

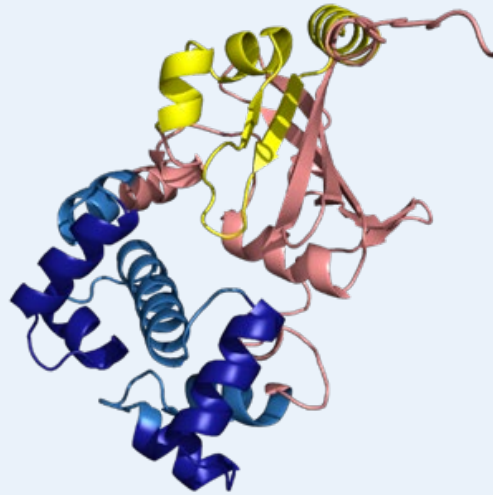
UNIVERSIDAD DE ZARAGOZA

FACULTAD DE CIENCIAS

Departamento de Bioquímica y Biología Molecular y Celular

TRABAJO FIN DE GRADO

Purificación y caracterización del
regulador transcripcional Al1651 de la
cianobacteria *Anabaena sp.* PCC7120



Gabriel Sánchez Atrián

Junio 2023



**Universidad
Zaragoza**



Facultad de Ciencias
Universidad Zaragoza

*Purificación y caracterización
del regulador transcripcional
All1651 de la cianobacteria
Anabaena sp. PCC7120*

UNIVERSIDAD DE ZARAGOZA

FACULTAD DE CIENCIAS

Departamento de Bioquímica y Biología Molecular y Celular

**Trabajo de Fin de Grado para optar al título de
Graduado en Biotecnología**

Alumno: Gabriel Sánchez Atrián

Directores: Dra. Emma Sevilla Miguel

Dra. María Teresa Bes Fustero

Jorge Guío Martínez

Agradecimientos

En las primeras líneas de este trabajo quiero dar las gracias a todas aquellas personas que lo han hecho posible y a todas las personas que me han acompañado durante estos 4 años.

A Jorge, por enseñarme, guiarme y ayudarme desde el primer día del verano que entré al laboratorio para hacer las prácticas.

A Emma Sevilla y a María Teresa Bes, por su ayuda en la realización y redacción del TFG.

A Marisa Peleato, por ayudarme y darme consejos cuando estaba perdido y no tenía claro nada.

A mis compañeras de laboratorio, Ángela, Inés y Marta, por el equipo que hicimos, por las risas y por enseñarme que el laboratorio puede ser una zona de confort entre tanto estrés.

A mis compañeros de clase y de vida, por todas las risas, las charlas, los cafés, las dianas y billares, la ayuda, las fiestas, las lágrimas... que hemos compartido. Porque la vida nos siga juntando y seamos nosotros mismos y felices.

A mi madre Carmina, mi padre Salvador y mi hermana Sandra, ya no solo por hacer posible que haya podido estudiar, sino por darme los valores que me hacen ser cómo soy y por ayudarme a que sea feliz, aunque a veces me resulte complicado.

Y, por último, un agradecimiento especial a mi hermana Sandra y a dos compañeros de vida, David y Javier, por ser luces que me guiaron cuando todo se nubló.

Índice

| | |
|--|-----------|
| 1. Resumen/Abstract | 1 |
| 1.1. Resumen..... | 1 |
| 1.2. Abstract..... | 1 |
| 2. Introducción | 2 |
| 2.1. Las cianobacterias..... | 2 |
| 2.2. Homeostasis de metales en cianobacterias: las proteínas FUR..... | 2 |
| 2.2.1. FurA..... | 3 |
| 2.2.2. FurB..... | 4 |
| 2.2.3. FurC..... | 4 |
| 2.3. Redes de regulación en <i>Anabaena sp.</i> PCC7120..... | 4 |
| 3. Objetivos | 6 |
| 4. Materiales y Métodos | 6 |
| 4.1. Microorganismos y condiciones de cultivo..... | 6 |
| 4.1.1. Cepas <i>Escherichia coli</i> | 6 |
| 4.1.2. Condiciones de cultivo..... | 6 |
| 4.2. Aislamiento y análisis de DNA..... | 7 |
| 4.2.1. Reacción en cadena de la polimerasa (PCR)..... | 7 |
| 4.2.2. Purificación y cuantificación de productos de PCR..... | 7 |
| 4.2.3. Electroforesis del DNA..... | 7 |
| 4.3. Sobreexpresión de proteínas recombinantes..... | 8 |
| 4.3.1. Transformación de <i>E.coli</i> mediante choque térmico..... | 8 |
| 4.3.2. Prueba de sobreexpresión a pequeña escala..... | 8 |
| 4.3.3. Prueba de solubilización..... | 8 |
| 4.4. Purificación de proteínas recombinantes..... | 9 |
| 4.4.1. Obtención de biomasa..... | 9 |
| 4.4.2. Lisis celular..... | 10 |
| 4.4.3. Purificación por cromatografía de afinidad a iones metálicos..... | 10 |
| 4.5. Técnicas analíticas de proteínas..... | 11 |
| 4.5.1. Precipitación con ácido tricloroacético (TCA)..... | 11 |
| 4.5.2. Electroforesis desnaturizante en geles de poliacrilamida..... | 11 |
| 4.5.3. Cuantificación de las proteínas por espectroscopía..... | 11 |
| 4.6. Ensayos de retardo en gel..... | 12 |
| 4.7. Técnicas bioinformáticas..... | 12 |
| 5. Resultados | 13 |
| 5.1. Análisis bioinformático del regulador All1651..... | 13 |
| 5.2. Prueba de sobreexpresión..... | 16 |
| 5.3. Análisis de la solubilización empleando diferentes tampones..... | 16 |
| 5.4. Purificación con cloruro de guanidinio..... | 17 |
| 5.5. Análisis de la solubilización empleando diferentes estirpes y condiciones de cultivo..... | 18 |
| 5.6. Purificación a partir de biomasa generada en la estirpe pG-KJE8..... | 19 |
| 5.7. Estudio de unión al DNA mediante ensayos de retardo en gel..... | 20 |

| | |
|---|-----------|
| 6. Discusión..... | 21 |
| 6.1. El regulador All1651 es un regulador de la familia AraC que contiene un dominio PAS..... | 21 |
| 6.2. La sobreexpresión en la estirpe pG-KJE8 permite la purificación del regulador All1651..... | 22 |
| 6.3. El regulador All1651 muestra actividad de unión al DNA..... | 23 |
| 7. Conclusiones/Conclusions..... | 23 |
| 8. Bibliografía..... | 24 |

Abreviaturas

Å: Ångstrom

Abs: Absorbancia

AIM: medio de autoinducción

ATP: Adenosín trifosfato

c: concentración

Cys: Cisteína

C-terminal: carboxilo terminal

DNA: Ácido desoxirribonucleico (ADN)

dNTPs: mezcla de los cuatro desoxirribonucleótidos trifosfato

DTT: 1,4-ditiotreitol

EDTA: Ácido etilendiaminotetraacético

EMSA: ensayo de retardo en gel (*Electrophoretic Mobility Shift Assay*)

FUR: Regulador de la absorción de hierro (*Ferric Uptake Regulator*)

IMAC: Cromatografía de afinidad por unión a metales (*Immobilized Metal Ion Affinity Chromatography*)

IPTG: Isopropil-β-D-tiogalactopiranosido

kDa: Kilodalton

l: Longitud de la cubeta

LB: Luria-Bertani

mA: Miliamperios

Motivo HTH: motivo hélice – giro – hélice (helix – turn – helix)

NADPH: Nicotinamida adenina dinucleótido fosfato

O/N: durante la noche (over night)

PAS: acrónimo de Per (proteína del periodo circadiano), Arnnt (proteína translocadora del receptor nuclear Ah), Sim (proteína *single-minded* que regula el desarrollo del sistema nervioso central de *Drosophila*), proteínas en las que se describió por primera vez este dominio

pb: Pares de bases

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

pI: Punto isoeléctrico

PMSF: Fluoruro de fenilmetilsulfonilo

PSA: Persulfato amónico

p/v: Porcentaje peso/volumen

RNA: Ácido ribonucleico

rpm: Revoluciones por minuto

SDS: Dodecilsulfato sódico

SDS-PAGE: Electroforesis desnaturalizante en gel de poliacrilamida

TBE: Tris Borato EDTA

TCA: Ácido tricloroacético

TEMED: N, N, N', N'-tetrametiletildiamina

Tris: Tris(hidroximetil)aminometano

U: Unidad de actividad enzimática

UV: Ultravioleta

V: voltios

v/v: Porcentaje volumen/volumen

ϵ : Coeficiente de extinción molar

1. Resumen / Abstract

1.1. Resumen

Las cianobacterias son microorganismos procariotas fotosintéticos cuyo estudio es de gran interés, debido a su importancia ecológica y sus numerosas aplicaciones biotecnológicas. Uno aspecto crucial en la fisiología de estos organismos es el mantenimiento de la homeostasis de metales, que en la mayoría de procariotas está controlado por la familia de reguladores transcripcionales FUR (Ferric Uptake Regulator). La cianobacteria *Anabaena* sp. PCC7120 presenta tres parálogos de proteínas FUR, FurA, FurB y FurC, los cuales, además de la homeostasis de metales, controlan diversos procesos celulares, como la fotosíntesis o el metabolismo del nitrógeno, actuando como reguladores globales. Recientemente, se ha descrito que los reguladores FUR en *Anabaena* sp. PCC 7120 son capaces de regular la expresión de reguladores transcripcionales, sistemas de dos componentes o factores sigma, orquestando redes de regulación transcripcional que les permitirían controlar muchos procesos celulares de forma indirecta.

En este trabajo se ha llevado a cabo la caracterización y purificación del regulador transcripcional All1651, el cual está regulado tanto por FurA como por FurC. Para ello se han realizado estudios bioinformáticos que ha permitido identificar su familia y dominios y conocer sus propiedades fisicoquímicas. Por otro lado, se han optimizado las condiciones de sobreexpresión de esta proteína en *E. coli* y se ha establecido un protocolo de purificación del regulador mediante el uso de cromatografía por afinidad a iones metálicos (IMAC). Finalmente, se ha estudiado su actividad de unión al DNA *in vitro* mediante ensayos de retardo en gel (EMSA).

1.2. Abstract

Cyanobacteria are photosynthetic prokaryotic microorganisms whose study is of great interest due to their ecological importance and their numerous biotechnological applications. A crucial aspect in the physiology of these organisms is the maintenance of metal homeostasis, which in most prokaryotes is controlled by the FUR (Ferric Uptake Regulator) family of transcriptional regulators. The cyanobacterium *Anabaena* sp. PCC7120 contains three FUR protein paralogs, FurA, FurB and FurC, which, in addition to metal homeostasis, control diverse cellular processes such as photosynthesis or nitrogen metabolism, acting as global regulators. Recently, it has been described that FUR regulators in *Anabaena* sp. PCC 7120 are able to control the expression of transcriptional regulators, two-component systems or sigma factors, orchestrating transcriptional regulatory networks that allow them to control many cellular processes indirectly.

In this work we have performed a characterization and purification of the transcriptional regulator All1651, which is regulated by both FurA and FurC. For this purpose, bioinformatic studies have been carried out to identify its family and domains and to determine its physicochemical properties. On the other hand, the conditions for overexpression of this protein in *E. coli* have been optimized and a protocol for purification of the regulator using metal ion affinity chromatography (IMAC) has been established. Finally, its *in vitro* DNA binding activity has been studied by electrophoretic mobility shift assays (EMSA).

2. Introducción

2.1. Las cianobacterias

Las cianobacterias son microorganismos procariotas que llevan a cabo la fotosíntesis oxigénica, es decir, son capaces de captar la energía luminosa y utilizar el agua como donador de electrones para obtener energía en forma de ATP y poder reductor en forma de NADPH, lo que permite la fijación de CO₂ a compuestos orgánicos (1). Puesto que en la fotosíntesis oxigénica se libera oxígeno, el surgimiento de las cianobacterias en la era Proterozoica hace unos 3500 millones de años ocasionó un aumento en los niveles de oxígeno en nuestro planeta, lo que permitió transformar la atmósfera reductora en una atmósfera oxidante y dio lugar al surgimiento de otras formas de vida (2). Además, según la teoría endosimbiótica, se considera que las cianobacterias son el ancestro evolutivo de los cloroplastos y por tanto responsables del surgimiento de los organismos eucariotas fotosintéticos (3).

Gracias a su enorme diversidad metabólica y su capacidad para tolerar una gran variedad de estreses ambientales, las cianobacterias han sido capaces de colonizar la mayor parte de los hábitats terrestres, desde ecosistemas marinos o fluviales hasta regiones polares, pasando por suelos, regiones desérticas géiseres o criocónitas (4). En estos hábitats, las cianobacterias desempeñan un papel ecológico fundamental, ya que son responsables de la fijación de carbono y la producción de oxígeno y por tanto los productores primarios en estos ecosistemas. Además, algunas cianobacterias son capaces de llevar a cabo la fijación de nitrógeno atmosférico, un proceso que en algunas especies tiene lugar en células especializadas que reciben el nombre de heterocistos, enriqueciendo el suelo y agua de compuestos nitrogenados que podrán ser consumidos por otros organismos (5).

Sin embargo, a pesar de tener un papel muy importante en la naturaleza, la población de cianobacterias ha de mantenerse regulada en los ecosistemas, ya que un crecimiento excesivo de algunos de sus géneros puede dar lugar a las conocidas floraciones o *blooms*. Las floraciones tienen un impacto negativo en la calidad del agua y pueden afectar a la biodiversidad de las comunidades biológicas, ya sea por la reducción del resto de flora acuática o por la producción de cianotoxinas (6).

El estudio de la fisiología de las cianobacterias es de gran interés, no solo por su enorme importancia ecológica, sino también debido a sus numerosas aplicaciones biotecnológicas. Su capacidad para fijar el nitrógeno atmosférico hace que tengan un gran potencial como biofertilizantes en agricultura sostenible y algunas especies como *Spirulina* pueden ser utilizadas en suplementos alimenticios. Además las cianobacterias son capaces de captar y almacenar una variedad de compuestos contaminantes, como metales pesados o xenobióticos, y permiten la obtención de biocombustibles como el hidrógeno o el biodiesel (7).

2.2. Homeostasis de metales en cianobacterias: las proteínas FUR

Uno de los aspectos más importante de la fisiología de las cianobacterias es el estudio de la homeostasis de metales, pues estos organismos tienen un requerimiento de metales mucho más alto que otros organismos procariotas (8). Esto se debe principalmente a dos procesos celulares: la fotosíntesis y, en especies fijadoras, la fijación de nitrógeno. Metales como el hierro, el manganeso, el magnesio y el cobre son cofactores esenciales de pigmentos fotosintéticos como la clorofila, así como de los fotosistemas I y II, del citocromo b₆f y de otros transportadores de electrones de la cadena fotosintética como la plastocianina, una metaloproteína de cobre (8). La fijación de nitrógeno también tiene una alta demanda de metales, ya que el complejo enzimático que lleva a cabo este proceso la nitrogenasa, contiene numerosos cofactores metálicos. En concreto, su componente dinitrogenasa reductasa contiene un centro sulfoférico 4Fe-4S y su componente dinitrogenasa contiene dos centros sulfoféricos P (Fe₈S₇) y dos cofactores Fe-Mo (Fe₇MoS₉C) (9).

Sin embargo, pese a ser esenciales, los metales resultan tóxicos a altas concentraciones, pues pueden catalizar la generación de especies reactivas de oxígeno como radicales hidroxilo o peróxido de hidrógeno por la reacción de Fenton (8). Por ello, es necesario que existan sistemas de regulación que garanticen que los metales estén en niveles apropiados y que por tanto haya una correcta homeostasis de metales en la cianobacteria.

Una de las familias de proteínas más relevantes en la regulación de la homeostasis de metales es la familia de reguladores transcripcionales FUR (Ferric Uptake Regulator) (10). Estas proteínas presentan un peso molecular de 13 a 20 kDa y se caracterizan por la presencia de varios dominios: un dominio de unión al DNA en la región amino terminal y un dominio de unión a metal regulador en la región carboxilo terminal que está también implicado en la dimerización (10). Contienen además dos motivos característicos, un motivo rico en histidinas en la región bisagra que conecta ambos dominios y uno o varios motivos CXXC, que pueden estar implicados en la unión de metal estructural o en intercambios tiol-disulfuro (10).

Dentro de la familia de proteínas FUR se encuentran las proteínas Fur (Ferric Uptake Regulator), Zur (Zinc Uptake Regulator), Mur (Manganese Uptake Regulator), Nur (Níquel Uptake Regulator), PerR (Peroxide Responsive Regulator) e Irr (Iron Response Regulator) (10). Los reguladores Fur, Zur, Mur y Nur regulan la homeostasis de hierro, zinc, manganeso y níquel, respectivamente, metales que al unirse al dominio de unión al metal regulador modulan su capacidad de interaccionar con el DNA y por tanto sus genes diana. No obstante, algunos miembros de la familia como Irr y PerR responden a otros estímulos como disponibilidad de hemo y estrés oxidativo, respectivamente (10).

En el caso de la cianobacteria objeto de estudio en este trabajo, *Anabaena sp.* PCC 7120 se han identificado tres parálogos de proteínas FUR: FurA, FurB y FurC (11).

2.2.1. FurA

FurA es el principal regulador de la homeostasis de hierro en la cianobacteria *Anabaena*, controlando a nivel transcripcional la expresión de transportadores de hierro, enzimas de síntesis de sideróforos o genes implicados en la síntesis de tetrapirrol, modulando así la captación, transporte, almacenamiento y utilización de este metal (12). No obstante, además de la homeostasis de hierro, se ha descrito que FurA controla muchos otros procesos celulares como el metabolismo del nitrógeno, la fotosíntesis, la morfología celular o el estado redox, actuando como un regulador global (13).

FurA actúa como un metaloregulador dependiente de hierro, que requiere de la presencia de este metal para unirse al DNA, así como de condiciones reductoras (14). Este regulador contiene cinco cisteínas, cuatro de las cuales forman dos motivos CXXC (C₁₀₁XXC₁₀₄ y C₁₄₁XXC₁₄₄) que muestran actividad disulfuro reductasa, lo que refleja una implicación de este regulador en las vías de señalización redox de la célula (15). En concreto, se ha descrito que FurA coexiste en varios estados redox en el interior celular que modulan su actividad, ya que sus cisteínas son capaces de formar puentes disulfuro (16), y que la interconversión entre estas formas está mediada por la tiorredoxina A y es dependiente de transiciones luz-oscuridad, lo que conecta a este regulador con la actividad de la cadena de transporte de electrones de la fotosíntesis (17).

Finalmente, FurA también se ha descrito como una proteína sensora capaz de detectar ciertos estreses y señales celulares mediante la interacción con determinados metabolitos, como ejemplo es la unión con el grupo hemo (18) y la interacción con el 2-oxoglutarato en deficiencia de nitrógeno (19).

2.2.2. FurB

FurB es el principal regulador de la homeostasis de zinc en *Anabaena* y ha sido propuesto como el regulador Zur en esta cianobacteria, siendo responsable de modular la expresión de genes como metaloproteínas dependientes de zinc y transportadores de este metal (20). Además de la homeostasis de zinc, FurB también desempeña un importante papel en la defensa frente al estrés oxidativo al regular genes como superóxido dismutasas, catalasas y peroxirredoxinas que permiten eliminar especies reactivas de oxígeno (21). Asimismo, protege al DNA ante dicho estrés uniéndose a él inespecíficamente (22). Recientemente, el análisis transcriptómico de una estirpe de delección de FurB permitió determinar que este regulador estaba implicado en control del tráfico de metales, el desarrollo de heterocistos y la formación de biofilms (23).

2.2.3. FurC

FurC fue propuesto como el regulador PerR en la cianobacteria *Anabaena*, pues se determinó que modulaba la expresión de genes implicados en la respuesta a estrés oxidativo como la peroxirredoxina A y la sulfirredoxina A y que su unión al DNA estaba modulada *in vitro* por un mecanismo de oxidación catalizada por metal (metal catalysed oxidation, MCO), propio de los reguladores PerR (24). Se describió que en presencia de metal correpresor FurC era capaz de unirse al DNA y reprimir la expresión de sus genes diana pero que en presencia de agentes oxidantes como el peróxido de hidrógeno se producía una oxidación de sus histidinas catalizada por este metal que daba lugar a un cambio conformacional que hacía que FurC se liberase del DNA y se expresasen sus genes diana (24).

Posteriormente se demostró *in vivo* que la actividad FurC estaba modulada por estrés oxidativo, pues se describió que producía la desrepresión de sus genes diana en presencia de peróxido de hidrógeno y también se encontró que modulaba otros procesos como la fotosíntesis y la división celular (25). El análisis transcripcional de una estirpe de sobreexpresión de este regulador mostró que FurC controlaba genes implicados en muchos procesos celulares, como el metabolismo del hierro o el metabolismo energético, actuando por tanto como un regulador global (26). En este trabajo, también se puso de manifiesto que FurC jugaba un papel clave en el metabolismo del nitrógeno y la diferenciación de heterocistos, pues era un regulador de genes importantes en estos procesos y su sobreexpresión impedía la formación de heterocistos, células especializadas en la fijación de nitrógeno (26).

2.3. Redes de regulación transcripcional en *Anabaena sp.* PCC 7120

La regulación de la transcripción en cianobacterias se lleva a cabo mediante cuatro mecanismos: reguladores transcripcionales, sistemas de dos componentes, factores sigma y RNA antisentido (27).

