



## **MÁSTER UNIVERSITARIO EN I+D EN BIOTECNOLOGÍA Y BIOMEDICINA**



VNIVERSITAT  
E VALÈNCIA

### **TRABAJO DE FIN DE MÁSTER**

## **Expresión recombinante de una RNA polimerasa bacteriana**

**AUTOR:**

**DAVID DALMAU MERCÉ**

**TUTORES:**

**MARIA LUISA GIL HERRERO**

**JOSÉ LUIS LLÁCER GUERRI**

**SEPTIEMBRE, 2020**



## **MÁSTER UNIVERSITARIO EN I+D EN BIOTECNOLOGÍA Y BIOMEDICINA**

### **TRABAJO DE FIN DE MÁSTER**

# **Expresión recombinante de una RNA polimerasa bacteriana**

**AUTOR:**

**DAVID DALMAU MERCÉ**

**TUTORES:**

**MARIA LUISA GIL HERRERO**

**JOSÉ LUIS LLÁCER GUERRI**

---



## **ABREVIATURAS**

2-OG: 2-oxoglutarato

$\alpha$ -CTD: Dominio C-terminal de la subunidad alfa de la RNA polimerasa

$\alpha$ -NTD: Dominio N-terminal de la subunidad alfa de la RNA polimerasa

ADP: Adenosín difosfato

ATP: Adenosín trifosfato

BSA: Albúmina sérica bovina

cAMP: Adenosín monofosfato cíclico

CRP/FNR: cAMP receptor protein-fumarate/nitrate reductase regulator (grupo de factores de transcripción)

DMSO: Dimetilsulfóxido

DTT: 1,4-ditio-D-treitol

EDTA: Ácido etilendiaminetetraacético

FPLC: Modalidad de cromatografía líquida para proteínas, de alta velocidad

GOGAT: Glutamato sintasa, glutamina: 2-Oxoglutarato aminotransferasa

GS: Glutamina sintetasa

GST: Glutación-S-transferasa

HTH: Motivo hélice-giro-hélice (Helix-Turn-Helix) de unión a DNA

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

LB: Luria-Bertani

PCR: Reacción en cadena de la polimerasa (Polymerase Chain Reaction)

PDB: Banco de datos de proteínas (Protein Data Bank)

PMSF: Fluoruro de fenilmetilosulfonilo

RNAP: RNA polimerasa

RPc: Conformación cerrada RNAP-promotor

RPo: Conformación abierta RNAP-promotor

RT: Temperatura ambiente (Room temperatura)

SDS-PAGE: Electroforesis en gel de poliacrilamida en condiciones desnaturalizantes

SDM: Mutagénesis de sitio dirigido (Site-directed mutagénesis)

TLD: Dominio de tipo Tudor (Tudor-Like Domain)

## ÍNDICE GENERAL

RESUMEN.....	I
ABSTRACT .....	II
<b>1. INTRODUCCIÓN .....</b>	<b>1</b>
1.1. Control de la asimilación de nitrógeno.....	1
1.2. Las cianobacterias .....	2
1.3. La regulación del metabolismo del nitrógeno en cianobacterias mediada por el factor de transcripción NtcA y la familia CRP/FNR .....	2
1.4. El papel de PipX en la regulación del metabolismo del nitrógeno .....	4
1.5. El sistema PII-PipX-NtcA.....	5
1.6. El complejo RNAP-NtcA-PipX-DNA .....	6
1.7. Plan de trabajo .....	7
1.8. Estado inicial del proyecto .....	8
<b>2. OBJETIVOS .....</b>	<b>9</b>
<b>3. MATERIALES Y MÉTODOS .....</b>	<b>10</b>
3.1. Diseño de la estrategia de clonación a seguir .....	10
3.2. Transformación en <i>E. coli</i> .....	10
3.3. Comprobación de plásmidos con el/los inserto/s deseados mediante PCR de colonias y electroforesis en gel de agarosa .....	11
3.4. Electroforesis en gel de agarosa .....	11
3.5. Purificación de los plásmidos con los constructos de interés y cuantificación .....	12
3.6. Purificación de fragmentos lineales de DNA y cuantificación .....	13
3.7. Comprobación de los insertos mediante secuenciación.....	13
3.8. Eliminación de una mutación en el gen <i>rpoC2</i> .....	14
3.8.1. Mutagénesis de sitio dirigido (SDM).....	14
3.8.2. Clonación del gen a partir del plásmido original con <i>rpoC2</i> .....	16
3.8.3. Eliminación de la mutación de <i>rpoC2</i> mediante su sustitución con un gen artificial .....	17
3.9. Clonación en tándem de los distintos genes en los plásmidos pQLink .....	18
3.9.1. Diseño de la estrategia de clonación a seguir y fundamento general de la clonación de genes en tándem en vectores pQLink.....	18
3.9.2. Construcción de plásmidos de co-expresión de las tres variantes RNAP en pQLink .....	20
3.10. Expresión y purificación de las proteínas NtcA y PipX .....	21
3.10.1. Expresión de NtcA y PipX .....	21
3.10.2. Purificación de NtcA y PipX mediante cromatografía de afinidad .....	23
<b>4. RESULTADOS Y DISCUSIÓN .....</b>	<b>25</b>
4.1. Eliminación de una mutación en el gen <i>rpoC2</i> .....	25

4.1.1.	Mutagénesis de sitio dirigido (SDM) .....	25
4.1.2.	Clonación del gen a partir del plásmido original con rpoC2.....	26
4.1.3.	Eliminación de la mutación de rpoC2 mediante su sustitución con un gen artificial ....	27
4.2.	Clonación en tándem de los distintos genes en los plásmidos pQLink.....	28
4.3.	Expresión y purificación de las proteínas NtcA y PipX .....	32
4.3.1.	Expresión de NtcA y PipX.....	32
4.3.2.	Purificación de NtcA mediante cromatografía de afinidad .....	33
4.4.	Perspectivas de futuro .....	35
4.4.1.	Co-expresión y purificación de una RNAP cianobacteriana .....	35
4.4.2.	Estrategia a seguir para el ensamblaje del complejo RNAP-NtcA-PipX-DNA para su estudio mediante Cryo-EM.....	36
5.	CONCLUSIÓN .....	38
6.	BIBLIOGRAFÍA.....	39
7.	ANEXO .....	43
7.1.	Anexo I.....	43
7.2.	Anexo II .....	48
7.3.	Anexo III .....	49

Figura 1. Esquema de la activación de la transcripción en promotores bacterianos de clase II. ....	3
Figura 2. Estructura del complejo NtcA-DNA con 2OG. ....	4
Figura 3. Complejo PipX-NtcA. ....	5
Figura 4. Esquema de los pasos requeridos para la fusión de los genes en el vector de co-expresión pQLink. ....	8
Figura 5. Estrategia de mutagénesis del gen rpoC2. ....	14
Figura 6. Sistema Link de clonación ....	19
Figura 7. Gel de agarosa 0,8 % del producto de PCR de la mutagénesis de rpoC2 ( $\beta'$ ). ....	25
Figura 8. Gel de agarosa 0.8 % de la PCR de rpoC2 ( $\beta'$ ) I. ....	26
Figura 9. Gel de agarosa 0.8 % de la PCR de rpoC2 ( $\beta'$ ) II. ....	26
Figura 10. Gel de agarosa 0.8 % de la digestión con Bbvcl y NgoMIV de pQLink N-rpoC2 ( $\beta'$ ) tras 1 hora de electroforesis. ....	27
Figura 11. Gel de agarosa 0.8 % de la digestión con Bbvcl y NgoMIV de pQLink N-rpoC2 ( $\beta'$ ) tras 30 minutos de electroforesis. ....	27
Figura 12. Gel de agarosa 0,8 % de la digestión con Bbvcl y NgoMIV de pUC19 con el gen artificial de 1389 pb. ....	28
Figura 13. Gel de agarosa 0.8 % de la PCR de colonia del constructo pQLink N + $\omega$ + $\gamma$ . ....	29
Figura 14. Comprobación del constructo rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ) mediante digestión con PaeI. ....	29
Figura 15. . Comprobación de la ligación del constructo rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ )- rpoD1 ( $\sigma_A$ ) ....	29
Figura 16. Gel de agarosa 0.8 % de la PCR de colonia del constructo pQLinkN + $\sigma_A$ + $\omega$ + $\gamma$ . ....	30
Figura 17. Gel de agarosa 0.8 % de la digestión pQLinkN + $\omega$ + $\gamma$ + $\sigma_A$ y PCR de colonia de los constructos pQLinkN + $\omega$ + $\gamma$ + $\sigma_B$ y pQLinkN + $\omega$ + $\gamma$ + $\sigma_C$ ....	31
Figura 18. Gel de agarosa 0.8 % de la PCR de colonia del constructo pQLink N + $\omega$ + $\gamma$ + $\sigma_B$ ....	31
Figura 19. Gel de poliacrilamida 10 % de la expresión de las proteínas NtcA y PipX antes y después de inducir la expresión con IPTG. ....	32
Figura 20. Gráfica de A280 de las diferentes fracciones obtenidas a diferentes concentraciones de tampón B. ....	33
Figura 21. Gel de poliacrilamida 10 % de la purificación de NtcA. ....	34
Figura 22. Secuencia del operador del gen glnA, glnB y glnN para su ensamblaje con NtcA y la RNAP con $\sigma_A$ , $\sigma_B$ y $\sigma_C$ . ....	37

Tabla 1. Reactivos utilizados en la PCR con NZYtaq II DNA polimerasa.....	11
Tabla 2. Ciclo térmico utilizado en la PCR con NZYtaq II DNA polimerasa .....	11
Tabla 3. Reactivos utilizados en la PCR con NZYSupreme DNA polimerasa .....	15
Tabla 4. Ciclo térmico en la PCR con NZYSupreme DNA polimerasa.....	15
Tabla 5. Tratamiento con mix KLD.....	16
Tabla 6. Condiciones para la ligación con T4 DNA ligasa.....	18
Tabla 7. Relación de las cantidades de plásmido en la clonación en tándem.....	21
Tabla 8. Componentes del gel de poliacrilamida .....	22
Tabla 9. Volumen de muestra utilizado para la cuantificación, absorbancia y concentración de proteína obtenida mediante el método de Bradford.....	33



**RESUMEN**

En cianobacterias, la regulación de la transcripción de genes relacionados con el metabolismo del nitrógeno está controlada, generalmente, por el factor de transcripción NtcA. Este factor de transcripción, perteneciente a la familia CRP/FNR, interacciona con una secuencia consenso en el promotor de los genes a los que regula a fin de reclutar a la RNA polimerasa (RNAP) al lugar de inicio de la transcripción y activarla. El modo de interacción y el mecanismo por el cual activa a la RNAP es desconocido. En condiciones de pobreza de nitrógeno, NtcA es co-activada por la proteína PipX. Se conoce que dicha co-activación se debe a la estabilización por parte de PipX de la forma activa del factor NtcA, pero no se sabe si PipX también interacciona con la RNAP y/o juega un papel activándola. El objetivo global del grupo de acogida, experto en determinación de estructuras macromoleculares, es estudiar mediante criomicroscopía electrónica (cryo-EM) la activación de la transcripción por los reguladores globales del metabolismo del nitrógeno NtcA y PipX obteniendo para ello la estructura de estas dos proteínas unidas simultáneamente a la RNAP y al promotor de un gen. Dentro de este plan se enmarca el proyecto de final de máster, cuyo objetivo principal será la obtención de tres variantes de la RNAP de la cianobacteria *Synechococcus elongatus* de forma recombinante en *Escherichia coli* utilizando vectores policistrónicos, compatibles para la co-expresión, que portan los genes que codifican las distintas subunidades de la RNAP y su posterior purificación. Además se expresarán y purificarán las proteínas NtcA y PipX.

**PALABRAS CLAVE**

Cianobacteria; Transcripción; RNAP; NtcA; PipX; Metabolismo del nitrógeno.

## ABSTRACT

In cyanobacteria, transcriptional regulation of nitrogen metabolism-related genes is mainly mediated by the transcription factor NtcA. This transcription factor, which belongs to the CRP/FNR family, interacts with a consensus sequence in the promoter of the genes that regulates in order to recruit RNA polymerase (RNAP) to the transcription start site and activate transcription. The interaction mode and the mechanism by which NtcA activates RNAP is unknown. In very poor nitrogen conditions, NtcA is co-activated by the PipX protein. We know that this co-activation is partly due to PipX stabilization of the active form of NtcA, but we do not know if PipX also interacts with the RNAP and/or plays a role in its activation. The overall aim of the host group, specialized in determination of macromolecular structures is to study the transcription activation by the global nitrogen metabolism regulators NtcA and PipX by electron cryomicroscopy (cryo-EM), obtaining for this purpose the structure of these two proteins bound simultaneously to the RNAP and to a gene promoter. The current project is part of this aim, being its main objective to obtain three variants of a recombinant cyanobacteria *Synechococcus elongatus* RNAP in *Escherichia coli*, using polycistronic vectors, suitable for co-expression, that carry the genes that encode the different subunits of RNAP and subsequent purification. Furthermore, we have expressed and purified the proteins NtcA and PipX.

## KEY WORDS

Cyanobacteria; Transcription; RNAP; NtcA; PipX; Nitrogen metabolism.

## 1. INTRODUCCIÓN

### 1.1. Control de la asimilación de nitrógeno

Las bacterias son los organismos más abundantes del planeta principalmente debido a su capacidad de adaptación a cualquier ecosistema. En este sentido, son ampliamente conocidas por su gran versatilidad para metabolizar nutrientes en hábitats terrestres o acuáticos y hasta en los más extremos como manantiales de agua caliente o las profundidades marinas. Para ello, debe existir un control metabólico que coordine el catabolismo y la asimilación de las fuentes de carbono y nitrógeno.

Precisamente, el control del nitrógeno es una facultad compartida por gran cantidad de microorganismos. El nitrógeno es uno de los elementos más importantes para el desarrollo de la vida, pues es necesario para la producción de aminoácidos, nucleótidos y aminoazúcares entre otros. En consecuencia, las bacterias han desarrollado varios mecanismos para la asimilación de este importante bioelemento a partir de diferentes fuentes. Iones inorgánicos como el nitrato y el amonio o compuestos orgánicos como los aminoácidos o la urea son ejemplos de compuestos que pueden ser utilizados por distintos organismos como fuentes de nitrógeno, pero no todos son igual de fácilmente asimilables. Dicho control consiste en la represión de las rutas de asimilación de ciertas fuentes de nitrógeno, mientras otras, asimilables de manera más sencilla, son favorecidas (Arcondeguy et al., 2001; Herrero et al., 2001; Leigh y Dodsworth, 2007). En la mayoría de las ocasiones, el amonio es la fuente preferida de nitrógeno, pues es fácilmente asimilable al ser el sustrato energéticamente menos costoso de procesar, por contra, las fuentes inorgánicas requieren ser reducidas previamente, mientras que las fuentes orgánicas necesitan ser degradadas (Arcondeguy et al., 2001). Existen dos rutas principales para la asimilación de amonio: el ciclo de la glutamina sintetasa (GS) y glutamato sintasa (GOGAT), presente en bacterias, y la ruta glutamato deshidrogenasa, energéticamente más eficiente pero con una baja afinidad por el amonio. Por tanto, destaca el sistema GS/GOGAT que cataliza la incorporación del amonio en una molécula de glutamato formando glutamina, que luego junto con 2-oxoglutarato (2OG) produce glutamato gracias a la GOGAT, que provee dos moléculas de glutamato para sintetizar otros aminoácidos (Muro-Pastor et al., 2005).

De esta forma, niveles altos de glutamina y bajos de 2OG significan concentraciones de nitrógeno elevadas, en cambio, una baja cantidad de glutamina y una alta abundancia de 2OG revelan niveles bajos de nitrógeno en la célula. Si a esto añadimos que 2OG es la conexión entre el metabolismo del nitrógeno, por proporcionar el esqueleto carbonado para la asimilación de amonio y el del carbono, por ser un intermediario del ciclo de Krebs, el 2OG se convierte en un compuesto de gran interés. De hecho, esta molécula cobra aún mayor relevancia en cianobacterias, como indicador clave del balance celular N/C, ya que poseen un ciclo de Krebs incompleto debido a la carencia de 2-oxoglutarato deshidrogenasa. Como consecuencia se acumula 2-oxoglutarato y los cambios en las reacciones de asimilación de carbono y nitrógeno se ven reflejados directamente en la

cantidad de dicho metabolito. Esta influencia implica una condición metabólica que puede tener efectos reguladores (Stanier y Cohen-Bazire, 1977; Steinhauser et al., 2012).

## **1.2. Las cianobacterias**

Las cianobacterias son procariotas gramnegativas pertenecientes al dominio bacteria. Su principal característica es la capacidad para realizar fotosíntesis oxigénica utilizando el agua como agente reductor. La forma fundamental de crecimiento de estos organismos es la fotoautotrofia, que fija el CO<sub>2</sub> a través del ciclo de Calvin. Constituyen una de las formas de vida más antiguas sobre la Tierra, estimándose que aparecieron hace 3000 millones de años. Desde el punto de vista evolutivo, se consideran muy importantes por la transformación de la atmósfera anóxica de la Tierra en una atmósfera aerobia. Asimismo, son antecesoras de las plantas y algas, pues todo indica que fueron el origen de los plastos de éstas por endosimbiosis (Hohmann-Marriott y Blankenship, 2011).

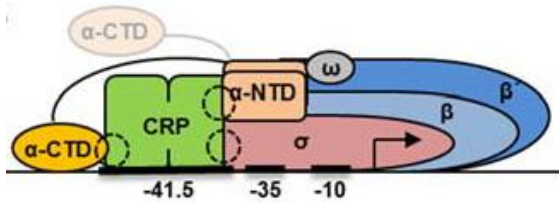
Las cianobacterias pueden utilizar como fuente de nitrógeno amonio, nitrato y nitrito, y en ocasiones, según la especie, también aminoácidos neutros, cianato, urea o dinitrógeno. Además, presentan una notable diversidad morfológica, y se pueden encontrar en forma unicelular o filamentosa, y en algunos casos, presentan potencial para la diferenciación celular, como la especialización de heterocistos fijadores de N<sub>2</sub> (Beck et al., 2012).

En particular, la cianobacteria *Synechococcus elongatus* PCC 7942, previamente conocida como *Anacystis nidulans* R2 fue la primera cianobacteria en la que se demostró la transformación (Shestakov y Khyen, 1970). Desde entonces, ha sido objeto de estudio en investigaciones sobre la adquisición de carbono inorgánico (Omata et al., 2001), la regulación de nitrógeno (Vázquez-Bermúdez et al., 2002; Lee et al., 2000) o la aclimatación a estreses (Sauer et al., 2001). Hoy en día, sigue siendo un sistema modelo de referencia de estos organismos empleado en análisis de los mecanismos de los ritmos circadianos en procariotas (Leypunskiy et al., 2017). *S. elongatus* es unicelular y fotótrofo obligado de agua dulce no fijador de nitrógeno (Rippka et al., 1979). En definitiva, se presenta como un organismo modelo apropiado para el trabajo por tener su genoma secuenciado y ser transformable de forma fiable.

## **1.3. La regulación del metabolismo del nitrógeno en cianobacterias mediada por el factor de transcripción NtcA y la familia CRP/FNR**

En estos organismos, el control del nitrógeno está mediado por NtcA, un factor de transcripción universal para cianobacterias, pues se han detectado secuencias homólogas en numerosas especies. Es conocido como el regulador global del nitrógeno pues controla la expresión de cientos de genes relacionados con su metabolismo como *nper* (asimilación de nitrato), *urt* (transporte de urea) o *amt* y *glnA* (asimilación de amonio), aunque también controla genes del metabolismo de azúcares, fotosintéticos o de defensa antioxidante (Herrero et al., 2001). Consiste en una proteína dimérica de 222 aminoácidos que se une a la secuencia consenso GTAN<sub>8</sub>TAC de los promotores de los genes que regula (Picossi et al., 2014) o al menos a la secuencia esencial GTN<sub>10</sub>AC (Herrero et al., 2004). En la mayoría de

las ocasiones, NtcA actúa como activador, casi siempre uniéndose a promotores cuyo sitio consenso está localizado a -41.5 nucleótidos del lugar de inicio de la transcripción como ocurre con los promotores bacterianos de clase II (Figura 1). Por el contrario, actúa como represor en unos pocos casos impidiendo la unión de la RNA polimerasa al promotor debido a que su sitio de unión se sitúa muy cerca del lugar de inicio de la transcripción.

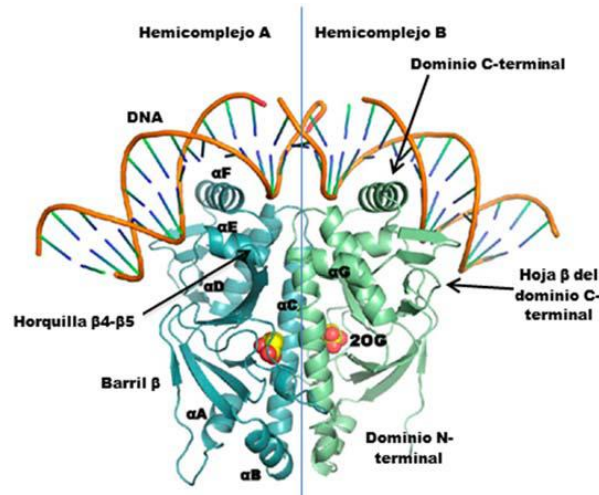


**Figura 1. Esquema de la activación de la transcripción en promotores bacterianos de clase II.** Los contactos entre CRP y la RNA polimerasa (subunidades  $\alpha$ -CTD,  $\alpha$ -NTD y  $\sigma$ ) representados como círculos con borde discontinuo. El DNA representado de forma lineal señalando las secuencias de reconocimiento de CRP a -41.5 y los elementos -35 y -10 (Forcada-Nadal, 2018).

La transcripción dependiente de NtcA se activa en presencia de fuentes pobres en nitrógeno. Activa su propia expresión, expresándose a niveles basales en presencia de amonio y aumentando su transcripción en su ausencia (Mitschke et al., 2011; Browning y Busby, 2004). Se trata de un factor de transcripción perteneciente a la familia CRP/FNR (cAMP receptor protein-fumarate/nitrate reductase regulator). Estos factores de transcripción se encuentran ampliamente distribuidos en procariontes y engloban a cientos de miembros. Este grupo está constituido por proteínas de unión a DNA que regulan un elevado número de genes e incluso a otros factores de transcripción, principalmente, activándolos. Los factores de esta familia se caracterizan por tener un dominio sensor N-terminal de unión a ligando y un dominio C-terminal de unión a DNA. También pueden responder a un amplio espectro de señales ambientales como pueden ser el cAMP, 2OG, anoxia, estado redox, temperatura, óxido nítrico, etc. (Körner et al., 2003; Browning y Busby, 2004).

Respecto a NtcA, cada subunidad está compuesta por un dominio C-terminal que contiene un motivo hélice-giro-hélice (HTH) que participa en la unión al promotor y por un dominio N-terminal que es el dominio regulador o de unión al efector que se une al 2OG, mostrando la estructura característica de los reguladores transcripcionales pertenecientes a la familia CRP-FNR (Figura 2). La molécula 2OG implicada en la señalización facilita la adopción de la conformación activa del factor llevando a cabo la unión al DNA y al parecer es indispensable para la activación de la transcripción (Llácer, 2012; Valladares et al., 2010). NtcA en ausencia de 2OG es capaz de unirse al DNA a su sitio en el promotor, pero no puede activar la transcripción (Llácer, 2012). Tras la unión del 2OG al dominio sensor de cada subunidad, NtcA sufre una transición desde un estado inactivo con baja afinidad para unir DNA, a un estado activo. Por similitud con CRP, NtcA en presencia de 2OG podría activar la transcripción a dos niveles, el primero de ellos reclutando con mayor eficiencia a la RNA polimerasa (RNAP) por interacción directa de la subunidad de NtcA más proximal con el dominio C-terminal de la subunidad  $\alpha$  de la RNAP (Figura 1), ya que, aunque la RNAP es capaz de unirse al promotor en ausencia del factor de transcripción, lo hace con muy baja afinidad. El segundo, posteriormente, favoreciendo la tasa de isomerización de complejo cerrado a abierto de RNAP, posibilitando así la formación de la burbuja de transcripción al

separar ambas hebras del DNA probablemente gracias a los contactos de la subunidad más distal de NtcA con las subunidades  $\sigma$  y el dominio N-terminal de la subunidad  $\alpha$  de la RNAP (Figura 1). De esta forma aumenta a su vez también la afinidad de NtcA por el DNA (Browning y Busby, 2004; Forcada-Nadal et al., 2018).

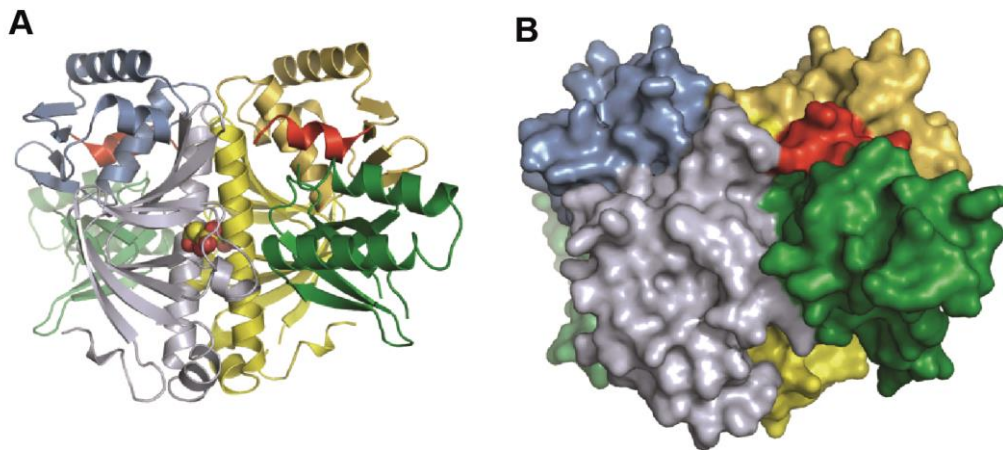


**Figura 2. Estructura del complejo NtcA-DNA con 2OG.** 2OG representado en esferas, una subunidad de NtcA coloreada en azul y la otra en verde. Se señalan algunos elementos de estructura secundaria (Forcada-Nadal, 2018).

