los componentes de identificación del ADN. El proceso de desarrollo de éstas últimas es costoso y laborioso ya que necesita técnicas de ingeniería proteica y del ensamblaje de sus dominios. Por el contrario, que la especificidad de edición esté determinada exclusivamente por el diseño del sgRNA hace que CRISPR/Cas9 sea un sistema accesible, simple y versátil (Bortesi & Fischer, 2015; Jinek et al., 2012). Otra notable característica del sistema CRISPR-Cas9 es la posibilidad de “programar” una misma proteína Cas9 con múltiples sgRNA dirigidos a distintas regiones del genoma (Li et al., 2013, Mao et al., 2013). Este hecho demuestra la capacidad de esta tecnología de generar ediciones en simultáneo en distintos sitios teniendo grandes implicancias. Por ejemplo, permite mutar múltiples miembros de una familia de genes, o genes funcionalmente relacionados

que interactúan para dar una característica compleja; discernir modelos subyacentes a las enfermedades multigénicas, entre otros (Bortesi & Fischer, 2015). Una característica general importante de los sistemas CRISPR /Cas que los diferencia de los ZFN y TALEN es su capacidad para dirigirse diferencialmente al ADN o al ARN. El sistema CRISPR-Cas de tipo III-B (*Pyrococcus furiosus*) es un sistema de silenciamiento de ARN único que implica la degradación dependiente de la homología del ARN complementario en presencia de crRNA modificado tanto *in vitro* como *in vivo* (Hale et al., 2012). Dicho control postranscripcional flexible de la expresión génica compensa los posibles errores de una tecnología dirigida exclusivamente al ADN, cuando la unión del complejo al ADN diana se ve obstaculizada por la estructura de la cromatina o por la presencia de otras proteínas unidas, o cuando se desea la eliminación de sólo una de varias variantes de empalme de un transcrito. Además, el sistema CRISPR/Cas9 puede escindir el ADN metilado en células humanas (Hsu et al., 2013), lo que permite modificaciones genómicas que están más allá del alcance de las otras nucleasas (Ding et al., 2013). El sistema CRISPR/Cas9 también presenta la ventaja frente a las otras nucleasas de determinar con facilidad mediante análisis de secuencia los sitios donde es probable que se generen mutaciones fuera del sitio blanco (“off-targets”). Esto es posible debido a que las interacciones que ocurren entre el sgRNA y el ADN blanco son del tipo Watson y Crick (Cho et al., 2014). En este sentido se han desarrollado diversos programas informáticos que asisten en el desarrollo de un experimento de edición genómica por CRISPR/Cas9 mediante la búsqueda de sgRNAs adecuados capaces de dirigir un DBS específico evitando ediciones en sitios no deseados (Concordet & Haeussler, 2018). Alterar la longitud del sgRNA también puede minimizar las modificaciones off-target: extendiendo el sgRNA agregando dos residuos de guanidina en el extremo 5' o acortándolo, generando un sgRNA truncado de solo 17 nucleótidos (Bortesi & Fischer, 2015) (**Figura 11C**).

El hecho de que los organismos eucariotas carecen del sistema CRISPR/Cas9, resulta ser una ventaja a la hora de utilizarlo como herramienta de edición genómica frente al ARN de interferencia (ARNi). CRISPR/Cas9 no necesita factores endógenos del hospedero como las proteínas DICER, o el complejo Argonauta, esto hace que sea un sistema más específico (Bortesi & Fischer, 2015). Además, el sistema CRISPR/Cas9 permite llevar a cabo múltiples aplicaciones, mientras que el ARNi, está restringido al silenciamiento de la expresión de los genes (Doudna & Charpentier, 2014).

### 1.3.7 Dificultades y soluciones potenciales de CRISPR/Cas9

Si bien el sistema CRISPR-Cas9 ha supuesto un avance en la obtención de herramientas de edición génica útiles y eficaces, presenta ciertas desventajas. Las posibles modificaciones no deseadas que se producen en el genoma del organismo diana constituyen el principal obstáculo que tienen que superar los investigadores. Aunque dicha tecnología permite modificar su sitio blanco, induciendo de manera eficiente DSB específicos, también puede crear escisiones no deseadas en sitios fuera del objetivo similares en secuencia a este, limitando así su utilidad en muchas aplicaciones como la terapia génica o celular (Koo et al., 2015; Cong et al., 2013). Estas modificaciones inespecíficas ocurren debido a que en los organismos procariontes que poseen el sistema CRISPR/Cas9 naturalmente, en el intento por adquirir inmunidad adaptativa contra ácidos nucleicos foráneos, puede ser útil el reconocimiento de ADN hipervariable (Koo et al., 2015). Asimismo, desde que los investigadores tuvieron constancia de estos inconvenientes, han centrado sus estudios en el desarrollo de estrategias para corregirlas. En el seno de estas investigaciones es donde nace las nickasas (nCas9), que son dos proteínas Cas9 las cuales han sufrido un proceso de modificación en uno de sus dominios catalíticos (RuvC o NHN) para que solo escinda una hebra del ADN blanco (Cho et al., 2014; Shen et al., 2014). De esta forma, es posible combinar dos proteínas nCas9 de las cuales una tendrá mutado el dominio RuvC para que solo realice el corte el dominio NHN en la cadena complementaria al sgRNA y viceversa. La característica esencial es la programación de dos sistemas sgRNA:nCas9 para que los cortes en cada una de las dos hebras del ADN se produzcan a una distancia cercana simulando la acción de la Cas9 cuando forma un DSB (**Figura 11A**) (Hryhorowicz et al., 2017). De esta manera resulta poco probable que el sistema pueda encontrar la misma secuencia nucleotídica a una distancia determinada en el genoma. Además, la generación de un nick fuera del objetivo suele repararse de forma inmediata sin mutaciones indeseables (Li et al. 2014). La contracara de esta modificación es que baja eficiencia de corte en comparación con la Cas9 convencional. A su vez, otra desventaja que tienen las nickasas es la necesidad de introducir en el organismo diana dos sistemas de expresión que contengan los dos complejos de edición génica (Hryhorowicz et al., 2017). Alternativamente, es posible la fusión entre la Cas9 catalíticamente inactiva (dCas9) y la nucleasa FokI, dando lugar a la fCas9 (Bortesi & Fischer, 2015). Debido a que FokI posee su actividad catalítica cuando se encuentra en forma de dímero, es necesario utilizar dos dCas9-FokI, con dos sgRNA, para que ambos dominios FokI se encuentren y cliven el ADN. La fusión de FokI a la dCas9 obliga que se produzca la dimerización que generará el corte entre dos sgRNAs separados por entre 15 y 25 pb, con sus PAMs orientados hacia afuera (**Figura 11B**) (Bortesi & Fischer, 2015). Esta estrategia aumenta mucho la especificidad y la eficiencia de corte, ya que reduce el número potencial de secuencias off-target (Bortesi & Fischer, 2015). Por consiguiente, diversas modificaciones sobre la estructura proteica de la Cas9 han permitido obtener variaciones de la Cas9 nativa como son las eSpCas9 (enhanced Specificity Cas9) y Cas9-HF (Cas9-High Fidelity) que mejoran sustantivamente la especificidad de la proteína (Lee et al., 2017).



**Figura 11.** Estrategias que reducen efectos off-target. **A)** El uso simultáneo de dos nickasas Cas9 que se unen a secuencias opuestas en ambas hebras de ADN simula un DSB. **B)** Fusión de las proteínas FokI y dCas9, para dar lugar a la fCas9. El clivaje del ADN dado por FokI depende de dimerización. La fusión de FokI a la Cas9 obliga que se produzca la dimerización que generará el clivaje entre dos sgRNAs. **C)** Modificar la longitud del sgRNA, mediante una extensión del sgRNA por la adición de dos residuos de guanina (G) en el extremo 5'; o por un acortamiento que da lugar a un ARN truncado de 17 nt. Fuente: Bortesi & Fischer, 2015.

En esta tesina se busca construir el sistema de etiquetado endógeno de AGO2 por *knock-in* utilizando la edición genómica por CRISPR/Cas9 dirigido por los componentes sintetizados *in vitro*. El trabajo busca diseñar y producir los componentes del sistema CRISPR/Cas9 y analizar la eficiencia del corte doble hebra sitio específico sobre el gen AGO2 *in vitro*. La demostración de que este gen puede ser clivado por este sistema proporciona un gran avance para poder, en trabajos posteriores, introducir el mismo en líneas celulares de CaP en las que trabaja nuestro laboratorio. La importancia en la generación de un etiquetado en AGO2 radica en la obtención de un epítotope a fines de que pueda ser inmunoprecipitado con un anticuerpo específico en la estrategia conocida como *pull-down* (en español, halar hacia abajo). En este caso, el etiquetado de AGO2 en una línea celular de CaP permitiría el análisis de ARNs mensajeros unidos a la proteína. Es importante mencionar que hasta el momento no se ha publicado ningún etiquetado endógeno de la proteína AGO2 mediante CRISPR/Cas9. Este método en conjunto con otras aproximaciones permite identificar los genes que están siendo regulados directamente por determinados miARNs. Dado que el sistema de CRISPR/Cas9 requiere de varios pasos de ajuste, es importante disponer de métodos para evaluar la eficacia de sus etapas. El *knock-out* de proteínas fluorescentes es ampliamente usado para determinar la eficiencia de la edición dentro de la célula. En base a esto, en este trabajo también se sintetizaron y evaluaron *in vitro*, los componentes del sistema CRISPR/Cas9 a fines de realizar un *knock-out* del gen de la proteína fluorescente mCherry con el propósito de utilizarlo como control de edición *in vivo*. Para alcanzar los objetivos, se comenzó con la producción y purificación de la SaCas9 (*Staphylococcus aureus* Cas9) en base a un protocolo ya establecido en el laboratorio. Esto se realizó a partir de un plásmido comercial para expresión recombinante en *E.coli*. Luego se procedió a la purificación de la proteína expresada mediante cromatografía de afinidad. Posteriormente se diseñaron los sgRNAs correspondientes para ambos genes. Se evaluó la actividad de la SaCas9 mediante ensayos *in vitro* donde se expone un fragmento de ADN blanco a ser digerido por la nucleasa guiada por los sgRNAs diseñados. Finalmente, se diseñaron los ADN donante para la edición por HDR del gen AGO2 y mCherry.

# OBJETIVOS

## 2.1 Objetivo general

Construir un sistema *in vitro* para el etiquetado endógeno del gen AGO2 humano y la inactivación del gen control mCherry mediante edición genómica por CRISPR/SaCas9 recombinante.

## 2.2 Objetivos específicos

1. Expresar y purificar la proteína recombinante SaCas9.
2. Diseñar y sintetizar ARNs guías para la edición de AGO2 y mCherry.
3. Obtener un corte *in vitro* eficiente y específico mediado por la SaCas9 en los genes AGO2 y mCherry.
4. Diseñar ADNs donantes para la edición de AGO2 y mCherry.



Inicialmente se realizó un pre-cultivo bacteriano con 25mL del medio 2xYT, al cual se le agregaron 25µL de Ampicilina (100mg/mL), y un inóculo de 25µL del vial de cepas Rosetta 2(DE3) transformantes congelado. Se dejó a 37°C, a 200 rpm overnight (ON). Al día siguiente, en dos matraces Erlenmeyer de 2L se preparó el cultivo, con 500mL de medio 2xYT, 500µL de ampicilina y 10mL del pre-cultivo. Los matraces se dejaron en un agitador con una velocidad de 200 rpm y una temperatura de 37°C, hasta que el cultivo alcanzó una densidad óptica (DO) entre 0,4 y 0,6 unidades de absorbancia (UA). Se procedió añadiendo 500µL de isopropil-β-D-1- tiogalactopiranosido (IPTG) 1M al medio de cultivo y se mantuvo en el incubador con agitación durante 24 horas a 18°C. El IPTG se utiliza como análogo no hidrolizable de la alolactosa con la finalidad de inducir la expresión de genes que se encuentren bajo la regulación de un promotor lac (en este caso T7-lac). Cuando es agregado al cultivo bacteriano, actúa inactivando al represor lac induciendo entonces la transcripción a partir del promotor. En este sentido, el IPTG induce la expresión de la proteína SaCas9 en aquellas cepas que portan el plásmido mencionado anteriormente. Luego se traspasó el cultivo a tubos Falcon de 50mL, y se los centrifugó por 20 minutos a 10000 rpm y 4°C. Se masó el peso húmedo de los pellets resultantes para calcular el rendimiento de los cultivos y se los guardó a -80°C. Posteriormente se lisaron las células químicamente; para ello se resuspendió el pellet en buffer fosfato en un volumen 1/20 respecto al cultivo y se prosiguió según el protocolo de lisis química explicado a continuación. Luego se centrifugó el lisado a 10000g por 3 minutos a 4°C, se guardó el pellet (fracción insoluble) a 20°C y el sobrenadante a 4° (fracción soluble) hasta la etapa de purificación. Finalmente se realizó una electroforesis en gel de poliacrilamida desnaturalizante por dodecil sulfato de sodio (SDS-PAGE), para verificar la presencia de SaCas9.

Protocolo empleado para lisis química de las bacterias transformantes luego de la inducción:

- 1) Se resuspendió el pellet de 100mL de cultivo en 5 mL de buffer fosfato obteniendo una concentración final de 1X.
- 2) Se agregó 50 µL de PMSF 100 mM para una concentración final de 1mM.
- 3) Se agregó 100 µL de Lisozima 50mg/mL para una concentración final de 1mg/mL.
- 4) Se agregaron perlas de vidrio y se agitó la muestra durante 30 minutos a temperatura ambiente.
- 5) Se agregó 4 µL de DNAsa 10mg/mL y RNAsa 10mg/mL quedando en una concentración final de 8 µg/mL.
- 6) Se agregó 5 µL de MgCl<sub>2</sub> 1M para una concentración final de 1mM.
- 7) Se agitó la muestra durante 30 minutos a temperatura ambiente.
- 8) Se quitaron las perlas de vidrio.



### 3.1.2 Purificación

La purificación de la SaCas9 se llevó a cabo a través de cromatografía de afinidad por columna de níquel (His-Trap HP, GE™) debido a que nuestra proteína recombinante presenta una etiqueta de seis residuos de histidina en su extremo N-terminal. Las histidinas tienen afinidad por el níquel y por lo tanto quedan retenidas en la columna junto con la proteína taggeada. Se utilizó un equipo de FPLC (del inglés, Fast Protein Liquid Chromatography) de Cytiva y modelo ÄKTA, configurando la colección de fracciones cada 4mL de fase móvil. Previo a la etapa de purificación, es necesario preparar la columna de níquel His-Trap mediante el procedimiento conocido como “stripping and recharging” si la misma ha sido utilizada previamente. La columna no tiene que ser despojada y recargada entre cada purificación si se va a purificar la misma proteína. Este método permite que el níquel se una a la matriz estacionaria de la columna de Sepharose™, y por consiguiente al pasar las proteínas (fase móvil), la SaCas9 quede retenida en la matriz. Para su preparación, en primer lugar, se pasaron diez volúmenes de agua por la columna utilizando una bomba peristáltica y una jeringa Microliter Hamilton™, debido a que para su almacenamiento se encuentra cargada con etanol. Posteriormente, se pasaron diez volúmenes de buffer de unión (binding buffer) con 50mM de EDTA. En segundo lugar, se realizó la limpieza de la columna pasando diez volúmenes de hidróxido de sodio (NaOH) 1M, diez volúmenes del buffer de unión y diez volúmenes de agua, en orden cronológico. El tercer y último paso, es la recarga de la columna, para ello se pasaron dos volúmenes de sulfato de níquel (NiSO<sub>4</sub>) y diez volúmenes de agua. En este punto la columna está lista para ser utilizada, y se almacenó en etanol 20% a 4°C. Previo al uso se retiró el etanol y se equilibró la columna en buffer de unión. Todas las soluciones incorporadas a la columna fueron filtradas en 0,45 µm y se trabajó a un máximo de 1 mL/min. Una vez que la columna estaba lista, se cargó con el producto de lisis correspondiente a la fracción soluble, descrito anteriormente. Luego se realizaron lavados con buffer de unión conteniendo 30mM de imidazol. El imidazol actúa como competidor de los residuos de histidina por la afinidad a la columna de níquel. Al comienzo una baja concentración de imidazol de 30mM permite la interacción específica de proteínas con la matriz (100% buffer de unión, **ver anexo**). En este lavado, se colecta aquella fracción de proteínas no unidas a la columna (FT, de flow-through). Con el objetivo de desplazar a las proteínas que estén unidas a la columna se procedió a incrementar la concentración de imidazol en tres lavados, incorporando cantidades crecientes de buffer de elución (300mM imidazol, **ver anexo**). El primer lavado se realizó en 53 mM de imidazol (5/95% buffer elución/unión), el segundo en 82 mM (11/89% buffer elución/unión) y el tercero con 124 mM de imidazol (20/80% buffer elución/unión). Para el lavado final o elución de la proteína SaCas9, se utilizó 500 mM de imidazol correspondiéndose con el 100% del buffer de elución. Posteriormente se utilizó una columna de gel filtración PD10 desalting column, Sephadex™ (N° de catálogo: 17-0851-01, Amersham), siguiendo las instrucciones del fabricante. El propósito de esta etapa es el de separar la SaCas9 de otros contaminantes como componentes del

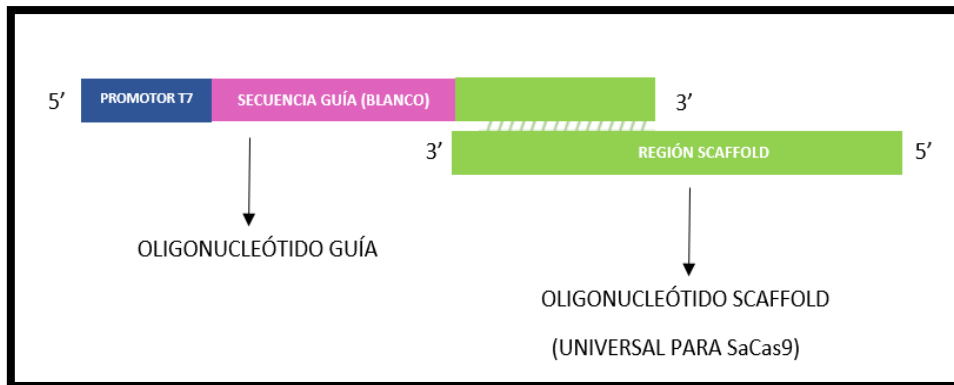
buffer de la cromatografía de afinidad, entre otros. Se utilizaron centricones MilliporeSigma™ Amicon™ Ultra Centrifugal Filtre Units (N° de catálogo: UFC901008, Thermo Fisher Scientific) para llevar a la proteína a una concentración que permitiera su utilización en las posteriores aplicaciones. El producto obtenido se conservó en 40% de glicerol y se almacenó a -20°C. Por último, se hizo un SDS-PAGE con los distintos lavados, incluyendo la elución final, con la finalidad de evaluar la pureza de la proteína purificada.

