



**BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA**

---

---



**Centro de Investigaciones en Ciencias Microbiológicas  
CICM-ICUAP**

“Análisis bioinformático y clonación del gen *cdgG* de  
*Azospirillum baldaniorum Sp245*”

Tesis presentada para obtener el grado de:  
Maestría en ciencias (microbiología) con opción en  
bioquímica y genética

PRESENTA:

**LBT. RUTH ABRIL VELA SANCHEZ**

Asesor:

**DR. ALBERTO RAMIREZ MATA**

**Noviembre, 2024**

## **Agradecimientos**

Un poco de informalidad antes de la inminente formalidad:

A todos y cada uno de ustedes, no tengo más que una página para expresar mi eterno agradecimiento y cariño en mi trayectoria profesional y ser mi apoyo cuando más necesitada, angustiada y débil estaba. Gracias por hacer más llevaderas las quimioterapias y mi separación. Los llevaré conmigo. Mi alma y cuerpo han sido sanados y me permiten seguir viviendo para continuar. GRACIAS INFINITAS Y TOTALES.

Al Dr. Alberto Ramírez Mata por permitirme incorporarme a su equipo de trabajo y por su apoyo y comprensión en cada etapa de mi vida personal y profesional. Gracias por sus lecciones inculcadas.

A la Dra. Sandra por su ayuda y orientación en la realización de los experimentos y por ser una amiga en mis etapas difíciles.

A la química Mary, por todo su apoyo en mis experimentos y ayudarme a pulir técnicas. Gracias por cuidarme y aconsejarme como si fuera su hija.

A cada uno de los amigos que hice en el laboratorio: Vania, por ser más que una amiga incondicional. David "confi", gracias por tu orientación y por tus consejos que me ayudaron. Ricardo "richis", por ayudarme en la parte bioinformática mientras nos reíamos. Uriel "nene", por tu sonrisa franca y tu preocupación sincera. Emiliano "nene", gracias por reírte de mis chistes malos y las buenas recomendaciones; Job "padrino", gracias por tu apoyo y carisma cuando más lo necesite. Mis niños de biomedicina: Dafne y Oscar, gracias por su sonrisa y abrazos cuando lo necesite. A mi MVZ querida, Daniela. Mi amiga y hermana de la fe. Gracias por cuidarme y enseñarme a cocinar.

A ustedes, gracias por cada momento que llevo en mi corazón y cada risa que tuvimos, cada experimento que realizamos y por todo su apoyo y amor en el momento más crítico de mi vida.

A mis padres, Ruth y Donato. Gracias por su amor, apoyo incondicional y por luchar a la par mía en mis momentos de mayor oscuridad. Me debo a ustedes.

A mi hermano, Misael. Gracias por tu amor, preocupación y por ser mi compañero "otaku" y comer ramen conmigo. Me debo a ti.

Gracias al que me sostuvo en todo momento. Mi amparo, escudo y fortaleza. Mi auxilio en las tribulaciones. Todo es por El, de El y Para El. A Dios la Honra y el poder por los siglos.

Feel pain. Contemplate pain. Accept pain. Know pain. One who does not know pain cannot possibly understand true peace.

## Tabla de imágenes

<b>Figura</b>	<b>Título</b>	<b>Página</b>
1	Diagrama del sistema de quimiotaxis en <i>E. coli</i>	8
2	c-di-GMP.	9
3	Un modelo para señalización de adentro hacia afuera a través de LapD.	13
4	Receptores del c-di-GMP	16
5	Diagrama de proteínas híbridas implicadas en el metabolismo del c-di-GMP de <i>A.baldaniorum</i> Sp 245.	18
6	Contexto genético del gen AZOBR_30001	22
7	Diagrama de la proteína CdgG donde se muestran los dominios EAL, GGDEF Y CBS.	23
8	Identificación de un probable promotor del gen <i>cdgG</i> de <i>A. baldaniorum</i> Sp245	24
9	Identificación de posibles factores transcripcionales que modularían la expresión del gen <i>cdgG</i>	25
10	Alineamiento múltiple de los motivos característicos de DGC de la proteína CdgG	26
11	Alineamiento múltiple de aminoácidos del dominio EAL de la proteína CdgG	27
12	Estructura tridimensional de la proteína CdgG generada por AlphaFold.	29
13	Empalme estructural entre los dominios CBS presentes en CdgG y la proteína DarB de <i>Bacillus subtilis</i> .	31
14	Acoplamiento del dímero de CdgG con el tetrámero de DarB de <i>Bacillus subtilis</i>	32
15	Análisis filogenético del gen <i>cdgG</i>	33
16	Esquema de trabajo para la clonación del gen <i>cdgG</i>	34
17	Condiciones para la PCR del producto CdgG	35
18	Producto de PCR.	34
19	Clonas obtenidas producto de la transformación en <i>E. coli</i> DH5a	37
20	Producto de la restricción	38

## Índice

.....	1
Tabla de imágenes.....	3
Glosario .....	5
Resumen.....	6
1.Introducción .....	8
2.Antecedentes .....	20
3.Justificación .....	23
4.Objetivos .....	23
5.Esquema de trabajo .....	24
6.Materiales y métodos .....	25
7.- Discusión .....	42
8.- Conclusión.....	48
9.-Perspectivas .....	48
10.- Bibliografía .....	49
11.-Anexos.....	54

## Glosario

DGC: di guanilato ciclasa

PDE: fosfodiesterasa

GTP: guanidina trifosfato

di-GMPc: Ácido di-(3'-5') guanosína monofosfato

c-di-GMP: cyclic di guanylate (di-guanilato cilcasa)

PGPR: Plant Growth Promoting Rhizobacteria (rizobacterias promotoras del crecimiento de plantas)

CdgG: Diguanilato ciclasa G

EPS: Exopolisacáridos

Km: Kanamicina

MPM: marcador de peso molecular

Tc: Tetraciclina

LB: Luria Bertani

Ap: Ampicilina

ARN: Ácido Ribonucleico

TAE: Tris-acetato-EDTA

ADN: Ácido desoxirribonucleico

PCR: Reacción en cadena de la polimerasa

MCP: Proteína aceptora de grupos metilo

SPE: Sustancias poliméricas extracelulares

## Resumen

*Azospirillum baldaniorum* Sp245, una rizobacteria promotora del crecimiento de las plantas, es una de las cepas de mayor interés debido a la gran variedad de beneficios que le confiere a la planta entre los que destacan la producción de fitohormonas, la fijación de nitrógeno, la protección contra patógenos y diversos tipos de estrés. Estos beneficios requieren de una colonización exitosa que inicia con la adherencia de la bacteria y mantenimiento de la interacción *Azospirillum*-planta a través de la formación de biopelículas sobre las raíces de la planta hospedera. La formación de biopelículas es a través de un proceso controlado por c-di-GMP. Los bajos niveles de c-di-GMP favorecen un estado de movilidad en la bacteria, mientras que cuando incrementan los niveles, se favorece la expresión de exopolisacáridos y adhesinas, los que desembocan en la formación de biopelícula. Los niveles intracelulares del c-di-GMP están influenciados por su síntesis y degradación.

El c-di-GMP se sintetiza a partir de 2 moléculas de GTP por la acción de las diguanilato ciclasas (DGC) y es degradado por fosfodiesterasas específicas de c-di-GMP (PDE). Estas enzimas que sintetizan y degradan al c-di-GMP están frecuentemente asociadas a diferentes dominios sensores, por lo que los niveles de este segundo mensajero son un reflejo de los cambios que tienen lugar en el ambiente intra y extracelular (primeros mensajeros) y que son detectados por los sistemas sensoriales bacterianos. La falta de información sobre los tipos de señales que controlan las enzimas que metabolizan el c-di-GMP limita la comprensión de las redes regulatorias de este nucleótido y dificulta el estudio de la adaptabilidad ambiental bacteriana. Por lo cual, la presente investigación se centra en el estudio de los dominios sensoriales capaces de detectar moléculas para modular la actividad de las enzimas y contribuir a una mejor comprensión de la regulación de c-di-GMP en *A. baldaniorum*. Para ello, se utilizaron programas bioinformáticos para conocer a mayor detalle la estructura genética y el producto proteico del gen *cdgG*. Este gen codifica para una proteína híbrida, probablemente implicada en el metabolismo del c-di-GMP, el producto génico está conformado por tres dominios característicos EAL-CBS-GGDEF, lo que sugiere su probable participación en la síntesis y degradación de c-di-GMP. Dado que los dominios catalíticos se encuentran conservados, nos centramos en analizar al dominio CBS que forman parte de la proteína CdgG y encontramos que la secuencia primaria del dominio CBS se encuentra poco conservada, mientras que la estructura terciaria del dominio se conserva. Esto se pudo observar gracias a la estructura cristalográfica que se encuentra depositada en el Protein Data Bank de la proteína DarB de *B. subtilis*, en donde, se ha

observado que al unir c-di-AMP inhibe la síntesis de la alarmona (p)ppGpp. La síntesis de la alarmona es parte de la respuesta astringente de células bacterianas bajo privación de aminoácidos o bajo condiciones de estrés.

## 1. Introducción

### ***Azospirillum baldaniorum***

*Azospirillum baldaniorum* Sp245 es una rizobacteria promotora del crecimiento vegetal (PGPR). Favorece el crecimiento de las plantas a través de diferentes mecanismos que incluyen la secreción de fitohormonas, la fijación biológica del nitrógeno, la solubilización del fosfato, entre otros (Spaepen et al., 2009). También benefician a la planta a través de efectos indirectos, asociados principalmente con la reducción del daño causado por patógenos, funcionando como agentes de control biológico al actuar directamente sobre el patógeno o inducir la resistencia sistémica en la planta (Pieterse et al.; 2014). Los efectos benéficos que aporta *Azospirillum* a la planta es un incremento en la longitud y número de raíces laterales, un mayor número de pelos radiculares así como en la respiración celular, entre otros (Pereyra et al., 2023).

*A. baldaniorum* Sp245 fue reclasificada (Ferreira et al., 2020) siendo considerada una cepa modelo, por la amplia variedad de estudios realizados en las últimas décadas en la promoción del crecimiento de plantas ya que se ha utilizado como inoculante en cultivos agrícolas en varios países (Cassán et al., 2020). Esta cepa fue aislada por primera vez en Rio Grande do Sul Brasil a partir de raíces de trigo (*Triticum aestivum* L.) superficialmente desinfectadas, dada su capacidad para colonizar las raíces de plantas endófitamente. Las células de la cepa Sp245 son bacilos Gramnegativos, ligeramente curvados o espirales con una longitud de 1.6-2.1  $\mu\text{m}$  y de ancho 0.5-0.7  $\mu\text{m}$ . Son altamente móviles por la presencia de un largo flagelo polar y varios flagelos laterales cortos en longitud. La cepa Sp245 es capaz de crecer entre 20 y 38 °C con una temperatura óptima de 30°C dependiente del medio de cultivo (Ferreira et al., 2020). Crece en una amplia variedad de fuentes de carbono como: D-fructosa, D-arabinol, L-fucosa, D-galactosa, gentiobiosa,  $\alpha$ -D-glucosa, melibiosa, D-psicosa, L-ramnosa, ácido fórmico, ácido D-galacturónico-lactona, ácido D-galacturónico, ácido  $\alpha$ -cetovalérico, glucuronamida y glucosa-6-fosfato (Ferreira et al., 2020).

### **Colonización *Azospirillum*-planta**

Para que *A. baldaniorum* pueda promover el crecimiento de la planta, tiene que existir una unión estable y duradera entre la bacteria y la planta. En la primera etapa de la colonización, la planta secreta exudados radiculares que contienen fuentes de carbono y nitrógeno como: azúcares, aminoácidos, compuestos fenólicos, metabolitos secundarios, polisacáridos, proteínas, entre otros (Cruz-Pérez et al., 2021). Estos exudados le permiten a la planta tener una interacción con la comunidad microbiana de la rizosfera, atrayéndolos hacia ésta. En este caso, *A. baldaniorum* es

quien detecta estos exudados, en donde la bacteria exhibe un comportamiento quimiotáctico que le permite navegar hacia las raíces de la planta (Wadhams y Armitage, 2004).

Con base a la señal que detecte la bacteria, se darán distintos tipos de taxis (migración hacia) como, por ejemplo: aerotaxis (gradientes de oxígeno), fototaxis (gradiente de luz), taxis hacia aceptores de electrones alternativos o quimiotaxis (sustancias químicas). *A. baldaniorum* es atraída principalmente por ácidos orgánicos (malato y succinato), azúcares (fructosa, ribosa, arabinosa y galactosa), aminoácidos (glutamato, aspartato, alanina, glutamina y asparagina) y aceptores de electrones alternativos para la respiración celular (nitrato, nitrito y dimetilsulfóxido) (Sierra-Cacho et al., 2021).

Una vez que *Azospirillum* encuentra la planta, prosigue a colonizar las raíces en dos fases denominadas: adsorción y anclaje. En la adsorción la bacteria se une a la raíz de forma reversible, proceso que toma de 30 minutos hasta 8 horas. Esta unión reversible se debe a la interacción de adhesinas y al flagelo polar de la bacteria, los cuales interaccionan con receptores específicos que se encuentran en la superficie de la raíz de la planta (Levanony, 1991; Michiels et al., 1989) Posteriormente, el anclaje se da por la presencia de exopolisacáridos (Camelo et al., 2011) que están implicados en la unión célula-célula y con la superficie en donde se adhiere la bacteria. Una vez realizado el anclaje, la bacteria prosigue a formar una biopelícula, que le conferirá resistencia ante las condiciones del ambiente y microorganismos competidores. De esta forma asegura su estancia y le permite generar beneficios en el crecimiento de la planta hospedera. Dentro de estas etapas, la quimiotaxis, la adherencia y la formación de biopelícula, son eventos cruciales en la interacción *Azospirillum*-planta (Sierra-Cacho et al., 2021).

### **Quimiotaxis**

La quimiotaxis bacteriana es un comportamiento fisiológico que permite a las bacterias navegar por su entorno y llegar a lugares que les son benéficos. Este movimiento se basa en la detección de atrayentes, sustancias químicas que las atraen, y repelentes, que las alejan de ambientes hostiles (Wadhams and Artimage., 2004). La quimiotaxis presenta un papel crucial en la vida de las bacterias, puesto que les permite localizar nutrientes esenciales para su crecimiento y supervivencia. Por ejemplo, las raíces de las plantas liberan exudados ricos en carbono y energía que atraen a las bacterias del suelo. La quimiotaxis también permite a las bacterias escapar de sustancias tóxicas o

condiciones desfavorables. La capacidad de las bacterias para moverse hacia zonas favorables les permite colonizar nuevos nichos ecológicos y expandir su población (Sierra-Cacho et al., 2021).

Las vías de señalización de la quimiotaxis han sido ampliamente estudiadas en bacterias como; *Escherichia coli* y *Salmonella entérica* (Hadjidemetriou et al.,2022; Stecher et al., 2004). Se ha mostrado que el movimiento comienza por los flagelos presentes en las bacterias, mismos que girarán a favor o en contra de las manillas del reloj. Cuando se da un cambio en la dirección en sentido contrario a las agujas del reloj o uno a favor, provoca el cambio de dirección del movimiento por parte de la bacteria. Este cambio de dirección del flagelo es regulado por una cascada de transducción de señales que involucra a un sistema de doble componente, y a un receptor de membrana (MCP, proteína aceptora de grupos metilo) el cual se encuentra en grupos localizados en los polos de la bacteria, una proteína cinasa sensora (CheA), una proteína adaptadora (CheW), un regulador de respuesta (CheY) y dos proteínas adaptadoras CheB y CheR. La presencia de repelentes o ausencia de atrayentes resulta en el incremento en la actividad de CheA, la cual incrementa los niveles de CheY fosforilada (CheY-P). Esta última interactúa con las proteínas del motor flagelar FliM y FliN para generar el cambio de dirección del rotor. Esto se detiene mediante la acción de CheZ que actúa como fosfatasa para remover el fosfato de CheY. CheA por su lado fosforila la metiltransferasa CheB, la cual removerá los grupos metilo de la MCP. La acción de CheB es contrarrestada por la metiltransferasa CheR. La proteína MCP con varios grupos metilo puede activar a CheA, la cual activara a CheB. De esta forma CheB y CheR pueden generar adaptaciones permitiendo a la célula responder a cambios químicos del ambiente (Hazelbauer et al., 2008; Wadhams et al., 2004) (**figura 1**).

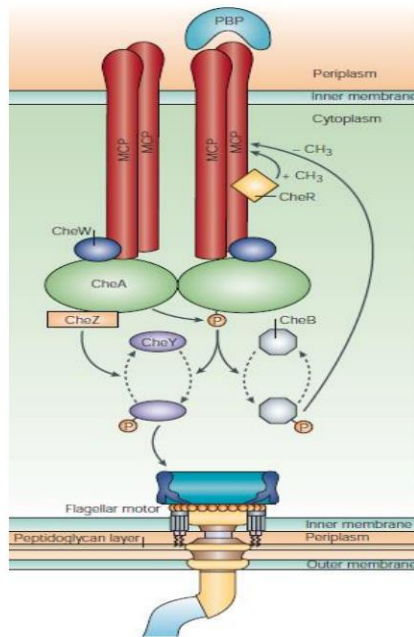


Figura 1. Diagrama del sistema de quimiotaxis en *E. coli* (Wadhams et al; 2004).

En el caso de *Azospirillum* se ha demostrado que mutantes deficientes en quimiotaxis presentan una colonización radicular disminuida en comparación con la cepa silvestre (Pereyra et al., 2023). La cepa *A. brasilense* Sp7 presenta cuatro vías de señalización, de las cuales únicamente las vías Che1 y Che4 juegan un papel esencial en la quimiotaxis controlando, respectivamente, la velocidad del flagelo y los cambios en la dirección de rotación del motor flagelar. Además, se ha demostrado la importancia de la vía Che4 para la colonización de raíces (Ganusova et al., 2021).