Los reguladores transcripcionales se unen a secuencias específicas del DNA para modular la expresión génica. En algunos casos dan lugar a un aumento en la transcripción, por ejemplo favoreciendo el ensamblaje de la RNA polimerasa, y reciben el nombre de activadores mientras que en otros casos disminuyen la transcripción de sus genes diana, por ejemplo impidiendo la elongación de la RNA polimerasa, y reciben el nombre de represores (28). Existen muchas familias de reguladores transcripcionales, aunque la mayoría suelen tener dos dominios: un dominio de unión a DNA, que suele presentar uno o varios motivos HTH (hélice-giro-hélice) de unión a DNA, y un dominio capaz de percibir diferentes señales. La mayoría de reguladores transcripcionales suelen unirse al DNA en forma de dímeros y reconocer secuencias palindrómicas de DNA (28).

Los sistemas de dos componentes son sistemas de transducción de señal que permiten a los organismos detectar y responder a cambios en su entorno y están presentes en una gran variedad de organismos desde bacterias a plantas y animales (29). Estos sistemas se componen de dos elementos una sensor o histidin quinasa, que es capaz de detectar señales específicas del entorno y, un regulador de respuesta, que desencadena respuestas celulares en respuesta a esta señal. La histidin quinasa actúa como un interruptor molecular que al percibir la señal se autofosforila y transmite el grupo fosfato al regulador de respuesta, lo que da lugar a cambios en su actividad (29). La mayoría de reguladores de respuesta actúan como factores de transcripción que modulan la expresión de genes, aunque algunos pueden mediar interacciones proteína-proteína o tener actividades enzimáticas o de unión a RNA (29).

El factor sigma (σ) es una de las subunidades de la RNA polimerasa, que guía a la holoenzima a las secuencias promotoras y permiten que se inicie la transcripción. Además del factor sigma canónico existen factores sigma alternativos que pueden reemplazarlo en diferentes estreses, modulando la interacción de la RNA polimerasa con los promotores y cambiando el patrón de expresión de genes (30). *Anabaena* sp. PCC 7120 tiene 12 factores sigma (31), los cuales pueden reemplazar al factor sigma canónico en condiciones como alta salinidad o deficiencia de nitrógeno y orquestar respuestas frente a estos estreses. Por su parte, los RNA antisentido son capaces de unirse a las moléculas de RNA mensajero, mediando su degradación o bloqueando su traducción, y en cianobacterias juegan un papel muy importante en la modulación del metabolismo del nitrógeno y la fotosíntesis (32).

Todos estos mecanismos de regulación transcripcional no actúan de forma independiente, sino que están perfectamente coordinados y regulados para permitir que en cada momento se expresen los genes apropiados, constituyendo redes de regulación transcripcional. En la cianobacteria *Anabaena* sp. PCC 7120 se ha descrito la existencia de una red de regulación transcripcional mediada por las proteínas FUR, en la que se han identificado un gran número de reguladores transcripcionales, sistemas de dos componentes y factores sigma que están directamente regulados por FurA, FurB y FurC (Jorge Guío, datos no publicados). De esta manera, las proteínas FUR de *Anabaena* sp. PCC 7120, además de regular sus genes diana de forma directa, estarían regulando muchos procesos de forma indirecta.

En concreto, se ha descrito que FurA modula cuatro reguladores transcripcionales y seis sistemas de dos componentes, FurB modula cuatro genes de sistemas de dos componentes y FurC modula tres reguladores transcripcionales y tres sistemas de dos componentes (Jorge Guío, datos no publicados). Además, se ha determinado que los 12 factores sigma de *Anabaena* están modulados al menos por un parálogo FUR (Jorge Guío, datos no publicados). No obstante, se desconoce la función y los procesos controlados por la mayor parte de los reguladores transcripcionales y sistemas de dos componentes que forman parte de esta red.

A partir de estudios del transcriptoma de mutantes de delección y sobreexpresión de los distintos homólogos FUR de la cianobacteria realizados en el grupo de trabajo, se había observado que el gen *all1651* estaba regulado de forma directa tanto por FurA como por FurC y por tanto podría estar desempeñando funciones importantes en la cianobacteria *Anabaena* sp. PCC7120. Así pues, con el fin de ampliar el conocimiento sobre el regulón de las proteínas FUR y caracterizar esta red de regulación, en este trabajo nos propusimos estudiar el regulador transcripcional All1651.

3. Objetivos

El propósito de este Trabajo Fin de Grado es purificar y estudiar el regulador All1651 de la cianobacteria *Anabaena* sp. PCC 7120. Para ello se establecieron los siguientes objetivos:

- Realizar un estudio bioinformático del regulador All1651 para conocer su familia y dominios, así como modelar su estructura y analizar sus relaciones filogenéticas.
- Optimizar la sobreexpresión de este regulador en *E. coli* para poder llevar a cabo su purificación mediante sobreexpresión heteróloga
- Establecer un protocolo de purificación para el regulador All1651 empleando cromatografía de afinidad a iones metálicos inmovilizados (IMAC)
- Determinar unas condiciones iniciales que permitan la unión de All1651 a DNA mediante ensayos de retardo en gel (EMSA)

4. Materiales y métodos

4.1 Microorganismos y condiciones de cultivo

4.1.1. Cepas de *Escherichia coli*

En este estudio se han empleado las cepas de *Escherichia coli* recogidas en la tabla 1.

| Cepa | Características | Origen |
|--------------------|---|------------------|
| BL21 (DE3) | F ⁻ , ompT, hsdS _B (r _B ⁻ , m _B ⁻), dcm, gal, λ (DE3) | Novagen |
| BL21 (DE3) pG-KJE8 | <i>E. coli</i> BL21 (DE3) conteniendo el plásmido pG-KJE8, que codifica para las chaperonas DnaK-DnaJ-GrpE y GroES-GroEL. Resistente a cloranfenicol (50 µg/mL) | Takara Bio Group |

Tabla 1. Cepas de *E. coli* utilizadas en el estudio.

La cepa BL21 (DE3) de *E. coli* es una cepa ampliamente utilizada en la sobreexpresión y producción de proteínas recombinantes, pues carece de las proteasas OmpT y Ion para evitar la degradación de las proteínas heterólogas (33, 34). Esta cepa contiene además el locus DE3 que codifica la RNA polimerasa del fago T7 bajo el control del promotor *lacUV5*, lo que permite que la sobreexpresión de la proteína de interés sea inducible por lactosa y por IPTG.

La cepa BL21 (DE3) pG-KJE8 consiste en la cepa BL21 de *E. coli* conteniendo el plásmido pG-KJE8, un plásmido que codifica cinco genes de chaperonas (*dnaK*, *dnaJ*, *grpE*, *groEL* y *groES*) bajo el control de los promotores inducibles de arabinosa (*araB*) y de tetraciclina (*Pzt-1*), controlados por los reguladores *araC* y *tetR*, respectivamente (35).

4.1.2. Condiciones de cultivo

Las cepas de *E. coli* se cultivaron en medio de cultivo Luria Bertani (LB), compuesto por 10 g/L de triptona, 5 g/L de extracto de levadura y 5 g/L de cloruro de sodio, que se disolvieron en agua destilada y la solución se esterilizó en un autoclave. Para optimizar las condiciones de sobreexpresión, también se cultivaron células de *E. coli* en medio de autoinducción (AIM). Este medio se compone de 10 g/L de triptona, 5 g/L de extracto de levadura, 0,5 g/L de glucosa, 2 g/L de lactosa, 3,3 g/L de (NH₄)₂SO₄, 6,8 g/L de KH₂PO₄, 7,1 g/L de Na₂HPO₄, 0,15 g/L de MgSO₄ y 0,03 g/L de elementos traza (50 mM FeCl₃, 20 mM CaCl₂·2H₂O, 10 mM MnCl₂·4H₂O, 10 mM ZnSO₄·7H₂O, 2 mM CoCl₂·6H₂O, 2 mM CuCl₂, 2 mM NiCl₂, 2 mM Na₂MoO₄, 2 mM Na₂SeO₃ y 2 mM H₃BO₃). El medio se preparó disolviendo 34,85 g de preparado comercial de medio LB de autoinducción con elementos traza de la casa comercial Formedium[®] en 1 L de agua destilada y se esterilizó en autoclave (36).

La diferencia entre el medio LB y el medio AIM reside en que mientras en el medio LB la expresión de las proteínas se da de forma rápida tras la adición del inductor (IPTG) en el medio AIM la expresión de las proteínas se induce poco a poco por la lactosa a medida que se va consumiendo la glucosa.

Los cultivos en medio líquido se incubaron a 37 °C con agitación continua a 200 rpm en un incubador orbital. Por su parte, los cultivos en medio sólido se incubaron a 37 °C en placas Petri empleando medio LB suplementado con agar al 1,5 % (p/v). Además, cuando fue necesario, los medios se suplementaron con kanamicina y/o cloranfenicol a una concentración final de 50 µg/mL a partir de un stock 1000 veces concentrado, que se preparó en agua destilada en el caso de la kanamicina o en etanol absoluto en el caso del cloranfenicol y se esterilizó con un filtro de 0,22 µm.

4.2. Aislamiento y análisis de DNA

4.2.1. Reacción en cadena de la polimerasa (PCR)

En este estudio, la PCR se utilizó para amplificar la región promotora del gen que codifica el regulador transcripcional All1651, con el fin de, una vez purificado dicho regulador, estudiar si se une a su propio promotor mediante ensayos de retardo en gel (EMSA). La mezcla de reacción contenía 72 µL de agua miliQ estéril, 10 µL de tampón 10x (Tris-HCl 75 mM pH 9,0, KCl 50 mM y $(\text{NH}_4)_2\text{SO}_4$ 20 mM), 3 µL de MgCl_2 50 mM, 2 µL de mezcla de dNTPs 10 mM, 2,5 µL de cebador directo 20 µM, 2,5 µL de cebador reverso 20 µM, 1 µL de Taq DNA polimerasa 1 U/µL (Biotools) y 100 ng de DNA genómico, en un volumen total de 100 µL. Las reacciones de amplificación se realizaron en un termociclador Thermal Cycler 2720 (Applied Biosystems). El programa de PCR y la secuencia de los cebadores se recogen en las secciones 9.1 y 9.2 de Material Suplementario.

4.2.2. Purificación y cuantificación de productos de PCR

Se llevó a cabo una purificación de los productos de PCR, para separar los productos de amplificación del resto de reactivos empleados en la reacción que pueden interferir en posteriores experimentos. Para ello se empleó el kit GFXTM PCR DNA and Gel band Purification de la compañía GE Healthcare. Finalmente, para determinar la concentración de los productos de PCR se utilizó el equipo NanoVue PlusTM de la compañía GE Healthcare.

4.2.3. Electroforesis de DNA

Tras la purificación de los productos de PCR, se llevó a cabo una electroforesis en gel de agarosa para verificar la correcta amplificación y la ausencia de productos inespecíficos. Para ello se preparó un gel de agarosa al 1 % (p/v) disolviendo 50 g de agarosa en 50 mL de tampón TBE (Tris-HCl 90 mM pH 8, ácido bórico 90 mM y EDTA 2 mM) por calentamiento en un microondas. Tras su disolución se dejó enfriar y antes de añadir la mezcla al molde se añadieron 50 µL de bromuro de etidio 0,5 g/mL.

Para analizar los productos de PCR por electroforesis, se mezclaron 4 µL del producto de PCR con 3 µL de tampón de carga (Tris-HCl 50 mM pH 8, 30 % (v/v) de glicerol y 0,25 % (p/v) de azul de bromofenol). La mezcla se cargó en un pocillo del gel y se realizó la electroforesis en tampón TBE a 90 V durante 30 min utilizando una fuente BioRad Powerac300. Pasados 30 min, el gel se reveló con un equipo GelDoc200 de la firma BioRad.

4.3. Sobreexpresión de proteínas recombinantes

El método elegido para la obtención de la proteína de interés fue su sobreexpresión heteróloga en *E. coli* y posterior purificación mediante cromatografía de afinidad a iones metálicos (IMAC).

4.3.1. Transformación de *E. coli* mediante choque térmico

Para llevar a cabo la expresión heteróloga de la proteína de interés se partió de células de *E. coli* BL21 (DE3) o BL21 (DE3) pG-KJE8 termocompetentes, que se transformaron mediante choque térmico con el plásmido pET28a:*all1651*, que contiene el gen que codifica para el regulador All1651 bajo el control del promotor T7 y el operador *lac* y una cola de histidinas en su extremo amino terminal para permitir su purificación. Para llevar a cabo el choque térmico, se descongelaron alícuotas de 100 μ L de células termocompetentes en un baño de hielo durante 10 min. Pasado el tiempo se añadieron, en condiciones de esterilidad, 100 ng del plásmido pET28a:*all1651*, se mezcló suavemente y se incubó 10 min en hielo. A continuación, las células se sometieron a un choque térmico a 42 °C durante 1 min y se incubaron 2 min en hielo. Posteriormente se añadieron 500 μ L de medio LB estéril sin antibiótico y se dejó crecer a las células durante 1 h a 37 °C y 200 rpm. Finalmente, el cultivo se sembró en una placa de LB agar suplementada con 50 μ g/mL de kanamicina (o 50 μ g/mL de cloranfenicol en el caso de las células BL21 (DE3) pG-KJE8) y se incubó O/N a 37 °C en una estufa.

4.3.2. Prueba de sobreexpresión a pequeña escala

Antes de llevar a cabo la sobreexpresión del regulador All1651 a gran escala, se realizó una prueba a pequeña escala para determinar las condiciones en las que la expresión de la proteína era máxima. Para realizar este estudio se seleccionó una de las colonias obtenidas tras el choque térmico, se inoculó en un tubo Falcon que contenía 2 mL de medio LB suplementado con 50 μ g/mL de kanamicina para mantener el plásmido pET28:*all1651* y se incubó a 37 °C y 200 rpm O/N. A la mañana siguiente el cultivo se refrescó en 10 mL de medio LB con 50 μ g/mL de kanamicina y se incubó a 37 °C y 200 rpm hasta que alcanzó una densidad óptica de 0,6 a 600 nm. Los 10 mL de cultivo se dividieron en dos tubos, de los cuales uno se dejó sin inducir y el otro se indujo con 5 μ L de IPTG 1 M.

Se ensayaron diferentes tiempos y temperaturas de inducción: 2 horas a 37 °C, 4 horas a 37 °C, 4 horas a 30 °C y O/N a 15 °C. Para analizar la sobreexpresión de la proteína en cada una de las condiciones se tomó 1 mL de cultivo y se centrifugó en un tubo Eppendorf durante 5 min a 16000 x g. Se descartó el sobrenadante y el precipitado se resuspendió en 75 μ L de agua miliQ estéril. Se añadieron 15 μ L de tampón Laemmli 6x (60 mM de Tris-HCl pH 6,8, 9,6 % (p/v) de SDS, 42 % (v/v) de glicerol, 0,06 % (p/v) de azul de bromofenol y 30 % (v/v) de β -mercaptoetanol), se hirvió a 95 °C durante 5 min y se centrifugó 1 min a 16000 x g. Finalmente, se tomaron 10 μ L de sobrenadante y se analizó la sobreexpresión de la proteína mediante electroforesis en geles de SDS-PAGE como se describe en el apartado 4.5.2.

4.3.3. Prueba de solubilización

Para encontrar unas condiciones óptimas de purificación del regulador All1651 se estudió su solubilidad empleando diferentes tampones y condiciones de inducción. Dado que el punto isoeléctrico teórico de la proteína era de 6,5, se eligió como tampón de trabajo Tris-HCl 50 mM pH 8, cuyo valor de pH está alejado del punto isoeléctrico, favoreciendo la solubilidad de la proteína. Se analizó la solubilidad de la proteína en tampón de trabajo suplementado con 500 mM de NaCl y 10 % de glicerol; con 500 mM de NaCl, 10 % de glicerol y 0,1 % de Tween ; y con 500 mM de NaCl, 10 % de glicerol y 1 U de lisozima. También se analizó su solubilidad en un tampón compuesto por 2 M de cloruro de guanidinio, 100 mM NaH₂PO₄ y 10 mM de Tris-HCl pH 8.

Con respecto a las condiciones de inducción, se prepararon preinóculos de células BL21 y células BL21 pG-KJE8 conteniendo el plásmido pET28:*all1651* y se llevó a cabo la sobreexpresión de la proteína en medio LB y en medio de autoinducción.

Para los ensayos de sobreexpresión en medio LB, en el caso de las células BL21 el preinóculo se refrescó en 10 mL de medio LB con 50 µg/mL de kanamicina y en el caso de las células BL21pG-KJE8, el medio LB se suplementó además con 50 µg/mL de cloranfenicol y con 1 mg/mL de L-arabinosa y 10 ng/mL de tetraciclina para inducir la expresión de las chaperonas *dnaK*, *dnaJ*, *grpE*, *groEL* y *groES*. Los cultivos se incubaron a 37 °C y 200 rpm hasta que alcanzaron una densidad óptica de 0,6 y se indujeron con 1 mM de IPTG.

Para los ensayos de sobreexpresión en medio de autoinducción no fue necesaria la adición de IPTG, sino que el cultivo de BL21 pET28:*all1651* se refrescó en 10 mL de medio AIM y se incubó a 200 rpm durante 24 horas a 25 °C o 48 horas a 15 °C para permitir la sobreexpresión de la proteína.

Para analizar la solubilidad de la proteína en cada una de las condiciones se tomaron 5 mL de cultivo en un tubo Falcon, se centrifugaron en una centrífuga Allegra X 30R a 4000 rpm durante 5 min, el precipitado se resuspendió en 1 mL del tampón que se quería analizar y se pasó a un tubo Eppendorf. A continuación, se añadieron 10 µL de inhibidor de proteasas PMSF 100 mM y se llevaron a cabo dos ciclos de congelación/descongelación con nitrógeno líquido para favorecer la rotura de las células. Las células se rompieron por sonicación llevando a cabo 10 ciclos de 45 s a amplitud 0,8 y ciclo 0,5 y 30 s de reposo en un sonicador UP 200S Dr. Hielscher, manteniendo las muestras en hielo para evitar el calentamiento de las muestras y evitar la desnaturalización de la proteína.

Después de la sonicación, las muestras se centrifugaron a 16000 x g y 4 °C durante 5 min y se separó el sobrenadante (fracción soluble) del precipitado (fracción insoluble), el cual se resuspendió en 1 mL de agua mili Q. En ambos casos se añadieron 200 µL de tampón Laemmli, se hirvió a 95 °C durante 5 min y se centrifugó 1 min a 16000 x g. Finalmente, se tomaron 10 µL de sobrenadante y se analizó la solubilidad de la proteína en cada condición mediante electroforesis en geles de SDS-PAGE como se describe en el apartado 4.5.2.

4.4. Purificación de proteínas recombinantes

La purificación del regulador All1651 se llevó a cabo mediante cromatografía de afinidad a iones metálicos (IMAC) a partir de biomasa de células de *E. coli* en las que se había sobreexpresado la proteína. Se emplearon dos aproximaciones, un protocolo estándar en el que se empleó como tampón de trabajo Tris 50 mM pH 8 y un protocolo basado en el uso de cloruro de guanidinio, agente caotrópico que a baja concentración permite purificar proteínas que formen cuerpos de inclusión o tengan poca solubilidad, sin llegar a desnaturalizarse (37).

4.4.1. Obtención de biomasa

Una vez optimizadas las condiciones de sobreexpresión, se inocularon colonias de BL21 o BL21 pG-KJE8 en Falcon con 10 mL de medio LB suplementados con 50 µg/mL de kanamicina (y 50 µg/mL de cloranfenicol en el caso de las células BL21 pG-KJE8). Los cultivos se incubaron a 37 °C y 200 rpm O/N y se utilizaron para iniciar cultivos en Erlenmeyer con 1 L de medio LB suplementado con 50 µg/mL de kanamicina (y 50 µg/mL cloranfenicol, 1 mg/mL de arabinosa y 10 ng/mL de tetraciclina en el caso de las células BL21 pG-KJE8). Los cultivos se incubaron a 37 °C y 200 rpm hasta que alcanzaron una densidad óptica de 0,6 y se indujeron con 1 mM de IPTG. Los cultivos inducidos se incubaron a 30 °C y 200 rpm durante 4 horas y se centrifugaron a 10000 x g y 4 °C durante 15 min. Se descartó el sobrenadante y se recogió el precipitado, que se congeló a -20 °C hasta su uso.

4.4.2. Lisis celular

Para poder purificar la proteína recombinante a partir de la biomasa obtenida es necesario romper las células y liberar su contenido. Para ello, se rompieron las células por sonicación, para lo cual se tomaron 10 g de biomasa y se resuspendieron en 50 mL de tampón de rotura (Tris-HCl 50 mM pH 8, 10 % (v/v) de glicerol, 500 mM de NaCl y 0,1 % (v/v) de Tween en el caso del protocolo estándar y 2 M de cloruro de guanidinio, 100 mM de NaH₂PO₄ y 10 mM de Tris-HCl pH 8 en el caso del protocolo con cloruro de guanidinio). Se añadieron 500 µL de PMSF 100 mM y se sonicó la muestra mediante 10 ciclos de 45 s a amplitud de 0,8 y ciclo 0,5 intercalados con 30 s de reposo, utilizando un baño de hielo para evitar el calentamiento de la muestra. Finalmente, la suspensión de células sonicadas se clarificó mediante 3 ciclos de centrifugación a 12 000 x g y 4 °C durante 15 min para obtener el extracto crudo, que se mantuvo a 4 °C.