## 3.2 sgRNA

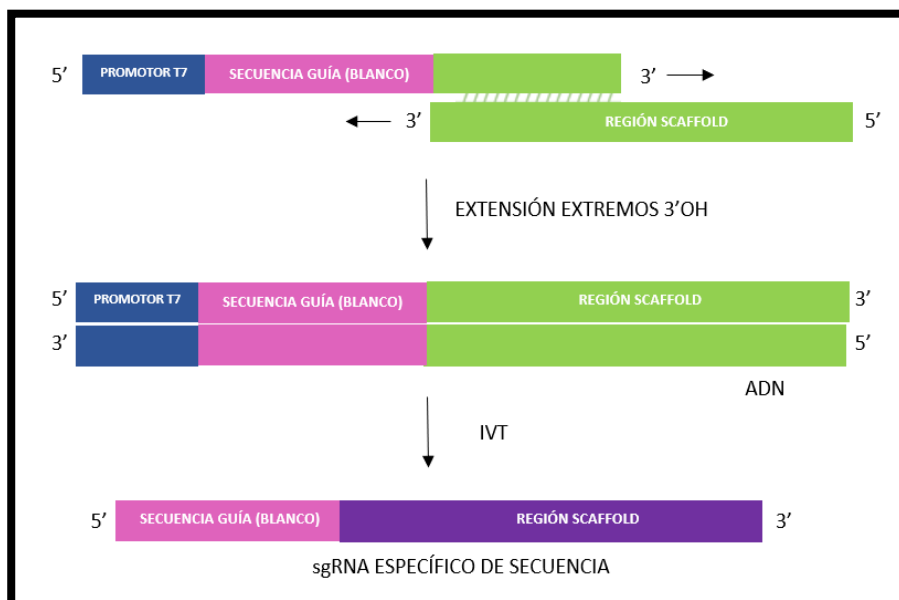
### 3.2.1 Diseño

Se utilizó el software informático CRISPOR (Concordet & Haeussler, 2018) para el diseño de sgRNAs capaces de dirigir un DSB en regiones de interés en los genes mCherry y AGO2. El sitio CRISPOR es una herramienta web dedicada al diseño de experimentos de edición genómica a través del sistema CRISPR/Cas9. El programa contribuye en la búsqueda de posibles sgRNAs según la secuencia de ADN blanco a ser editada y los evalúa en base a distintos parámetros identificando los mejores candidatos para el ensayo de edición. Para esto, el software asiste en el diseño de dos oligonucleótidos parcialmente complementarios que mediante un proceso de extensión de sus extremos 3'OH darán lugar al molde de ADN que será utilizado para la síntesis del sgRNA en un proceso de transcripción *in vitro* (IVT) (Figura 13 y 14). No obstante, cabe destacar que todos los resultados de CRISPOR se evaluaron manualmente para verificar la calidad del diseño automático, puesto que aún existen varios criterios en discusión en la literatura. Por un lado, se diseñó un oligonucleótido denominado “andamio” (Scaffold, en inglés) que forma las estructuras necesarias para la interacción con la SaCas9 cumpliendo el rol del tracrRNA siendo específico de cada Cas9, por lo tanto, puede ser utilizado con esta proteína indistintamente del gen blanco a editar (Figura 13). Por otro lado, se diseñó el otro oligonucleótido que incluye la secuencia complementaria a la región a editar y mediante complementariedad de bases dirige a la SaCas9 hacia el sitio específico de corte, simulando así el crRNA. Este oligonucleótido contiene en su extremo 5' la secuencia del promotor T7, que es reconocida exclusivamente por la ARN polimerasa T7 y en su extremo 3' una región de complementariedad con el extremo 3' del primer oligonucleótido andamio (Figura 13). Para el diseño de este oligonucleótido guía, es necesario introducir las secuencias de los genes a editar y un genoma de referencia. Para la secuencia del gen AGO2 humano se utilizó la base de datos NCBI y para el caso del gen control mCherry, se utilizó la secuencia contenida en el plásmido pEZX-MR02 (<https://www.genecopoeia.com/wp-content/uploads/oldpdfs/product/mirna/pEZX-MR02.pdf>, N. de catálogo: CmiR0001-MR02, GeneCopoeia). El sitio CRISPOR identifica sitios PAM de la SaCas9 (5'-NNGRRT-3') en las secuencias indicadas y analiza la secuencia de homología de 20 nucleótidos hacia el 5' de las posibles regiones PAM, buscando características óptimas de diseño del oligonucleótido con la secuencia guía

(contenido GC cercano a 50% y ausencia de estructuras secundarias). A su vez, es importante diseñar este oligonucleótido guía de modo de reducir de la mayor forma posible la presencia de sitios adicionales de corte (off targets) en el genoma de referencia que se utilice. Considerando los parámetros mencionados, se seleccionaron los oligonucleótidos adecuados para ejercer un corte de interés en mCherry y AGO2; y se enviaron a sintetizar a la empresa MACROGEN (<https://www.dna.macrogen.com/>).



**Figura 13.** Esquema del diseño del sgRNA a partir de oligonucleótidos solapantes. En el oligonucleótido guía, se indica en azul el promotor T7 y en rosado la región complementaria al sitio de edición del gen de interés. El oligonucleótido universal se muestra en verde, indicando la región que codifica para el andamio del sgRNA. Las regiones en verde unidas mediante líneas en ambos oligonucleótidos indican el sitio de complementariedad.



**Figura 14.** Esquema del procedimiento llevado a cabo para la generación del sgRNA. Mediante un ensayo de extensión a partir de los extremos 3'OH de los oligonucleótidos solapantes se genera el ADN molde para la IVT. El resultado final es una molécula sgRNA capaz de dirigirse hacia un sitio específico del ADN.

## 3.2.2 Síntesis

### 3.2.2.1 ADN molde para IVT

El ADN que actúa como molde para los sgRNAs en un posterior proceso de IVT, se genera mediante un ensayo de extensión a partir de los extremos 3'OH de los oligonucleótidos solapantes descritos en la sección anterior. Se trata de una reacción de síntesis de ADN, donde el extremo 3'OH de cada oligonucleótido es extendido mediante una ADN polimerasa, sirviéndose ellos mismos de molde (Bassett et al., 2013) (**Figura 13**). Se utilizó una ADN polimerasa de alta fidelidad, “Phusion™ High-Fidelity DNA Polymerase” (N° de catálogo: F-530XL, ThermoFisher Scientific), con la finalidad de evitar posibles errores en el direccionamiento del DSB. Además, las ADN polimerasas convencionales, a diferencia de la utilizada, poseen actividad adenil transferasa terminal, y por lo tanto incorporan dATP al finalizar su extensión. El ADN molde para la producción de los sgRNAs debe ser muy preciso y la adición de la adenina final podría afectar la eficiencia de la producción por IVT. De esta forma, se sintetizaron los ADN molde de IVT para los sgRNAs de mCherry y AGO2 según se especifica en la **Tabla 1** y el programa de termociclado mostrado en la **Tabla 2**. Posteriormente, se purificaron los productos mediante el método fenol-cloroformo y posterior precipitación en etanol. La concentración de ADN obtenida (Absorbancia a 260nm) y la pureza (Absorbancia a 260/280) se midieron en un espectrofotómetro de microvolúmenes. Por último, se analizaron los productos por electroforesis en gel de agarosa 1%.

REACTIVOS	VOLUMEN	CONCENTRACIÓN FINAL
5X Phusion HF Buffer	20 µL	1X
dNTP Mix (10mM each)	2 µL	200 µM each
Oligonucleótido guía (10 µM)	5 µL	0,5 µM
Oligonucleótido universal (10 µM)	5 µL	0,5 µM
Phusion DNA Polymerase (2 U/ µL)	1 µL	0,02 U/µL
H2O milliQ csp 100 µl	67 µL	-

**Tabla 1.** Reactivos y condiciones utilizados para generar el ADN molde de las IVT.

ETAPAS	TEMPERATURA	TIEMPO	NRO. DE CICLOS
Desnaturalización inicial	98°C	30 segundos	1
Desnaturalización	98°C	10 segundos	35
Annealing	60°C	30 segundos	35
Extensión	72°C	15 segundos	35
Extensión final	72°C	10 minutos	1

**Tabla 2.** Programa de termociclado utilizado para generar el ADN molde para las IVT.

### 3.2.2.2 Transcripción *in vitro*

La síntesis de los sgRNAs se llevó a cabo mediante ensayos de IVT, utilizando el ADN molde generado y purificado en el apartado anterior. Para esto, se utilizó el “HiScribe™ T7 High Yield RNA synthesis Kit” (N° de catálogo: E2040S, New England Biolabs (NEB)). Se preparó la reacción de IVT según la **Tabla 3** y se incubaron las reacciones a 37°C durante 16 horas (Bassett et al., 2013). Luego, los sgRNAs obtenidos se purificaron utilizando el “Monarch® RNA Cleanup Kit” (N° de catálogo: T2040S, New England Biolabs (NEB)) y se cuantificó su concentración y pureza midiendo la absorbancia a una longitud de onda de 260 y 260/280nm respectivamente en un espectrofotómetro de microvolúmenes. Los sgRNAs se almacenaron a -80°C.

REACTIVOS	VOLUMEN	CONCENTRACIÓN FINAL
10X buffer de reacción	2 µL	1X
ATP (100mM)	2 µL	10mM
UTP (100mM)	2 µL	10mM
CTP (100mM)	2 µL	10mM
GTP (100mM)	2 µL	10mM
ADN molde	-	1 µg
T7 ARN polimerasa mix	2 µL	-
H2O RNAsa free csp para 20 µL	-	-

**Tabla 3.** Reactivos y condiciones utilizados en las IVT.

### 3.3 ADN blanco

#### 3.3.1 mCherry

El ADN blanco para el ensayo de actividad de la SaCas9 del gen mCherry, se obtuvo mediante PCR a partir del plásmido pEZX-MR02 que contiene la secuencia de este gen. Los cebadores utilizados (**Tabla 4**) están diseñados para amplificar un producto de 360pb. Para estas reacciones de PCR se utilizó la enzima “DreamTaq DNA Polymerase” (N° de catálogo: EP0705, ThermoFisher Scientific), siguiendo las instrucciones detalladas en la **Tabla 5** y **6**. Posteriormente, el producto fue purificado mediante el método de extracción acuosa con fenol-cloroformo seguido de una precipitación con etanol. La concentración de ADN obtenida (Abs. 260 nm) y la pureza (Abs (260/280nm)) fueron medidas en un espectrofotómetro de microvolúmenes. Los fragmentos amplificados se analizaron mediante una electroforesis en gel de agarosa 1%.

CEBADORES		% GC	TM (°C)
FORWARD	5'-CACGAGTTCGAGATCGAGGG-3'	60%	64,8°C
REVERSE	5'-AGCCCATGGTCTTCTTCTGC-3'	55%	65,3°C

**Tabla 4.** Oligonucleótidos cebadores utilizados para amplificar el gen mCherry. Se muestra el contenido GC de cada uno de ellos (GC%) y la temperatura de melting (TM).

REACTIVO	VOLUMEN	CONCENTRACIÓN FINAL
10X DreamTaq Buffer	5 µL	1X
dNTP Mix (10mM each)	1 µL	0,2 mM each
Forward primer (10 µM)	1 µL	0,2 µM
Reverse primer (10 µM)	1 µL	0,2 µM
ADN plásmido (30 ng)	5 µL	-
DreamTaq DNA Polymerase (5U/ µL)	0,25 µL	0,025U/µL
H2O milliQ csp 50 µL	36,75 µL	-

**Tabla 5.** Reactivos y condiciones utilizados en la PCR del gen mCherry.

ETAPAS	TEMPERATURA	TIEMPO	NRO. DE CICLOS
Desnaturalización inicial	95°C	3 minutos	1
Desnaturalización	95°C	30 segundos	35
Annealing	60°C	30 segundos	35
Extensión	72°C	1 minuto	35
Extensión final	72°C	10 minutos	1

**Tabla 6.** Programa de termociclado utilizado para amplificar el gen mCherry.

### 3.3.2 AGO2

Se llevaron a cabo tres reacciones de PCR diferentes para amplificar la región de interés de AGO2 a partir de ADN genómico humano. Los cebadores utilizados para cada una de estas reacciones se detallan en la **Tabla 7**. Para los tres pares de cebadores PAR1, PAR2 y PAR3, se espera que se amplifique un fragmento de 613, 879 y 1755pb respectivamente. La enzima utilizada es la “DreamTaq DNA Polymerase” (N° de catálogo: EP0705, ThermoFisher Scientific), siguiendo las instrucciones detalladas en **Tabla 8 y 9**. Alternativamente, se realizó otra PCR con el PAR3 de cebadores utilizando la enzima “KAPA HiFi HotStart DNA Polymerase” (Roche). En la **Tabla 8 y 9** se detalla el procedimiento y el programa de termociclado utilizado. Alternativamente, se hizo la síntesis *in vitro* del ADN blanco de AGO2 para el ensayo de actividad mediante el diseño de oligonucleótidos parcialmente complementarios. Por un lado, se diseñaron dos oligonucleótidos (F1 y F2) que al hibridarse de forma independiente con un tercer oligonucleótido universal (R) luego de un ensayo de extensión de extremos 3’OH, darán lugar a dos pequeñas moléculas de ADN (A1 y A2). Ciertamente, estas moléculas contienen la secuencia PAM reconocida por la SaCas9 y, inmediatamente corriente arriba, la región de complementariedad con la secuencia guía del sgRNA diseñado para AGO2. Una vez diseñados los oligonucleótidos, se mandaron sintetizar a MACROGEN para luego formar las moléculas de ADN doble hebra según se muestra en la **Tabla 10 y 11**. Los productos de PCR obtenidos fueron purificados mediante fenol-cloroformo seguido de una precipitación en etanol. Algunos de los productos del solapamiento de los oligonucleótidos se purificaron mediante el mismo método y otros con el “Monarch® PCR & DNA Cleanup Kit” (N° de catálogo: T1030S, New England BioLabs (NEB)). La concentración de ADN obtenida (Abs. 260 nm) y la pureza (Abs (260/280nm)) fueron medidas en un espectrofotómetro de microvolúmenes. Los resultados obtenidos se evaluaron en geles de agarosa al 1 y 2%.

CEBADORES		% GC	TM (°C)	
PAR 1	FORWARD	5'-GCGCTCCGGGCTCGGGCCTAC-3'	82	80
	REVERSE	5'-CCTCCTCCTTCCCTCCCCGC-3'	75	71
PAR 2	FORWARD	5'-GCGCTCCGGGCTCGGGCCTAC-3'	82	80
	REVERSE	5'- GCGGCTTCTCTGGGTGGCAG-3'	70	70
PAR 3	FORWARD	5'-GAGACCGCCTTCTACTCTGC- 3'	60	65
	REVERSE	5'-GTCCGCTGATCTGATCCCTG- 3'	60	65

**Tabla 7.** Cebadores utilizados en las tres reacciones de PCR del gen AGO2. Se muestra el contenido GC (GC%) y la temperatura de melting (TM).

CEBADORES UTILIZADOS	REACTIVO	VOLUMEN	CONCENTRACIÓN FINAL
PAR 1, 2 y 3	10X DreamTaq Buffer	5 µL	1X
	dNTP Mix (10mM each)	1 µL	0,2 mM each
	Forward primer (10 µM)	0,5-1 µL	0,1-0,2 µM
	Reverse primer (10 µM)	0,5-1 µL	0,1 -0,2 µM
	ADN genómico (42 ng)	5 µL	-
	DreamTaq DNA Polymerase (5U/µL)	0,25 µL	0,025U/µL
	H2O csp para 50 µL	-	-
PAR 3	5X KAPA HiFi GC Buffer	5 µL	1X
	dNTP Mix (10mM each)	1 µL	0,3 mM each
	Forward primer (10 µM)	0,75 µL	0,3 µM
	Reverse primer (10 µM)	0,75 µL	0,3 µM
	ADN genómico (13ng)	5 µL	-
	KAPA HiFi HotStart DNA Pol. (1 U/µL)	0,5 µL	0.5 U/ µL
	H2O milliQ csp para 25 µL	12 µL	-

**Tabla 8.** Condiciones utilizadas en las PCRs del gen AGO2 para las distintas enzimas empleadas.

ENZIMA	CEBADORES UTILIZADOS	ETAPAS	TEMPERATURA	TIEMPO	NRO. DE CICLOS
DreamTaq DNA Polymerase	PAR 1, 2 y 3	Desnaturalización inicial	95°C	3 minutos	1
		Desnaturalización	95°C	30 segundos	35
		Annealing	*(58°C-68°C)	30 segundos	35
		Extensión	72°C	+	35
		Extensión final	72°C	10 minutos	1
KAPA HiFi HotStart DNA	PAR 3	Desnaturalización inicial	95°C	5 minutos	1
		Desnaturalización	98°C	20 segundos	35
		Annealing	60°C	30 segundos	35
		Extensión	72°C	1 minuto y 30 segundos	35
		Extensión final	72°C	10 minutos	1

**Tabla 9.** Programa de termociclado utilizado para amplificar el gen AGO2 según las distintas enzimas utilizadas. \* Se probaron distintas temperaturas en un gradiente de 58°C a 68°C para optimizar la PCR. + En lo que respecta al par de cebadores 1 y 2 se aplicó un tiempo de extensión de un minuto. Para el par 3 se utilizó dos minutos.



REACTIVO	VOLUMEN	CONCENTRACIÓN FINAL
10X DreamTaq Buffer	5 $\mu$ L	1X
dNTP Mix (10mM each)	1 $\mu$ L	0,2 mM each
Oligonucleótido F1 o F2 (10 $\mu$ M)	1 $\mu$ L	0,2 $\mu$ M
Oligonucleótido R (10 $\mu$ M)	1 $\mu$ L	0,2 $\mu$ M
DreamTaq DNA Polymerase (5U/ $\mu$ L )	0,25 $\mu$ L	0,025U/ $\mu$ L
H2O milliQ csp para 50 $\mu$ L	41,75 $\mu$ L	-

**Tabla 10.** Condiciones utilizadas en el ensayo de extensión de los oligonucleótidos solapantes para generar la secuencia de interés del gen AGO2.

ETAPAS	TEMPERATURA	TIEMPO	NRO. DE CICLOS
Desnaturalización inicial	95°C	3 minutos	1
Desnaturalización	95°C	30 segundos	35
Annealing	62°C	30 segundos	35
Extensión	72°C	30 segundos	35
Extensión final	72°C	10 minutos	1

**Tabla 11.** Programa de termociclado utilizado para la síntesis in vitro de ADN de AGO2 a partir de los oligonucleótidos solapantes.

