



---

---

Benemérita Universidad Autónoma de Puebla  
Facultad de Medicina

**Tesis:**

Diseño de estrategias para la detección de proteínas que interactúan con la diguanilato  
ciclase A de *Azospirillum brasilense* Sp7

**Para Obtener el Título:**

Licenciada en Biomedicina

**Alumna:**

Zaira Melina López Juárez

**Asesor Experto:**

D.C. Alberto Ramírez Mata

**Asesor Metodológico:**

D.C. Beatriz E. Baca

---

---

Diciembre, 2018

---

---



**BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA.**  
**FACULTAD DE MEDICINA.**  
**COORDINACIÓN DE EFICIENCIA TERMINAL EN PREGRADO**

**FORMATO "A" DE AUTORIZACIÓN DE TEMA.**  
**LICENCIATURA EN BIOMEDICINA.**

**DATOS DEL SOLICITANTE.**

Utilizar los renglones para señalar los datos que se indican en relación con el solicitante:  
**Nombre Completo:** Zaira Melina López Juárez      **Matrícula:** 201229150  
**Campus:** Puebla      **Correo Electrónico:**  
zairam2005@hotmail.com  
**Domicilio Actual:** Primera Privada Puebla #3, Col. Independencia      **Teléfono:** 2227644908

**Firma:** \_\_\_\_\_

**NOMBRE DEL TEMA:**

"Diseño de estrategias para la detección de proteínas que interactúan con la diguanilato ciclasa A de *Azospirillum brasilense* Sp7"

**DIRECTOR EXPERTO:**

**Nombre:** D.C. Alberto Ramírez Mata      **No. De trabajador:** 100521262  
**Especialidad:** Doctor en Ciencias      **Correo Electrónico:** alberto.ramirez@correo.buap.mx  
**Domicilio:** Av. San Claudio, Cd Universitaria, Instituto en Ciencias Microbiológicas.  
**Teléfono:** 2295500 ext. 2519

**Firma:** \_\_\_\_\_

**OBJETIVO GENERAL:**

Diseñar estrategias que permitan detectar proteínas que interactúan con la proteína CdgA de *Azospirillum brasilense* Sp7.

**JUSTIFICACIÓN:**

Los niveles de di-GMPc tienen un papel importante en los diferentes estilos de vida de las bacterias, uno fundamental es la transición de célula plantónica a célula sésil y viceversa, por ello, entender el mecanismo por el cual este segundo mensajero interviene en las funciones de las células es relevante. El estudio del sistema de señalización del di-GMPc permitirá conocer y comprender las interacciones, adaptaciones metabólicas y fisiológicas del organismo en respuesta al entorno. El diseño de estrategias para determinar la participación de la proteína CdgA en la cascada de señalización del di-GMPc permitirá responder las siguientes preguntas, ¿Las diguanilato ciclasas intervienen en las vías regulatorias de transición de estilo de vida y no sólo participar en la catálisis del di-GMPc? ¿Habrán diferencias de interacción proteína-proteína entre dos estilos de vida bacteriana?

**TIPO DE ESTUDIO:** Experimental de ciencia básica

**INSTITUCIÓN EN LA QUE SE REALIZARÁ:** Centro de Investigación en Ciencias Microbiológicas, ICUAP- BUAP

**No. de Folio de Registro:**

**Director Metodológico:** D.C. Alberto Ramírez Mata      **Teléfono:** 2295500 ext. 2519  
**Horario de Localización:** 10 am a 3 pm  
**Domicilio:** Av San Claudio, Cd Universitaria, Instituto en Ciencias Microbiológicas.  
**Correo electrónico:** alberto.ramirez@correo.buap.mx

**Fecha y Firma de Autorización:** \_\_\_\_\_



**COORDINADORA DE TITULACIÓN Y E.T**  
**MASS IRMA ORTEGA SANCHEZ**



**BUAP**

Oficio No SIEP / C.I. / 067/2018  
Asunto: Constancia de Registro

D.C. ALBERTO RAMÍREZ MATA  
D.C. BEATRIZ EUGENIA BACA  
D.C. LUCÍA SOTO URZÚA  
ZAIRA MELINA LÓPEZ JUÁREZ

**PRESENTES:**

El Comité de Investigación de la Facultad de Medicina de la B.U.A.P., a través de la Secretaría de Investigación y Estudios de Posgrado hace **C O N S T A R** que el Proyecto de Investigación presentado en autoría Colectiva por:

- ZAIRA MELINA LÓPEZ JUÁREZ
- D.C. ALBERTO RAMÍREZ MATA
- D.C. BEATRIZ EUGENIA BACA
- D.C. LUCÍA SOTO URZÚA

Titulado:

"DISEÑO DE ESTRATEGIAS PARA LA DETECCIÓN DE PROTEÍNAS QUE INTERACTÚAN CON LA DIGUANILATO  
CICLASA DE AZOSPIRILLUM BRASILENSE SP7"

Ha sido registrado en esta Secretaría con los siguientes datos:

Fecha de registro: 28 de mayo de 2018.

Número de Libro: 2

Número de Hoja: 64

Número de Registro: 579

Vigencia:

Inicio 28 de mayo 2018

Termino 28 de septiembre de 2018

ATENTAMENTE

"PENSAR BIEN, PARA VIVIR MEJOR"

H. PUEBLA DE Z., A 28 DE MAYO DE 2018

M.C. JOSE LUIS GANDARA RAMIREZ  
PRESIDENTE DEL COMITÉ DE INVESTIGACIÓN

D.C. JORGE ALEJANDRO CEBADA RUIZ  
SECRETARIO DEL COMITÉ DE INVESTIGACIÓN

c.c.p. archivo  
c.c.p. minutarío  
cc "CEC" "vial"





**BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA**  
**FACULTAD DE MEDICINA**  
**COORDINACIÓN DE EFICIENCIA TERMINAL EN PREGRADO**

**FORMATO "C" DE AUTORIZACIÓN DE TESIS**  
**LICENCIATURA EN BIOMEDICINA**

**DATOS DEL SOLICITANTE.**

Nombre Completo: Zaira Melina López Juárez

Matricula: 201229150

Correo Electrónico: zairam2005@hotmail.com

Teléfono: 2227644908

Firma: 

Folio otorgado por el comité de investigación FMBUAP: 010/2018

**TÍTULO DE LA TESIS:**

"Diseño de estrategias para la detección de proteínas que interactúan con la diguanilato ciclasa A de *Azospirillum brasilense* Sp7"

**ÁREA DEL ESTUDIO:** Microbiología

Presentar la tesis elaborada en computadora, con letra Arial, numero 12 a espacio y medio firmada por los asesores de la Tesis y coordinación de titulación (original y cuatro portadas).