4.4.3. Purificación por cromatografía de afinidad a iones metálicos

La cromatografía de afinidad a iones metálicos (IMAC) permite la purificación de proteínas que presentan una región rica en histidinas gracias a la afinidad de estos residuos por los metales bivalentes inmovilizados en la columna. En el caso del regulador All1651 esto fue posible porque el gen se clonó en el plásmido pET28 con una cola de histidinas en el extremo amino terminal.

En primer lugar, se preparó y equilibró la columna, para lo que se añadieron 10 mL de matriz *Chelating Sepharose Fast Flow* sobre una columna de cristal. La matriz primero se empaquetó por gravedad y después se empaquetó con 10 volúmenes de agua miliQ a un flujo de 3 mL/min, velocidad que se empleó a lo largo de todo el proceso de purificación. A continuación, se aplicaron 3 volúmenes de CoCl₂ 0,25 M para activar la columna con Co²⁺ y se lavó el exceso de metal con 5 volúmenes de agua mili Q. Finalmente, se equilibró la columna con 10 volúmenes de tampón de purificación. En el caso de la purificación en condiciones estándar se empleó como tampón de purificación Tris-HCl 50 mM pH 8, 500 mM de NaCl, 10 % (v/v) de glicerol y en la purificación con cloruro de guanidinio se empleó 2 M de este reactivo, 100 mM de NaH₂PO₄ y 10 mM de Tris-HCl pH 8.

Tras empaquetar y equilibrar la columna se llevó a cabo la purificación. Para ello se aplicó el extracto crudo, recogiendo el eluido de la columna (volumen obtenido durante el fijado del extracto crudo en la columna) para posteriormente verificar que la proteína de interés se había fijado a la columna. A continuación, se pasaron 15 volúmenes del tampón de lavado (Tris-HCl 50 mM pH 8, 200 mM de NaCl, 0,1 % Tween, 10 % glicerol, 40 mM de imidazol en el caso de la purificación estándar y 10 volúmenes de (NH₄)₂SO₄ 0,5 M en tampón de purificación seguidos de 5 volúmenes de glicina 35 mM en tampón de purificación en el caso de la purificación con cloruro de guanidinio). Para verificar que los lavados habían eliminado todas las proteínas unidas a la columna de forma inespecífica se midió la absorbancia del eluido a 280 nm y se comprobó que era inferior a 0,1.

Finalmente se llevó a cabo la elución de la proteína de interés, para lo que se aplicaron 50 mL de un gradiente creciente de imidazol de 40 mM a 1 M en tampón de purificación. El eluido se recogió en fracciones de 1 mL, que se analizaron mediante electroforesis desnaturante en geles de SDS-PAGE como se describe en el apartado 4.5.2. En el caso de la purificación con cloruro de guanidinio, previo a la electroforesis, se llevó a cabo una precipitación con TCA para eliminar el cloruro de guanidinio, como se describe en el apartado 4.5.1. Las fracciones que contenían la proteína de interés se combinaron y se sometieron a dos ciclos de diálisis en 3 L de tampón de diálisis. En la purificación estándar se utilizó Tris-HCl 50 mM pH 8, 100 mM de NaCl y 10% glicerol y en la purificación con cloruro de guanidinio se utilizó acético acetato 10 mM pH 4. Finalmente, las columnas se regeneraron con 5 volúmenes de agua mili Q, 5 volúmenes de Strip Buffer (500 mM NaCl, 50 mM EDTA, pH 8), 5 volúmenes de NaCl 500 mM, 3 volúmenes de agua mili Q y 5 volúmenes de etanol al 20 %.

4.5. Técnicas analíticas de proteínas

4.5.1. Precipitación con ácido tricloroacético (TCA)

El cloruro de guanidinio es un compuesto que interfiere con la electroforesis desnaturante en geles SDS-PAGE. Por tanto, para analizar las fracciones obtenidas en la purificación con cloruro de guanidinio es necesario llevar a cabo una precipitación con ácido tricloroacético (TCA). La precipitación con TCA se realizó mezclando 15 μL de muestra con 85 μL de agua mili Q y 100 μL de TCA al 10 % (v/v) e incubando en hielo durante 15 min. Pasado el tiempo las muestras se centrifugaron a 16000 x g durante 10 min y 4 °C. Una vez centrifugadas, se eliminó el sobrenadante de las muestras y se lavó el precipitado con 500 μL de acetona fría, tras lo cual se centrifugaron de nuevo en las mismas condiciones. Finalmente, se eliminó la acetona, se dejaron los Eppendorfs abiertos durante 10 min para que se evaporase cualquier resto de acetona que pudiera quedar, y se resuspendieron en 20 μL de agua mili Q. De esta manera se eliminó el cloruro de guanidinio y se analizaron las muestras mediante electroforesis desnaturante en geles de SDS-PAGE (apartado 4.5.2.).

4.5.2. Electroforesis desnaturante en geles de poliacrilamida

La electroforesis desnaturante en gel de poliacrilamida (SDS-PAGE) es una técnica que permite separar proteínas según su movilidad electroforética. En este trabajo, se utilizó para analizar el nivel de expresión de las proteínas en la prueba de sobreexpresión (apartado 4.3.2), su solubilidad al emplear diferentes tampones o condiciones de sobreexpresión (apartado 4.3.3.) y para analizar las fracciones de la purificación (apartado 4.4.)

Para llevar a cabo esta técnica se prepararon geles desnaturantes de SDS-PAGE que presentaban una región concentradora de 5 % de acrilamida/bisacrilamida y una región separadora de 17 % de acrilamida/bisacrilamida. La composición de los geles se recoge en el apartado 9.3. de Material Suplementario.

Las muestras se prepararon mezclando 15 μL de solución proteica con 3 μL de tampón Laemmli 6x (60 mM de Tris-HCl pH 6,8, 9,6 % (p/v) de SDS, 42 % (v/v) de glicerol, 0,06 % (p/v) de azul de bromofenol y 30 % (v/v) de β -mercaptoetanol). A continuación, las muestras se hirvieron a 95 °C durante 5 min, se centrifugaron a 16000 x g durante 2 min y se cargaron 10 μL de muestra en cada pocillo de los geles de SDS-PAGE. La electroforesis se llevó a cabo durante 45 min con amperaje constante de 35 mA por gel en un tampón compuesto por Tris-HCl 25 mM pH 8,8, glicina 200 mM y SDS 3,5 mM.

Los geles se tiñeron con una solución de azul de Coomassie (0,25 % (p/v) de Coomassie Blue R-250, 10 % (v/v) de ácido acético, 45 % (v/v) de metanol y 45 % de agua) durante 30 min en agitación. Finalmente se eliminó el exceso de colorante con una disolución decolorante (10 % (v/v) ácido acético, 25 % (v/v) metanol y 65 % de agua).

4.5.3. Cuantificación de las proteínas por espectroscopía

Tras la purificación, se obtuvo un espectro UV-Visible del regulador All1651 empleando un espectrofotómetro Specord 250 Plus (Analytik Jega AG), utilizando como blanco el tampón empleado en la diálisis. La proteína se cuantificó a partir de su absorbancia a 280 nm, aplicando la ley de Beer-Lambert ($Abs = \epsilon \cdot C \cdot l$), que permite relacionar la absorbancia obtenida (Abs) con la concentración proteica (c), en base al coeficiente de extinción molar de la proteína (ϵ). En este caso se utilizó el valor teórico $\epsilon_{280} = 44600 \text{ M}^{-1} \text{ cm}^{-1}$.

4.6. Ensayos de retardo en gel

El ensayo de retardo en gel o EMSA es una técnica que permite estudiar *in vitro* si un regulador transcripcional se une a la región promotora de un gen. En este trabajo esta técnica se usó para estudiar la unión del regulador All1651 a su propio promotor. Para llevar este ensayo se prepararon geles de poliacrilamida no desnaturalizantes al 4 %, los cuales presentan un grado de entrecruzamiento adecuado para permitir la migración de los complejos DNA-proteína. La composición de estos geles se encuentra en el apartado 9.4. de Material Suplementario.

Una vez preparados los geles se precorrieron a 60 V durante 30 min a 4 °C en un tampón compuesto por 50 mM de Tris-HCl pH 8 y 380 mM de glicina, lo cual equilibra el gel en el tampón de electroforesis y consigue que todo el sistema esté a la temperatura adecuada.

Las mezclas de reacción se prepararon mezclando distintas concentraciones de proteína con 25 ng de la región promotora del gen *all1651*, 25 ng de un fragmento de DNA inespecífico (*ifpkn22*) en tampón de unión (Bis Tris 10 mM pH 7,5, KCl 40 mM, glicerol 5 % (v/v) y BSA 50 µg/mL). Además, se analizaron diferentes condiciones como la presencia de DTT 1 mM, MnCl₂ 100 µM o EDTA 200 µM. Una vez preparadas, las mezclas se incubaron durante 30 min a temperatura ambiente para, a continuación, añadirles 3 µL de tampón de carga (Tris-HCl 50 mM pH 8, glicerol 30 % (v/v) y azul de bromofenol 0,05% (p/v)) y cargarlas en el gel. La electroforesis se llevó a cabo a 90 V durante 90 min y a 4 °C en el mismo tampón empleado para precorrer los geles. Pasado el tiempo, los geles se tiñeron con 1 µg/mL de SyBr[®] Safe Fluorescent Dye durante 10 mins en agitación y se revelaron con ayuda de un transiluminador.

4.7. Técnicas bioinformáticas

En este trabajo se emplearon técnicas bioinformáticas para obtener información sobre el regulador All1651. Se determinaron las propiedades biofísicas de la proteína a partir de la secuencia de aminoácidos mediante ProtParam de Expasy (<https://web.expasy.org/protparam/>) y se estudió la existencia de dominios funcionales mediante la herramienta Prosite de Expasy (<https://prosite.expasy.org/>). También se predijo su estructura secundaria mediante PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) y su estructura terciaria mediante Alpha-Fold (<https://alphafold.com/>) y Phyre2 (<http://www.sbg.bio.ic.ac.uk/~phyre2/>). Las estructuras terciarias se visualizaron mediante el programa PyMol (<https://pymol.org/>)

También se llevaron a cabo estudios filogenéticos. Para ello, se buscó proteínas homólogas en otras cianobacterias y en otros procariotas llevando a cabo un BLAST (Basic Local Alignment Search Tool), en el servidor BLASTP del NCBI (<https://blast.ncbi.nih.gov/Blast.cgi>) y se construyeron árboles filogenéticos en el portal Phylogeny.fr. (<http://www.phylogeny.fr>), para lo cual se realizó alineamientos múltiples utilizando la herramienta ClustalW (http://www.phylogeny.fr/one_task.cgi?task_type=clustalw), se curaron los alineamientos con GBlocks (http://www.phylogeny.fr/one_task.cgi?task_type=gblocks) y finalmente, se construyó el árbol con PhyML (http://www.phylogeny.fr/one_task.cgi?task_type=phyml) que se visualizó mediante TreeDyn (http://www.phylogeny.fr/one_task.cgi?task_type=treedyn). Los alineamientos múltiples se visualizaron con Jalview (<https://www.jalview.org/>).

5. Resultados

5.1. Análisis bioinformático del regulador All1651

En primer lugar se llevó a cabo un análisis de los parámetros biofísicos de la proteína a partir de su secuencia de aminoácidos empleando el servidor ProtParam. Así, se obtuvo que su peso molecular es de 32,0 kDa, su pI teórico es de 6,5 y su coeficiente de extinción molar (ϵ), asumiendo que todas las cisteínas están oxidadas, es de $44600 \text{ M}^{-1} \text{ cm}^{-1}$.

Para determinar la existencia en la secuencia de aminoácidos de rasgos compatibles con una proteína de unión al DNA y obtener información acerca de su posible función se estudió la presencia de dominios funcionales en la proteína mediante un análisis de su secuencia con el servidor Prosite (figura 1). Esto permitió localizar dos dominios: un dominio PAS (Period-Arnt-Sim) entre los aminoácidos 21 y 65; y un dominio de unión a DNA de tipo AraC, que abarca desde el aminoácido 171 hasta el 269, en el cual hay dos motivos HTH de unión al DNA (aminoácidos 188-209 y 236-259)

```
MFSNQNISSTNTYSQSSLSEPVSLIQHLINQVKEAIFCLDYQGHFYINDAACSLLGHSRQKLLNMTIDEVEMD
FLPSSWSYYWQLLKQQTSLTFKKVCSSQEYQSIIEVTIKSLEHGSEVLGCILAHFTSTKNPDISSDQGLYLDEKN
QPVHPSNKSLSSADFY PNCVQIRPIFEFIEQNYHLPI SLNDVAQAVGYSPAYLTNLVKKRRTQLTIIDWILERKM
LEARSLLINSDKSVTEIAMAVGFTDAYYFSRRFSQYHKLSPRSWRQKYQSLQTMN
```

Figura 1. Predicción de dominios en la secuencia primaria del regulador All1651. Aparece coloreado en amarillo el dominio PAS y en azul el dominio AraC, dentro del cual se resaltan subrayados y en negrita los motivos HTH de unión al DNA.

El dominio PAS (Period-Arnt-Sim) es un dominio localizado en el extremo N-terminal de algunas proteínas bacterianas y presenta una estructura tridimensional conservada. En lo relacionado a su función, el dominio PAS participa en múltiples sistemas de señalización celular, ya sea promoviendo la interacción proteína-proteína, la transferencia de señales o la detección directa de estímulos (38).

Los reguladores transcripcionales de la familia AraC se encargan de regular diversos sistemas genéticos mediante la presencia de un dominio de unión al DNA en el extremo C-terminal, el cual es funcional en forma de dímero. En cuanto a los procesos que controlan destacan aquellos que forman parte del catabolismo de azúcares, la respuesta a estreses y la virulencia (39).

Tras haber identificado los dominios de la proteína se llevó a cabo una predicción de estructura secundaria a partir de la secuencia de aminoácidos mediante el servidor PSIPRED, cuyo resultado se muestra en la figura 2.

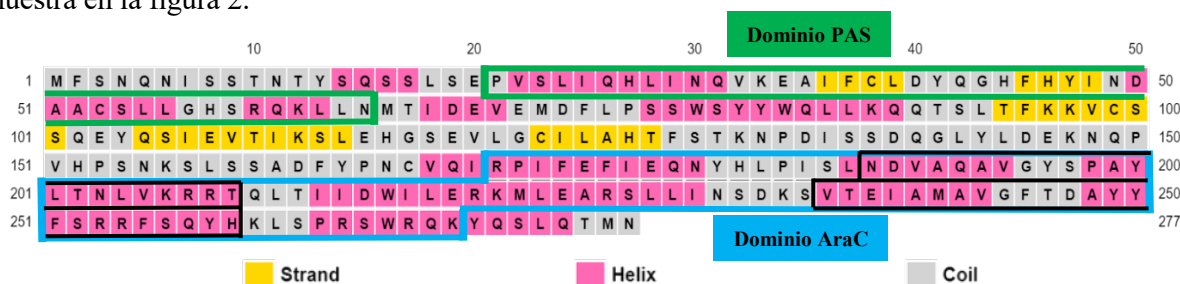


Figura 2. Predicción de la estructura secundaria del regulador All1651. Aparecen sombreados en color rosa los residuos que se predice que forman parte de una alfa hélice, en amarillo los que forman parte de una lámina beta y en gris los que forman bucles. Se recuadra en verde el dominio PAS y en azul el dominio AraC, con los motivos HTH recuadrados en negro.

La predicción de estructura secundaria (figura 2) indica que el dominio PAS está formado por tres hélices alfa y dos láminas beta, separadas por bucles, mientras que el dominio AraC se compone sólo de hélices alfa separadas por un bucle. Además, se puede observar que los motivos HTH se componen de una secuencia de hélice-giro-hélice (figura 2).

A continuación, se llevó a cabo una predicción de estructura terciaria utilizando el servidor AlphaFold. El resultado se muestra en la figura 3 A. Como se puede observar, tanto el dominio PAS como el dominio AraC contienen los motivos de estructura secundaria descritos anteriormente. En concreto el dominio AraC contiene siete hélices alfa y el dominio PAS tres hélices alfa y dos láminas beta. Además, como se observa en la figura 3 B, en esta estructura hay dos cisteínas (Cys38 y Cys123) que quedan a una distancia de 3,8 Å, compatible con la posibilidad de formación de un puente disulfuro.

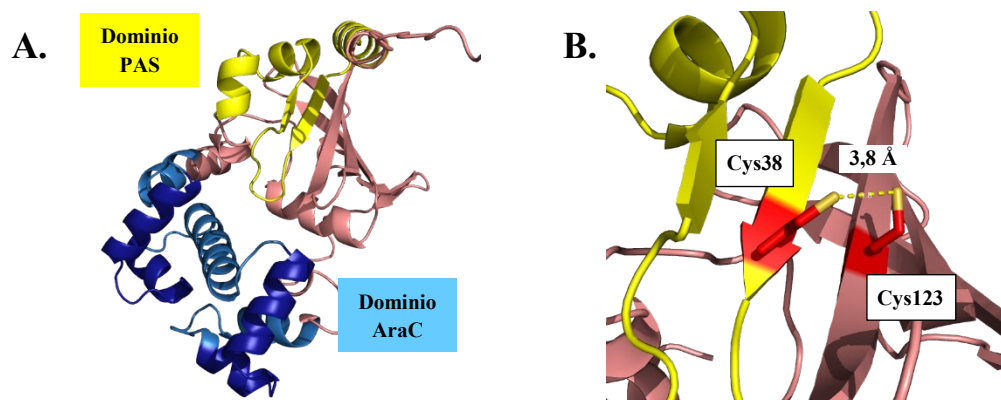


Figura 3. Predicción de la estructura terciaria del regulador All1651 empleando el servidor AlphaFold A. Estructura terciaria del regulador All1651, en la que se indica en color amarillo el dominio PAS y en azul el dominio AraC, con los motivos HTH de unión a DNA resaltados en azul oscuro. **B.** Detalle del dominio PAS, en el que hay dos cisteínas (Cys 38 y Cys 123) que quedan muy próximas y podrían formar un puente disulfuro.

Además de usar Alpha Fold, se llevó a cabo una predicción de la estructura terciaria del regulador All1651 con el servidor Phyre2, que permite llevar a cabo modelados por homología. Así, se obtuvo un modelo de All1651 en base a la estructura del regulador Rns de *E. coli* (PDB: 6XIV) (figura 4).

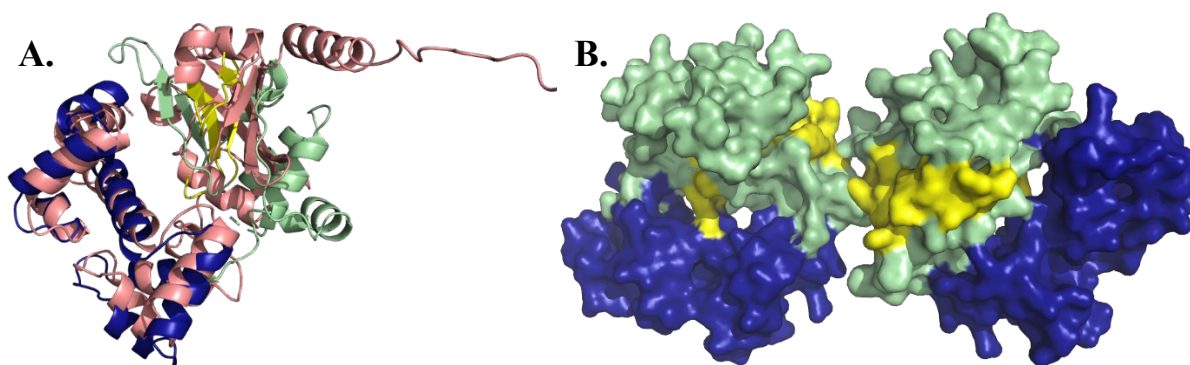


Figura 4. Predicción de la estructura terciaria del regulador All1651 empleando el servidor Phyre 2 A. Superposición del modelo construido por Phyre2 (verde), en el que se resalta en azul el dominio AraC y en amarillo el dominio PAS, sobre el modelo de Alpha Fold (rosa) **B.** Predicción de los dímeros del regulador All1651, obtenida al superponer el modelo de Phyre2 sobre la estructura de Rns de *E. coli* (PDB:6XIV)

Al superponer el modelo predicho por Alpha Fold con el modelo predicho por Phyre2 se observó que la predicción para el dominio AraC era muy similar, pero la predicción para la región amino terminal, donde se encuentra el dominio PAS, cambiaba notablemente (figura 4 A). Además, dado que el regulador Rns dimeriza y el regulador All1651 mostraba homología de secuencia con esta proteína, se llevó a cabo la superposición de la estructura predicha por Phyre2 para el regulador All1651 sobre cada una de las dos unidades del dímero de Rns para obtener un modelo del posible dímero de All1651 (figura 4 B). En la predicción se observa que los dominios PAS quedan expuestos y los dominios de unión a DNA quedan bastante separados entre sí.