### 3.4 Evaluación de la actividad de la SaCas9

Con la finalidad de comprobar la funcionalidad del sistema CRISPR/SaCas9 generado en este trabajo, se evaluó la capacidad de la proteína SaCas9 de ejercer un DSB específico *in vitro* en los genes mCherry y AGO2. Para esto, se realizó una incubación de los componentes generados para cada sistema durante 1 hora a 37°C, según Burle-Caldas et al., 2018. Se utilizó una relación molar sgRNA:SaCas9 de 3:1, respectivamente. Los resultados de las digestiones para mCherry fueron analizados en geles de agarosa al 2% y en geles de poliacrilamida 15% para AGO2. En la **Tabla 12** se detalla el procedimiento empleado.

REACTIVOS	CONCENTRACIÓN INICIAL
SaCas9	6 µg
sgRNA	6 µg
ADN blanco	2 µg
Buffer SaCas9*	1 µL
H2O RNAsa free csp 20 µL	-

**Tabla 12.** Procedimiento empleado para el ensayo de actividad de la SaCas9 producida tomado de Burle-Caldas et al., 2018. **\*ver anexo**

Luego de la incubación, las reacciones de digestión fueron tratadas con Proteinasa K (800 U/mL, NEB) y ARNasa A (20mg/mL, NEB) de la siguiente manera:

- 1) Se agregó 1 µL de Proteinasa K para 20 µL de reacción de digestión y se incubó a 37°C durante 30 minutos.
- 2) Se realizó la inactivación de la proteínasa K mediante una incubación a 95 °C grados durante 10 minutos.
- 3) Se dejó enfriar las muestras en hielo durante 10 minutos.
- 4) Se agregó 2 µL de ARNasa y se incubó durante 8 minutos a 50 °C.

### 3.5 Purificación del ADN: fenol- cloroformo

Como se mencionó anteriormente, para la purificación de los productos de ADN obtenidos se utilizó el método fenol-cloroformo con posterior precipitación en etanol. A continuación, se detalla el procedimiento llevado a cabo en estas dos etapas para un volumen final de 100 uL de PCR:

- 1) Se añadieron 100  $\mu$ L de fenol-cloroformo al tubo de PCR y homogeneizó hasta formar una emulsión.
- 2) Se centrifugó a 12000g durante 1 minuto y medio a temperatura ambiente.
- 3) Se transfirió la fase acuosa (arriba) a un nuevo tubo, luego se descartó la fase orgánica (abajo) y la interfase.
- 4) Se repitieron los pasos 1, 2 y 3 la cantidad de veces posible, para obtener un mayor rendimiento.
- 5) Se añadió un volumen de cloroformo y se repitieron los pasos 2 y 3.
- 6) Se repitió el paso 5 con la finalidad de reducir al máximo posible los restos de fenol.
- 7) PRECIPITACIÓN:
  - A. Se agrego un volumen de Acetato de sodio a una concentración final de 0.3 M y se homogeneizó la muestra mediante vortex.
  - B. Se precipitó la muestra con dos volúmenes de etanol 95%, se homogeneizo y se almacenó en el freezer durante dos horas para la formación del precipitado.
  - C. Se centrifugo el ADN a máxima velocidad durante 20 minutos a 0°C.
  - D. Se removió el líquido sobrenadante.
  - E. Se realizó un lavado con etanol 70% y se centrifugó a 12000g durante dos minutos.
  - F. Se repitió el paso D y E
  - G. Se descarto el líquido sobrenadante
  - H. Se hizo una dilución del precipitado en 20uL de agua milli-Q.

## 3.6 Análisis de resultados mediante electroforesis

### 3.6.1 Electroforesis en gel de poliacrilamida

En lo que respecta a la expresión y purificación de la proteína SaCas9, los resultados se analizaron a través de SDS-PAGE. En primer lugar, se mezclaron los componentes requeridos para la formación del gel separador 10% según especifica la **Tabla 13**. Luego se introdujo la mezcla en los vidrios utilizados como molde hasta alcanzar aproximadamente  $\frac{3}{4}$  de su altura. Se termina de llenar el molde con alcohol 70% hasta que el gel separador polimerice y se remueve. Posteriormente, se mezclaron los componentes del gel concentrador como se indica en la misma tabla. Se vertió la mezcla por encima del gel separador, y luego se insertó el peine para la formación de los pocillos. El gel concentrador alinea la proteínas en una fina banda que entra al mismo tiempo y velocidad en el gel separador, donde se separan según su peso molecular. Una vez preparado el gel, se colocó en la cuba de electroforesis, se completó la misma con buffer de corrida 1X (**ver anexo**) y se procedió a cargar las muestras. Anteriormente, las muestras fueron hervidas durante 8 minutos para desnaturalizarlas luego de agregarles buffer de carga 1X (**ver anexo**). El gel se corrió a 100V mientras las muestras atravesaban el gel concentrador, y a 120V cuando ingresaron al gel separador. Al finalizar la corrida, se tiñó el gel durante una hora con agitación suave con azul de Coomassie (**ver anexo**), un colorante que forma complejos fuertes no covalentes con las proteínas. Su incorporación es proporcional a la cantidad de proteínas permitiendo detectar bandas de hasta 100 ng de proteína. Por último, se reemplazó la solución de tinción por solución de desteñido (**ver anexo**) y se incubó con agitación suave durante una hora.

REACTIVOS	GEL SEPARADOR 10% ACRILAMIDA	GEL CONCENTRADOR
Acrilamida: Bisacrilamida 30%	1,6mL	243,5μL
Tris-HCl 1.5 M pH 8 SDS 0.4%	1,3mL	-
Tris-HCl 0,5M pH6,8 SDS 0,4%	-	468,5μL
APS 10%	30,0μL	15μL
TEMED	6,0μL	3μL
Agua mQ	2,1mL	1,14mL

**Tabla 13.** Preparación de gel separador al 10% y gel concentrador en condiciones desnaturalizantes.

Para el análisis del ensayo de actividad de la SaCas9 en AGO2, se preparó un gel de poliacrilamida al 15% bajo condiciones no desnaturizantes para impedir la desnaturalización de la doble hebra del ADN (**Tabla 14**). El procedimiento seguido fue el mismo con la diferencia que no se requiere de gel concentrador y los geles separadores se prepararon utilizando buffer TBE 5X (**ver anexo**). Una vez preparado el gel, se completó la cuba con buffer de corrida TBE 1X y se procedió a cargar las muestras luego de añadirles buffer de carga para ADN (**ver anexo**). El gel se corrió en un rango de 80-100 V. Luego se sumergió en un baño de BrEt para su visualización seguido de un lavado con buffer de corrida 1X. Finalmente se introdujo el gen en un transiluminador ultravioleta para visualizar los resultados.

REACTIVOS	15% ACRILAMIDA
Acrilamida: Bisacrilamida 30%	2,5 mL
Buffer TBE 5X	1 mL
APS 10%	30µL
TEMED	15µL
Agua	1,45 mL

**Tabla 14.** Preparación de gel separador 15% no desnaturizante.

### 3.6.2 Electroforesis en gel de agarosa

Muchos resultados obtenidos se evaluaron en electroforesis en gel de agarosa. En este trabajo se prepararon geles de 1 y 2%. En el caso de ADN, se mezcló la agarosa con buffer TAE 1X (**ver anexo**) según se muestra **Tabla 15**. Luego de realizar esta mezcla se procedió a calentar hasta generar una solución homogénea. Posteriormente se agregó bromuro de etidio (BrEt) y se vertió el gel en la cuba de electroforesis con buffer de corrida TAE 1X. Por último, se colocó el peine para la formación de los pocillos. Una vez pronto el gen, se procedió a cargar las muestras las cuales previamente fueron preparadas con buffer de carga para ADN. La corrida electroforética se realizó a 100V. Al finalizar la corrida, se reveló el gel en un transiluminador ultravioleta. Para la visualización de los sgrNAs, se siguió el mismo procedimiento con algunas modificaciones: se preparó el gel añadiendo 1g de agarosa en 47,5mL de buffer MOPS (1X) (**ver anexo**), se disolvió la agarosa calentando, y se agregaron 2,5mL de formaldehído y BrEt. Las muestras fueron preparadas con buffer de carga para ARN 1X (**ver anexo**), y se calentaron por 10 minutos a 65°C.

REACTIVOS	1%	2%
Agarosa	1g	2g
Buffer TAE 1X	100mL	100mL

**Tabla 15.** Preparación gel de agarosa 1 y 2%.

### 3.7 Diseño ADN donante

Se diseñaron moléculas de ADN donante para la incorporación de secuencias mediante HDR luego del corte de CRISPR/Cas9. Al momento de diseñar estas moléculas de ADN donante, se tuvieron en cuenta las siguientes consideraciones:

- Longitud brazos de homología suficientemente largos para optimizar la eficiencia de recombinación por HDR (Li et al., 2014).
- La asimetría del ADN donante hacia ambos lados en referencia al sitio de corte tiene una eficiencia de recombinación superior en comparación con aquellos simétricos (Richardson et al., 2016).
- Inclusión de un sitio de enzima de restricción, ya sea mediante la inserción de la secuencia o realizar cambios de nucleótidos cerca del inserto para crear un sitio de restricción y facilitar la detección de la edición por HDR.
- Interrupción del reconocimiento del complejo CRISPR/SaCas9 mediante modificación de la secuencia PAM o la región de complementariedad con el sgRNA para evitar ediciones luego de la incorporación del fragmento de interés en el gen.

Dependiendo del objetivo que se quiera alcanzar en cada gen se diseñan distintos ADN donantes. Se utilizó el software SnapGene 1.1.3.

# RESULTADOS Y DISCUSIÓN

## 4.1 Diseño de experimento de edición CRISPR/SaCas9 por HDR

### 4.1.1 sgRNA

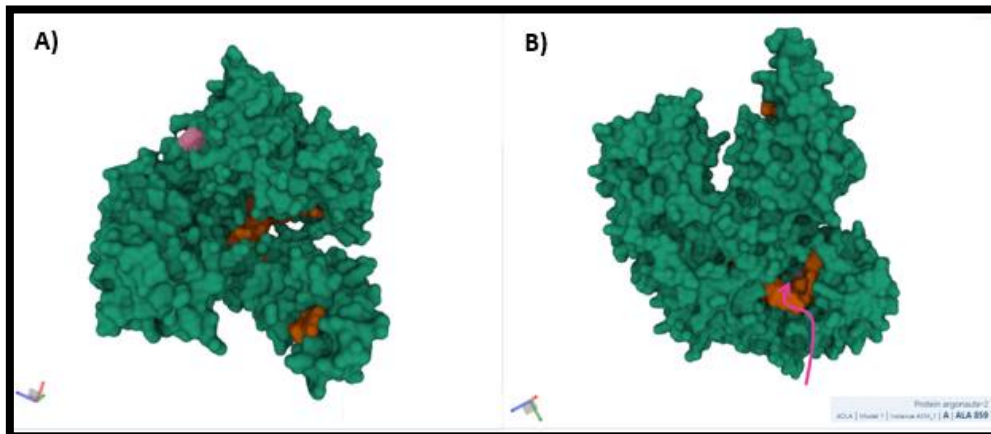
En los experimentos de edición genómica CRISPR/Cas9, la especificidad y eficiencia de escisión dependen principalmente del diseño de los ARN guía. En este sentido, es fundamental que el diseño de estas moléculas cumpla con los criterios establecidos en la literatura. En este trabajo, se llevó a cabo el diseño de los sgRNAs utilizando el recurso informático CRISPOR (Concordet & Haeussler, 2018). No obstante, a pesar de que existen consideraciones que los software de diseño toman en cuenta, los ARNs diseñados *in silico* fueron verificados de forma empírica. La generación de cada sgRNA implicó el diseño de dos oligonucleótidos parcialmente complementarios, que se muestran en la **Tabla 16**. Como se explica en el apartado 3.2.1, uno de los oligonucleótidos debe ser complementario a la región blanco del gen a editar (oligonucleótido guía), mientras que el otro será el encargado de formar las estructuras necesarias para la interacción con la proteína SaCas9 (oligonucleótido de andamiaje o andamio). Este último es universal, pudiendo ser utilizado en diferentes experimentos de edición con la proteína SaCas9.

OLIGONUCLEÓTIDOS		%GC
MCherry	5'-AAGCTAATACGACTCACTATAGGCAAGTAGTCGGGGATGTCGGGTTTGTAGTACTCTGGAACAGAA-3'	57
AGO2	5'-AAGCTAATACGACTCACTATAGGAGGACTCACCGGGGCCGGCTCGTTTGTAGTACTCTGGAACAGAA-3'	80
Universal SaCas9	5'-TAAAAAATCTCGCCAACAAGTTGACGAGATAAACACGGCATTTCCTGTTTTAGTAGATTCTGTTCCAGAGTACTAAAAC -3'	-

**Tabla 16.** Oligonucleótidos diseñados para la síntesis de los sgRNAs de mCherry y AGO2. En rojo se marca la secuencia correspondiente al promotor T7, en azul la secuencia complementaria al sitio blanco de edición de estos genes (secuencia guía) y en verde se representa la futura región andamio de interacción con la SaCas9. Se especifican los valores de contenido GC para cada secuencia guía (GC).

El sgRNA para el gen AGO2 se diseñó de forma tal que sea capaz de dirigirse hacia la región que codifica para el extremo N-terminal a fines de incorporar un péptido etiqueta en futuros experimentos de edición *in vivo* por HDR. La importancia en la generación de un etiquetado en AGO2 radica en la obtención de un epítipo a fines de que pueda ser inmunoprecipitado con un anticuerpo específico en la estrategia conocida como *pull-down* (en español, halar hacia abajo). La elección de dicho extremo se debe a que, en la estructura tridimensional producto de la cristalografía

de esta proteína, dicha región se halla en la superficie molecular (**Figura 15A**). De esta forma, se vería facilitado el reconocimiento por parte del anticuerpo. Por el contrario, el extremo C-terminal de la proteína AGO2, se encuentra en el interior de la proteína, dificultando la unión al anticuerpo y posiblemente comprometiendo la funcionalidad (**Figura 15B**), dado que en el extremo C-terminal ocurre la interacción con las proteínas GW18 que es fundamental en la función reguladora de AGO2 (Lian et al., 2009). En este sentido, se consideró importante evitar posibles alteraciones en la región C-terminal de la proteína AGO2.

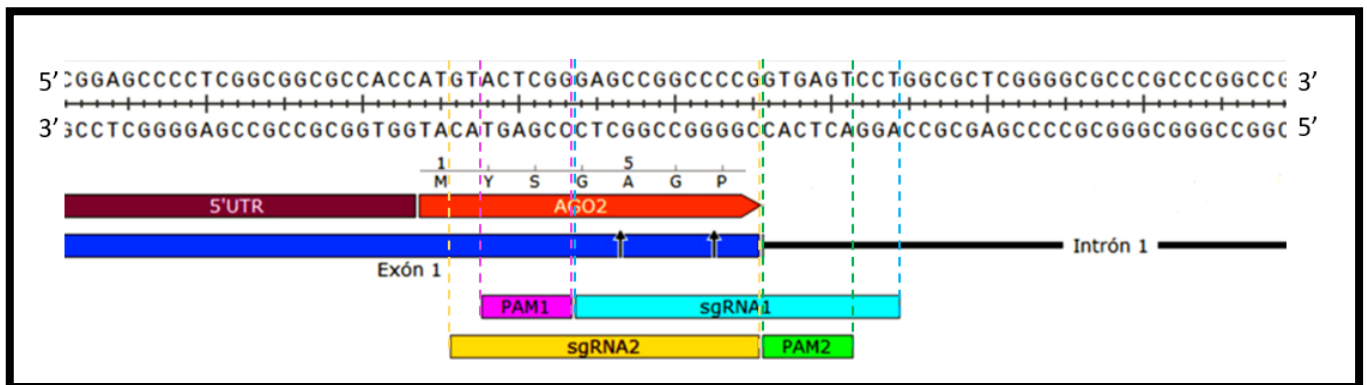


**Figura 15.** Estructura tridimensional de la proteína AGO2 (verde) en interacción con un miARN (naranja). **A)** Extremo N-terminal de AGO2 (rosado) **B)** Extremo C-terminal de AGO2 (flecha rosada) (Sehna et al. 2021).

Afortunadamente se encontraron dos sitios PAM específicos de SaCas9 en el primer exón del gen AGO2, el cual codifica parte de la 5' UTR y el comienzo de la secuencia codificante, es decir, la región N-terminal de la proteína. Se evaluaron las características de la secuencia complementaria del oligonucleótido guía (secuencia guía) para los dos posibles sgRNAs. En la **Figura 16** se muestra un esquema de alineamiento de estas secuencias con un fragmento de la región de ADN de AGO2 mencionada y la secuencia proteica que codifica. Según CRISPOR, ambas secuencias guías poseen un contenido GC de 80%, una puntuación de especificidad de aproximadamente 100/100 y se predice un valor de eficiencia de 73/100 para el sgRNA1 y de 76/100 para el sgRNA2. Adicionalmente, la secuencia guía del sgRNA1 presenta veinte posibles sitios off-targets en genoma humano, mientras que el sgRNA2 presenta doce. No obstante, se decidió optar por el primero porque guiará un corte doble hebra más cercano al sitio de inicio de la proteína (ATG) debido a que esta reportado que aumenta la eficiencia de HDR cuanto más cercano a la región de inserción se produzca el corte (Liang et al. 2017). El alto porcentaje de nucleótidos guanina y citosina (GC) de la región N-terminal de la proteína AGO2 impone el diseño de sgRNAs con un contenido GC superior al recomendado. Varios estudios concuerdan acerca de que aquellos sgRNAs que contienen un contenido GC muy bajo o alto son

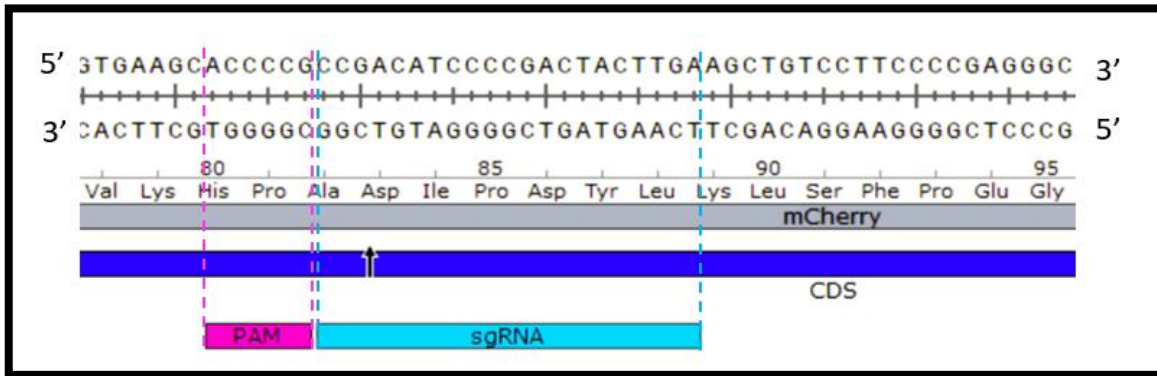


menos efectivos (Doench et al., 2014). Sin embargo, el análisis sobre el contenido óptimo de GC varía mucho entre diferentes organismos. En líneas de células animales, se informó un contenido de GC preferible del 40-60% (Liu et al., 2016). En plantas, el análisis de una cantidad significativa de estos ARNs validados reveló un amplio contenido de GC que oscilaba entre el 30% y el 80% (Liang et al., 2016). En base esto, el contenido de GC del sgRNA que se diseñó para dirigirse hacia AGO2 humano, podría reducir la eficiencia de corte en los ensayos de actividad de la SaCas9.