#### 1.4. El papel de PipX en la regulación del metabolismo del nitrógeno

La unión de 2OG por NtcA favorece además la unión de otra proteína llamada PipX a NtcA. La consecuencia de dicha interacción es una activación de la transcripción génica, por lo que PipX actúa como co-activador en condiciones de bajo nitrógeno. PipX es una proteína monomérica de 89 aminoácidos exclusiva de cianobacterias. La expresión de PipX está controlada, al menos en *Anabaena* sp., por NtcA como la mayoría de genes relacionados con el metabolismo del nitrógeno. A su vez la disponibilidad de PipX para interactuar con NtcA está controlada por la proteína PII. PipX se une a PII cuando hay bajos niveles de 2OG dejando los dominios de PipX que interactúan con NtcA ocultos impidiendo su unión con el factor de transcripción (Llácer et al., 2010). PII es una pequeña proteína de transducción de señal y es indudablemente una de las proteínas señaladoras más conservadas en el dominio bacteria, donde desarrolla una importante función en la coordinación del metabolismo del nitrógeno.

Para que el 2OG se una a la proteína es estrictamente necesario que también esté unida una molécula de ATP. En ausencia de ATP, PII se une a ADP o AMP, lo que imposibilita la unión del 2OG. Esta competición entre el ATP y el ADP/AMP, permite a PII controlar el estado energético celular (Zeth et al., 2014). Puede interactuar con múltiples dianas, entre las cuales destaca PipX. En la estructura cristalina de PipX se observa que está formada por un dominio de plegamiento  $\beta$  seguido de dos hélices  $\alpha$  C-terminales (Figura 3). En el dominio  $\beta$  se observa el plegamiento de un dominio de tipo tudor (TLD) que participa en la interacción de la proteína con PII, NtcA y otras moléculas de PipX (Llácer et al., 2010; Llácer, 2012). En cianobacterias, la expresión de la proteína PII también depende del factor de transcripción NtcA (Espinosa et al., 2006).



**Figura 3. Complejo PipX-NtcA.** Representación en cintas (A) o en superficie (B). PipX en verde. NtcA con una subunidad plateada y otra amarilla (dominios de unión a DNA coloreados en tonos más oscuros). La hélice C-terminal, se destaca en rojo por ser el único elemento no presente en CRP. 2OG se representa en esferas (Llácer et al. 2012).

### 1.5. El sistema PII-PipX-NtcA

Según los niveles de 2OG, PII forma un complejo con PipX cuya función es el secuestro y la liberación controlada de PipX en función de la disponibilidad de nitrógeno asimilable. De hecho, PipX fue inicialmente identificada como un receptor de PII en *S. elongatus* (Burillo et al., 2004). PII interacciona y secuestra a PipX en situaciones de abundancia de nitrógeno (bajos niveles de 2OG) y de esta manera se evita la transcripción innecesaria de genes dependientes de NtcA, mientras que, en condiciones de extrema pobreza de nitrógeno PipX es liberada y se induce la transcripción incrementada de dichos genes dependientes de NtcA. Cuando la disponibilidad de amonio disminuye mucho y ocurre la liberación de PipX, a juzgar por las concentraciones de 2OG, que resultan ser más elevadas que las concentraciones de la misma molécula necesarias para promover la unión de NtcA al DNA, la interacción de PipX con NtcA debe considerarse un estadio de hiperactivación de NtcA (Vazquez-Bermudez et al., 2002; Espinosa et al., 2006).

Gracias a la determinación de las estructuras cristalinas de NtcA en su forma activa y/o unida a DNA y del complejo NtcA-PipX (Figura 2, Figura 3) se sabe cómo NtcA reconoce su secuencia específica en el DNA, cómo se dispone en su forma activa gracias a 2OG y por qué su interacción con PipX implica una hiperactivación de la expresión de los genes que regula debido a la estabilización por PipX de la conformación activa de NtcA y también por la posible ayuda de PipX reclutando a la RNA polimerasa. Asimismo, se ha comprobado que NtcA produce una curvatura en el DNA que sería indispensable para su ulterior interacción con la RNAP y que PipX no interacciona con el DNA. Sin embargo, dichas estructuras no han logrado explicar a nivel molecular cómo la RNAP es activada por NtcA ni si PipX interacciona directamente con la RNAP (Llácer et al., 2010; Zhao et al., 2010).

### 1.6. El complejo RNAP-NtcA-PipX-DNA

En este contexto, el objetivo principal del proyecto que engloba este trabajo es la obtención de la estructura RNAP-NtcA-PipX-DNA mediante criomicroscopía electrónica (cryoEM), con la intención de entender la activación de la transcripción a nivel molecular/estructural mediada por el complejo NtcA-PipX, utilizando para ello factores sigma alternativos. Este estudio puede ser además, de gran ayuda para entender los mecanismos estructurales de la activación de la RNAP por factores de transcripción bacterianos. Por consiguiente, se requiere la obtención de las proteínas NtcA y PipX, que se pueden adquirir siguiendo el protocolo de Llácer et al. 2010, y la producción de una RNAP cianobacteriana, para lo cual se realizará la clonación en vectores de expresión bacterianos de las distintas subunidades de la RNAP de la cianobacteria *Synechococcus elongatus*, objetivo del presente trabajo.

El núcleo de la RNAP, capaz de emplear ribonucleótidos para la síntesis de RNA a partir de una secuencia de DNA molde, está formada por varias subunidades. Para el inicio de la transcripción se debe incorporar a este núcleo de la RNAP una subunidad  $\sigma$  que aumenta la afinidad para unirse a regiones promotoras específicas. De esta manera se forma el holoenzima que al asociarse a proteínas accesorias dará lugar al complejo de transcripción listo para llevar a cabo la síntesis del RNA.

Esta enzima cianobacteriana posee un core relativamente similar al de otras bacterias compuesto por cuatro subunidades,  $\alpha$  (2 subunidades),  $\beta$ ,  $\beta'$ , y  $\omega$  (Imamura y Asayama, 2009). El componente  $\alpha$  consiste en dos dominios plegados independientemente que forman un dímero y están codificados por el gen *rpoA*. El dominio amino-terminal de  $\alpha$  está implicado en la interacción con el resto de subunidades de RNAP mientras que su dominio carboxilo-terminal interacciona con el molde de DNA y los activadores transcripcionales (Ebright, 1995). Por su parte, *rpoB*, que codifica para  $\beta$ , posee una zona conservada evolutivamente que le sirve para interaccionar con un factor  $\sigma$  que sufre un cambio conformacional tras el contacto facilitando su unión con la caja -35 (Kuznedelov et al., 2002). Una particularidad de los organismos fotosintéticos es que la subunidad  $\beta'$  se encuentra dividida en dos partes,  $\gamma$  (*rpoC1*) y  $\beta'$  (*rpoC2*). Por tanto, el núcleo real de la RNAP que nos atañe es  $\alpha(2)\beta\beta'\gamma\omega$  (Allison et al., 1985; Schneider y Hasekorn, 1988). La pequeña subunidad  $\omega$ , producida por *rpoZ*, no resulta ser esencial pero contribuye a la formación y estabilidad de la polimerasa favoreciendo el plegamiento de la subunidad  $\beta'$  (Minakhin et al., 2001). Sin embargo, no está claro si la RNAP cianobacteriana contiene la subunidad  $\omega$ , a pesar de que su gen está codificado en el genoma (Mukherjee y Chatterji, 1997).

Por otro lado, en cianobacterias el factor  $\sigma$  más habitual y que es esencial para la viabilidad celular es  $\sigma_A$  (*rpoD1*), comparable a  $\sigma_{70}$  en bacterias. Sin embargo, con la intención de tener capacidad de respuesta a diferentes situaciones de estrés, se pueden incorporar factores sigma alternativos que confieren al holoenzima especificidad para otros promotores. Al menos otros dos de los factores  $\sigma$  del grupo II,  $\sigma_B$  y  $\sigma_C$ , codificados por *rpoD2* y *rpoD5* respectivamente, están implicados en la transcripción de genes del regulón de NtcA (Ghosh et al., 2010; Imamura y Asayama, 2009). La RNAP debe unirse de forma específica al promotor del gen que va a transcribir para que la transcripción tenga lugar y esto lo consigue gracias a las subunidades  $\sigma$ . Dichas subunidades facilitan la separación de la doble hebra de DNA en el sitio de inicio de la transcripción, aseguran el reconocimiento



de la secuencia promotora específica y facilitan la unión de la RNAP a esta secuencia (Feng et al., 2016).

En muchos casos, adicionalmente, se logra con la ayuda de factores de transcripción que contribuyen en el reclutamiento de la RNAP y constituyen el complejo de iniciación RNAP-promotor cerrado (RPC). La RNAP junto con los factores de transcripción facilita el desenrollamiento de 12-14 pb del dúplex de DNA y permite la unión de la hebra molde al sitio activo de la RNAP formando el complejo RNAP-promotor abierto (RPO) (Boyaci et al., 2019). Cuando esto ocurre da comienzo la síntesis de RNA y la subunidad  $\sigma$  se disocia antes de empezar la etapa de elongación (Murakami, 2015; Ghosh et al., 2010).

### 1.7. Plan de trabajo

Para la consecución exitosa del proyecto es necesaria la obtención de la RNAP de forma soluble en cantidades apreciables y elevada pureza. Sin embargo y con el objeto de evitar la más que probable insolubilidad de las subunidades cuando estas se expresan por separado, deberán ser clonadas en vectores compatibles para su co-expresión en *Escherichia coli*. Por esta razón los genes de los distintos componentes se clonan en el plásmido bicistrónico pETCola (Novagen), y en el plásmido de co-expresión pQLink (Addgene), donde teóricamente no hay restricción en cuanto al número de subunidades a clonar.

La elección de los plásmidos destino no es aleatoria. Los vectores donde se clonan las correspondientes subunidades son policistrónicos y compatibles para su co-expresión en *E.coli*. Es decir, que contienen diferentes marcadores seleccionables, en este caso, resistencias a antibióticos distintos, y diferentes orígenes de replicación. Respecto a la primera condición pETCola resiste a kanamicina mientras que la ampicilina nos sirve para seleccionar los pQLink. Asimismo, son compatibles para su transformación en la cepa de expresión *E.coli* Rosetta, debido a que esta contiene un plásmido que codifica para tRNAs de codones raros en *E.coli* y que posee resistencia a cloranfenicol. En relación a los diferentes orígenes de replicación, en pQLink está ColE1Ori, en cambio, para pETCola es Cola. Por otro lado, el motivo por el cual se usan ambos plásmidos en lugar de solamente pQLink es para evitar posibles problemas en la transformación del plásmido debido a que se generaría un plásmido de un tamaño demasiado grande.

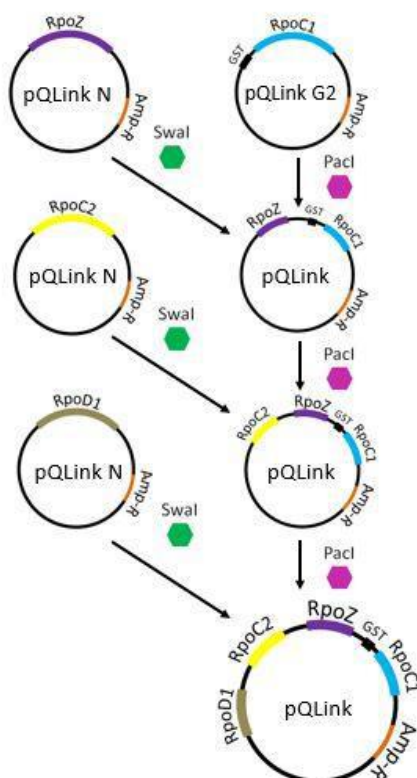
Por su parte, el sistema LINK es un procedimiento de clonación en tándem estandarizado y robusto que permite la generación de clones de co-expresión en *E.coli*. En principio, este método admite un número ilimitado de genes para ser co-expresados. Se generan vectores pQLink con los genes para cada una de las subunidades de la RNAP (excepto para dos de las subunidades, que se clonan en el plásmido pETCola) mediante clonación convencional. Así, se produce un plásmido pQLink con los genes para 4 subunidades de la RNAP gracias a la escisión con el enzima PacI de un plásmido e inserción en el sitio generado con Swal en otro plásmido mediante la técnica LIC (Ligation Independent Cloning). En este trabajo se utilizan los plásmidos pQLinkG2, que contienen una etiqueta GST, y pQLinkN. Para lograr la clonación de los genes en tándem en el plásmido pQLink final se utilizan unos enzimas de restricción raros (PacI y Swal), y luego se fusionan dos plásmidos pQLink (con diferentes subunidades de la RNAP), evitando la

reacción en cadena de la polimerasa (PCR) que puede introducir mutaciones. Repitiendo el proceso de fusión entre el producto anterior y un nuevo plásmido o constructo se pueden clonar en tándem los genes de manera que los vectores permiten que las proteínas se expresen individualmente en cantidades estequiométricas (Scheich et al., 2007).

### 1.8. Estado inicial del proyecto

En el laboratorio del grupo de acogida se encuentran los genes para 4 de las subunidades del núcleo principal de la RNAP ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\omega$ ), así como los factores sigma ( $\sigma_A$ ,  $\sigma_B$  y  $\sigma_C$ ) clonados en los plásmidos pQLinkG2 y pQLinkN, por lo que parte del trabajo de clonaje ya está realizado. Las secuencias de los genes utilizados se recogen en el Anexo I. El gen para la subunidad  $\beta'$  (*rpoC2*) también se encuentra clonado en pQLinkN pero posee una mutación (inserción de 49 pb) que se debe corregir. La intención es preparar tres combinaciones distintas del plásmido pQLink en el cual clonar en tándem tres de las subunidades del núcleo de la RNAP y que el cuarto gen clonado en tándem sea el de cada una de las tres subunidades sigma, mediante fusiones (Figura 4). Las otras dos subunidades  $\alpha$  y  $\beta$  pertenecientes al núcleo de la RNAP, ya se encuentran clonadas en pETCola. Esta combinación permitirá la purificación de la RNAP mediante cromatografía de afinidad empleando una cola de polihistidina en el extremo N-terminal del gen clonado en la posición 1 del plásmido pETCola (*rpoA*) y una etiqueta GST en el extremo N-terminal del gen *rpoC1* clonado en pQLinkG2. La elección de las dos subunidades de la RNAP clonadas junto a las etiquetas no es aleatoria sino que su elección se basa principalmente en criterios estructurales, ya que tras el análisis *in silico* de la RNAP bacteriana de *Thermus aquaticus* (código PDB: 1L9Z), se observa que la región N-terminal de las subunidades  $\alpha$  y  $\gamma$  queda expuesta hacia el solvente, y lejos del resto de la RNAP, por lo que las etiquetas introducidas no deberían dificultar el ensamblaje en la forma nativa de la RNAP.

Además el laboratorio de acogida también cuenta con los clones para la expresión de los reguladores transcripcionales NtcA y PipX, para lo cual seguiremos los protocolos para su expresión y purificación descritos por Llácer, 2012 y Forcada-Nadal, 2018.



**Figura 4. Esquema de los pasos requeridos para la fusión de los genes en el vector de co-expresión pQLink.** Orden de clonación y disposición de los genes en distintos colores. Para *rpoD2* y *rpoD5* se procederá de igual manera que con *rpoD1*. También se señalan elementos como los vectores con sus resistencias a ampicilina y las endonucleasas Swal (verde) y PacI (morado).

## 2. OBJETIVOS

El objetivo principal del presente trabajo es lograr la expresión recombinante de una RNA polimerasa (RNAP) cianobacteriana, concretamente de *Synechococcus elongatus* y la obtención de las proteínas NtcA y PipX. Dichos elementos se requieren para el posterior estudio de la estructura del complejo RNAP-NtcA-PipX-Promotor mediante criomicroscopía electrónica. Para ello se propusieron los siguientes objetivos específicos aunque la consecución de los mismos se ha visto desafortunadamente afectada por la pandemia causada por el SARS-CoV-2.

1. Diseño de la estrategia de clonación a seguir, en un mínimo número de pasos, para poder expresar y purificar la RNAP en cantidad considerable, elevada pureza y de forma soluble.
2. Obtención de las subunidades  $\beta'$ ,  $\gamma$ ,  $\omega$  junto a  $\sigma_A$ ,  $\sigma_B$  y  $\sigma_C$  clonadas en tándem en plásmidos en los plásmidos policistrónicos pQLink, a una concentración y pureza adecuadas, adquiriendo tres combinaciones distintas (una para cada subunidad  $\sigma$ ).
3. Diseño de una estrategia de mutagénesis para eliminar una inserción de 49 pares de bases en el gen rpoC2 que se produjo durante la clonación de este gen.
4. Expresión y purificación de NtcA obteniendo altas concentraciones y elevada pureza de proteína.
5. Expresión y purificación de PipX obteniendo altas concentraciones y elevada pureza de proteína.
6. Diseño de oligonucleótidos cuya secuencia corresponda a genes bajo el control de NtcA para el ensamblaje del complejo RNAP-NtcA-PipX-DNA para su estudio mediante cryo-EM.
7. Expresión y purificación de la RNAP cianobacteriana. Este objetivo ha resultado completamente inabordable por la detención prematura de la parte experimental del TFM debido a la situación de pandemia causada por el SARS-CoV-2.

### 3. MATERIALES Y MÉTODOS

#### 3.1. Diseño de la estrategia de clonación a seguir

Los genes *rpoA* ( $\alpha$ ) y *rpoB* ( $\beta$ ) ya se encuentran clonados en el vector pETCola. Los genes *rpoZ* ( $\omega$ ), *rpoD1* ( $\sigma_A$ ), *rpoD2* ( $\sigma_B$ ) y *rpoD5* ( $\sigma_C$ ) se encuentran clonados en vectores pQLinkN (ADDGENE) y el gen *rpoC1* ( $\gamma$ ) se encuentra clonado en el vector pQLinkG2 (ADDGENE). El gen *rpoC2* ( $\beta'$ ) también está clonado en un vector pQLinkN, pero contiene una mutación por inserción que debe eliminarse. Esta variación indeseada fue generada probablemente por la aparición de estructuras secundarias en el DNA molde durante la amplificación del gen pues esta inserción es una duplicación que sólo difiere en una base de la secuencia anterior de 49 pb.

De este modo, el plan inicial es corregir la mutación del gen *rpoC2* a la vez que se lleva a cabo la clonación en tándem de las subunidades *rpoZ* y *rpoC1*. Una vez obtenido este plásmido pQLink-*rpoZ-rpoC1*, se clonaría el gen *rpoC2* con el error enmendado para finalmente añadir al constructo cada una de las tres unidades  $\sigma$ , *rpoD1*, *rpoD2* y *rpoD5*, de forma independiente. Este orden de clonación se realiza con la intención de maximizar la eficacia de la expresión posterior y minimizar los pasos para obtener las tres variantes de la RNAP (ver apartado 3.9).

#### 3.2. Transformación en *E. coli*

Las transformaciones por choque térmico se realizan en células competentes de *E. coli* TopTen (NEB) con los plásmidos correspondientes. En este método se produce la transformación debido al aumento de la permeabilidad de la membrana por cambios bruscos en la temperatura. Para ello, se incuban 50  $\mu$ L de células Top Ten y 5  $\mu$ L de vector en un tubo eppendorf durante 30 minutos en hielo, seguido de 45 segundos a 42°C y finalmente 5 minutos más en hielo.

Las transformaciones por electroporación, que suelen ser más eficaces, se realizan en células electrocompetentes DH5 $\alpha$  (Clontech) con los constructos indicados. En primer lugar, se utiliza luz UV para desinfectar las cubetas de electroporación (en el caso de que sean cubetas reutilizadas). En ellas se mezclan 3  $\mu$ L de vector y 40  $\mu$ L de células de *E. coli* DH5 $\alpha$ . Se introduce la cubeta de electroporación en un Electroporator 2510 (Eppendorf) a un voltaje de 1700 V para aumentar la permeabilidad de la membrana celular permitiendo la entrada de los plásmidos de interés a la bacteria.

En ambos casos las células se recuperan del estrés ocasionado, añadiendo a la mezcla 200  $\mu$ L de LB e incubando el microtubo en una estufa SingleDisplay (Memmert) durante 1 hora a 37°C. Finalmente se siembran con 100 - 200  $\mu$ L placas de LB agar con ampicilina (a 0.1 mg/mL) y se incuban en agitación durante la noche a 37°C.

### 3.3. Comprobación de plásmidos con el/los inserto/s deseados mediante PCR de colonias y electroforesis en gel de agarosa

Con el objetivo de verificar que el tamaño del inserto de los transformantes es el esperado, se realiza una PCR de las colonias crecidas en la placa. Seleccionamos 15 colonias aleatoriamente y se pican, una a una, con una punta estéril, y esta se sumerge a continuación en un microtubo de PCR de 0.2 mL con 25  $\mu$ L de mezcla de reacción "Master Mix" (NZYTech). Luego, con la misma punta, se siembra en una placa dividida en 15 zonas, para una ulterior recuperación en caso de que dicha colonia contenga el constructo deseado. Esta placa se incuba a 37°C durante toda la noche en la estufa. Igualmente por cada prueba se incluye un control negativo con una punta estéril. En este experimento se utiliza una DNA polimerasas termoestable con alta procesividad para amplificaciones rutinarias NZYTaq II DNA polymerase.

Para la preparación de los tubos para la PCR seguimos las instrucciones recomendadas por el fabricante (NZYTech) (Tabla 1). También seguimos dichas instrucciones para el protocolo de amplificación (Tabla 2) variando únicamente los tiempos de extensión en función del tamaño del inserto.

**Tabla 1. Reactivos utilizados en la PCR con NZYTaq II DNA polimerasa**

Componente	Concentración	Volumen
2x Green Master Mix	1x	12.5 $\mu$ L
10 $\mu$ M Cebador pQTEV3U	0.5 $\mu$ M	1.25 $\mu$ L
10 $\mu$ M Cebador pQTEV3L	0.5 $\mu$ M	1.25 $\mu$ L
Agua	-	Hasta 25 $\mu$ L

**Tabla 2. Ciclo térmico utilizado en la PCR con NZYTaq II DNA**

Etapa del ciclo	Temperatura	Tiempo	Número de ciclos
Desnaturalización inicial	95 °C	3 min	1
Desnaturalización	94 °C	40 s	35
Hibridación	50 °C	45 s	
Extensión	72 °C	15-30 s/kb	
Extensión final	72 °C	7 min	1

### 3.4. Electroforesis en gel de agarosa

Una vez obtenidos los productos de PCR se realiza una electroforesis de DNA en gel de agarosa 0.8 % para verificar que el tamaño del amplicón es el esperado. Con este paso se separan los fragmentos de DNA según su tamaño en una cubeta Labnet Gel XL Ultra V-2 con tampón TAE 1x. Tras depositar la muestra con el tampón de carga 6x, preparado en el laboratorio (30% glicerol, 0.25% azul de bromofenol y 0.25 % xylene cyanol ff), en los pocillos, se pueden separar las moléculas al atravesar la matriz porosa del gel de agarosa gracias a un campo eléctrico de 100 V durante 25 minutos. Se cargan 5  $\mu$ L de cada una de las muestras y 10  $\mu$ L de marcador de 1 kb elaborado en el laboratorio a partir de plásmidos pPSU (Henrici et al., 2017). Una vez transcurrido el tiempo, se revela durante 15 minutos en disolución 1x de GelRed® (Biotium). Por último, se efectúa la detección en un equipo

con sistema de captación de imagen Ultima 16si Plus (Isogen) con el software Proxima AQ-4, y se comprueba si la altura de la banda es la esperada según el tamaño teórico del gen de interés amplificado o en este caso a la suma del tamaño de los genes en tándem más el espacio de plásmido situado entre ellos.

En el caso de las PCR de colonia la muestras se pueden cargar directamente en el gel de agarosa para su comprobación por electroforesis sin necesidad de incorporar el tampón de carga pues el NZYTaQ II 2× Green Master Mix ya contiene colorantes para monitorizar el progreso de la separación.

### 3.5. Purificación de los plásmidos con los constructos de interés y cuantificación

Todas las colonias de *E. coli* de las que se amplificó un inserto del tamaño esperado (juzgado mediante electroforesis en gel de agarosa, ver apartado anterior) se pican con un asa bacteriológica en la placa y se inoculan en 10 mL de medio LB líquido con 10 µL de ampicilina (a 100 µg/mL) en tubos de cultivo y se incuban toda la noche a 37°C en agitación. Para el aislamiento y purificación del DNA plasmídico de *E. coli* se sigue un protocolo comercial de lisis alcalina (Miniprep de NZYTech) en el cual la adsorción del DNA plasmídico a una membrana de sílice especialmente tratada tiene lugar en presencia de alta concentración de sales, de manera que impurezas como proteínas, sales, nucleótidos y oligonucleótidos son lavados.

Primero, se recuperan las células en un pellet del cultivo LB de *E. coli* tras centrifugar en una centrífuga 5810R con refrigeración y rotor basculante (Eppendorf) durante 10 minutos a 4000 xg. Se retira el sobrenadante y se elimina todo el material posible. A continuación, se produce la lisis celular al resuspender el pellet en 250 µL de tampón A1, con RNAsas, y mezclar vigorosamente. Se transfiere la solución a un microtubo de 1.5 mL donde se añaden otros 250 µL de tampón A2 (con NaOH 0.2M y SDS 1%) y se mezcla 8 veces por inversión dejando incubar a temperatura ambiente no más de 4 minutos. Además, se adicionan 300 µL de tampón A3 (neutralizante; con acetato sódico pH 4.8 y el agente caotrópico cloruro de guanidinio) que se mezclan mediante la inversión del tubo 8 veces.