Director Experto de la Tesis: D. C. Alberto Ramírez Mata

No. de Trabajador: 100521262

Teléfono: 2295500 Ext. 2519  
alberto.ramirez@correo.buap.mx

Correo Electrónico:

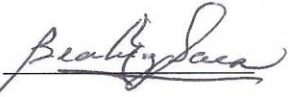
Firma de autorización: 

Director Metodológico de la Tesis: D.C. Beatriz Eugenia Baca

No. de Trabajador: 100004344

Teléfono: 2295500 Ext. 2525  
beatriz.baca@correo.buap.mx

Correo Electrónico:

Firma de autorización: 

Revisor de la Tesis. D.C Lucía Soto Urzúa

No. de Trabajador: 100101922

Teléfono: 2295500 Ext. 2535

Correo Electrónico: lucia.soto@correo.buap.mx

Firma de autorización: 

Visto bueno y se autorizará la impresión de la Tesis.

  
COORDINADORA DE TITULACION Y E/T  
MASCARRA IRMA ORTEGA SANCHEZ

Fecha de Autorización





---

---

Benemérita Universidad Autónoma de Puebla  
Facultad de Medicina

**Tesis:**

Diseño de estrategias para la detección de proteínas que interactúan con la diguanilato  
ciclase A de *Azospirillum brasilense* Sp7

**Para Obtener el Título:**

Licenciada en Biomedicina

**Alumna:**

Zaira Melina López Juárez

**Asesor Experto:**

D.C. Alberto Ramírez Mata

**Asesor Metodológico:**

D.C. Beatriz E. Baca

---

---

Diciembre, 2018

*Agradecimientos...*

*“Aunque mi alma se ponga en tinieblas, se alzaré en perfecta luz; he amado mucho las estrellas para ser temeroso de la noche...” - Sarah Williams*

*Agradezco a mi papá y mamá, que me han amado, regañado, consentido, escuchado, aconsejado y, sobre todo, han demostrado que la perseverancia trae grandes recompensas. Ustedes me han enseñado tantas cosas, que creo que una vida entera no bastaría para agradecerles y honrarles todo lo que han hecho por mí. Sin ustedes, yo hoy no estaría aquí.*

*A mis hermanos Raphael y Mario quienes con su constante apoyo, liderazgo y amor me han guiado a cumplir otro reto más. Sin sus sabios consejos y grandes retos estilo “lucha libre”, hoy no sería la persona que soy. Me forjaron para ser una mujer fuerte :3*

*A mi familia materna y paterna por acobijarme entre sus brazos y darme palabras de aliento para seguir por más difícil que estuviera el camino, en especial a mis abuelitas (Lupita y Amalia) que me han criado y cuidado desde pequeña, y a mi abuelito Temo quien, aunque ya no esté presente sé que me cuida y guía desde el cielo.*

*A mis amigos, Yuyis, Isaías y Pao, no sé qué hubiera hecho sin ustedes para levantarme cada vez que pensaba que iba a tropezar y caer... en verdad, ¡los amo! Ya pasamos la preparatoria juntos y ahora la universidad, ahora nos toca toda una vida por delante.*

*A mis amigos de la carrera, Soph, Adri, Anny, Ivonne, Law, Paquito, Ángel, Job, y al honorable traidor Jorge, ustedes hicieron que 2 años y medio volaran entre risas y lágrimas con el proclamado “Club de los Poetas de la Muerte” mis pequeños y valientes biscochos. Decirles que los amo no es suficiente, porque en casi seis años han logrado que nuestra amistad a pesar de la distancia siga día tras día igual o más fuerte, están en mi corazón y espero sigan en mi vida muchos, pero muchos años por venir...*

*A mis amigos y colegas del Laboratorio de la Interacción-Bacteria Planta... Iris, Rox, Dan, Paquito, Daniel, Ricardo, Iván, Jocelyn, Ibelith, Andy y Saul... ustedes me han acompañado a lo largo de los últimos años y se han vuelto como una segunda familia para mí. Y bueno, Miguel, Miguelito, hermanito de laboratorio :’3 nunca creí volverme cercana a ti a pesar de cursar la misma carrera, pero la vida te da muchas sorpresas y tu fuiste una de ellas. Sin ti, no hubiera hecho muchas cosas y probablemente mis experimentos no hubieran funcionado. En verdad aprecio tanto su ayuda, apoyo, enseñanzas y buenos momentos que he compartido con ustedes. Por esto y más les agradezco a cada uno, me quedo con un pedacito de cada uno, ¡gracias!*

*A los Dr. Alberto, Dra. Baca y Mtra. Mary por sus enseñanzas, sus críticas y sus consejos para poder llevar acabo este trabajo. Aprendí muchísimo de ustedes y por ello, quedaré infinitamente agradecida.*

*A la VIEP de la Benemérita Universidad Autónoma de Puebla, por autorizar el proyecto 100521262\_VIEP2018 que financió este trabajo.*

*A todos... les agradezco este momento de mi vida...*

*Agradecimientos*

*Esta investigación fue realizada gracias al apoyo del Consejo de Ciencia y Tecnología del  
Estado de Puebla.*