**Figura 16.** Esquema ilustrativo del alineamiento de las secuencias guías de los dos posibles sgRNAs para AGO2. En celeste se representa la secuencia guía correspondiente al sgRNA junto a la secuencia PAM, simbolizada en rosado. En amarillo se representa la secuencia guía del sgRNA2 y en verde la correspondiente secuencia PAM. Las flechas izquierda y derecha indicadas sobre el exón 1 muestran el sitio de corte de la Cas9 en el primer exón para el sgRNA1 y 2 respectivamente. En ese sitio se modificará la secuencia de codones de la proteína por inserción del ADN donante.

Por otro lado, en este trabajo se diseñó un sgRNA para realizar un *knock-out* del gen de la proteína fluorescente mCherry *in vivo* mediante *knock-in* con un ADN donante que al integrarse por HDR provoca la aparición de un codón stop. En nuestro laboratorio contamos con una línea celular que expresa el gen reportero mCherry de forma constitutiva, de modo que la inactivación de este gen fluorescente otorgaría una validación sencilla del sistema CRISPR/SaCas9 debido a la pérdida de fluorescencia en las células editadas, que puede cuantificarse rápidamente por citometría de flujo o microscopía de fluorescencia. Para esto, se buscaron secuencias PAM-SaCas9 en la secuencia de ADN de dicho gen a fines de generar un clivaje *in vivo*. Dentro de las numerosas posibilidades, se optó por la secuencia complementaria que agrupaba las mejores características con intenciones de generar un sgRNA preciso y eficaz. En la **Figura 17** se muestra un esquema del alineamiento de los guías obtenidos en CRISPOR con el segmento de secuencia de ADN de mCherry a editar. Se seleccionó una secuencia guía que exhibe las siguientes condiciones: contenido GC de 57%, puntaje de especificidad de 78/100, eficiencia estimada de 45% y el mérito de presentar un único posible sitio off-target en el genoma humano.



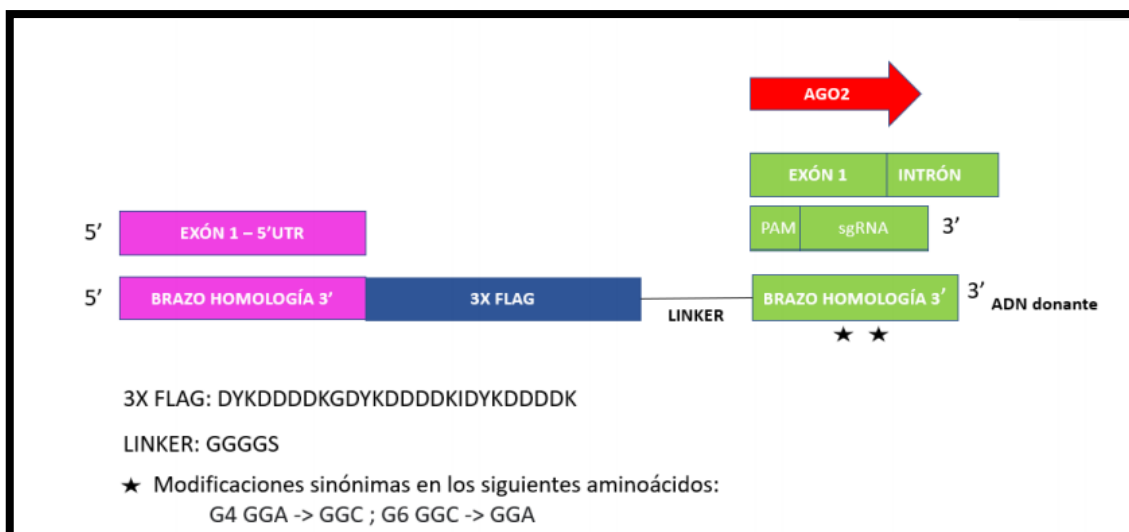
**Figura 17.** Esquema ilustrativo del alineamiento de la secuencia guía seleccionada para el sgRNA del gen mCherry. Se representa en celeste la secuencia guía del sgRNA adyacente a la correspondiente secuencia PAM, simbolizada en rosado. La flecha indica el sitio de edición en la secuencia del gen, representado en azul, y el congruente sitio de edición en la proteína que codifica la cual se muestra en gris.

#### 4.1.2. ADN donante

Una vez generado el clivaje in vivo en un sitio específico del genoma mediado por la Cas9, la célula repara este DSB mediante NHEJ o HDR. La primera de estas crea inserciones o deleciones cercanas al punto de corte mientras que HDR utiliza un ADN donante con cierta homología para insertar o eliminar con precisión nucleótidos en el sitio afectado (Ceccaldi et al., 2016). La fase del ciclo celular desempeña un papel importante a la hora de determinar qué vía será utilizada por la célula para reparar el DSB (Lin et al., 2014). Las células en las fases G1, S o G2 tienden a emplear NHEJ, mientras que la HDR se produce en las fases S y G2 del ciclo celular después de que se completa la replicación del ADN, ya que las cromátidas hermanas pueden entonces servir como moldes de reparación (Orthwein et al., 2014). Las células disminuyen la expresión de genes involucrados en el mecanismo de reparación HDR en la fase M y G1 (Orthwein et al., 2014). En este sentido, al momento de diseñar un experimento de edición por HDR, no solo es importante incorporarle a la célula una molécula de ADN donante homóloga a la región del sitio de corte, sino también la sincronización del ciclo celular en las fases S o G2. Adicionalmente, la introducción de los componentes CRISPR/Cas9 en complejos RNP preensamblados puede aumentar la eficiencia de HDR.

En la **Figura 18** se muestra un esquema ilustrativo del diseño del ADN donante para la edición del extremo N-terminal del gen AGO2. Como se puede observar, el ADN donante contiene la secuencia que codifica para el octapéptido etiqueta "FLAG" (Ácido Aspártico (D), Tirosina (Y), Lisina (K), D, D, D, D, K) repetida tres veces y

separadas por los aminoácidos Glicina (G) e Isoleucina (I). Esta secuencia corta de aminoácidos es reconocida específicamente por el anticuerpo anti-etiqueta FLAG y no se encuentra en ninguna proteína nativa humana. Debido a su pequeño tamaño, la etiqueta FLAG no se espera que afecte las propiedades bioquímicas, el transporte o las modificaciones postraduccionales de la proteína etiquetada o su interacción con otras moléculas. En el diseño del ADN donante, es necesario interrumpir el reconocimiento por el complejo CRISPR/SaCas9 a fines de que no se pueda ejercer un corte en el locus del gen luego de la edición. Para esto, se puede modificar la secuencia PAM reconocida por la SaCas9 o modificar la secuencia de complementariedad con el sgRNA. En este caso, no pudo ser posible la modificación del sitio PAM sin afectar a la proteína. Por un lado, se podría haber cambiado el primer codón que codifica para el aminoácido Tirosina (Y) por su sinónimo, sin embargo, este cambio no hace más que generar otro sitio de reconocimiento alternativo para la SaCas9. Por otro lado, los cambios sinónimos del siguiente codón que codifica para el aminoácido Serina (S), tienen una frecuencia de uso baja en *Homo sapiens*, con relación al original. Por este motivo, se decidió alternativamente modificar la secuencia de complementariedad con el sgRNA generando cambios sinónimos en el aminoácido Glicina (G) en las posiciones 4 y 6, evitando así que la sgRNA pueda reconocer la secuencia ya editada, tal como se ilustra en la **Figura 18**.



**Figura 18.** Esquema ilustrativo del diseño del ADN donante para edición en AGO2. El brazo de homología 5' es homólogo a la región 5' UTR (rosado). En el interior, se encuentra la secuencia que codifica para el epítipo etiqueta FLAG (azul) seguido de una secuencia linker (línea negra). El brazo de homología 3' es complementario al exón y al intrón (verde). Las estrellas negras indican los cambios sinónimos para modificar la región de complementariedad del sgRNA.

En lo que respecta al diseño del ADN donante para el gen mCherry, se siguió la misma lógica de construcción variando la secuencia de interés a incorporar. En este caso, el inserto que contiene tres codones STOP en los tres marcos de lectura posible a fines de inactivar la expresión por la acción de un STOP prematuro (**Figura 19**). Además,

dicho inserto contiene un sitio blanco para tres enzimas de restricción diferentes. De esta forma, se podría amplificar el gen mCherry por PCR y luego de un ensayo de digestión se podría evaluar la incorporación del fragmento mediante ensayos de detección basados en PCR, como el polimorfismo de longitud de fragmentos de restricción (RFLP); o de forma alternativa por escisión del producto de PCR en el sitio de la enzima de restricción recién introducido.



**Figura 19.** Esquema ilustrativo del ADN donante para la edición del gen mCherry. Se observan en los extremos las secuencias homólogas a las regiones flanqueantes al sitio de corte (rosado y verde). En el interior se observa la secuencia de interés a incorporar, la cual contiene tres codones stop para los tres marcos de lectura (azul).

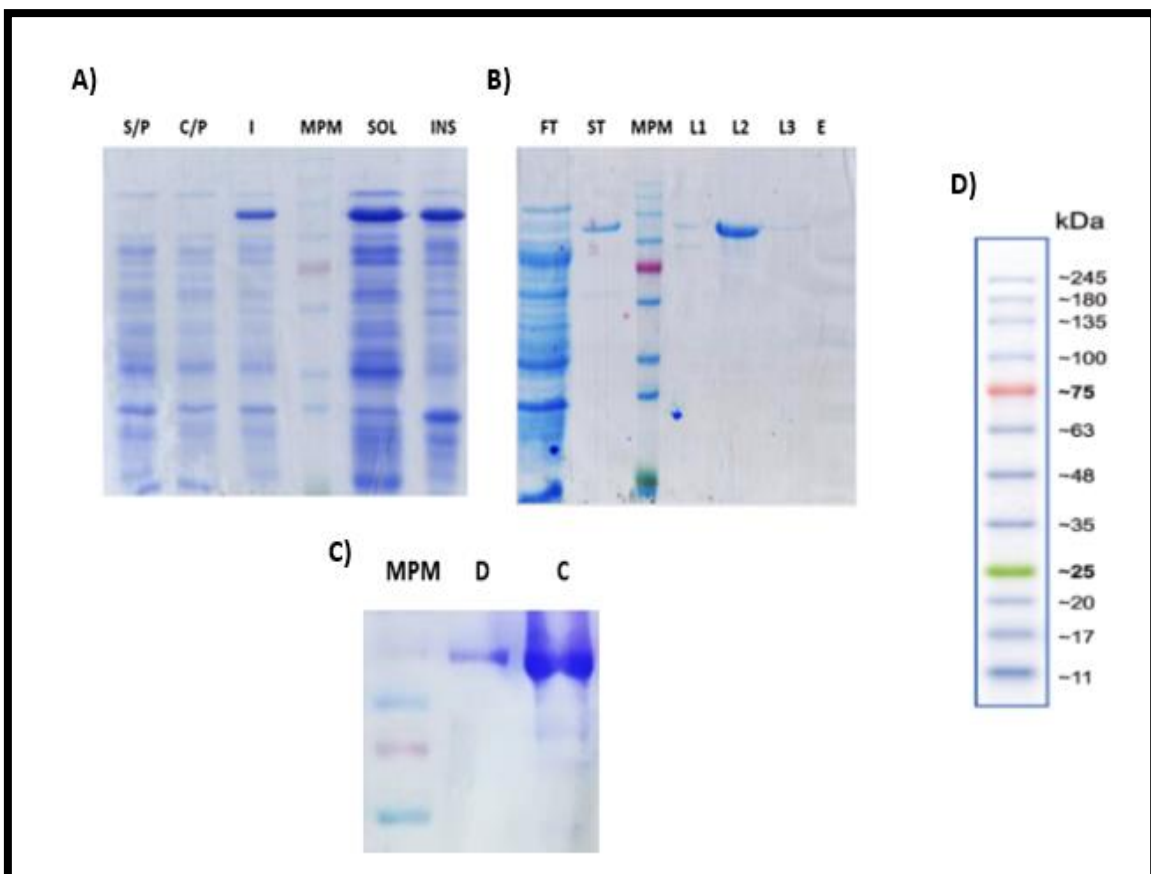
Para el diseño de moléculas de ADN donantes, es importante tener en cuenta aquellas variables que juegan un papel importante en el aumento de la tasa de HDR. Una de estas variables, se trata de la longitud de los brazos de homología 3' y 5'. Según la literatura, la eficiencia de recombinación aumenta a medida que aumenta la longitud de los brazos de homología (Song & Stieger, 2017). Estudios afirman que el largo del inserto contribuye a disminuir la eficacia de recombinación siendo compensada con el aumento del tamaño del brazo de homología. Se ha demostrado que para insertos cortos (menor o igual a 100pb), brazos de homología de 50pb son suficientes, de modo que el ADN donante pueda ser sintetizado como oligonucleótido. Por otro lado, ADNs donantes con brazos asimétricos parecen tener una mejora modesta en la recombinación sobre los brazos simétricos de acuerdo con la literatura (Li et al., 2014). En este trabajo se diseñaron los correspondientes ADNs donantes a partir de oligonucleótidos. En el caso de AGO2, el inserto contiene 70 pb y se encuentra flanqueado por brazos de homología de 70 pb en la región 5' y 50 pb en 3'. El ADN donante para la edición del gen mCherry, contiene un inserto de 20pb con un brazo de homología 5' de 76pb y otro 3' de 57pb.

## 4.2 Síntesis de los componentes CRISPR/SaCas9

### 4.2.1 SaCas9

La proteína Cas9 más comúnmente utilizada en los ensayos de edición genómica suele ser la de *Streptococcus pyogenes* (SpCas9). No obstante, en este trabajo se optó por producir de forma recombinante la Cas9 de *Staphylococcus aureus* (SaCas9). La elección de esta última se debe a que estudios previos han sugerido que la masa molecular podría afectar la eficiencia de transfección del complejo ribonucleoprotéico (RNP) Cas9/sgRNA y que a su vez sería menos tóxica para la célula receptora (Burle-Caldas et al., 2018; Medeiros et al., 2017). Además, ambas nucleasas presentan eficiencias de edición equivalentes, el hecho de que la SaCas9 posea un tamaño menor que la SpCas9 (124 vs 163 kDa), favorece la eficiencia de transfección de los complejos RNP en los ensayos de edición genómica (Burle&Caldas et al., 2018; Medeiros et al., 2017; Nishimasu et al., 2015). Es importante destacar que la secuencia PAM de la SaCas9 (5'-NNGRRT-3') tiene un tamaño mayor que la reconocida por la SpCas9 (5'-NGG-3'). Esta particularidad contribuye a reducir los posibles eventos off-target en el genoma en los experimentos de edición *in vivo* (Medeiros et al., 2017). La expresión de la proteína SaCas9 se llevó a cabo cultivando cepas *Rosetta*<sup>TM</sup> 2(DE3) pLysS (Novagen) de *E. coli* transformadas con el plásmido comercial inducible (p6xHis\_NLS-SaCas9). A partir de la inducción con IPTG de 5.4 L de cultivo se obtuvieron seis pellets de 9,0 g de peso humedo correspondientes a 900mL de cultivo cada uno, lo que correspondió a un rendimiento de 10g/L de cultivo. En la **Figura 20A** se muestra una electroforesis SDS-PAGE al 10% donde se observa el perfil proteico de las diferentes etapas del proceso de expresión y obtención de la SaCas9. Se puede apreciar la expresión de la proteína SaCas9 (aprox.130 kDa) en la muestras correspondientes a bacterias transformadas e inducidas con IPTG (**carril I**), y en las muestras proteicas solubles (**carril SOL**) e insoluble (**carril INS**) obtenidas a partir de la lisis química. En los tres casos, la SaCas9 es más abundante que otras proteínas. La muestra de proteínas no transformadas sin inducir (**carril S/P**) no expresan la proteína SaCas9 al igual que las bacterias transformadas sin inducir (**carril C/P**), indicando que el plásmido no tiene una expresión espuria de SaCas9 sino que el gen está reprimido en ausencia de IPTG. Esto permite un buen crecimiento bacteriano, porque no hay estrés traduccional o toxicidad previo a la inducción con IPTG. Es así que no se observa alteración en el perfil proteico en ninguna de las muestras a excepción de la expresión de la SaCas9 luego de la inducción con IPTG (I, Sol, Ins). Asimismo, el perfil electroforético de la fracción soluble (Sol) e insoluble (Ins) del extracto proteico inducido muestra que una proporción importante de la proteína queda soluble en las condiciones ensayadas, lo que evita la necesidad de usar protocolos adicionales de solubilización o de cambiar la estrategia de producción. Posteriormente se procedió a la purificación de la fracción soluble obtenida, mediante cromatografía de afinidad utilizando una columna de Níquel (*HisTrap HP* de 1mL, GE<sup>TM</sup>). En la **Figura 20B** se puede observar las distintas etapas correspondientes al proceso de purificación. En las primeras eluciones con 30 mM de imidazol se colectaron las fracciones de

proteínas no unidas a la columna, denominado *Flow-through* (**carril FT**). El primer lavado (**carril L1**) se realizó en 53 mM de imidazol, el segundo en 82 mM y el tercer lavado (**carril L3**) se realizó con 124 mM de imidazol. Para el lavado final o elución (**carril E**) se utilizó 500 mM de imidazol. Se puede observar una proteína SaCas9 producida en el laboratorio de la Profesora Santuza Teixeira (Universidad Federal de Minas Gerais) que fue donada para uso como control (**carril ST**). En el análisis por SDS-PAGE se puede apreciar que la SaCas9 purificada coincide en peso molecular de la SaCas9 control (130 kDa). Se obtuvo la mayor concentración de SaCas9 en el lavado correspondiente a 82 mM de imidazol (**carril L2**). En esta elución también se pueden observar bandas de menor tamaño que podrían corresponder a contaminantes que co-purifican con la SaCas9. También se observa la presencia de la SaCas9 en los carriles L1 y en L3, en concentraciones considerablemente menores comparado a L2. En la elución final con 500 mM de imidazol (**carril E**) no se visualizan bandas. En la **Figura 20C** se muestra la proteína SaCas9 luego de atravesar la columna PD-10 en el proceso de desalado mediante purificación por gel filtración (**carril D**). Además, se puede apreciar la proteína SaCas9 luego de ser concentrada mediante centrífuga Milipore (carril C). En total se obtuvieron 24mg de proteína (2,4 ug/uL) a partir de 54g de bacterias.