Después, se centrifuga 5 - 10 minutos dependiendo del volumen de cultivo inicial a 11000 xg a temperatura ambiente para clarificar el lisado. Se coloca la columna NZYTech en un tubo de recolección de 2 mL y se pasa el sobrenadante para que tenga lugar la unión del DNA a la columna de sílice. De nuevo se centrifuga 1 minuto a 11000 xg y se descarta el eluido. A continuación, se agregan al filtro 500 µL de tampón AY para el lavado de la membrana de sílice para la eliminación de los altos niveles de endonucleasas. Se vuelve a centrifugar a 11000 xg durante 60 segundos y se desecha el eluido. Más adelante, se centrifuga otro minuto a 11000 xg tras sumar 600 µL de Buffer A4 (con etanol) y se elimina el eluido de nuevo. Con la intención de secar la columna se centrifuga 2 minutos a 11000 xg recogiendo el eluido en un tubo recolector de 2 mL nuevo. Posteriormente, se coloca la columna NZYTech ya seca en un microtubo de 1.5 mL y se añaden 50 µL de Buffer AE justo a la parte central de la columna. Finalmente se incuba 1 minuto a temperatura ambiente y se centrifuga otro minuto a 11000 xg provocando la elución del DNA plasmídico con alta pureza. Tras el aislamiento del DNA de interés se determina la concentración en la muestra purificada midiendo su absorbancia a 260 nm mediante el NanoDrop™ 1000

Spectrophotometer (Thermo Scientific). También se considera el ratio de absorbancia 260/280 que nos indica la calidad de la muestra, siendo generalmente aceptado un valor  $\sim 1,8$  para un DNA puro. Para la medición, primero se lava y calibra la zona de medición con 1  $\mu\text{L}$  de agua Milli-Q. Seguidamente se utiliza como blanco el tampón AE o TE (Turbo Elution) usados y se añade 1  $\mu\text{L}$  de muestra para la medición.

### 3.6. Purificación de fragmentos lineales de DNA y cuantificación

El aislamiento y purificación de fragmentos de DNA desde soluciones enzimáticas o desde geles de agarosa se realiza siguiendo el protocolo GeneClean Turbo Kit (MP Biomedicals). Este procedimiento permite el desalado, la eliminación de proteínas de la reacción enzimática y la exclusión de cebadores y nucleótidos no incorporados. La adición de sales caotrópicas a una alta concentración rompe la estructura del agua alrededor de la sílice, permitiendo la formación de un enlace con las cargas negativas de los fosfatos del esqueleto del DNA.

Siguiendo las instrucciones indicadas por la empresa, se mide el volumen de la muestra para purificar y se añade 5 volúmenes de la solución Turbo Salt mezclando el resultado, de forma que quedan ajustadas las condiciones. El segundo paso es transferir a la columna en cantidades menores a 600  $\mu\text{L}$  de DNA/Solución salina. Se centrifuga a menos de 14000  $\times g$  hasta que todo el líquido pase a través del filtro (5 segundos) en una microcentrifuga de sobremesa con refrigeración 5415R (Eppendorf). Luego, se vacía el tubo recolector y se repite el procedimiento si fuera necesario hasta transferir toda la muestra al filtro. Seguidamente, se añaden 500  $\mu\text{L}$  del preparado Turbo Wash para llevar a cabo el lavado, se centrifuga a menos de 14000  $\times g$  durante 5 segundos y se vacía el tubo recolector de nuevo. Repitiendo el procedimiento de lavado tal como se ha detallado se consigue aumentar la pureza. Se vuelve a centrifugar durante 4 minutos adicionales a la misma aceleración para favorecer el secado del Turbo Cartridge.

Finalmente, se inserta la columna en un microtubo de 1.5 mL y se añaden 30  $\mu\text{L}$  de la solución Turbo Elution directamente sobre la membrana. Se incuba durante 5 minutos a temperatura ambiente. Cuando la concentración de sales disminuye, la rehidratación de la matriz permite romper la atracción entre la matriz de sílice y el DNA. Así se eluye el DNA al centrifugar otra vez a menos de 14000  $\times g$  durante 5 segundos. La concentración de DNA se determina al igual que en el caso del DNA plasmídico midiendo su absorbancia a 260 nm (ver apartado anterior).

### 3.7. Comprobación de los insertos mediante secuenciación

Se secuenciaron las muestras resultantes de la mutagénesis y los constructos obtenidos tras las fusiones en el Servicio Secuenciación y Análisis de ADN del Instituto de Biomedicina de Valencia. El servicio requiere por secuenciación un mínimo de 4  $\mu\text{L}$  de muestra a una concentración adecuada (mayor a 50  $\text{ng}/\mu\text{L}$ ) y 5  $\mu\text{L}$  de cebador a una concentración de 5 mM. Para saber si los genes se encuentran realmente en el plásmido aislado y que estos genes se han clonado sin ninguna modificación, se llevaron a cabo reacciones de secuenciación automática usando los cebadores especificados en el Anexo III. Se trata de los mismos cebadores empleados para realizar las PCR de colonia. Además para las secuencias más largas se diseñan nuevos cebadores internos con la ayuda del

programa para análisis de secuencias Gene Runner. Cuando se reciben los archivos (Anexo III) con una longitud de lectura de 600 – 800 pb con las secuencias se efectúa un alineamiento de los ficheros adquiridos para cerciorarse de la fidelidad de la secuencia, comparando, mediante Clustal Omega, con los distintos genes de la base de datos que codifican la RNA polimerasa (Anexo I).

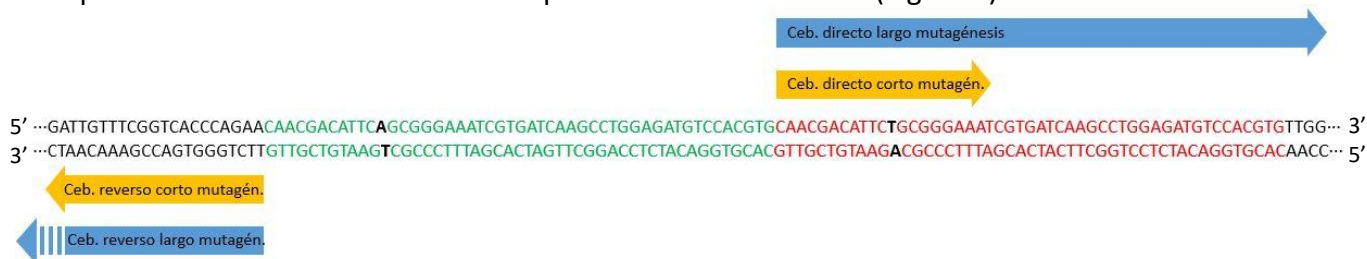
### 3.8. Eliminación de una mutación en el gen *rpoC2*

El gen *rpoC2*, que se encuentra clonado en el vector pQLinkN, contiene una inserción de 49 pb que se desea eliminar mediante mutagénesis. Cabe destacar que en la clonación previa del gen en pQLinkN surgieron numerosas dificultades, debidas en parte por el gran tamaño del gen *rpoC2* (3957 pb) que deriva en poco rendimiento de las amplificaciones en cuanto a cantidad de producto e inespecificidad y aparición de numerosas mutaciones puntuales. Por tanto se decide inicialmente tratar de eliminar la mutación en vez de empezar la clonación de nuevo. Con tal fin se propone utilizar una estrategia de mutagénesis de sitio dirigido (SDM) mediante reacción en cadena de la polimerasa (PCR).

#### 3.8.1. Mutagénesis de sitio dirigido (SDM)

En primer lugar, se diseña un experimento de SDM siguiendo el protocolo Q5 Site-directed mutagenesis de New England Biolabs. Este procedimiento utiliza cebadores de oligonucleótidos diseñados a medida para corregir una mutación *in vitro*. Los cebadores directo y reverso diseñados son estándar, no mutados y con orientación consecutiva (back-to-back) de forma que flanquean la región a eliminar. El primer paso es la amplificación exponencial utilizando los cebadores no solapantes y un master mix que contiene la DNA polimerasa (Q5 Hot Start High-Fidelity). El segundo paso implica la incubación con una mezcla enzimática (que contiene kinasa, ligasa y de restricción DpnI) Estas enzimas permiten la circularización del producto de la PCR y la eliminación del DNA molde. El último paso es una transformación de alta eficiencia en células competentes.

Los cebadores propuestos (Anexo II), seleccionados mediante el software Gene Runner y que se utilizan para amplificar el gen *rpoC2* sin la secuencia mutada, son de dos tipos: unos más cortos y otros más largos. Los cebadores largos son oligonucleótidos de 52 pb de longitud de modo que uno de ellos cubre la mutación y unas pocas bases más que le proporcionan suficiente especificidad para diferenciar la mutación de la secuencia correcta. En cambio, los cebadores cortos tienen longitudes de 18 pb y 21 pb y para eliminar la inserción se trataría de explotar el hecho de que la duplicación de 49 pb no es perfecta sino que además contiene una mutación puntual de una sola base (Figura 5).



**Figura 5. Estrategia de mutagénesis del gen *rpoC2*.** Secuencia nucleotídica correcta en verde y en rojo la mutación. En negrita se destaca la base puntual en que difiere. Las flechas azules corresponden a los cebadores largos (el reverso continúa hacia 3') y las amarillas a los cortos.



Primero se prueba con los cebadores largos pues si se consigue obtener producto de la PCR sabiendo la dificultad de obtener resultados (ya que el tamaño de los cebadores obliga a utilizar temperaturas de hibridación cercanas a la temperatura de extensión) la probabilidad de que la mutación haya sido corregida es mayor que utilizando los cebadores cortos. También se intenta corregir la mutación empleando los cebadores cortos teniendo en cuenta que es más fácil obtener producto de la PCR, aunque, la probabilidad de que se repare la mutación es menor, porque se pretende que la diferencia en una única base sea suficiente para amplificar la secuencia correcta y no la inserción. Con este método se realizaron varias PCR usando la DNA polimerasa de alta fidelidad NZY Supreme (NZYTech) (en lugar de Q5) y variando las concentraciones de DNA molde y de DMSO (previene la formación de estructuras secundarias y facilita la hibridación). Los componentes de las reacciones y las condiciones generales de las PCR se recogen en la Tabla 3 y Tabla 4.

**Tabla 3. Reactivos utilizados en la PCR con NZYSupreme DNA polimerasa**

Componente	Concentración	Volumen
Polimerasa NZY Supreme	0.025 U/ $\mu$ L	0.25 $\mu$ L
5x Reaction Buffer	1x	5 $\mu$ L
10 $\mu$ M dNTPs	0.2 $\mu$ M	0.5 $\mu$ L
10 $\mu$ M F. primer MFL	0.5 $\mu$ M	1.25 $\mu$ L
10 $\mu$ M R. primer MRL	0.5 $\mu$ M	1.25 $\mu$ L
DNA	154 ng	1 $\mu$ L *
DMSO 100%		2 $\mu$ L *
Agua		hasta 25 $\mu$ L

\* Condiciones variables según el experimento

**Tabla 4. Ciclo térmico en la PCR con NZYSupreme DNA polimerasa**

Etapas del ciclo	Temperatura	Tiempo	Número de ciclos
Desnaturalización inicial	96 °C	4 min	1
Desnaturalización	96 °C	30 s	25 *
Hibridación	72 °C (ceb. largos) * 55 °C (ceb. cortos) *	30 s	
Extensión	72 °C	4 min 30 s *	
Extensión final	72 °C	7 min	1

\* Condiciones variables según el experimento

Una vez obtenido el producto de PCR, este se trata con la endonucleasa DpnI que escinde cuando su sitio de reconocimiento se encuentra metilado, así se elimina el molde aislado de *E. coli* y queda el DNA sintetizado *in vitro*. Para ello se digieren 25  $\mu$ L de producto de PCR con 10 U de DpnI (NEB) durante 1 hora a 37°C. También se efectúa, alternativamente, un tratamiento con DpnI seguido de una incubación, a 15°C durante 5

horas, de la mezcla obtenida con 2 U de T4 DNA ligasa (NEB) y 2  $\mu\text{L}$  de tampón ligasa 10x llevándolo después a un volumen de 20  $\mu\text{L}$  con agua. Este paso se realiza para favorecer la unión entre los extremos creados por los cebadores no consecutivos al realizar la delección. Ante los resultados adversos, y con el mismo propósito, se utiliza el mix KLD (Kinase-Ligase-Dpnl) de New England Biolabs, en las condiciones recomendadas por el fabricante (Tabla 5).

**Tabla 5. Tratamiento con mix KLD**

Componente	Concentración	Volumen
Producto de PCR		1 $\mu\text{L}$
2X KLD Reaction Buffer	1x	5 $\mu\text{L}$
10X KLD Enzyme Mix	1x	1 $\mu\text{L}$
Agua	-	Hasta 10 $\mu\text{L}$

A continuación, se realiza una electroforesis de DNA en gel de agarosa 0.8 %, en el cual, se carga 5  $\mu\text{L}$  muestra para comprobar que el tamaño del amplicón es el esperado (ver apartado 3.4). Seguidamente, se transforma *E. coli* DH5 $\alpha$  por electroporación con el producto de PCR obtenido (ver apartado 3.2), y se realiza una PCR de colonia para comprobar los transformantes (ver apartado 3.3) antes de confirmar mediante una electroforesis en gel de agarosa 0.8 % que el inserto es del tamaño esperado. Posteriormente, se purifica y se cuantifica el DNA plasmídico de las colonias que contenían el inserto esperado (ver apartado 3.5). Finalmente para verificar si se ha conseguido el gen rpoC2 correcto, se secuencian los fragmentos de interés (ver apartado 3.7) donde se encuentra la mutación en las muestras purificadas en el Servicio de Secuenciación y Análisis de ADN del Instituto de Biomedicina de Valencia usando 5  $\mu\text{L}$  de cada muestra purificada y 4  $\mu\text{L}$  del cebador interno de rpoC2, PIB3 (Anexo II) (5 mM).

### 3.8.2. Clonación del gen a partir del plásmido original con rpoC2

Originalmente y previamente a la clonación en el plásmido pQLinkN, el gen rpoC2 se clonó en un plásmido pLIC a partir de DNA genómico de *S. elongatus*. Decidimos clonar de nuevo rpoC2 en pQLinkN, pero con el objeto de no proceder a todo el proceso de clonación sin saber previamente si es posible amplificar el gen de forma correcta sin la inserción, se procede a secuenciar los amplicones resultantes de la amplificación del gen rpoC2 a partir del plásmido original pLIC, usando diversas concentraciones de DNA molde (10ng; 20ng), DMSO (0 $\mu\text{L}$ ; 0.5 $\mu\text{L}$ ; 1 $\mu\text{L}$ ) y variando la temperatura de extensión. Estas amplificaciones por PCR se llevan a cabo usando la DNA polimerasa NZY Supreme (NZYTech) bajo las mismas condiciones que la Tabla 3. Se purifican los productos de PCR con el kit GeneClean y se secuencian (ver apartados 3.6 y 3.7). Dado que obtenemos una banda del tamaño esperado (ver Figura 9 en sección Resultados) y con el objeto de ganar algo de tiempo en el proceso de clonaje en el caso de que la secuenciación del amplicón demuestre que la inserción indeseada ya no se produce, en paralelo se digieren 1.76  $\mu\text{g}$  del amplicón con 36 U de las enzimas NotI y HindIII y 4.1  $\mu\text{L}$  de tampón CutSmart que se lleva a un volumen de 41  $\mu\text{L}$ . Luego, tras una nueva purificación con el kit de GeneClean, se realiza la ligación con la T4 DNA ligasa del fragmento con el vector pQLinkN, previamente digerido. El procedimiento seguido para la ligación se describe con mayor detalle en el apartado siguiente.

Desafortunadamente, antes de poder completar el clonaje, los resultados de secuenciación indican que también está la mutación. Es entonces cuando se decide enviar a secuenciar rpoC2 el plásmido inicial pLIC, descubriendo por los resultados que ya estaba mutado en origen (poseíamos 3 plásmidos aislados y teóricamente secuenciados, y para nuestra sorpresa en todos estaba la inserción).

Por tanto, en tercer lugar, se propone llevar a cabo una escisión de la parte incorrecta del gen mediante digestión encontrando sitios de restricción únicos (que no corten el resto del gen o el plásmido pQLinkN) cercanos por ambos lados a la mutación, para sustituirlo posteriormente con el fragmento del gen corregido.

### 3.8.3. Eliminación de la mutación de rpoC2 mediante su sustitución con un gen artificial

Dicho método consiste en la búsqueda de dos endonucleasas que solamente tengan un único sitio de restricción en pQLinkN-rpoC2, lo más aproximado posible a la mutación por ambos lados. De este modo se sustituye el fragmento liberado con la mutación por su forma correcta. Mediante la herramienta NEBCutter v2.0 se observó que las enzimas de restricción BbvCI y NgoMIV (New England Biolabs) pueden escindir un fragmento de 1438 pb, dejando linealizado el resto de gen con el plásmido (7327 pb). Por otra parte se diseña un gen sintético de 1389 pb (Anexo I), del mismo fragmento pero sin la inserción, y se recibe ya clonado en el plásmido de alto número de copias pUC19 (Proteogenix).

Para ello, el primer paso es digerir 25  $\mu\text{L}$  del plásmido pQLinkN-rpoC2 (154  $\text{ng}/\mu\text{L}$ ) con 36 U de la enzima de restricción BbvCI (2  $\text{U}/\mu\text{L}$ ) y 18 U de NgoMIV (10  $\text{U}/\mu\text{L}$ ) según las recomendaciones del fabricante durante 2 horas a 37°C. Una vez finalizada la digestión, se realiza una electroforesis en gel de agarosa 0.8 % cargando los 50  $\mu\text{L}$  de producto de digestión y 10  $\mu\text{L}$  de marcador de 1 kb (Thermo Fisher) para purificar el fragmento de 7327 pb. En este caso, se utiliza un peine como molde modificado para unir los pocillos y que quepa el mayor número de microlitros de muestra posible. También se mantiene corriendo durante 1 hora (mayor tiempo que el habitual), favoreciendo así la separación de las diferentes bandas para la posterior purificación de la banda de interés.

Una vez finalizada la electroforesis, se lleva a cabo la purificación desde gel de agarosa, con la ayuda de un bisturí y gracias a un transiluminador Vilber Lourmat Super Bright TFX-20 MX. Se escoge la sección de gel que contiene pQLinkN-rpoC2 linealizado (7327 pb) y sin la región errónea ya liberada. Las secciones se introducen en un tubo de cultivo vacío, y este habrá sido previamente pesado en una balanza de precisión PB3002-S DeltaRange (Mettler Toledo), para que, tras sumar los pedazos de gel, por diferencia de masa, se pueda determinar los gramos de gel que han sido recuperados. Sabiendo esta medida, se puede empezar con la purificación mediante GeneClean Turbo Kit (MP Biomedicals). Así, se procede adicionando 100  $\mu\text{L}$  de la solución Turbo Salt por cada 0.1 g de porción de gel. Seguidamente, se incuba la solución durante 5 minutos a 55°C en el baño de agua Grant TX150-P5 para que se derrita el gel, invirtiendo el tubo para favorecer la homogeneización. A partir de esta combinación DNA/solución salina se puede transferir una cantidad menor a 600  $\mu\text{L}$  a cada columna siguiendo con el protocolo detallado en el apartado 3.6 a partir del segundo paso.

Por otra parte, se digieren 2.5 µg del plásmido pUC19 con el fragmento de gen artificial (200 ng/µL) con 25 U de la enzima de restricción BbvCI (2 U/µL) y 13 U de NgoMIV (10 U/µL) según las recomendaciones del fabricante durante 1 hora a 37°C. Una vez finalizada la digestión, se realiza una electroforesis en gel de agarosa 0.8 % cargando los 31 µL de producto de digestión y 10 µL de marcador de 1 kb (Thermo Fisher). A continuación, se lleva a cabo la purificación desde gel de agarosa de la banda correspondiente al fragmento de 1389 pb del gen sin la mutación siguiendo el mismo procedimiento utilizado para purificar el resto de plásmido pQLinkN-rpoC2. El siguiente paso es realizar la ligación del fragmento de gen de rpoC2 sin la mutación purificado y el resto de plásmido pQLinkN-rpoC2 purificado empleando la ligasa de DNA del bacteriófago T4 (1 U/µL) y el tampón de reacción, T4 DNA Ligasa Buffer 10x (New England Biolabs) atendiendo a las condiciones de la Tabla 6. Se realizan dos intentos de ligación con proporciones plásmido:inserto de 1 y 1.5 en un volumen total de 20 µL añadiendo la máxima cantidad posible de plásmido e inserto. Una vez finalizada la reacción, se transforman células de *E. coli* para finalmente poder obtener el gen rpoC2 clonado en pQLinkN libre de mutaciones.

**Tabla 6. Condiciones para la ligación con T4 DNA ligasa**

<b>Componentes</b>	<b>Relación 1/1</b>	<b>Relación 1/1.5</b>
Vector	11.36 µL (53.4 ng)	9.7 µL (45.6 ng)
Inserto	5.94 µL (53.46 ng)	7.6 µL (68.4 ng)
T4 DNA ligasa	1 µL	1 µL
ATP 0.1 M	0.2 µL	0.2 µL
Buffer 10x	1.5 µL	1.5 µL
Agua	Hasta 20 µL	Hasta 20 µL

### 3.9. Clonación en tándem de los distintos genes en los plásmidos pQLink

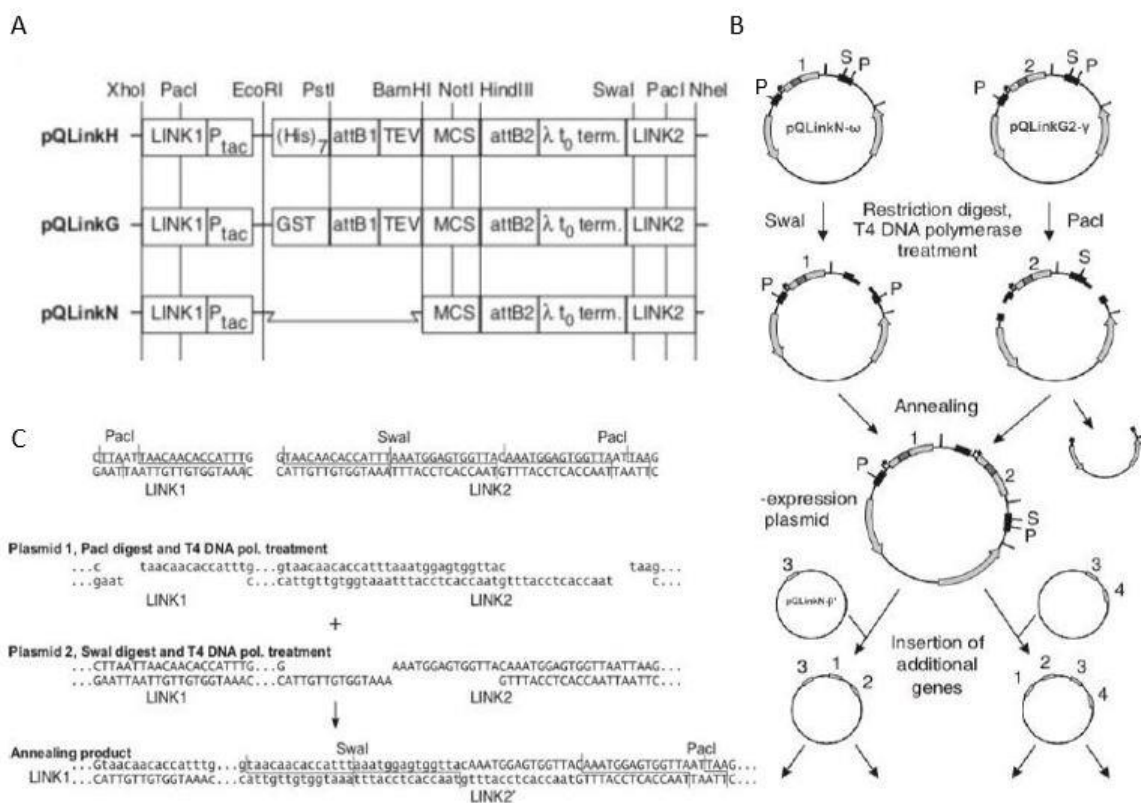
#### 3.9.1. Diseño de la estrategia de clonación a seguir y fundamento general de la clonación de genes en tándem en vectores pQLink

La clonación de los distintos genes de la RNAP en tándem se desarrolla según el procedimiento explicado en Scheich et al. (2007) que permite la generación de clones de co-expresión pQLink en *E. coli*. En primer lugar, se “fusionan” dos de los plásmidos pQLink siguiendo el protocolo mencionado. Es importante mencionar que tras la “fusión” de los dos plásmidos pQLink, el plásmido no se duplica, y el plásmido resultante es de un solo pQLink y dos genes clonados en tándem. A continuación, al producto resultante de la fusión de las dos subunidades se le adiciona otra de las subunidades siguiendo el mismo protocolo anterior, para obtener un nuevo pQLink con tres subunidades clonadas en tándem, y así sucesivamente se procedería si quisiésemos adicionar más subunidades. En nuestro caso, a este pQLink de tres subunidades se le adicionaría una cuarta subunidad, las subunidades  $\sigma$  (rpoD1, rpoD2, rpoD5), para obtener los tres distintos sets de plásmidos pQLink (Figura 4).

En mayor detalle, el método consiste en que, tras la clonación de forma individual de cada subunidad en plásmidos pQLink (Figura 6, panel A), se utilizan las endonucleasas PacI y SwaI para cortar cada uno de los dos plásmidos a fusionar, respectivamente. Mientras que PacI corta en ambos extremos del gen de interés, SwaI escinde una única vez

en el extremo C-terminal del gen de interés dejando el vector linealizado. Es por tanto muy importante cerciorarse de que estos enzimas de restricción no digieran accidentalmente ninguno de los genes a clonar, lo cual no es sencillo teniendo en cuenta que se trata de genes largos que en algún caso se aproximan a los 4000 pares de bases de longitud. En el caso de que corten, aún así se podría realizar la clonación en tándem para dicho gen con el resto de genes si es solo uno de los dos enzimas el que corta; para ello se ha de tener en cuenta el orden de adición del gen y su lugar en el constructo final.