Por otro lado, se llevó a cabo una búsqueda de homólogos de All1651 en cianobacterias y se construyó un alineamiento múltiple y un árbol filogenético (Secciones 9.5., 9.6. y 9.7. de Material Suplementario). Se observó que mientras el dominio AraC estaba altamente conservado en los homólogos de todas las cianobacterias el dominio PAS solo estaba presente en algunas de ellas, que eran precisamente aquellas que eran filogenéticamente más próximas en el árbol filogenético (figura 5 A) En concreto, las secuencias de *Calothrix*, *Cyanothece* y *Microcystis* contenían un dominio PAS de alta identidad con la secuencia de *Anabaena* y las secuencias de *Microcoleus*, *Trchocoleus* y *Gloeotrichia* contenían también un dominio PAS, pero con menor identidad (figura 5 B)

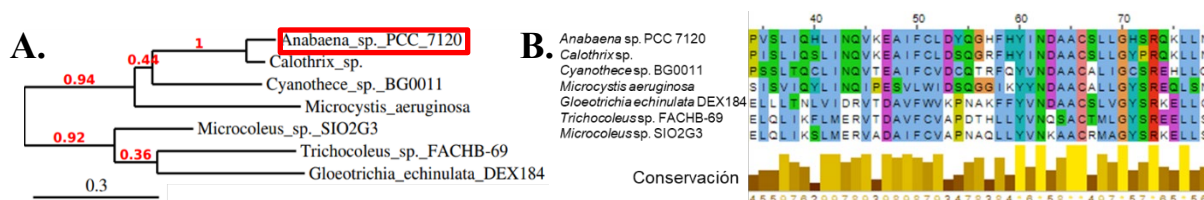


Figura 5. A. Árbol filogenético de homólogos del regulador All1651 en cianobacterias, en el que se muestran solo las secuencias filogenéticamente más próximas a la de *Anabaena sp. PCC 7120*. El árbol filogenético completo se muestra en la sección 8.6. de Material Suplementario B. Alineamiento múltiple del regulador All1651 en *Anabaena sp. PCC 7120* con reguladores de la familia AraC de otros procariontas, en el que se muestra la región correspondiente al dominio PAS de las secuencias filogenéticamente más próximas. El alineamiento completo se muestra en la sección 8.7. de Material Suplementario

Finalmente, se comparó la secuencia del regulador All1651 con la secuencia de reguladores de la familia AraC de otros procariontas, para lo que se construyó un alineamiento múltiple y un árbol filogenético (Secciones 9.8., 9.9. y 9.10. de Material Suplementario). Se obtuvo que los reguladores filogenéticamente más próximos eran HptR de *Staphylococcus aureus* y LacR de *Staphylococcus xylosus*. Además, se observó que mientras que el dominio AraC era similar al de otros procariontas, la secuencia del dominio PAS presentaba mucha divergencia con respecto a la región correspondiente en el resto de reguladores de la familia AraC (figura 6).

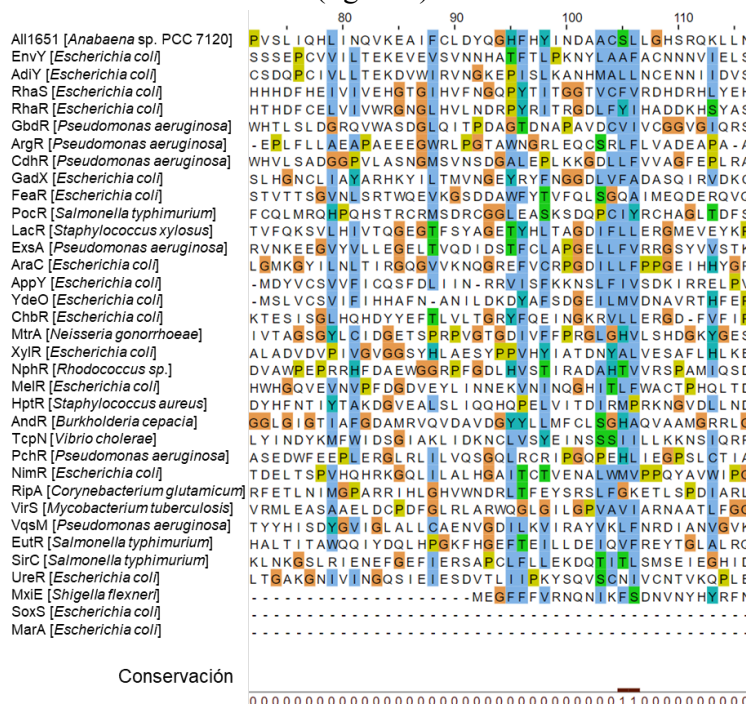


Figura 6. Detalle del alineamiento múltiple del regulador All1651 en *Anabaena sp. PCC 7120* con reguladores de la familia AraC de otros procariontas, en el que se muestra la región correspondiente al dominio PAS. El alineamiento completo se muestra en la sección 9.10. de Material Suplementario

5.2. Prueba de sobreexpresión del regulador All1651

Puesto que se pretendía purificar el regulador All1651, en primer lugar se determinaron las condiciones de máxima expresión de esta proteína en células de *E. coli* BL21 transformadas con el plásmido pET28:all1651. Para ello, se indujo la sobreexpresión de la proteína a una densidad óptica a 600 nm de 0,6 con 1 mM de IPTG y mediante electroforesis desnaturalizante en geles de SDS-PAGE se analizó el nivel de expresión a diferentes tiempos y temperaturas de inducción: 2 horas a 37 °C, 4 horas a 37 °C, 4 horas a 30 °C y O/N a 15 °C. Los resultados se recogen en la figura 7.

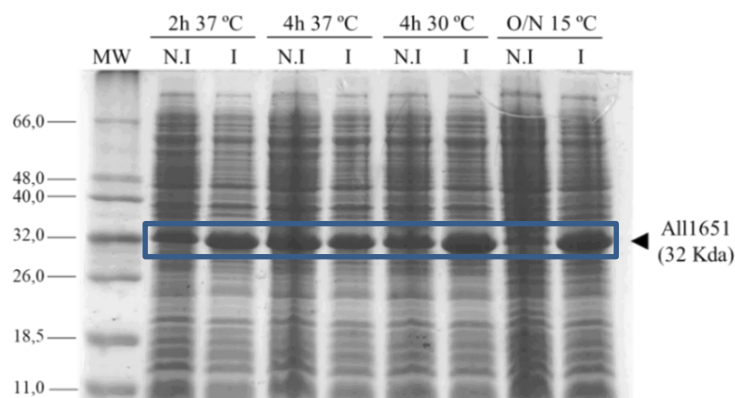


Figura 7. Prueba de sobreexpresión del regulador All1651. Se analizó el contenido proteico de extractos de cultivos de *E.coli* BL21 transformados con el plásmido pET-28:all1651 y sometidos a 4 condiciones de inducción: 2 h a 37 °C, 4 h a 37 °C, 4 h a 30 °C y O/N a 15 °C. Para cada condición de inducción, parte del cultivo se dejó sin inducir (N.I.) y parte se indujo con 1 mM de IPTG (I). La banda que aparece a una altura de 32 kDa corresponde al regulador All1651. El primer carril (MW) corresponde al marcador de peso molecular Low Molecular Weight Protein Marker II de NZYTech.

Como se observa en la figura 7, en todas las condiciones estudiadas se obtuvo un nivel considerable de expresión del regulador All1651. No obstante, el nivel de proteína fue ligeramente superior cuando las células inducidas se cultivaron a 30 °C durante 4 horas. Por tanto se eligieron estas condiciones de inducción para obtener la biomasa de células de *E.coli* conteniendo el regulador All1651.

5.3. Análisis de la solubilización del regulador All1651 empleando diferentes tampones

Una vez establecidas las condiciones de inducción que permitían obtener la máxima expresión del regulador All1651, se preparó biomasa de células de *E. coli* BL21 sobreexpresando esta proteína y se llevó a cabo una primera purificación mediante un protocolo estándar de cromatografía de afinidad a iones metálicos con Co^{2+} (IMAC- Co^{2+}) empleando como tampón de purificación Tris 50 mM pH 8. Sin embargo, tras analizar las muestras de los diferentes pasos de purificación mediante electroforesis desnaturalizante, se observó que apenas se obtenía proteína en las fracciones y que la totalidad de la proteína estaba en forma de cuerpos de inclusión en el precipitado obtenido al clarificar el extracto crudo (figura S-5 de Material Suplementario).

Esto sugería que el regulador All1651 no era capaz de solubilizarse en estas condiciones, por lo que se decidió estudiar su solubilidad añadiendo compuestos que favorecen la ruptura celular y la estabilidad proteica como NaCl, glicerol, el detergente tween o lisozima. En concreto, se analizó la solubilidad de la proteína empleando tampón de purificación suplementado con 10 % de glicerol y 500 mM de NaCl (Tampón 1); suplementado con glicerol 10 %, 500 mM de NaCl y 0,1 % de Tween (Tampón 2); y suplementado con glicerol 10 %, 500 mM de NaCl y 1 U de lisozima (Tampón 3). También se analizó la solubilidad en un tampón compuesto por 2 M de cloruro de guanidinio, 100 mM NaH_2PO_4 y 10 mM de Tris-HCl pH 8 (Tampón 4), pues en trabajos anteriores se había descrito que permitía la purificación de reguladores transcripcionales de forma exitosa (37).

Los resultados de la prueba de solubilización se recogen en la figura 8.

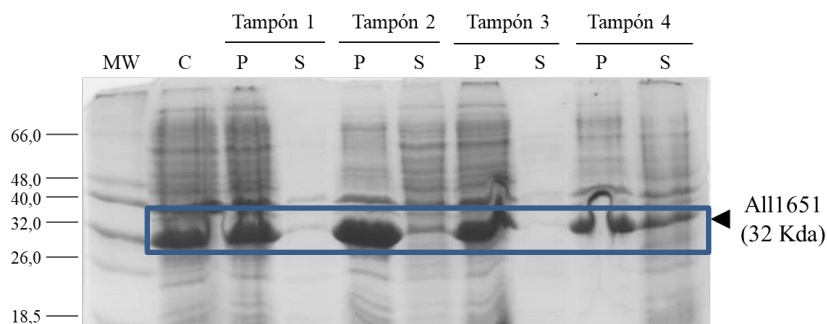


Figura 8. Prueba de solubilidad del regulador All1651. Se analizó la solubilidad del regulador All1651 en muestras de 5 mL de cultivo de *E.coli* BL21 con el plásmido pET-28:all1651 e inducidas con 1 mM de IPTG durante 4 h a 30 °C. Una de las muestras se trató con Laemli Buffer para liberar todo el contenido celular (C) y el resto se sonicaron en cuatro tampones diferentes. En cada tampón estudiado, se analizó el precipitado (P) y el sobrenadante (S) después de centrifugar el sonicado para determinar que fracción de la proteína de interés se había solubilizado. La banda que aparece a una altura de 32 kDa corresponde al regulador All1651. El primer carril (MW) corresponde al marcador de peso molecular Low Molecular Weight Protein Marker II.

En la figura 8 se observa que tanto en el tampón 1 (Tris 50 mM pH 8 con 10 % de glicerol y 500 mM de NaCl) como en el tampón 3 (Tris 50 mM pH 8 con 10 % de glicerol, 500 mM de NaCl y 1 U de lisozima) la totalidad del regulador All1651 se encuentra en el precipitado. En el caso del tampón 2 (Tris 50 mM pH 8 con glicerol 10 %, 500 mM de NaCl y 0,1 % de Tween) se observa que hay una pequeña fracción del regulador All1651 en el sobrenadante. La mayor solubilización del regulador All1651 se obtuvo al utilizar el tampón 4, que contenía cloruro de guanidinio, un agente caotrópico que a la concentración de trabajo (2 M) interacciona con las proteínas facilitando su purificación (37).

5.4. Purificación con cloruro de guanidinio

Dado que la mayor solubilización de la proteína se conseguía al emplear un tampón conteniendo cloruro de guanidinio y además existían indicios previos de que la purificación mediante cromatografía de afinidad a iones metálicos empleando un tampón conteniendo cloruro de guanidinio permitía purificar reguladores transcripcionales en una cantidad considerable y sin que perdieran su actividad (37), se decidió llevar a cabo la purificación del regulador All1651 empleando cromatografía de afinidad a iones metálicos en presencia de cloruro de guanidinio. Los resultados del análisis de los diferentes pasos de purificación mediante electroforesis desnaturante en geles de SDS se muestran en la figura 9.

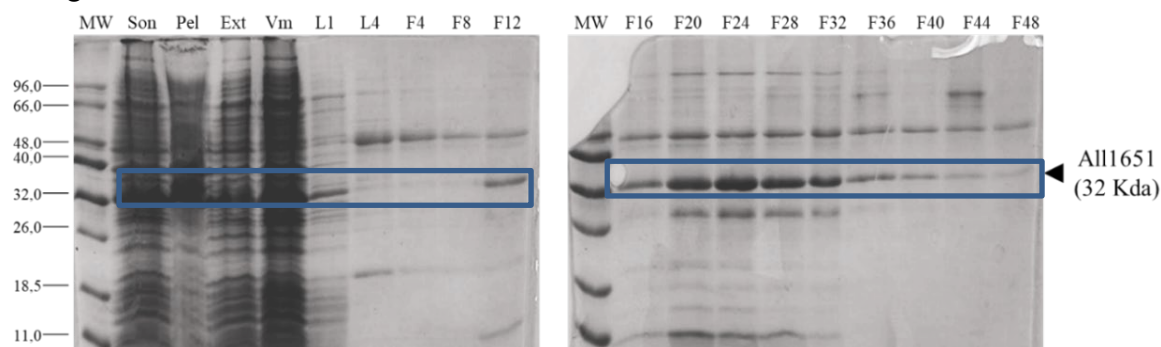


Figura 9. Purificación del regulador All1651 con cloruro de guanidinio. En el gel se observan las diferentes etapas del proceso de purificación: biomasa sonicada (Son), precipitado tras la clarificación (Pel), extracto crudo (Ext), volumen tras el fijado del extracto crudo (Vm), lavado 1 (L1), lavado 4 (L4) y fracciones eluidas del gradiente de imidazol (F4-F48). La banda que aparece a una altura de 32 kDa corresponde al regulador All1651. El primer carril (MW) corresponde al marcador de peso molecular Low Molecular Weight Protein Marker II de NZYTech.

Como se observa en la figura 9, la purificación en presencia de cloruro de guanidinio permitió obtener fracciones con una cantidad considerable de proteína y con un relativo grado de pureza. Por tanto, se seleccionaron las fracciones comprendidas entre la 18 y la 34 y se dializaron. Sin embargo, toda la proteína precipitó durante la diálisis, lo que mostraba que este método no era adecuado para poder purificar el regulador All651. Dado que en la prueba de solubilidad se observó que al emplear el resto de tampones no se obtenía una cantidad considerable de proteína soluble, se decidió intentar solubilizar el regulador All651 empleando otras estirpes de *E. coli* y otras condiciones de cultivo.

5.5. Análisis de la solubilización empleando diferentes estirpes y condiciones de cultivo

Se llevó a cabo la sobreexpresión del regulador All651 en diferentes estirpes de *E. coli* y con diferentes condiciones de cultivo. En concreto se sobreexpresó la proteína en células de *E. coli* BL21 y en *E. coli* pG-KJE8, que contiene un plásmido que codifica para las chaperonas *dnaK*, *dnaJ*, *grpE*, *groEL* y *groES*, las cuales favorecen la solubilización de las proteínas. En ambos casos, las células se cultivaron en medio LB hasta alcanzar una densidad óptica a 600 nm de 0,6, momento en el que se indujo su expresión O/N a 15 °C con 1 mM de IPTG. También se sobreexpresó la proteína en células de *E. coli* BL21 cultivadas en medio de autoinducción durante 24 horas a 25 °C y durante 48 horas a 15 °C. Este medio permite que la sobreexpresión de la proteína tenga lugar de forma más lenta y por tanto sea más probable que proteínas que presentan problemas de solubilización se plieguen de forma correcta. Como tampón para solubilizar las proteínas se eligió Tris 50 mM pH 8 con glicerol 10 %, 500 mM de NaCl y 0,1 % de Tween, pues este tampón, después del cloruro de guanidinio, era el que permitía obtener mayor solubilización de la proteína (Figura 8). Los resultados de esta prueba se recogen en la figura 10.

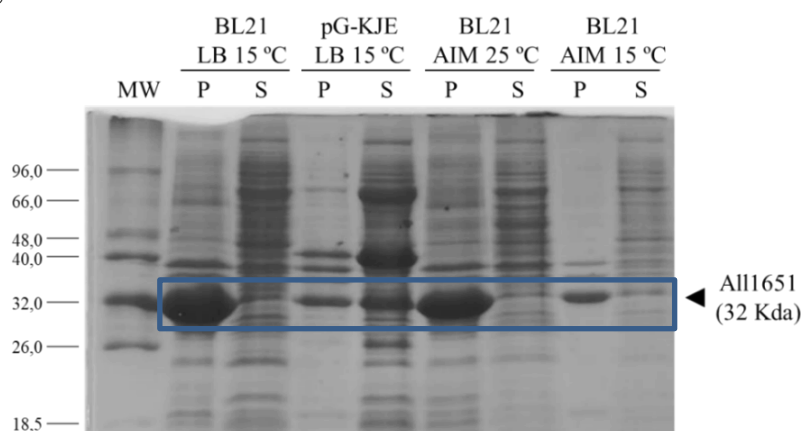


Figura 10. Prueba de solubilidad del regulador All651 en estirpes y condiciones de inducción. Se analizó la solubilidad del regulador All651 en muestras de 5 mL de cultivo de *E. coli* BL21 inducidas con 1 mM de IPTG O/N a 15 °C, *E. coli* BL21pG-KJE8 inducidas con 1 mM de IPTG O/N a 15 °C, *E. coli* BL21 cultivadas 24 h en medio de autoinducción a 25 °C y *E. coli* BL21 cultivadas 48 h en medio de autoinducción a 15 °C. Todas las muestra se sonicaron en Tris 50 mM pH 8 con glicerol 10 %, 500 mM de NaCl y 0,1 % de Tween y en todos los casos se analizó el precipitado (P) y el sobrenadante (S) después de centrifugar el sonificado para determinar que fracción de la proteína de interés se había solubilizado. La banda que aparece a una altura de 32 kDa corresponde al regulador All651. El primer carril (MW) corresponde al marcador de peso molecular Low Molecular Weigh Protein Marker II de NZYTech.

Como se observa en la figura 10, la sobreexpresión en células BL21 no permitió solubilizar la proteína, ni al cultivarlas en medio LB ni al cultivarlas en medio de autoinducción. Sin embargo, la estirpe BL21 pG-KJE8 sí que permitió una notable solubilización del regulador All651. Por tanto, se decidió utilizar esta cepa para generar biomasa y llevar a cabo la purificación de la proteína.

5.6. Purificación a partir de biomasa generada en la estirpe pG-KJE8

Tras determinar que la sobreexpresión del regulador All1651 en la estirpe BL21 pG-KJE8 permitía obtener una notable solubilización de la proteína se generó biomasa de esta estirpe sobreexpresando el regulador All1651 y se llevó a cabo una purificación de la proteína utilizando cromatografía de afinidad a metales inmovilizados con Co (IMAC-Co) y empleando como tampón base Tris-HCl 50 mM pH 8, 500 mM de NaCl y 10 % glicerol. Los resultados de la purificación se recogen en la figura 11.

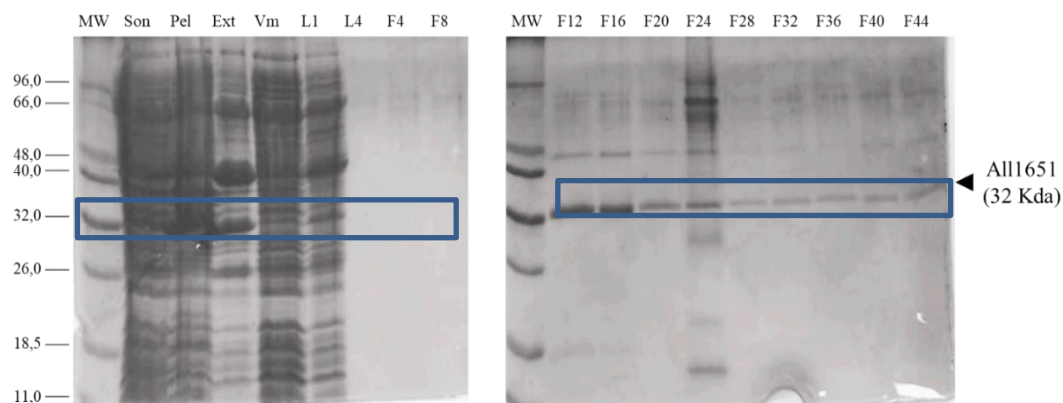


Figura 11. Purificación del regulador All1651 con Tris 50 mM pH 8, NaCl 500 mM y 10 % glicerol a partir de biomasa de células pG-KJE8. En el gel se observan las diferentes etapas del proceso de purificación: biomasa sonicada (Son), precipitado tras la clarificación (Pel), extracto crudo (Ext), volumen tras el fijado del extracto crudo (Vm), lavado 1 (L1), lavado 4 (L4) y fracciones eluidas del gradiente de imidazol (F4-F44) La banda que aparece a una altura de 32 kDa corresponde al regulador All1651. El primer carril (MW) corresponde al marcador de peso molecular Low Molecular Weigh Protein Marker II de NZYTech.