**Figura 20.** Análisis por SDS-PAGE de las proteínas obtenidas a largo del proceso de producción de SaCas9. **A)** Perfil proteico expresado de cepas Rosetta con o sin plásmido (S/P, C/P), después de la inducción con IPTG (I) y de las fracciones soluble e insoluble obtenidas luego de la lisis

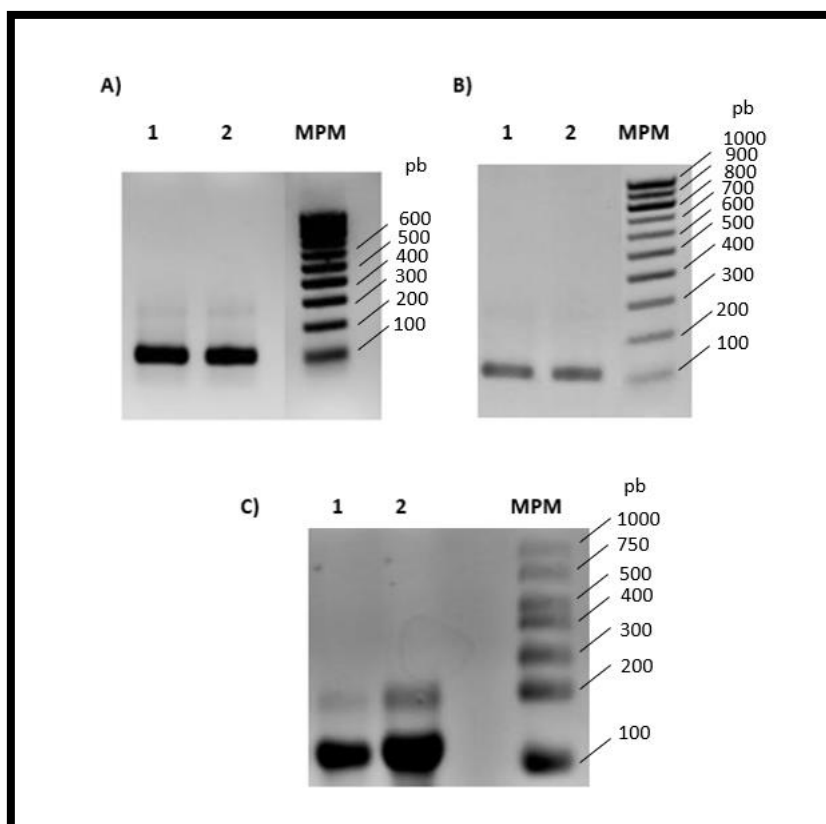
química (SOL, INS). **B)** Proceso de purificación de SaCas9 mediante cromatografía de afinidad. Se observa la fracción proteica no retenida (FT) (fracción no retenida), los lavados realizados a 53,82 y 124 mM de imidazol (L1, L2 y L3), el eluyente final a 500 mM de imidazol (500mM de imidazol) y una alícuota de SaCas9 control (ST). **C)** Se puede observar la proteína SaCas9 luego del desalado mediante cromatografía por gel-filtración (D) y luego de ser concentrada mediante centrífuga (C). SDS-PAGE 10%. **D)** Nombre del marcador de peso molecular utilizado (MPM): AccuRuler RGB PLUS (GeneBio Systems, INC™).

Con respecto a la cantidad de SaCas9 producida es posible afirmar que se obtuvieron rendimientos satisfactorios en los cultivos bacterianos lo que permitió generar, luego de la purificación, 24mg proteína, que constituye una cantidad suficiente para los ensayos *in vitro* de este trabajo y para posteriores ensayos *in vivo*. Como se puede observar en la **Figura 20A**, la proteína SaCas9 se obtuvo en la fracción soluble e insoluble. La presencia de SaCas9 en la fracción insoluble podría explicarse debido a errores en el plegamiento de la proteína, siendo incorporada en agregaciones proteicas denominadas cuerpos de inclusión (Mohan et al., 2005). Estas agregaciones se forman en el espacio citoplasmático y periplásmico de las células de *E. coli* cuando se obtienen altos niveles de expresión de proteínas heterólogas (más del 2% del total de proteínas bacterianas), entre otras cosas (Mohan et al., 2005). El hecho de obtener mayor concentración de proteína en la fracción soluble (evaluada cualitativamente) significó un contribuyente a que la purificación sea más sencilla. La fracción insoluble fue descartada ya que el rendimiento de la fracción soluble fue satisfactorio. El grado de pureza de la proteína obtenida no fue calculado, pero de modo cualitativo según el patrón de bandas observado en los geles teñidos con azul de Coomassie, parece aceptable para proseguir en su uso para el protocolo de edición. En lo que refiere a la purificación de la SaCas9, el hecho del no requerir altas concentraciones de imidazol (53mM y 82mM) para la elución podría deberse a una interacción débil entre la columna y el *His-Tag* fusionado a SaCas9. Tal vez la existencia de un posible plegamiento particular del extremo N-terminal de la proteína podría de alguna manera enmascarar el *His-Tag*, haciendo que éste no se encuentre lo suficientemente expuesto para su interacción con el níquel de la columna. A su vez, es importante destacar que el *His-Tag* no se ubica precisamente en el extremo N-terminal ya que existen otros elementos allí presentes. Por ejemplo, el primer elemento presente es el tag de fusión *Trx* que se utiliza comúnmente para ayudar a solubilizar las proteínas de expresión ectópica en *E. coli*. Este elemento de 109 aminoácidos de extensión podría estar involucrado en la dificultad entre el reconocimiento columna- *His-Tag* por la columna.

### 4.2.2 sgRNAs

Luego del diseño y síntesis de los oligonucleótidos solapantes mostrados en **Tabla 16**, se procedió a sintetizar las moléculas de sgRNAs en dos etapas consecutivas. En primer lugar, se realizaron ensayos de extensión a partir de los extremos 3' para generar moléculas de ADN doble hebra. Posteriormente se hicieron transcripciones *in vitro* (IVT) utilizando como molde el ADN previamente generado.

En la **Figura 21A** y **21B** se observa el resultado de la síntesis de los ADN molde para las IVTs de los sgRNAs de mCherry (**carril 1**) y AGO2 (**carril 2**) antes y después de purificados respectivamente. En el caso de mCherry, se obtuvo una concentración de 168.8 ng/ $\mu$ L y una pureza de 1,7. En lo que respecta a AGO2, la concentración fue de 195 ng/ $\mu$ L y una pureza de 1,82. Los productos mostrados en el gel coinciden con el tamaño esperado de 120 pb según el MPM utilizado. Los ARNs sintetizados por IVT a partir de estos moldes de ADN, se muestran en la **Figura 21C**. Las concentraciones alcanzadas para los sgRNAs de mCherry (**carril 1**) y AGO2 (**carril 2**) fueron de 652 y 1272 ng/ $\mu$ L. El valor de pureza obtenido fue de 2,25 para ambos casos. La presencia de dos bandas en el gel correspondiente a los sgRNAs, probablemente se debe a la formación de dímeros o multímeros y podría resolverse mediante desnaturalización térmica antes del experimento de escisión con la SaCas9 (Liu et al.,2016).



**Figura 21.** Síntesis de los sgRNAs. Análisis de los ADN molde para IVT de los sgRNA de mCherry (1) y AGO2 (2) antes (**A**) y después (**B**) de la purificación por fenol-cloroformo. Electroforesis en



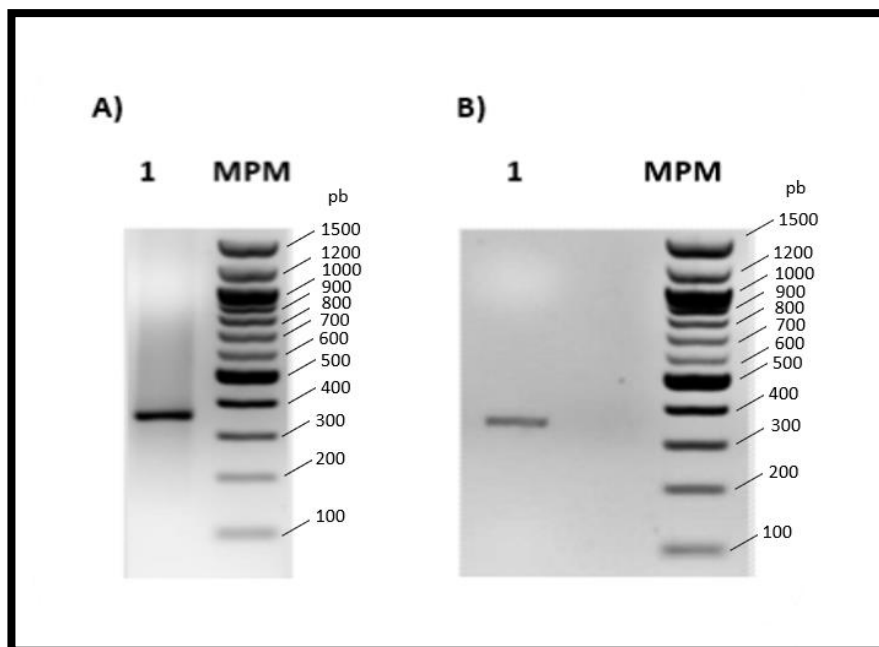
gel de Agarosa 1%, teñido con BrEt. MPM: Perfect DNA 100 bp Ladder, Novagen. **C)** Análisis de los sgRNAs de mCherry (1) y AGO2 (2) sintetizados por IVT usando los ADN moldes de A y B y purificados. Electroforesis en gel de Agarosa 2% y Buffer MOPS. MPM: Century™-Plus RNA Markers, ThermoFisher Scientific.

En suma, se obtuvieron sgRNA para ambos genes en concentración y pureza adecuados para su uso en el ensayo de evaluación de la SaCas9.

### 4.2.3 ADN blanco

Con el propósito de evaluar la actividad de la proteína SaCas9 producida y verificar el diseño de los sgRNAs sintetizados para los genes mCherry y AGO2, es necesario la obtención de las moléculas de ADN blanco a ser clivadas por la enzima *in vitro*.

La **Figura 22A** y **22B** muestran el resultado de las reacciones de PCR para la amplificación de una secuencia de 360 pb perteneciente al gen mCherry (**carril 1**) antes y después de la purificación. Como se puede observar, se obtuvo un fragmento único coincidente con el tamaño esperado. El producto se obtuvo a una concentración de 182 ng/μl y pureza de 1,76.



**Figura 22.** Análisis de la amplificación del ADN blanco del gen mCherry (1) mediante PCR antes **A)** y luego **B)** de la purificación mediante fenol cloroformo. Agarosa 1%. Marcador de peso molecular (MPM): 100 pb DNA Ladder, NEB.

Las reacciones de PCR de mCherry a partir de plásmido no presentaron dificultades. Se logró una amplificación específica y eficiente con los cebadores y el programa de termociclado descrito en materiales y métodos. Sin embargo, la baja concentración de ADN luego de la purificación en relación con la cantidad requerida para los ensayos de evaluación de la SaCas9 (2ug/reacción) hizo que este sea un factor limitante.

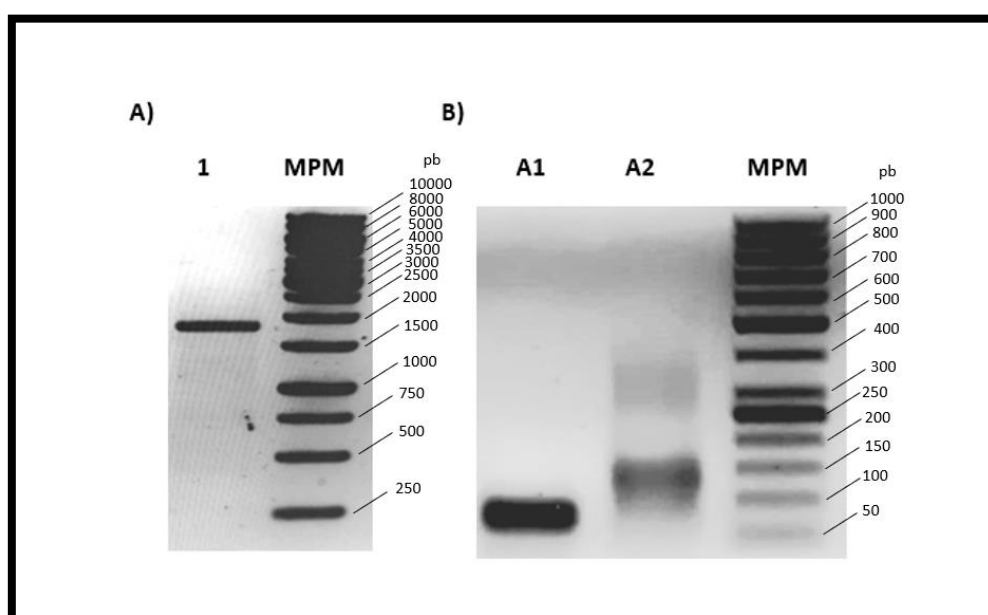
Para la amplificación de la secuencia de interés en el extremo N-terminal del gen AGO2 a partir de ADN genómico humano se realizaron varios ensayos de PCR utilizando tres juegos de cebadores distintos según se describe en 3.3.2. Inicialmente, se llevó a cabo una PCR utilizando el PAR 1 de cebadores para los cuales se esperaba un fragmento de aproximadamente 613pb. La abundancia de productos inespecíficos y la ausencia del amplicón de interés condujo a modificar variadas condiciones, incluyendo temperaturas de annealing en un termociclador de gradiente, concentración de los cebadores, entre otras. Sin embargo, no se logró amplificar el fragmento esperado (resultado no mostrado). Consecuentemente, se procedió a hacer otra PCR con el PAR 2 de cebadores esperando un producto de amplificación de 874pb aproximadamente. Si bien se obtuvieron varios productos inespecíficos, a diferencia de la PCR con el PAR1 de cebadores, en este se obtuvo un fragmento que parecía tener el tamaño esperado. Se decidió probar de purificarlo para utilizarlo en los ensayos de clivaje, por lo que se extrajo del gel de agarosa utilizando el “Monarch® DNA Gel Extraction Kit (NEB #T1020)”. Para corroborar si este fragmento tenía la secuencia de interés de AGO2, se digirió el mismo con enzimas de restricción capaces de clivar esta secuencia, y en paralelo se digirió un plásmido para confirmar que las condiciones de digestión eran correctas. Aunque para el plásmido control se obtuvieron los fragmentos correspondientes al producto de corte, en el producto de PCR no se visualizaron los fragmentos esperados para ninguna de las enzimas de restricción. Además, se mandó el fragmento a secuenciar por Sanger, lo que confirmó que no se trataba de la secuencia esperada, sino que correspondía a una región del genoma compuesta de múltiples repetidos (resultados no mostrados). A continuación, se llevó a cabo una PCR utilizando el PAR 3 de oligonucleótidos, los cuales amplifican para un producto esperado de 1755 pb. Se realizó una PCR en un termociclador de gradiente a diferentes temperaturas de annealing. Nuevamente, no se observó producto de amplificación bajo ninguna condición. Finalmente se hizo la misma PCR, pero utilizando la enzima “Kappa HotStart Hifi” y, consecuentemente, se logró amplificar el producto esperado (**Figura 23A**).

Dada la dificultad de obtener una amplificación específica de la región del inicio del gen de AGO2, en paralelo se sintetizaron *in vitro* dos secuencias de ADN conteniendo la región de complementariedad con el sgRNA generado para AGO2 seguido de la secuencia PAM correspondiente. Estas secuencias se obtuvieron luego de un ensayo de extensión de extremos 3'OH a partir de oligonucleótidos solapantes previamente diseñados (**Tabla 17**). Los oligonucleótidos F1 Y R darán lugar a la molécula de ADN

blanco de 82pb (A1) mientras que el F2 y R a una molécula blanco de 139 (A2). En la **Figura 23B** se muestran los resultados obtenidos para la síntesis de las mismas luego del proceso de purificación. Se observa que los fragmentos coinciden con los tamaños esperados. Para el ADN blanco A1 se obtuvo una concentración de 300 ng/μL a una pureza de 1,67. Por otro lado, para el A2 se obtuvo una concentración de 348 ng/μL y una pureza de 2,23.

OLIGONUCLEÓTIDOS		%GC
F1	5'- CCCTCGGCGGGCGCCACCATGTACTCGGGAGCCGGCCCCGGTGAGTCCTGG- 3'	76
F2	5'GGCTCGCGGCGTCTGGGGCCCGTGGCGCGCGCGCGGCCGCCCTCGGCCCGGAGCCCCCTCGGCGGCG CCACCATGTACTCGGGAGCCGGCCCCGGTGAGTCCTGGCGC-3'	83,5
R	5'AGCGCGCGAACGGCCGGGGCGGGCGCCCCGAGCGCCAGGACTCACGGGGCCGGCTCCCGAGTACATGG-3'	79

**Tabla 17.** Oligonucleótidos diseñados para dar lugar a dos moléculas de ADN blanco correspondiente al gen AGO2. Se detalla el contenido GC para cada uno de ellos (%GC).



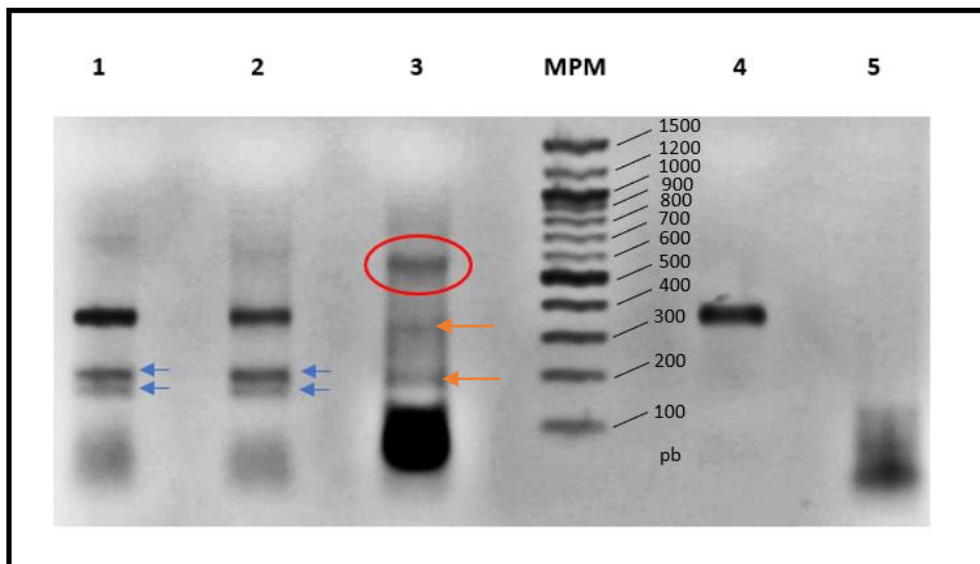
**Figura 23.** Análisis de la obtención del ADN blanco del gen AGO2. **A)** Producto de PCR obtenido con el PAR3 de cebadores y la enzima Kappa HotStart Hifi (1). Agarosa 1% teñida con BrEt. MPM: GeneRuler 1kb DNA ladder, Thermo Scientific. **B)** Producto purificado de ensayo de extensión de los extremos 3'OH a partir de oligonucleótidos solapantes (A1 y A2). Agarosa 2%. MPM: GeneRuler 50bp DNA ladder, Thermo Scientific.