## Contenido

Definiciones y Abreviaturas .....	10
Resumen .....	13
1.- Introducción .....	14
1.1.- El di-GMPc .....	14
1.2.- Enzimas diguanilato sintasas (DGC) y fosfodiesterasas (PDE) .....	16
1.3.- Diguanilato Ciclasas .....	17
1.5.- Dominios PAS .....	20
2.- Antecedentes .....	22
2.1.- <i>Azospirillum brasilense</i> .....	22
2.2.- Biopelícula .....	22
2.3.- Interacciones proteína-proteína entre DGCs y efectores del di-GMPc .....	24
3.- Antecedentes directos .....	26
4.- Hipótesis .....	28
5.- Justificación .....	28
6.- Objetivos .....	28
6.1.- Objetivo General .....	28
6.2.- Objetivos Particulares .....	28
7.- Esquema de Trabajo .....	29
8.- Material y Métodos .....	30
8.1.- Material biológico .....	30
8.2.- Inducción de la proteína recombinante CdgA .....	30
8.3.- Lisis celular y purificación de la proteína CdgA por cromatografía de afinidad .....	30
8.4.- Crecimiento celular de <i>A. brasilense</i> Sp7 en estado platónico y sésil .....	31
8.5.- Lisis celular de extractos de <i>A. brasilense</i> Sp7 .....	32
8.6.- Acoplamiento de la proteína CdgA a la resina de afinidad Ni-NTA para ensayos Pull Down .....	32
8.7.- Ensayos Pull Down (interacción proteína-proteína) .....	33
8.8.- Producción de anticuerpos policlonales en contra de la proteína CdgA recombinante .....	33
8.9.- Geles de poliacrilamida desnaturizantes o SDS-PAGE al 10% .....	34
8.10.- Western Blot: Transferencia a membrana .....	35
8.11.- Western Blot: Inmunodetección anti-CdgA (anticuerpos policlonales de conejo) .....	36
9.- Resultados .....	37

<b>9.1.- Modificación de los protocolos de inducción, lisis y purificación de la proteína CdgA recombinante .....</b>	<b>37</b>
<b>9.2.- Ensayos Pull Down.....</b>	<b>42</b>
<b>9.3.- Anticuerpos policlonales Anti-CdgA de <i>A. brasilense</i> Sp7 .....</b>	<b>45</b>
<b>9.4.- Análisis bioinformático estructural de la proteína CdgA de <i>A. brasilense</i> Sp7.....</b>	<b>49</b>
<b>9.4.1- Modelo estructural del dominio GGDEF de la proteína CdgA .....</b>	<b>49</b>
<b>9.4.2.- Análisis estructural comparativo entre los dominios GGDEF de las Diguanylateo ciclasas GcbC de <i>Pseudomonas fluorescens</i>, PleD de <i>C. vibrioides</i> y CdgA de <i>A. brasilense</i> Sp7 .....</b>	<b>51</b>
<b>9.5.- Análisis bioinformático de interacción proteína-proteína de la diguanilato ciclasas CdgA con otras proteínas de <i>A. brasilense</i> Sp7 por medio del programa bioinformático STRING .....</b>	<b>54</b>
<b>9.5.1.- Predicción de interacción proteína-proteína sometiendo el contexto genético de CdgA de <i>A. brasilense</i> Sp7 .....</b>	<b>55</b>
<b>9.5.2.- Predicción de interacción proteína-proteína utilizando el banco de datos del programa STRING .....</b>	<b>56</b>
<b>9.5.3.- Predicción de interacción proteína-proteína de CdgA y las presuntas DGCs <i>A. brasilense</i> Sp7.....</b>	<b>60</b>
<b>9.5.4.- Predicción de interacción proteína-proteína retando el contexto genético para las proteínas de dominio híbrido GGDEF-EAL de <i>A. brasilense</i> Sp7.....</b>	<b>66</b>
<b>9.5.5.- Predicción de interacción proteína-proteína de CdgA con PDEs de dominio EAL de <i>A. brasilense</i> Sp7 .....</b>	<b>69</b>
<b>10.- Discusión .....</b>	<b>78</b>
<b>11.- Conclusión .....</b>	<b>85</b>
<b>12.- Perspectivas .....</b>	<b>85</b>
<b>13.- Referencias .....</b>	<b>86</b>
<b>14.- ANEXOS.....</b>	<b>90</b>