Como se observa en la figura 11, la purificación del regulador All1651 mediante IMAC-Co a partir de biomasa obtenida de células BL21 pG-KJE permitió obtener una cantidad aceptable de proteína purificada y con un cierto grado de pureza. Por ello, se seleccionaron las fracciones comprendidas entre la 10 y la 18, que se combinaron y se sometieron a dos ciclos de diálisis en tampón Tris-HCl 50 mM pH 8 con 100 mM de NaCl y 10% glicerol. Aunque una cierta cantidad de proteína precipitó durante la diálisis gran parte de ella permaneció soluble.

La proteína obtenida tras la diálisis se analizó mediante espectrofotometría, obteniendo un espectro UV-Visible de la misma entre 250 y 700 nm que se muestra en la figura 12.

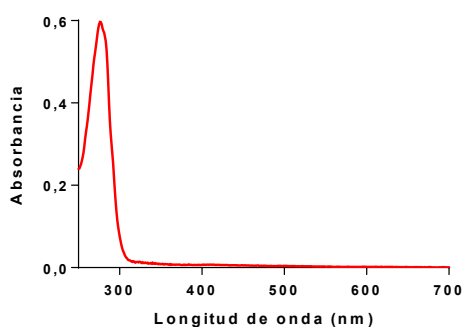


Figura 12. Espectro UV-visible de All1651 en tampón Tris-HCl 50 mM pH 8 con 100 mM de NaCl y 10 % de glicerol, obtenido tras la purificación a partir de células pG-KJE8

Finalmente, la proteína se cuantificó a partir de la absorbancia a 280 nm mediante la ley de Beer-Lambert, utilizando el coeficiente de extinción teórico ($\epsilon_{280\text{ nm}} = 44600 \text{ M}^{-1} \text{ cm}^{-1}$). Así, se determinó que se había obtenido un total de 8 mL de solución de proteína a una concentración de 12 μM .

5.7. Estudio de unión al DNA mediante ensayos de retardo en gel

Tras haber purificado el regulador All1651, el último paso fue comprobar que la proteína obtenida presentaba actividad de unión al DNA y establecer las condiciones óptimas de unión. Para ello, dado que la mayoría de reguladores transcripcionales están autorregulados, se llevaron a cabo ensayos de retardo en gel (EMSA) con su propia región promotora. Se estudió la unión del regulador All1651 al DNA en diferentes condiciones. Dado que la proteína contenía 5 cisteínas, y al realizar el modelado de la proteína se había observado que 2 de ellas podían formar un puente disulfuro, se estudió si la unión al DNA se veía afectada por la presencia de DTT, un agente reductor de puentes disulfuro. Por otra parte, dado que muchos reguladores transcripcionales son metalorreguladores, se quiso determinar si la unión del regulador All1651 variaba en presencia de Mn^{2+} , un metal que típicamente modula la unión de reguladores transcripcionales al DNA *in vitro*, y de EDTA, un quelante de metales. Los resultados de los ensayos de retardo en gel se recogen en la figura 13.

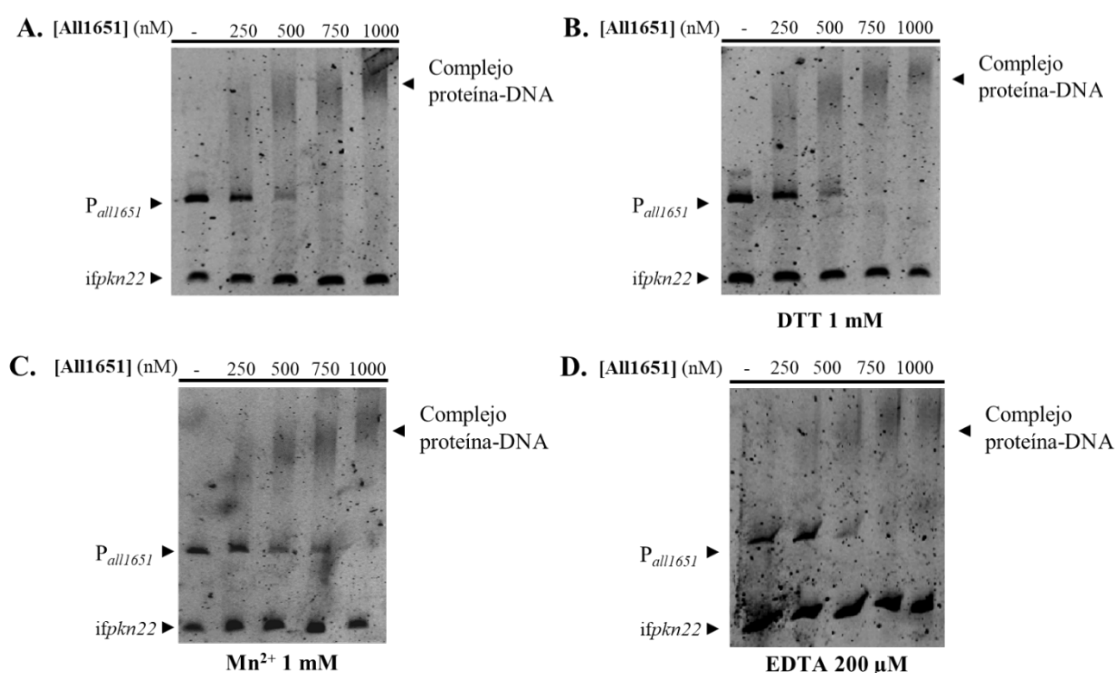


Figura 13. Ensayos de retardo en gel del regulador All1651 con su propia región promotora en geles de poliacrilamida al 4%. El carril 1 corresponde al promotor del gen all1651 en ausencia de proteína y los carriles 2, 3, 4 y 5 corresponden a la región promotora incubada con 250, 500, 750 y 1000 nM de proteína. Los ensayos se realizaron en presencia de un fragmento de DNA inespecífico (ifpkn22) que sirvió como control y en ausencia de agentes reductores y metal (A), con 1 mM de DTT (B), 1 mM de Mn^{2+} (C) o 200 μM de EDTA (D).

Como se observa en la figura 13 A., el regulador All1651 fue capaz de unirse a su propia región promotora en ausencia de agentes reductores y sin necesidad de añadir ningún metal, formándose un complejo DNA-proteína cuya intensidad se incrementaba al aumentar la concentración de regulador transcripcional. Además, esta unión mostró ser específica, pues no se observa disminución en la intensidad de la banda del DNA inespecífico (ifpkn22). De hecho, la adición de 1mM de DTT a la mezcla de reacción no provocó cambios significativos en la actividad de unión al DNA del regulador (figura 13 B). Del mismo modo, la adición de 1 mM de Mn^{2+} (Figura 13 C) o de 200 μM de EDTA (Figura 13 D) tampoco produjo un aumento o disminución significativa de los complejos proteína-DNA. Así pues, la proteína obtenida mediante el protocolo de purificación puesto a punto en este trabajo era una proteína activa con capacidad de unirse al DNA.

6. Discusión

Las proteínas FUR modulan la expresión de otros reguladores transcripcionales, haciendo que su actividad reguladora vaya más allá de la regulación directa de los genes diana. Un ejemplo de estos reguladores es el estudiado en este trabajo, All1651, que resultaba de interés porque su expresión estaba regulada por dos de las proteínas FUR de *Anabaena* sp. PCC 7120, FurA y FurC (Jorge Guío, datos no publicados). En las bases de datos, el gen *all1651* aparecía anotado como un regulador transcripcional, aunque no había ningún estudio que lo confirmase. En este trabajo se han analizado sus características, se ha optimizado un protocolo de purificación y se han descrito sus condiciones de unión a DNA *in vitro*.

6.1. El regulador All1651 es un regulador de la familia AraC que contiene un dominio PAS

El análisis bioinformático del regulador All1651 mostró que presentaba un dominio de unión a DNA propio de los reguladores transcripcionales de la familia AraC localizado en el extremo carboxilo terminal (figura 1) (40). Estos dominios se caracterizan por estar compuestos por siete hélices alfa y contener dos motivos HTH (hélice-giro-hélice) de unión a DNA, el primero compuesto por las hélices 2 y 3 y el segundo compuesto por las hélices 5 y 6 (40), distribución de motivos que se observa tanto en la predicción de estructura secundaria como en los modelos de estructura terciaria (figuras 2 y 3).

Además de un dominio carboxilo terminal conservado responsable de la unión a DNA, los reguladores de la familia AraC, presentan un dominio amino terminal variable responsable de la unión a ligandos (41). En este caso, el regulador All1651 presenta en su región amino terminal un dominio PAS. Como se ha comentado previamente, se ha descrito que estos dominios están implicados en el reconocimiento de ligandos (38). Al igual que en el dominio de unión a DNA, la predicción de estructura secundaria y los modelos de estructura terciaria (figuras 2 y 3) mostraron que la disposición de motivos de estructura secundaria del dominio PAS del regulador All1651 coincide con la estructura canónica del dominio PAS, compuesta por una hoja beta antiparalela y entre dos y tres hélices alfa (38). Así, se concluye que el regulador All1651 es un regulador de la familia AraC con un dominio sensor de tipo PAS.

El alineamiento del regulador All1651 con otros reguladores de la familia AraC en procariotas (figura 6) reveló que el dominio PAS no presentaba similitud con las regiones amino terminal de ninguno de los reguladores de la familia AraC analizados. También se llevó a cabo un alineamiento del dominio PAS con la secuencia de dominios PAS descritos en otras proteínas (38) (datos no mostrados) que puso de manifiesto que el dominio PAS del regulador All1651 no era similar a ningún dominio PAS descrito hasta la fecha. Además, en el dominio PAS del regulador All1651 también se observó la existencia de dos cisteínas que debido a su proximidad podrían establecer un puente disulfuro entre ellas (figura 3 B). Esto puede ser un indicio de que este dominio podría responder al estado redox de la célula, ya que según las condiciones redox el puente disulfuro puede estar formado o no, lo que podría afectar a su capacidad de reconocer ligandos a través del dominio PAS o a su actividad de unión al DNA.

Por otra parte, al estudiar la secuencia de reguladores transcripcionales homólogos a All1651 en otras cianobacterias, solo se observa la presencia del dominio PAS en aquellas secuencias filogenéticamente más próximas a *Anabaena* (figuras 5 A y 5 B), estando especialmente conservado en las secuencias de *Calothrix* y *Cyanothece*, donde tiene un 87 % y 64 % de identidad. Estos resultados sugieren que el dominio PAS del regulador All1651 es exclusivo de algunas cianobacterias y en concreto parece ser muy similar al de la secuencia de *Calothrix*, fijadora de nitrógeno en heterocistos, y en *Cyanothece*, también fijadora de nitrógeno aunque no formadora de heterocistos. Por tanto este dominio podría estar implicado en el reconocimiento de señales específicas de las cianobacterias y, dado que parece estar más conservado en cianobacterias fijadoras de nitrógeno, es posible que fueran moléculas señal de la deficiencia de nitrógeno, como el 2-oxoglutarato (9).

Finalmente, atendiendo al modelo de la estructura del dímero del regulador All1651 (figura 4 B) se observa que los dominios PAS están en la zona central del dímero y expuestos hacia el exterior, lo que les permitiría reconocer la señal, mientras que los dominios de unión a DNA quedan muy distantes. Esta disposición de los dominios de unión a DNA coincide con estudios previos de reguladores de la familia AraC (41), en los que se ha descrito que en ausencia del ligando, el regulador se une a regiones distantes del DNA, lo que hace que el DNA se doble y no se pueda unir la RNA polimerasa. Por el contrario, en presencia del ligando, los dominios AraC quedan más próximos entre sí y se unen a regiones del DNA más próximas, activando la expresión de sus genes diana (41).

6.2. La sobreexpresión en la estirpe pG-KJE8 permite la purificación del regulador All1651

Para obtener el regulador All1651 se optó por su sobreexpresión heteróloga en *E. coli* y posterior purificación mediante cromatografía de afinidad a iones metálicos inmovilizados (IMAC), ya que el gen *all1651* se había clonado en el plásmido pET28 con una cola de histidinas en el extremo amino terminal. La sobreexpresión de All1651 en células de *E. coli* BL21 resultó satisfactoria (figura 7), por lo que se decidió llevar a cabo una purificación mediante IMAC en Tris 50 mM pH 8. Sin embargo, se observó que toda la proteína se encontraba en el precipitado obtenido tras sonicar la biomasa, lo que era indicativo de que la proteína, posiblemente, estuviera formando cuerpos de inclusión. Al estudiar la solubilidad de la proteína en diferentes condiciones (figura 8) se observó que al añadir 2 M de cloruro de guanidinio al tampón de purificación se lograba solubilizar parcialmente el regulador All1651, por lo que se decidió llevar a cabo una purificación empleando un tampón de purificación con cloruro de guanidinio, pues trabajos anteriores habían descrito que el uso de este agente caotrópico permitía purificar reguladores transcripcionales propensos a la agregación sin perder su actividad de unión a DNA (37).

Aunque la purificación con cloruro de guanidinio permitió recuperar una cantidad considerable de esta proteína (figura 9), al llevar a cabo la diálisis en tampón acético/acetato pH 4, que en el caso de otros reguladores que mostraban problemas de precipitación a pH 8 permitía mantener las proteínas en solución (37), el regulador All1651 precipitaba en su totalidad, por lo que no resultó un protocolo satisfactorio. En las pruebas de solubilización se había observado que la adición de agentes que favorecen la ruptura celular o que impiden la agregación no lograban solubilizar la proteína (figura 8), lo que sugería que el regulador All1651 podía estar formando cuerpos de inclusión. Por ello, se decidió modificar las condiciones de sobreexpresión.

Por una parte, se ensayó su sobreexpresión empleando un AIM que permite una sobreexpresión más lenta, lo que favorece el correcto plegamiento de proteínas propensas a formar cuerpos de inclusión y con ello su solubilización (36). Por otra parte, se llevó a cabo la sobreexpresión en la estirpe pG-KJE8, en la que se co-expresan chaperonas que favorecen el correcto plegamiento de las proteínas (42). Mientras que el medio de autoinducción no mejoró la solubilización de la proteína, la co-expresión con chaperonas sí lo hizo (figura 10), por lo que, a partir de la biomasa obtenida en la estirpe pG-KJE8 se llevó a cabo la purificación de este regulador mediante IMAC. Aunque la cantidad obtenida y el grado de pureza no fueron muy altos, eran aceptables para poder estudiar sus condiciones de unión al DNA mediante ensayos de retardo en gel. No obstante, son necesarios futuros estudios para optimizar la purificación de esta proteína. Una opción podría ser el método establecido por Klinker *et al.* (43) para la purificación de histonas, que también son proteínas de unión a DNA que forman cuerpos de inclusión y en las que la desnaturalización, purificación por intercambio iónico y posterior renaturalización permite purificar estas proteínas muy puras y en alta cantidad.

6.3. El regulador All1651 muestra actividad de unión al DNA

Una de las características intrínsecas de un regulador transcripcional es su capacidad de unirse de forma específica a sus promotores diana para inducir o reprimir su expresión. Por lo tanto, una vez se pudo purificar parcialmente el regulador All1651 se procedió a estudiar su capacidad de unión al DNA mediante el ensayo de retardo en gel o EMSA, para lo cual se utilizó la región promotora de su propio gen, ya que la mayoría de reguladores transcripcionales de la familia AraC están autorregulados (39).

Para estudiar su unión al DNA se intentaron imitar las condiciones fisiológicas a las que actúa el regulador, para lo cual los EMSA se realizaron con las condiciones en las que otros reguladores de *Anabaena* sp. PCC 7120 habían mostrado actividad de unión al DNA. Se ensayó si su unión se veía afectada por la presencia o ausencia de DTT, de metal (Mn^{2+}) o de quelante de metales y se observó la formación de complejos proteína-DNA en todas las condiciones analizadas (figura 13). Esto indicaba que el protocolo de purificación era satisfactorio al mantener la actividad de unión a DNA de la proteína de estudio, y que All1651 era capaz de unirse a su propio promotor, modular por tanto su propia expresión. Además, los resultados sugieren que All1651 no era un metalorregulador y que las condiciones redox no afectan a su actividad de unión al DNA. Futuros estudios irán encaminados a identificar los ligandos reconocidos por el dominio PAS y estudiar su posible influencia sobre la actividad de unión al DNA de All1651, así como identificar genes de *Anabaena* sp. PCC 7120 que estén modulados por este regulador.

7. Conclusiones

Tras la realización de este Trabajo Fin de Grado se han obtenido las siguientes conclusiones:

- Los estudios bioinformáticos muestran que All1651 es un regulador transcripcional de la familia AraC con un dominio PAS en su extremo amino terminal, el cual parece ser exclusivo de cianobacterias.
- La sobreexpresión heteróloga de All1651 en células pG-KJE8 y su purificación mediante IMAC- Co^{2+} permite la obtención de esta proteína con una pureza y en una cantidad aceptables.
- Los ensayos de retardo en gel indican que All1651 tiene actividad de unión al DNA y es capaz de unirse a su propia región promotora *in vitro* de forma específica
- La actividad de unión al DNA de All1651 no se ve afectada por la presencia de metal ni por las condiciones redox, sugiriendo que no se trata de un metalorregulador ni de un regulador redox.

Conclusions

After finishing this Final Degree Project, the following conclusions have been obtained:

- Bioinformatics studies show that All1651 is a transcriptional regulator of the AraC family with a PAS domain at its amino terminus, which seems to be exclusive to cyanobacteria.
- The heterologous overexpression of All1651 in pG-KJE8 cells and its purification by IMAC- Co^{2+} allows obtaining this protein with acceptable purity and quantity.
- Electrophoretic mobility shift assay revealed that All1651 has DNA binding activity, and it is capable of binding to its own promoter region *in vitro* in a specific manner.
- The DNA binding activity of All1651 is not affected by the presence of metal or by redox conditions, suggesting that it is neither a metalloregulator nor a redox regulator.

8. Bibliografía

1. Stanier RY, Cohen-Bazire G. Phototrophic prokaryotes: the cyanobacteria. *Annual Review of Microbiology*. 1977;31(1):225-74.
2. Whitton BA, Potts M. Introduction to the Cyanobacteria. In: Whitton BA, editor. *Ecology of Cyanobacteria II: Their Diversity in Space and Time*. Dordrecht: Springer Netherlands; 2012. p. 1-13.
3. Green BR. After the primary endosymbiosis: an update on the chromalveolate hypothesis and the origins of algae with Chl c. *Photosynthesis Research*. 2011;107(1):103-15.
4. Gaysina LA, Saraf A, Singh P. Chapter 1 - Cyanobacteria in Diverse Habitats. In: Mishra AK, Tiwari DN, Rai AN, editors. *Cyanobacteria*: Academic Press; 2019. p. 1-28.
5. Falkowski P. Ocean Science: The power of plankton. *Nature*. 2012;483(7387):S17-S20.
6. Havens KE. Cyanobacteria blooms: effects on aquatic ecosystems. In: Hudnell HK, editor. *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs*. New York, NY: Springer New York; 2008. p. 733-47.
7. Gupta V, Ratha SK, Sood A, Chaudhary V, Prasanna R. New insights into the biodiversity and applications of cyanobacteria (blue-green algae)—Prospects and challenges. *Algal Research*. 2013;2(2):79-97.
8. Shcolnick S, Keren N. Metal Homeostasis in Cyanobacteria and Chloroplasts. Balancing Benefits and Risks to the Photosynthetic Apparatus. *Plant Physiology*. 2006;141(3):805-10.
9. Flores E, Herrero A. Assimilatory Nitrogen Metabolism and Its Regulation. In: Bryant DA, editor. *The Molecular Biology of Cyanobacteria*. Dordrecht: Springer Netherlands; 1994. p. 487-517.
10. Fillat MF. The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. *Archives of Biochemistry and Biophysics*. 2014;546:41-52.
11. Hernández JA, López-Gomollón S, Bes MT, Fillat MF, Peleato ML. Three fur homologues from *Anabaena* sp. PCC7120: exploring reciprocal protein-promoter recognition. *FEMS Microbiology Letters*. 2004;236(2):275-82.
12. González A, Bes MT, Valladares A, Peleato ML, Fillat MF. FurA is the master regulator of iron homeostasis and modulates the expression of tetrapyrrole biosynthesis genes in *Anabaena* sp. PCC 7120. *Environmental Microbiology*. 2012;14(12):3175-87.
13. González A, Bes MT, Peleato ML, Fillat MF. Expanding the Role of FurA as Essential Global Regulator in Cyanobacteria. *PLOS ONE*. 2016;11(3):e0151384.
14. Hernández JA, López-Gomollón S, Muro-Pastor A, Valladares A, Bes MT, Peleato ML, et al. Interaction of FurA from *Anabaena* sp. PCC 7120 with DNA: A Reducing Environment and the Presence of Mn²⁺ are Positive Effectors in the Binding to *isiB* and *furA* Promoters. *Biomaterials*. 2006;19(3):259-68.
15. Botello-Morte L, Bes MT, Heras B, Fernández-Otal Á, Peleato ML, Fillat MF. Unraveling the Redox Properties of the Global Regulator FurA from *Anabaena* sp. PCC 7120: Disulfide Reductase Activity Based on Its CXXC Motifs. *Antioxidants & Redox Signaling*. 2013;20(9):1396-406.
16. Botello-Morte L, Pellicer S, Sein-Echaluce VC, Contreras LM, Neira JL, Abián O, et al. Cysteine Mutational Studies Provide Insight into a Thiol-Based Redox Switch Mechanism of Metal and DNA Binding in FurA from *Anabaena* sp. PCC 7120. *Antioxidants & Redox Signaling*. 2015;24(4):173-85.
17. Guío J, Bes MT, Balsera M, Calvo-Begueria L, Sevilla E, Peleato ML, et al. Thioredoxin Dependent Changes in the Redox States of FurA from *Anabaena* sp. PCC 7120. *Antioxidants*. 2021;10(6).
18. Hernández JA, Peleato ML, Fillat MaF, Bes MT. Heme binds to and inhibits the DNA-binding activity of the global regulator FurA from *Anabaena* sp. PCC 7120. *FEBS Letters*. 2004;577(1):35-41.
19. Guío J, Sarasa-Buisan C, Velázquez-Campoy A, Bes MT, Fillat MF, Peleato ML, et al. 2-oxoglutarate modulates the affinity of FurA for the *ntcA* promoter in *Anabaena* sp. PCC 7120. *FEBS Letters*. 2020;594(2):278-89.
20. Napolitano M, Rubio MÁ, Santamaría-Gómez J, Olmedo-Verd E, Robinson NJ, Luque I. Characterization of the Response to Zinc Deficiency in the Cyanobacterium *Anabaena* sp. Strain PCC 7120. *Journal of Bacteriology*. 2012;194(10):2426-36.
21. Sein-Echaluce VC, González A, Napolitano M, Luque I, Barja F, Peleato ML, et al. Zur (FurB) is a key factor in the control of the oxidative stress response in *Anabaena* sp. PCC 7120. *Environmental Microbiology*. 2015;17(6):2006-17.