El empleo de diferentes técnicas de microscopía ha permitido la visualización de la colonización de la rizosfera y raíces, gracias a lo cual ha permitido establecer que *A. brasilense* coloniza preferentemente los pelos radiculares y las zonas de elongación radicular en plantas de trigo. La respuesta quimioatrayente se observó segundos después de exponer a la rizobacteria a raíces de trigo; respuesta que no se produjo en cepas mutantes en quimiotaxis (O' Neal et al., 2020). En cambio, se observó una respuesta quimiorrepelente en la zona del ápice radicular de trigo. Las cepas secuenciadas de *A. brasilense* presentan aproximadamente 50 proteínas quimiorreceptoras y dos de ellas, Tlp1 y Aer, son importantes para la colonización de raíces de trigo (O'Neal et al., 2019; O'Neal et al., 2020). El quimiorreceptor Aer juega un papel en la quimiotaxis energética de *A. brasilense* y promueve respuestas de quimiorrepelencia frente a especies reactivas de oxígeno. Por otro lado, Tlp1 es responsable de promover quimiotaxis atrayente hacia ácidos orgánicos como el malato, piruvato y succinato, además de respuestas quimiorrepelentes frente a especies reactivas

de oxígeno. Así, una mutante en *tlp1* es deficiente en la colonización de pelos radiculares y zonas de elongación radicular, pero coloniza eficientemente ápices radiculares (O'Neal et al., 2019).

### **Biopelículas**

Para establecer una interacción bacteria-planta exitosa y prolongada, *A. baldaniorum* debe formar biopelículas, es decir, comunidades bacterianas incrustadas en una matriz de fabricación propia formada por sustancias poliméricas extracelulares (SPE) que proporcionan condiciones favorables para la supervivencia. Las SPE están constituidas por exopolisacáridos, proteínas, ADN extracelular (ADN) y lípidos.

Este estilo de vida le proporciona a la comunidad bacteriana protección contra condiciones adversas, como la presencia de antibióticos, desecación, radiación, ataque de bacteriófagos, etcétera. La formación de la biopelícula consta de 5 etapas: 1) La unión inicial donde participan proteínas de superficie (adhesinas), flagelos, fimbrias de tipo I, IV y Tad, así como también, polisacáridos capsulares, permitiendo a las células unirse al sustrato. 2) Formación de una monocapa de células; una vez que la bacteria se ha adherido a la superficie, comienza a dividirse y las células hijas se extienden alrededor del sitio de unión, formando una microcolonia junto con la secreción de SPE, 3) Migración para formar colonias en multicapa, 4) Producción de la matriz extracelular y síntesis de las SPE para generar una estructura dinámica con canales en su interior que permiten el flujo de nutrientes, 5) Maduración de la biopelícula y dispersión de las células, rompiéndose la matriz conformada por las SPE para la formación de una nueva biopelícula y, de esta forma, repetir el ciclo desde el estado de vida planctónico (Wei et al., 2013).

Los factores de competencia y diversidad de condiciones exigen mecanismos de señalización y adaptación eficientes para garantizar esta supervivencia. Estos sistemas suelen componerse de señales que son acopladas a mecanismos complejos de transducción de señales donde se emplean moléculas denominadas segundos mensajeros, cuya finalidad es amplificar y transmitir la información a los sensores intracelulares y, a distintos efectores cuya actividad es regulada por la concentración de estas moléculas. Existen una gama amplia de segundos mensajeros: alarmona guanosín tetra/pentafosfato, (p)ppGpp, diguanosín monofosfato cíclico (c-di-GMP) y diadenosín monofosfato cíclico (c-di-AMP), entre otros (Jenal et al., 2017).

## c-di-GMP y formación de biopelículas

El segundo mensajero di-(3'-5') guanosín monofosfato cíclico ó c-di-GMP es responsable de controlar múltiples funciones celulares, las cuales son importantes para la supervivencia bacteriana debido a que regula una amplia variedad de procesos bacterianos, entre los que destacan la virulencia, resistencia al estrés oxidativo, progresión del ciclo celular, la internalización del microorganismo hacia su hospedero, el ensamblaje de sistemas de secreción, la reparación del ADN, movilidad y formación de biopelículas (Cruz-Pérez et al., 2021).

La regulación dependiente de la señalización del c-di-GMP tiene lugar a nivel transcripcional, postranscripcional y postraduccional (Cruz-Pérez et al., 2021). Por ejemplo, al unirse al factor transcripcional VpsT en *Vibrio cholerae*, potencia la formación de biopelícula. A niveles bajos de c-di-GMP favorecen un estado de movilidad en la bacteria, mientras que cuando incrementan los niveles, se favorece la expresión de exopolisacáridos y adhesinas, los que desembocan en la formación de biopelícula (Jenal et al., 2017) (Figura 2).

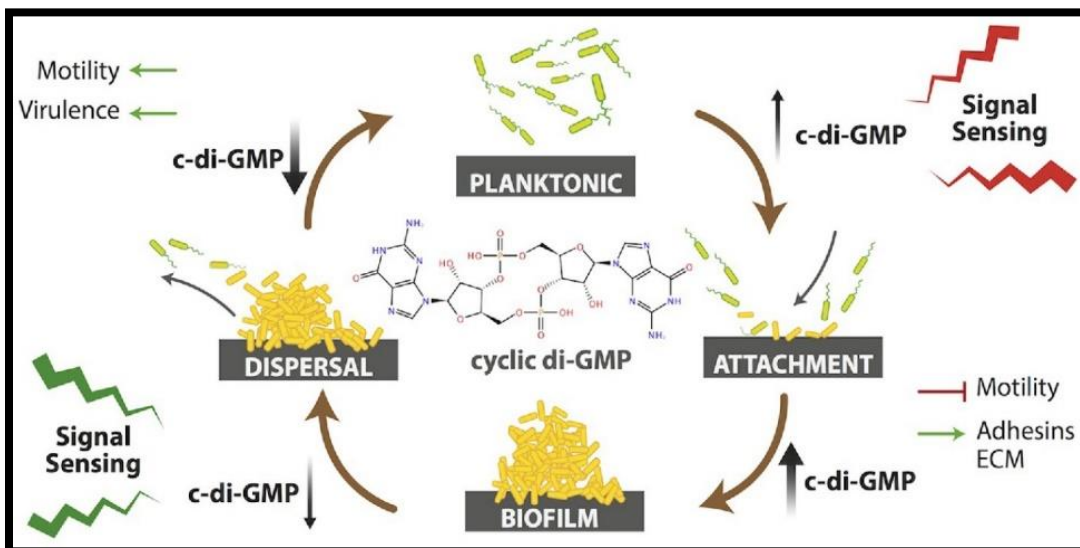


Figura 2. c-di-GMP. Este segundo mensajero tiene la capacidad regular distintas actividades fisiológicas bacterianas mediante su interacción con proteínas efectoras. (Rinaldo., 2018).

Si bien los componentes corrientes abajo de las vías de señalización dependientes de c-di-GMP han sido ampliamente estudiados, la comprensión de las señales ascendentes que modulan los niveles de c-di-GMP sigue siendo un área de investigación activa. La falta de información completa sobre los tipos de señales que controlan las CME limita nuestra comprensión de las redes regulatorias de c-di-GMP y dificulta el estudio de la adaptabilidad ambiental bacteriana. En vista de la diversidad de las capacidades sensoriales de las CME de señalización y la escasez de revisiones exhaustivas sobre este tema, es necesario un análisis en profundidad de las vías de señalización ascendente que controlan los niveles bacterianos de c-di-GMP. Este análisis debe centrarse en los dominios sensoriales capaces de detectar moléculas para modular la actividad de las CME y contribuir a una mejor comprensión de la regulación de c-di-GMP en las bacterias (Yu et al., 2023).

La formación de la biopelícula es un proceso altamente regulado por el segundo mensajero (3'5') guanosín-monofosfato-cíclico (c-di-GMP) del cual depende en gran medida la transición del estilo de vida móvil (planctónica) a sésil (biopelícula) de la bacteria. Estos niveles del c-di-GMP son regulados por enzimas con actividad de DGC y/o PDE, que participan en su síntesis y degradación (Römling U et al, 2013; Jenal U et al., 2017)

### **Metabolismo del c-di-GMP**

Los niveles intracelulares del c-di-GMP están influenciadas por su síntesis y degradación. El c-di-GMP se sintetiza a partir de 2 moléculas de GTP por la acción de las DGC y es degradado por PDE específicas de c-di-GMP (Hengge, 2009).

### **Diguanilato ciclasas y síntesis del c-di-GMP**

La reacción de síntesis es catalizada por los dominios GGDEF de las Diguanilato ciclasas (DGC). Las proteínas con actividad DGC usualmente actúan como homodímeros, donde cada monómero contribuye con la unión de una molécula GTP para posteriormente unirlos mediante un enlace fosfodiéster utilizando  $Mg^{2+}$  o  $Mn^{2+}$  como cofactores (Römling U., 2013). Además de los dominios GGDEF, estas proteínas (**Figura 3**) contienen varios motivos proteicos importantes que contribuyen a la actividad catalítica y a su propia regulación. En primer lugar, se encuentra el sitio activo (sitio A), que realiza la actividad catalítica para sintetizar el c-di-GMP y se caracteriza por poseer el motivo de aminoácidos conservados GG (D/E) EF (Gly-Gly-Asp/Glu-Glu-Phe) donde las dos glicinas intervienen en la unión de la molécula GTP, mientras que el tercer y cuarto aminoácido están involucrados en la catálisis y coordinación del metal, respectivamente. Por otro lado, se encuentra

el sitio de inhibición, sitio I: Este sitio I consta de un motivo conservado RXXD localizado 5 aminoácidos río arriba del sitio A, el cual puede unir al c-di-GMP, inhibiendo de manera alostérica su síntesis (Römling, 2013).

Amplios estudios bioinformáticos y experimentales demuestran que las proteínas con dominio GGDEF poseen en su extremo N-terminal dominios de señalización citoplasmáticos, transmembranales o periplásmicos como son: los PAS, GAF, REC, GCS, MHYT o CACHEC (Ramírez-Mata et al., 2018). Las DGCs funcionan en forma de dímero, que en ciertos casos pueden ser activadas de manera alostérica alostéricamente por acción de alguno de estos dominios de señalización. Como ocurre con la proteína PleD de *Caulobacter crescentus*, donde la formación del dímero y activación del dominio GGDEF se lleva a cabo por la fosforilación de su dominio REC que posee en el extremo N-terminal de la proteína (Chan et al., 2004).

A la fecha, también se han caracterizado DGCs con dominios GGDEF que no tienen función enzimática y que se les denomina dominios degenerados ya que no poseen conservados los motivos de aminoácidos indispensables para que se lleve a cabo la actividad enzimática. Sin embargo, estos dominios funcionan como receptores y/o efectores de c-di-GMP mediante la interacción de este en el sitio-I (Römling, 2013).

### **Fosfodiesterasas y degradación del c-di-GMP**

La hidrolisis del c-di-GMP se lleva a cabo por enzimas que se les conoce como fosfodiesterasas (PDEs) y existen dos tipos de éstas, unas que se caracterizan por poseer un motivo de aminoácidos conservado EAL y otras que poseen el motivo de aminoácidos conservado HD-GYP (Hengge, 2009).

El producto de la hidrolisis del c-di-GMP será diferente dependiendo el tipo de PDE que esté realizando la catálisis, siendo una molécula lineal pGpG cuando el c-di-GMP sea hidrolizado por PDE con el dominio EAL o dos moléculas de GMP cuando sea hidrolizado por enzimas con el dominio HD-GYP (Römling, 2013; Rao, 2008).

El dominio EAL consta aproximadamente de 250 aminoácidos. Este dominio requiere de cationes divalentes como el  $Mg^{2+}$  o  $Mn^{2+}$  para que se dé su actividad enzimática, mientras que el  $Ca^{2+}$  y  $Zn^{2+}$  la inhiben (Römling, 2013).

Además del motivo EAL conservado en las PDEs, hay otros aminoácidos conservados que intervienen en la actividad enzimática. Uno de estos es el motivo de aminoácidos conservado DDFGTGYS conocido como "lazo 6" (loop 6 en inglés) que recibe este nombre por la conformación

que adquiere en la proteína. En este motivo los dos ácidos aspárticos al igual que el ácido glutámico del motivo EAL funcionan como receptores de  $Mg^{2+}$  o  $Mn^{2+}$  para que se lleve a cabo el ataque nucleofílico en la hidrólisis del c-di-GMP. Por otra parte, la tirosina y serina conservadas, cuatro aminoácidos río abajo de los dos ácidos aspárticos del “lazo 6” intervienen en la formación del dímero entre dos dominios EAL, interaccionando una molécula de c-di-GMP en cada monómero, aumentando la estabilidad y actividad enzimática de la proteína. Sin embargo, el dominio EAL puede ser activo en forma de monómero también (Römling, 2013).

Los dominios EAL se pueden clasificar en tres tipos: catalíticamente activos (clase 1), catalíticamente activos con algunos residuos de aminoácidos degenerados (clase 2), y los catalíticamente inactivos (clase 3). La mayoría de las proteínas con dominio EAL estudiadas son de clase 1, las cuales se caracterizan por poseer dominios de señalización en el extremo N-terminal, poseen actividad catalítica aun en estado no activado y poseen conservado el lazo 6. Los de clase 2 poseen actividad catalítica, aunque algunos de sus aminoácidos no estén conservados como el motivo de dimerización en el lazo 6; esta clase de proteínas no han sido ampliamente descritas. Las proteínas de clase 3 no poseen actividad enzimática ya que no presentan los aminoácidos conservados necesarios para que se lleve a cabo la hidrólisis del c-di-GMP. Sin embargo, tienen la capacidad de interaccionar con el c-di-GMP, funcionando como receptores y/o efectores (Römling, 2013).

Dentro de la clase 3 de fosfodiesterasas, se encuentra un tipo de estas proteínas que no funcionan como receptores de c-di-GMP. Estas proteínas en cambio funcionan en la interacción proteína-proteína, como ocurre con la proteína YdiV de *Salmonella* que interaccionan con el regulador maestro de flagelo FlhDC en complejo con la proteasa ClpXP para que se lleve a cabo la degradación del regulador (Takaya et al., 2012) mostrando la gran variedad de funciones que pueden desarrollar este tipo de dominios proteicos y su implicación en la transducción de señales.

### **Proteínas con dominio GGDEF-EAL (híbridas)**

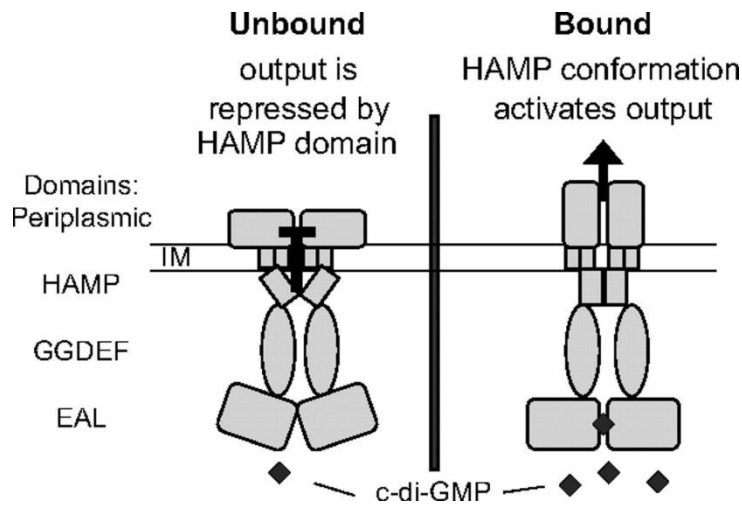
Así como se encuentran proteínas con dominio único GGDEF o EAL, a menudo los genomas de bacterias codifican para proteínas con ambos dominios GGDEF-EAL formando parte de la misma proteína. Esto resulta en una regulación aún más compleja de los niveles de c-di-GMP en la célula, además de que la mayoría de estas proteínas también poseen en el extremo N-terminal, dominios de señalización citoplasmáticos o peri plasmáticos como PAS o CACHE (Chao et al., 2013).

A la fecha se ha reportado que existen 3 escenarios de la función que poseen estas proteínas híbridas, ya que pueden sintetizar como degradar al c-di-GMP.

El primer escenario es cuando ambos dominios son enzimáticamente activos pero la función de cada uno de ellos será activada dependiendo de los estímulos ambientales que reciba la célula, como ocurre en la proteína DcpG de *Paenibacillus dendritiformis* donde esta proteína que posee los dominios GCS-GGDEF-EAL presenta ambas actividades (Patterson et al., 2021). La proteína DcpG contiene un dominio sensor globina, denominado GCS, a través del cual se han podido identificar las moléculas que habilitan a cada una de las actividades de DGC y PDE. El dominio GCS está constituido por un grupo hemo que tiene la capacidad de unir gases; el dominio GGDEF se activa cuando une a la molécula Fe(II)-NO (óxido nítrico), mientras que el dominio EAL se habilita en presencia de Fe(II)-O<sub>2</sub>. Estos datos demuestran el mecanismo por el cual un solo dominio regulatorio puede controlar selectivamente a dos dominios con actividades enzimáticas opuestas (Patterson et al., 2021).

El segundo escenario es cuando se presenta una regulación alostérica de uno de los dominios por la interacción con un ligando. En este escenario uno de los dominios es enzimáticamente activo y el otro dominio es inactivo debido a que está degenerado. Sin embargo, este dominio funciona como regulador de la actividad del dominio enzimáticamente activo. Tal es el caso del dominio GGDEF de la proteína PdeA de *Caulobacter crescentus* la cual une GTP al dominio EAL y este cumpla su función de PDE (Christen et al., 2006). Se conoce que PdeB, de *Shewanella oneidensis* MR-1 es una proteína con dominio GGDEF degenerado EAL, conservado la cual funciona solo como PDE (Chao et al., 2013).

El tercer escenario es cuando ambos dominios están degenerados y, por lo tanto, son enzimáticamente inactivos. Sin embargo, estas proteínas funcionan como receptoras o efectoras del c-di-GMP continuando con la cascada de señalización y afectando a un sitio o molécula blanco en específico mediante la interacción proteína-proteína. Este mecanismo ocurre con la proteína de membrana LapD de *Pseudomonas fluorescens* que posee ambos dominios GGDEF/EAL degenerados y carece de ambas actividades enzimáticas, sin embargo, tiene la capacidad de unir c-di-GMP mediante su dominio EAL provocando su dimerización y mediante un dominio accesorio HAMP que posee en el dominio N-Terminal, conduce la señal hacia el dominio periplásmico secuestrando a la proteasa denominada LapG; de esta manera, se evita la degradación de la adhesina llamada LapA, que participa en la formación de biopelícula como se muestra en la **figura 3** (Newell et al., 2009).