## Definiciones y Abreviaturas

- Ácidos grasos. – Biomolécula lipídica formada por un grupo carboxilo seguido de una cadena larga hidrocarbonada lineal con diferentes longitudes o número de átomos de carbono unidas por medio de enlace covalente sencillo o doble.
- Ácidos orgánicos. – Son compuestos que poseen al menos un grupo ácido.
- Alfa-hélices. – Estructura secundaria de las proteínas que mantiene su forma por la presencia de los puentes de hidrógeno que se forman entre los átomos de oxígeno del grupo carbonilo de un aminoácido y el átomo de hidrógeno del grupo amino de otro aminoácido, optando por una estructura helicoidal dextrógira.
- Aminoácidos. – Molécula orgánica conformada por un átomo de carbono alfa unido a un hidrógeno, un grupo R o cadena lateral, con un grupo amino (-NH<sub>2</sub>) y un grupo carboxilo (-COOH).
- Antirabbit-IgG-HRP conjugado. – Anticuerpo conjugado con peroxidasa generado en cabra en contra de IgG de conejo.
- Anticuerpo. – Es una proteína producida por el sistema inmunitario del cuerpo cuando detecta sustancias dañinas, llamadas antígenos.
- Antígeno. – Es cualquier sustancia que provoca que el sistema inmunitario produzca anticuerpos contra sí mismo.
- Auxinas. – Grupo de fitohormonas que actúan como reguladoras del crecimiento vegetal.
- Bacteria. – Organismo microscópico unicelular, carente de núcleo, que se multiplica por división celular sencilla o por esporas.
- Bioinformática. – Es una disciplina que permite aplicar herramientas de informática al estudio y la gestión de datos de la biología.
- Bioquímica. – Parte de la química que estudia los elementos que forman parte de la naturaleza de los seres vivos.
- Buffer. – Solución amortiguadora o reguladora constituido por un ácido débil y una base conjugada, o por una base y su ácido conjugado que tiene capacidad amortiguadora.
- Catálisis. – Variación en la velocidad de una reacción química producida por la presencia de un catalizador químico o biológico.
- Citoplasma. – Parte de la célula que rodea el núcleo y que está limitada por la membrana externa.
- Diguanilato ciclasa. – Enzima catalizadora del di-GMPc
- Dímero. – Especie química que consiste en dos subunidades estructuralmente similares denominadas monómeros unidas por enlaces que pueden ser fuertes o débiles.
- Dominio proteico. – Término genérico que designa una región de una proteína de interés biológico funcional o estructural.
- DO . – Densidad óptica.
- Enlaces fosfodiéster. – Enlace covalente que se produce entre un grupo fosfato (H<sub>3</sub>PO<sub>4</sub>) y un grupo hidroxilo (-OH), se puede encontrar en los fosfolípidos y nucleótidos.
- Enzimas. – Proteínas que catalizan reacciones químicas de ciertas sustancias, acelerando el proceso.
- Espectrofotometría. – Método científico utilizada para medir la absorbancia de luz de una sustancia química cuando un haz luminoso pasa a través de la solución muestra.
- Estructura secundaria. – Plegamiento regular local entre residuos aminoácidos de la cadena polipeptídica de las proteínas. Da lugar a la formación de alfa hélices, beta plegadas y giros o azas.
- Exopolisacáridos (ESP). – Son macromoléculas de carbohidratos, sintetizadas y secretadas por bacterias.
- Fitohormonas. – Son sustancias conocidas como hormonas vegetales, cuya función es similar a las hormonas presentes en los animales. Regulan la fisiología de la planta.
- Flavina. – Es una base nitrogenada cuya cadena principal es una sustancia heterocíclica nitrogenada de tres anillos y dos grupos oxo, conocida como isoaloxazina. Su nombre viene el latín flavus o amarillo, ya que su forma oxidada es amarilla y su forma reducida es incolora.
- Fosfodiesterasa. – Enzimas encargadas de hidrolizar las uniones formadas por los enlaces fosfodiéster del di-GMPc. Hay dos tipos, aquellas con dominio EAL y aquellas con dominio HD-GYP.

- Gen. – Segmento corto de ADN que funciona como unidad de información ya que codifica para un producto funcional como una macromolécula con función celular específica.
- GTP. – Mejor conocida como guanosín trifosfato, nucleótido trifosfatado utilizado en el metabolismo celular.
- Hojas beta-plegadas. – Estructura secundaria que puede adoptar la cadena polipeptídica por el posicionamiento paralelo de dos cadenas de aminoácidos dentro de la misma proteína, en las cuales los grupos N-H de una cadena forman enlaces de hidrógeno con los grupos C=O de la opuesta.
- h. –Horas
- Imidazol. – Intermediario de la biosíntesis de la histidina que se forma desde el imidazol glicerol fosfato con la pérdida de agua. Pertenece a la familia de compuestos químicos aromáticos, caracterizada a disponer de un anillo pentagonal nitrogenado.
- IPTG. – Reactivo isopropil- $\beta$ -D-1-tiogalactopiranosido utilizado comúnmente en la investigación como análogo no hidrolizable de la alolactosa para inducir la expresión génica en el laboratorio.
- Microorganismo. – Ser vivo o sistema biológico que sólo puede visualizarse con el microscopio.
- mM . – milimolar
- Molécula. – Agrupación definida y ordenada de átomos que constituye la porción más pequeña de una sustancia pura y conserva todas sus propiedades.
- M . – molar
- $\mu$ g . – microgramo
- $\mu$ l . – microlitro
- mg . – miligramo
- ml . – mililitro
- Monómero. – Unidad básica de una molécula.
- Motivo. – Secuencia de aminoácidos conservados asociada a una función en concreto en una proteína.
- Nucleótido. – Compuesto químico orgánico fundamental de los ácidos nucleicos, constituido por una base nitrogenada, un azúcar y una molécula de ácido fosfórico unida por enlaces covalentes.
- PGPR. – Por sus siglas en inglés que significa Rizobacteria Promotora del Crecimiento Vegetal.
- Proteína. – Son macromoléculas formadas por cadenas lineales de aminoácidos con diversas funciones celulares.
- Resina de afinidad Ni-NTA . – resina de afinidad a níquel ácido nitrilotriacético (Qiagen 30210)
- rpm . – revoluciones por minuto.
- Seg . – segundo.
- Segundo mensajero. – Toda molécula que transduce señales hasta producir un cambio fisiológico en un efector dentro de la célula.
- Sustrato (bioquímica). – Molécula sobre la cual actúa una enzima



## Resumen

Las diguanilato ciclasas (DGC) son enzimas encargadas de sintetizar al di-GMPc, molécula señal utilizada para regular diversas funciones fisiológicas en las bacterias a través del uso de riboswitches o proteínas efectoras que son el blanco de este dinucleótido. Este grupo de enzimas ha sido estudiado en diversos modelos bacterianos y se caracteriza por la presencia del dominio GGDEF, el cual se encarga de la síntesis del dinucleótido; y la presencia o no de dominios accesorios que contribuyen a regular la actividad catalítica. Sin embargo, poco se conoce sobre otras funciones que pudieran estar empleando las DGC en relación con la red de señalización del di-GMPc.