22. López-Gomollón S, Sevilla E, Bes MT, Peleato ML, Fillat María F. New insights into the role of Fur proteins: FurB (All2473) from *Anabaena* protects DNA and increases cell survival under oxidative stress. *Biochemical Journal*. 2009;418(1):201-7.
23. Oliván-Muro I, Sarasa-Buisan C, Guío J, Arenas-Busto J, Sevilla E, Fillat MF. Unbalancing Zur (FurB)-mediated homeostasis in *Anabaena* sp. PCC7120: Consequences on metal trafficking, heterocyst development and biofilm formation. *Environmental Microbiology*. 2023.
24. Yingping F, Lemeille S, Talla E, Janicki A, Denis Y, Zhang C-C, et al. Unravelling the cross-talk between iron starvation and oxidative stress responses highlights the key role of PerR (alr0957) in peroxide signalling in the cyanobacterium *Nostoc* PCC 7120. *Environmental Microbiology Reports*. 2014;6(5):468-75.
25. Sevilla E, Sarasa-Buisan C, González A, Cases R, Kufryk G, Peleato ML, et al. Regulation by FurC in *Anabaena* Links the Oxidative Stress Response to Photosynthetic Metabolism. *Plant and Cell Physiology*. 2019;60(8):1778-89.
26. Sarasa-Buisan C, Guío J, Broset E, Peleato ML, Fillat MF, Sevilla E. FurC (PerR) from *Anabaena* sp. PCC7120: a versatile transcriptional regulator engaged in the regulatory network of heterocyst development and nitrogen fixation. *Environmental Microbiology*. 2022;24(2):566-82.
27. Rachedi R, Foglino M, Latifi A. Stress Signaling in Cyanobacteria: A Mechanistic Overview. *Life*. 2020;10(12).
28. Pis Diez CM, Juncos MJ, Villarruel Dujovne M, Capdevila DA. Bacterial Transcriptional Regulators: A Road Map for Functional, Structural, and Biophysical Characterization. *International Journal of Molecular Sciences*. 2022;23(4).
29. Stock AM, Robinson VL, Goudreau PN. Two-Component Signal Transduction. *Annual Review of Biochemistry*. 2000;69(1):183-215.
30. Srivastava A, Summers ML, Sobotka R. Cyanobacterial sigma factors: Current and future applications for biotechnological advances. *Biotechnology Advances*. 2020;40:107517.
31. Yoshimura H, Okamoto S, Tsumuraya Y, Ohmori M. Group 3 sigma factor gene, *sigJ*, a key regulator of desiccation tolerance, regulates the synthesis of extracellular polysaccharide in cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Research*. 2007;14(1):13-24.
32. Kopf M, Hess WR. Regulatory RNAs in photosynthetic cyanobacteria. *FEMS Microbiology Reviews*. 2015;39(3):301-15.
33. Jeong H, Barbe V, Lee CH, Vallenet D, Yu DS, Choi S-H, et al. Genome Sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *Journal of Molecular Biology*. 2009;394(4):644-52.
34. Phue J-N, Lee SJ, Trinh L, Shiloach J. Modified *Escherichia coli* B (BL21), a superior producer of plasmid DNA compared with *Escherichia coli* K (DH5 α). *Biotechnology and Bioengineering*. 2008;101(4):831-6.
35. Zhang H, Li Z, Pereira B, Stephanopoulos G. Engineering *E. coli*-*E. coli* cocultures for production of muconic acid from glycerol. *Microbial Cell Factories*. 2015;14(1):134.
36. Studier FW. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*. 2005;41(1):207-34.
37. Pellicer S, Bes MT, González A, Neira JL, Peleato ML, Fillat MF. High-recovery one-step purification of the DNA-binding protein Fur by mild guanidinium chloride treatment. *Process Biochemistry*. 2010;45(2):292-6.
38. Henry JT, Crosson S. Ligand-Binding PAS Domains in a Genomic, Cellular, and Structural Context. *Annual Review of Microbiology*. 2011;65(1):261-86.
39. Martin RG, Rosner JL. The AraC transcriptional activators. *Current Opinion in Microbiology*. 2001;4(2):132-7.
40. Rodgers ME, Schleif R. Solution structure of the DNA binding domain of AraC protein. *Proteins: Structure, Function, and Bioinformatics*. 2009;77(1):202-8.
41. Bustos SA, Schleif RF. Functional domains of the AraC protein. *Proceedings of the National Academy of Sciences*. 1993;90(12):5638-42.
42. Nishihara K, Kanemori M, Yanagi H, Yura T. Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in *Escherichia coli*. *Applied and Environmental Microbiology*. 2000;66(3):884-9.
43. Klinker H, Haas C, Harrer N, Becker PB, Mueller-Planitz F. Rapid Purification of Recombinant Histones. *PLOS ONE*. 2014;9(8):e104029.

**MATERIAL
SUPPLEMENTARIO**

9. Material Suplementario

9.1. Programa utilizado en las reacciones de PCR

1. Desnaturalización inicial del DNA: 95 °C, 5 min
2. Desnaturalización: 95 °C, 30 s
3. Hibridación: 50 °C, 30 s
4. Extensión: 72 °C, 45 s
5. Elongación final: 72 °C, 10 min
6. Conservación del material amplificado: 4 °C

Repitiendo los pasos 2, 3 y 4 durante 30 ciclos.

9.2. Oligonucleótidos utilizados para la amplificación de promotores mediante PCR

Los oligonucleótidos utilizados en este trabajo fueron los siguientes:

| Cebador | Secuencia 5' → 3' | Secuencia amplificada |
|--------------------|-------------------------|---|
| <i>Pall1651</i> up | GAAACCCGTCACACGCAGTG | Promotor del gen <i>all1651</i> de <i>Anabaena</i> sp. PCC 7120 |
| <i>Pall1651</i> dw | GCTGCATATAGTGTTAATTTAGG | |
| <i>ifpkn22</i> up | AAAGATGAATTACTGGCG | Fragmento interno del gen <i>pkn22</i> de <i>Anabaena</i> sp. PCC7120 |
| <i>ifpkn22</i> dw | CTGCAAACACTGTGGCAGAATA | |

Tabla S-1. Oligonucleótidos utilizados para amplificar mediante PCR la región promotora del gen *all1651* y el DNA inespecífico utilizado en los ensayos de EMSA

9.3. Composición de los geles de SDS-PAGE empleados en la electroforesis desnaturalizante

| Reactivo | Gel concentrador al 5 % | Gel separador al 17 % |
|---------------------------------------|-------------------------|-----------------------|
| Agua destilada | 0,36 mL | 1,4 mL |
| Tris-HCl 1,5 M pH 8,8 | 2,24 mL | - |
| Tris-HCl 0,5 M pH 6,8 | - | 0,25 mL |
| Acrilamida/bisacrilamida (30:0,8 p/v) | 3,4 mL | 0,33 mL |
| SDS 10 % (p/v) | 30 µL | 10 µL |
| PSA 10 % (p/v) | 20 µL | 10 µL |
| TEMED | 10 µL | 10 µL |

Tabla S-2. Composición de los geles desnaturalizantes SDS-PAGE

9.4. Composición de los geles nativos al 4% empleados en los ensayos de retardo en gel (EMSA)

| Reactivo | Gel 4 % EMSA |
|---|--------------|
| Agua destilada | 6,23 mL |
| Acrilamida/bisacrilamida (30:0,8 p/v) | 1,33 mL |
| Glicerol 50 % (v/v) | 1,4 mL |
| Running buffer 10x (30,28 g/L Tris-HCl y 142 g/L glicina, pH 8) | 0,93 mL |
| PSA 10 % (p/v) | 50 µL |
| TEMED | 30 µL |

Tabla S-3. Composición de los geles no desnaturalizantes para EMSA

9.5. Secuencias de los homólogos del regulador All1651 de *Anabaena* sp. PCC 7120 en otras cianobacterias

>Anabaena_sp._PCC_7120

MFSNQNISSTNTYSQSSLEPVSILIQHLINQVKEAIFCLDYQGHFHYINDAACSLLGHSRQ
KLLNMTIDEVEMDFLPSSWSYYWQLLKQQTSLTFKKVCSSQEYQSIEVTIKSLEHGSEVLG
CILAHTFSTKNPDISSDQGLYLDEKNQPVHPSNKSLSADFPNCVQIRPIFEFIEQNYHL
PISLNDVAQAVGYSPAYLTNLVKRRQTQTLTIIDWILERKMLEARSLLINSDKSVTEIAMAVG
FTDAYYFSRRFSQYHKLSPRSWRQKYQSLQTMN

>Atlanticothrix_silvestris

MALIPQNCSSKSLAGETVKNILVIEDETKIRNLFNLCKSEGFCITGAENGIIGIQRARECS
PDLVLCDISMPELDGYGVLNLTQRDPVTAVIPFIFLTVTLTKSELRKGMNLGADDYLTQPC
SVEELLGAIARLEKQTALKQWYLNQHQSLKKSQSTDTATFYPTSSRLSKVFQYIEENYHQ
QINLSDVAQAVGYSPAYLTNLVKRQTQRSVHDWIVERMTEARYLLINTDETVNRIATKVG
YVDAGYFIRLFRQIHKLPKWRNTSRIQLVEECFSASSLN

>Coleofasciculus_sp._S288

MTKILVIEDEAQRNSFLKCLNAEGFKAIGAENFIGIQRAQEQLSDLVICDIIMPELDGY
TVLTTLRQNPVTAIIPFIFLTAKPSKAEFRQAMELGADDYLTQKSTAEELLRAIAAQLEKQ
AALRQWYALEYQQVPEPPTSETVKLADSQPFPTCPKLGEVDFIEANHHQPITLNDVARA
VGYSPAYLTDLVRRTQGTQTVQRWIKRRMAAACSLLLNTNQANQIAEAVGYNDKNYFFYQ
FRQYHGMPQMWKRTQRTPFSTQELKKF

>Cyanothece_sp._BG0011

MRSNPSFDLINSQTQERFSEPSLLTQCLINQVTEAIFCVCQTRFQYVNDAAACALIGCSRE
HLLQTTINDIEIDFFSQTWSEYWALIQEKSSVNFQTELDTEQVLDVTINFLEYDSEVY
GCILVRPLAKVSPQPLLRVVPNHQQSCQKQTSFYPNCTQLKPIFEFIEENHYTQPISLNDVA
EAVGYSPAYLTNLVKRRTRKTIINWILERMLEARSLLLNTQDSITKIAMSIGFSDAYYFS
RRFSQLHEISPKRWRQQHQMQSA

>Desmonostoc_muscorum

MQNILVIEDETRNLFVNSLKSkgfytiaaengTVGIQRAEEYSPDLVLCVQMPPELDGY
GVLNKLQRNPMTAIIIPFIFLTVKLTAKELRKGMLGADDYLTQKPCSVEELVGAIAARLEKQ
KTLKQWYLNQNLAKESQSTETTFYPTSSRLSKVFQFIEDNYHQQISLSDVAQAAGYSNA
YLTHLVKRQTKRTVHDWIVERMRTQARSLLLNTDES VNQIATKVG YLDAGYFIRSFRQIHK
LPPKEWRNASSL

>Fischerella_sp._CENA71

MKKILLIEDEIKTKNFFLNCLNIEGFHTISAENGLTGIQSAKEQLPDLVICDVVMQELDGY
SVLNKLRQDPATAIIPFIFLTGKASYTDIRRGME LGADDYIIPKCKIEEILKAITTQLEKQ
KIIKQWYTHEIQCANSKLLTGSQSIFPNCSQLNQVFQFIEENYYRPIKLSDIALVVGYS
PAYLTNLVKRKTNRSVHCWVTERRMAEARFLLINTDEPVNQIAAKVGYPDAGYFIRQFRQLYK
IPPKAWRKAQSIDS AIR

>Gloeotrichia_echinulata_DEX184

MVVDTCISVQRASQVKQLEKELLLTNLVIDRVTDAVFWVKPNAKFFVNDAAACSLVGYSR
KELLCMSIQDLNLEFLIEVWLQHWETIKQGSLSYFTSLYCNQEGLSLPVEITVTYMEEEGK
EYGCLLIRKLNRRQTPFEILASLPPAVAQTAKPSLSIFPANPLLSQIFEFIEANYHQSIS
LCDVATAIGYSPAYLTNLVRRHTGQPVNHWI IERRMAEARALLIETNQSVHEIAETVGYQN
EGHFFRQFRQHHKTPQVWRKAQREEHLKIVH

>Hyella_patelloides

MKTILVIEDEAQRKVLNCLKFEFGFKAIEADNGKSGLKLAQECHPDLIVCDIMPEMDGY
EVLSALRKHSSTLAIPIIFLTAKVSMGLDLRGLMDLGAEDYLTQKPCNVDRFLTAITTRLK
EELKKWYVSPNTSSEFFESIDLKCSKVERVFKFIEENYYKPLELKDVAQAVGYS
PAYLTNLIRQKTERTVKQWIIERRMNEARELLLNTEQSITIEIAEKTGYVDTGYFIRQFKRLHGT
TPNQWRNSPEIVEAA

>Leptolyngbya_ohadii

MRKILVIEDEVETRELFLRCLTFEGFWAIGAAGGAIGVKLAQMHTPDLIVCDIMPDLDGY
SVLSQLRQQSATAGIPFIFLTAKVTMADLRGME LGADDYLTQKCTIEQFLTAITVRLQRQ

DELKTSYDKATPALQEDQSTVFPNCPRLREVFEFIEMHYHLPISLADVAQAAGYSPPYLTT
LAQEQTGRTIKQWITERRMVQARQLLNTSOSIGQITKAVGYQDASYFIRQFRQIHGISPO
AWRQSTPTIA

>Leptothermofonsia_sichuanensis

MTFKYEQAFSLANAALQVSGDPMKTIILVIEDETQTRDIFVRSITFEGFRAIGAESGSIG
IRLAQTHLPDLIVCDIMMPDRDGYEVLSTLRQKKATAAIPFIFLTAKVTMADLRGMELGA
DDYLTKPCTVEQFLAAIATRLEQQTLKQWYTCNLGTGANAGESSTSPSAPDLAEIFPDCPRL
KPVFHFIEHYQPISLRDVAQAVGYSPAYLTNLVQNLGTGRVTKQWITERRMVQARELLMK
SAQSIGQIAEAVGYGDASYFIRQFRQVHGVSPOVWRNTSTAVEYQAV

>Mastigocladus_laminosus

MKKILLIEDEIQTRKFFLNCLNIEGFHTISAENGTGIESAKEQLPDLVICDVVMQELDGY
SVLSKLRQHPTTAIIPFIFLTGKASYSDIRGMELGADDYIIPCKIEEILKAITTQLEKQ
KIIKQWYAHEVQCADSKLLTDSQSIFPNCSQLNEVFQFIEENYRPIKLTDIALVVGYS
PAYLTNLVKKRKTQRSVYCWVTERRMAQARFLLINTDEPVNQIAAKVGYPDAGYFIRQFRQLYK
IPPKAWRKAQSIDHM

>Microcoleus_sp._SIO2G3

MAFNQGVLEQGDQSTPHPEEELQLIKSLMERVADAIFCVAPNAQLLYVNKAACRMAGYSR
KELLSMTIHDVNSDFSPKVVLEYWKRIKQONSIFYESQHWKDGQSFVEITVTYLKYENR
DYGCI FVRNIARCKSAELGLQKVHVKLEGRVTERIAKLRDTNEQLEKQLVFSQWWVTECQQ
LQQSQPTGTARLPDSQSIFLGCSQLNKVFQFIEDNYHKPINLCDVAQAVGYSPAYLTNLV
KRRTRTIHDWIVNRRMAEARSLLLETDSVNHIAAAVGYPDAGHFTRQFRRLHENS
PKVWRNRHHTLKVICALRAT

>Microcystis_aeruginosa

MLLTPNCNPLNSLNFSHSTANFLNAVPPQEIASEISVIQYLINQIPESVLWIDSQGGIKYY
NDAACALLGYSREQLSNLTIEEAAQDFLNSDWSRYWKLIIQRGFGRLITHHPYYGDHQLD
ITVNLLKYGDQDYISIVAHSLNFLESEKDGACRIHPSVNSKQAPQSLIHPSSTPQNF
FD SLASFYPNSSQLKPVFEFIEANYSQPIITLNDVAGSAGYSPAYLTNLVKKRRTGK
TIIHWIVERRMLEARSLLLKTHKSITQIAATVGFDPYYSRRESQLYKVS
PKRWREKHQTQVA

>Nodosilinea_sp._FACHB-141

MKTILIIEDDETQTRHLFLKCLEFEGFRAVGASDGTGTVAMAQQHTPNLVVCDIMMPDMDGY
SVLSALRKRQTALIIPLIFLTAKVTMADLRGMELGADDYLTKPCTVEQFLAAINSRLRQ
RQEQLSKLYADEGSHPQEPSQASLSTNIFPVEPKLANVFCFIEAHYQKPI
SLNDVAQNVGYSPAYLTNLVQSHTGRTIKQWIIERRLAEAKTLLATTANSIRQIAEASGYS
DAGYFTRQFRQFHGVSPQAWRQVSVSETTD

>Petrachloros_mirabilis

MVIEDELQTRNIFLRCLELEGFRAVVASDGTGIALAKQHS PDLVVCDIMMPDMDGYAVLS
ALRKSQRQTALIIPLIFLTAKVTMADLRGMELGADDYLTKPCTVEQFLAAINSRLTRQEQ
LS QLYGGEQTAVLEEQFPSTSPDNIFPSDPKLDVFRFIEAHYRQPI
SLNDVAQEAGYSPAYLTSLVQSQTGRTIKQWIIERRMAEARNLLVKTTQSVRHIAETSGY
TDAGYFTRQFRQFHGAS PQTWRQKSVAKLTK

>Phormidium_sp._FACHB-1136

MPAKTILVIEDEVQTRNIFLRCLEFEGFQAVGASDGTGTVAMAQQHQPDLVVCDIMMPDMD
GYAVLSTLRKVPQTAGIPLIFLTAKVAMADLRGMELGADDYLTKPCTVEQFLAAINTRLH
RQEQLSALLNAGKTSPPGQLNTPPKGGRGDPTEASGPENQPDFPDPKLPVVFQFIEAH
YQRPISLNDVAQAAGYSPAYLTNLVQTHGTGRVTKQWIIERRMAQARQLLAQTNQ
SIRTVAA TVGYGDAGYFTRQFRQCHGVSPQLWRQESVTEITKI

>Romeria_gracilis

MSKILLIEDSVRTASLFAKCLRSEGYCTFAALNGLEGVQLIKSEQPDVICDVLMQPOLDGY
GVLSHLRSDPATAVTPIFLTARCSREEVRYGMELGADDYLTKPCSFDEILKAVSTRLDKQ
KSLQWCSAECQIRQSPPALNESQSPPEPASFFPALPKLEKIFTYIEAHFCQFITLDDVA
QATNYSPAYMTHLVKKRKTQRSVYCWILERRMVEARKLLEDRLSVKEVALQIGYSDPGYFT
RQFKQYHQVSPKTTWRQQNTDPQR

>Trichocoleus_sp._FACHB-69

Material Suplementario

MTFNQSIPESSNLQFTFPQPEEELQLIKFLMERVTDVAVFCVAPDTHLLYVNSACTMLGYSR
EELLSLTMDDDIDPDFLLKVRNSNHWKTIKQQGSFIFESQHRTKDGSFPVEITVTYEYHSR
EYNCIFARALTQRKQVEVTLQKANKAVESRLEERTALEEALRQSKVEFGTFPHLNEVFHFI
EANYHQPITLSDVAQAVGYSPAYLTDLVRRTGQTVYRWI IKRRMAEAHRLLLMTSQSVQQ
IAEAVGYLDHGHFIRQFRFRNSMTPQAWRKAYLSNLTSG

>Calothrix_sp.