**Figura 3.** Un modelo para señalización de adentro hacia afuera a través de LapD. (Izquierda) En ausencia de unión de c-di-GMP, la producción periplásmica se reprime a través del dominio HAMP. (Derecha) Cuando c-di-GMP está unido por LapD, el dominio HAMP asume una conformación que activa la salida. LapD se representa como un dímero, porque se sabe que otras proteínas del dominio HAMP son dímeros (Nawell et al., 2009).

Algunos reportes de diversos genomas bacterianos hechos *in silico*, han demostrado la presencia de un amplio número de genes que codifican para proteínas híbridas. Esta redundancia proteica confiere al microorganismo la capacidad de responder de forma rápida y efectiva hacia los diversos estímulos mediante la interacción del c-di-GMP con los efectores (Rybtke et al., 2017; Nicastro et al., 2020; Pérez-Mendoza et al., 2014).

Dado el basto número de proteínas DGC o PDE que se codifican en un mismo genoma, se especula que estas enzimas forman una compleja red de señalización, que inicia con la detección de estímulos, culminando con la generación de alguna respuesta.

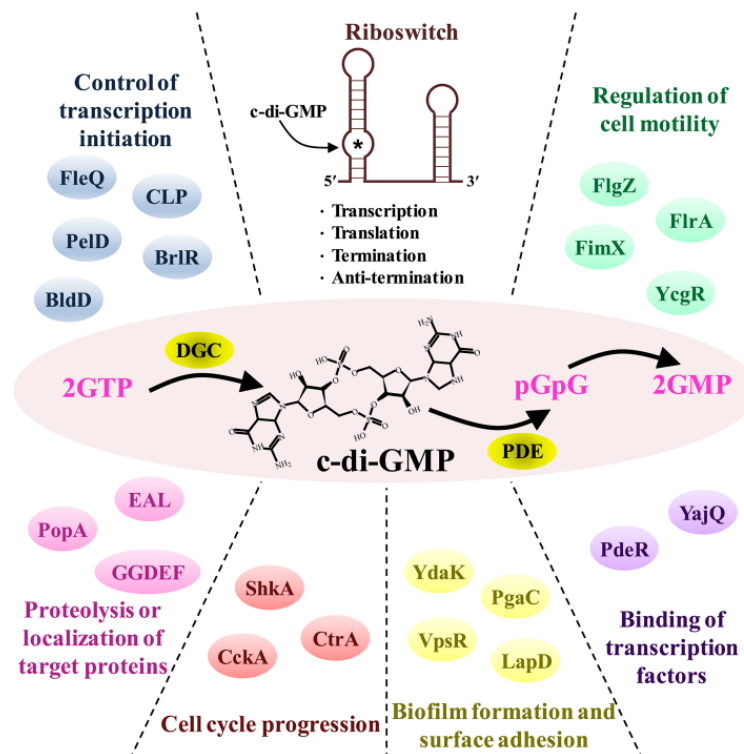
### Receptores y/o efectores del c-di-GMP

Se ha reportado que el c-di-GMP interacciona con efectores específicos, como son proteínas que contienen dominios PilZ, proteínas con dominios GGDEF y EAL degenerados, factores transcripcionales y riboswitches (Jeong et al., 2023). A la fecha se han reportado una gran variedad de proteínas que contienen dominios PilZ, las cuales tienen la capacidad de formar un complejo con el c-di-GMP mediante la interacción del nucleótido con los motivos RxxR y (D/N)x(S/A)xxG presentes en estos dominios (Jeong et al., 2023). Entre los efectores con el dominio PilZ están Alg44 en *P. aeruginosa* y YcgR en *E. coli*. Alg44 está implicado en la síntesis de alginato mientras que YcgR

regula la movilidad bacteriana tras unirse al c-di-GMP. Este complejo, YcgR/c-di-GMP interactúa con el motor flagelar frenando al flagelo para detener la motilidad bacteriana (Paul et al., 2019). Mientras que las proteínas con dominios GGDEF y EAL degenerados pueden unir c-di-GMP a pesar de no presentar actividad catalítica, como previamente se describió (Ramírez-Mata et al., 2019).

En el caso de los factores transcripcionales dependientes del c-di-GMP, se ha reportado que estos pueden regular la expresión de genes implicados en la síntesis del flagelo, así como de exopolisacáridos y factores de virulencia. Por ejemplo, VpsT y FleQ (figura 4) son reguladores transcripcionales de *V. cholerae* y *P. aeruginosa*, respectivamente, los cuales están implicados en la movilidad y formación de biopelícula (Povolotsky y Hengge., 2016; Kun et al; 2020).

Los riboswitches son secuencias de RNA no codificantes, con estructuras secundarias específicas ubicadas en los extremos 5' de RNAs mensajeros de genes que codifican para proteínas implicadas en el metabolismo del c-di-GMP o reguladas por éste. La unión del nucleótido a estas estructuras conlleva a una modificación en la conformación tridimensional del riboswitch, lo que provoca cambios en la expresión de genes que se encuentren río abajo de estas estructuras (Povolotsky y Hengge., 2016; Kunz et al., 2020).



**Figura 4. Receptores del c-di-GMP.** La interacción del de c-di-GMP con diferentes proteínas receptoras que controlan diversas funciones biológicas en bacterias (Jeong et al., 2023).

## 2. Antecedentes

Todos los procesos anteriormente descritos, se han estudiado ampliamente en bacterias patógenas de humanos y plantas, y muy pocos estudios existen en bacterias promotoras del crecimiento vegetal, tal es el caso de *A. baldaniorum*, quien transita del estado planctónico a sésil en su interacción con plantas de interés agronómico. En este sentido, la transducción de señales mediadas por el c-di-GMP es poco entendida. Actualmente, se sabe que una PDE (denominada ChsA de *A. brasilense* Sp7) participa en quimiotaxis, cuya mutación afectó la formación de biopelícula, disminuye los niveles intracelulares de c-di-GMP, en respuesta a la concentración de quimioattractantes y oxígeno, modulando de manera transitoria la movilidad, a través de un receptor proteico de quimiotaxis y aerotaxis denominado Tlp1 (Russell et al., 2013). Por otra parte, también se ha caracterizado funcionalmente a una DGC (DgcA) de dominio único en *A. brasilense* Sp7, encontrando que la mutante isogénica disminuye la producción de biopelícula y exopolisacáridos con respecto a la cepa silvestre y a través de ensayos *in vitro* se determinó que la proteína ostenta actividad de DGC (Ramírez-Mata et al., 2016).

Los dos genes de *A. brasilense* Sp7 identificados en nuestro laboratorio fueron empleados para buscar en el genoma totalmente secuenciado de *A. baldaniorum* Sp245, genes que presentaran homología con genes cuyos productos de traducción presuntamente contienen los dominios GGDEF y EAL. En este análisis bioinformático se utilizaron la base de datos Uniprot y MiST2 (microbial signal transduction database), además de programas de análisis de dominios conservados Prosite, Pfam database y SMART, con los cuales se obtuvieron 35 proteínas con dominios GGDEF y EAL; 20 DGCs de dominio único, perfectamente conservadas, 5 PDEs de dominio único; 3 conservadas y 2 degeneradas y 10 proteínas híbridas, de las cuales 7 son degeneradas; 6 en el dominio GGDEF y una en el dominio EAL. Además, algunas de éstas incluyen en su estructura dominios de señalización (REC, PAS, MHYT, CACHE, CHASE, HAMP, etc.) (Ramírez-Mata et al., 2018). Por otra parte, analizó el contexto genético en donde se ubican los genes que las codifican. Varios forman presuntos operones con cinasas histidínicas (proteínas que constituyen los sistemas de doble componente en bacterias). Algunos genes se localizan en cromosomas. Datos recientes de nuestro grupo de trabajo, muestran funcionalmente 3 de estos genes; tal es el caso de la proteína CdgC la cual es una DGC funcional que está conformada por un dominio REC y un dominio GGDEF, la cual está involucrada

en la formación de biopelícula, producción de exopolisacáridos en condiciones de fijación de nitrógeno y en la colonización a plantas de trigo, dado que la mutante disminuye la capacidad de *Azospirillum* en colonizar los tejidos internos de la raíz en comparación con la cepa silvestre, mientras que la cepa complementada restauró el fenotipo (Sierra-Cacho et al., 2021).

En cuanto a los genes que codifican proteínas híbridas se ha caracterizado a la proteína CdgD la cual está conformada por los dominios CHASE-PAS-GGDEF-EAL. A través de ensayos de PCR en tiempo real se logró determinar que el gen *cdgD* se expresa en mayor proporción cuando el cultivo de *Azospirillum* crece con  $\text{KNO}_3$  como fuente de nitrógeno y en menor proporción en  $\text{NH}_4\text{Cl}$ . La mutante del gen *cdgD* disminuye su capacidad de formar biopelícula y colonización en plantas de trigo ya que bajo estas condiciones funciona como DGC en comparación con la cepa silvestre y restaurando el fenotipo con la cepa complementada. En cuanto a su actividad de PDE, a la fecha se están realizando experimentos que permitan dilucidar su función. Sin embargo, datos preliminares y bioinformáticos predicen que también podría degradar al c-di-GMP (**Figura 5**) (Cruz-Pérez et al., 2021).

Finalmente, se ha demostrado que la proteína CdgB es una proteína híbrida con ambas actividades de DGC y PDE. Esta proteína contiene los dominios MHYT-PAS-SGDEF-EAL, a través de análisis fenotípicos como son: la formación de biopelícula, producción de exopolisacáridos y movilidad; se logró determinar que la proteína es bifuncional. A través de fusiones traduccionales se determinó que la proteína se ancla a la membrana citoplasmática mediante su dominio MHYT y potencialmente se localiza en los polos de la célula, sugiriendo que esta proteína participaría en la generación de pozos subcelulares de c-di-GMP (Viruega-Góngora et al., 2022).

Estos análisis confirman la expresión e importancia que tienen las proteínas híbridas (proteínas con dominios GGDEF y EAL) en los genomas bacterianos y que probablemente los dominios sensores participan en la modulación de estas actividades opuestas. Tales hallazgos sugieren que estas proteínas deben tener un papel importante en la fisiología, metabolismo y adaptación de *Azospirillum* al ambiente.

AZOBR\_30001 CdgG

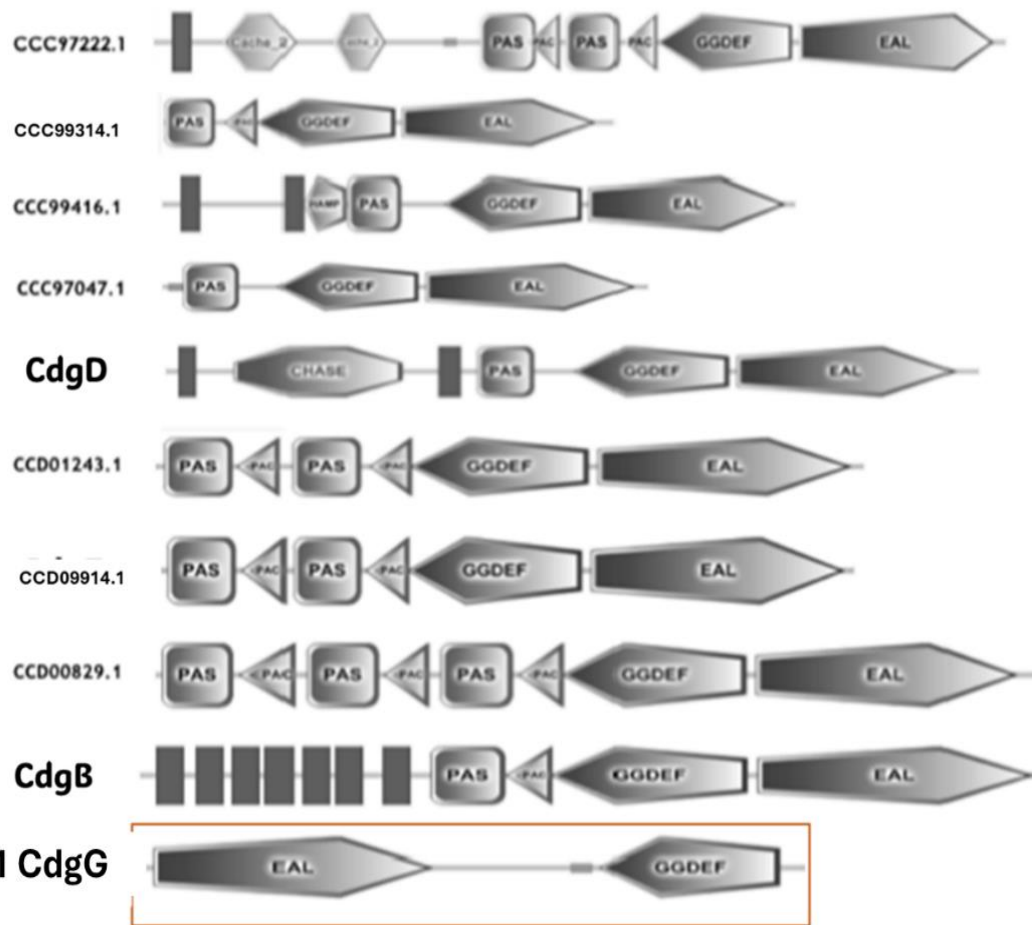


Figura 5. Diagrama de las proteínas híbridas implicadas en el metabolismo del c-di-GMP de *A. baldaniorum* Sp245.

### 3. Justificación

Las bacterias del género *Azospirillum* son una de las PGPRs más estudiadas en muchos países debido a su aplicación como inoculante en cultivos agrícolas. Conforme a esto, es necesario avanzar en el conocimiento de los mecanismos por los cuales estas bacterias interactúan en la rizosfera para otorgar beneficios a las plantas. Para entender estos procesos en la interacción planta-bacteria, es necesario comprender el papel del c-di-GMP, el cual interviene en diversos mecanismos y actividades bacterianas como la movilidad, formación de biopelícula, transición del estado de vida móvil a sésil, y en la colonización de raíces de las plantas, entre otros aspectos.

Consecuentemente, este trabajo se propone la clonación del gen **AZOBR\_30001 (cdgG)**, que posee un dominio EAL en el extremo N-terminal, y un dominio GGDEEF en el extremo c-terminal, fundamentales para la síntesis y degradación del c-di-GMP. Así mismo, se muestra los resultados de los estudios bioinformáticos realizados al gen y la proteína codificada denominada como *CdgG*.

### 4. Objetivos

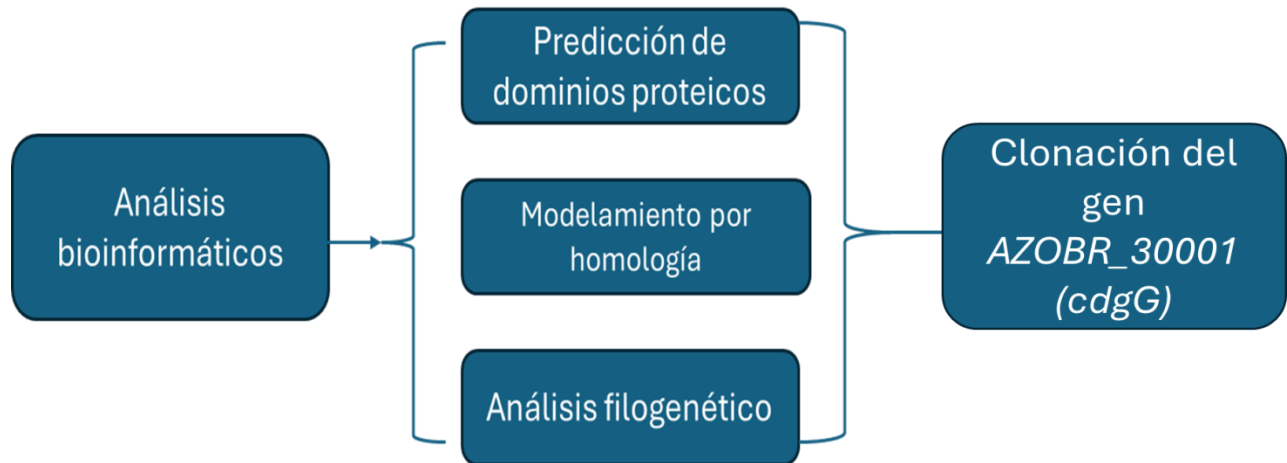
#### 4.1 objetivo general

Estudiar bioinformáticamente y clonar el gen *AZOBR\_30001 (cdgG)* de *A. baldaniorum* Sp245.

#### 4.2 objetivos particulares

- Analizar bioinformáticamente el gen *AZOBR\_30001* (LOCUS WP\_041811385), la región promotora y la proteína codificada *CdgG* de *A. baldaniorum* Sp245.
- Clonar el gen *AZOBR\_30001*.

## 5. Esquema de trabajo



## 6. Materiales y métodos

### 6.1 Material biológico

Cepa	Características	Referencia
<i>Escherichia coli</i> DH5α	<i>fhA2Δ(argFA-lacZ)</i> U169 <i>phoA</i> <i>glnV44Φ80Δ(lacZ)</i> M15 <i>gyrA</i> <i>recA thi-1 hsdR17</i>	Hanahnan,1983
<i>A.baldaniorum</i> Sp245	Cepa silvestre	Baldani et al., 1987

### 6.2 Vectores

Vectores	Características	Referencia
pGEM-T-Easy	<i>F1 ori, ori, Amp<sup>r</sup>, lacZα,</i> promoter: T7, <i>lac</i> , SP6	Promega

### 6.3 Iniciadores

INICIADORES GEN CDGG		LONGITUD	TM	%GC
delantero: <i>Bam</i> HI CDGG	5'- <i>ggatccatggcctcggtgccca</i> -3'	21 nucleotidos	64.5°	66.7%
reverso: <i>Not</i> I CDGG	5'- <i>ctcgagctaaccctccccgctcc</i> -3'	24 nucleotidos	66.7°	70.8%

### 6.4 Condiciones de cultivo

La cepa de *E. coli* fue sembrada en medio líquido Luria-Bertani (LB) e incubadas a 37°C a 150 rpm por 18 horas. *E. coli* DH5a fue sembrada sin antibiótico.