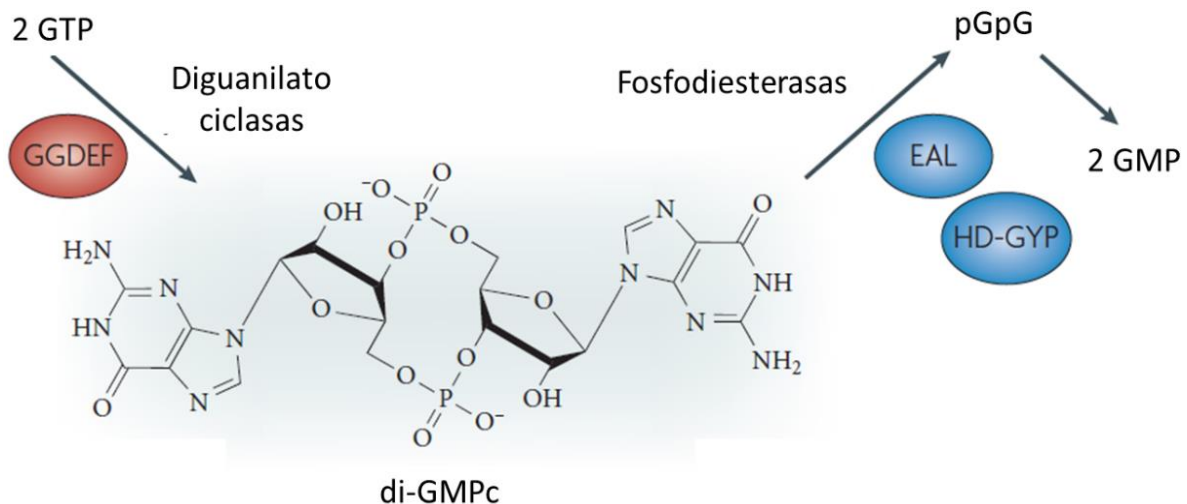
El diseño de técnicas que permitan la detección de interacción proteína-proteína para la DGC CdgA de *A. brasilense* Sp7 permitió la identificación de posibles proteínas que pudieran estar implicadas en la red de señalización del di-GMPc en dos diferentes estados de crecimiento de la bacteria (células plantónicas y células sésiles) a un día de crecimiento. Por otro lado, se detectó la presencia de la proteína CdgA en células plantónicas y células sésiles a cuatro diferentes tiempos de crecimiento mediante el uso de anticuerpos policlonales en contra de CdgA. Finalmente, se obtuvo un modelo estructural por homología del dominio GGDEF confiable de esta, que permite respaldar los resultados obtenidos de interacción proteína-proteína.

# 1.- Introducción

## 1.1.- El di-GMPc

El di-GMPc o di-3', 5' guanosín-monofosfato es una molécula de señalización intracelular considerada como segundo mensajero involucrado en regular diferentes funciones de las bacterias, entre las cuales están: Regulación de secreción de adhesinas, señalización celular, transición de estilo de vida, plantónico a sésil o formación de biopelícula, virulencia, movilidad, entre otros. Fue descrito por Benziman y colaboradores en la década de los 80's como un activador alostérico de la síntesis de celulosa en la bacteria *Gluconacetobacter xylinus* (Ross, *et al.*, 1987; D'Argenio & Miller, 2004).

Esta molécula se encuentra constituida por dos moléculas de GTP (guanosín trifosfato), las cuales son cicladas y unidas por enlaces fosfodiéster, perdiendo dos grupos pirofosfato en el proceso. La catálisis de la reacción es realizada por enzimas denominadas diguanilato ciclasas, las cuales cuentan con un dominio característico denominado dominio GGDEF que le permite anclar al sustrato y llevar a cabo el proceso de síntesis. Por otro lado, las enzimas responsables de degradar el di-GMPc son denominadas fosfodiesterasas y se consideran dos tipos, aquellas con dominio EAL que catalizan la hidrólisis de un enlace fosfodiéster y linearizan el di-GMPc a una molécula pGpG o aquellas con dominios HD-GYP que eliminarán ambos enlaces fosfodiéster del di-GMPc produciendo dos moléculas de guanosín monofosfato (GMP) como se puede observar en la Figura 1 (Schirmer, *et al.* 2009).



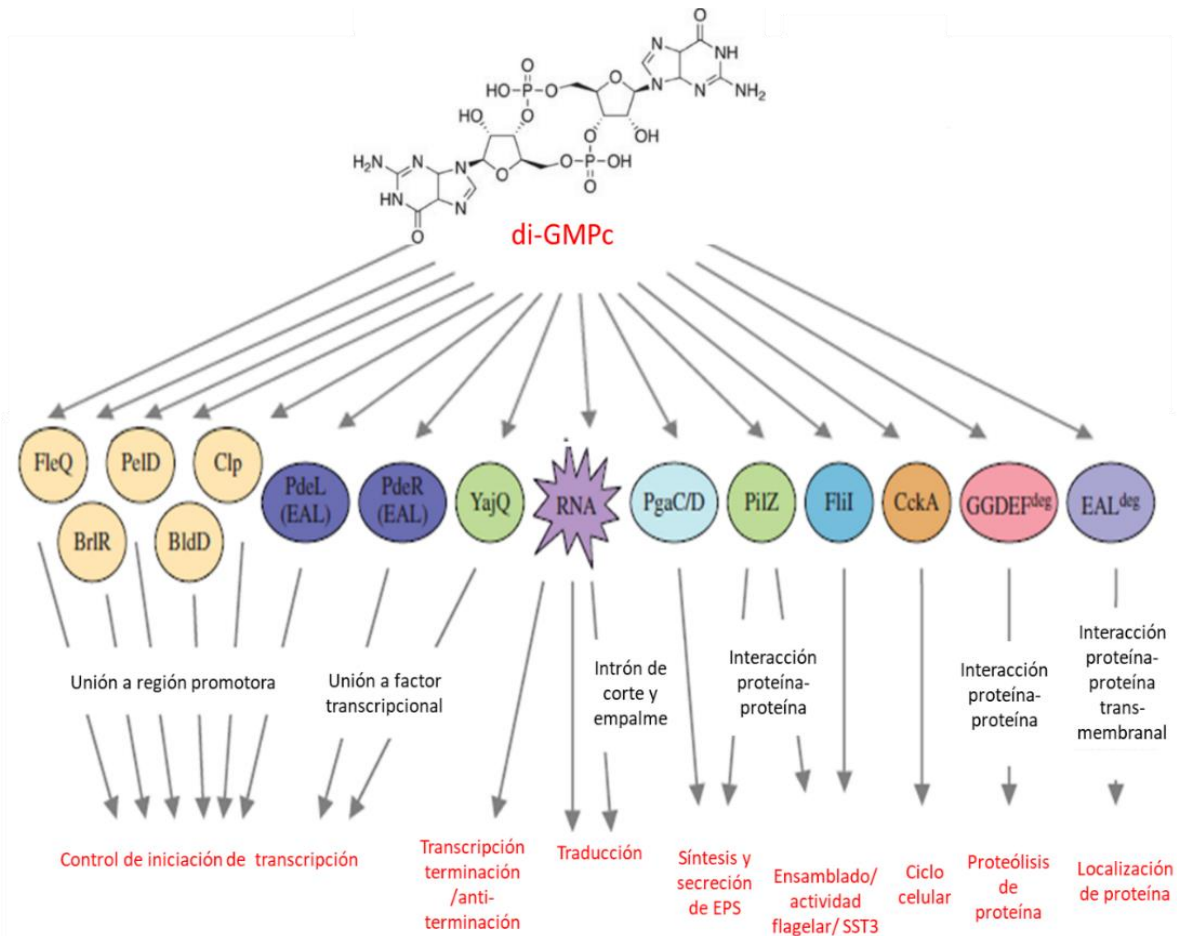
**Figura 1.- Síntesis y degradación del di-GMPc.** A partir de dos moléculas de GTP como sustrato, las diguanilato ciclasas (óvalo rojo) sintetizarán una molécula de di-GMPc. Por otro lado, las fosfodiesterasas