MLSSLNIPSVNSYSQPKLVEPISLIQSLINQVKEAIFCLDSQGRFHYINDAACSLGYPQ
KLLNMTIDELEIDFLPSTWSYYWKLQQTSLTFKKVCSKLYQSIEVTIKTLEHGSQILG
CILARTFSTKILDISAEEGLYLEQKNQPIPPSAKFLSSEFYPNCQQLRPIFEFIEENYHL
PISLNDVAQAVGYSPAYLTNLVKKRRTQLTIIDWILERKMLEARALLINSNKS SVTEIAMAVG
FTDAYYFSRRFSQYHKISPRSWRQKYQLQSA

>Westiellopsis_prolifica_IICB1

MKILLIEDEIQTRKFFLHCLNIEGFHTISAENGITGIEFAKEQLPDLVICDVVMQELDGY
GVLKSLRQHPATAIIPFIFLTGKASYDIRRMELGADDYI IKPCKIEEILKAITTQLEKQ
KIIKQWYAHEVQCADSKLLTDSQWIFPNCSQLNEVFQFIEENYYRPIKLTDIALLVVGYS
PA YLTNLVKKRKTGRSVYCWVTERRMAQARFLLINTDEPNQIAAKVGYPDAGYFIRQFRQLYK
IPPKAWRKAQSIDYM_

>Nodularia_spumigena

MKYESAKKILVIEDDTVTRNLYVKGLQAKGFDTISAENGLAGIQQAQEHLPDLVVCDITMP
DIDGYSVLTMLRQEPVTGIIPFIFITGSCTRGSRKGMELGADDYVTKPFTLDELLRAIT
T QLQRQAMLQNWCAKTFQEAQKSSADKTSAFNQPSIFPCVPQLKKVDFDIEAHYHEGITL
CDVAEAVGYSPAYLTNQVGKQTGETINSWIVKRRMQGARFLLQNDHQTVEQIARALGYQDV
SHFSRQFRQHHGLPPHAWRKEHQVGYKVKVSS

>Synechococcus_sp._Nb3U1

MSKILVIEDELTRQIFLHCLQLEGFEVFGAENGRIGIEQARRQRPDLICDIMPEVDGY
GVLHILQQDPQTARIPLIFLTAKVAKQDVRQGMVLGADDYLTCPCTVNEFLTAINARLRKQ
ALWRRSSSSALPAPSPTDTTLTDFQYPNQAKLTAVFAFIEAHYAEP IELTDVARAVGYSP
AYLTLLRQKQGRSVKDWITERRMAAARLLRNSQASIKQIGQAVGYVDPAYFARHFRQRH
GIPPLLWREQQGSFFQA

>Synechocystis_sp._PCC_7509

MSQQSSAKILVIEDDASTRNI FMDCLNCEGYTGMGAEDGLEGITQAQKHL PDLVICDVAMP
KLDGYGVLTTLRQNPLTAITSFIFLTASVSKAELRKGME LGADDYLTCPPTIGELLKAIAT
RLEKRACLRSWYANTS AQKPVTEKVNPSIFPSAPELQEVFKFIEANYHQPITLCDVAQAVG
YSPAYLTNRVGSQTGQTVNRWIVERRMVAARSLQNTDQSVQVANALGYQNTCHF SRQFR
QHHGAPPQAWRKVQQKPSVPVLSMV

9.6. Árbol filogenético del regulador All1651 de *Anabaena* sp. PCC 7120 con sus homólogos en otras cianobacterias

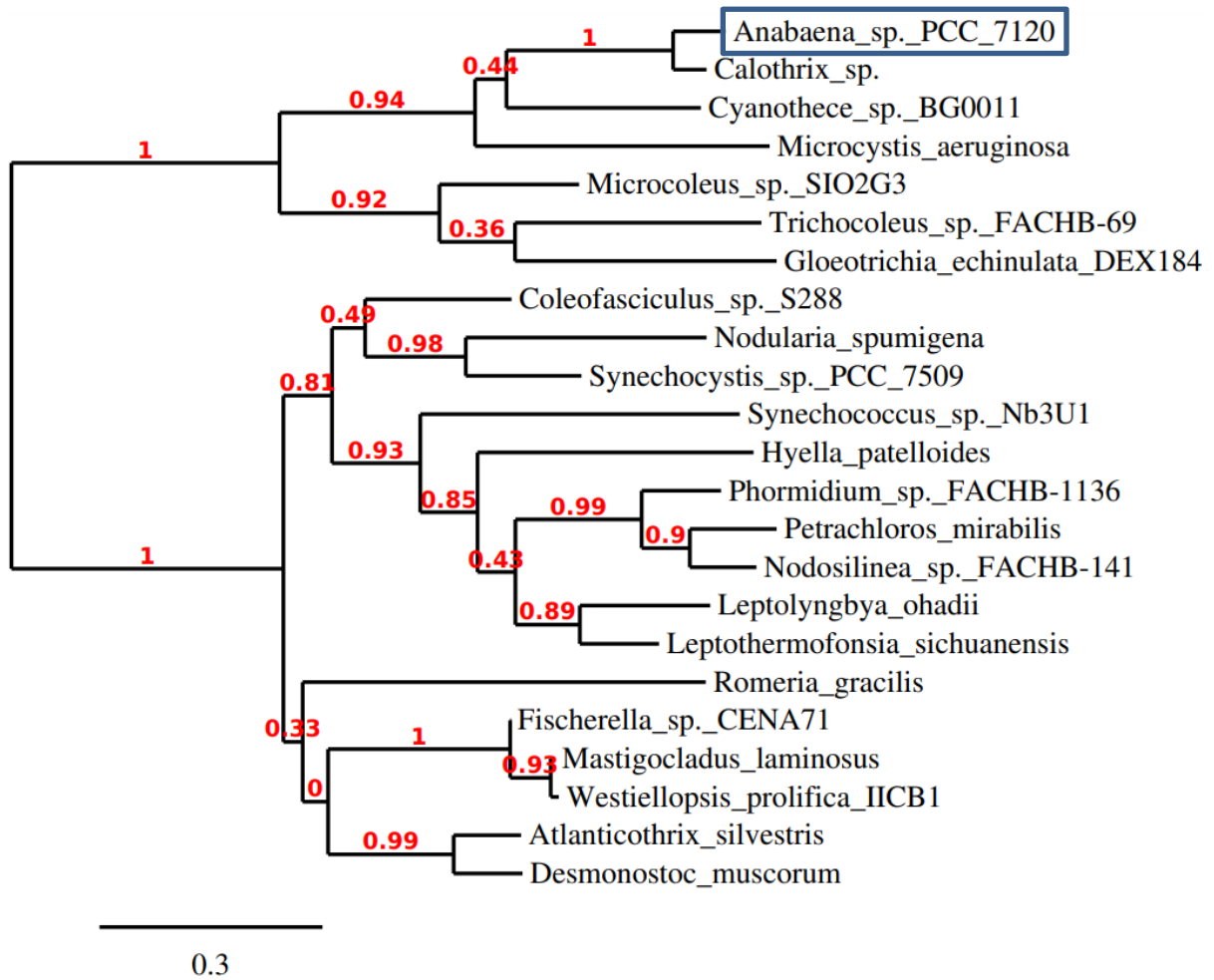


Figura S-1. Árbol filogenético del regulador All1651 de *Anabaena* sp. PCC 7120 con sus homólogos en otras cianobacterias. Se recuadra en color azul la secuencia de All1651 de *Anabaena* sp. PCC7120

9.7. Alineamiento múltiple del regulador All1651 de *Anabaena* sp. PCC 7120 con sus homólogos en otras cianobacterias

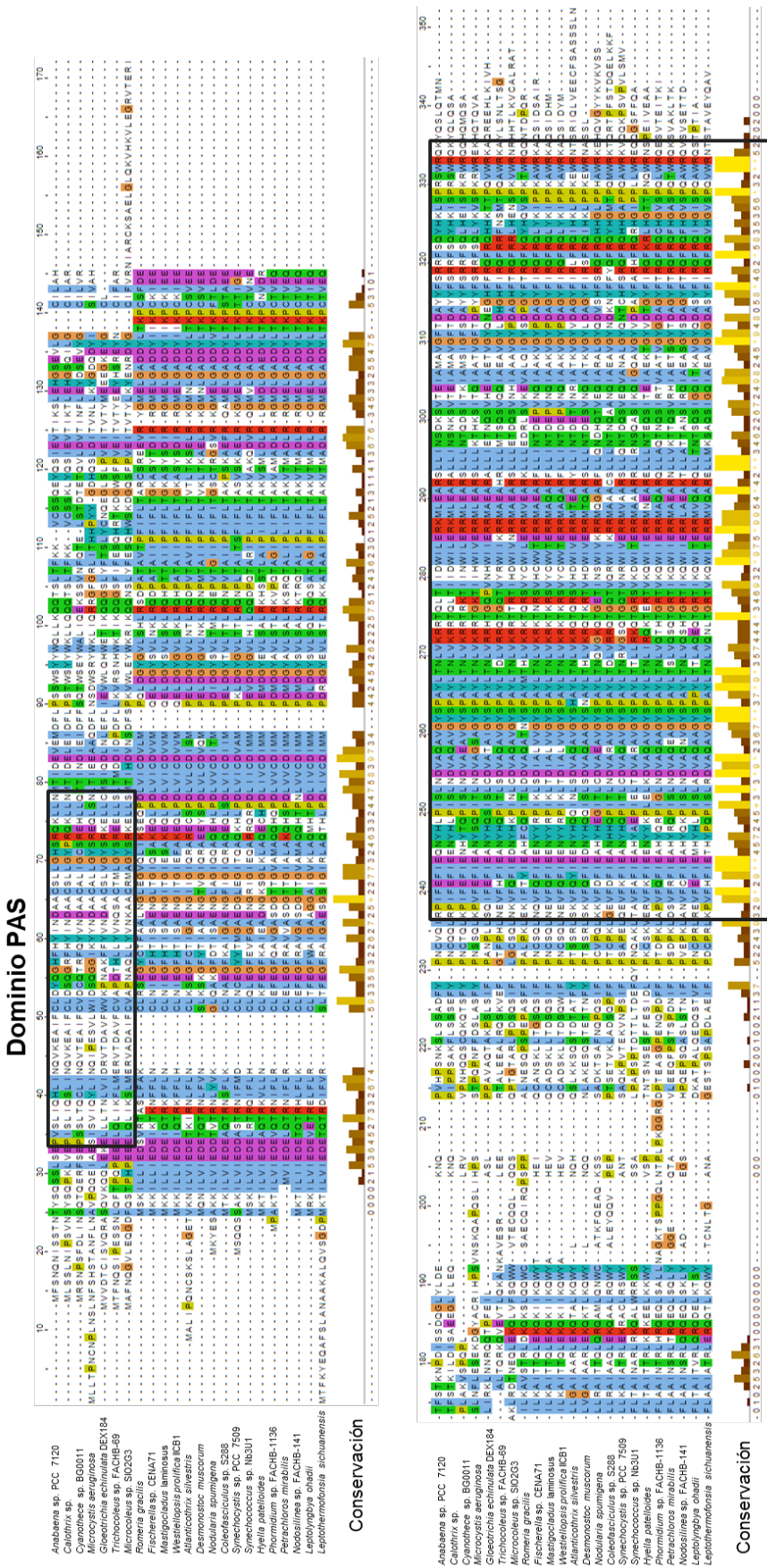


Figura S-2: Alineamiento múltiple de genes homólogos a All1651 en cianobacterias. Imagen generada mediante JalView del alineamiento múltiple de las secuencias de aminoácidos de los homólogos de All1651 en cianobacterias. Aparecen coloreadas siguiendo un código Clustal, y se encuentran recuadrados los dominios PAS, y de unión de DNA tipo AraC.

9.8. Secuencias de reguladores transcripcionales de la familia AraC en otros procariontes

```

>All1651_[Anabaena_sp._PCC_7120]
MFSNQNISSTNTYSQSSSLSEPVSLIQHLINQVKEAIFCLDYQGHFHYINDAACSLLGHSR
QKLLNMTIDEVEMDFLPSWSYQWQLLQKQTSLSLTFKKVCSSQYQSIIEVTIKSLEHGSEV
LGCILAHFTSTKNPDISSDQGLYLDEKNQPVHPSNKSLSADFPNCVQIRPIFEFIEQN
YHLPISLNDVAQAVGYSPAYLTNLVKKRRTQTLTIIDWILERKMLEARSLLINSKSVTEIA
MAVGFTDAYYFSRRFSQYHKLSPRSWRQKYQSLQTMN
>RhaS_[Escherichia_coli]
MTVLHSVDFPSPGNASVAIEPRLPQADFPPEHHDFHEIVIVEHGTGIHVFNGQPYTITGG
TVCFVRDHRHLYEHTDNLCLTNVLYRSPDRFQFLAGLNQLLPQELDGQYPSHWRVNHVS
LQQVRQLVAQMEQQEGENDLPSTASREILFMQLLLLLLRKSSLQENLENSASRLNLLLAWL
EDHFADEVNWDVAVADQFSLSLRTLHRQLKQQTGLTPQRYLNRLRLMKARHLLRHSEASVT
DIAYRCGFSDSNHFSTLFRREFNWSRPDIRQGRDGFLO
>AraC_[Escherichia_coli]
MAEAQNDFLLPGYSFNAHLVAGLTPIEANGYLDFFIDRPLGMKGYILNLTIRGQGVVKNQ
GREFVCRPGDILLFPPGEIHHYGRHPEAREWYHQWVYFRPRAYWHEWLNWPSIFANTGFF
RPDEAHQPHFSDLFGQIINAGQGEGRYSELLAINLLEQLLLRRMEAINESLHPPMDNRVR
EACQYISDHLADSNFDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRI SQAKLLLS
TTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAGCEEKVNDVAVKLS
>ExsA_[Pseudomonas_aeruginosa]
MQGAKSLGRKQITSCHWNIPTFEYRVNKEEGVYVLLLEGELTVQDIDSTFCLAPGELLFVR
RGSYVSTKGDSTRILWIPLSAQFLQGFVQRFGALLSEVERCDEPVPGIIAFAATPLLAG
CVKGLKELLVHEHPPMLACLKIEELLMLFAFSPQGPLLMSVLRQLSNRHVERLQLFMEKH
YLNEWKLSDFSREFGMGLTTFKELFGSVYGVSPRAWISERRILYAHQLLLNSDMSIVDIA
MEAGFSSQSYFTQSYRRRFGCTPSRSRQKDECRANN
>GbdR_[Pseudomonas_aeruginosa]
MTTYAPGVPPQNRNPQSIGFLLLDNFTLISLASEVEPLRMANQLSGRELYRWHTLSLDGR
QVWASDGLQITPDAGTDNAPAVDCVIVCGGVGIQRSVTRHVTFLQAQARQGRRLGAVCT
GSWALARAGLLDGYDCSVHWECLAAMQEAFFPRVAMSTRLFSDRNRFTSSGGTAPMDMML
HLIGREHGRELAAISEMFIYERIRNEQDHQRVPLKHMLGTNQPQLQEI VALMEANLEEF
IDLDELAVYVNVSRRLERLQKYLHCSPSRYYLKLRLIRARQLLKQTSMSIEVASVCG
FVSTPHFSKCYREYFGIPPRDERQGGPLGQPVVLMPI PQDLALMPNSSALSALSQAQGES
TFASVRI
>RhaR_[Escherichia_coli]
MAHQLKLLKDDFFASDQQA VAVADRYPDVFAEHTHDFCELVIVWRGNGLHVLNDRPYRI
TRGDLFYIHADDKHSYASVNDLVLQNIICYPERLKLNLDWQGAIPGFNASAGOPHWRLGS
MGMAQARQVIGQLEHESSQHVPFANEMAELLFGQLVMLLNHRHYTSDSLPPTSSETLLDK
LITRLAASLKSFPALDKFCDEASCSEVLRQQFRQQTGMTINQYLRQVRVCHAQYLLQHS
RLDISDISTECGFEDSNYFSVVFTRETGMTPSQWRHLNSQKD
>SoxS_[Escherichia_coli]
MSHQKIIQDLIAWIDEHIDQPLNIDVVAKKSGYSKWYLQRMFRTVTHQTLGDYIRQRLL
LAAVELRITTERPIFDIAMDLGYVSQQTFSRVFRRQFDRTPSDYRHRL
>MarA_[Escherichia_coli]
MSRRNTDAITIHSILDWIEDNLESPLSLEKVSERSGYSKWHLQRMFKKETGHS LGQYIRS
RKMTEIAQKLKESNEPILYLAERYGFESQQTLTRTFKNYFDVPPHKYRMTNMQGESRFLH
PLNHYS
>GadX_[Escherichia_coli]
MQSLHGNCLIA YARHKYILTMVNGEYRYFNGGDLVFADASQIRVDKCVENFVFSRDTLS
LFLPMLKEEALNLHAHKVSSLLVHHCSDIPVFQEVQALSQNKNLRYAEMLRKRALIFA
LLSVFLEDEHFIPLLLNVLPNMRTRVCTVINNNIAHEWTLARIASELLMSPSLLKKKLR
EEETSYSQLLTECRMQRALQLLIVIHGFSIKRVAVSCGYHSVSYFIYVFRNYYGMTPTEQ
ERSAQRLSNRDSAASIVAQGNFYGTDRSAEGIRL
>VirS_[Mycobacterium_tuberculosis]

```

MELGSLIRATNLWGYTDLMLRELGADPLPFLRRFDIPPGIEHQEDAFMSLAGFVRMLEASA
AELDCPDFGLRLARWQGLGILGPVAVIARNAATLFGGLEAIGRYLYVHSPALTLTVSSTT
ARSNVRFVGYEVEPIPIYPLQGYELSMANAARMIRLLGGPQARARVFSFRHAQLGTDAAY
REALGCTVRFGRTWCGFEVDHRLAGRPIDHADPETKRIATKYLESQYLPSDATLSERVVG
LARLLPTGQCSAEAIADQLDMHPRTLQRRLAEEGLRCHDLIERERRAQAARYLAQPGLY
LSQIAVLLGYSEQSALNRSRWFGMTPRQYRAYGGVSGR
>NphR_[Rhodococcus_sp.]
MAEREQSNDSARTDVPAIVSLRTRELDTGEGRMQWASTLERLYCETDVAWPEPRRHFD
WGGRPFGLHVVSTIRADAHTVVRSPAMIQSDSGEGYLVCLVTDGSVEVRQSGRATVVEPG
SFALLDCAAPFVHSPAPFRQVVVRSPREVLTSRLPGRIVEHGTARS IHGDTGAGGLVGR
LFVDIADMDAPMSQGAAVSFASSAVDMLATALTEGLLATSAADLHRTEDLTRVQRVIEQN
LHDADITLSDIAAAAGMSLRVTVHKL FNAEGTTTRAWLYQARLEAARRYLLTTDLSVADVS
ECAGFRDVSHFSRLFRSTFGSSPGLYRKEHARIGS
>ArgR_[Pseudomonas_aeruginosa]
MTAQPPRIGFLLWPATRALTLSLAEEALRAARRLHPEALYEPLFLLAEPAAEEEGWRLPG
TAWNGRLEQCSRLFLVADEAPAAVSPALGLALKQLARSGAAIGALSAGIYPLAQLGLLDG
YRAAVHWRWHDDFTERFPKVIATNHLFEWDRDRMTACGMAVLDLLLALLSRDHGAELAG
AVSEELVVERIREGNERQRIPLKNRLGSSHPKLTQAVLLMEANIEEPLTTDEIAQHVCVS
RRQLERIFKQYLNRVPSQYYLELRLNRARQMLMQTSKSI IQIGLSCGFSSGPHFSSAYRN
FFGVTPREDRNQRRGGSASFETTFTPVERG
>AppY_[Escherichia_coli]
MDYVCSVVVICQSFDLIINRRVISFKKNSLFIIVSDKIRRELPVCP SKLRIVDIDKKTCLS
FFIDVNNELPGKFTLDKNGYIAEEEPPLSLVFSLFEGIKIADSHSLWLKERLCISLLAMF
KKRESVNSFILTNINTFTCKITGIISFNIERQWHLKDIAELIYTSESLIKKRLRDEGTSF
TEILRDTRMRYAKKLITSNSYSINVVAQKCGYNSTSYFICAFKDYGVTPSHYFEKIIGV
TDGINKTID
>MelR_[Escherichia_coli]
MNTDTFMCSSDEKQTRSPLSLYSEYQRMIEIFRAPHIMPTSHWHGQVEVNVFPDGDVEYL
INNEKVNINQGHITLFWACTPHQLTDTGTCSMAIFNLP MHLF LSWPLDKDLINHVTHGM
VIKSLATQQLSPFEVRRWQOELNSPNEQIRQLAIDEIGLMLKRFSLSGWEPILVNKTSRT
HKNSVSRHAQFYVSQMLGFIAENYDQALTINDVAEHVKLNANYAMGIFQRMVQLTMKQYI
TAMRINHVRRALLSDTDKSIDIALTAGFRSSSRFYSTFGKYVGMSPQQYRKLSSQRRQTF
PG
>TcpN_[Vibrio_cholerae]
MIGKKSFTQTNVYRMSKFDTYIFNNLYINDYKMFWIDSGIAKLIDKNCLVSYEINSSSIIL
LKKNSIQRFSLTSLSDENINVSVITISDSFIRSLKSYILGDL MIRNLYSENKDLLLWNCE
HNDIAVLSEVVNGFREINYSDEFLKVVFFSGFFSKVEKKYNSIFITDDL DAMEKISCLVKS
DITRNWRWADICGELRTNRMILKKELESRGVKFRELI NSIRISYSISLMKTGEFKIKQIA
YQSGFASVSYFSTVFKSTMNVPSEYLFMLTGVAEK
>EnvY_[Escherichia_coli]
MQLSSSEPCVVILTEKEVEVSVNNHATFTLPKNYLAAAFACNNNVIELSTLNHVLIITHINR
NIINDYLLFLNKNLTCVKPWSRLATPVIACHSRTPEVFR LAANHSKQPPSRPCEAELTRA
LLFTVLSNFLEQSRFIALLMYILRSSVRDSVCRI IQSDIQHYWNLRIVASSLCLSPSLK
KKLKNENTSYSQIVTECRMRYAVQMLLMDNKNITQVAQLCGYSSTSYFISVFKAFYGLTP
LNYLAKQRQKVMW
>ChbR_[Escherichia_coli]
MMQPVINAPEIATAREQQLFNGKNFHVFIYNKTESISGLHQHDYEF TLVLTGGRYFQEIN
GKRVLLERGDVFIPLGSHHQSFYEFGATRILNVGISKRF FEQHYP LLLPYCFVASQVYR
TNNAF LTYVETVISSLNFR ETGLEEFVEMVTFYVINRLRHYREEQV IDDPQWLKSTVEK
MHDKEQFSESALENMVALSAKSQEYLTRATQRYYGKTPMQI INEIRINFAKKQLEMTNYS
VTDIAFEAGYSSPSLFIKTFKKLTSFTPKSYRKKLTFE FNQ
>AdiY_[Escherichia_coli]
MRICSDQPCIVLLTEKDVWIRVNGKEPISLKHANMALLNCENNIIDVSSLNNTLVVAHISH