La cepa silvestre de *A. brasilense* Sp245 fue cultivada en medio K-malato a 30°C en agitación a 150 rpm.

### 6.5 Extracción y manipulación del ADN genómico

Se sembraron en tubos con 5 ml de medio LB\* con 1 o 2 colonias de un cultivo proveniente de placas con RC y se incubaron 18h a 30°C y 150 rpm. En tubo eppendorf (1.5 ml) se concentraron de 3 a 5 ml del cultivo a 10,000 rpm por 3 min; se decantó el medio y se obtuvo el paquete celular, el cual se resuspendió en 500 µl de buffer TE, 30 µl de SDS 20% y 3 µl de pronasa E (10 mg/ml) y se incubó 1h a 37°C. Se agregó 100 µl de NaCl 3M y 80 µl de una solución alcalina de CTAB al 10 % y se incubó a

65°C, 10 min. Posteriormente, se agregaron 500 µl de una mezcla de cloroformo-alcohol isoamílico (24:1) y se centrifugó 5 min a 12,000 rpm. Se transfirió la fase acuosa a otro tubo limpio y se repitió lo anterior, agregando 500 µl de cloroformo-alcohol isoamílico (24:1) y centrifugando 5 min a 12,000 rpm. Se volvió a transferir la fase acuosa a otro tubo eppendorf y se lavó nuevamente con cloroformo-alcohol isoamílico (24:1). Una vez limpio el DNA, se agregaron 1.5 volúmenes de Etanol 100 %, y se incubó a -80°C por 15 min. Se centrifugó a 10,000 rpm durante 15 min y se eliminó el sobrenadante por decantación, cuidando no perder el DNA. Obtenido el DNA, se lavó con 800 µl de etanol al 70 % centrifugando a 10,000 rpm durante 10 min, 2 veces, se eliminó el etanol y se evaporó al colocar el tubo eppendorf con la tapa abierta, en un termoblock a 37°C por 5-10 min. Quedando únicamente el DNA al fondo del tubo, se resuspendió en 30 µl de RNAsa (100 µg/ml) y se incubó a 37 °C por 30 min. Finalmente se realizó una electroforesis con 3-5 µl de muestra de DNA para corroborar la calidad de la extracción genómica y el resto de la muestra se almacenó a -20 °C.

#### **6.6 Extracción y manipulación de ADN plasmídico**

Con 1 o 2 colonias tomadas de un cultivo en RC, se inocularon 5 ml de LB\* y se incubó por 18h a 30°C en agitación constante (150 rpm), del cultivo, se tomaron de 3 a 5 ml y se centrifugó a 10,000 rpm por 3 min; se decantó el sobrenadante y se obtuvo el paquete celular en el fondo del tubo. Se resuspendió el paquete celular en 300 µl de solución de lisis alcalina I (Tris-HCl 10 mM (pH 8.0); EDTA 1 mM) y se agregaron 400 µl de solución de lisis alcalina II (NaOH 200 mM; SDS 1%), se mezcló bien y se colocó el tubo a -20°C 10 min. Se agregaron 300 µl de solución alcalina III (Acetato de potasio 3.0 M pH 5.5) fría, se mezcló y mantuvo a -20 °C por 15 min. Posteriormente se centrifugó a 12,000 rpm, 15 min, y el sobrenadante se transfirió a un nuevo tubo, se agregaron 200 µl de fenol y 200 µl de cloroformo:alcohol isoamílico (24:1), mezclando bien y centrifugando por 5 min a 10,000 rpm. El sobrenadante se colocó en otro tubo eppendorf y se agregó 400 µl de cloroformo:alcohol isoamílico (24:1), se mezcló por inversión y se centrifugó a 10,000 rpm durante 5 min; este procedimiento se repitió hasta que la interfase era clara. Posteriormente se separó el sobrenadante en un tubo eppendorf y se agregaron 0.6 volúmenes de isopropanol, se mezcló bien la solución, invirtiendo el tubo varias veces y se dejó precipitar por 10 min a temperatura ambiente. Transcurrido el tiempo, se centrifugó a 10,000 rpm durante 15 min y se recuperó el DNA. Se eliminó el sobrenadante y el DNA que permanece al fondo del tubo, se lavó, agregando 800 µl etanol al 70% y centrifugando 5 min a 10,000 rpm. Este paso se repitió una vez más. Se eliminó por decantación el etanol y, se evaporó al colocar el tubo con la tapa abierta, en un termoblock a 37 °C de 5 a 10 min. Dejando únicamente el DNA en el tubo, éste se resuspendió en 30 µl de RNAsa (100 µg/ml) y se

incubó 30 minutos a 37°C. Finalmente se realizó la electroforesis con 3-5 µl de muestra de DNA plasmídico, para comprobar la calidad de la extracción y el resto de muestra se almacenó a -20°C (Sambrook y Russell, 2001).

### **6.7 Amplificación del gen *cdgG***

Se llevó a cabo la amplificación del gen *cdgG* usando los iniciadores *Bam*HI CdgG Y *Not*I CdgG, Bajo las condiciones siguientes: DNA 2ng, dNTP's 4 mM, DMSO 1.5X, iniciador delantero 0.5 µM, iniciador reverso 0.5 µM, Taq Pol 0.5 U, Buffer 2.5 µl, MgCl<sub>2</sub> 1.2mM, H<sub>2</sub>O (7.3 µl).

### **6.8 Clonación de *cdgG* en vector pGEM-T-Easy**

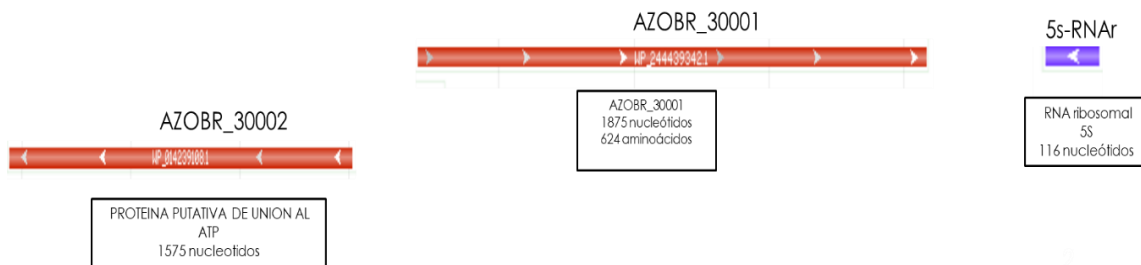
Determinadas las condiciones óptimas para la PCR, se realizó una amplificación empleando *Taq pol Platinum*, este amplificado se clonó en el plásmido pGEM-T-Easy.

## 7.Resultados

### 7.1 Análisis bioinformático

#### 7.1.1 Contexto genético del gen AZOBR\_30001 (*cdgG*)

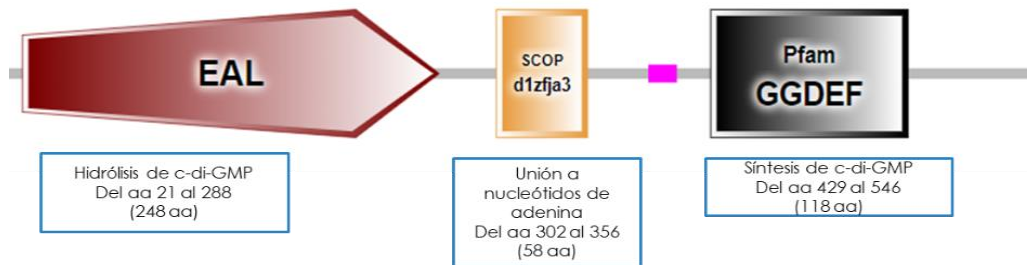
El WP\_041811385 al que a partir de ahora denominaremos *cdgG*, se encuentra codificado en el cromosoma de *A. baldaniorum* Sp245; posee una longitud de 1877 nucleótidos, y codifica para una presunta proteína híbrida diguanilato ciclasa/fosfodiesterasa con una longitud de 624 aminoácidos. En 2021 hubo una actualización del genoma de *A. baldaniorum* Sp245 en el NCBI donde renombraron al gen AZOBR\_30001 (*cdgG*) como AZOBR\_RS02585, el cual se encuentra flanqueado río arriba por el gen AZOBR\_RS02590, que posee una longitud de 1575 nucleótidos y codifica para una presunta proteína de unión a ATP, y río abajo por el gen *AZOBR\_5s\_rRNA* de 116 nucleótidos que codifica para la subunidad 5S ribosomal. Entre el gen *cdgG* y el gen río arriba hay una distancia intergénica de 194 pares de bases, mientras que entre el gen *cdgG* y el gen río abajo hay 149 pares de bases. Los genes aledaños a *cdgG* difieren en la dirección de la transcripción por lo que no forman parte de un operón (**figura 6**).



**Figura 6.** Contexto genético del gen WP\_041811385 (*cdgG*)

El gen *cdgG* codifica para una presunta proteína híbrida denominada CdgG, la cual presenta en su extremo N-terminal un dominio EAL de 248 aminoácidos de longitud y un dominio GGDEF con una longitud de 118 aminoácidos. La arquitectura obtenida en el 2018 con el programa SMART solo mostraba ambos dominios, sin ningún otro dominio accesorio o sensor. Con las actualizaciones en las bases de datos, en el año 2021 se reportó la presencia de un dominio adicional, denominado CBS (Cistiotina  $\beta$ -sintetasa), el cual consta de 58 aminoácidos de longitud y se encuentra ubicado entre

el dominio EAL y GGDEF. Esto se confirmó con ayuda de la base de datos del Conserved Domains de NCBI (**Figura 7**).



**Figura 7.** Diagrama de la proteína CdgG donde se muestran los dominios EAL, GGDEF Y CBS (obtenido por la plataforma SMART-PROTEIN 2021)

### 7.1.2 Búsqueda de promotores y factores transcripcionales que regularían la expresión del gen *cdgG*

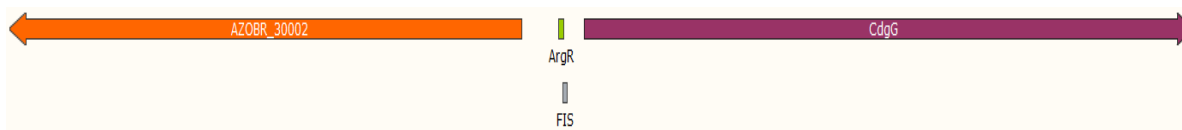
Con el objetivo de identificar probables regiones promotoras y sitios de unión para factores transcripcionales que regulen la expresión del gen *cdgG*, utilizamos la región intergénica de 194 nucleótidos presente entre el gen AZOBR\_RS02590 y el gen *cdgG*. Para este análisis utilizamos diferentes programas bioinformáticos que predicen la presencia de probables promotores, como son BPROM-Softberry e IPro54; para factores transcripcionales se utilizó el programa PRODORIC.

En cuanto al probable promotor, se localizaron cajas -10 (CGCTACAAG) y -35 (TTGTCA). Dichos elementos corresponden a promotores del factor sigma 70, por lo que se sugiere que el gen *cdgG* podría estar regulado por el factor sigma 70, **figura 8**.

```
> test sequence
Length of sequence-          194
Threshold for promoters -    0.20
Number of predicted promoters -    1
Promoter Pos:      154 LDF-    0.45
-10 box at pos.    139 cgctacaag Score    30
-35 box at pos.    114 ttgtca   Score    53
```

**Figura 8.** Identificación de un probable promotor del gen *cdgG* de *A. baldaniorum* Sp245 (BPROM ©, Softberry).

Por otro lado, con ayuda del programa PRODORIC identificamos secuencias probables de unión para 2 diferentes factores transcripcionales (**figura 9**); el factor transcripcional Fis y el ArgR.



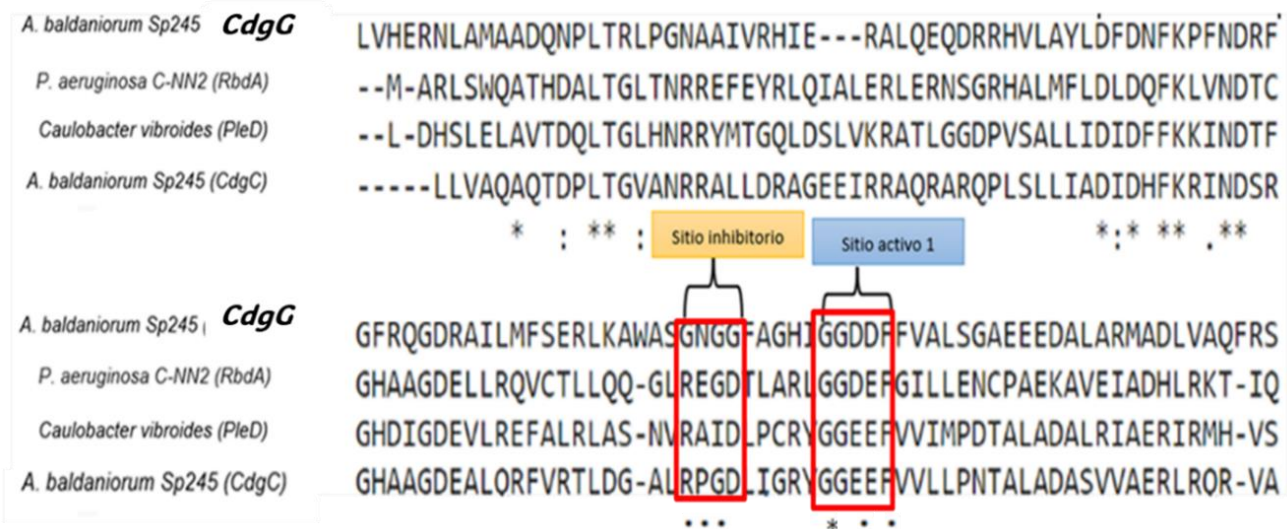
**Figura 9.** Identificación de posibles factores transcripcionales.

### 7.1.3 Alineamiento de secuencias de aminoácidos de la proteína CdgG

Con el objetivo de analizar si los dominios sensor y catalíticos de la proteína CdgG se encuentran conservados, se realizaron alineamientos múltiples con proteínas que ya se ha probado experimentalmente su actividad enzimática. Las proteínas utilizadas en estos análisis fueron: CdgC de *A. baldaniorum*, la proteína PleD de *C. crescentus* y la proteína híbrida RbdA de *P. aeruginosa* para GGDEF.

#### Dominio GGDEF

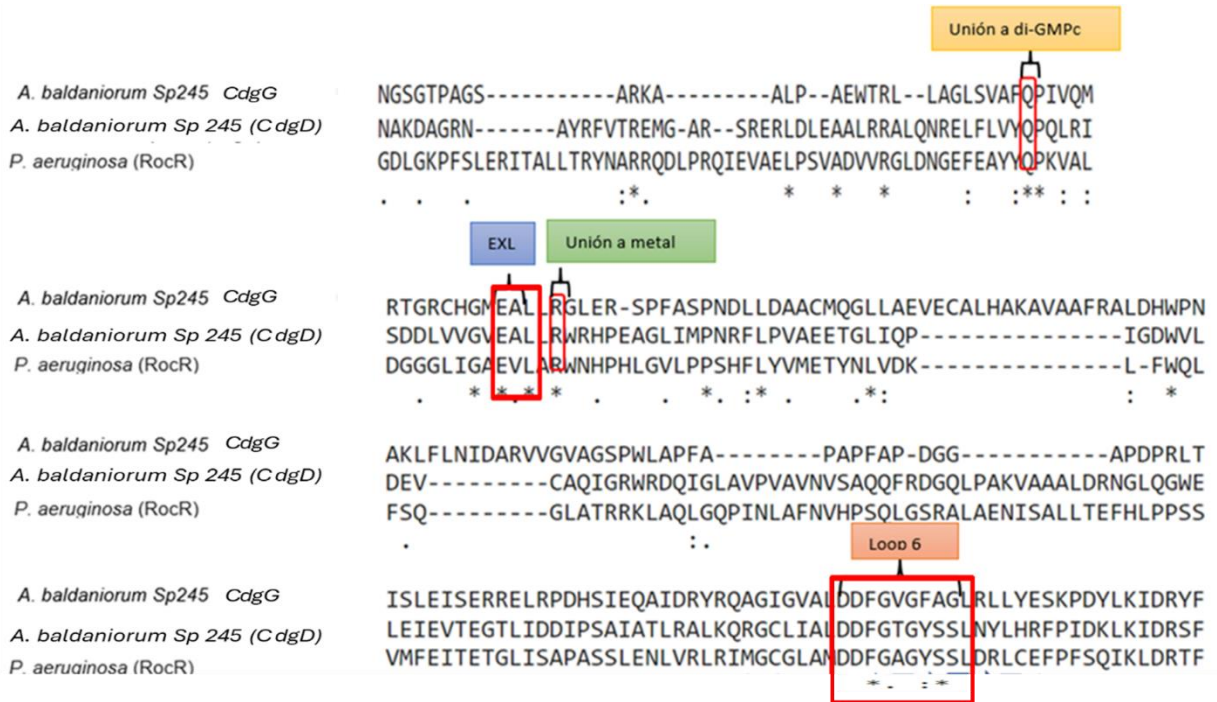
Las DGCs enzimáticamente activas se caracterizan por contener los motivos conservados; sitio activo GG(D/F)EF denominado sitio activo A1, el sitio activo 2 (DxxxxNxxxxD) y el sitio inhibitorio I (RxxD). En la **figura 10** se observa que los sitios activos I y 2 se encuentran conservados en todas las proteínas analizadas. En cuanto al motivo RxxD de la proteína CdgG este se encuentra degenerado en su totalidad, por lo que se sugiere que esta proteína no se regularía a través de este mecanismo.



**Figura 10.** Alineamiento múltiple de los motivos característicos de las DGCs con el dominio GGDEF de la proteína CdgG

### Dominio EAL

Para el análisis del dominio EAL de la proteína CdgG, se utilizaron los dominios EAL de la proteína RocR de *P. aeruginosa* y CdgD de *A. baldaniorum* Sp245. Como se observa en la **figura 11**, el sitio de unión al c-di-GMP está totalmente conservado, así como el motivo EAL seguido del segundo sitio de unión a metal. El motivo DDFG(T/A)GYSS, que forma el loop 6 se encuentra parcialmente conservado. Se observa un cambio de las dos últimas serinas por alanina y glicina, por lo que en estudios futuros se analizará si estos cambios impactan en la actividad de PDE, o en la dimerización de esta proteína.