(óvalo azul) se encargarán de linealizar el sustrato di-GMPc en moléculas lineares como pGpG o 2 moléculas de GMP, dependiendo el dominio catalítico que posean (Imagen tomada y modificada de Hengge, 2009).

Se le considera como un segundo mensajero universal en bacterias debido a que ha sido encontrado en gran parte de éstas y se ha relacionado a esta molécula con diversas características de las células bacterianas como ya fue descrito (Feirer, *et al.*, 2015).

Por lo que, se ha reportado que la diversidad de moléculas efectoras que pueden estar respondiendo a la interacción con el di-GMPc, para su activación o represión, intervienen a nivel transcripcional, post-transcripcional, traduccional y post-traduccional. Algunos de estos ejemplos se muestran en la Figura 2, donde se presenta por categorías una gran diversidad de moléculas efectoras (riboswitches, reguladores transcripcionales, proteínas con dominios PilZ y proteínas con dominios GGDEF o EAL degenerados, entre otras), que interactúan con el segundo mensajero para llevar a cabo diversas tareas celulares implicadas en control de transcripción, síntesis y secreción de EPS (exopolisacáridos), ensamblaje flagelar, ciclo celular entre otros. Sin embargo, se desconoce la cantidad de moléculas efectoras o mecanismos celulares en los cuales el segundo mensajero esté involucrado (Römling, *et al.*, 2013; Hengge, 2016).

Por otro lado, la localización del dinucleótido a nivel celular es desconocida. Estudios realizados en *Pseudomonas aeruginosa* y otros modelos bacterianos, proponen dos mecanismos teóricos posibles para la localización del di-GMPc intracelular. La primera teoría sugiere la existencia de una poza global que distribuya y module la actividad o represión a todos los efectores de manera general, mientras que el segundo mecanismo involucraría la existencia de una concentración subcelular del dinucleótido, en una zona específica de la bacteria de distribución y modulación a un efector específico, lo que es denominado micro-compartmentación (Dahlstrom, *et al.*, 2015).



**Figura 2.- El di-GMPc y ejemplos de moléculas efectoras implicadas en mecanismos celulares a nivel transcripcional, traduccional o post-traduccional.** Debido a que el di-GMPc participa extensamente en la fisiología celular de las bacterias, la diversidad de efectores presentes que involucren a esta molécula señal varían dependiendo del microorganismo. Algunos ejemplos de proteínas reguladas alostéricamente por parte del di-GMPc son FleQ, PelD, regulador BrlR presentes en *Pseudomonas aeruginosa*. El regulador de virulencia Clp en *Xanthomonas campestris* y regulador maestro del desarrollo BldD de *Streptomyces*. YajQ y PilZ son pequeñas proteínas que unen al di-GMPc, son adaptadores versátiles o factores importantes entre la unión del di-GMPc y complejos de proteínas efectoras con diversas funciones. El operón PgaC/D es un sistema de secreción y sintetasa del exopolisacárido PGA en *Escherichia coli*, ambas proteínas mantienen actividad al unir el di-GMPc. FliI une al di-GMPc en el cuerpo basal del flagelo como un exportador de ATPasas, compartiendo cierta similitud con el sistema de secreción Tipo III en *Pseudomonas fluorescens*. Las proteínas con dominios totalmente degenerados GGDEF o EAL son inactivas catalíticamente, sin embargo, presentan la capacidad de unión al di-GMPc y funcionan como moléculas efectoras del di-GMPc como lo son PdeR y PdeL de *E. coli* (Modificada de Hengge, 2016).

## 1.2.- Enzimas diguanilato sintasas (DGC) y fosfodiesterasas (PDE)

La presencia de motivos y dominios característicos de las DGCs y PDEs se han determinado por medio de programas bioinformáticos, genéticos, bioquímicos y moleculares de enzimas cristalizadas. Las DGCs contienen el dominio GGDEF mientras que las PDE presentan dos tipos de dominio, EAL o HD-GYP. Las enzimas pueden

presentar sólo el dominio característico GGDEF, o bien el EAL y se consideran de dominio único. Un tercer grupo de enzimas que abarca un tercio del total de las enzimas, son denominadas proteínas híbridas ya que cuentan con un dominio GGDEF y un dominio EAL, generalmente, continuo en el C-terminal de la proteína. Los dominios pueden presentarse conservados, parcialmente conservados o totalmente degenerados, en uno o ambos dominios. Estas modificaciones producen como efecto que ambos dominios sean activos, que un dominio sirva como regulador para el otro, o que ambos dominios sean inactivos (Schirmer, *et al.*, 2009).