Teniendo todo esto en cuenta, en el proceso de clonación de cada uno de los genes clonados anteriormente en vectores pQLink en un único vector pQLink, se debe cumplir la premisa de que rpoC1 ( $\gamma$ ) sea digerido con PacI. Esto es debido a que la región codificante para la etiqueta GST es diana de la enzima de restricción SwaI. Por tanto, se incluye el gen rpoC1 en el extremo más C-terminal (4ª posición) para que la GST no se vea comprometida. En cuanto al orden del resto de genes, en tercer lugar quedaría el gen rpoZ de la pequeña subunidad  $\omega$  y es preferible que rpoC2 ( $\beta'$ ) quede en segunda posición para garantizar una conveniente expresión pese a su gran tamaño. Con la intención de minimizar el número de pasos, se sitúa en el extremo más N-terminal (1ª posición) el factor  $\sigma$  que es la subunidad variable de la RNAP (Figura 4 y Figura 6, panel B). De este modo una vez obtenido el núcleo  $\beta'$ - $\omega$ - $\gamma$ , solamente sería necesario incorporar cada subunidad  $\sigma$  ( $\sigma_A$ ,  $\sigma_B$  o  $\sigma_C$ ) a este núcleo una sola vez, mientras que de otro modo implicaría dos eventos de clonación extra.



**Figura 6. Sistema Link de clonación. (A)** Elementos genéticos de los vectores pQLinkH, pQLinkG y pQLinkN (MCS: Multiple cloning site, TEV: sitio de escisión de proteasas, term: terminador de la transcripción). **(B)** Construcción del plásmido de coexpresión a partir de los vectores pQLinkN y pQLinkG2 con dos insertos distintos nombrados como 1 ( $\omega$ ) y 2 ( $\gamma$ ), el plásmido originado acepta insertos adicionales (S: SwaI, P: PacI). **(C)** Las secuencias Link, su digestión con PacI y SwaI, su tratamiento con T4 DNA pol. y su posterior ligación. Editada de Scheich et al., (2007).

La producción de un plásmido de co-expresión requiere 0.2-0.5 µg de cada uno de los dos plásmidos pQLink implicados en la fusión. Uno de ellos se digiere durante 3 horas a 37°C con 5 unidades de Pacl, mientras que el otro se escinde con 5 unidades de Swal a 25°C durante el mismo tiempo. Como se ha comentado es necesario que el plásmido pQLinkG2, que contiene la etiqueta GST se digiera con Pacl. En todos los casos se adicionan las cantidades de plásmido en proporciones molares incorporando 0.5 µg del que posee menor masa molar. A continuación, se inactivan los enzimas durante 20 minutos a 65°C.

Tras la digestión e inactivación, el siguiente paso se basa en la generación de extremos cohesivos mediante el uso de la T4 DNA polimerasa (técnica LIC; Ligation Independent Cloning) (Figura 6, panel C). Así, a 5 µL de cada solución se les añade 1.3 unidades de T4 DNA polimerasa (NEB) con el tampón adecuado (50 mM Tris-HCl, PH 8.0, 10 mM MgCl<sub>2</sub>, 5 µg/mL BSA, 5 mM DTT) y se adiciona 2.5 mM dCTP al producto de digestión con Pacl y 2.5 mM dGTP al producto de digestión con Swal. Tras 30 minutos de incubación a 25°C y una inactivación por calor a 65°C por 20 minutos, se mezclan los dos plásmidos. El producto se calienta a 65°C durante 4 minutos para eliminar cualquier estructura secundaria que pudiese dificultar la ligación posterior y se enfría a temperatura ambiente para la hibridación. Justo después se añaden 2 µL de 25 mM de EDTA y se lleva a cabo la transformación de 5 µL de la mezcla final en *E. coli* TopTen usando el procedimiento habitual (ver apartado 3.2).

La comprobación de los transformantes que contengan los genes correspondientes clonados se efectúa mediante PCR de colonia con los oligonucleótidos comerciales específicamente diseñados para la comprobación, pQTEV3U y pQTEV3L (Anexo II), con NZYtaq II 2× Green Master Mix (ver apartado 3.3). Cuando se consigue al menos una colonia positiva, juzgado mediante electroforesis en gel de agarosa 0.8 % (ver apartado 3.4), se inocula un cultivo líquido en 10 mL de LB y 10 µL de ampicilina 100 mg/mL (concentración final, 0,1 mg/mL) y se crece durante la noche a 37°C. Al día siguiente, el plásmido se purifica y cuantifica (ver apartado 3.5). Adicionalmente, se digiere el plásmido aislado con Pacl (corta en ambos extremos del gen o genes clonados en tándem) para cerciorarse de que el tamaño del fragmento liberado es el esperado (al menos la suma de los tamaños de los dos o más genes, juzgado mediante electroforesis en gel de agarosa 0.8 % del inserto digerido). En última instancia, se envían las muestras para su secuenciación (ver apartado 3.7).

### 3.9.2. Construcción de plásmidos de co-expresión de las tres variantes RNAP en pQLink

En primer lugar se digiere con Pacl 2.5 µL del plásmido pQLinkG2-rpoC1 (77.6 ng/µL) y 8 µL del plásmido pQLinkN-rpoZ (39.5 ng/µL) con Swal. Tal y como se describe en el apartado anterior, tras la digestión e incubación, se produce el tratamiento con T4 DNA polimerasa, en presencia de dCTPs para el plásmido digerido pQLinkG2-rpoC1 y dGTPs para el plásmido digerido pQLinkN-rpoZ. Posteriormente, se produce la ligación de los dos constructos digeridos con extremos cohesivos, transformación de *E. coli* con el producto de ligación, aislamiento de los plásmidos y comprobación mediante secuenciación de que el constructo pQLink-rpoZ-rpoC1 es correcto. Con posterioridad, conforme a lo planeado se debía proceder a clonar rpoC2, pero debido a que aún no se había podido corregir la mutación indeseada en dicho gen, se decidió clonar cada una de las subunidades  $\sigma$ , rpoD1,

rpoD2 y rpoD5 en el siguiente paso. Para situar los factores  $\sigma$  en tándem con rpoZ y rpoC1, y obtener los constructos [pQLink-rpoD1-rpoZ-rpoC1], [pQLink-rpoD2-rpoZ-rpoC1] y [pQLink-rpoD5-rpoZ-rpoC1] se repite el mismo procedimiento anteriormente utilizado. La reacción de corte de pQLink-rpoC1-rpoZ (100 ng/ $\mu$ L) se realiza con Pacl. Mientras, las escisiones con Swal de pQLinkN con cada una de los genes de las subunidades  $\sigma$  se llevan a cabo según las relaciones de la Tabla 7.

**Tabla 7. Relación de las cantidades de plásmido en la clonación en tándem**

	Incorporación de rpoD1 ( $\sigma_1$ )	Incorporación de rpoD2 ( $\sigma_2$ )	Incorporación de rpoD5 ( $\sigma_3$ )
pQLink-rpoC1-rpoZ	4.3 $\mu$ L (0.43 $\mu$ g)	4.1 $\mu$ L (0.41 $\mu$ g)	4.3 $\mu$ L (0.43 $\mu$ g)
Concentración muestra	100 ng/ $\mu$ L	100 ng/ $\mu$ L	100 ng/ $\mu$ L
pQLinkN-rpoD_	4.34 $\mu$ L (0.5 $\mu$ g)	7.62 $\mu$ L (0.5 $\mu$ g)	3.7 $\mu$ L (0.5 $\mu$ g)
Concentración muestra	115.8 ng/ $\mu$ L	65.6 ng/ $\mu$ L	134.5 ng/ $\mu$ L

### 3.10. Expresión y purificación de las proteínas NtcA y PipX

La expresión y purificación de las proteínas NtcA y PipX se realiza conforme a los procedimientos utilizados en Llácer (2012) y Forcada-Nadal (2018). En ambos casos se emplean protocolos de expresión y purificación partiendo de glicerizados de células *E. coli* Rosetta (Novagen) que portan los vectores pET15b y pTrc99A con los genes *ntca* y *pipx*, respectivamente. La cepa Rosetta es elegida por su idoneidad para incrementar la expresión de proteínas recombinantes de organismos diferentes a *E. coli*, ya que poseen un plásmido con genes codificantes para tRNAs poco usados por *E. coli*. Este plásmido confiere resistencia al cloranfenicol mientras que tanto pET15b como pTrc99A confieren resistencia a ampicilina.

#### 3.10.1. Expresión de NtcA y PipX

En primer lugar, a partir de los glicerizados correspondientes (conservados a  $-80^{\circ}\text{C}$ ) se inoculan cultivos de 5 mL de medio LB líquido y 3  $\mu$ L de ampicilina a 100 mg/mL y 10  $\mu$ L de cloranfenicol a 35 mg/mL y se incuban durante 6 horas a  $37^{\circ}\text{C}$ . Seguidamente, se realiza un segundo precultivo transfiriendo 1 mL del precultivo inicial a un matraz con 25 mL de LB y los mismos antibióticos a la misma concentración, que se deja creciendo durante la noche en agitación a  $37^{\circ}\text{C}$ . Al día siguiente, se vierte todo el volumen del precultivo en un matraz con 1 L de LB al que se le adiciona 600  $\mu$ L de ampicilina 100 mg/mL y 2 mL de cloranfenicol 35 mg/mL y se deja crecer a  $37^{\circ}\text{C}$  en agitación, hasta que se alcanza un valor de densidad óptica  $\text{OD}_{600}$  de 0.6 - 0.8, momento en el cual se consigue una densidad celular suficiente para la inducción de la expresión de la proteína de interés antes de que el cultivo llegue a fase estacionaria.

Previo a la inducción se toma una alícuota de 100  $\mu$ L de cada cultivo en un eppendorf, y se centrifuga 10 minutos a  $4^{\circ}\text{C}$ . Tras descartar el sobrenadante se conservan los pellets para usarlos posteriormente como control de la expresión a tiempo 0. La

expresión se induce mediante la adición al cultivo de IPTG (isopropil- $\beta$ -D-1-tiogalactopiranosido), análogo no hidrolizable de la alolactosa. La cantidad utilizada para la expresión de NtcA es de 0.5 mM mientras que la expresión de PipX se promueve con 0.25 mM en matraces Erlenmeyer de 1 L. Dicha variación en la concentración de IPTG se debe a la menor solubilidad de PipX; como norma general, a menor estabilidad de una determinada proteína a expresar, menor es la concentración de IPTG añadida, ya que así la síntesis de la proteína es más lenta y se facilita su plegamiento. Del mismo modo, con una disminución de la temperatura de incubación tras la inducción se consigue un efecto similar sobre la solubilidad. En nuestro caso, se incuban durante la noche a 18°C en el caso de NtcA y a 25°C en el caso de PipX.

Tras la expresión, se toma otra muestra de 100  $\mu$ L en un eppendorf de los cultivos inducidos y se procede de la misma forma que con las primeras alícuotas a tiempo 0. A continuación, se divide la muestra en dos botes de polipropileno de 1 L que se centrifugan a 4000 rpm durante 15 minutos a 4°C en la centrifuga Sorvall RC-5C PLUS. Después de descartar el sobrenadante se resuspenden los pellets en 25 mL de tampón A (ver apartado siguiente) y se centrifugan a 4000 rpm en tubos Falcon de 50 mL durante 30 minutos a 4°C en la centrifuga Eppendorf 5425R. Con el propósito de conservarlos para el proceso de purificación, tras descartar el sobrenadante se conservan los pellets con las células a -20°C.

#### Verificación de la expresión mediante electroforesis desnaturizante SDS-Page

El paso previo a la purificación de ambas proteínas es verificar que efectivamente se ha producido una inducción de la expresión bajo las condiciones utilizadas y que no ha habido ningún problema en su expresión. Para ello se compara mediante una electroforesis desnaturizante SDS-Page los controles recogidos antes y después de la inducción con IPTG. La separación se realiza en un gel de acrilamida 15 %, preparando los geles concentrador (stacking) y separador (separating) atendiendo a los reactivos de la Tabla 8. Se preparan las muestras resuspendiendo el pellet en 10  $\mu$ L en tampón de carga 2x (Tris-HCl, glicina, SDS y agua Milli-Q) y calentando durante 5 minutos a 100°C en el termobloque (Eppendorf). En el gel se cargan los 10  $\mu$ L de muestra y 3  $\mu$ L de patrón de proteínas PageRuler (Thermo Scientific). La electroforesis tiene lugar durante 50 minutos a 200 V. Seguidamente se tiñe el gel con Azul de Coomassie G250 durante 10 minutos para poder observar las bandas correctamente después de desteñirlo con una solución que contiene 10% EtOH absoluto y 10% ácido acético glacial.

**Tabla 8. Componentes del gel de poliacrilamida**

SEPARATING GEL 10%		STACKING GEL	
REACTIVOS	VOLUMEN	REACTIVOS	VOLUMEN
Acrilamida 40 %	1.25 mL	Acrilamida 40 %	0.23 mL
Agua Milli-Q	2.45 mL	Agua Milli-Q	1.5 mL
Separating gel 4 %	1.25 mL	Stacking gel 4 %	0.5 mL
APS 10 %	56 $\mu$ L	APS 10 %	14 $\mu$ L
TEMED	2.5 $\mu$ L	TEMED	14 $\mu$ L
<b>Total</b>	<b>3.8835 mL</b>	<b>Total</b>	<b>2.258 mL</b>



### 3.10.2. Purificación de NtcA y PipX mediante cromatografía de afinidad

La purificación de NtcA se realiza por cromatografía de afinidad en una columna HisTrap HP (1-5 mL) (GE Healthcare) montada en un sistema FPLC (ÄKTA). La elección de dicha columna se debe a que NtcA posee una etiqueta de polihistidinas en su región N-terminal que se une con gran afinidad al níquel de dicha columna.

En primer lugar, se descongelan los pellets de la expresión conservados a  $-20^{\circ}\text{C}$  y se resuspenden mediante agitación en medio de sonicado para NtcA (25 mM citrato sódico pH 6.5, 0.5 M NaCl, 5 mM  $\text{MgCl}_2$ , 1 mM DTT, 1 mM PMSF y 5 mM imidazol) previamente filtrado. En este, como se hará en el resto de tampones, se adiciona en último lugar y justo antes de utilizarlo el DTT (para evitar su oxidación) y el inhibidor de proteasas (PMSF), pues tiene una vida media corta.

El contenido se vierte en un recipiente metálico para proceder al lisado de las células mediante sonicación en el Soniprep 150 (MSE-SANYO), realizando 18 ciclos (15"ON-45"OFF-15"ON) de 11  $\mu\text{m}$  de amplitud, manteniendo la muestra en hielo y mezclándola entre ciclos con una pipeta Pasteur. Seguidamente, se centrifuga la muestra sonicada en la centrífuga Eppendorf 5810R durante 30 minutos a 15000 rpm y  $4^{\circ}\text{C}$  de temperatura. En paralelo se extrae una alícuota de 50  $\mu\text{L}$  del sonicado y se centrifuga a 15000 rpm durante 30 minutos a  $4^{\circ}\text{C}$  en la centrífuga de mesa Eppendorf 5425R. En ambos casos, de este modo se consigue separar las proteínas solubles (sobrenadante) del resto de componentes celulares no solubles presentes en la muestra (pellet). Después de centrifugar, se recupera el sobrenadante que contiene la proteína y se pasa por un filtro de jeringa de nitrocelulosa de 0.45  $\mu\text{m}$  de poro para eliminar restos indeseados que podrían obturar la columna. Por otra parte, el sobrenadante y pellet provenientes de la alícuota se conservan en hielo hasta su análisis mediante electroforesis SDS-Page. De este modo se consiguió separar las proteínas solubles del resto de componentes celulares no solubles presentes en la muestra. Paralelamente, se extrae una alícuota de 50  $\mu\text{L}$  del sonicado con el propósito de tener un control de la eficiencia de sonicación y estimar el grado de solubilidad.

Antes de someter la muestra a cromatografía de afinidad con iones de níquel a los que se unirá NtcA mediante las colas de poli-histidinas, se lava la columna con 5 CV de  $\text{H}_2\text{O}$  Milli-Q y 5 CV de tampón A de NtcA (25 mM HEPES pH 7.5, 0.5 M NaCl, 5 mM  $\text{MgCl}_2$ , 20 mM imidazol). La muestra se introduce en la columna HisTrap HP mediante el uso una bomba peristáltica Gilson Minipuls 3 a un flujo de 5 mL/min, tras lo cual la columna se monta en el sistema FPLC (ÄKTA) y se realiza un nuevo lavado con varios volúmenes de columna con tampón A hasta obtener un eluido cuya absorbancia a 280 nm sea constante y baja.

La elución de la proteína de interés se produce al aumentar la concentración de imidazol, que compite con las colas de histidinas por la unión a la columna de níquel. Para ello se crea un gradiente de concentración con soluciones tamponadas frías con tampón A (que posee 25 mM de imidazol) y tampón B de NtcA (25 mM HEPES pH 7.5, 0.5 M NaCl, 5 mM  $\text{MgCl}_2$ , 500 mM imidazol). Durante el proceso se monitoriza la absorbancia a 280 nm del efluente, longitud de onda de absorción máxima de las proteínas, y se recogen fracciones de 2 mL situadas en un colector sobre hielo fundente. Se seleccionan las fracciones de interés correspondientes con los picos observados en el cromatograma y se caracteriza su pureza en NtcA mediante electroforesis SDS-PAGE en geles de 15% de

poliacrilamida, discerniendo qué fracciones contienen proteínas del tamaño esperado de NtcA (~25 kDa).

A continuación, se realiza una cuantificación mediante el método de Bradford (A=595 nm) para determinar la concentración de proteína de forma reproducible y rápida de cada fracción de interés. Para ello, se elabora una curva patrón midiendo la absorbancia a dicha longitud de onda con diferentes concentraciones de BSA (0, 2, 4, 6 y 8 µg) en espectrofotómetro DR 1900 (Hach), añadiendo en cada caso 1 mL de reactivo de Bradford (Bio-Rad) e incubando 10 minutos a temperatura ambiente. Una vez realizada la curva patrón, se añadió 1 mL de reactivo de Bradford a una cantidad determinada de muestras (rango entre 2 y 20 µL) y se mide su absorbancia a 595 nm para poder estimar la concentración de proteína en cada fracción.

Teniendo en cuenta las cantidades de proteína en cada fracción de interés se cargan en el gel de acrilamida además de las muestras del pellet del sonicado, sobrenadante del sonicado, no retenida en la columna y lavado de la columna. En general, se depositan entre 2 µL y 12 µL de muestra junto a la cantidad de tampón de carga 2X correspondiente y llevando la mezcla a 10 µL o 15 µL de forma que cantidad de proteína total en las fracciones cargadas sean lo más similares posible.

Atendiendo a los resultados obtenidos en el gel de acrilamida, se juntan las fracciones de proteína de 2 mL que contienen NtcA pura y se concentra la proteína mediante ultrafiltración en la centrífuga Eppendorf 5810R a 4°C y 4000 rpm con filtros Amicon Ultra (Millipore) de corte nominal 10 kDa hasta un volumen de 3 mL alcanzando un compromiso entre solubilidad y concentración de proteína. Debido a la mezcla de fracciones se calcula que la concentración de imidazol en la muestra es de 308 mM. Una vez concentrada la muestra se resuspende en tampón de conservación hasta 15 mL reduciendo la concentración de imidazol hasta 60 mM y se adiciona glicerol hasta una concentración del 10%, que funciona como crioprotector. Igualmente se cuantifica la concentración de proteína por Bradford y se realizan alícuotas de 25, 50, 100 y 200 µL que se congelan en nitrógeno líquido para conservarlas a -80°C para ensayos posteriores.

Para la purificación de PipX se realizaría el mismo procedimiento, aunque variando algunos componentes de los tampones de sonicado y conservación que favorecen a la estabilidad de la proteína debido a su mayor insolubilidad comparada con NtcA. En ambos tampones, se utilizaría tampón MES pH 6 (50mM) en sustitución al tampón Na Citrato, además de añadir glicerol (10% v/v) con el objeto de aumentar la estabilidad de la proteína. En cuanto a la etapa de concentración de la proteína una vez purificada, se debería utilizar un dispositivo Amicon de un corte nominal menor a 10 kDa ya que PipX tiene un tamaño de 10.5 kDa, y por lo tanto se perdería una gran parte de proteína.

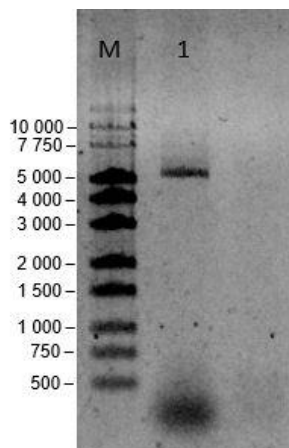
## 4. RESULTADOS Y DISCUSIÓN

### 4.1. Eliminación de una mutación en el gen *rpoC2*

#### 4.1.1. Mutagénesis de sitio dirigido (SDM)

Como se ha explicado en el apartado de Materiales y métodos, se intenta eliminar la inserción indeseada en el gen mediante SDM, haciendo varias pruebas en las que se varía el tipo de cebador utilizado (largo o corto), la concentración de DMSO, las concentraciones de DNA molde (154 ng; 51 ng) y los parámetros de amplificación en la reacción de PCR, cambiando tanto la temperatura de hibridación como el número de ciclos.

Tras la digestión con DpnI y la ligación del plásmido, se examinan los resultados mediante electroforesis en gel de agarosa. Utilizando los cebadores largos, solo se obtiene producto de PCR cuando la temperatura de hibridación se disminuye de 72 a 68 °C para que no coincida con la temperatura de extensión, eso sí, de un tamaño aparentemente inferior al deseado, pero que podría ser el correcto debido a que al fragmento de DNA no es lineal sino un plásmido recircularizado (Figura 7). Sin embargo, tras proceder a la ligación y pese a que se transforma *E.coli* con mayor cantidad de la habitual de producto de digestión (10 µL), no aparecen colonias. Por esta razón, se aumenta hasta 40 el número de ciclos de amplificación con la intención de incrementar la cantidad de producto final y favorecer la transformación. Aunque en el gel de agarosa (no mostrado) aparecen además otras bandas anchas inespecíficas decidimos transformar en *E. coli* TopTen, también con resultado negativo.

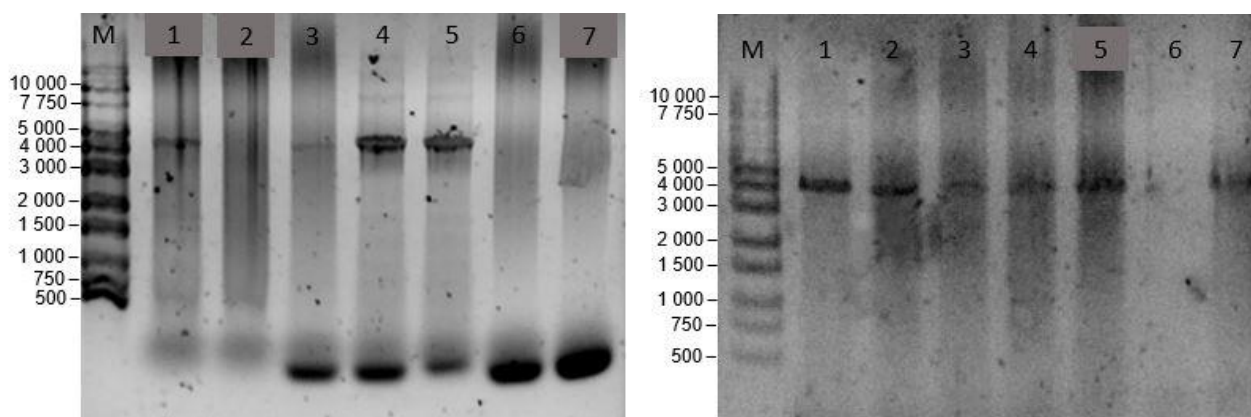


**Figura 7. Gel de agarosa 0,8% del producto de PCR de la mutagénesis de *rpoC2* ( $\beta'$ ).** M: 10 µL de marcador de peso molecular 1 kb (pPSU), 2: 5 µL de producto de PCR.

Finalmente tras un nuevo intento utilizando esta vez el tratamiento KLD del Kit Q5 Site-directed mutagenesis, que combina digestión con DpnI, fosforilación y ligación del producto de PCR obtuvimos 10 transformantes. Desafortunadamente, y tras el aislamiento de los plásmidos observamos mediante secuenciación que ninguna presentaba la mutación corregida. Por otro lado, tras la PCR usando los cebadores cortos y el tratamiento KLD del Kit Q5 Site-directed mutagenesis, de nuevo, se obtiene una banda observada en el gel (no mostrada) inferior a la esperada de 8,7 kb, pero que podría ser correcta debido a que el plásmido está circularizado (DNA superenrollado) y por este motivo se decide transformar aunque tampoco se obtienen colonias tras la transformación.

#### 4.1.2. Clonación del gen a partir del plásmido original con rpoC2

Dados los resultados negativos, se decide, como se explica en Materiales y métodos, intentar la amplificación del gen rpoC2 a partir del plásmido original pLIC. La primera prueba, que se lleva a cabo utilizando 6 condiciones (Figura 8), resulta positiva al observar una banda esperada para un gen de tamaño 3957 kb en el gel de electroforesis en las condiciones con 0.5  $\mu$ L de DMSO 100%. En esta ocasión utilizamos como control positivo una condición sin DMSO (que es la que aparentemente genera la inserción indeseada). También se aumenta la temperatura de extensión para incrementar la especificidad y evitar la formación de estructuras secundarias. Para ello se hacen dos grupos a 72°C y 74°C de temperatura de extensión. Tras un primer intento en el que no hay resultados positivos, decidimos repetir el experimento porque solo hay bandas en algunas de las carreras y/o son difusas, probablemente por un problema en el estado de los dNTPs. Tras cambiar estos sí se obtienen las bandas esperadas en todos los casos como se observa en la Figura 9.