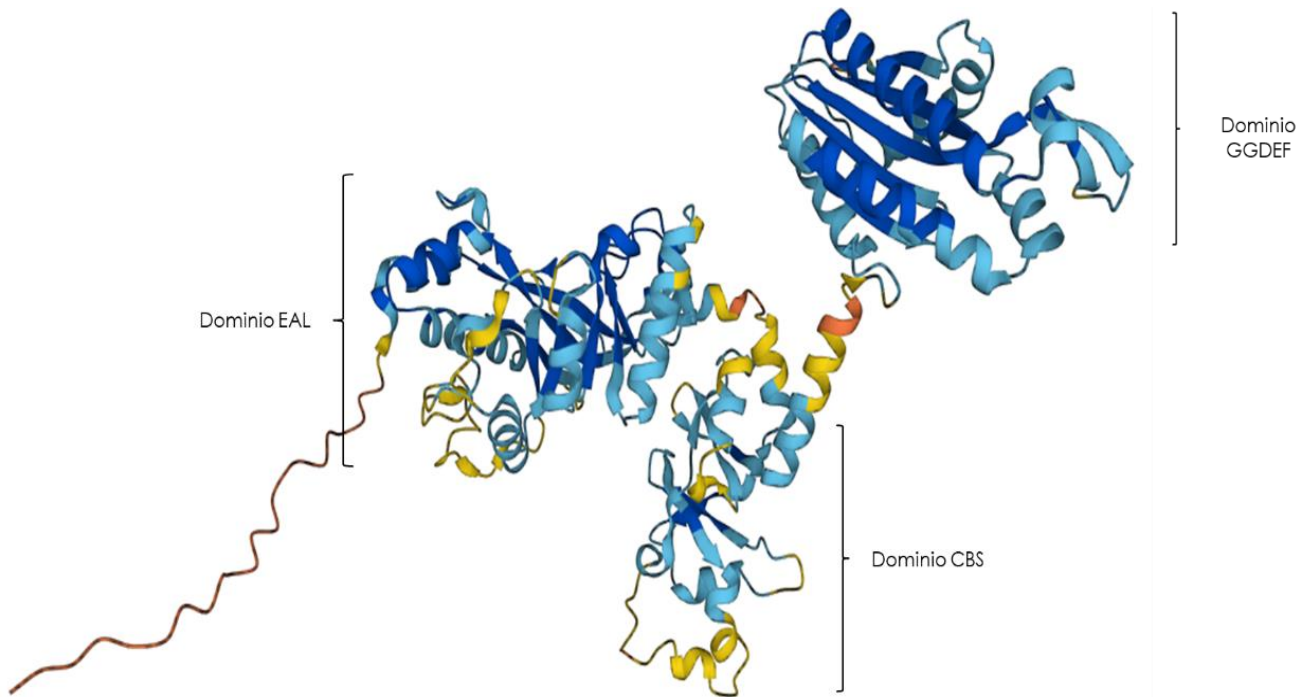


**Figura 11.** Alineamiento múltiple de aminoácidos del dominio EAL de la proteína CdgG con proteínas que tienen función de PDE.

#### 6.1.4 Estructura de la proteína CdgG: Topología del dominio CBS

A la fecha, el genoma de *A. baldaniorum* Sp245 se encuentra totalmente secuenciado y depositado en bases de datos de libre acceso. Mientras que en la base de datos UniProt se encuentran disponibles los modelos estructurales de cada una de las proteínas codificadas en el genoma de este microorganismo, dichos modelos fueron generados a través de la plataforma AlphaFold (<https://www.uniprot.org/uniprotkb/A0A9P1JNS7/entry>)(<https://alphafold.ebi.ac.uk/entry/G8AH69>) . Con estos modelos estructurales disponibles, nos dimos a la tarea de analizar el modelo estructural de la proteína CdgG (PDB: 2v0n). Con ayuda del visualizador de estructuras proteicas UCSF Chimera confirmamos la presencia de los dominios que conforman a la proteína en estudio. Se corroboró que el dominio EAL se encuentra ubicado en el extremo N-terminal de la proteína CdgG. Luego se encuentran dos dominios CBS dispuestos en tándem, y finalmente en el extremo C-terminal se encuentra ubicado el dominio GGDEF (**figura 12**). AlphaFold produce una puntuación de confianza por residuo (denominado Plddt) el cual maneja un rango de 0 a 100. AlphaFold proporciona dos métricas adicionales para evaluar la precisión de sus predicciones: la puntuación de modelado de plantilla predicha (pTM) y la puntuación de modelado de plantilla predicha de interfaz. Ambos se derivan de una medida llamada puntuación de modelado de plantilla (TM), que mide la precisión de la estructura global de la proteína y es relativamente insensible a imprecisiones localizadas. Esto quiere decir que puntuaciones mayores al 70 puntos representado en colores azul, nos indican mayor confianza en cuanto a una predicción exacta. Algunas regiones con bajo pLDDT pueden no estar estructuradas. Siguiendo estas reglas se califican los modelos estructurales en esta plataforma (Xhu y Zhang, 2010).

De acuerdo con el modelo estructural de la proteína CdgG obtenido en AlphaFold; los dominios catalíticos EAL y GGDEF muestran valores altos de confianza (pLDDT > 90, regiones en azul), principalmente en los sitios activos. Mientras que para el dominio CBS, en general, se observó un valor pLDDT > 70, otorgándole a las 3  $\beta$ -plegadas un valor mayor (estructuras secundarias importantes en los dominios CBS). Si bien hay regiones de la proteína con baja confianza, estas regiones corresponden a los extremos de la proteína o a regiones que funcionan de uniones entre los dominios, los cuales normalmente son regiones no bien estructuradas, otorgándoles valores de baja puntuación (pLDDT > 50 o < 50, regiones en color naranja y amarillo, respectivamente).



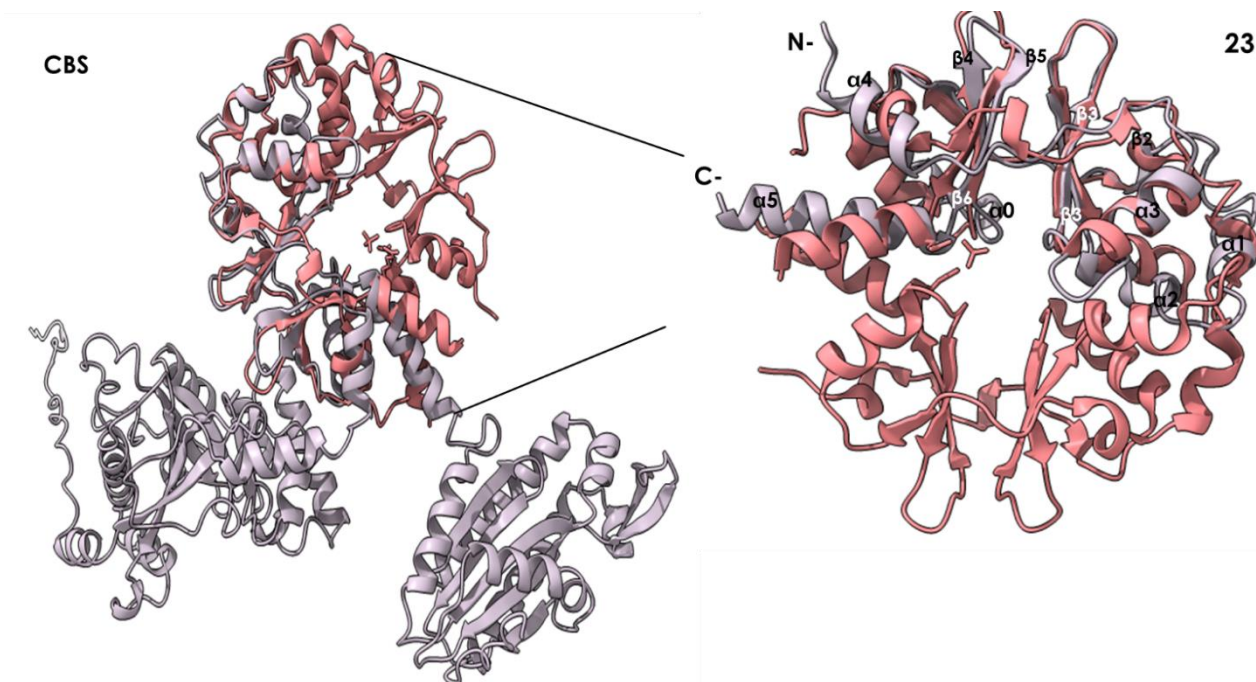
**Figura 12.** Estructura 3D de la proteína CdgG en AlphaFold.

### Topología de los dominios CBS

El dominio CBS consiste en una estructura  $\beta$ -plegada de tres hebras y dos  $\alpha$ -hélices empaquetadas de acuerdo con la secuencia  $\beta 1$ -  $\alpha 1$ -  $\beta 2$ -  $\beta 3$ -  $\alpha 2$ . Todos los dominios CBS siempre están precedidos por un enlazador de longitud variable que incluye un giro N-terminal de alfa hélice ( $\alpha 0$ -1). Como ya se mencionó previamente, los dominios CBS casi siempre se asocian en ensamblajes diméricos denominados módulos Bateman (Baykov et al., 2011).

Dado que los dominios catalíticos se encuentran conservados, nos centramos en analizar a los dominios CBS. Para tal fin, se realizó un empalme estructural entre los dominios CBS presentes en CdgG y la proteína DarB de *Bacillus subtilis*, la cual ya se ha cristalizado y está depositada en el Protein Data Bank con el número PDB 6YJA. Este cristal tiene una resolución de 1.76 Å y se cristalizó

con una molécula de c-di-AMP. Se puede notar el empalme estructural entre el dominio CBS de CdgG con la proteína DarB, lo que sugiere una conservación estructural entre los dominios CBS de ambas proteínas (Heidemann et al., 2022).

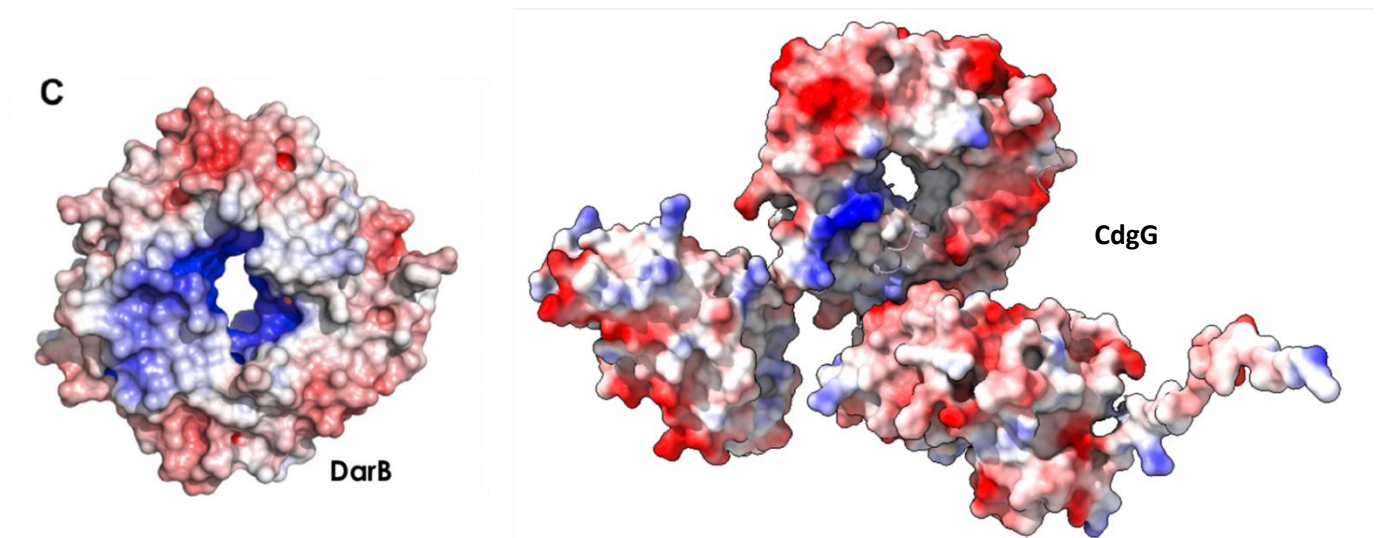


**Figura 13.** Empalme estructural entre los dominios CBS presentes en CdgG y la proteína DarB de *Bacillus subtilis*, en rosa se muestra lo correspondiente al dominio CBS la proteína Dar, en color gris lo correspondiente a la proteína CdgG. Visualizado en Chimera X.

### 6.1.5 Análisis de potencial electrostático de la proteína CdgG

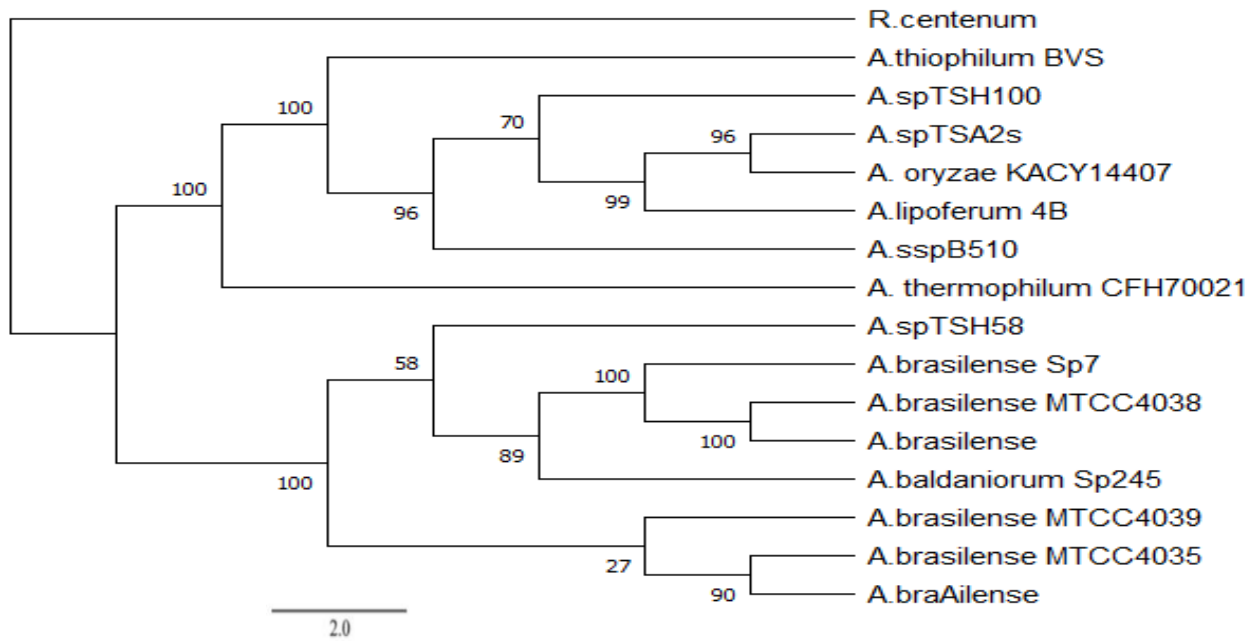
Un análisis del potencial electrostático de una proteína permite determinar la carga neta, por lo que es probable inferir la carga de la molécula que puede unir, y más aún, si se tiene un cristal acoplado con su ligando. De acuerdo con estos datos y para saber cuál es la carga neta de los dominios CBS, llevamos a cabo un análisis de potencial electrostático, tanto de los dominios CBS de CdgG, así como de la proteína DarB, en los que observamos que DarB forma una estructura que asemeja a una rosquilla, la cual muestra una superficie externa con predominante carga negativa, mientras que la superficie de su poro central está cargado positivamente; como se observa en la **figura 14**, en donde se une la adenina que forma parte del c-di-AMP. En cuanto al potencial electrostático de la proteína CdgG conserva de manera importante las mismas cargas electrostáticas que DarB como se observa en la **figura 14**. Con estos datos podemos sugerir que los dominios CBS de CdgG podrían unir

ligandos semejantes a los que une DarB, (Heidemann et al., 2022) y de esta manera, activar o reprimir las actividades de DGC y/o PDE de CdgG. Sin embargo, estos datos se analizarán en futuras investigaciones, para determinar si estas hipótesis son correctas o no.



**Figura 14.** Acoplamiento del dímero de CdgG con el tetrámero de DarB. Se analizó la carga neta de la proteína y se comparó con la proteína DarB. El homodímero DarB en forma de rosquilla tiene una superficie externa predominantemente cargada negativamente, mientras que la superficie de su poro central está cargada positivamente.

Adicionalmente se realizó un análisis filogenético encontrando los genes homólogos a *cdgG* en la clase de las proteobacterias (**figura 15**). Se encontraron 1120 ortólogos de los cuales se analizaron 16 secuencias que corresponden a alfa proteobacterias, sugiriendo que la arquitectura característica de la proteína y su potencial actividad dual DGC y/o PDE pudiera dar a estas bacterias una ventaja selectiva o la capacidad de adaptarse en función a las señales ambientales que pudieran detectar los dominios CBS. El análisis filogenético se hizo con la secuencia de aminoácidos, y arrojó el árbol filogenético mostrado en la **figura 15**.



**Figura 15.** Análisis filogenético del gen *cdgG*. El gen *cdgG* de *A. baldaniorum* Sp245 está conservado en la clase alfa proteobacteria. Las secuencias de nucleótidos del gen *cdgG* y los genes homólogos fueron alineadas con Clustal Omega. Los archivos del árbol filogenético fueron generados por el software MEGA versión VII. El árbol filogenético se produjo utilizando el método de Máxima Verosimilitud y el modelo de sustitución de Tamura-Nei, con 1000 réplicas bootstrap. Los números en las ramas del filograma indican los valores de bootstrap mayores al 50%. La barra de escala corresponde a 2.0 sustituciones por posición de nucleótido.