Los tres tipos de enzimas pueden presentar dominios accesorios en el N-terminal de la proteína, que les otorgan diversas funciones ya sea para activar, regular o inactivar la producción de di-GMPc. Los dominios accesorios permiten captar señales extra o intracelulares, y pueden estar ubicadas tanto en el citoplasma o estar ancladas a la membrana como proteínas de unión a membranas (Valentini, *et al.*, 2016).

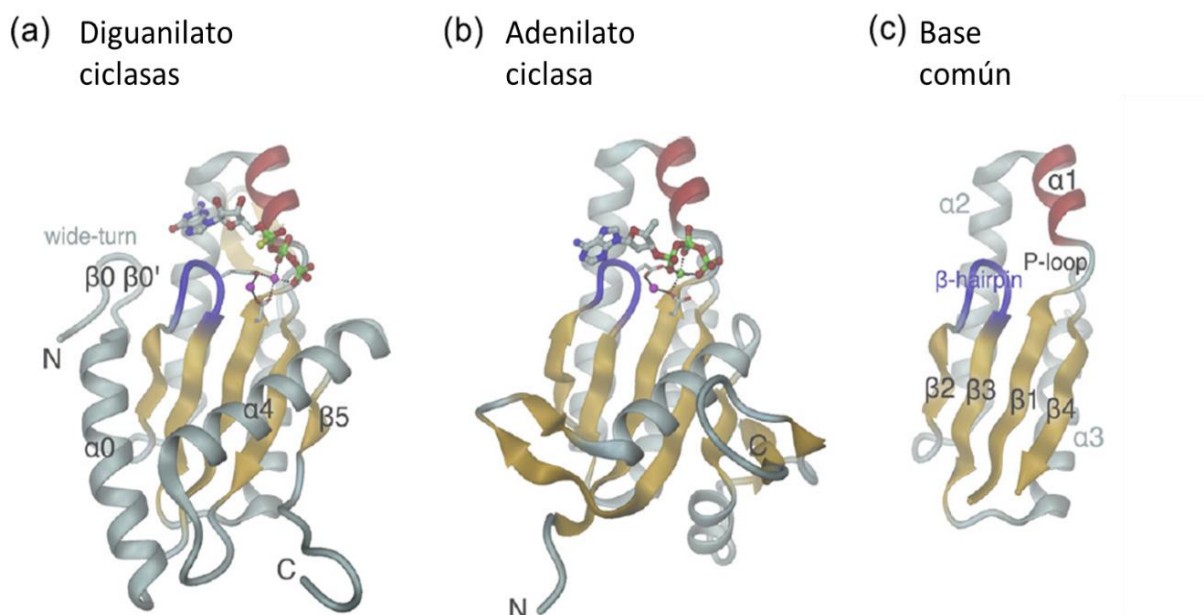
Algunos ejemplos de dominios accesorios son los dominios PAS, REC, GAF, HAMP, CACHE, CHASE, MASE y otros más, capaces de detectar diversas señales como: la concentración de oxígeno, potencial redox, luz, unión de mononucleótidos y receptores de respuestas de regulación, entre otras (Whiteley, *et al.*, 2015; Dahlstrom, *et al.*, 2017).

Así mismo, se ha podido correlacionar la cantidad de DGC y PDE presentes en un microorganismo con la capacidad de adaptación ambiental donde se desarrolle, entre más nichos pueda establecerse, más enzimas involucradas con la síntesis y regulación del di-GMPc se presentarán en el genoma del microorganismo (Krasteva, *et al.*, 2012).

### **1.3.- Diguanilato Ciclasas**

El dominio característico GGDEF presenta dos motivos importantes para la catálisis e inhibición de la enzima. El sitio A o motivo GG(D/E)EF es el sitio de unión al sustrato GTP con aminoácidos altamente conservados, y el sitio I o dominio RXXD se encuentra cinco aminoácidos río arriba del motivo GG(D/E)EF y está involucrado en la regulación de la actividad enzimática (Deepthi, *et al.*, 2014).

En el motivo GG(D/E)EF, la presencia del ácido aspártico o ácido glutámico permiten la coordinación de cationes divalentes que intervienen en la reestructuración del sitio activo e inician la catálisis enzimática mientras que la presencia de las glicinas permite que el sustrato GTP cuente con el espacio suficiente para su anclaje (Römmling, *et al.*, 2017).