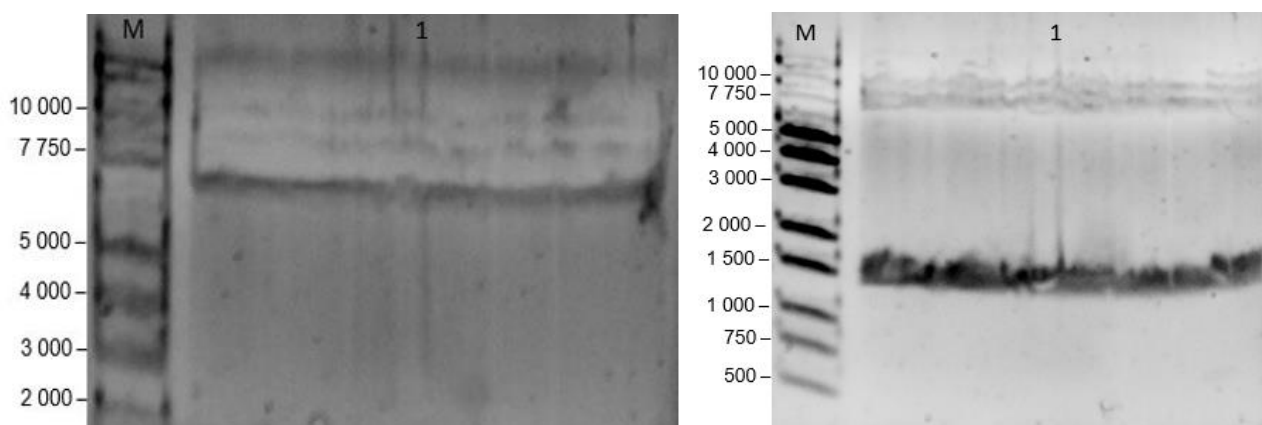
**Figura 8 (Izq.). Gel de agarosa 0.8 % de la PCR de rpoC2 (6') I.** M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1: Control positivo, 2-6: 5  $\mu$ L de producto de PCR (2: 10 ng/ $\mu$ L de DNA molde + 0.5  $\mu$ L de DMSO; 3: 10 ng/ $\mu$ L de DNA molde + 1  $\mu$ L de DMSO; 4: 20 ng/ $\mu$ L de DNA molde; 5: 20 ng/ $\mu$ L de DNA molde + 0.5  $\mu$ L de DMSO; 6: 20 ng/ $\mu$ L de DNA molde + 1  $\mu$ L de DMSO, 7: Control negativo.

**Figura 9 (Der.). Gel de agarosa 0.8 % de la PCR de rpoC2 (6') II.** M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1-5,7: 5  $\mu$ L de producto de PCR (20 ng/ $\mu$ L de DNA molde + 0.5  $\mu$ L), 6: Control negativo.

Se purifican las muestras cargadas en este gel junto con las pocas muestras que mostraron una banda clara en el gel del primer intento y poseían DMSO. Tras la purificación se obtiene una concentración de 220 ng/ $\mu$ L y un ratio  $A_{260/280}$  de 1.96 que se secuencia con el cebador PIB3 (Anexo II). Mientras se espera a los resultados de la secuenciación, se realiza la digestión del producto de PCR con los enzimas HindIII y NotI y se purifica consiguiendo una concentración de 157 ng/ $\mu$ L y un ratio  $A_{260/280}$  de 1.84. Desafortunadamente, los resultados frustrados de la secuenciación (Anexo III), que indican que esta muestra también contiene la mutación, nos obligan a cambiar de estrategia durante el ensayo de ligación. Es entonces cuando decidimos revisar la secuencia del plásmido original pLIC-rpoC2 mediante secuenciación con el cebador interno PIB3 y se constata que la mutación ya estaba en origen. En este momento y dadas las dificultades, se decide variar la estrategia para eliminar la mutación mediante su sustitución con un gen artificial.

#### 4.1.3. Eliminación de la mutación de *rpoC2* mediante su sustitución con un gen artificial

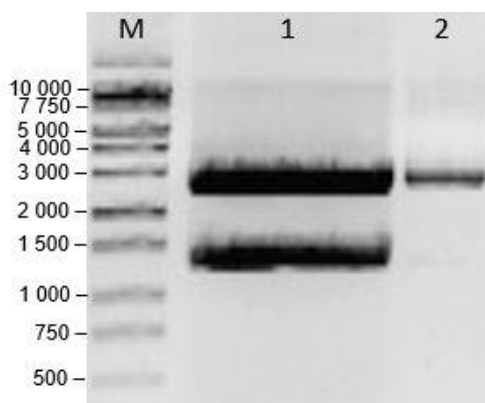
El primer paso para la corrección de la mutación es la digestión del constructo pQLinkN-*rpoC2*, que contiene la mutación, con los enzimas de restricción BbvCI y NgoMIV, según se ha descrito en Materiales y métodos. La electroforesis se realiza durante 1 hora a 100 V (el doble de tiempo habitual) y tal y como se espera, se obtiene una banda intensa que corresponde a pQLinkN-*rpoC2* sin el fragmento de 1438 pb que contiene la inserción (7327 pb). La banda que coincide con el fragmento de 1438 pb escindido no aparece en el gel de la Figura 10, pero previamente nos cercioramos de su existencia tomando una imagen del gel tras sólo 30 minutos de tiempo de electroforesis (Figura 11). En la Figura 10 puede observarse además la aparición de varias bandas menos intensas mayores a 7750 pb que posiblemente correspondan al plásmido sin digerir, o escindido por uno solo de los enzimas (el plásmido lineal es de 8763 pb), por lo que concluimos que deberíamos haber empleado más tiempo en la reacción de digestión del constructo. Por esta razón se incrementó el tiempo de electroforesis, para facilitar la purificación posterior de la banda de interés (la banda inferior de las mostradas en la Figura 10) a partir de gel y separarla del resto.



**Figura 10 (Izq.).** Gel de agarosa 0.8 % de la digestión con BbvCI y NgoMIV de pQLink N-*rpoC2* (6') tras 1 hora de electroforesis. M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1: 50  $\mu$ L de digestión de pQLink N-*rpoC2*.

**Figura 11 (Der.).** Gel de agarosa 0.8 % de la digestión con BbvCI y NgoMIV de pQLink N-*rpoC2* (6') tras 30 minutos de electroforesis. M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1: 50  $\mu$ L de digestión de pQLink N-*rpoC2*.

Una vez seleccionada, se purifica la sección de 7327 pb consiguiendo una muestra de 150  $\mu$ L a una concentración de 4,7 ng/ $\mu$ L (valor normal teniendo en cuenta que se recupera sobre un 10% con la purificación de gel) y elevada pureza a juzgar por su  $A_{260/280}$  de 1.87. Del mismo modo se procede tras recibir el gen artificial clonado en un plásmido pUC19; tras transformar *E. coli* con el mismo, se procede a purificar más cantidad de plásmido, obteniéndose dos muestras de 50  $\mu$ L a 151 ng/ $\mu$ L y 200 ng/ $\mu$ L. Posteriormente, se digiere con las enzimas NgoMIV y BbvCI para liberar el fragmento de gen libre de la mutación, tal y como puede observarse por la aparición de una banda a la altura teórica esperada de aproximadamente 1389 pb (Figura 12). Se purifica dicha banda desde el gel obteniéndose el DNA a una concentración de 9 ng/ $\mu$ L y un ratio  $A_{260/280}$  1.76.



**Figura 12.** Gel de agarosa 0,8 % de la digestión con *BbvCI* y *NgoMIV* de *pUC19* con el gen artificial de 1389 pb. M: 10 µL de marcador de peso molecular 1 kb (*pPSU*), 1: 31 µL de digestión de *pUC19* con el gen artificial, 2: 100 ng de plásmido *pUC19* con el gen artificial sin digerir

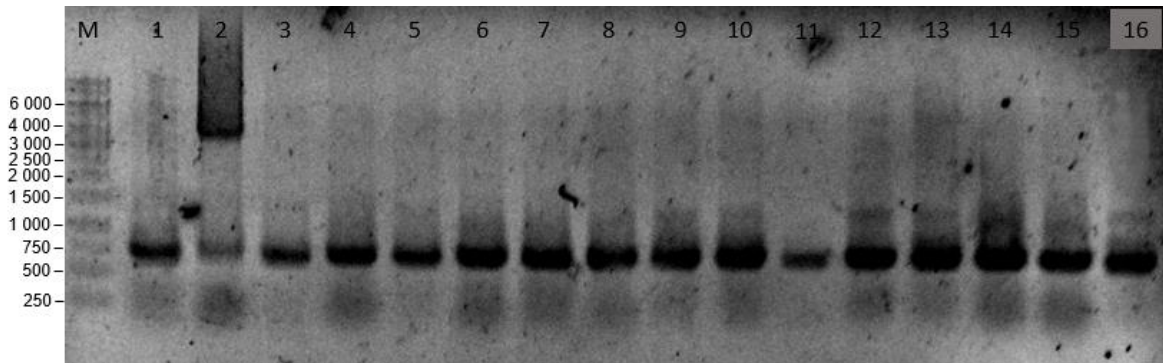
Estos dos fragmentos resultantes de la digestión *BbvCI*-*NgoMIV* se ligan y se procedería a transformar *E. coli* con el producto de ligación. Por falta de tiempo, esta etapa del proceso no pudo realizarse, pero de haberlo hecho hubiésemos obtenido casi con toda seguridad el constructo *pQLinkN-rpoC2* libre de mutaciones y se podría proceder a la última etapa de clonaje del gen *rpoC2* en los constructos *pQLink-rpoD1-rpoZ-rpoC1*, *pQLink-rpoD2-rpoZ-rpoC1* y *pQLink-rpoD5-rpoZ-rpoC1* (ver próximo apartado).

#### 4.2. Clonación en tándem de los distintos genes en los plásmidos *pQLink*

En este apartado se muestran los resultados de las fusiones realizadas con las subunidades correspondientes para obtener el plásmido *pQLink* definitivo. Aunque se deseaba situar el factor  $\sigma$  en la 1ª posición, por ser la subunidad variable de la RNAP, con la intención de minimizar el número de pasos y así, una vez obtenido el núcleo  $\beta'$ - $\omega$ - $\gamma$ , solamente se necesitaba incorporar cada subunidad  $\sigma$  ( $\sigma_A$ ,  $\sigma_B$  o  $\sigma_C$ ) a este núcleo, no pudo ser así. Esto fue debido a los problemas para eliminar la mutación que contenía el gen *rpoC2*, pues en el momento de la clonación en tándem aún no se había podido corregir. Se decidió clonar entonces *rpoD1*, *rpoD2* y *rpoD5* en 2ª posición junto a *rpoZ* y *rpoC1* (3ª y 4ª posición, respectivamente), por lo que el gen *rpoC2* ( $\beta'$ ) quedará en 1ª posición.

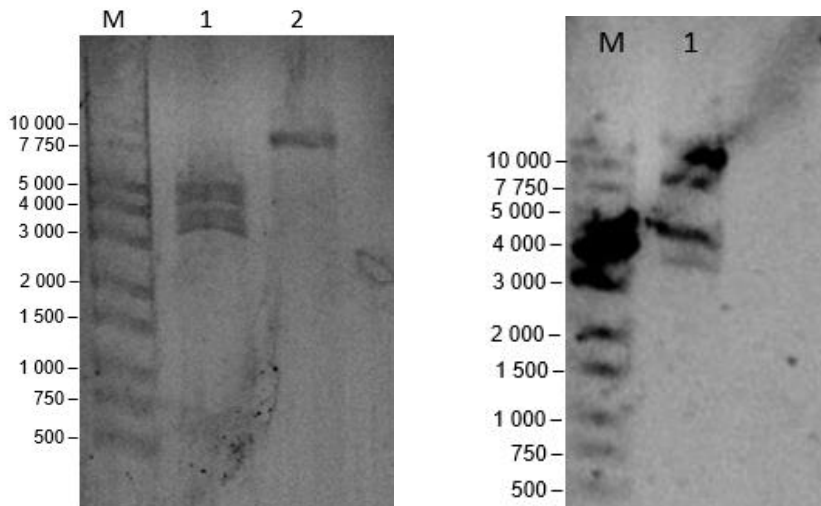
En primer lugar, se efectúa la clonación en tándem de los genes que quedarán en tercera y cuarta posición, esto es, *rpoZ* ( $\omega$ ) y *rpoC1* ( $\gamma$ ) y se transforma el producto de reacción en *E. coli* (ver apartado 3.2). Tras un cribado mediante PCR de colonias (ver apartado 3.3) se obtiene una única colonia con el tamaño de amplicón esperado de aproximadamente 3300 pb (Figura 13). El plásmido de esta colonia se purifica por Miniprep (ver apartado 3.5) consiguiendo una concentración de 100 ng/µL y un ratio  $A_{260/280}$  de 1.82. Después, se envía a secuenciar utilizando el cebador *pQTEV3U* (mismo que para la PCR) ya que se sitúa en posición "upstream LINK1", donde empieza omega. Puesto que omega (*rpoZ*) es de pequeño tamaño (231 pb) y la longitud de la lectura es superior, se debe secuenciar también parte de gamma (1875 pb), verificando de esta forma la presencia de ambas. Sin embargo, no es así, y solo aparece la secuencia de *rpoZ*, seguramente debido a que la etiqueta GST ( $\approx$  800 pb) se encuentra en la región N-terminal de *rpoC1* (además de

la región intergénica en el plásmido). Por tanto, se secuencian la misma muestra con el cebador interno de *rpoC1*, PIG1 (Anexo II) y se confirma su presencia (Anexo III).



**Figura 13.** Gel de agarosa 0.8 % de la PCR de colonia del constructo pQLink N +  $\omega$  +  $\gamma$ . M: 4  $\mu$ L de marcador de peso molecular O'GeneRuler (1kb), 1-16: 5  $\mu$ L de producto de PCR de colonia de pQLinkN-rpoZ-rpoC1. Colonia 2 positiva.

Los pasos consiguientes son la unión del constructo pQLink-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ) obtenido con cada una de las subunidades  $\sigma$ . Dado que en el procedimiento de clonación en tándem (ver apartado 3.9.1) para la fusión tras la digestión previa se requiere 5  $\mu$ L de cada una de las mezclas de plásmidos, se utilizan los otros 5  $\mu$ L de la solución digerida de pQLink-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ) con *PacI* para cerciorarnos de nuevo de que realmente están bien clonados los genes de rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ). La banda correspondiente al inserto se encuentra con un tamaño mayor a lo esperado por la unión de los dos genes y esto es debido a que está también incluida la etiqueta GST y la región intergénica (Figura 14). Una vez confirmado se procede a la "fusión" del constructo pQLink-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ) con pQLinkN-rpoD1 ( $\sigma_A$ ). Desafortunadamente, tras la transformación de *E. coli* no aparecen colonias en las placas, así que utilizamos 5  $\mu$ L de la mezcla final para comprobar mediante un gel de agarosa que se haya producido la ligación (Figura 15).

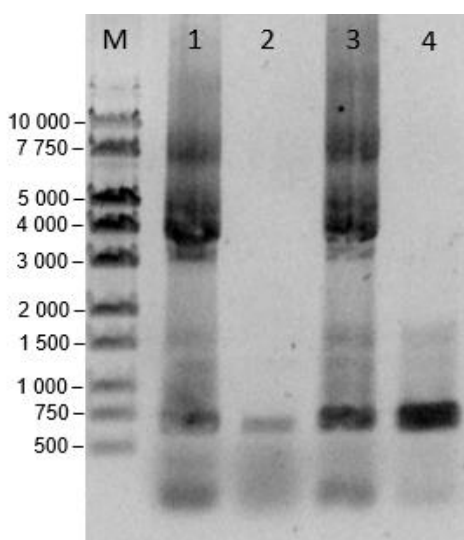


**Figura 14 (Izq.).** Comprobación del constructo rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ) mediante digestión con *PacI*. M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1: 5  $\mu$ L de producto de digestión con *PacI* de pQLink N +  $\gamma$  +  $\omega$ , 2: 5  $\mu$ L de plásmido pQLink N +  $\gamma$  +  $\omega$  sin digerir.

**Figura 15 (Der.).** Comprobación de la ligación del constructo rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ )- rpoD1 ( $\sigma_A$ ). M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1: 5  $\mu$ L de producto de fusión frustrada.

Observamos tres bandas cuyo tamaño coincide con el de los dos plásmidos individuales digeridos, pero no al del producto de “fusión”, por tanto la ligación no funcionó. Surgen dos posibles hipótesis para explicar los resultados: (1) no se deja el tiempo suficiente para enfriar a temperatura ambiente cuando se juntaron las dos soluciones tras la incubación a 65°C o (2) la transformación no funciona de forma correcta. Por este motivo, se repite el paso en el cual se calienta la mezcla a 65°C durante 4 minutos y se deja enfriar a temperatura ambiente el tiempo conveniente para favorecer el “annealing”. Paralelamente, se trata 13 µL de la mezcla con 1 µL de T4 DNA ligasa y se incuba a 15 °C durante 4 horas. En ambos casos se transforman 10 µL del producto obtenido tras la ligación por choque térmico y 3 µL por electroporación. Únicamente, se observa el crecimiento de dos colonias en cada una de las placas pertenecientes a la transformación mediante electroporación, por lo que para las “fusiones” del constructo pQLink-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ) con pQLinkN-rpoD2 ( $\sigma_B$ ) y con rpoD5 ( $\sigma_C$ ), respectivamente, las transformaciones de los productos de reacción se llevan a cabo en células *E. coli* DH5 $\alpha$  mediante electroporación, obteniéndose en ambos casos decenas de colonias.

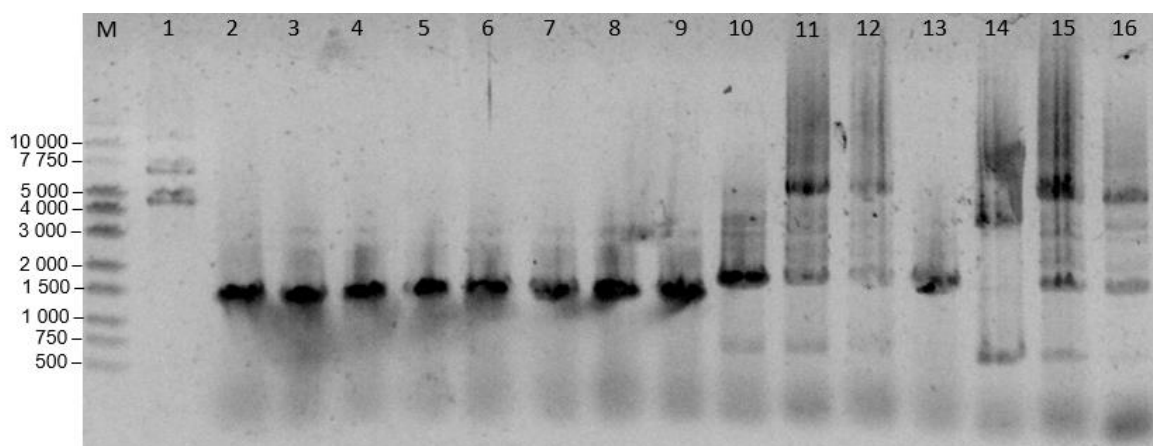
Para los tres posibles constructos, los transformantes se confirman mediante PCR de colonia seguida de una electroforesis de DNA en gel de agarosa 0.8 %. Como los genes son de longitud similar, los tamaños de inserto esperados son parecidos por lo que se utiliza en los tres casos el mismo tiempo de extensión de 3 minutos en la reacción de PCR.



**Figura 16. Gel de agarosa 0.8 % de la PCR de colonia del constructo pQLinkN +  $\sigma_A$  +  $\omega$  +  $\gamma$ .** M: 10 µL de marcador de peso molecular 1 kb (pPSU), 1-4: 5 µL de producto de PCR de colonia de pQLinkN-rpoD1-rpoz-rpoC1. Colonias 1 y 3 positivas.

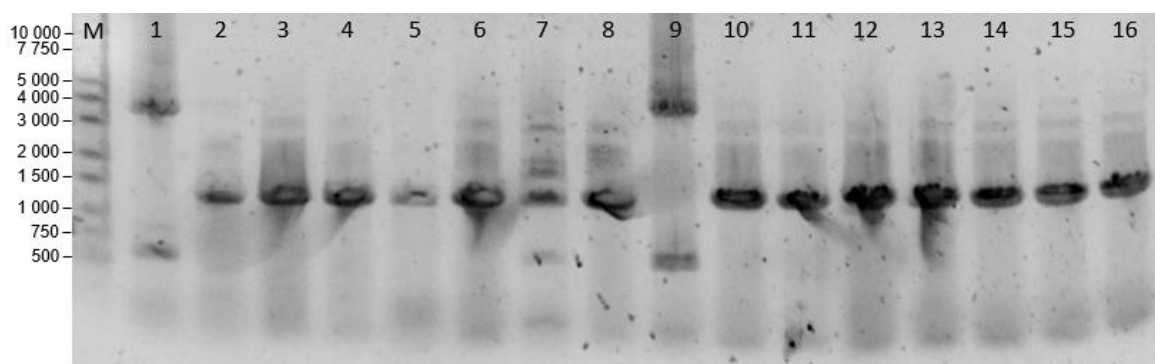
La aparición de bandas del tamaño esperado en el gel de agarosa de la Figura 16 confirma la presencia de rpoD1 ( $\sigma_A$ ) en el fragmento con los genes deseados en las colonias 1 y 3, pues las bandas tienen un tamaño aproximado de 4700 pb que coincide con la suma de  $\gamma$  (1875 pb),  $\omega$  (231 pb),  $\sigma_A$  (1200 pb), tag GST ( $\approx$  800 pb), regiones intergénicas y los fragmentos entre gen y sitio de hibridación de los cebadores. Después de la purificación por Miniprep (ver apartado 3.5) se logran plásmidos concentrados a 78 ng/µL y 51.5 ng/µL y ratios  $A_{260/280}$  de 1.78 y 1.96, respectivamente. Con la intención de confirmar que el inserto es realmente correcto y aprovechando que aún se le debe adicionar una subunidad más al plásmido (falta rpoC2), se digiere el constructo con PaeI. El fragmento liberado tiene el tamaño esperado (Figura 17) y además la presencia de las tres subunidades se verifica mediante secuenciación usando los cebadores pQTEV3U y PIG1.





**Figura 17.** Gel de agarosa 0.8 % de la digestión pQLinkN +  $\omega$  +  $\gamma$  +  $\sigma_A$  y PCR de colonia de los constructos pQLinkN +  $\omega$  +  $\gamma$  +  $\sigma_B$  y pQLinkN +  $\omega$  +  $\gamma$  +  $\sigma_C$ . M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1: 5  $\mu$ L de producto de PCR de colonia de la digestión con PaeI de pQLink N +  $\sigma_A$  +  $\gamma$  +  $\omega$ , 2-9: 5  $\mu$ L de producto de PCR de colonia de pQLink N +  $\sigma_B$  +  $\omega$  +  $\gamma$ , 10-16: 5  $\mu$ L de producto de PCR de colonia de pQLink N +  $\sigma_C$  +  $\omega$  +  $\gamma$ .

Por su parte, la comprobación de la fusión con rpoD2 ( $\sigma_B$ ) (el constructo pQLink-rpoD2 ( $\sigma_B$ )-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ )) se lleva a cabo mediante PCR en un total de 8 colonias. Los resultados, visualizados en un gel de agarosa (Figura 17) muestran bandas de unos de 1500 pb de longitud, aproximadamente, por lo que deben pertenecer a plásmidos no digeridos o religados (los plásmidos donantes de la subunidad rpoD2 que se han cortado con SwaI). Se repite la PCR de colonia con 16 colonias más y se disminuye también el tiempo de extensión a 2 minutos y 30 segundos (para evitar la aparición de bandas inespecíficas). Esta vez aparecen dos colonias (colonias 1 y 9) que pueden presentar el plásmido esperado correcto ya que la banda mayoritaria en el gel de agarosa es de  $\approx$  4400 pb (Figura 18). Los plásmidos se aíslan a concentraciones de 30.9 ng/ $\mu$ L y 40.5 ng/ $\mu$ L, con ratios  $A_{260/280}$  de 1.42 y 1.68 respectivamente.



**Figura 18.** Gel de agarosa 0.8 % de la PCR de colonia del constructo pQLink N +  $\omega$  +  $\gamma$  +  $\sigma_B$ . M: 10  $\mu$ L de marcador de peso molecular 1 kb (pPSU), 1-16: 5  $\mu$ L de producto de PCR de colonia de pQLinkN-rpoD2-rpoZ-rpoC1. Colonias 1 y 9 positivas.

En lo que respecta al plásmido de fusión con rpoD5 ( $\sigma_C$ ) (el constructo pQLink-rpoD5 ( $\sigma_C$ )-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ )), se distinguen 4 colonias que podrían ser positivas, de un total de 7 colonias analizadas (Figura 17). Las colonias 10 y 13 parecen contener plásmidos religados como ocurre en la “fusión” con rpoD2 y las colonias 11, 12, 14 y 15 presentan el amplicón del tamaño esperado de  $\approx$  4700. Se purifican mediante Miniprep (ver apartado 3.5) los plásmidos pertenecientes a las colonias 11 y 12 de la fusión pQLink-rpoD5 ( $\sigma_C$ )-rpoZ ( $\omega$ )-

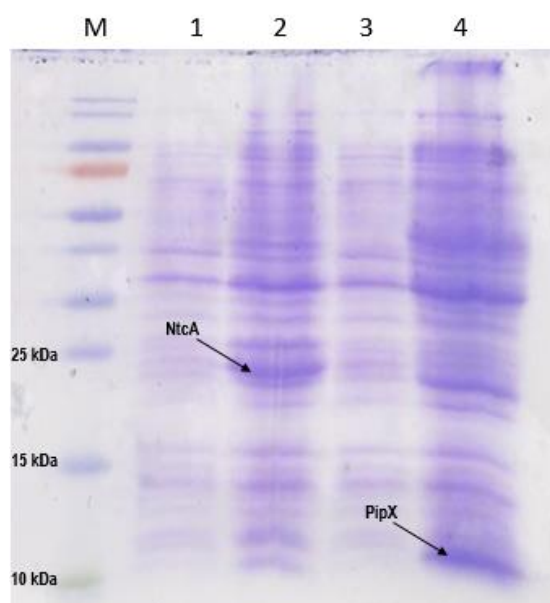
rpoC1 ( $\gamma$ ) (la banda obtenida en el gel era más clara e intensa que las de las colonias 15 y 16, ver Figura 17) se purifican obteniendo dos tubos de 34 ng/ $\mu$ L y 30.9 ng/ $\mu$ L, con un ratio de absorbancia  $A_{260/280}$  de 2.08 y 1.76. Al igual que en los casos anteriores, se secuencian con el cebador pQTEV3U, por el lado de la subunidad rpoD ( $\sigma$ ) y con el primer interno de  $\gamma$ , PIG1 para confirmar que el constructo es el correcto.