DI IKDYLRFLNKDLSQIPVWQRSATPILTLPLCLTPDVFRVAAQHSMMPAETESEKERTRA
 LLFTVLSRFLDSKFKVSLMMYMLRNCVSDSVYQIIIESDIHKDWNLSMVASCLCLSPSLLK
 KKLKSENTSYSQIITTCRMRVAVNELMMDGKNISQVSQSCGYNSTSYFISVFKDFYGMTP
 LHYVSQHRERTVA
 >PchR_[*Pseudomonas_aeruginosa*]
 MTITIIAPPQADAAAPAPGNRPGVAHIDPNMKLVGTGFCASSEDWFEEPLERGLRLILVQ
 SGQLRCRIPGQPEHLIEGPSLCTIANDGDFTSAQIYGTDKPLRYTIVQLGVEALDSRLGW
 LPEQLIRRPGGDPRIMSCPAPRAMQALASQIATCQMLGPTRDLYLGGKALELAALSAQFL
 SGEGRPVEEPRITCSEVERIHAARDLLVGALQEPPSLDTLASRVGMNPRKLTAGFRKVFV
 ASVFGYLRQYRLREAHRLCDEEANVSTVAYRVGYSPAHSIAFRKRYGISPSEIR
 >YdeO_[*Escherichia_coli*]
 MSLVCSVIFIHAFNANILDKDYAFSDGEILMVDNAVTRTHFEPYERHFKEIGFTENTIKK
 YLQCTNIQTVTVPVPAKFLRASNVPTGLLNEMIAYLNSEERNHHNFSSELLLFSCLSIFAA
 CKGFITLLTNGVLSVSGKVRNIVNMKPAHPWKLDICDCLYISESLKKKLQEQTTFSQ
 ILLDARMQHAKNLRVEGVSVNKIAEQCGYASTSYFIYAFRKHFGNSPKRVSKEYRCQSHT
 GMNTGNTMNALAI
 >NimR_[*Escherichia_coli*]
 MMHRLNLNGYEPDRHHEAAVAFCIHAGTDELTSVPVQHRKQQLILALHGAICTVENALW
 MVPPQYAVWIPGGVEHSNQVTANAELCFLFIEPSAVTMTPTCCTLKISPLCRELILTLAN
 RTTTQRAEPMTRRLIQVLFDELPOQPQQQLHLPVSSHPIKIRTMVEMMAKGPVEWGALGQW
 AGFFAMSERNLARLIVKETGLSFRQWRQQLQALIMALQGLVKGDTVQKVAHTLGYDSTTAF
 ITMFKKGLGQTPGRYIARLTTVSPQSAKPDPRQ
 >PocR_[*Salmonella_typhimurium*]
 MISASALNSELINKIAQDFAQATGLAVVVVNIHGDEISELNFNTPFCQLMRQHPQHSTRC
 RMSDRCGGLEASKSDQPCIYRCHAGLTDFSIPLVIAGHLVGFVLCGQVRLSNDVELVNIL
 NVDDRWDADPELLNEFRNVPEMDYSRVIASADLLKLIVENCLKKQLNFVVIKDNPQQSEA
 NKTTRGPTPHDSKMKKALRYIDAHLSDDLRLEDVASHVYLSPPYFYSKLFKQYQIGFNAW
 VNRQRMVSARELLCHSDWSIASIARNLGFSTSYFCKVFRQTYQVTPQAYRQQINENSHP
 PSL
 >HptR_[*Staphylococcus_aureus*]
 MFKVVICDDERIIREGLKQIIPWGDYHFNTIYTAKDGEALSIIQHQPELVIDTIRMPR
 KNGVDLLNDIAHLDCNVIILSSYDDFEYMKAGIQHHVLDYLLKPVDDHAQLEVLGRLVRT
 LLEQQSQNGRSLASCHDAFQPLKVEYDDYVNVQIVDQIKQSYQTKVTVSDLIQHIDVSE
 SYAMRTFKDHVGITIVDYLNRIRILQSLQLLDRHYKHYEIADKVGFSYKMFYHFKKYL
 QMSPSDYCKQAK
 >FeaR_[*Escherichia_coli*]
 MNPAVDNEFQQWLSQINQVCGNFTGRLLTERYTGVLDTHFAGLKLSTVTTSGVNLRSRTW
 QEVKGSDDAWFYTVFQLSGQAIMEQDERQVQIGAGDITLLDASRPSLYWQESSKQISLL
 LPRTLLEQYFPHQKPICAERLDADLPMVQLSHRLLQESMNNPALSETESEAALQAMVCLL
 RPVLHQRESVQPRRERQFQKVVTLIDDNIREEILRPEWIAGETGMSVRSLYRMFADKGLV
 VAQYIRNRRLDFCADAIRHAADDEKLAGIGFHWGFSQSHFSTVFKQRFGMTPEYRRKF
 R
 >AndR_[*Burkholderia_cepacia*]
 MSPTSFEPLALRAHRLFESRDLDRETRERISRVMPHALLPSGRMQGAAHMDFVRLGGLGI
 GTIAFGDAMRVQVDAVDGYLLMFCLSGHAQVAAMGRRLGVDGQTVGLCAPGERFDAVLS
 ADCEQFVLRIDAATVGLTGNPRATLDPVLHISDAALAAWHQQMLMLVARSPALLARANAN
 PRVAAQLEHLLIDLLIEGHPPAAASMRHDPAGFLLRAQEFVNAHYAQPLQLADIVVAAN
 VPERTLRDAFLQFRGMSPMQYLRAATRLEHARELLRRSQPERRIADVALDCGFTHLGRFAI
 AYREKFGESPSETVGVRR
 >RipA_[*Corynebacterium_glutamicum*]
 MSSASLLWCHSGVSTVRFGERIFTLVAGDLLFAPEEAQVADDSQGLVLNIRFETLNIMGP
 ARRIHLGHVWNRDLTFEYSRSLFGKETLSPDIARLFTDRVPTPLPAPRKARAVAQVLVS
 NPADQTSLEEFAEIQGVSARTLQRQFLKSTGYSFSEWRAAQRVCVAASLLAHDFSISVVA

NLVGFAATSSSLTRAFRRHTGATPSTFTTGGQIGMGSAGHPPRI PATTTFAEAHQDQQLWIY
SGTATVTTTPGYCRFMGQDMVTI PAGTQTRIDVAAGSIAFPVPVGLDEWGM DLTRVVAVN
NQPKPLTILEQSEWSKLSEELLNTPVPVQM
>CdhR_[*Pseudomonas_aeruginosa*]
MSQDFWFLLLPGFSVMGFVSAVEPLRVANRFHADLYRWHVLSADGGPVLASNGMSVNSDG
ALEPLKKGDLLFVVAGFEPLRAVTPALVQWLRKLDNRNGVTLGGIDTGSVVLAEAGLLDGR
RATLHWEAIDAFQESYPQLSVTQELFEIDGPRI TSAGGTASIDLMLDLIAQAHGPQLAVQ
VSEQFVLGRIRPRQDHQRLQVATRYGVSNRKLVQVIGEMERHTEPPLTTLELAERIQVTR
RQLERLFRVHLD DTPSNFYLGRLDKARQLLRQTDLSVLQVSLACGFESPSYFSRSYRAR
FAASPSQDRAVLPLKAPAATPPGAPAGHRTPraERG
>VqsM_[*Pseudomonas_aeruginosa*]
MVTEHIFDFEWRPLQNYLRDRDIAIPFAAEQESRSDALYNWIVSKGYSGKEGLNLGTY
HISDYGVIGLALLCAENVGDILKVIRAYVKLFNRDIANVGKPRENYEVEIHSVNFKPE
WNDASRQFHVNVIASATYKLIMDLLGNDFPISAL TVPVHGADKTIYEGFFSLPVKHEGND
IIFSFPAHQLNRTLATANPAVFQSALTMAGESFNALLEVEMGGLRQRIELFLDSIPDQYP
SLVTTAKYLRMNERTVRRRLADEGCTYRQIVDKARKERIAALLLNSSIPVDRI SDILGYS
ETASFRHAFRRWTGQSTTEFRNSFH
>MtrA_[*Neisseria_gonorrhoeae*]
MDILDKLVDLAQLTGSADVQCLLGGQWSVRHETLQCEGLVHIVTAGSGYLCIDGETSPRP
VGTGDIVFFPRGLGHVLSHDGKYGESLQPDIRQNGTFMVKQCGNGLDMSLFCARFRYDTH
ADLMNGLPETVFLNIAHPSLQYVVSMLQLESEKPLTGTVSVVNALPSVLLVLI LRAYLEQ
DKDVELSGVLKGWQDKRLGHLIQKVIDKPEDEWNIDKMVAANMSRAQLMRRFKSQVGLS
PHAFVNHIRLQKGALLLKKTPDSVLEVALSVGFQSETHFGKAFKRQYHVSPGQYRKEGGQ
K
>EutR_[*Salmonella_typhimurium*]
MKKTRTANLHHLIHEALPEDVKLTPRVEVDNVHQRRTTDVYEHAL TITAWQQIYDQLHPG
KFHGEFTEILLDEIQVFREYTG LALRQSCLVWPNSFWFGIPATRGEQGFIGAQGLGSAEI
ATRPGGTEFELSTPDDYITILGVVISEDVISRQATFLHNPERVLHMLRNQLALEVKEQHK
ALWGFVQALATFSESPETLHQPAVRKVLSDNLLLAMGMTMLEEAKPIHSAESISHQGYRR
LLSRAREYVLENMSEPLTVLDLCNQLHVSRRTLQNAFHAILGIGPNAWLKRIRLNAVRRE
LISPWSQSATVKDAAMQWGFWHLGQFATDYQQLFAEKPSLTLHQMRQWA
>SirC_[*Salmonella_typhimurium*]
MVLPSMNKSVEAISNNHLQQPNKFPLINGLADVRDYYVANCLL FKLNGSLRIENEFGEF
IERSAPCLFLLKEDQTITLSMSEIEGHIDFSSLEVS YDLMQKFYKVFYSTRNYNDRELSL
KTKPKYFFHADLLPGMSDTFDSILHG VACPRVCSNVSIDDHDYSYFSLMYLISAFVRKPG
GFDFLERAIKITTEKVVYNI IISDLTRKWSQAEVAGKLFMSVSS LKRKLAAEEVSFSKIY
LDARMNQAIKLLRMGAGNISQVATMCGYDTPSYFIAIFKRHFKITPLSFMRTMNH
>LacR_[*Staphylococcus_xylosus*]
MQVLWKKFQKKLIDANLAECGIEIGV PNVGYN YTVFQKSVLHIVTQEGGTFSYAGETYHL
TAGDIFLLERGMEVEYKPSFSNPW TYYWVG MNGKQILNYLSRCSIVDSHVILGQD TTDIK
NIIQLICKLSQSIESNNSNDILIMQYLYQLVYTLQEKFPKIFSVQVDIVNEDIQHAVDFI
NTNYQKHITVEDVAKSVNITRSHLYKLFKKNLGCSPKEYLTYIRMYHASQLLIHTSTLIS
DISRQVGYKDPLLFSKNFTKHFEISASEYRHHFSINN KQ
>MxiE_[*Shigella_flexneri*]
MEGFFFVRNQNIKFSDNVNYHYRFNINSCAKFLAFWDYFSGALVEHSHA EKCIHFYHEND
LRDSCNTE SMLDKLMLRFIFSSDQNVSNALAMIRMTESYHLVLYLLRTIEKEKEVRIKSL
TEHYGVSEAYFRSLCRKALGAKVKEQLNTWRLVNGLLDVFLHNQTITSAAMNNGYASTSH
FSNEIKTRLGFSARELSNITFLVKKINEKI
>XylR_[*Escherichia_coli*]
MFTKRHRITLLFNANKAYDRQVVEGVGEY LQASQSEWDIFIEEDFRARIDKIKDWLGDGV
IADFDDKQIEQALADVDVPIVGVGGSYHLAESYPPVHYIATDNYALVESAF LHLKEKGVN
RFAFYGLPESSGKRWATEREYAFRQLVAEEKYRGVVYQGLE TAPENWQHAQNRLADWLQT
LPPQTGIIAVTDARARHILQVCEHLHIPVPEKLCVIGIDNEELTRYLSRVALSSVAQGAR

QMGYQAAKLLHRLLDKEEMPLQRILVPPVRVIERRSTDYRSLTDPAVIQAMHYIRNHACK
GIKVDQVLDAVGISRSNLEKRFKEEVGETIHAMIHAEKLEKARSLLISTTLSINEISQMC
GYPSLQYFYFVFKKAYDTPKEYRDVNSEVML
>UreR_[Escherichia_coli]
MEYKFFLSESQMVLKAFYIESAMIAMLTGAKGNIVINGQSIEIESDVTLIIPKYSQVSCN
IVCNTVKQPLELHTLSMSAEELQAVFLLKTLMKSGAPLTSHQPIYHISPPETVRDNFSL
LKQCLPLKKQSASQEALLMKQSLYFILMAIYQEGIDILNLFNFYDEPKNQAITHLITQE
PQKKWHLDDVAKALFTTPSTLRRHLNREGVSRQQLLDVRMGMALNYLTFSNYSVFQISH
RCGFGSNAYFCDVFKRKYNMTPSQFRLQSRQSNPNFITNLSLRSNPIEFDKEIDE

9.9. Árbol filogenético del regulador All1651 de *Anabaena* sp. PCC 7120 con reguladores transcripcionales de la familia AraC en otros procariotas

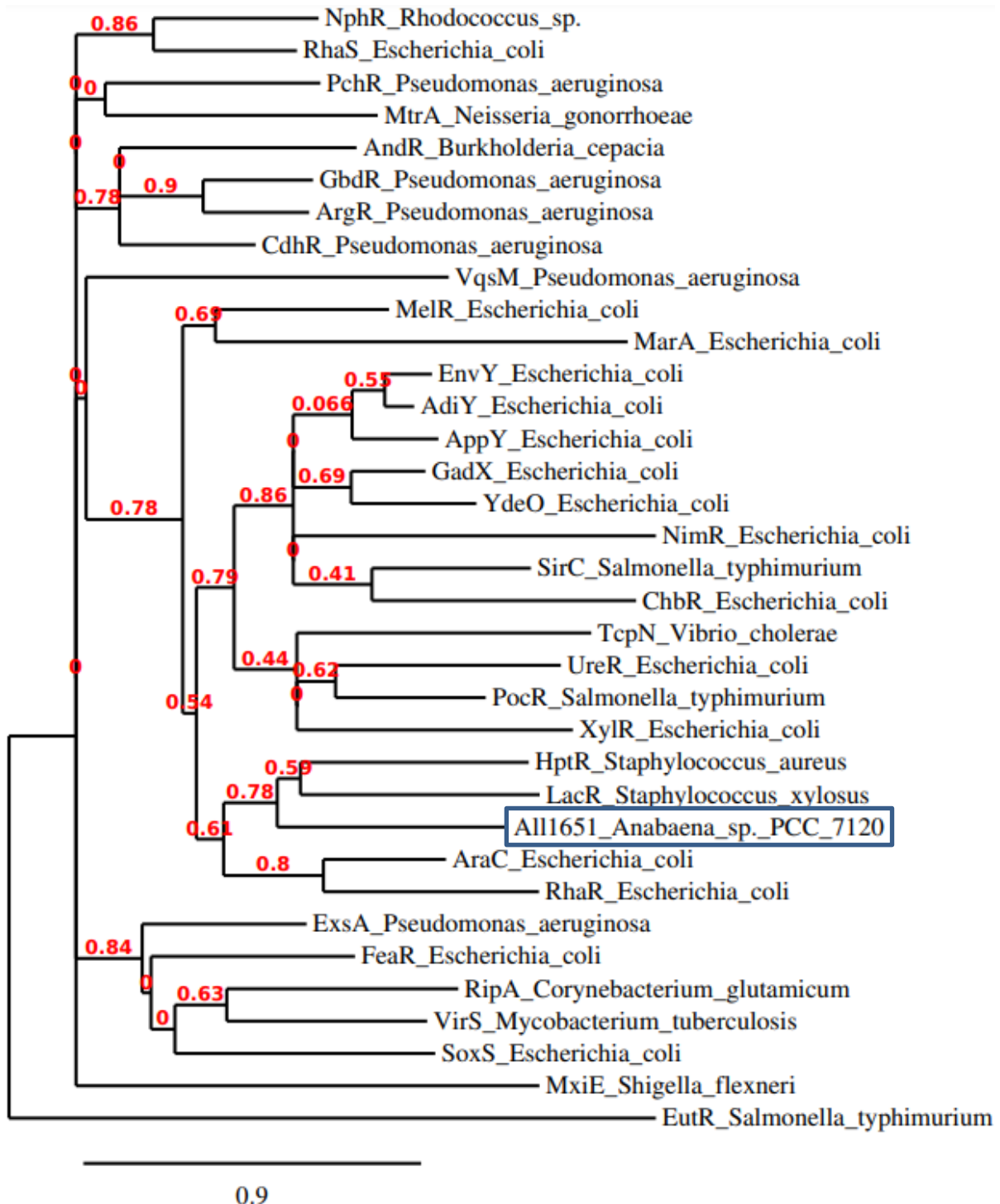


Figura S-3. Árbol filogenético del regulador All1651 de *Anabaena* sp. PCC 7120 con reguladores de la familia AraC en otros procariotas. Se recuadra en color azul la secuencia de All1651 de *Anabaena* sp. PCC7120

9.11. Purificación del regulador All1651 mediante cromatografía de afinidad a metales en Tris 50 mM pH 8

El regulador All1651 se intentó purificar en primer lugar a partir de biomasa de células BL21 sobreexpresando esta proteína y mediante cromatografía de afinidad a iones metálicos con Co^{2+} (IMAC- Co^{2+}) empleando como tampón base Tris 50 mM pH 8. Tras la purificación se analizaron todos los pasos de purificación y se observó que gran parte de la proteína quedaba en el precipitado resultante de clarificar el extracto crudo (Figura S-5).

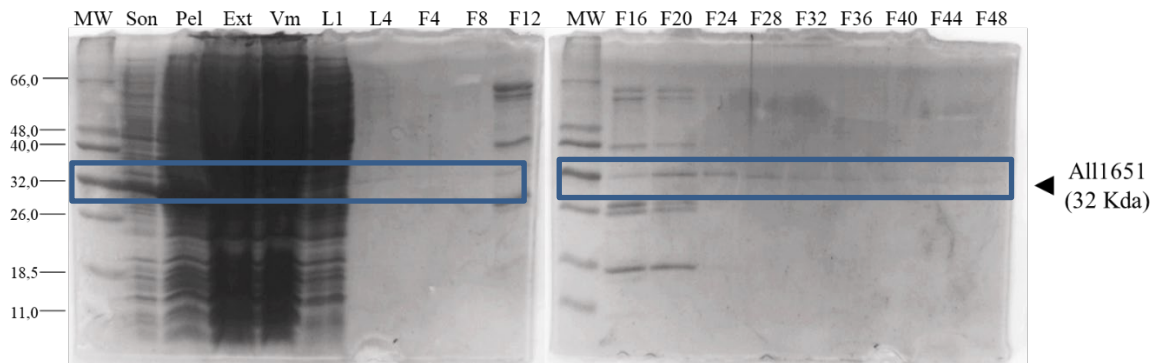


Figura S-5. Purificación del regulador All1651 con Tris 50 mM pH 8. En el gel se observan las diferentes etapas del proceso de purificación: biomasa sonicada (Son), precipitado tras la clarificación (Pel), extracto crudo (Ext), volumen muerto (Vm), lavado 1 (L1), lavado 4 (L4) y fracciones eluidas del gradiente de imidazol (F4-F48) La banda que aparece a una altura de 32 kDa corresponde al regulador All1651. El primer carril (MW) corresponde al marcador de peso molecular Low Molecular Weigh Protein Marker II de NZYTech.