### Clonación del gen *cdgG* en el vector pGEM-TEasy.

Para la clonación del gen *cdgG* de *A. baldaniorum* Sp245 al vector de mantenimiento pGEM-TEasy (figura 16), se diseñaron oligonucleótidos que amplificaron el ORF de *cdgG* (tabla 1). Una vez que se obtuvieron los cebadores correspondientes al ORF *cdgG*, se procedió a realizar la PCR con las condiciones mostradas en la figura 17. En la figura 18 se muestra el producto de PCR.

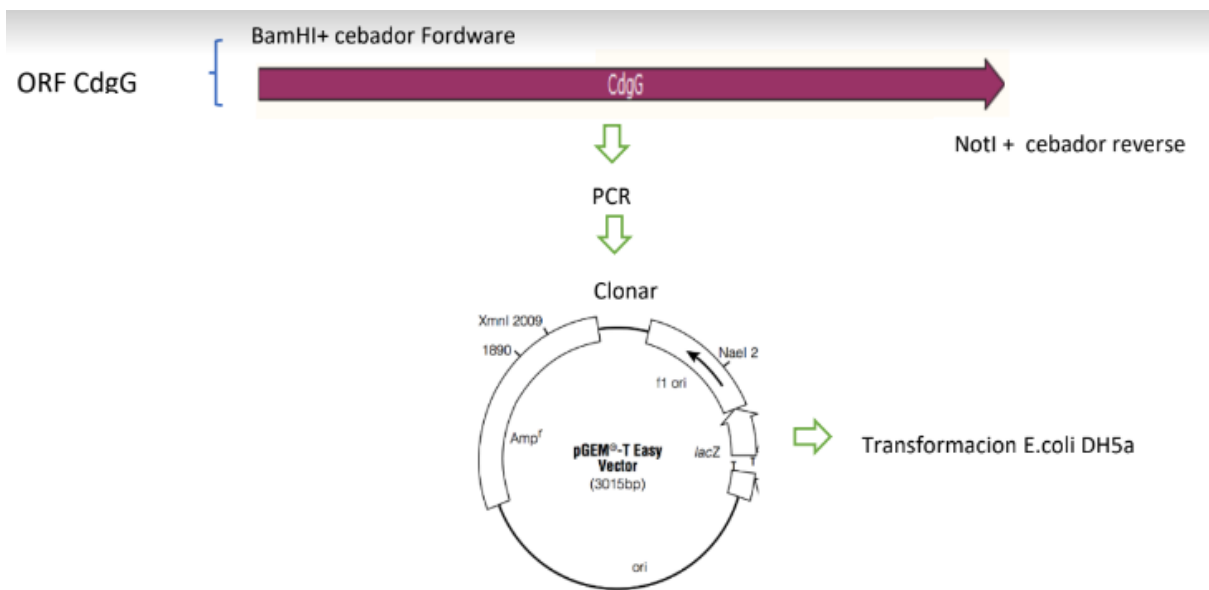


Figura 16. Esquema de trabajo para la clonación del gen *cdgG*

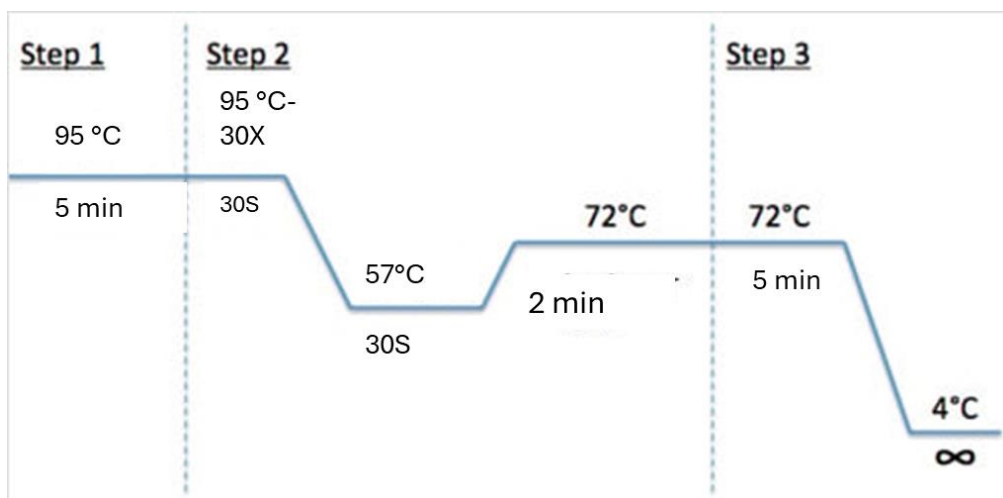
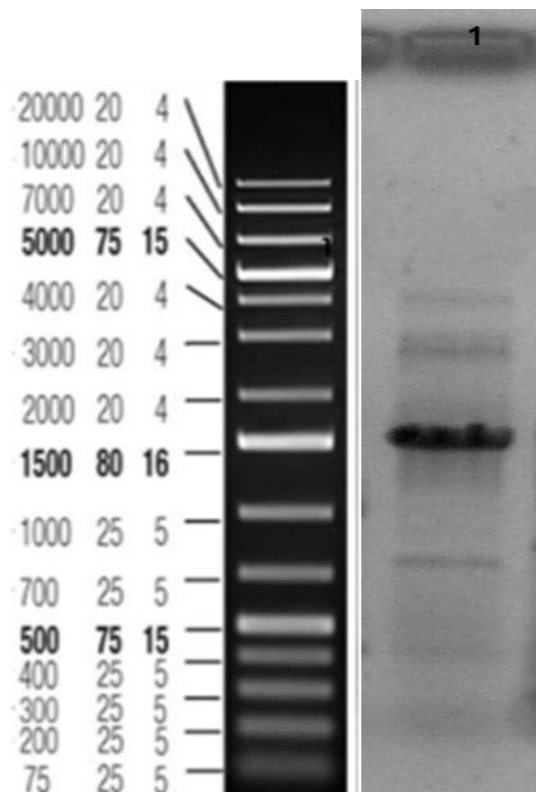


Figura 17. Condiciones para la PCR del gen *cdgG*.

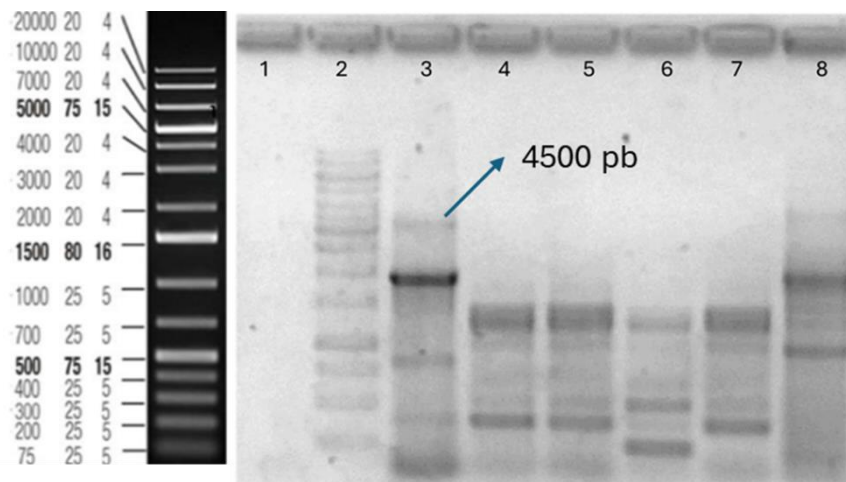
Primers	
Delantero: <i>Bam</i> HI CdgG	5'-ggatccatggcctcggtgcca-3'
Reverso: <i>Not</i> I CdgG	5'-ctcgagctaaccctccccgctcc-3'

Tabla 1. Iniciadores que abarcan el ORF del gen *cdgG* y las secuencias de restricción *Bam*HI y *Not*I, respectivamente.



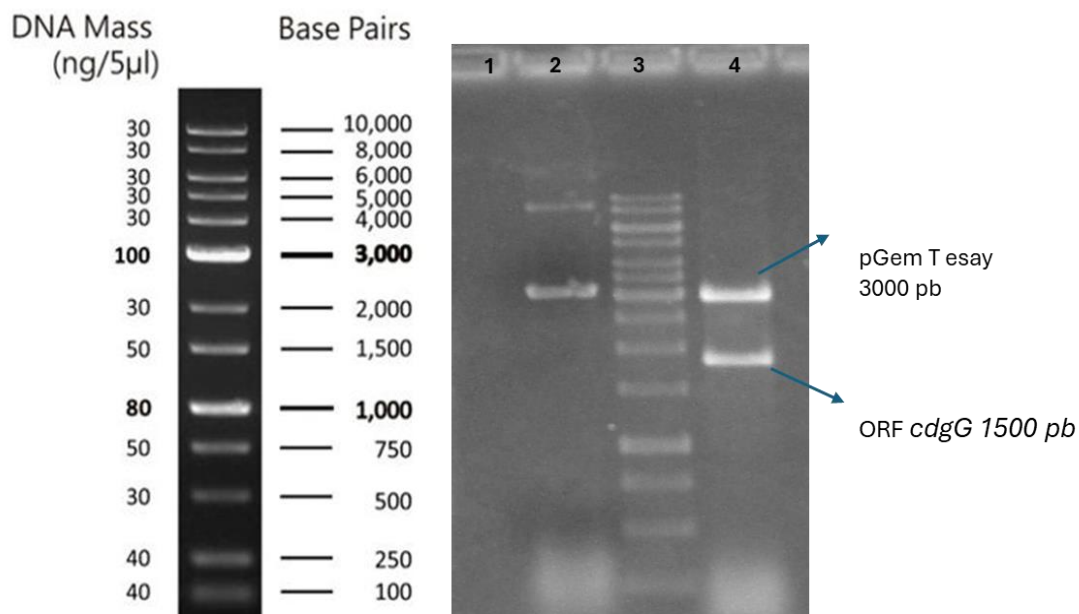
**Figura 18.** Producto de PCR. Carril 1 producto de 1890 pb.

Con el producto de la PCR, se llevó a cabo la ligación al vector pGEM-Teasy y el vector resultante se transformó en *E. coli DH5a*. La figura 19 muestra un total de 6 colonias obtenidas a las que se les realizó PCR en colonia para determinar si poseían el producto clonado (**figura 19**).



**Figura 19.** Clonas obtenidas producto de la transformación en *E. coli DH5a*. Se muestra las bandas del producto de la PCR en colonia. Carril 1, control negativo de la PCR en colonia, carril 2 MPM, carril 3 al 8 se muestran las colonias siendo un total de 6. Las bandas adicionales a 4500 pb podrían sugerir que son secuencias inespecíficas.

Se seleccionó la colonia del carril 3 a la cual se le realizó la restricción con *NotI* y *BamHI* la cual se muestra en la **figura 20**, confirmado que se tiene clonado el ORF del gen *cdgG* en pGEM-Teasy.



**Figura 20.** Producto de la restricción. Carril 1 Control negativo, carril 2 pGem-*cdgG*, carril 3 MPM, carril 4 producto de la restricción. Se observa los 2 fragmentos liberados por los cortes con *NotI* y *BamHI*. Un producto liberado de aproximadamente 3000 pb y otro de 1500 pb.

## 7.- Discusión

Los sistemas de traducción de señales participan en redes de señalización celular que permiten traducir los estímulos externos (Mascher et al., 2006). Las bacterias emplean principalmente segundos mensajeros de origen nucleotídico (Stülke J and Krüger L., 2020). Entre los segundos mensajeros nucleotídicos bacterianos se encuentran la guanina (penta o tetra) fosfato ((p)ppGpp), el monofosfato de adenosina cíclico (cAMP), el di-AMP cíclico (c-di-AMP), el di-GP cíclico (c-di-GMP) y el monofosfato de guanisina cíclico-monofosfato de adenosina (cGAMP), entre otros. Estos mensajeros actúan como intermediarios cruciales en la conexión entre la percepción de señales bacterianas y las respuestas celulares (Burroughs, 2015). Como resultado, se observan alteraciones en las concentraciones de nucleótidos específicos, las cuales asumen el papel de segundos mensajeros. Las variaciones en los niveles de los segundos mensajeros modulan su unión a receptores y/o efectores posteriores, regulando así funciones fisiológicas específicas en las bacterias. Esta compleja red de señalización permite a las células bacterianas adaptar sus respuestas a una amplia gama de estímulos ambientales (Orr et al., 2016).

Al recibir una señal específica, estas enzimas pueden modificar rápidamente las concentraciones celulares del segundo mensajero respectivo, lo que a su vez altera las propiedades (actividad, conformación y/o estado oligomérico) de sus receptores. Los receptores celulares para segundos mensajeros nucleotídicos suelen ser abundantes y diversos, lo que le permite regular las vías descendentes de múltiples maneras (expresión génica, actividad enzimática y propiedades de unión). Estos sistemas de transducción de señales permiten a las bacterias participar en la detección continua y/o la regulación coordinada de un solo proceso biológico, asegurando una respuesta precisa y eficiente a los estímulos ambientales (Yu et al., 2023) (Orr et al., 2016).

Las diguanilatociclasa y las fosfodiesterasas actúan como intermediarios clave en la transducción de señales, transformando estímulos ambientales en cambios en la concentración de c-di-GMP, lo que a su vez modula el comportamiento bacteriano. Estas enzimas de señalización se clasifican en dos categorías según su modo de percepción de señales:

En señalización al integrar dominios sensoriales y un dominio enzimático en una sola proteína, permitiendo una respuesta directa a una amplia gama de estímulos, como gases, luz, estado redox, temperatura y compuestos químicos.

Sistemas de dos componentes: En contraste, algunas enzimas carecen de dominios sensoriales propios y dependen de mecanismos alternativos para percibir señales:

Interacciones proteína-proteína: algunas reciben señales a través de interacciones con proteínas sensoriales que contienen dominios sensoriales específicos como en el caso de DevR en *Mycobacterium smegmatis* (Hu et al., 2019).

Regulación corriente abajo: Otras actúan como reguladores de respuesta corriente abajo en sistemas de dos componentes o redes de fosforrelvo multicomponente más complejas, modificando su actividad enzimática en respuesta a la transferencia de grupos fosforilo desde dominios receptores de fosfato corriente arriba (Motaleb et al., 2007) (Hu et al., 2019).

A pesar de las diferencias en los mecanismos de detección de señales, todas las DGC PDE y enzimas de señalización comparten la característica común de ajustar sus actividades catalíticas tras recibir señales, alterando así las concentraciones de c-di-GMP en las bacterias. Estas fluctuaciones son detectadas por receptores específicos corriente abajo, que a su vez regulan múltiples funciones fisiológicas (Valentini et al., 2019).

En el presente trabajo utilizamos programas bioinformáticos para conocer a mayor detalle la estructura genética y producto proteico del gen *cdgG*. Este gen codifica para una proteína híbrida probablemente implicada en el metabolismo del c-di-GMP, el producto génico está conformado por tres dominios característicos EAL-CBS-GGDEF, lo que sugiere su probable participación en la síntesis y degradación de c-di-GMP. Con base en esta arquitectura inusual de contener en la región N-terminal un dominio EAL o de PDE y un dominio GGDEF o DGC en el carboxilo terminal, surgió la pregunta ¿Sí este gen se encuentra conservado en el género *Azospirillum* y en bacterias filogenéticamente emparentadas o es exclusivo de la cepa 245? Después de hacer la búsqueda en las bases de datos logramos mostrar que el gen *cdgG* se encuentra codificado también en otras cepas de *Azospirillum* y en otras bacterias filogenéticamente emparentadas como es el caso de *Rhodospirillum centenum*, lo que sugiere que debe tener un papel importante en el metabolismo del c-di-GMP.

Posteriormente, indagamos por métodos bioinformáticos si este gen se encuentra formando parte de un operón y, debido a que los genes aledaños a *cdgG* difieren en la dirección de la transcripción y existen espacios intergénicos largos, se propone que no formaría parte de un operón. Así mismo, se buscó la presencia de probables promotores y factores transcripcionales que regulen la expresión

del gen *cdgG*, encontrando en su región intergénica 5' de 194 pb elementos -10 y -35 que sugieren se podría regular por el factor sigma 70. Cabe mencionar, que en dicha región, se buscó elementos reguladores de otros factores. Sin embargo, no se halló secuencias involucradas en otros factores. En esa misma región intergénica identificamos secuencias probables de unión a los factores transcripcionales Fis y ArgR. Fis es un regulador global descrito en *E.coli*; se ha observado que tiene un papel crítico en coordinar la expresión de genes de *E. coli*, así como también regula la condensación del DNA en la célula bacteriana. Fis es una proteína abundante asociada a nucleóide y está relacionada a la regulación de la expresión de genes en células de rápido crecimiento y puede variar su función ya que puede funcionar como represor o activador transcripcional de acuerdo con la posición de su sitio de unión. Se ha observado que Fis tiene preferencia de unión a regiones ricas de AT (Monteiro et al., 2020). En otras bacterias como *Dickeya dadantii*, se demostró que Fis modula la producción de enzimas extracelulares, incluidas celulasas, pectatoliasas y proteasas, además de participar en la formación de biopelículas (Prigent-Combaret et al., 2012).

En cuanto al factor transcripcional ArgR, se ha reportado que en bacterias como *P. aeruginosa* y *E. coli*, el metabolismo y el transporte de arginina están controlados por el regulador transcripcional ArgR. En presencia del aminoácido arginina, ArgR actúa como represor de los genes involucrados en la biosíntesis de arginina y activa la expresión de genes en la vía catabólica de la succiniltransferasa de arginina, así como en los operones que codifican los sistemas básicos de transporte de aminoácidos. ArgR de *E. coli* forma trímeros en ausencia de arginina que se estabilizan como hexámeros por L-arginina. Se ha propuesto que este cambio conformacional altera la interacción de la proteína con su sitio de unión al DNA, por la caja ARG presente en el operador de los genes regulados por ArgR (Cho et al., 2015).

ArgR es un represor transcripcional que controla la expresión de genes que participan en la síntesis de arginina, pirimidinas, metabolismo del nitrógeno, nitrato, producción de antibióticos y utilización de nitrato en bacterias (Cho et al., 2015).