Por tanto, a falta de confirmar mediante secuenciación, los resultados obtenidos hasta el momento son compatibles con los tamaños esperados para los constructos pQLink-rpoD2 ( $\sigma_B$ )-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ) y pQLink-rpoD5 ( $\sigma_C$ )-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ). Confiamos haber obtenido estos dos constructos además del ya confirmado pQLink-rpoD1 ( $\sigma_A$ )-rpoZ ( $\omega$ )-rpoC1 ( $\gamma$ ), por lo que únicamente faltaría la clonación del gen rpoC2 ( $\beta'$ ) en la posición 1 en cada uno de los tres constructos obtenidos para empezar a realizar ensayos de expresión y purificación de la RNAP de *Synechococcus elongatus*.

### 4.3. Expresión y purificación de las proteínas NtcA y PipX

#### 4.3.1. Expresión de NtcA y PipX

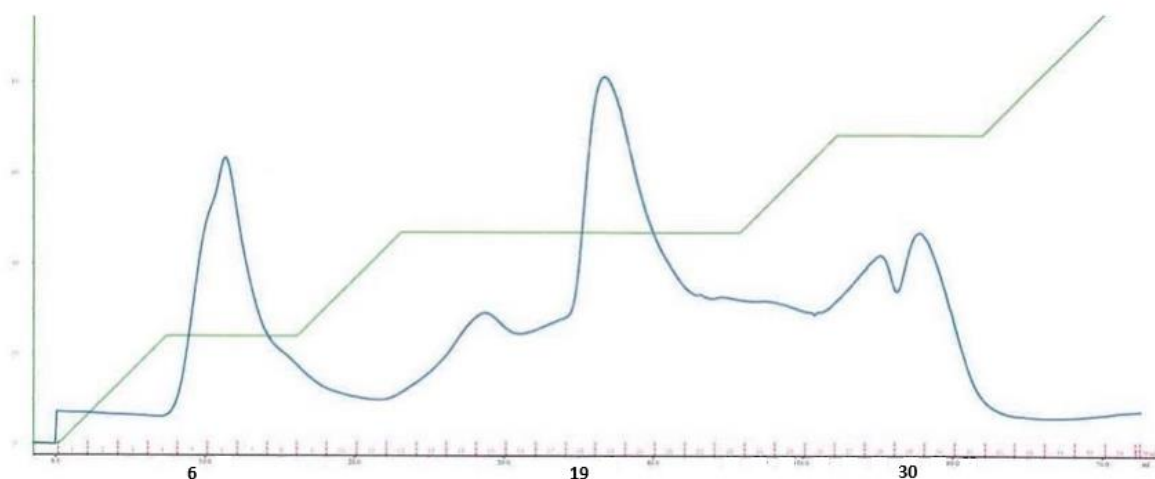
Los resultados derivados de la inducción de la expresión de NtcA y PipX mediante IPTG se observan en la Figura 19 correspondiente a la electroforesis desnaturizante SDS-Page en gel de poliacrilamida 10 %. Cabe destacar que NtcA es una proteína dimérica, está formada por dos monómeros de 25 kDa, sin embargo, solo se pueden observar los monómeros a 25 kDa debido a la electroforesis desnaturizante que se ha llevado a cabo. En el carril 2, correspondiente a la expresión de NtcA tras la inducción con IPTG, se distingue una banda que destaca a  $\approx$ 25 kDa que coincide con el tamaño de NtcA junto a la cola de polihistidinas (etiqueta para la posterior purificación) en comparación con la muestra de la expresión de NtcA sin inducir con IPTG en el carril 1. Lo mismo ocurre para PipX, cuya expresión después de la inducción con IPTG en el carril 4 se aprecia de forma pronunciada en una banda a 11-12 kDa correspondiente con el tamaño de la proteína PipX junto a la cola de polihistidinas.



**Figura 19.** Gel de poliacrilamida 10 % de la expresión de las proteínas NtcA y PipX antes y después de inducir la expresión con IPTG. M: 3  $\mu$ L de marcador PageRuler™ (Thermo Scientific), 1: 10  $\mu$ L de NtcA antes de inducir con IPTG, 2: 10  $\mu$ L de NtcA tras inducir con IPTG, 3: 10  $\mu$ L de PipX antes de inducir con IPTG, 4: 10  $\mu$ L de PipX tras inducir con IPTG.

#### 4.3.2. Purificación de NtcA mediante cromatografía de afinidad

Tal y como se ha descrito en material y métodos, el pellet de células de la expresión de NtcA se resuspende en tampón de sonicado, y tras el lisado de las células, la fracción soluble se carga en una columna de níquel His-trap. Tras el lavado de la columna mediante tampón A, se procede a eluir NtcA. En la Figura 20 se observan las diferentes fracciones de proteína eluida obtenidas tras haber finalizado el gradiente de imidazol creado en el sistema FPLC a diferentes concentraciones de tampón A y B de NtcA.



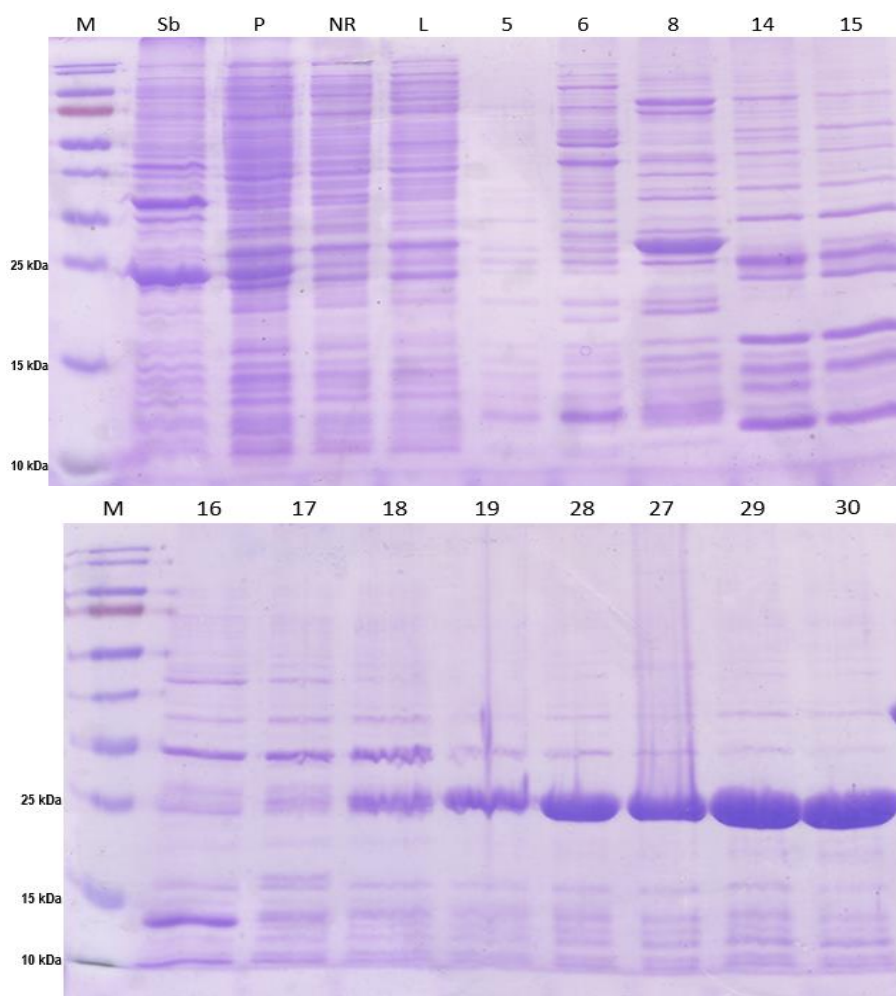
**Figura 20.** Gráfica de A<sub>280</sub> de las diferentes fracciones obtenidas a diferentes concentraciones de tampón B. En el eje Y se representa el porcentaje de tampón B, en el eje X se representan las diferentes fracciones de 2 mL. Los picos corresponden a la absorbancia obtenida para cada fracción.

En este gráfico se representa la A<sub>280</sub> para las diferentes fracciones adquiridas. Destacan tres picos de absorbancia A<sub>280</sub> pertenecientes a las fracciones 6, 19 y 30 que podrían corresponderse con la proteína NtcA. Por ello, se estima la concentración de proteína mediante el método Bradford tal y como se explica en Materiales y métodos de estas tres fracciones junto con la de las alícuotas tomadas del pellet del sonicado, el sobrenadante de sonicado, la fracción no retenida de la columna y el lavado de la columna. La concentración de proteína se determina midiendo la A<sub>595</sub> (Tabla 9) e interpolando con los valores de la curva patrón ( $y=0,0532x + 0,0018$ ) realizada con diferentes concentraciones de BSA (0, 2, 4, 6, 8  $\mu$ g).

**Tabla 9.** Volumen de muestra utilizado para la cuantificación, absorbancia y concentración de proteína obtenida mediante el método de Bradford.

	Volumen ( $\mu$ L)	A <sub>595</sub>	[Proteína] (mg/mL)
No retenido	1	0.293	5.47
Pellet	1	0.143	2.66
Sobrenadante	0.8	0.200	5.13
Lavado	1	0.305	5.69
Fracción 6	1	0.158	2.94
Fracción 19	3	0.124	0.77
Fracción 30	1	0.204	3.80

Estos datos nos permiten inferir cómo se ha desarrollado el proceso de purificación y si ha ocurrido algún contratiempo. Se observa una alta concentración de proteína total en el pellet del lisado, lo que podría indicar que o bien nuestra proteína de interés es muy insoluble o bien que deberían haberse realizado más ciclos de sonicación. Llegamos a la conclusión de que es principalmente debido a lo segundo ya que tras observar el resultado del gel de poliacrilamida (Figura 21), se ve que la mayor parte de NtcA se encuentra en la fracción soluble. Por otra parte, las fracciones 6 y 30 contienen una concentración bastante alta de proteína, que pueden corresponderse con NtcA, pues en las purificaciones de esta proteína se obtienen concentraciones elevadas (Llácer, 2012; Forcada-Nadal, 2018). La presencia de NtcA se confirma mediante la observación de las distintas fracciones en el gel de poliacrilamida 10 % realizado en condiciones desnaturalizantes (SDS-Page). La banda esperada 25-26 kDa que se corresponde con NtcA no se observa de manera pronunciada en las fracciones 5-18. Sin embargo, a partir de la fracción 19 se aprecia una banda evidente y en las fracciones 27-30 se contempla nítidamente sin apenas contaminación. Entonces, se decide seleccionar las fracciones de 2 mL de los tubos 23-32, que se juntan y se lleva a cabo la concentración de la proteína mediante ultrafiltración mediante el filtro Amicon Ultra de 10 kDa. El volumen total de 20 mL se reduce hasta un volumen de 11 mL logrando una concentración de proteína de 3.5 mg/mL. Por tanto, se consiguen 38.5 mg de NtcA pura, que es más que suficiente para los ensayos requeridos.



**Figura 21. Gel de poliacrilamida 10 % de la purificación de NtcA.** M: 3  $\mu$ L de marcador PageRuler<sup>TM</sup> (Thermo Scientific), Sb: Sobrenadante del sonicado, P: Pellet del sonicado, NR: No retenido en la columna, L: Lavado de la columna, 5-30: Fracciones de interés.

A pesar de no poder purificar la proteína PipX por falta de tiempo, en conformidad con los experimentos realizados por Llácer et al. (2010) para la obtención de PipX purificada, es muy posible que la cantidad total y concentración tras el procedimiento de purificación sea más baja que la de NtcA debido a su alta insolubilidad. En cualquier caso la cantidad obtenida será más que suficiente para los estudios de interacción y de cryo-EM del complejo RNAP-NtcA-PipX-Promotor. De hecho para la formación del complejo la proteína limitante no será PipX, sino la RNAP debido a la gran complejidad de su expresión y purificación a tenor de lo explicado en el siguiente apartado.

#### 4.4. Perspectivas de futuro

Tras la obtención de los tres constructos pQLink con las cuatro subunidades, descritos en el apartado 4.2 (o tras la obtención de al menos uno de ellos) la siguiente fase del proyecto correspondería a la expresión y purificación de las tres variantes de la RNAP, para luego realizar el ensamblaje del complejo RNAP-NtcA-PipX-DNA para su estudio mediante Cryo-EM.

##### 4.4.1. Co-expresión y purificación de una RNAP cianobacteriana

Como ya se ha mencionado en la Introducción, el holoenzima de la RNAP bacteriana se compone de seis subunidades  $\alpha_2\beta\gamma\beta'\omega$ , junto a una subunidad sigma (en nuestro caso  $\sigma_A$ ,  $\sigma_B$  ó  $\sigma_C$ ). Una vez finalicemos con la clonación de las 4 subunidades en los plásmidos pQLink, el siguiente paso será co-transformar células *E. coli* BL21 (cepa deficiente en proteasas, ideal para expresión) con cada uno de los tres plásmidos pQLink-rpoC2-rpoD (1, 2 ó 5)-rpoZ-rpoC1 y el plásmido pETCola-rpoA ( $\alpha$ )-rpoB ( $\beta$ ) y se podrá empezar a realizar pruebas de expresión de la RNAP.

Para ello se cultivarán las células teniendo en cuenta que el vector pETCola tiene resistencia a kanamicina y pQLink tiene resistencia a ampicilina. Tras la inducción de los cultivos, en el caso de que alguna de las proteínas presente problemas de bajos niveles de expresión, se co-transformará *E. coli* Rosetta (Novagen) con ambos plásmidos. Esta cepa de *E. coli*, posee un plásmido adicional con resistencia a cloranfenicol que codifica para tRNAs que reconocen codones poco frecuentes en *E. coli*.

Las pruebas de expresión se desarrollarán mediante la inducción durante al menos 3 horas de la expresión génica mediante el uso de IPTG. De forma habitual y para proteínas con gran solubilidad, estas se expresan a 37°C y usando una concentración de IPTG de 1 mM. Anticipamos que no va a ser el caso con la RNAP y probablemente sea poco soluble, por lo que la temperatura de expresión se deberá variar entre los 15°C y 37°C y también la concentración de IPTG, que se disminuirá. Se ajustará el tiempo de inducción atendiendo a la temperatura elegida, hasta un máximo de 16 horas si la temperatura es de 15°C. Si surgen más problemas de solubilidad añadiremos chaperonas moleculares o induciremos las carabinas propias de *E. coli* mediante shocks combinados de etanol/frío para favorecer la estabilidad de las diferentes subunidades.

El paso previo a la purificación es la verificación de la expresión mediante electroforesis desnaturante SDS-Page. Si se pueden apreciar las bandas para cada una de las diferentes subunidades y la proteína obtenida es más o menos soluble se procederá

a purificar a gran escala partiendo de cantidades mayores de cultivo, desde 1 L hasta 6 L de cultivo (o incluso más si fuese necesario). Seguidamente se producirá la purificación de la holoenzima de RNAP mediante cromatografía de afinidad teniendo presente que tenemos dos etiquetas de afinidad. Primero, usaremos la cola de polihistidinas en el N-terminal de *rpoA* para la purificación en una columna de níquel en un sistema AKTA FPLC (GE Healthcare) de forma similar a como se procede con NtcA. Esto se debe a que estas columnas permiten flujos elevados disminuyendo por tanto el tiempo en el que la proteína de interés se encuentra en contacto con un lisado altamente impuro y rico en proteasas. La elución de la proteína se realizará mediante un gradiente de imidazol. La RNAP eluida se inyectará en una columna GSTrap FF mediante el mismo sistema de cromatografía líquida a la que se unirán las etiquetas GST del N-terminal de *rpoC1*. La elución de la proteína se realizará mediante glutatión reducido.

Por último, y si es necesario debido a la existencia de numerosas impurezas en la muestra o a una incorrecta estequiometría de las distintas subunidades, añadiremos una etapa más de cromatografía mediante exclusión molecular (gel-filtración). Presumiblemente las subunidades  $\alpha$  y  $\gamma$  se encontrarán purificadas en exceso en comparación al resto ya que poseen las etiquetas. Utilizaremos una columna Superdex200 (Cytiva), capaz de separar moléculas de entre 10 y 600 kDa, porque el tamaño de la RNAP es de alrededor de 560 kDa. Priorizaremos durante la purificación de la RNAP, la obtención de una muestra con una correcta estequiometría de sus componentes frente al rendimiento en cantidad, ya que con cerca de 30  $\mu$ g de material puro es suficiente para los posteriores estudios de microscopía.

#### 4.4.2. Estrategia a seguir para el ensamblaje del complejo RNAP-NtcA-PipX-DNA para su estudio mediante Cryo-EM

En el momento en que obtengamos la RNAP, NtcA y PipX se requerirá el uso de oligonucleótidos correspondientes a genes bajo el control de NtcA. Por ello, se diseñan tres operadores cuya secuencia corresponde a *glnA*, *glnB* y *glnN* (Imamura y Asayama, 2009). Estos genes tienen diferentes requerimientos de 2OG, NtcA, PipX y factores  $\sigma$  para activar la transcripción (Forcada-Nadal et al., 2014).

Las secuencias son diseñadas (Figura 22) a partir de las secuencias con código de identificación: AAB87965.1 (ENA) para *glnA*, Synpcc7942\_0321 (Ensembl) para *glnB* y AF251806.1 (ENA) para *glnN*, y contendrán la secuencia consenso de unión de NtcA al promotor en la posición -41.5 y el resto de secuencias consenso (región -35, región -10) necesarias para la formación del complejo activación de la transcripción. La longitud de las secuencias diseñadas es de unos 85 pb (de -70 a +15 del sitio de inicio de transcripción), basándonos en los procedimientos utilizados por Feng et al., (2016) y Shi et al., (2020). Además se añadirían algunas bases no complementarias (de -13 a +5 del inicio de la

transcripción) en la hebra complementaria para facilitar la formación de la burbuja de transcripción (Shi et al., 2020).

	<u>Sitio de unión de NtcA</u>	<u>Caja -10</u>	<u>Posible sitio de inicio</u>
<i>GlnA</i>	GTGCCAATCGCCTCCATTTTAT <b>GTATCAGCTGTTAC</b> AAAAGTGCCGTTTCGGGCTACC	<b>TAGGAT</b> GAAAG <b>CG</b> GTCAGCAATGCTTG	
<i>GlnB</i>	ATGTAGACAGCGAATTTTCGATTGCT <b>GTAGCAGTAAC</b> TACA	AACTGTGGTCTAGTCAGCGG <b>TGTTAC</b> CAAA <b>GA</b> GTCGCTGACCGGGG	
<i>GlnN</i>	TCAAAATCTGGCCCTTGGTTCT <b>GTATCTTTCTAGC</b> GATCGAGCTGGTCACCATTGAG	<b>TACGAT</b> CAATT <b>GA</b> CTAGCTTTTTTGGG	

**Figura 22. Secuencia del operador del gen *glnA*, *glnB* y *glnN* para su ensamblaje con NtcA y la RNAP con  $\sigma A$ ,  $\sigma B$  y  $\sigma C$ . En cada secuencia se resalta el sitio de unión de NtcA (en rojo), la caja -10 (en azul) y el posible sitio de inicio de la transcripción (en verde)**

Seguidamente se realizarán ensayos de interacción con los distintos componentes, mediante ultrafiltración centrífuga usando filtros de tamaño de poro 100 kDa, aprovechando la gran diferencia de tamaño de la RNAP con NtcA y PipX de forma similar a los procedimientos utilizados en Llácer et al. (2007), en los que se usa una proteína de 200 kDa y otra de 37 kDa. También, se realizarán experimentos de SEC-MALS [cromatografía de exclusión molecular (SEC) acoplada con dispersión de luz multiángulo (MALS)], permitiendo analizar tamaño, forma y homogeneidad del complejo obtenido (Some et al., 2019).

Este tipo de experimentos de interacción se realizarán en paralelo con los tres constructos de la holoenzima de la RNAP junto a NtcA, PipX y las tres secuencias obtenidas de los operadores de los genes *glnA*, *glnB* y *glnN*. Para ello, durante el proceso, se variarán las concentraciones de los componentes y las condiciones de los tampones utilizados, modificando mayoritariamente el pH, la concentración de NaCl y la adición de algún detergente como CHAPSO que ha demostrado mejorar considerablemente la obtención de estructuras de RNAP mediante Cryo-EM (Boyaci et al., 2019). Una vez identificadas las condiciones óptimas se seleccionarán los complejos más prometedores para la obtención de estructuras tridimensionales mediante Cryo-EM.

Se espera que estos estudios estructurales sirvan para entender en detalle como un activador transcripcional de clase II (NtcA en nuestro caso) promueve el paso de la conformación cerrada RNAP-promotor a la conformación abierta (con el DNA desenrollado) y por lo tanto entender como activan la transcripción. Además no existe ningún otro ejemplo en la literatura de co-activación por parte de otro factor de transcripción, que en nuestro caso sería PipX, por lo que sería interesante conocer si PipX contacta con alguna de las subunidades de la RNAP, y el efecto que esto pueda tener sobre el complejo de iniciación de la transcripción.

## 5. CONCLUSIÓN

Tras la realización del presente trabajo podemos concluir que:

1. Se ha diseñado una estrategia de clonación que funciona correctamente y permitiría clonar en tándem todas los genes que codifican las diferentes subunidades de la RNAP cianobacteriana. Únicamente falta la clonación del gen *rpoC2* ( $\beta'$ ) para empezar a realizar ensayos de expresión y purificación de la RNAP de *Synechococcus elongatus*.
2. Los plásmidos obtenidos concuerdan con los tamaños esperados para las tres combinaciones de constructos pQLink + *rpoD1* ( $\sigma_A$ ) + *rpoZ* ( $\omega$ ) + *rpoC1* ( $\gamma$ ), pQLink + *rpoD2* ( $\sigma_B$ ) + *rpoZ* ( $\omega$ ) + *rpoC1* ( $\gamma$ ) y pQLink + *rpoD5* ( $\sigma_C$ ) + *rpoZ* ( $\omega$ ) + *rpoC1* ( $\gamma$ ). Por tanto, se logran en concentración y pureza adecuadas los constructos definitivos sin *rpoC2* ( $\beta'$ ), a falta de confirmar mediante secuenciación los dos últimos.
3. La estrategia de mutagénesis para eliminar la mutación de 49 pb del gen *rpoC2* mediante PCR no ha resultado exitosa. Sin embargo, se ha hallado una alternativa que consiste en la sustitución del fragmento de gen que contiene mutación por un fragmento artificial del mismo gen pero con la secuencia correcta.
4. Se ha conseguido expresar y purificar la proteína NtcA obteniendo una concentración y pureza adecuadas mediante cromatografía de afinidad.
5. Se ha conseguido expresar la proteína PipX correctamente como se evidencia en el gel de poliacrilamida.
6. Se ha llevado a cabo el diseño de oligonucleótidos con la secuencia correspondiente a los genes *glnA*, *glnB* y *glnN*, genes bajo el control de NtcA para el estudio mediante Cryo-EM del complejo RNAP-NtcA-PipX-DNA.



## 6. BIBLIOGRAFÍA

- ALLISON, L.; MOYLE, M.; SHALES, M.; INGLES, C.J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell*, 42(2), 599-610.
- ARCONDEGUY, T.; JACK, R.; MERRICK, M. (2001). PII Signal Transduction Proteins, Pivotal Players in Microbial Nitrogen Control. *Microbiology And Molecular Biology Reviews*, 65(1), 80-105.
- AUSUBEL, F. M. (1998). *Current Protocols in Molecular Biology*. Editorial John Wiley & Sons. (Libro)
- BECK, C.; KNOOP, H.; AXMANN, I.; STEUER, R. (2012). The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms. *BMC Genomics*, 13(1), 56.
- BOYACI, H., CHEN, J., JANSEN, R., DARST, S. A., & CAMPBELL, E. A. (2019). Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature*, 565(7739), 382-385.
- BROWNING, D.; BUSBY, S. (2004). The regulation of bacterial transcription initiation. *Nature Reviews Microbiology*, 2(1), 57-65.
- BURILLO, S.; LUQUE, I.; FUENTES, I.; CONTRERAS, A. (2004). Interactions between the Nitrogen Signal Transduction Protein PII and N-Acetyl Glutamate Kinase in Organisms That Perform Oxygenic Photosynthesis. *Journal Of Bacteriology*, 186(16), 5555-5555.
- EBRIGHT, R. (1995). The Escherichia coli RNA polymerase alpha subunit: structure and function. *Current Opinion In Genetics & Development*, 5(2), 197-203.
- ESPINOSA, J.; FORCHHAMMER, K.; BURILLO, S.; CONTRERAS, A. (2006). Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. *Molecular Microbiology*, 61(2), 457-469.
- FORCADA-NADAL, A. (2018). Hacia el entendimiento del control de la expresión génica del regulón cianobacteriano del nitrógeno mediado por el factor de transcripción NtcA y la proteína reguladora PipX. Tesis Doctoral en Ciencias Biológicas. Univ. de Valencia
- FORCADA-NADAL, A.; LLÁCER, J.; CONTRERAS, A.; MARCO-MARÍN, C.; RUBIO, V. (2018). The PII-NAGK-PipX-NtcA Regulatory Axis of Cyanobacteria: A Tale of Changing Partners, Allosteric Effectors and Non-covalent Interactions. *Frontiers In Molecular Biosciences*, 5.
- FENG, Y.; ZHANG, Y.; EBRIGHT, R. (2016). Structural basis of transcription activation. *Science*, 352(6291), 1330-1333.
- GHOSH, T.; BOSE, D.; ZHANG, X. (2010). Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiology Reviews*, 34(5), 611-627.
- HENRICI, R.C., PECEN, T.J., JOHNSTON, J.L., TAN, S. (2017). The pPSU Plasmids for Generating DNA Molecular Weight Markers. *Sci Rep* 7, 2438.