Debido al alto contenido GC de la región N-terminal del gen AGO2, y la dificultad que esto representó, se llevaron a cabo distintas estrategias con propósito de amplificar la secuencia localizada en un contexto genómico desafiante. En la PCR con el PAR 1 de cebadores se variaron diferentes condiciones con la finalidad de obtener nuestro producto de interés y evitar amplificaciones inespecíficas. Por ejemplo, se decidió modificar la relación de concentración entre los cebadores (0,5 FP- 1 RP y viceversa) para detectar si los amplicones inespecíficos eran generados por uno de estos. Por otro lado, también se aumentaron las temperaturas de annealing con la finalidad de favorecer la hibridación específica de los cebadores. A pesar de éstas y otras modificaciones realizadas, no fue posible amplificar la secuencia de interés ni tampoco evitar amplificación de productos no deseados. Estudios de secuencia tipo BLAST sugieren que la formación de productos inespecíficos de amplificación se debía a la capacidad del cebador reverso de hibridarse en múltiples regiones del genoma. En base a esto, se optó por realizar otra PCR manteniendo el mismo oligo forward y utilizando un nuevo oligo reverso disponible en el laboratorio, para dar lugar al PAR 2 de cebadores, que sería capaz de amplificar un producto único de aproximadamente 800 pb correspondiente a AGO2 . No obstante, en este caso también se obtuvieron varios productos de amplificación y, aunque uno de ellos coincidía con el tamaño esperado, no correspondía a la secuencia de AGO2. Posteriormente se diseñó el PAR 3 de cebadores para amplificar una región más grande conteniendo la región requerida para generar el corte en AGO2. Sin embargo, no se obtuvieron productos de amplificación. La imposibilidad de amplificar AGO2 con polimerasas convencionales y variando algunas condiciones de la PCR, condujo a pensar en ideas alternativas. Las secuencias de ADN con alto contenido GC (65%) pueden afectar la eficacia de la PCR debido a que tienden a plegarse en estructuras secundarias complejas, no solo provocando la detención de la ADN polimerasa sino obstaculizando la hibridación con los cebadores durante la fase de annealing del ciclo de PCR, dando como resultado una amplificación incompleta o no específica. Además, los cebadores utilizados para amplificar regiones ricas en GC a menudo tienen una gran capacidad para formar dímeros propios y cruzados y una fuerte tendencia a plegarse en estructuras de tallo-bucle que pueden impedir el progreso de la ADN polimerasa a lo largo del ADN molde (Flores-Juárez et al., 2016). Para abordar estos problemas, es aconsejable utilizar aditivos adicionales en la reacción de PCR, tales como dimetilsulfóxido (DMSO), formamida, glicerol y betaína. Estos se utilizan a menudo debido a su capacidad de reducir el efecto de los enlaces de hidrógeno, desestabilizar las estructuras secundarias y mejorar la especificidad de la reacción de amplificación (Mammedov et al., 2008). Por otro lado, se demostró que es posible emplear variantes de la PCR convencional como *HotStart PCR* o *stepdown PCR* para la amplificación de regiones de ADN ricas en GC (Lorenz, 2012). Por ejemplo, *HotStart PCR* es una técnica que reduce la amplificación inespecífica y dímeros de cebadores (Green & Sambrook, 2018). La posibilidad de utilizar una enzima especializada en la amplificación de productos con un alto contenido GC ya que contiene un buffer particular conteniendo algunos de los aditivos mencionados anteriormente y cumple con la propiedad *HotStart* permitió

amplificar exitosamente la región esperada del gen AGO2. De esta forma, se destaca la importancia de modificar las condiciones de la PCR cuando se amplifican regiones del genoma con secuencias de ADN de ricas en GC y repetidos. Desafortunadamente, la escasa cantidad de enzima disponible no permitió generar la cantidad de ADN molde necesario para el ensayo *in vitro* de actividad de la SaCas9 ya que se requiere de 2 $\mu$ g de ADN para cada una de las reacciones. Por este motivo, se condujo a optar por la síntesis *in vitro* del ADN blanco de AGO2 mediante oligonucleótidos solapantes.

### 4.3 Evaluación de la actividad de la SaCas9

Una vez obtenidos todos los componentes del sistema de edición CRISPR/Cas9, se evaluó *in vitro* la actividad de la proteína SaCas9 producida de forma recombinante en este trabajo. Con el propósito de comprobar si esta proteína era capaz de generar un DSB en los genes mCherry y AGO2 se la incubo con el ADN blanco y el sgRNA respectivo, según se detalla en 3.4. Los resultados fueron analizados por electroforesis en gel de agarosa 2% (mCherry) y poliacrilamida 15% no desnaturizante (AGO2). De esta forma, considerando el marcador de peso molecular correspondiente, conociendo el tamaño del ADN blanco, y la ubicación del sitio de corte en el mismo, se podría verificar la presencia de los fragmentos esperados comprobando así la funcionalidad de la SaCas9. En la **Figura 24** se muestra el resultado del ensayo de actividad de la SaCas9 para el gen mCherry. Para las reacciones de digestión (SaCas9+ADNblanco+sgRNA) se utilizó la mitad de concentración de los componentes en relación a la requerida por el protocolo (*Burle-Caldas et al., 2018*), obteniendo así un volumen final de 10 uL en cada una. Se llevaron a cabo tres digestiones (**carriles 1, 2 y 3**), dos de ellas tratadas con proteinasa K y ARNasa A (**carriles 1 y 2**). Por otro lado, se realizó una incubación control solamente con el ADN blanco (**carril 4**) y otra con el sgRNA (**carril 5**) a efectos de evaluar la integridad y el patrón de migración de estos componentes individualmente.



**Figura 24.** Ensayo de actividad de la SaCas9 para el gen mCherry. Se observa una digestión duplicada utilizando proteinasa K y ARNasa (1 y 2), una digestión sin proteinasa K y ARNasa (3), el ADN blanco solo (4) y el sgRNA solo (5). Las flechas azules indican los productos de corte. El círculo rojo muestra el “shift” mencionado en el texto. Las flechas naranjas indican fragmentos de tamaño superior a los del producto de corte. Agarosa 2%. MPM: 100 pb DNA Ladder, NEB.

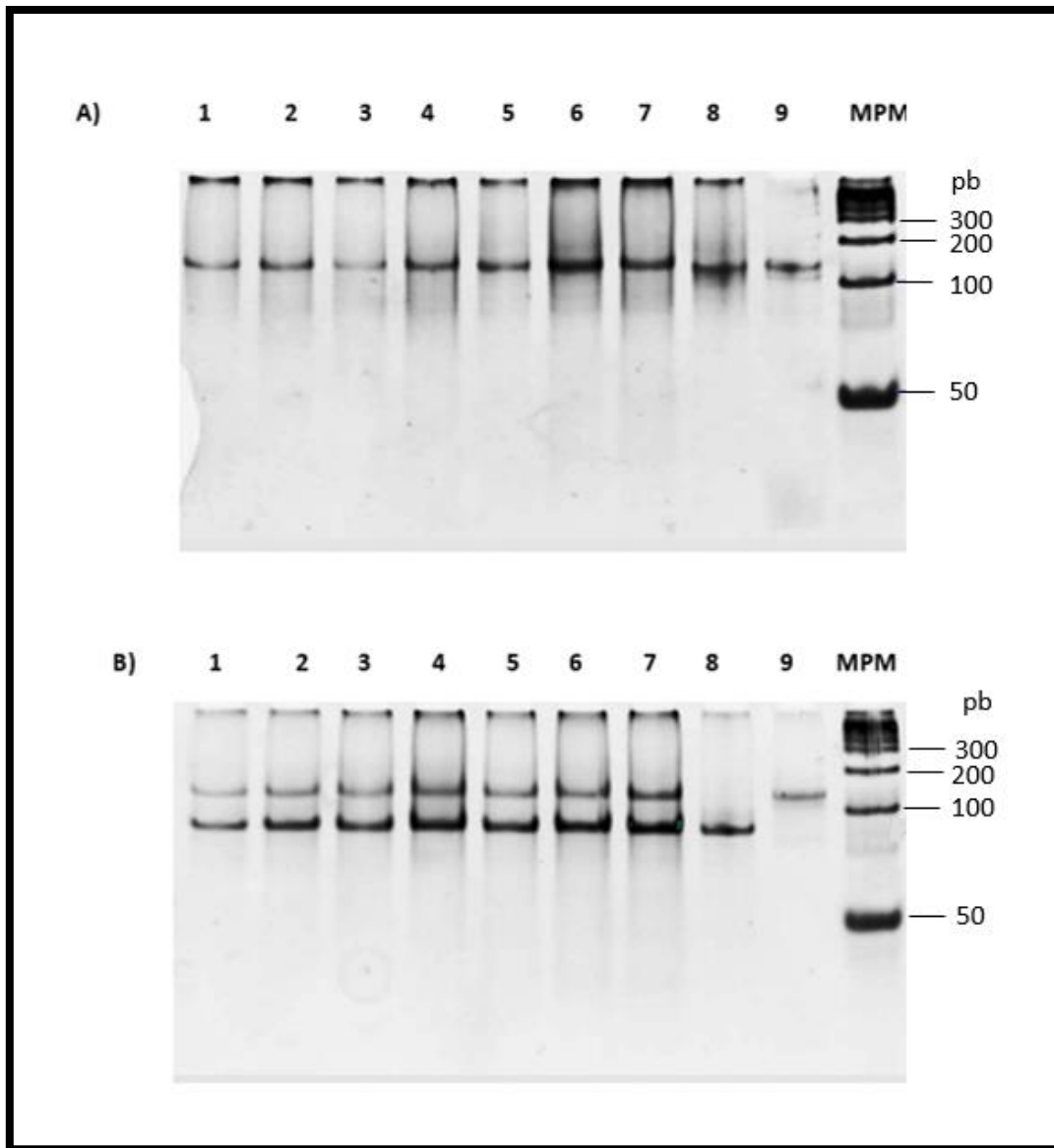
Se puede observar en la reacción de digestión sin proteinasa K y sin RNasa (**carril 3**) cuatro bandas: Una banda de un tamaño mayor al ADN blanco de mCherry (shift) (aprox. 550pb), dos bandas con un tamaño mayor del esperado para los productos de corte y, una banda intensa de aproximadamente 100 pb corresponde al sgRNA sin RNAsa. Por otro lado, en las reacciones tratadas con proteinasa K y RNasa A (**carriles 1 y 2**), se observa una banda correspondiente al tamaño del ADN blanco sin cortar (360pb) y otras dos bandas coincidentes con los productos luego del corte con la SaCas9 (200 y 160pb). El ADN blanco (**carril 4**) y el sgRNA (**carril 5**) migran acorde a su tamaño. La diferencia de intensidad del sgRNA entre el **carril 3** y **5** se debe a que en este último se cargó menos concentración. El patrón de migración diferencial o “shift” observado en la digestión sin proteinasa K puede explicarse por la asociación de la proteína SaCas9 con los fragmentos de ADN luego del corte. Estudios con SpCas9 demuestran que luego de generar un DSB la proteína puede permanecer unida a los fragmentos de corte generando una banda del mismo tamaño que una dSpCas9 inactiva catalíticamente. Esta proteína es capaz de reconocer la secuencia PAM, pero no de ejercer un corte (Sternberg et al., 2014). De esta forma, luego del tiempo de incubación requerido para el ensayo de actividad de la SaCas9, se hizo un tratamiento con proteinasa K a las digestiones. Efectivamente, la degradación resultó en la pérdida del shift. De esta forma, resulta lógico concluir que los fragmentos de mayor tamaño de los esperados luego del corte del **carril 3** corresponden a la SaCas9 unida a cada fragmento luego del corte.

En base a los resultados obtenidos, se puede afirmar que la SaCas9 producida en este trabajo logró ejercer *in vitro* un DSB específico y parcial en el gen mCherry. Esto no solo evidencia la funcionalidad de la proteína SaCas9 sino también el acertado diseño del sgRNA capaz de guiar un corte hacia una región esperada. Es importante destacar que se trata de un trabajo cualitativo, en el cual se consideró suficiente la visualización de las bandas de corte del tamaño adecuado en el gel, para confirmar que la SaCas9 realizó el DSB en el ADN blanco. Sin embargo, para una confirmación definitiva sería necesario la secuenciación de los fragmentos de ADN generados a partir de los cortes por la nucleasa. Dado que la proteína SaCas9 que se produjo aquí fue capaz de cortar el 100% de otros ADN moldes con sgRNA específicos, sin diferenciarse de una enzima comercial (resultados no mostrados), creemos que la reducida eficiencia del corte *in vitro* en este trabajo puede deberse al diseño del sgRNA. De requerirse una mejor eficiencia, debería diseñarse y testarse sgRNA adicionales. Dado que la eficiencia de la edición es fácilmente medible en este sistema de reportero, convendría hacer el ensayo *in vivo* antes de descartar el sgRNA ya sintetizado.

En lo que respecta al ensayo de actividad de la SaCas9 en AG02, la dificultad inicial para obtener suficiente ADN blanco a partir de la amplificación de ADN genómico humano por PCR, nos condujo a sintetizar el mismo a partir de oligonucleótidos solapantes. Estos oligonucleótidos dieron lugar a dos moldes de ADN conteniendo la región PAM para la SaCas9 y, enseguida corriente arriba de la misma, la secuencia de complementariedad con el sgRNA diseñado para AGO2. Considerando el corto tamaño de los fragmentos esperados luego del corte ejercido por la SaCas9, las digestiones se analizaron en geles de poliacrilamida al 15% teñidos con BrEt.

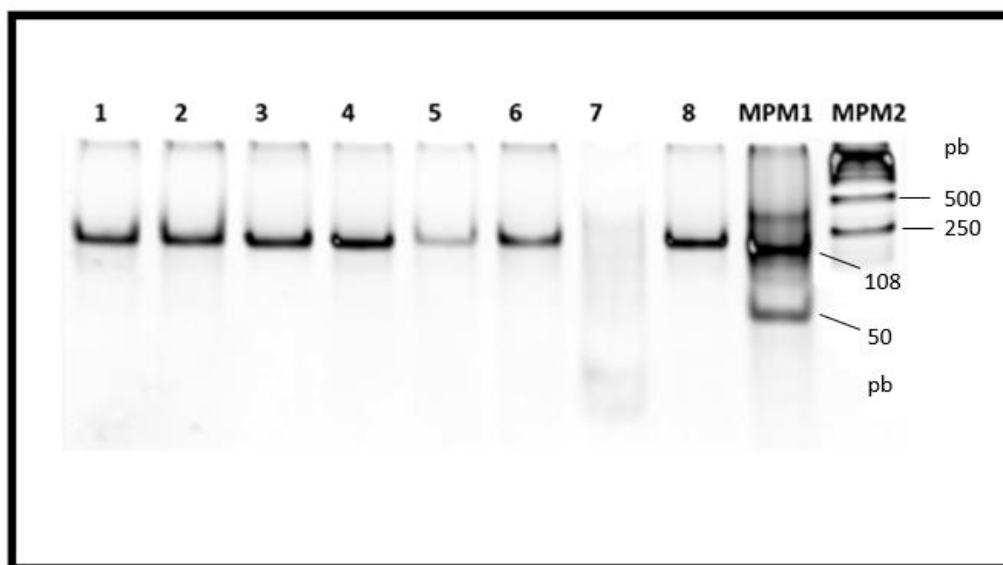
En la **Figura 25** se muestra el resultado del ensayo de actividad de la SaCas9 para ambos ADNs blanco sintetizados *in vitro* correspondientes a AGO2 (A1 y A2) purificados con fenol cloroformo. Se esperan fragmentos de 30 y 52 pb para A1 luego del corte y ,de 52 y 87pb para A2. Se llevaron a cabo las mismas reacciones para ambos, pero en las digestiones para A1 se utilizó RNasa A. Todas las reacciones de digestión fueron tratadas con proteinasa K. Para cada uno de ellos, se llevaron a cabo cuatro reacciones de digestión. Primero, se hicieron digestiones con una alícuota de SaCas9 previamente utilizada en otros experimentos *in vitro* (**carriles 1 y 2**). Por otro lado, se llevaron a cabo digestiones con otra alícuota de la SaCas9 nunca utilizada (**carriles 3 y 4**), para controlar la posible pérdida de actividad por la manipulación de la anterior. También se hicieron digestiones utilizando una SaCas9 comercial (**carriles 5 y 6**). Dichas digestiones se realizaron utilizando la mitad de concentración de los componentes del protocolo de incubación, obteniendo un volumen final de 10  $\mu$ L. A partir de estas digestiones se cargaron 250 ng (**carriles 1,3 y 5**) y 500 ng (**carriles 2,4 y 6**) en un gel de poliacrilamida 15%. Paralelamente se hizo una reacción de digestión con la SaCas9 producida ya utilizada en otros ensayos, en un volumen final de 20  $\mu$ L utilizando la concentración original de los componentes (**carril 7**). Para esta última se cargaron 500ng de ADN en el gel. Por otro lado, se incubaron y cargaron en el gel el ADN blanco (**carril 8**) y sgRNA (**carril 9**) individualmente. No se observan bandas correspondientes a los productos de corte esperados en ninguna de las digestiones ensayadas. Tampoco se observan diferencias entre las distintas concentraciones de ADN cargadas en el gel para cada digestión. La banda de tamaño de migración menor al ADN blanco sin cortar en A1 (aprox. 120pb) corresponde al sgRNA intacto, ya que en este caso no se utilizó RNasa A.





**Figura 25.** Ensayo de actividad de la SaCas9 sobre AGO2 utilizando el sgRNA sintetizado y A1 como molde **A)** y A2 **B)**. Se cargaron 250 ng de ADN para las siguientes digestiones: SaCas9 utilizada (1), SaCas9 sin utilizar (3), SaCas9 comercial control (5) y 500 ng de ADN para las mismas digestiones en el mismo orden (2, 4, 6). Se cargó 500 ng de ADN para cada digestión utilizando la totalidad de concentración de componentes y SaCas9 utilizada (7). Se muestra el ADN blanco solo (8) y el sgRNA (9). Poliacrilamida no desnaturalizante 15%. MPM: Perfect DNA 100 bp Ladder (Novagen), con el agregado de un oligonucleótido de 50pb.