El regulador transcripcional ArgR está ampliamente distribuido en bacterias, actuando principalmente como represor de los genes para la captación y biosíntesis de arginina y ornitina en *Pseudomonas* (Szwajkajzer et al., 2021). ArgR interactúa con el DNA en secuencias específicas conocidas como cajas ARG (ARG box). *E. coli* está formada por dos secuencias de 18 pb separadas por 3 pb y se encuentra cerca de los promotores de los genes controlados por ArgR. En *E. coli* la

unión de ArgR a las cajas ARG depende de la L-arginina que funciona como un correpresor (Barrientos-Moreno et al., 2020).

En cuanto al análisis bioinformático de la proteína CdgG se determinó que en el dominio de fosfodiesterasa EAL se encuentran conservados los motivos importantes para la actividad de fosfodiesterasa, mientras que el dominio GGDEF se encuentran conservados también los sitios catalíticos, no así el sitio inhibitorio ya que este se encuentra completamente degenerado, por lo que la actividad de DGC no podría ser regulada por el mecanismo de retroalimentación positiva como ocurre en otras proteínas DGC tanto de dominio único PleD *C. crescentus* (Christen et al., 2006) como híbridas, tal es el caso de RbdA de *P. aeruginosa* (Liu et al., 2018).

Por otro lado, y dado que los dominios catalíticos se encuentran conservados, nos centramos en analizar al dominio CBS que forman parte de la proteína CdgG y encontramos que la secuencia primaria del dominio CBS se encuentra poco conservada, mientras que la estructura terciaria del dominio se conserva. Esto lo pudimos observar gracias a la estructura cristalográfica que se encuentra depositada en el PDB de la proteína DarB de *B. subtilis* (Heidemann et al., 2022).

Los dominios CBS son pequeños módulos proteicos (60 residuos de longitud) descritos por primera vez en la proteína cistationina  $\beta$ -sintasa en humanos y bacterias (Baykov et al., 2011). Posteriormente, se han encontrado también en proteínas membranales y citoplasmáticas de todos los reinos de la vida. Los dominios CBS están presentes en una gran variedad de proteínas, incluidos transportadores de cloro, cinasas activadas por AMP, inosin-5-monofosfato deshidrogenasa (Biemans-Oldehinkel et al., 2006; Baykov et al., 2011).

Los dominios CBS presentan baja conservación en su estructura primaria, pero son uniformes en longitud. Estos dominios usualmente se encuentran adyacentes y separados por un motivo DRTGG (que se encuentra presente en la proteína CdgG) en aproximadamente el 1% de las proteínas reportadas (principalmente en proteínas de bacterias). Los dominios CBS suelen presentarse en pares y con menos frecuencia en cuartetos. Los dominios CBS que se encuentran en pares forman una estructura ajustada denominada dominio Bateman o módulo Bateman. Estos dominios adoptan un plegamiento con estructura  $\alpha/\beta$  con simetría interna. Ellos consisten en arreglos de  $\beta$ - $\alpha$ - $\beta$ - $\alpha$ , con hélices pequeñas adicionales y cadenas  $\beta$  cortas en la región N-terminal de la proteína y forma un enlace con el dominio CBS vecino.

Se ha demostrado que los dominios CBS se unen principalmente a ligandos con un grupo adenosilo como AMP, ATP, S-AdoMet (S-adenosilmetionina), c-di-AMP, pero también pueden unirse a iones metálicos o ácidos nucleicos. Los dominios CBS regulan las funciones celulares en respuesta a los cambios en el estado energético (es decir, funcionan como sensores de energía), participando en el transporte de iones  $Mg^{2+}$ , osmorregulación, modulación del tráfico intracelular de los canales de cloruro, transporte de nitrato o en la actividad de la pirofosfatasa (Biemans-Oldehinkel et al., 2006).

Recientemente, una proteína con dominio CBS denominada DarB en *B. subtilis* fue identificada unida a c-di-AMP (Krüger et al., 2021). DarB tiene una longitud de 147 aminoácidos y forma un repetido en tándem CBS que comprende dos dominios CBS sin ningún otro dominio adicional. DarB se une al GTP de la pirofosfocinasa Rel y estimula la síntesis del nucleótido (p)ppGpp de una manera independiente de ribosoma. Mientras que esta interacción es suprimida cuando DarB se une a c-di-AMP. La síntesis de (p)ppGpp es parte de la respuesta astringente de células bacterianas bajo privación de aminoácidos o bajo otras condiciones de estrés. Adicionalmente, la versión apo-DarB se une a la enzima piruvato carboxilasa para estimular el reinicio del ciclo del ácido cítrico (Christen et al., 2006; Heidemann et al., 2022). Se ha demostrado que DarB se une específicamente a c-di-AMP con una constante de disociación en el rango nanomolar y a AMP con mucha más baja afinidad. DarB también se une a cGAMP *in vitro*; sin embargo, como cGAMP no existe en *B. subtilis*, esta interacción puede no tener relevancia fisiológica, pero puede revelar posibilidades interesantes para el desarrollo de blancos terapéuticos. Para entender la especificidad y afinidad a diferentes ligandos se han determinado cuatro estructuras cristalográficas de DarB, en su forma apo y en unión a 3 nucleótidos c-di-AMP, cGAMP y AMP, respectivamente. Los cuatro dominios CBS de la estructura homodimérica de DarB tiene la capacidad de unión a dos moléculas de c-di-AMP; sin embargo, solo una adenina de cada c-di-AMP es específicamente reconocida por DarB, mientras que la segunda adenina sobresale del homodímero con forma de dona. Esta adenina expuesta es probable que inhiba la interacción DarB-Rel, ya que no hay cambios conformacionales en DarB una vez que une c-di-AMP, lo cual podría causar la pérdida de la interacción proteína-proteína (Heidemann et al., 2022) (Ainelo et al., 2023).

Finalmente, es necesario ahondar en los análisis bioinformáticos de interacción molecular como el docking, para identificar entre otras cosas, posibles ligandos de esta proteína, y del mismo modo, abordarlos de manera experimental que permitan o no corroborar la hipótesis de que el dominio CBS es un dominio sensor que podría regular las funciones de la CdgG como PDE o DGC y determinar

¿cuál es su impacto en la interacción con la planta? También es interesante especular que esta proteína pudiera ser un nodo de interacción de dos nucleótidos (c-di-AMP y c-di-GMP), para modular el comportamiento bacteriano. Sin embargo, en primer lugar, hay que mostrar que en *Azospirillum* se sintetiza c-di-AMP y que funciona como segundo mensajero como ocurre en *B. subtilis*, para sugerir que existe una conexión entre estos dos nucleótidos.

## 8.- Conclusión

De acuerdo con los análisis bioinformáticos realizados a la proteína CdgG, se observó que el dominio EAL se encuentra ampliamente conservado. En el caso del dominio GGDEF; los sitios catalíticos se encuentran conservados. Sin embargo, el sitio inhibitorio no está del todo conservado. En cuanto al dominio CBS tiene alta identidad estructural con la proteína DarB de *B. subtilis*.

## 9.-Perspectivas

- Secuenciar el fragmento clonado obtenido en la parte experimental
- Generar la mutante AZOBR\_30001 mediante la inserción de un cassette de resistencia a kanamicina y el factor fluorescente mCherry en el vector suicida PJMS-mcherryKan.
- Analizar fenotípicamente la cepa mutante y silvestre, en formación de biopelícula, movilidad y colonización con plantas de trigo.

## 10.- Bibliografía

- Ainelo A, Caballero-Montes J, Bulvas O, Ernits K, Coppieters 't Wallant K, Takada H, Craig SZ, Mazzucchelli G, Zedek S, Pichová I, Atkinson GC, Talavera A, Martens C, Hauryliuk V, Garcia-Pino A. The structure of DarB in complex with Rel<sup>NTD</sup> reveals nonribosomal activation of Rel stringent factors. *Sci Adv.* 2023 Jan 18;9(3):eade4077. doi: 10.1126/sciadv.ade4077.
- Alexander A. Baykov, Heidi K. Tuominen, and Reijo Lahti. *ACS Chemical Biology* **2011** 6 (11), 1156-1163.DOI: 10.1021/cb200231c
- Barrientos-Moreno L, Molina-Henares MA, Ramos-González MI, Espinosa-Urgel M. Arginine as an environmental and metabolic cue for cyclic diguanylate signalling and biofilm formation in *Pseudomonas putida*. *Sci Rep.* 2020 Aug 12;10(1):13623. doi: 10.1038/s41598-020-70675-x.
- Baykov AA, Tuominen HK, Lahti R. The CBS domain: a protein module with an emerging prominent role in regulation. *ACS Chem Biol.* 2011 Nov 18;6(11):1156-63. doi: 10.1021/cb200231c.
- Biemans-Oldehinkel E, Mahmood NA, Poolman B. A sensor for intracellular ionic strength. *Proc Natl Acad Sci U S A.* 2006 Jul 11;103(28):10624-9. doi: 10.1073/pnas.0603871103.
- Burroughs AM, Zhang D, Schäffer DE, Iyer LM, Aravind L. Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res.* 2015 Dec 15;43(22):10633-54. doi: 10.1093/nar/gkv1267. Epub 2015 Nov 20.
- Camelo, M., Vera, S.P. and Bonilla, R. (2011) Mecanismos de acción de las rizobacterias promotoras del crecimiento vegetal. *Rev. Corpoica-Ciencia y Tecnol. Agropecuaria*, 12, 159-166.
- Carreño-López, R., Sánchez, A., Camargo, N. *et al.* Characterization of *chsA*, a new gene controlling the chemotactic response in *Azospirillum brasilense* Sp7. *Arch Microbiol* 191, 501–507 (2009). <https://doi.org/10.1007/s00203-009-0475-x>
- Cassán F, Coniglio A, López G *et al.* Everything you must know about *Azospirillum* and its impact on agriculture and beyond. *Biol Fertil Soils* 2020
- Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, Schirmer T. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A.* 2004 Dec 7;101(49):17084-9. doi: 10.1073/pnas.0406134101. Epub 2004 Nov 29. PMID: 15569936; PMCID: PMC535365.
- Chao L, Rakshe S, Leff M, Spormann AM. PdeB, a cyclic Di-GMP-specific phosphodiesterase that regulates *Shewanella oneidensis* MR-1 motility and biofilm formation. *J Bacteriol.* 2013 Sep;195(17):3827-33. doi: 10.1128/JB.00498-13.
- Cho S, Cho YB, Kang TJ, Kim SC, Pálsson B, Cho BK. The architecture of ArgR-DNA complexes at the genome-scale in *Escherichia coli*. *Nucleic Acids Res.* 2015 Mar 31;43(6):3079-88. doi: 10.1093/nar/gkv150
- Christen B, Christen M, Paul R, Schmid F, Folcher M, Jenoe P, Meuwly M, Jenal U. Allosteric control of cyclic di-GMP signaling. *J Biol Chem.* 2006 Oct 20;281(42):32015-24. doi: 10.1074/jbc.M603589200.

- Cintia Mariana Pereyra, Claudia Cristina Dal Lago, Cecilia Mónica Creus, María Alejandra Pereyra, *Azospirillum baldaniorum* Sp 245 inoculation affects cell wall and polyamines metabolisms in cucumber seedling roots, *FEMS Microbiology Letters*, Volume 370, 2023, fnad005, <https://doi.org/10.1093/femsle/fnad005>
- Cruz-Pérez, J. F., Lara-Oueilhe, R., Marcos-Jiménez, C., Cuatlayotl-Olarte, R., Xiqui-Vázquez, M. L., Reyes-Carmona, S. R., Baca, B. E., & Ramírez-Mata, A. (2021). Expression and function of the *cdgD* gene, encoding a CHASE–PAS–DGC–EAL domain protein, in *Azospirillum brasilense*. *Scientific Reports*, 11(1), 520. <https://doi.org/10.1038/s41598-020-80125-3>
- Ferreira, Natalia Dos Santos; Sant’Anna, Fernando Hayashi; Reis, Veronica Massena; Ambrosini, Adriana; Volpiano, Camila Gazolla; et al.; Genome-based reclassification of *azospirillum brasilense* SP245 as the type strain of *azospirillum baldaniorum* sp. nov; Society for General Microbiology; International Journal of Systematic and Evolutionary Microbiology; 70; 12; 10-2020; 6203-6212
- Ganusova EE, Vo LT, Mukherjee T, Alexandre G. Multiple CheY Proteins Control Surface-Associated Lifestyles of *Azospirillum brasilense*. *Frontiers in Microbiology*. 2021 ;12:664826. DOI: 10.3389/fmicb.2021.664826.
- Hadas, R., Okon, Y. Effect of *Azospirillum brasilense* inoculation on root morphology and respiration in tomato seedlings. *Biol Fert Soils* 5, 241–247 (1987). <https://doi.org/10.1007/BF00256908>
- Hadjidemetriou K, Kaur S, Cassidy CK, Zhang P. Mechanisms of *E. coli* chemotaxis signaling pathways visualized using cryoET and computational approaches. *Biochem Soc Trans*. 2022 Dec 16;50(6):1595-1605. doi: 10.1042/BST20220191.
- Hengge, R. Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* 7, 263–273 (2009). <https://doi.org/10.1038/nrmicro2109>
- Heidemann JL, Neumann P, Krüger L, Wicke D, Vinhoven L, Linden A, Dickmanns A, Stülke J, Urlaub H, Ficner R. Structural basis for c-di-AMP-dependent regulation of the bacterial stringent response by receptor protein DarB. *J Biol Chem*. 2022 Jul;298(7):102144. doi: 10.1016/j.jbc.2022.102144
- Hu Q, Zhang J, Chen Y, Hu L, Li W, He ZG. Cyclic di-GMP co-activates the two-component transcriptional regulator DevR in *Mycobacterium smegmatis* in response to oxidative stress. *J Biol Chem*. 2019 Aug 23;294(34):12729-12742. doi: 10.1074/jbc.RA119.008252.
- Jenal U, Reinders A, Lori C. Cyclic di-GMP: second messenger extraordinaire. *Nat Rev Microbiol*. 2017 May;15(5):271-284. doi: 10.1038/nrmicro.2016.190. Epub 2017 Feb 6.
- Jeong, GJ., Khan, F., Tabassum, N. et al. Motility of *Acinetobacter baumannii*: regulatory systems and controlling strategies. *Appl Microbiol Biotechnol* 108, 3 (2024). <https://doi.org/10.1007/s00253-023-12975-6>
- Jinrui Xu, Yang Zhang, How significant is a protein structure similarity with TM-score = 0.5?, *Bioinformatics*, Volume 26, Issue 7, April 2010, Pages 889–895,
- Kreiling V, Thormann KM. Polarity of c-di-GMP synthesis and degradation. *Microlife*. 2023 Apr 5;4:uqad014. doi: 10.1093/femsml/uqad014.
- Kunz S, Tribensky A, Steinchen W, Oviedo-Bocanegra L, Bedrunka P, Graumann PL. Cyclic di-GMP Signaling in *Bacillus subtilis* Is Governed by Direct Interactions of Diguanylate Cyclases and Cognate Receptors. *mBio*. 2020 Mar 10;11(2):e03122-19. doi: 10.1128/mBio.03122-19

- Krüger L, Herzberg C, Wicke D, Bähre H, Heidemann JL, Dickmanns A, Schmitt K, Ficner R, Stülke J. A meet-up of two second messengers: the c-di-AMP receptor DarB controls (p)ppGpp synthesis in *Bacillus subtilis*. *Nat Commun*. 2021 Feb 22;12(1):1210. doi: 10.1038/s41467-021-21306-0. PMID: 33619274; PMCID: PMC7900238.
- Mata AR, Pacheco CM, Cruz Pérez JF, Sáenz MM, Baca BE. In silico comparative analysis of GGDEF and EAL domain signaling proteins from the *Azospirillum* genomes. *BMC Microbiol*. 2018 Mar 9;18(1):20. doi: 10.1186/s12866-018-1157-0.
- Mascher, T., Helmann, J. D. & Uden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiology and Molecular Biology Reviews*, 70(4), 910-+. doi: <https://doi.org/10.1128/MMBR.00020-06>
- Maxwell Burroughs, D. Z. (2015). Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Research*, 10633–10654.
- Michiels, K., de Troch, P., Onyeocha, I., van Gool, A., Elmerich, C., & Vanderleyden, J. (1989). Plasmid localization and mapping of two *Azospirillum brasilense* loci that affect exopolysaccharide synthesis. *Plasmid*, 21(2), 142–146. [https://doi.org/10.1016/0147-619X\(89\)90058-9](https://doi.org/10.1016/0147-619X(89)90058-9)
- Monteiro LMO, Sanches-Medeiros A, Westmann CA, Silva-Rocha R. Unraveling the Complex Interplay of Fis and IHF Through Synthetic Promoter Engineering. *Front Bioeng Biotechnol*. 2020 Jun 18;8:510. doi: 10.3389/fbioe.2020.00510.
- Motaleb MA, Miller MR, Li C, Charon NW. Phosphorylation assays of chemotaxis two-component system proteins in *Borrelia burgdorferi*. *Methods Enzymol*. 2007;422:438-47. doi: 10.1016/S0076-6879(06)22022-6.
- Nicastro, G.G., Kaihami, G.H., Pulschen, A.A. *et al.* c-di-GMP-related phenotypes are modulated by the interaction between a diguanylate cyclase and a polar hub protein. *Sci Rep* 10, 3077 (2020). <https://doi.org/10.1038/s41598-020-59536-9>
- Newell PD, Monds RD, O'Toole GA. LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1. *Proc Natl Acad Sci U S A*. 2009 Mar 3;106(9):3461-6. doi: 10.1073/pnas.0808933106
- O'Neal L, Vo L, Alexandre G. Specific Root Exudate Compounds Sensed by Dedicated Chemoreceptors Shape *Azospirillum brasilense* Chemotaxis in the Rhizosphere. *Appl Environ Microbiol*. 2020 Jul 20;86(15):e01026-20. doi: 10.1128/AEM.01026-20. PMID: 32471917
- Orr MW, Galperin MY, Lee VT. Sustained sensing as an emerging principle in second messenger signaling systems. *Curr Opin Microbiol*. 2016 Dec;34:119-126. doi: 10.1016/j.mib.2016.08.010.
- Paul K, Nieto V, Carlquist WC, Blair DF, Harshey RM. The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. *Mol Cell*. 2010 Apr 9;38(1):128-39. doi: 10.1016/j.molcel.2010.03.001
- Patterson DC, Ruiz MP, Yoon H, Walker JA, Armache JP, Yennawar NH, Weinert EE. Differential ligand-selective control of opposing enzymatic activities within a bifunctional c-di-GMP enzyme. *Proc Natl Acad Sci U S A*. 2021 Sep 7;118(36):e2100657118. doi: 10.1073/pnas.2100657118.