- HERRERO, A.; MURO-PASTOR, A.; FLORES, E. (2001). Nitrogen Control in Cyanobacteria. *Journal Of Bacteriology*, 183(2), 411-425.
- HERRERO, A.; MURO-PASTOR, A.; VALLADARES, A.; FLORES, E. (2004). Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiology Reviews*, 28(4), 469-487.
- HOHMANN-MARRIOTT, M.; BLANKENSHIP, R. (2011). Evolution of Photosynthesis. *Annual Review Of Plant Biology*, 62(1), 515-548.
- IMAMURA, S.; ASAYAMA, M. (2009). Sigma Factors for Cyanobacterial Transcription. *Gene Regulation And Systems Biology*, 3, GRSB.S2090.
- KIMPLE, M.; BRILL, A.; PASKER, R. (2013). Overview of Affinity Tags for Protein Purification. *Current Protocols In Protein Science*, 9.9.1-9.9.23.
- KÖRNER, H.; SOFIA, H.; ZUMFT, W. (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiology Reviews*, 27(5), 559-592.
- KUZNEDELOV, K.; KORZHEVA, N.; MUSTAEV, A.; SEVERINOV, K. (2002). Structure-based analysis of RNA polymerase function: the largest subunit's rudder contributes critically to elongation complex stability and is not involved in the maintenance of RNA-DNA hybrid length. *The EMBO Journal*, 21(6), 1369-1378.
- LEE, H.; FLORES, E.; FORCHHAMMER, K.; HERRERO, A.; TANDEAU DE MARSAC, N. (2000). Phosphorylation of the signal transducer PII protein and an additional effector are required for the PII-mediated regulation of nitrate and nitrite uptake in the cyanobacterium *Synechococcus* sp. PCC 7942. *European Journal Of Biochemistry*, 267(2), 591-600.
- LEIGH, J.; DODSWORTH, J. (2007). Nitrogen Regulation in Bacteria and Archaea. *Annual Review Of Microbiology*, 61(1), 349-377.
- LEYPUNSKIY, E.; LIN, J.; YOO, H.; LEE, U.; DINNER, A.; RUST, M. (2017). The cyanobacterial circadian clock follows midday in vivo and in vitro. *Elife*, 6.
- LLÁCER, J.L. (2012). Bases Estructurales de la Señalización mediada por la proteína PII. Tesis Doctoral en Ciencias Biológicas. Univ. de Valencia.
- LLÁCER, J.; ESPINOSA, J.; CASTELLS, M.; CONTRERAS, A.; FORCHHAMMER, K.; RUBIO, V. (2010). Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proceedings Of The National Academy Of Sciences*, 107(35), 15397-15402.
- MINAKHIN, L., BHAGAT, S., BRUNNING, A., CAMPBELL, E. A., DARST, S. A., EBRIGHT, R. H., & SEVERINOV, K. (2001). Bacterial RNA polymerase subunit  $\omega$  and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proceedings of the National Academy of Sciences*, 98(3), 892-897.
- MITSCHKE, J.; VIOQUE, A.; HAAS, F.; HESS, W.; MURO-PASTOR, A. (2011). Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proceedings Of The National Academy Of Sciences*, 108(50), 20130-20135.

- MUKHERJEE, K.; CHATTERJI, D. (1997). Studies on the omega Subunit of Escherichia Coli RNA Polymerase. Its Role in the Recovery of Denatured Enzyme Activity. *European Journal Of Biochemistry*, 247(3), 884-889.
- MURAKAMI, K. (2015). Structural Biology of Bacterial RNA Polymerase. *Biomolecules*, 5(2), 848-864.
- MURO-PASTOR, M.; REYES, J.; FLORENCIO, F. (2005). Ammonium assimilation in cyanobacteria. *Photosynthesis Research*, 83(2), 135-150.
- OMATA, T.; GOHTA, S.; TAKAHASHI, Y.; HARANO, Y.; MAEDA, S. (2001). Involvement of a CbbR Homolog in Low CO<sub>2</sub>-Induced Activation of the Bicarbonate Transporter Operon in Cyanobacteria. *Journal Of Bacteriology*, 183(6), 1891-1898.
- PICOSSO, S.; FLORES, E.; HERRERO, A. (2014). ChIP analysis unravels an exceptionally wide distribution of DNA binding sites for the NtcA transcription factor in a heterocyst-forming cyanobacterium. *BMC Genomics*, 15(1), 22.
- RIPPKA, R.; STANIER, R.; DERUELLES, J.; HERDMAN, M.; WATERBURY, J. (1979). Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology*, 111(1), 1-61.
- SAUER, J.; SCHREIBER, U.; SCHMID, R.; VÖLKER, U.; FORCHHAMMER, K. (2001). Nitrogen Starvation-Induced Chlorosis in *Synechococcus* PCC 7942. Low-Level Photosynthesis As a Mechanism of Long-Term Survival. *Plant Physiology*, 126(1), 233-243.
- SCHEICH, C.; KUMMEL, D.; SOUMAILAKAKIS, D.; HEINEMANN, U.; BUSSOW, K. (2007). Vectors for co-expression of an unrestricted number of proteins. *Nucleic Acids Research*, 35(6), e43-e43.
- SCHNEIDER, G.; HASEKORN, R. (1988). RNA polymerase subunit homology among cyanobacteria, other eubacteria and archaebacteria. *Journal Of Bacteriology*, 170(9), 4136-4140.
- SHESTAKOV, S.; KHYEN, N. (1970). Evidence for genetic transformation in blue-green alga *Anacystis nidulans*. *MGG Molecular & General Genetics*, 107(4), 372-375.
- SHI, W., JIANG, Y., DENG, Y., DONG, Z. AND LIU, B., (2020). Visualization of two architectures in class-II CAP-dependent transcription activation. *PLOS Biology*, 18(4), p.e3000706.
- SOME, D., AMARTELY, H., TSADOK, A. AND LEBENDIKER, M., (2019). Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *Journal of Visualized Experiments*, (148).
- STANIER, R.; COHEN-BAZIRE, G. (1977). Phototrophic Prokaryotes: The Cyanobacteria. *Annual Review Of Microbiology*, 31(1), 225-274.
- STEINHAUSER, D.; FERNIE, A.; ARAÚJO, W. (2012). Unusual cyanobacterial TCA cycles: not broken just different. *Trends In Plant Science*, 17(9), 503-509.
- VALLADARES, A.; RODRIGUEZ, V.; CAMARGO, S.; MARTINEZ-NOE L, G.; HERRERO, A.; LUQUE, I. (2010). Specific Role of the Cyanobacterial PipX Factor in the Heterocysts of *Anabaena* sp. Strain PCC 7120. *Journal Of Bacteriology*, 193(5), 1172-1182.

- VÁZQUEZ-BERMÚDEZ, M.; HERRERO, A.; FLORES, E. (2002). 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus* glnA promoter. *FEBS Letters*, 512(1-3), 71-74.
- WHITMAN, W.; COLEMAN, D.; WIEBE, W. (1998). Prokaryotes: The unseen majority. *Proceedings Of The National Academy Of Sciences*, 95(12), 6578-6583.
- ZETH, K., FOKINA, O., & FORCHHAMMER, K. (2014). Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. *Journal of Biological Chemistry*, 289(13), 8960-8972.
- ZHAO, M.; JIANG, Y.; HE, Y.; CHEN, Y.; TENG, Y.; CHEN, Y. (2010). Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. *Proceedings Of The National Academy Of Sciences*, 107(28), 12487-12492.

## 7. ANEXO

### 7.1. Anexo I

#### Secuencias de los genes que codifican las diferentes subunidades de la RNA polimerasa de *Synechococcus elongatus* y gen artificial.

>Synpcc7942\_2209 DNA-directed RNA polymerase subunit alpha  
 GTGACATTTCAAGTCGAATGCGTCGAGAGCCGCACTGAAGCCGACCAAGGGCAATATGGC  
 CGCTTTTCCATCGAGCCTCTCGCTCGTGGGCAAGGTACAACGGTTGGCAATGCGCTCCGG  
 CGTGTGTTGCTCTCCAACCTAGAAGGAACAGCCGTCACAGCTGTCCGGATCGGTGGTGTG  
 AACCACGAATTGCAACCATCCCCGGCGTGCAGGAGGATGTCTCGACATCTTGCTCAAC  
 GTCCGTGAGTTGGTGGTGCATGCCACAGTCTCAGCCCCAGATTGGTCCGGTGCAGGGT  
 GTGGGTCCGGCAACTGTCACGGCTGCTGATGTTGACTTTGGCCCTGAAGTTGAAGTCATC  
 AATCCCAATCACTACATCGCTTCGCTCTCTGAGGGTGCCACCCTCGAGATGGAGTTGAAA  
 GTTGTAGTGGGGCACGGGCTACCGCGCCATCGACCGCAGCCATGACGAAACGACTGCCCTA  
 GACTTCTCGCACTTGATGCTGTATTTCATGCCGGTGCCTCGGTTGAACTACAGCGTTGAA  
 GATGGCGCGCTCGGGAAATCCACGGCCATCGACCCGCTCGTACTCGAGGTTTGGACCAAT  
 GGCAGTCTTCGCCCCAGGAAGCACTGAGCCAAGCCGCTCTTGTCTCGTCGCGTGTTC  
 GAGCCGTTGAAAAACGTCAGTGTTGGCAGTACCCACACCCGCGATCCTGAGCCACGCC  
 GAAAGCCAAACGCCGATCGAAGATCTGCAGCTTTCGGTTCGGGCTTACAACGCTGAAG  
 CGGGCGCAAGTGAACCTCCGTTGCTGACTTGTGAGCTACACCTACGAAGATCTACTGGAG  
 ATCAAAAACCTTCGGCCAGAAATCGGCTGAAGAGGTGGTGAAGCTCTCGAGCGCATCGGC  
 ATTAAGCTCCAGGAATCGAAAGTTTCTGA

>Synpcc7942\_1522 DNA-directed RNA polymerase subunit beta  
 ATGGCTGAGCAAACGCAACTCGCTCCCCTGCCTTCCATTTACCTGATCTCGTTGCCATT  
 CAACGGAACAGCTTCCGTTGGTTCCTAGAAGAAGGGCTCATTGAGGAGCTGGAAAGCTTT  
 TCACCAATCACCGACTATACCGGTAAGCTTGAGCTTCACTTCCCTCGGTAAGCAGTATAAG  
 CTCAAGCGCCCCGAAGTACGACGTTGATGAAGCAAAACGGCGTGACGGCACCTACTCGGTT  
 CAAATGTATGTGCCGACCCGTTTGTATCAACAAAGAAACGGGCGAAATCAAGGAGCAGGAA  
 GTCTTTATTGGCGATCTGCCCTGATGACTGATCGCGGCACCTTCATCATCAACGGTGCC  
 GAACGGGTGATCGTCAACCAGATCGTCCGTAGCCCCGGCGTCTACTACAAATCAGAGCGC  
 GACAAGAACGGTGCCTCACCCACAATGCCAGCTGATCCCCAACCGTGGCGCTTGGCTG  
 AAATTTGAAACCGATAAAAACGGTTTAGTTTGGGTGCGCATCGACAAGACCGGTAAGTTG  
 TCGGCGCAGGTACTGCTGAAAGCCTTGGGCTGAGCGACAACGAAATTTACGACAAGCTC  
 CGTACCCCTGAGTATTACCAGAAGACCATCGATAAAGAAGGTCAGTTCAGCGAAGACGAA  
 GCGCTGATGGAGCTCTACCAGCAAGCTCCGTCGGGCGAACCCGCCACGGTCTCTGGCGGT  
 CAGCAATTGCTGGAATCGCGGTTCTTCGATCCCAAACGCTACGACCTAGGTCCGGTGGGC  
 CGCTACAAGCTCAATAAGAAGCTGGGTCTCAACGTCGCTGATACGGTGCAGGACGCTGACA  
 TCCGAAGATATTTGGCGGCGATCGACTACCTGATTAACCTCGAGCTTGACTTGGGTGGC  
 TGTGAAGTCGATGACATCGACCACCTCGGCAACCGTCCGGTGCATCGGTGGGCGAGCTG  
 CTGCAAAACCAAGTGCAGGTCGGCTCAACCGCTAGAGCGGATCATTCCGGGAACGGATG  
 ACGGTGTCGGATTCCGACAGTCTTTCGCCGGCTTCCCTTGGTCAACCCCAAACCGCTGGTG  
 GCTGCGATCAAAGAATCTTTGGTTCCTCGCAACTCTCGCAGTTCATGGACCAAACCAAC  
 CCCTTGGCGGAGTTGACCATAAACGACGCTCGAGTGCCCTCGGTCCCGGTGGTTTGACA  
 CGGGAGCGGGCTGGCTTTGCGGTGCGGGACATTCACCCAGCCACTACGGCCGGATTTGC  
 CCGATTGAAACGCCGGAAGGCCCGAACCGGGTCTGATTGGTTCCTTAGCGACCCACGCC  
 CGGTC AACGACTACGGCTTTATTGAAACGCCGTTCTGGCGCGTGAAGAAGGACGGGTT  
 CGCAAGGACTTGGCGCCGGTCTACATGACTGCTGACCAGGAAGATGACCTGCGGGTGGCT  
 CCGGAGACGTGGCTACGGATGACGCGGGCTACATCCTGGGAACCAATTCGGTACGT  
 TATCGCCAGGACTTACCACCACGACGCGGGGAGCGGGTGGACTACGTTGCGCTCTCGCCG  
 GTGCGAGTTATCTCGGTGGCAACGTCGTTGATTCCTTTCTTGAACACGATGACGCCAAC  
 CGTGCCCTGATGGGCTCGAACATGCAACGGCAGGCGGTGCCGCTGTTGCGGCCAGAGCGG  
 CCTTTGGTTCGGGACGGGTCTGGAGCCCCAAGCGGCCCGTGACTIONAGGGATGGTATCACC  
 AGCCCGGTGCGATGGCACGATCTCCTACGTCGATGCCACCCACATTGAGGTGACGGCTGAC  
 ACAGGTGAGAAGTATGGCTACGCCCTACAGAAGTACCAGCGCTCCAACCAAGATACTTGT  
 CTGAACCAACGGCCGATCGTGTTTGAAGGCGATCGCGTTCAACGGGGTCAAGGTGATTGCT  
 GACGGTCTGCCACCGAAAAGGTGAGCTGGCCCTGGGGCAAAACATCCTCGTCGCTAC  
 ATGCCCTTGGGAAGGCTACAACCTACGAGGACGCGATTCTGATCAGCGAGCGGCTGGTCTAT  
 GACGACGCTCTATACCTCGATCCACATTGAAAATTCGAGATTGAGGCGCGTCAGACCAAG

TTAGGCCCTGAGGAGATTACCCGCGAAATTCCCAACGTTGGCGAAGATGCTCTGCGCCAA  
CTCGACGAAAACGGGATTATCCGCGTCGGCGCATGGGTGGAGTCCGGTGACATCTTGGTT  
GGTAAGGTGACGCCAAAGGCGAATCGGATCAGCCGCCAGAAGAAAACTGCTGCGGGCA  
ATCTTCGGTGAGAAAAGCGCGGGACGTCCGCGATAACTCCCTGCGGGTGCCGAATGGTGAG  
AAAGGCCGCGTTGTTGATGTGCGTCTCTTACCCCGTGAGCAGGGTGATGAGTTGCCGCCG  
GGCGCAACATGGTTGTCCGGGTTTACGTGGCTCAGAAACGCAAGATCCAAGTCGGCGAC  
AAGATGGCGGGTCGCCACGGCAACAAGGGGATCATTTCCGCGATTCTGCCCTGCGAAGAC  
ATGCCTTACCTGCCGGATGGCACGCCGTAGACATCGTGTCAATCCCTTGGGTGTACCC  
TCGCGGATGAACCGGCAAGATCCAGGTCTACGACGGTCGGACAGGGGAACCCCTTCGATCGC  
CCAGTGACGGTGGGTGCTGCCTACATGCTTAAGCTGGTTCACTTGGTTCGACGACAAGATC  
CACGTCGCTCGACAGGTCCCTACTCCTTGGTGACCCAACAACCCCTTGGGTGGTAAAGCC  
CAACAAGGGGGTCAGCGCTTCGGGGAGATGGAAGTTTGGGCACTGGAAGCCTACGGTGCT  
GCCTACATCCTGCAAGAGTTGCTGACGGTCAAGTCGGACGACATGCAAGGGCGGAATGAA  
GCTCTCAATGCGATCGTTAAGGGCAAAGCGATTCCCTCGCCCCGGCACACCAGAGTCTTTC  
AAGGTGTTGATGCGAGAGCTGCAGTCGCTCTGTCTCGACATTGCGGTCTACAAGGCCTCA  
ACCGAGGACTATGAAGAAGACAAAGAAGTGGATCTGATGGCGGATGTCAACCAGCGTCGG  
ACTCCCTCGCGCCCCGACCTACGAGTCGATGTCTGTGGGCGACATTGATGACGATGACGAC  
TAA

>Synpcc7942\_1523 DNA-directed RNA polymerase subunit gamma  
ATGGCGAAGCAGGAACAGCGGTTTACTACGTCAAGATCGCACTTGCCTCACCAGGAGCGT  
ATCCGGCAGTGGGGTGAGCGCACGTCGCCAATGGTCAAGTAGTCGGTGAAGTACCAAG  
CCTGAAACCATCAACTACGACATCGGACACTGAAGCCGGAGATGGACGGGCTGTTTTGCGAAAA  
ATCTTTGGTCCCTGCGAAGGACTGGGAATGCCACTGCGGTAAGTACAAACGTGTTTGGGCAC  
CGCGGCATTGTCTGCGAGCGTTGTGGCGTGGAAAGTCACCGAATCACGGGTGCGACGGCAC  
CGTATGGGCTTCATTAACCTGGCTGCTCCGGTTGCCACGTTTGGTACCTCAAAGGCATC  
CCGAGCTACATCGCGATTCTGCTCGACATGCCCCGTCGAGACGTTGAGCAAATCGTCTAT  
TTCAACTCCTACGTCGTCCTCAACCCAGGCAACCACAGCGAGCTGCAGTACAAGCAGCTA  
CTGAACGAAGACCAGTGGATGGAGATCGAGGACCAAATCTACGCTGAGGAATCTGACCTC  
GAAGGTATTGAGGTTGGCATCGGTGCAGAAGCCCTGCAACAGTTGCTGCAGGATTTGAAC  
CTCAATGAAGAGTCGGAACAACTCCGTCAAGAGATTGCTGAGTCGAAGGGCCAAAAACGG  
GCCAAGTTAATCAAACGCTGCGTGTGATTGATAACTTCATTGGCACAGAGTCTCGGCC  
GAGTGGATGGTGTGAATGTATCCCGGTGATCCCTCCCGATCTCCGCCCGATGGTGCAG  
TTGGATGGTGGCCGTTTGGCCACAGTGACCTCAACGATCTCTACCGTCGGGTGATCAAC  
CGGAATAACCGGTTAGCGCGTCTGCAGGAGATTCTCGCACCTGAGATTATCGTCCGTAAC  
GAAAAACGGATGTTGCAGGAAGCAGTAGACGCTCTGATCGACAACGGTCGTCGTGGCCGT  
ACGGTCGTCGGTGCCAAACACCGACCCCTCAAGTCCCTCTCGGACATCATCGAAGGGAAA  
CAGGGTCGCTTCCGGCAAACTTGTGGGTAACGGGTTGATTACTCCGGTTCGGTTCGGTC  
ATCGTTCGTCGGCCCCAACCAGAAATTCACCAGTGTGGTCTGCCACGGGAAATGGCGATC  
GAGCTGTTCCAGCCGTTTGTATCCATCGCCTGATCAAAAATCACAGCATCAACAACATC  
AAGCAGGCCAAAAAATCATTTAGAAGAACGATCCCTTGATTGGGATGTCTTGGAAAGAG  
TGATTGAAGGCCACCTGTGATGCTGAACCGGGCCCCGACGCTGCACCGTTTGGGGATT  
CAAGCCTTCGAACCGATCTTGGTAGAAGCCGTGCGATCCAGCTTCATCCCTTGGTCTGT  
CCAGCGTTTAAACGCTGACTTTGATGGTGACCAAATGGCGGTTACGTCGCCCTCTCGATT  
GAAGCCCAGGCTGAAGCTCGTATGCTGATGCTGGCTTCGGGCAACATTCTGTGCGCCGCA  
ACGGGGCAACCCATTGTCACGCCGTCCCAGGATATGGTCTTGGGCTGCTACTACTTGACC  
GCCGAGAATCCTGGTGTCTCAGAAAGGGGCTGGACGCTACTTTGCCAACCTCGAAGATGCC  
ATCCGCGCCTTTGAGCAGGGCTCTGTGATCTTCATGCCTGGGTCTGGGTGCGCTTTGAT  
GGCGAAGTGAAAAGCGAGGGTAAAAGCGACGAACCGGAGTCGGTGGTTCGACGCCGATGAT  
GGCACGGTCACCAAGACCTATCGTTTCCGCCGAATTCGCGGAGACTGAAGACGGGCAGCGT  
CTTCCCAGTACGTGAAAACAACGCCCGGTGCGATTCTCTTCAACAACACTGTTTCAGACT  
CGTTGATCCACTAG

>Synpcc7942\_1524 DNA-directed RNA polymerase subunit beta'  
ATGGCAGAAGCAGAAAAGCGCGCCGATTTTCCGCAACCGGGTCATCGATAAGAAGCAGCTG  
AAGAAGCTGATTGGCTGGACCTTTGCTCACTATGGGACAGCAAAAACAGCCGTAGTTGCC  
GACGATCTCAAAGCCCTGGGATTCGGTTACGCCACCCGTGCTGGGGTTTCAATCAGTATT  
GATGACCTCAAGGTGCCGGTTCTAAGGCGGAGCTGCTAGAGTCTGCTGAAAAGCGGATT