Por otro lado, en la **Figura 26** se observa el resultado de los ensayos de actividad de la SaCas9 en AGO2 utilizando el ADN blanco A1. A diferencia del utilizado para las digestiones anteriores, éste fue purificado a través del “(Monarch® PCR & DNA Cleanup Kit [N° de catálogo: T1030S, New England BioLabs (NEB)])” Todas las reacciones fueron tratadas con proteinasa K y ARNasa A. En las digestiones realizadas se utilizaron las siguientes proteínas: SaCas9 producida en este trabajo y utilizada en otros ensayos *in vitro* (**carril 1**), SaCas9 producida en este trabajo nunca utilizada (**carril 2**), SaCas9 comercial (**carril 3**). Para estas digestiones se cargaron 250ng de ADN en el gel y se realizaron utilizando la mitad de concentración de los componentes del protocolo de incubación, obteniendo un volumen final de 10 µL. También se llevaron a cabo las mismas digestiones mencionadas, pero cargando 500ng de ADN en el gel siguiendo el mismo orden que las anteriores (**carriles 4, 5 y 6**). En el gel también se cargaron los siguientes componentes: sgRNA (**carril 7**) y ADN blanco (**carril 8**). Según los resultados, no se observan fragmentos correspondientes a los esperados luego del corte en ninguna de las digestiones.



**Figura 26.** Ensayo de evaluación de la actividad de SaCas9 utilizando A1 purificado por kit como ADN blanco. Se cargaron 250 ng de ADN de las siguiente digestiones : SaCas9 utilizada (1), SaCas9 no utilizada (2) y SaCas9 comercial control (3). Se cargaron 500 ng de las mismas digestiones siguiendo el mismo orden (4,5 y 6). El carril 7 y 8 corresponde al ADN blanco solo y el sgRNA respectivamente. Poliacrilamida 15%. Se utilizó un marcador de peso molecular formado por oligonucleótidos de 50pb 68pb y 108pb (MPM1) y el GeneRuler 1kb DNA ladder, Thermo Scientific (MPM2).

Como se observa en los resultados, la SaCas9 no fue capaz de generar un DSB *in vitro*. Existen diversas razones que podrían explicar estos resultados. Una vez que la proteína Cas9 se asocia con el sgRNA sufre un cambio conformacional provocando la unión del complejo al ADN (Jinek et al., 2014). Cuando la Cas9 encuentra un sitio blanco con el PAM apropiado, desencadena la formación de un bucle R en el que la región guía del sgRNA forma un híbrido ARN: ADN con la cadena complementaria, desplazando la cadena no complementaria (Jiang et al., 2016). Esta estructura de bucle R, que se asemeja a las que se encuentran dentro de la burbuja de transcripción, es estabilizada por la proteína Cas9 mediante la interacción con la hebra no diana y el heteroduplex (Zeng et al., 2018). La formación del bucle R es de vital importancia ya que provoca un reordenamiento que posiciona los dominios catalíticos de la Cas9 en ambas hebras del ADN para ejercer el DSB (Jinek et al., 2014; Anders et al., 2014; Jiang et al., 2016). Como se mencionó anteriormente, el alto contenido GC de la región N-terminal de AGO2 implicó el diseño de un sgRNA que superó ampliamente el máximo porcentaje GC recomendado (40-60%) resultando en un 80%. Estudios basados en análisis termodinámicos de sgRNAs con diferente contenido GC, indican que la accesibilidad estructural de la secuencia guía se correlaciona con la funcionalidad del sgRNA (Wong et al., 2015). En general, la estabilidad estructural del ARN puede aproximarse por el contenido GC de la secuencia. En base a esto, estos estudios evidencian que la mayoría de sgRNAs no funcionales contenían un contenido de nucleótidos GC más alto en promedio en comparación con aquellos sgRNAs funcionales. A su vez, en comparación con las secuencias guías de los sgRNAs no funcionales, aquellos funcionales estaban desprovistos de nucleótidos repetitivos (Wong et al., 2015). Adicionalmente, las secuencias guías con cuatro nucleótidos de guanina consecutivos (GGGG) podrían formar una estructura secundaria especial llamada tétrada de guanina, lo que hace que la secuencia guía sea menos accesible para el reconocimiento de la secuencia de ADN blanco (Wong et al., 2015). En base a esto y, teniendo en cuenta la secuencia guía del sgRNA diseñado para AGO2, es lógico pensar que tal vez se estén formando estructuras adicionales, que pueden interferir en el apareamiento ADN-ARN (Tabla 16). A su vez, esto podría influir en la formación del bucle R contribuyendo a que la SaCas9 no pueda clivar las hebras de ADN. Alternativamente, las secuencias de ADN con alto porcentaje GC, tienden a exhibir interacciones más fuertes, lo que lleva a superestructuras autocomplementarias estables como horquilla y dímero, entre otras (Jensen et al., 2010). Las secuencias que contienen muchas repeticiones de guanina también pueden generar un complicado plegamiento entre cadenas debido a los enlaces de hidrógeno entre guaninas adyacentes. En este sentido, sería lógico pensar que la formación de estas estructuras en el ADN tal vez podría entorpecer el reconocimiento de la región PAM por la proteína Cas9. Además, el hecho de que estas estructuras sean más resistentes a ser desnaturalizadas podría repercutir en la apertura de las hebras por la nucleasa. Por otro lado, dichas estructuras secundarias podrían influir en el reconocimiento de la secuencia PAM por parte de la SaCas9 y también en la hibridación con la secuencia guía del sgRNA en la región blanco. A modo de ejemplo, existen estudios dedicados al

impacto de la formación de estructuras secundarias en el ADN, utilizando como modelo la formación de estructuras G-cuádruplex (GQ) formadas en segmentos de ADN ricos en guanina. Estas investigaciones sugieren que la estabilidad del complejo CRISPR/Cas9 es significativamente menor cuando existe un GQ de alta estabilidad en la hebra complementaria al sgRNA, donde la formación del bucle R debe competir con el plegamiento del GQ (Balci et al.,2021). Estos datos demuestran que existen limitaciones en el tipo y la estabilidad de las estructuras secundarias que CRISPR-Cas9 y el bucle R pueden mantener en un estado desplegado.

Por otro lado, se podría pensar que el ADN blanco formado por los oligonucleótidos de menor tamaño no otorga un contexto adecuado para la SaCas9 y que ésta precise de una molécula de ADN blanco de mayor tamaño. En este sentido, se llevó a cabo una digestión utilizando el escaso producto de 1755pb amplificado por PCR de ADN genómico con el PAR 3 de cebadores. Si bien no se pudo visualizar claramente la migración de las bandas correspondientes al corte esperado, se observó una disminución en la cantidad de ADN blanco en el carril correspondiente a la digestión en comparación con el ADN blanco solo y la aparición de señal de ADN a la altura aproximada de los fragmentos esperados (resultado no mostrado). Este resultado preliminar sugiere que hubo una digestión parcial por la Cas9 usando el sgRNA diseñado. Como se mencionó anteriormente, la falta de reactivo imposibilitó la obtención suficiente de ADN molde de 1755 pb, por lo que el ensayo de digestión no se hizo con las concentraciones adecuadas.

Finalmente, resulta lógico pensar en la funcionalidad de la propia proteína SaCas9. Esta hipótesis se descarta debido a los controles con la SaCas9 comercial. Además, ambas proteínas fueron utilizadas en otros trabajos del laboratorio donde demostraron eficiencias totales y parciales, dependiendo más que nada de los pares sgRNA/blanco específicos. Por otro lado, la idea de utilizar una alícuota de la misma SaCas9 producida recombinante no utilizada previamente radica en descartar posibles alteraciones generadas por condiciones cambiantes del ambiente de la proteína con la manipulación. Una explicación posible a la ausencia de digestión por Cas9, podría ser la presencia de fenol o cloroformo arrastrados de la purificación de ADN blanco. Para descartar esta posibilidad se realizó la purificación del ADN blanco por columna. No se obtuvo cortes en las diferentes condiciones ensayadas, lo que sugiere que la ausencia de actividad de corte es debida al diseño.

## CONCLUSIONES Y PERSPECTIVAS

En este trabajo se diseñaron y sintetizaron los componentes del sistema CRISPR/SaCas9 con la finalidad de generar un corte doble hebra *in vitro* en el gen que sirve como control del procedimiento (mCherry) y una proteína que interesa etiquetar con fines de investigación (AGO2). El clivaje *in vitro* de estos genes por el sistema CRISPR/Cas9 en forma de complejos ribonucleoproteicos representa un método simple y rápido para validar la funcionalidad y eficiencia relativa del sistema previo a su utilización en sistemas celulares. Es especialmente versátil en un laboratorio que trabaja en diferentes organismos modelo, puesto que es indiferente al organismo a editar. Además, tiene la ventaja de no requerir clonado. Por otro lado, en esta tesis se incursionó en el diseño de ADN donantes para la futura edición de estos genes mediante HDR.

Se consiguió producir una Cas9 recombinante soluble con un rendimiento y pureza compatible con la edición *in vitro*. Además, se diseñaron y sintetizaron *in vitro* dos sgRNAs en cantidad y calidad adecuadas para la reacción de edición. Por último, se generaron por PCR de ADN genómico los dos ADN blanco de corte sintéticos para los genes objeto de esta tesis. El ensayo de evaluación de la actividad de la SaCas9 utilizando la secuencia de ADN del gen mCherry como blanco, demostró el funcionamiento de la misma. Esta proteína producida de forma recombinante generó un corte doble hebra de eficiencia parcial en el sitio de interés demostrando la especificidad mediada por el sgRNA diseñado. Este resultado proporciona una plataforma para, en trabajos posteriores, incorporar el sistema en líneas celulares del laboratorio que expresan mCherry. De esta forma sería posible generar un *knock-out* por *knock-in* de este gen fluorescente con el propósito de establecer un control de edición *in vivo* que permita evaluar la eficiencia de los reactivos y métodos que se utilicen.

Los ensayos de incubación de los componentes generados para la edición de AGO2 no generaron un clivaje *in vitro*. Esto podría demostrar la posible limitación del sistema CRISPR/Cas9 respecto a secuencias de ADN complejas, con muchos repetidos y con un alto contenido de bases guanina y citosina. Estas pueden ser propensas a plegarse en estructuras secundarias influenciando así la actividad del complejo CRISPR/Cas9. No obstante, sería conveniente repetir a futuro el mismo ensayo de actividad utilizando una secuencia de ADN a partir de ADN genómico con mayor cantidad de ADN blanco (Burle-Caldas *et al.*, 2018). Por otro lado, también se podrían generar moléculas de ARN guía alternativas cercanas al sitio de edición establecido. La demostración de que este gen puede ser clivado *in vitro* por este sistema proporciona un buen avance para, en trabajos posteriores, introducir el mismo en líneas celulares de CaP en las que trabaja nuestro laboratorio y provocar la edición genómica por

incorporación del ADN molde con neo-epitope etiqueta (*knock-in*), posibilitando las aplicaciones discutidas previamente.

Como perspectiva inmediata se propone repetir el ensayo de edición con AGO2 amplificada de ADN genómico, probar la edición *in vivo* de mCherry y AGO2 con la Cas9 y los guías sintetizados durante esta tesis y evaluar la edición por la pérdida de la fluorescencia seguida de digestión de la región editada y la incorporación del tag en la secuencia endógena.

## REFERENCIAS BIBLIOGRÁFICAS

- Abrahimi, P., Chang, W. G., Kluger, M. S., Qyang, Y., Tellides, G., Saltzman, W. M., & Pober, J. S. (2015). Efficient gene disruption in cultured primary human endothelial cells by CRISPR/Cas9. *Circulation Research*, 117(2), 121–128. <https://doi.org/10.1161/CIRCRESAHA.117.306290>
- Aghdam, A. M., Amiri, A., Salarinia, R., Masoudifar, A., Ghasemi, F., & Mirzaei, H. (2019). MicroRNAs as diagnostic, prognostic, and therapeutic biomarkers in prostate cancer. *Critical Reviews in Eukaryotic Gene Expression*, 29(2), 127–139. <https://doi.org/10.1615/CritRevEukaryotGeneExpr.2019025273>
- Altuvia, Y., Landgraf, P., Lithwick, G., Elefant, N., Pfeffer, S., Aravin, A., Brownstein, M. J., Tuschl, T., & Margalit, H. (2005). Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research*, 33(8), 2697–2706. <https://doi.org/10.1093/nar/gki567>
- Anders, C., Niewoehner, O., Duerst, A., & Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*, 513(7519), 569–573. <https://doi.org/10.1038/nature13579>
- Aqeilan, R. I., Calin, G. A., & Croce, C. M. (2010). MiR-15a and miR-16-1 in cancer: Discovery, function and future perspectives. *Cell Death and Differentiation*, 17(2), 215–220. <https://doi.org/10.1038/cdd.2009.69>
- Barrangou, R., Coûteé-Monvoisin, A. C., Stahl, B., Chavichvily, I., Damange, F., Romero, D. A., Boyaval, P., Fremaux, C., & Horvath, P. (2013). Genomic impact of CRISPR immunization against bacteriophages. *Biochemical Society Transactions*, 41(6), 1383–1391. <https://doi.org/10.1042/BST20130160>
- Barrangou, R., Fremaux, C., Deveau, H., Richardss, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Against Viruses in Prokaryotes. *Science*, 315(March), 1709–1712.
- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004 Jan 23;116(2):281-97. doi: 10.1016/s0092-8674(04)00045-5. PMID: 14744438. Bartel, D. P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, 136(2), 215–233. <https://doi.org/10.1016/j.cell.2009.01.002>
- Bartel, D. P. (2018). Metazoan MicroRNAs. *Cell*, 173(1), 20–51. <https://doi.org/10.1016/j.cell.2018.03.006>
- Bassett, A. R., Tibbit, C., Ponting, C. P., & Liu, J. L. (2014). Highly Efficient Targeted Mutagenesis of *Drosophila* with the CRISPR/Cas9 System. *Cell Reports*, 6(6), 1178–1179. <https://doi.org/10.1016/j.celrep.2014.03.017>
- Bhaya, D., Davison, M., & Barrangou, R. (2011). CRISPR-cas systems in bacteria and archaea: Versatile small RNAs for adaptive defense and regulation. *Annual Review of Genetics*, 45, 273–297. <https://doi.org/10.1146/annurev-genet-110410-132430>

- Bortesi, L., & Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, 33(1), 41–52. 70  
<https://doi.org/10.1016/j.biotechadv.2014.12.006>
- Bubendorf, L., Schöpfer, A., Wagner, U., Sauter, G., Moch, H., Willi, N., Gasser, T. C., & Mihatsch, M. J. (2000). Metastatic patterns of prostate cancer: An autopsy study of 1,589 patients. *Human Pathology*, 31(5), 578–583. <https://doi.org/10.1053/hp.2000.6698>
- Burle-Caldas, G. A., Soares-Simões, M., Lemos-Pechnicki, L., DaRocha, W. D., & Teixeira, S. M. R. (2018). Assessment of two CRISPR-Cas9 genome editing protocols for rapid generation of *Trypanosoma cruzi* gene knockout mutants. *International Journal for Parasitology*, 48(8), 591–596. <https://doi.org/10.1016/j.ijpara.2018.02.002>
- Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., & Croce, C. M. (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24), 15524–15529. <https://doi.org/10.1073/pnas.242606799>
- Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. *Cell*. 2015;163(4):1011-1025. doi:10.1016/j.cell.2015.10.025
- Ceccaldi, R., Rondinelli, B., & D'Andrea, A. D. (2016). Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in Cell Biology*, 26(1), 52–64. <https://doi.org/10.1016/j.tcb.2015.07.009>
- Chandrasekaran, A. P., Song, M., Kim, K. S., & Ramakrishna, S. (2018). Different Methods of Delivering CRISPR/Cas9 Into Cells. In *Progress in Molecular Biology and Translational Science* (1st ed., Vol. 159, Issue June). Elsevier Inc. <https://doi.org/10.1016/bs.pmbts.2018.05.001>
- Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S., & Kim, J. S. (2014). Analysis of offtarget effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Research*, 24(1), 132–141. <https://doi.org/10.1101/gr.162339.113>
- Clark, P. M., Loher, P., Quann, K., Brody, J., Londin, E. R., & Rigoutsos, I. (2014). Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. *Scientific Reports*, 4, 1– 11. <https://doi.org/10.1038/srep05947>
- Concordet, J. P., & Haeussler, M. (2018). CRISPOR: Intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research*, 46(W1), W242–W245. <https://doi.org/10.1093/nar/gky354>
- Costa, P. M., & De Lima, M. C. P. (2013). Micrnas as molecular targets for cancer therapy: On the modulation of micrna expression. *Pharmaceuticals*, 6(10), 1195–1220. <https://doi.org/10.3390/ph6101195>
- Crawford, E. D., Stone, N. N., Yu, E. Y., Koo, P. J., Freedland, S. J., Slovin, S. F., Gomella, L. G., Berger, E. R., Keane, T. E., Sieber, P., Shore, N. D., & Petrylak, D. P. (2014). Challenges and recommendations for early identification of metastatic disease in prostate cancer. *Urology*, 83(3), 664–669. <https://doi.org/10.1016/j.urology.2013.10.026>



- Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J., & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature Biotechnology*, 32(12), 1262–1267. <https://doi.org/10.1038/nbt.302671>
- Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPRCas9. *Science*, 346(6213). <https://doi.org/10.1126/science.1258096>
- Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C., & Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer*, 127(12), 2893–2917. <https://doi.org/10.1002/ijc.25516>
- Flores-Juárez, C. R., González-Jasso, E., Antaramian, A., & Pless, R. C. (2016). PCR amplification of GC-rich DNA regions using the nucleotide analog N4-methyl-2'-deoxycytidine 5'-triphosphate. *BioTechniques*, 61(4), 175–182. <https://doi.org/10.2144/000114457>
- Fort, R. S., Mathó, C., Oliveira-Rizzo, C., Garat, B., Sotelo-Silveira, J. R., & Duhagon, M. A. (2018). An integrated view of the role of miR-130b/301b miRNA cluster in prostate cancer. *Experimental Hematology and Oncology*, 7(1). <https://doi.org/10.1186/s40164-018-0102-0>
- Gandhi, J., Afridi, A., Vatsia, S., Joshi, G., Joshi, G., Kaplan, S. A., Smith, N. L., & Khan, S. A. (2018). The molecular biology of prostate cancer: Current understanding and clinical implications. *Prostate Cancer and Prostatic Diseases*, 21(1), 22–36. <https://doi.org/10.1038/s41391-017-0023-8>
- Gasiunas, G., Barrangou, R., Horvath, P., & Siksnyš, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 109(39), 2579–2586. <https://doi.org/10.1073/pnas.1208507109>
- Gebert, L. F. R., & MacRae, I. J. (2019). Regulation of microRNA function in animals. In *Nature Reviews Molecular Cell Biology* (Vol. 20, Issue 1). <https://doi.org/10.1038/s41580-018-0045-7>
- Green, M. R., & Sambrook, J. (2018). Hot start polymerase chain reaction (PCR). *Cold Spring Harbor Protocols*, 2018(5), 346–349. <https://doi.org/10.1101/pdb.prot095125>
- Hale, C. R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., Olson, S., Resch, A. M., Glover, C. V. C., Graveley, B. R., Terns, R. M., & Terns, M. P. (2012). Essential Features and Rational Design of CRISPR RNAs that Function with the Cas RAMP Module Complex to Cleave RNAs. *Molecular Cell*, 45(3), 292–302. <https://doi.org/10.1016/j.molcel.2011.10.023>
- Hamilton, M. P., Rajapakshe, K. I., Bader, D. A., Cerne, J. Z., Smith, E. A., Coarfa, C., Hartig, S. M., & McGuire, S. E. (2016). The Landscape of microRNA Targeting in Prostate Cancer Defined by AGO-PAR-CLIP. *Neoplasia (United States)*, 18(6), 356–370. <https://doi.org/10.1016/j.neo.2016.04.008>
- Hanahan, D. & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646- 74.
- Hanahan, D. & Coussens, L. M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 21, 309-22.