- Pérez-Mendoza D, Aragón IM, Prada-Ramírez HA, Romero-Jiménez L, Ramos C, Gallegos MT, Sanjuán J. Responses to elevated c-di-GMP levels in mutualistic and pathogenic plant-interacting bacteria. *PLoS One*. 2014 Mar 13;9(3):e91645. doi: 10.1371/journal.pone.0091645.
- Pieterse, C. M. J., Zamioudis, C., Berendsen, R. L., Weller, D. M., van Wees, S. C. M., & Bakker, P. A. H. M. (2014). Induced Systemic Resistance by Beneficial Microbes. *Annual Review of Phytopathology*, 52(1), 347–375. <https://doi.org/10.1146/annurev-phyto-082712-102340>
- Stijn Spaepen, Jos Vanderleyden, Yaacov Okon, Chapter 7 Plant Growth-Promoting Actions of Rhizobacteria, *Advances in Botanical Research*, Academic Press, Volume 51, 2009, Pages 283-320, ISSN 0065-2296, ISBN 9780123748348, [https://doi.org/10.1016/S0065-2296\(09\)51007-5](https://doi.org/10.1016/S0065-2296(09)51007-5)
- Povolotsky TL, Hengge R. 2016. Genome-Based Comparison of Cyclic Di-GMP Signaling in Pathogenic and Commensal *Escherichia coli* Strains. *J Bacteriol* 198: <https://doi.org/10.1128/jb.00520-15>
- Prigent-Combaret C, Zghidi-Abouzid O, Effantin G, Lejeune P, Reverchon S, Nasser W. The nucleoid-associated protein Fis directly modulates the synthesis of cellulose, an essential component of pellicle-biofilms in the phytopathogenic bacterium *Dickeya dadantii*. *Mol Microbiol*. 2012 Oct;86(1):172-86. doi: 10.1111/j.1365-2958.2012.08182.x.
- Rao F, Yang Y, Qi Y, Liang Z 2008. Catalytic Mechanism of Cyclic Di-GMP-Specific Phosphodiesterase: a Study of the EAL Domain-Containing RocR from *Pseudomonas aeruginosa*. *J Bacteriol* 190: <https://doi.org/10.1128/jb.00165-08>
- Römling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. *Microbiol Mol Biol Rev* 77: <https://doi.org/10.1128/mmb.00043-12>
- Russell MH, Bible AN, Fang X, Gooding JR, Campagna SR, Gomelsky M, Alexandre G. Integration of the second messenger c-di-GMP into the chemotactic signaling pathway. *mBio*. 2013 Mar 19;4(2):e00001-13. doi: 10.1128/mBio.00001-13.
- Rybtke, M., Chua, S.L., Yam, J.K.H., Givskov, M., Yang, L., Tolker-Nielsen, T. (2017). Gauging and Visualizing c-di-GMP Levels in *Pseudomonas aeruginosa* Using Fluorescence-Based Biosensors. In: Sauer, K. (eds) *c-di-GMP Signaling. Methods in Molecular Biology*, vol 1657. Humana Press, New York, NY. [https://doi.org/10.1007/978-1-4939-7240-1\\_8](https://doi.org/10.1007/978-1-4939-7240-1_8)
- Sierra Cacho D, Zamorano Sánchez DS, Xiqui-Vázquez ML, Viruega Góngora VI, Ramírez-Mata A, Baca BE. CdgC, a Cyclic-di-GMP Diguanylate Cyclase of *Azospirillum baldaniorum* Is Involved in Internalization to Wheat Roots. *Front Plant Sci*. 2021 Oct 20;12:748393. doi: 10.3389/fpls.2021.748393
- Stecher B, Hapfelmeier S, Müller C, Kremer M, Stallmach T, Hardt WD. Flagella and chemotaxis are required for efficient induction of *Salmonella enterica* serovar Typhimurium colitis in streptomycin-pretreated mice. *Infect Immun*. 2004 Jul;72(7):4138-50. doi: 10.1128/IAI.72.7.4138-4150.2004.
- Szwajkajzer D., Dai L., Fukayama J.W., Abramczyk B., Fairman R., Carey J. Quantitative analysis of DNA binding by the *Escherichia coli* arginine repressor. *J. Mol. Biol.* 2001; **312**:949–962.
- Stülke J, Krüger L. Cyclic di-AMP Signaling in Bacteria. *Annu Rev Microbiol*. 2020 Sep 8;74:159-179. doi: 10.1146/annurev-micro-020518-115943.

- Takaya, A., Erhardt, M., Karata, K., Winterberg, K., Yamamoto, T. and Hughes, K.T. (2012), YdiV: a dual function protein that targets FlhDC for ClpXP-dependent degradation by promoting release of DNA-bound FlhDC complex. *Molecular Microbiology*, 83: 1268-1284. <https://doi.org/10.1111/j.1365-2958.2012.08007.x>
- Valentini, Martina, filloux, Alain. Multiple roles of c-di-GMP signaling in bacterial pathogenesis. In: *Annual review of microbiology*, 2019, vol. 73, p. 387–406. doi: 10.1146/annurev-micro-020518-115555.
- Viruega-Góngora VI, Acatitla-Jácome IS, Zamorano-Sánchez D, Reyes-Carmona SR, Xiqui-Vázquez ML, Baca BE, Ramírez-Mata A. The GGDEF-EAL protein CdgB from *Azospirillum baldaniorum* Sp245, is a dual function enzyme with potential polar localization. *PLoS One*. 2022 Nov 23;17(11):e0278036. doi: 10.1371/journal.pone.0278036
- Wadhams, G. H., & Armitage, J. P. (2004). Making sense of it all: bacterial chemotaxis. *Nature Reviews Molecular Cell Biology*, 5(12), 1024–1037. <https://doi.org/10.1038/nrm1524>
- Wei Q, Ma LZ. Biofilm Matrix and Its Regulation in *Pseudomonas aeruginosa*. *International Journal of Molecular Sciences*. 2013; 14(10):20983-21005. <https://doi.org/10.3390/ijms141020983>
- Yu Z, Zhang W, Yang H, Chou SH, Galperin MY, He J. Gas and light: triggers of c-di-GMP-mediated regulation. *FEMS Microbiol Rev*. 2023 Jul 5;47(4):fuad034. doi: 10.1093/femsre/fuad034.
- Zhaoqing Yu, W. Z.-H. (2023). Gas and light: triggers of c-di-GMP-mediated regulation. *FEMS Microbiology Reviews*, 1–30.

## 11.-Anexos

### Diseño de una estrategia de mutación del gen *cdgG* en *A. baldaniorum* Sp245

Con base en los resultados bioinformáticos obtenidos en este trabajo, se diseñó una estrategia de mutación del gen *cdgG* en *A. baldaniorum* Sp245.

Esta mutación se llevará a cabo mediante reemplazo alélico, al intercambiar el ORF del gen *cdgG* por un gen reportero *mCherry* y un gen de resistencia a kanamicina. Para esto, se utilizará el plásmido suicida pJMS-Km. Una vez identificado el promotor de *cdgG* a través de métodos bioinformáticos, se amplificarán por PCR dos fragmentos de homología de 700 pb, que flanquean el ORF del gen de interés. Los brazos de homología a amplificar se denominarán fragmento A que abarcará una región de 700 pb río arriba del ORF de *cdgG*. El fragmento B, abarcará 700 pb río debajo del ORF de *cdgG*. Para esta estrategia se utilizará el kit de New England Biolabs, NEBuilder HiFi DNA Assembly. Este método permite la unión continua de múltiples fragmentos de ADN, sin importar la longitud del fragmento o la compatibilidad de los extremos. Se ha utilizado para ensamblar oligonucleótidos de cadena simple o fragmentos de ADN de diferentes tamaños con superposiciones variables (15- 30 pb). El kit incluye diferentes enzimas que trabajan en el mismo tampón o buffer.

A) La exonucleasa crea extremos protuberantes que facilitan la hibridación de fragmentos que comparten complementariedad de bases en un extremo.

B) La polimerasa rellena los huecos dentro de cada fragmento hibridado.

C) La ligasa de ADN sella las muescas en el ADN ensamblado.

El resultado final es una molécula de ADN bicatenario completamente sellado que puede servir como plantilla para PCR o una variedad de otras aplicaciones de biología molecular, incluida la transformación directa en *E. coli*.

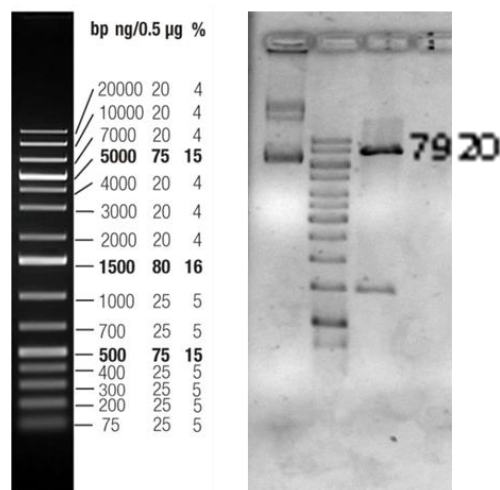
Los primers para la PCR deben tener dos componentes: una secuencia de superposición, requerida para ensamblar fragmentos adyacentes, y una secuencia específica del gen requerida para la preparación de la plantilla durante la PCR. La secuencia superpuesta no cebadora se añade en el extremo 5' del primer, siendo homóloga al fragmento adyacente para ser ensamblado. La longitud

dependerá de la cantidad GC de las secuencias. A continuación, se presenta la tabla de los primers requeridos para amplificar cada fragmento, así como para llevar a cabo el ensamble (**Figura 21**).

Nombre	longitud	Secuencia	% GC	Tm
CdgG-fragA-F	38	GGTACCGTTAACTCGAGCCCTGGTCATGGCGCCCAGCG	66	75
CdgG-fragA-R	28	TGCTCACCATGCTTCGTCCACCGCCGGC	68	76
MCheKm-F	28	TGGACGAAGCATGGTGAGCAAGGGCGAG	61	60.6
mCheKm-R	32	AGGGGATCAATCAGAAGAACTCGTCAAGAAGG	59	56
CdgG-fragB-F	28	GTTCTTCTGATTGATCCCCTCCCCTGCG	71	76.6
CdgG-fragB-R	39	AGCTCAGGCCTACGCGTCCCGGTGTCGGAGTTTTCGAG	64	66.4

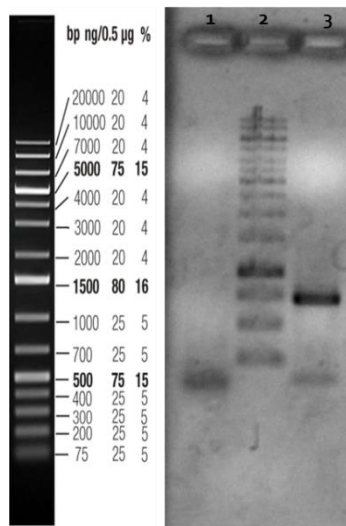
**Figura 21.** Cebadores utilizados para NEBuilder HIFI. El esquema se muestra en colores para mostrar de forma más simplificada la homología que deben compartir entre cebadores

El vector por utilizar es el pJMS-Km que posee sitios de corte para la enzima de restricción SmaI para poder linealizarlo. Para tal fin, se realizó la extracción del vector y se comprobó su autenticidad a través de los cortes con SmaI (**Figura 22**).



**Figura 22.** Electroforesis en gel de agarosa. Carril 1 Plásmido pJMS-Km extraído por miniprep. Carril 2 marcador de peso molecular de ADN de una Kb. Carril 3 Plásmido pJMS-Km digerido con la enzima SmaI.

Una vez que los primers se sintetizaron, se procedió a realizar la PCR con la Q5 HotStar High Fidelity polymerase. El producto esperado a amplificar es de 700 pb tanto para el fragmento A y B (Figuras 23 y 24)

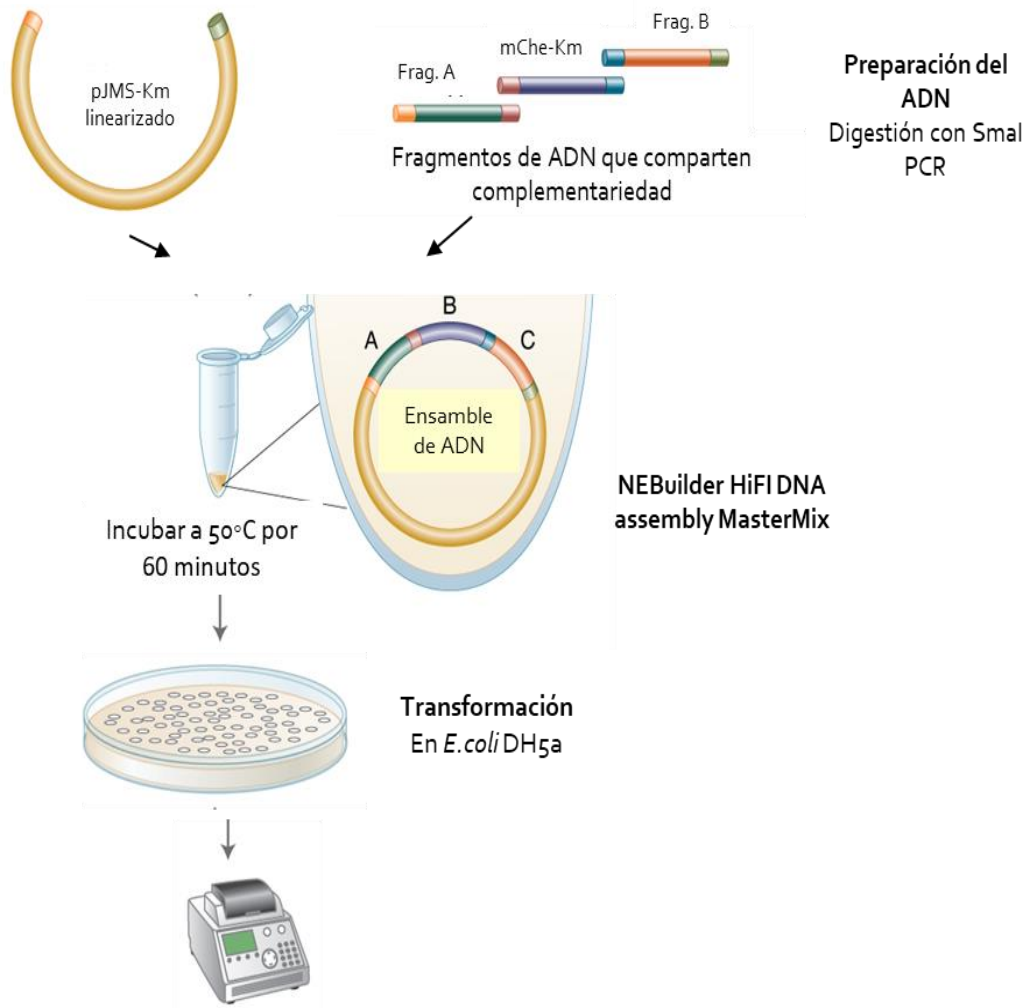


**Figura 23:** PCR de alta fidelidad del fragmento A:  
1: Control, 2: MPM, 3: Fragmento A



**Figura 24:** PCR de alta fidelidad del fragmento B y Cherry-Km: 1: Fragmento B, 2: MPM, 3: mCherry-Km

Una vez llevado a cabo lo anterior, procedimos a realizar el ensamblaje de acuerdo con el diagrama mostrado en la (Figura 25).

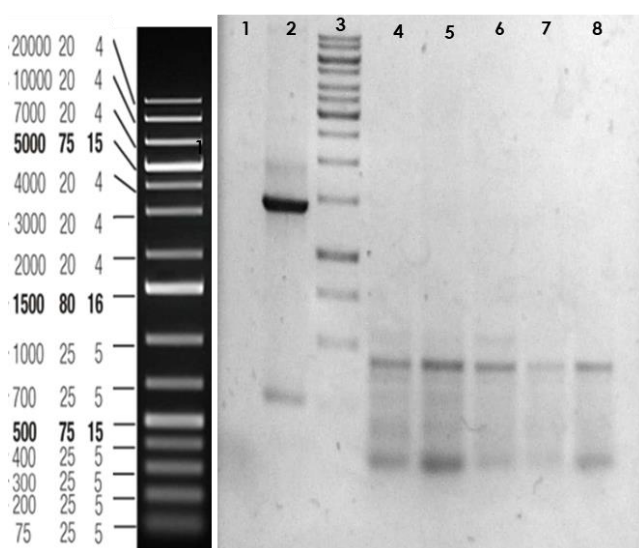


**Figura 25.** Metodología de NEBuilder HiFi DNA assembly.

Posteriormente se realizó una PCR en colonia para confirmar que se realizó el ensamblaje de forma correcta, la cual se transformó en células competentes de *E. coli* DH5 $\alpha$ . Se manejaron las condiciones mostradas en la **Tabla 1**. El producto de PCR se muestra en la **figura 26**.

**Tabla 1. Reactivos para PCR en colonia**

Reactivos	Volumen 1X	Volumen 6x
Agua para PCR	2 $\mu$ l	12 $\mu$ l
DreamTaq	5 $\mu$ l	30 $\mu$ l
CdgG-fragA-F	0.5 $\mu$ l	3 $\mu$ l
CdgG fragB-R	0.5 $\mu$ l	3 $\mu$ l



**Figura 26.** Producto de PCR en colonia donde el producto es de 1407 pb. Carril 1 control, carril 2 control positivo, carril 3 MPM, carril 4 al 8 producto de PCR de las clonas transformadas.

Tras varios ensayos de PCR, los amplificados resultantes eran de menor tamaño, por lo que no correspondía a lo esperado, dado que se esperaba un fragmento de amplificación de 3,400 pb. Con base en estos resultados, es necesario llevar a cabo ensayos adicionales para obtener la construcción deseada.