CAAGAGACTGAAGATCGCTACACCCGTGGAGAAATCACCGAGGTGGAGCGCTTCCAGAAG  
GTGATCGACACCTGGGCGAACACCAACGATGAGCTGACCGATCGCGTGGTCAAAAACCTTC  
CGCGAGTCCGATCCGCTCAACTCCGTCTACATGATGGCGTTCTCGGGTGCCCGGGGAAT  
ATCTCCCAGGTGCGACAGCTGGTGGGGATGCGGGGTCTGATGGCCAACCCCCAAGGCGAG  
ATTATTGACTTGCCGATTAAGACCAACTTCCGGGAAGGTTTGACCGTTACGGAGTACATC  
ATCTCCTCTACGGGGCGCGTAAGGGCCTGGTTGATACGGCTCTGCGGACGGCGGACTCG  
GGTTACTTGACCCGTCGTCTGGTGGACGTTTCGCAGGACGTGATCATTACGAAGTGGAC  
TGCGGAACCAGCCGGGGCCTATTTGTGGAAGCGATGACCGATGGCGATCGCATTTTGATT  
CCGATTTTCGCAGCGCTTGTGGGTGGGTTACGGCTGAGGCCGTGCTAGATCCCAGTACG  
GATGAAGTGCCTGGCTGAAGCAGGCCAAGACATCAATGAAGATCTGGCCAATCGAATTGAG  
AAAGCAGGCATCAAAAAGGTGAAGGTGCGATCGCCCTGACCTGCGAAGCCGCACGATCG  
GTTTGCCAAAAGTGTACGGCTGGAGCCTTGCCACGCCAGATGGTGGACATGGGTGAA  
GCTGTGCGGGATCATCGCAGCCAGTCGATCGGTGAGCCGGGGACCCAGCTGACCATGCGG  
ACCTTCCACACGGGTGGGGTATTTACGGGGGAAACTGCCCGTCTGCTGCGGGCACCTGTA  
GCGGGCACGATTAAGCTCGGCAAAAAGCCCGGACTCGTCCCTACCGGACCCGCCACGGC  
GAGGAAGCGCTGCTGGCAGAGGCCAACTTTGACCTCGTGCTGGAAGGGAAAGGTGCGAAG  
GAAACCTTCGCAATTCGCAAGGTCGACGATTTTTGTGCAGGACGGCGATAAAGTCGCA  
GCCGAAGCAATCTTGCCGAGGTGCCAGTCAGCGGTCTGTACCAAGCGCACCGTAGAAAA  
GCCACCAAGGACGTGGCGACTGATTTGGCGGGTGAATTCGCTTCCAAGACATCGTTCCG  
GAAGAGAAAACCGATCGCCAAGGCAACACTACCCGTATTGCTCAGCGGGGTGGTCTGCTT  
TGGGTCTTGGCTGGCGATGTCTACAACCTTGTGCCGGGTGCTGAGCCGACCGTCAAAAAC  
GGCGATCGCGTAGAGGTGGGAGATGTCTTGGCTGAAACCAAGCTGACCACTGAGCGCGGC  
GGTACCCTGCGCATGGGTGAGGACAACGGGAGCAGTACCCACCGGGAAGTCGAAATCATC  
ACGGCGTCAGTGGTACTAGATACCGCAACGGTGAAGCCGAGGCGAGCCAAGGCCGTGAG  
CACTACGTGATTGAGACCAAGGTGGCCAGCGGTTAACTTGTGGCAGCACCGGGAAC  
AAGTACAGACTGCGCCACGTGGTGGCTGAGCTAATCGACAGCCGCTACCGGACTCAAAC  
GGTGGTCTGCTCAAATACTCCGGCGTGAATCTCAAAAAGGCCGGGCTAAGCGGAAA  
CAGGGCTACGAGGTGACCAAGGGCGGCACCCTGCTCTGGATTCCCGAAGAGACCCACGAG  
GTCAACAAGGACATCTCCTTGCTGAACGTGCAAGACGGTCAGTTGGTGGAAAGCTGGCAG  
GAAGTCGTCAAAGACATCTTTTGCCAGACGACTGGGATTGTTTCGGTCACCCAGAACAAC  
GACATCTGCGGGAAATCGTGATCAAGCCTGGAGATGTCCACGTGTTGGATGACCCAGAC  
ACGGCGGCGAAGTACGACGAAGGTGCTTGGTTAACGCCGGCGAAGAAGTCTTCCCAGGT  
TTGACGGCCGAACAACCTGGTCTGGGCAGAAGCCGTTGATGGTACGGACGGCCCCGCTCCTC  
TTGCTGCGACCGGTGCAAGAGCTAGTCATTCTGATGAGCCGCCAGTACCGAGCCAAGAC  
TCCTCGCAAGAGTCAAGCAGCCGTTTCGATTGCTGCGCTGCGGTCCAACGTCTGCAGTTC  
CAAGATGGTGAACGGATCAAGTCCGTTGAGGGAGTAGATCTCCTGCGGACCCAGTTAGTG  
CTGGAGTCTGAAGAAGGTTCTCCCAGCTCTCGGCTGACATTGAACTGCTGCCGGACAGC  
AAGGATCCTGAAACGCTGCGCCTGCAGTTAGTAATCATTGAGCCTGTGGTGAATTCGTGCG  
GACGTGCTTCTGACACCACCCACGGCAGTACGCACACCCGAGTTGCGCGTTAAAGATGGC  
CAAAGGTGAAGCCGGGCGCCGTCATTGCTTGCACGCAAATTCAGTGCAAGGAAGCCGGT  
GTGGTTGCGGGCATCAAGAGGGCAGCGAAGCGGTTGCTGCGCTCTTGGTTCGAGCGGGAA  
CGGGACTGCGTCACCTTGACCTTGATGTCACTGCGGCAACGCAACTCCAGCCTGGCTCG  
CTGATGTGCGGGTACCCAGTTGGTAGATGCGCATCATCGACCCGGAATCCGGTGGGTT  
CGTGGCATCGCACCGGGCAACTGCAACTGCGGATTGCTCGTCCCTACCGTGTGTGCGCAG  
GGTGCAGTCCGACGTTGAAGATAAAGGTCTGGTGCAGCGGGGTGACAACCTTAGTGCTG  
CTCGTGTGTTGAACGAGCGAAAACCTGGGGACATCATTAGGGTCTACCGCGGATTGAGGAG  
CTCCTCGAGGCTCGTAAACCAAGGAAGCCTGCATTTTGCCCGCCGCTCCGGGTGTGGCG  
CACATCAACTACAGCGATGACGATGCAATCGACATCCAAGTGATTGAAGCAGACGGTACC  
CAAGCTGATTATCCAGTTGGTCTGCTGAGCCGCTGATCATCAGTGACGGTGAACCCGTTG  
GATGCAGGGCAAGCCCTGACGGATGGGCCAGCAAACCCCATGACTTGCTAGAAATCTAC  
TACGACTACTTCCGCGAGCAGTTGGGCGAAGACTACGAAGCAGCGCTAGAGAGCCTGCGT  
CGGGTGAAGCTCTGCTGGTGAACGAGGTGCAGTCGGTCTATCAATCCCAAGGGATTGAT  
ATCTCCGACAAACACATCGAGGTGATCGTGCGGCAGATGACTTCTAAGGTGCGGATCGAC  
GATGGCGGCGACACGATCATGCTGCCGGGCGAATTGCACGAACTGCGAGAGGTCTACAAC  
TCCAACAACACCATGGCTTTGACGGGCATGGCCCCGGCTCAGTTCACACCGGTTCTGTTG  
GGGATTACCAAGGCCTCACTCAACACCAACAGCTTCATCTCGGCAGCGAGCTTCCAAGAG  
ACGACACGGGTGCTCACGGAAGCGGCAATCGAAGGCAAGTCCGATTGGCTGCGCGGTCTC  
AAAGAGAACGTGATTATCGGTCTGATTCCGGCAGGGACTGGTTTCAAGGCCTACGAA  
GAGTCTTGTGACCGACGTTGATGGTGGCTACGAAGATCGGGTCTACGACGACGATTTG  
GCCGATGTGGTGTGATGATCGCGCCGCTCGCAGCTACACCCTCAATGAAGGTGCGGAC  
TTCAGTCTTCGATGACCTTTGCCGAAGGCGAATCGATGATTCTCGACGACGGTGAAGAG

TTGATCGATGACAGCAGTGCCAGCCTCCGCAACCTGGTGGATGTTGACGAAGACTAG

>Synpcc7942\_1710 DNA-directed RNA polymerase subunit omega  
 ATGCTCCAGCGCTTCGATCTGGATTCTCAAGATCTGCTCTTCAAAGCTGAGAGCCTGATT  
 GTTAACTCCACGAATCGCTATCACGTCACCCCTACAAATCGCCCGTCGCGCCAAGCAGGCT  
 CGCTACGAAGAGATGAAAAACCTCAGCGAAGAACTGGCATTAAAGCCCGTGTGCGCGCC  
 ATTTTAGAAATGTCAGACGAGCTCAATCAACCGGAGATTATCGGCGGGTGA

>Synpcc7942\_0649 RNA polymerase sigma factor RpoD  
 ATGACGCAGCTGATTTCAATCGATAAGGAGCAAGAGGAAGCCGGCATGACTCAGGCCACC  
 GAACTACTCGACCCCGCGCTCAAACCGGCGGAAACAAAGGCCAAACGGAGTAGTCGCAAAA  
 AAAGCGACAACCTGCTGTTGTTGAGCCCGCAACCACGATCGCTCCGACTGCAGATGTCGAC  
 GCGATCGACGACGAGGATTCCGGTTGGTGAGGATGAAGATGCGGCGGCCAAAGCCAAAGCT  
 AAAGTCCGCAAAACCTATACAGAAGACTCAATTCCGGCTCTATCTCCAAGAGATTGGTTCGA  
 ATTCGGTTGCTGCGGGCCGATGAAGAAATGAGCTCGCCCGTCAAATCGCTGACTTGCTT  
 GCTCTAGAACGCATTCGCGATGAGTTGCTCGAACAGCTCGATCGTTTGCCAGCGACGCC  
 GAGTGGGCCGCTGCTGTTGACTCGCCGCTGGATGAATTCGCGCGCCGCTGTTCCGGGGT  
 CGCCGCGCCAAGGACAAGATGGTGCAGTCAACCTGCGGTTGGTGGTGTGATCGCCAAG  
 AAATACATGAACCGGGTCTGTCTTCCAAGACCTGATTCAGGAAGGTAGCTTGGGTCTG  
 ATTCGGGCTGCCGAAAAGTTTGACCACGAGAAGGGCTACAAGTTCTCGACCTACGCGACT  
 TGGTGGATTTCGCCAAGCCATCACCCGCGGATCGCGGATCAAAGCCGCACGATCCGCGCTG  
 CCTGTTACCTCTACGAAACCATTTCTCGCATTAAGAAAACCACCAAGCTGCTTTCTCAG  
 GAAATGGGTTCGTAAGCCCACCGAGGAAGAGATCGCCACTCGCATGGAGATGACGATCGAG  
 AAGCTGCGGTTTATCGCTAAGTCAGCCAGCTGCCAATTCGCTGGAAACCCCAATCGGT  
 AAAGAAGAAGACTCGCGTCTGGGTGACTTCATCGAAGCTGATGGCGAAACCCCGAAGAT  
 GAAGTCGCAAAAAACCTGCTGCGCGAAGACCTGAAGGCGTCTTAGCACTCTCAGCCCC  
 CGCGAGCGCATGTGCTGCGCTGCGCTATGGCTTGGATGACGCGCCGATGAAAACGCTG  
 GAAGAAATCGGCCAGTTGTTTAAATGTCACCCGCGAACGCATCCGCCAGATCGAGGCGAAA  
 GCACGCGGAAACTGCGTCAATCCCAACCGCAACAGCATCTGAAAGAGTACATCCGCTAG

>Synpcc7942\_1746 group2 RNA polymerase sigma factor SigB  
 ATGTCTACCTCCTCCGCTCAATACAGTCCTGATTTAGTCCGTGCTTACCTGCAAGAAATT  
 GGCCGGGTGCGGCTGCTGACCGCTGAAGAAGAGTTGTGCTTTGGCCGACAGGTTACGCGC  
 CTGATGATGCTCTTGGATGCTCAAACCGAGCTGCGCGATCGCTTGGGGCATGAGCCTAGC  
 AAAGAAGAATGGGCTGCTGCTGTTGACCTGAATCTCGAAGATCTCGATCGCCAAATTGAG  
 CAAGTTCAGCGGGCGAAACGCAAAATGATTGAGGCCAACCTGCGCCTCGTTGTTTCGATC  
 GCCAAGAAATACCAGAAGCGTCACATGGAATTCCTGGATCTCATCCAAGAAGGCACACTG  
 GGGCTGGAGCGTGGTGTGCAAAAAATTCGATCCCTCCAAGGCTACAAGTTCTCTACCTAT  
 GCCTACTGGTGGATTTCGCCAAGCTATTACCCGTGCGATCGCCCAACAATCCCGCACGATT  
 CGTCTGCCATATTCACATCACTGAAAAGCTGAACAAGCTCAAAAAACCCAGCGCGAACTC  
 TCGCAGCAATTGGGCCGCGAGTGCCACAGCTTCAGAGCTGGCAGAAGTGTGGAGCTGCCA  
 CTGGAGCAGGTGCGGGAATACATCCAAATGAACCGCCAGCCAGTATCGCTCGATGTCAA  
 GTAGGTGACAGCCAAGACACTGAATTCAGGAACCTGCTGGAAGACGAGCAGTCTTCGCCC  
 TCAGATTACGTGGAACAAGAATCGCTACGCGCGATCTACGCAATTTAATGGCAGAACTG  
 ACCCCCCGCAACAGGCCGTGATTGCCCTGCGCTACGGCTTGGATGAAGGTGACAGTCTG  
 TCGTTGGCCAAAAGTCGGCGAACGCCGTAATATCAGCCGCGAACCGTGCAGGAACTAGAG  
 CGCCAAGCCATGGATCACCTACGTGCGCGCAGTCTGCTTCTGGCGGAATACGCAGCTAGC  
 TAG

>Synpcc7942\_1849 RNA polymerase sigma factor SigC  
 ATGACGGCTGCCTATTTCACTCATTCCACGGAATACGAAGCAGTTTCGTCTGACGTCGCT  
 AACTTGTGCAAACTCCAAGTTCTGAAGCCGTGGAAACAGCGGTCTCTAGCCGCCAGTCA  
 ACCGATCTGATTCGGCTGTATTTACAAGACATCGGGCGAGTCCGTCTCCTTGGGCGCGAT  
 GAAGAGGTGAGTGAGGCACGGTGGGTGCAGCGCTACATGCAGTTGTTGGCGGTTTCGGAAT  
 TGGCTGCGCGTTGAAGAAGGAGGGGTTCTCAAGGATTACGTTGACCTGCAGGAAGGCCGC  
 GATCGCCTGACCGCGCAGCTGGGTACCGTCTTCTTTGAAGCTTGGGCCAAGATTGTC  
 GATTGTTCCCTGACTGAATTACGCCAAGCACTACAACCTGGGGCTGCATCGCTGGGCTGAA  
 GTCTGCCAGCTCGATCTTGAAGAAATTAAGCAGGTCTGCAAGAAGGCCAGCGAGCTAAA  
 GATCACATGATCAAGGCCAACCTCCGTTTAGTTGTTTCCGTGCTAAAAAATATCAAAAC  
 CGAGGACTAGAGCTTCTCGATCTCGTTCAAGAAGGCACCCTAGGACTAGAGCGGGCTGTT  
 GAGAAGTTCGACCCGACGAAAGGCTATCGATTTAGTACCTATGCCTATTGGTGGATTCTG



CAGGGAATTACCCGCGCGATCGCAACCCAAAGTCGGACGATTTCGCTTGCCCCGTCCATGTG  
ACGGAGAAATTAACCGGATCAAAAAAGCCCAGCGCAAAATTGCGCAAACAAAGGGCAGA  
ACTGCCAACCTTGAGGAGATTGCTCAAGAACTTGAGATGACCCCGATCAAGTGCGGGAA  
GTCCTGTTGCGGGTGCCCAAATCGATTTCCCTAGAAACCCGTGTTGGCAAGGAAAAGGAT  
ACGGAGCTAGGGGATTTGCTGGAAACCACCGATTTACCCCAAGAAGACGATTGACTCAG  
GAATCCCTCAGCAATGATCTGCAGCGGTTATTGCTCGAACTCAATCACCGCGAACGCGAT  
GTAATTCGGATGCGGTTTGGTCTGCGCGATGGGCAGATTTATTCCCTCGCTGAAATTGGG  
CAAGTGCTCGACCTATCGCGTGAGCGAGTTCGCCAGATTGAGGCTAAGGCCCTGCAAAAG  
CTGCGGCATCCGCGACGACGCAATCAGATCCGTGACTATTTGGAGTTGTTGCACTAG

Gen\_artificial\_RpoC2

GCTGAGGCCGTGCTAGATCCCAGTACGGATGAAGTGTGGCTGAAGCAGGCCAAGACATCAATGAAGATCTGG  
CCAATCGAATTGAGAAAGCAGGCATCAAAAAGGTGAAGGTGCGATCGCCCCTGACCTGCGAAGCCGCACGATC  
GGTTTGCCAAAAGTGCTACGGCTGGAGCCTTGCCCACGCCAGATGGTGGACATGGGTGAAGCTGTCGGGATC  
ATCGCAGCCCAGTCGATCGGTGAGCCGGGGACCCAGCTGACCATGCGGACCTTCCACACGGGTGGGGTATTTA  
CGGGGGAAACTGCCCGTCTGCTGCGGGCACCTGTAGCGGGCACGATTAAGCTCGGCCAAAAAGCCCGGACTCG  
TCCCTACCGGACCCGCCACGGCGAGGAAGCGCTGCTGGCAGAGGCCAACTTTGACCTCGTGCTGGAAGGGAAA  
GGTCGGAAGGAAACCTTCGCAATTCTGCAAGGTTTCGACGATTTTTGTGCAGGACGGCGATAAAGTCGCAGCCG  
AAGCAATCTTGGCCGAGGTGCCAGTCAGCGGTCGTACCAAGCGCACCGTAGAAAAAGCCACCAAGGACGTGGC  
GACTGATTTGGCGGGTGAATTCGCTTCCAAGACATCGTTCCGGAAGAGAAAACCGATCGCCAAGGCAACACT  
ACCCGTATTGCTCAGCGGGTGGTCTGCTTTGGGTCTTGGCTGGCGATGTCTACAACCTGCTGCCGGGTGCTG  
AGCCGACCGTCAAAAACGGCGATCGCGTAGAGGTGGGAGATGTCTTGGCTGAAACCAAGCTGACCACTGAGCG  
CGGCGGTACCGTGCGCATGGGTGAGGACAACGGGAGCAGTACCCACCGGGAAGTCGAAATCATCACGGCGTCA  
GTGGTACTAGATACCGCAACGGTGAAAGCCGAGGCGAGCCAAGGCCGTGAGCACTACGTGATTGAGACCAAGG  
GTGGCCAGCGGTTTAACTTGTGGCAGCACCGGGAACAAAGGTACGACTGGCCACGTGGTGGCTGAGCTAAT  
CGACAGCCGCTACCGGACTCAAACCGGTGGTCTGCTCAAATACTCCGGCGTGGAATCTCCAAAAAGGCCGG  
GCTAAGGCGAAACAGGGCTACGAGGTGACCAAGGGCGGCACCCTGCTCTGGATTCCCAGAGAGACCCACGAGG  
TCAACAAGGACATCTCCTTGCTGAACGTGCAAGACGGTCAGTTGGTGGAAGCTGGCACGGAAGTCGTCAAAGA  
CATCTTTTGCCAGACGACTGGGATTGTTTCGGTCACCCAGAACAACGACATTCTGCGGGAAATCGTGATCAAG  
CCTGGAGATGTCCACGTGTTGGATGACCCAGACACGGCGGCGAAGTACGACGAAGGTCGCTTGGTTAACGCCG  
GC

## 7.2. Anexo II

Tabla de cebadores utilizados y sus características

Nombre	Secuencia	Longitud	% CG	Posición inicio
pQTEV3U	TATAAAAATAGGCGTATCACGAGG	24	37.5 %	-
pQTEV3L	CCAGTGATTTTTTTCTCCATTTT	23	30.4 %	-
MutRpoC2_LD	CAACGACATTCTGCGGGAAATCGTGATCAAGCCTGGAGATGTCCACGTGTTGG	52	52.8 %	262
MutRpoC2_LR	TTCTGGGTGACCGAAACAATCCCAGTCGTCTGGCAAAGATGTCTTTGACGA	52	48.1 %	161
MutRpoC2_CD	CAACGACATTCTGCGGGA	18	56 %	262
MutRpoC2_CR	TTCTGGGTGACCGAAACAATC	21	48 %	192
PIB3	TGGGCTTCATTAAGTGG	20	60 %	1780
PIG1	AGCCGCTACCGACTCAAAC	18	44.4 %	305

\*LD: Largo Directo, LR: Largo Reverso, CD: Corto Directo, CR: Corto Reverso, PIB: Primer Interno Beta', PIG: Primer Interno Gamma

### 7.3. Anexo III

#### Resultados de secuenciación

Secuenciación con PIB3 de mutagénesis de rpoC2 no corregida

```
GGTGGAAATCTCCAAAAAAGGCCGGGCTAAGGCGAAAACAGGGCTACGAGGTGACCAAGGGCGGCACCCTGCT
CTGGATTCCCGAAGAGACCCACGAGGTCAACAAGGACATCTCCTTGCTGAACGTCGAAGACGGTCAGTTGGTG
GAAGCTGGCAGCGGAAGTCGTCAAAGACATCTTTTGCCAGACGACTGGGATTGTTTCGGTCACCCAGAACAACG
ACATTCAGCGGGAAATCGTGATCAAGCCTGGAGATGTCCACGTGCAACGACATTCTGCGGGAAATCGTGATCA
AGCCTGGAGATGTCCACGTGTTGGATGACCCAGACACGGCGGCGAAGTACGACGAAGGTCGCTTGGTTAACGC
CGGCGAAGAAGTCTTCCAGGTTTGACGGCCGAACAACCTGGTCTGGGCAGAAGCCGTTGATGGTACGGACGGC
CCGCTCCTCTTGCTGCGACCGGTGCAAGAGCTAGTCATTCTGATGAGCCGCCAGTACCGAGCCAAGACTCCT
CGCAAGAGTCAAGCAGCCGTTTCGATTTCGCCTGCGTGCCTGCAACGTCTGCAGTTCCAAGATGGTGAACGGAT
CAAGTCCGTTGAGGGAGTAGATCTCCTGCGGACCCAGTTAGTGCTGGAGTCTGAAGAAGGTTTCTCCAGCTC
TCGACTGACATTGAACTGCTGCCGGACAGCAAGGATCCTGAAACGCTGCGCCTGCAGTTAGTAATCATTGAGC
CTGTGGTGATTCGTGCGGACGTGCTTCTGACACCACCCACGGCAGTACGCACACCGAGTTGCGCGTTAAAGA
TGGCCAAAAGGTGAAGCCGGGCGCCGTATTGCTTGCACGCAAATTCAGTGCAAGGAAGCCGGTGTGGTTTCGC
GGCATTCAAGAGGGCAGCGAAGCGGTTTCGTGCTCCTTTGGTTCGAGCGGGAACGGGACTGCGTCACCCTTGACC
TTGATGTCACTGC
```

Secuenciación con pQTEV3U de clonación pQLink+RpoZ+RpoC1

```
CACCATTTGTGCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAATTGT
GAGCGGATAACAATTTACACAGAATTCATTAAGAGGAGAAATTAAGTATGGGATCCATGCTCCAGCGCTTC
GATCTGGATTCTCAAGATCTGCTCTTCAAAGCTGAGAGCCTGATTGTTAACTCCACGAATCGCTATCACGTCA
CCCTACAAATCGCCCGTTCGCGCCAAGCAGGCTCGCTACGAAGAGATGGAAAACCTCAGCGAAGAACTGGCAT
TAAGCCCGTGTGCGCGCCATTTTAGAAATGTGACAGCAGCTCAATCAACCCGAGATTATCGGCGGGTGAAAG
CTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTTCAG
AACGCTCGGTTGCCGCGGGCGTTTTTTTATTGGTGAGAATCCAAGTAACAACACCATTTGTGCGAGAAATCATA
AAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAATTGTGAGCGGATAACAATTTACACA
GAATTCATTAAGAGGAGAAATTAAGTATGTCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAAC
CCACTCGACTTCTTTTGAATATCTTGAAGAAAAATATGAAGAGCATTGTATGAGCGCGATGAAGGTGATAA
ATGGCGAAACAAAAGTTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGATGGTGATGTTAAA
TTAACACAGTCTATGGCCATCATACTGTTATATAGCTGACAAGCACAAACATGTTGGGGTGGTTGTCCAAAAGAG
CGTGCAGAGATTTCAATGCTTGAAGAGGAGCGGTTTTGGATATTAGATACGGTGTTCGAGAAATTCATATA
GTAAAGACTTT
```

Secuenciación con PIG1 de clonación pQLink+RpoZ+RpoC1

```
GGCATCCCGAGCTACATCGCGATTCTGCTCGACATGCCCTGCGAGACGTTGAGCAAATCGTCTATTTCAACT
CCTACGTCGTCCTCAACCCAGGCAACCACAGCGAGCTGCAGTACAAGCAGCTACTGAACGAAGACCAGTGGAT
GGAGATCGAGGACCAAATCTACGCTGAGGAATCTGACCTCGAAGGTATTGAGGTTGGCATCGGTGCAGAAGCC
CTGCAACAGTTGCTGCAGGATTTGAACCTCAATGAAGAGTCGGAAAACCTCCGTCAAGAGATTGCTGAGTCGA
AGGGCCAAAACGGGCAAGTTAATCAAACGCCTGCGTGTGATTGATAACTTCATTGGCACAGAGTCTCGGCC
CGAGTGGATGGTGTGAATGTCATCCCGGTGATCCCTCCCGATCTCCGCCCGATGGTGCAGTTGGATGGTGGC
CGCTTTGCCACCAGTGACCTCAACGATCTCTACCGTCCGGTGTGATCAACCCGAATAACCGGTTAGCGCGTCTGC
AGGAGATTCTCGCACCTGAGATTATCGTCCGTAACGAAAACGGATGTTGCAGGAAGCAGTAGACGCTCTGAT
CGACAACGGTTCGTGTTGGCCGTACGGTTCGTGCGTGCCTGCAACAACCGACCCCTCAAGTCCCTCTCGGACATCATC
GAAGGGAAACAGGGTTCGCTTCCGGCAAAAATTTGCTGGGTAACGGGTTGATTACTCCGGTCCGGTCCGGTTCATCG
TCGGTTCGGCCCCAACCCTGAAAATTCACCAGTGTGGTCTGCCACGGGAAATGGCGATCGAGCTGTTCCAGCCGT
TTGTTATCCATCGCTGATCAAAAATCACAGGCATCAACAACATCAAGCAGGCCAAAAAACTCATTGAGAAG
AACGATCCCTTGATTTGGGGATGTCTTGGAAAGAGGTGATTG
```

Secuenciación con PIG1 de clonación pQLink+RpoZ+RpoC2+RpoD1

ATCCCGAGCTACATCGCGATTCTGCTCGACATGCCCCCTGCGAGACGTTGAGCAAATCGTCTATTTCAACTCCT  
ACGTCGTCCTCAACCCAGGCAACCACAGCGAGCTGCAGTACAAGCAGCTACTGAACGAAGACCAGTGGATGGA  
GATCGAGGACCAAATCTACGCTGAGGAATCTGACCTCGAAGGTATTGAGGTTGGCATCGGTGCAGAAGCCCTG  
CAACAGTTGCTGCAGGATTTGAACCTCAATGAAGAGTCGGAAAAACTCCGTCAAGAGATTGCTGAGTCGAAGG  
GCCAAAAACGGGCCAAGTTAATCAAACGCCTGCGTGTGATTGATAACTTCATTGGCACAGAGTCTCGGCCCGA  
GTGGATGGTGTGAATGTCATCCCGGTGATCCCTCCCGATCTCCGCCCGATGGTGCAGTTGGATGGTGGCCGC  
TTTGCCACCAGTGACCTCAACGATCTCTACCGTCGGGTGATCAACCGGAATAACCGGTTAGCGCGTCTGCAGG  
AGATTCTCGCACCTGAGATTATCGTCCGTAACGAAAAACGGATGTTGCAGGAAGCAGTAGACGCTCTGATCGA  
CAACGGTCGTCTGGCCGTACGGTCGTCCGTGCCAACCAACCGACCCCTCAAGTCCCTCTCGGACATCATCGAA  
GGGAAACAGGGTCGCTTCCGGCAAAACTTGCTGGGTAAACGGGTTGATTACTCCGGTCGGTCGGTCATCGTCG  
TCGGCCCCAACCTGAAAATTCACCAGTGTGGTCTGCCACGGGAAATGGCGATCGAGCTGTTCCAGCCGTTTGT  
TATCCATCGCCTGATCAAAAATCACAGCATCAACAACATCAAGCAGGCCAAAAAACTCATTGAGAAGAACGAT  
CCCTTGATTTGGGATGTCTTGAAGGAGGTGATTGAAGGGCCACCCCTGTGATGCTGAACCGGGCCCCCGACG  
GCTGCACCGTTTTGGGGATTCAAGGCCTTTCGAAACCGATCTTGGGTAGAAAGCCGTGCG