- Hayes, J., Peruzzi, P. P., & Lawler, S. (2014). MicroRNAs in cancer: Biomarkers, functions and therapy. *Trends in Molecular Medicine*, 20(8), 460–469. <https://doi.org/10.1016/j.molmed.2014.06.005>
- Hryhorowicz, M., Lipiński, D., Zeyland, J., & Słomski, R. (2017). CRISPR/Cas9 Immune System 72 as a Tool for Genome Engineering. *Archivum Immunologiae et Therapiae Experimentalis*, 65(3), 233–240. <https://doi.org/10.1007/s00005-016-0427-5>
- Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G., & Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, 31(9), 827–832. <https://doi.org/10.1038/nbt.2647>
- Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., Ménard, S., Palazzo, J. P., Rosenberg, A., Musiani, P., Volinia, S., Nenci, I., Calin, G. A., Querzoli, P., Negrini, M., & Croce, C. M. (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer Research*, 65(16), 7065–7070. <https://doi.org/10.1158/0008-5472.CAN-05-1783>
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakamura, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isoenzyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, 169(12), 5429–5433. <https://doi.org/10.1128/jb.169.12.5429-5433.1987>
- Jansen, R., Van Embden, J. D. A., Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 43(6), 1565–1575. <https://doi.org/10.1046/j.1365-2958.2002.02839.x>
- Jensen, M. A., Fukushima, M., & Davis, R. W. (2010). DMSO and betaine greatly improve amplification of GC-rich constructs in de novo synthesis. *PLoS ONE*, 5(6), 1–5. <https://doi.org/10.1371/journal.pone.0011024>
- Jiang, F., Taylor, D. W., Chen, J. S., Kornfeld, J. E., Zhou, K., Thompson, A. J., Nogales, E., & Doudna, J. A. (2016). Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science*, 351(6275), 867–871. <https://doi.org/10.1126/science.aad8282>
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA – Guided. 337(August), 816–822.
- Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nature Reviews Genetics*, 16(7), 421–433. <https://doi.org/10.1038/nrg3965>
- Kent, O. A., & Mendell, J. T. (2006). A small piece in the cancer puzzle: MicroRNAs as tumor suppressors and oncogenes. *Oncogene*, 25(46), 6188–6196. <https://doi.org/10.1038/sj.onc.1209913>
- Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J. S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, 24(6), 1012–1019. <https://doi.org/10.1101/gr.171322.113>

- Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li, Y., Gao, N., Wang, L., Lu, X., Zhao, Y., & Liu, M. (2013). Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nature Biotechnology*, 31(8), 681–683. <https://doi.org/10.1038/nbt.2661>
- Lian SL, Li S, Abadal GX, Pauley BA, Fritzler MJ, Chan EK. The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. *RNA*. 2009 May;15(5):804-13. doi: 10.1261/rna.1229409. Epub 2009 Mar 26. PMID: 19324964; PMCID: PMC2673069.
- Liang, X., Potter, J., Kumar, S., Ravinder, N., & Chesnut, J. D. (2017). Enhanced CRISPR/Cas9-73 mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *Journal of Biotechnology*, 241, 136–146. <https://doi.org/10.1016/j.jbiotec.2016.11.011>
- Lin, S., Staahl, B. T., Alla, R. K., & Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *ELife*, 3, e04766. <https://doi.org/10.7554/eLife.04766>
- Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L., & Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, 305(5689), 1437–1441. <https://doi.org/10.1126/science.1102513>
- Liu, X., Homma, A., Sayadi, J., Yang, S., Ohashi, J., & Takumi, T. (2016). Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Scientific Reports*, 6, 1–9. <https://doi.org/10.1038/srep19675>
- Lorenz, T. C. (2012). Polymerase chain reaction: Basic protocol plus troubleshooting and optimization strategies. *Journal of Visualized Experiments*, 63, 1–14. <https://doi.org/10.3791/3998>
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct*, 1, 1–26. <https://doi.org/10.1186/1745-6150-1-7>
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science*. 2013 Feb 15;339(6121):823-6. doi: 10.1126/science.1232033. Epub 2013 Jan 3. PMID: 23287722; PMCID: PMC3712628.
- Mamedov, T. G., Pienaar, E., Whitney, S. E., TerMaat, J. R., Carvill, G., Goliath, R., Subramanian, A., & Viljoen, H. J. (2008). A fundamental study of the PCR amplification of GC-rich DNA templates. *Computational Biology and Chemistry*, 32(6), 452–457. <https://doi.org/10.1016/j.compbiolchem.2008.07.021>
- Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F., & Zhu, J. K. (2013). Application of the CRISPR/Cas system for efficient genome engineering in plants. *Molecular Plant*, 6(6), 2008–2011. <https://doi.org/10.1093/mp/sst121>
- Medeiros, L. C. S., South, L., Peng, D., Bustamante, J. M., Wang, W., Bunkofske, M., Perumal, N., Sanchez-Valdez, F., & Tarleton, R. L. (2017). Rapid, selection-free, high-efficiency genome editing in protozoan parasites using CRISPR-cas9 ribonucleoproteins. *MBio*, 8(6). <https://doi.org/10.1128/mBio.01788-17>

Miyahira, A. K., Lang, J. M., Den, R. B., Garraway, I. P., Lotan, T. L., Ross, A. E., Stoyanova, T., Cho, S. Y., Simons, J. W., Pienta, K. J., & Soule, H. R. (2016). Multidisciplinary intervention of early, lethal metastatic prostate cancer: Report from the 2015 CoffeyHolden Prostate Cancer Academy Meeting. *Prostate*, 76(2), 125–139. <https://doi.org/10.1002/pros.23107>

Mojica FJ, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol*. 2000 Apr;36(1):244-6. doi: 10.1046/j.1365-2958.2000.01838.x. PMID: 10760181.

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, 60(2), 174–182. <https://doi.org/10.1007/s00239-004-0046-374>

Nguyen, D. X., & Massagué, J. (2007). Genetic determinants of cancer metastasis. *Nature Reviews Genetics*, 8(5), 341–352. <https://doi.org/10.1038/nrg2101>

Niaz, S., & Hussain, M. U. (2018). Role of GW182 protein in the cell. *International Journal of Biochemistry and Cell Biology*, 101, 29–38. <https://doi.org/10.1016/j.biocel.2018.05.009>

Nishimasu, H., Cong, L., Yan, W. X., Ran, F. A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., & Nureki, O. (2015). Crystal Structure of *Staphylococcus aureus* Cas9. *Cell*, 162(5), 1113–1126. <https://doi.org/10.1016/j.cell.2015.08.007>

Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., Ishitani, R., Zhang, F., & Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, 156(5), 935–949. <https://doi.org/10.1016/j.cell.2014.02.001>

Oliveto, S., Mancino, M., Manfrini, N., & Biffo, S. (2017). Role of microRNAs in translation regulation and cancer. *World Journal of Biological Chemistry*, 8(1), 45. <https://doi.org/10.4331/wjbc.v8.i1.45>

Orthwein, A., Fradet-Turcotte, A., Noordermeer, S. M., Canny, M. D., Brun, C. M., Strecker, J., Escibano-Díaz, C., & Durocher, D. (2014). Mitosis inhibits DNA double-strand break repair to guard against telomere fusions. *Science*, 344(6180), 189–193. <https://doi.org/10.1126/science.1248024>

Ost, P., Bossi, A., Decaestecker, K., De Meerleer, G., Giannarini, G., Karnes, R. J., Roach, M., & Briganti, A. (2015). Metastasis-directed therapy of regional and distant recurrences after curative treatment of prostate cancer: A systematic review of the literature. *European Urology*, 67(5), 852–863. <https://doi.org/10.1016/j.eururo.2014.09.004>

Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF. Mutation detection using Surveyor nuclease. *Biotechniques*. 2004 Apr;36(4):702-7. doi: 10.2144/04364PF01. PMID: 15088388.

Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., Zetsche, B., Shalem, O., Wu, X., Makarova, K. S., Koonin, E. V., Sharp, P. A., & Zhang, F. (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*, 520(7546), 186–191. <https://doi.org/10.1038/nature14299>

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. (2013). Double nicking by RNA-guided CRISPR

cas9 for enhanced genome editing specificity. *Cell*, 154(6), 1–10.  
<https://doi.org/10.1016/j.cell.2013.08.021>

Reid, G. (2015). MicroRNAs in mesothelioma: From tumour suppressors and biomarkers to therapeutic targets. *Journal of Thoracic Disease*, 7(6), 1031–1040.  
<https://doi.org/10.3978/j.issn.2072-1439.2015.04.56>

Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., & Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology*, 34(3), 339–344.  
<https://doi.org/10.1038/nbt.3481>

Salsman, J., & Dellaire, G. (2017). Precision genome editing in the CRISPR era. *Biochemistry 75 and Cell Biology*, 95(2), 187–201. <https://doi.org/10.1139/bcb-2016-0137>

Schirle, N. T., & MacRae, I. J. (2012). The crystal structure of human argonaute2. *Science*, 336(6084), 1037–1040. <https://doi.org/10.1126/science.1221551>

Shaham, L., Binder, V., Gefen, N., Borkhardt, A., & Izraeli, S. (2012). MiR-125 in normal and malignant hematopoiesis. *Leukemia*, 26(9), 2011–2018. <https://doi.org/10.1038/leu.2012.90>

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., & Skarnes, W. C. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature Methods*, 11(4), 399–402.  
<https://doi.org/10.1038/nmeth.2857>

Sheu-Gruttadauria, J., & MacRae, I. J. (2017). Structural Foundations of RNA Silencing by Argonaute. *Journal of Molecular Biology*, 429(17), 2619–2639.  
<https://doi.org/10.1016/j.jmb.2017.07.018>

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, 507(7490), 62–67.  
<https://doi.org/10.1038/nature13011>

Schirle NT, Sheu-Gruttadauria J, MacRae IJ. Structural basis for microRNA targeting. *Science*. 2014 Oct 31;346(6209):608-13. doi: 10.1126/science.1258040. PMID: 25359968; PMCID: PMC4313529.

Song F, Stieger K. Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. *Mol Ther Nucleic Acids*. 2017 Jun 16;7:53-60. doi: 10.1016/j.omtn.2017.02.006. Epub 2017 Feb 28. PMID: 28624224; PMCID: PMC5363683.

Sun, Y. M., Lin, K. Y., & Chen, Y. Q. (2013). Diverse functions of miR-125 family in different cell contexts. *Journal of Hematology and Oncology*, 6(1), 1–8. <https://doi.org/10.1186/1756-8722-6-6>

Svoronos, A. A., Engelman, D. M., & Slack, F. J. (2016). OncomiR or tumor suppressor? The duplicity of MicroRNAs in cancer. *Cancer Research*, 76(13), 3666–3670.  
<https://doi.org/10.1158/0008-5472.CAN-16-0359>

Terns, M. P., & Terns, R. M. (2011). CRISPR-based adaptive immune systems. *Current Opinion in Microbiology*, 14(3), 321–327. <https://doi.org/10.1016/j.mib.2011.03.005>

- Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, 65(2), 87–108. <https://doi.org/10.3322/caac.21262>
- Van Eijl, R. A. P. M., Van Den Brand, T., Nguyen, L. N., & Mulder, K. W. (2017). Reactivity of human AGO2 monoclonal antibody 11A9 with the SWI/SNF complex: A case study for rigorously defining antibody selectivity. *Scientific Reports*, 7(1), 1–11. <https://doi.org/10.1038/s41598-017-07539-4>
- Vanacore, D., Boccellino, M., Rossetti, S., Cavaliere, C., D’Aniello, C., Di Franco, R., Romano, F. J., Montanari, M., La Mantia, E., Piscitelli, R., Nocerino, F., Cappuccio, F., Grimaldi, G., Izzo, A., Castaldo, L., Pepe, M. F., Malzone, M. G., Iovane, G., Ametrano, G., ... Facchini, G. (2017). Micrnas in prostate cancer: An overview. *Oncotarget*, 8(30), 50240–50251. <https://doi.org/10.18632/oncotarget.1693376>
- Ventura, A., Young, A. G., Winslow, M. M., Lintault, L., Meissner, A., Erkeland, S. J., Newman, J., Bronson, R. T., Crowley, D., Stone, J. R., Jaenisch, R., Sharp, P. A., & Jacks, T. (2008). Targeted Deletion Reveals Essential and Overlapping Functions of the miR-17~92 Family of miRNA Clusters. *Cell*, 132(5), 875–886. <https://doi.org/10.1016/j.cell.2008.02.019>
- Wang, X., Julio, M. K. De, Economides, K. D., Walker, D., Yu, H., Halili, M. V., Hu, Y. P., Price, S. M., Abate-Shen, C., & Shen, M. M. (2009). A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature*, 461(7263), 495–500. <https://doi.org/10.1038/nature08361>
- Wong, N., Liu, W., & Wang, X. (2015). WU-CRISPR: Characteristics of functional guide RNAs for the CRISPR/Cas9 system. *Genome Biology*, 16(1), 1–8. <https://doi.org/10.1186/s13059-015-0784-0>
- Yong, S. L., & Dutta, A. (2009). MicroRNAs in cancer. *Annual Review of Pathology: Mechanisms of Disease*, 4, 199–227. <https://doi.org/10.1146/annurev.pathol.4.110807.092222>
- Zeng, Y., Cui, Y., Zhang, Y., Zhang, Y., Liang, M., Chen, H., Lan, J., Song, G., & Lou, J. (2018). The initiation, propagation and dynamics of CRISPR-SpyCas9 R-loop complex. *Nucleic Acids Research*, 46(1), 350–361. <https://doi.org/10.1093/nar/gkx1117>

## **AGRADECIMIENTOS**

A María Ana y Juan, quienes han tenido una gran dedicación en este trabajo. Gracias por su constante orientación, sugerencias y enseñanzas en cada una de las etapas de esta tesina que, sin dudas, contribuyeron enormemente a mi crecimiento en esta carrera.

A todos los integrantes del LIM, quienes aportaron a este trabajo con valiosos consejos e hicieron que las horas en el laboratorio sean muy amenas y divertidas.

A mis padres quienes me motivaron constantemente para alcanzar mis anhelos. Gracias por su amor incondicional, por brindarme la mejor educación posible y los valores que hoy definen mi vida.

A mi madrina, tíos, abuelos y nonnos por el amor y enseñanzas transmitidas.

A mis primos, en especial a Carolina quien ha sido mi gran compañera lo largo de la vida.

A mi fan número uno, Fernando, por su amor, comprensión y apoyo incondicional.

A mis amigas, quienes hacen que mi vida sea más divertida y bonita.

# ANEXO

## Expresión SaCas9

### Medio de cultivo 2xYT

- 16 g Triptona
- 10 g Extracto de levadura
- 5 g Cloruro de sodio (NaCl)
- Agua destilada csp 1L

## Purificación SaCas9 (*Burle-Caldas et al., 2018*):

### Buffers fosfato base

- Solución A: 27,6g fosfato de sodio monohidratado ( $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ ) en 1L de agua destilada.
- Solución B: 53,65g fosfato sódico hepta-hidratado ( $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ ) en 1L de agua destilada.

### Buffer fosfato 0,1M (pH 7,4)

- 19mL solución A
- 81mL solución B
- Llevar hasta 200mL con agua milli-Q

### Buffer fosfato 20mM \*

- 200mL de Buffer fosfato
- 29,22g cloruro de sodio (NaCl)
- Llevar a 1L con agua milli-Q)

### Buffer de unión (binding)

- 500mL de \*
- 30mM imidazol (1,02g)
- Ajustar pH a 7,5
- Filtrar por 0.22 $\mu\text{m}$

### Buffer de elución

- 500ml de \*
- 500mM imidazol (17,0g)
- Llevar a 100mL con agua milli-Q
- Ajustar pH a 7,5



- Filtrar por 0.22µm

## SDS-PAGE

### Buffer de carga (4X)

- 25mL Tris-HCl 0,5M pH 6.8 SDS 0,4%
- 4g SDS
- 20 mL glicerol
- 2mL β-mercaptoetanol
- 1mg azul de bromofeno
- Agua destilada csp 50 mL

### Buffer de corrida (5X)

- 15,15g Tris-base
- 72g Glicina
- 5g SDS
- Agua destilada csp 1L

### Tinción Azul de Coomassie

- 80 mL ácido acético glacial
- 250 mL etanol
- 1g de azul de Coomassie
- Agua destilada csp 1L

### Solución de desteñir

- 80 mL ácido acético glacial
- 250 mL etanol
- Agua destilada csp 1L

## PAGE-NATIVO ADN

### Buffer de corrida TBE (5X)

- 54 g Tris-base
- 27,5 g Ácido bórico
- 20 mL EDTA 0,5M pH 8,0
- Agua destilada csp 1L

### Buffer de carga ADN (6X)

- 25 mg Azul de bromofenol
- 25 mg Xileno cianol FF
- 4g Sacarosa

- Buffer TAE 1X csp 10 mL

## Electroforesis agarosa (ARN)

### **Buffer de corrida MOPS(10X)**

- 41,8g MOPS (sigma)
- Ajustar pH a 7 con NaOH concentrado o ácido acético.
- 16,6mL NaAc 3M.
- 20mL EDTA 0,5M pH 8.
- H<sub>2</sub>O RNAsa-free csp 1L

### **Buffer de carga (3X)**

- 750 µL formamida
- 150 µL buffer de corrida 10x (MOPS)
- 240 µL formaldehido
- 100 µL H<sub>2</sub>O RNAsa-free
- 10 µL glicerol
- 80 µL azul de bromofenol 10%

## Electroforesis agarosa (ADN)

### **Buffer de corrida TAE(10X)**

- 48,4 g Tris-base
- 11,4 mL ácido acético glacial
- 3,7 g EDTA
- Agua destilada csp 1L

### **Buffer de carga (6X)**

- 25 mg Azul de bromofenol
- 25 mg Xilene cianol FF
- 4g Sacarosa
- Buffer TAE 1X csp 10 mL

## Evaluación SaCas9

### **Buffer para SaCas9**

- 2 µL HEPES 200mM
- 2µL KCl 1,5M
- 2µL MgCl<sub>2</sub> 100mM
- 2µL DTT 5Mm