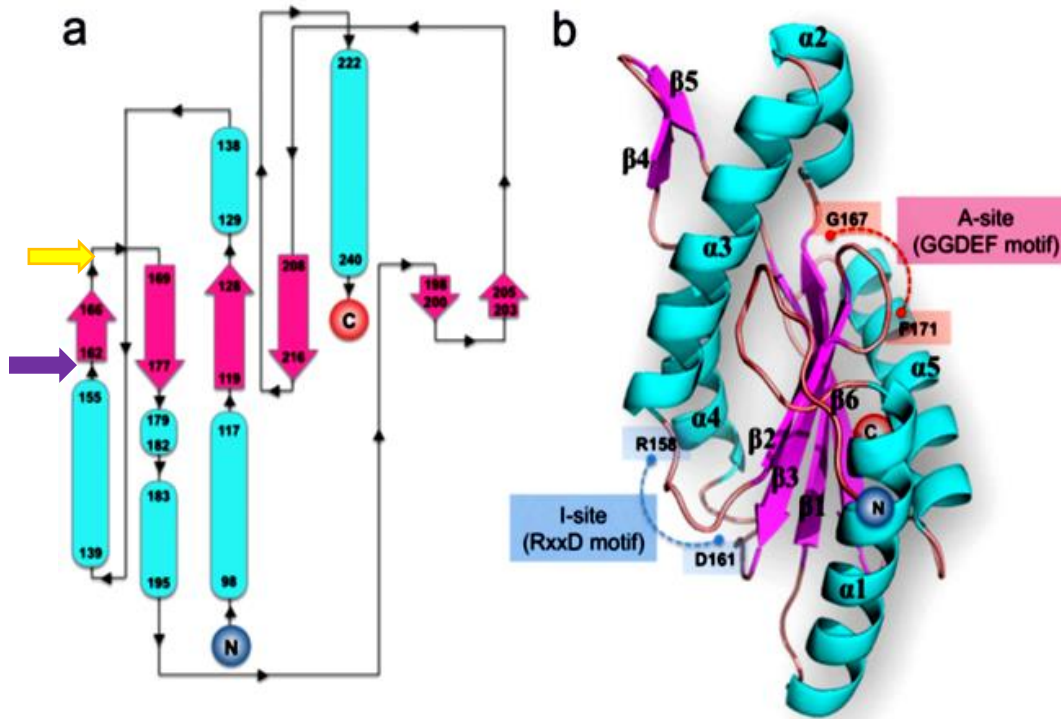


**Figura 3.- Similitud de la estructura secundaria de las enzimas Diguanylatociclasa y Adenilato ciclasa.**

**A)** Estructura conformacional del dominio GGDEF, las alfas-hélices se indican en gris, en amarillo las hojas beta-plegadas, el aza que posee el motivo catalítico se presenta en azul, en morado se presentan los iones de magnesio y la estructura del di-GMPc. **B)** La estructura de adenilato ciclasa indica a las alfas-hélices en gris, en amarillo las hojas beta-plegadas, el aza que posee el motivo catalítico se presenta en color azul, en morado se presenta el ion de magnesio, en verde el ion de calcio y la estructura del di-AMPc. **C)** Estructura similar (desarrollada a partir de las estructuras secundarias de diguanylatociclasa y adenilato ciclasa) conformada por cuatro hojas beta-plegadas, de las cuales destacan las hojas  $\beta_2$ - $\beta_3$  por poseer el motivo catalítico de las enzimas (Esquema modificado de Schirmer, 2016).

Los dominios GGDEF comparten una estructura secundaria similar a la presente en los dominios catalíticos de adenilato y guanylatociclasas, y comparten la presencia del motivo catalítico en la estructura secundaria central de  $\beta$ -plegadas. Análisis estructurales de varias proteínas cristalizadas han definido un núcleo altamente conservado de estructura secundaria con el siguiente orden de  $\alpha$ -hélices o  $\beta$ -plegadas,  $\beta\alpha\alpha\beta\beta\alpha\beta$  o  $\beta_1$ - $\alpha_1$ - $\alpha_2$ - $\beta_2$ - $\beta_3$ - $\alpha_3$ - $\beta_4$ , en el cual se localizan la presencia de los motivos importantes para la catálisis, proporcionado por el aza formada entre las dos hojas  $\beta$ -plegadas consecutivas centrales (Figura 3, aza azul). En las DGCs el motivo GG(D/E)EF se presenta en el lazo formado entre las hojas  $\beta_2$  y  $\beta_3$ -plegadas centrales y cinco aminoácidos río arriba, se forma otra aza

entre la  $\alpha_x$ - $\beta_2$  para el motivo RXXD. Aunque no todas las DGCs presentan la misma cantidad de  $\alpha$ -hélices o  $\beta$ -plegadas, presentan una secuencia similar respecto a la presencia del motivo GG(D/E)EF en el aza formado entre dos hojas  $\beta$ -plegadas consecutivas centrales como se puede observar en la Figura 4. (Whiteley, *et al.*, 2015; Schirmer, 2016).

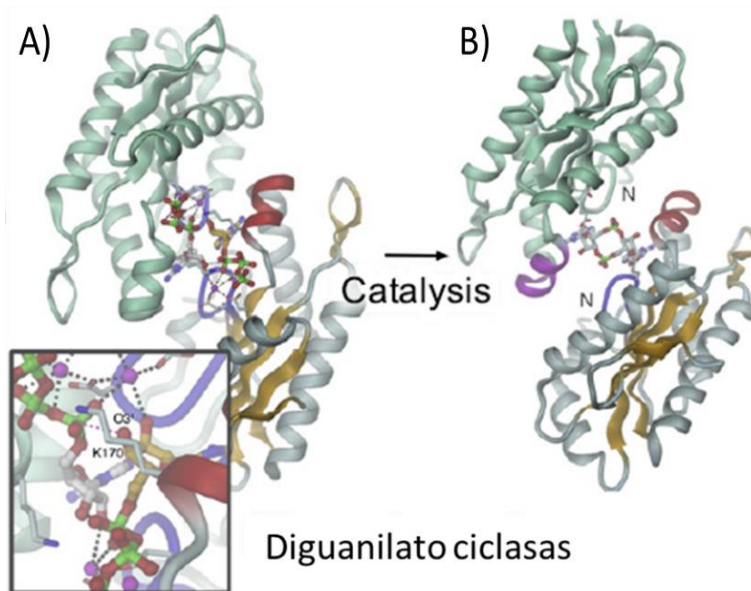


**Figura 4.- Arquitectura de la proteína tDGC de *Thermotoga maritima*.** A) Topología del dominio GGDEF de la proteína tDGC. El N-terminal y C-terminal se indican en círculos. Las beta-plegadas están marcadas en color rosa y las alfa-hélices se presentan en color azul. La flecha amarilla señala al motivo GG(D/E)EF y la flecha morada al motivo RXXD. B) El motivo GG(D/E)EF o sitio A (caja rosa) entre el aza de las hojas  $\beta_2$  y  $\beta_3$ -plegadas y sitio I (motivo RXXD en caja azul) presente en el lazo formado entre  $\alpha_3$ -hélice y  $\beta_2$ -plegada de las 5  $\alpha$ -hélices y 6  $\beta$ -plegadas que conforman la estructura secundaria de la proteína tDGC (Tomado de Deepthi, *et al.*, 2014).

La actividad catalítica de las DGCs requiere de la formación de dímeros que les permita conducir la reacción entre los motivos GG(D/E)EF de ambos monómeros que incluyen el sustrato GTP. Esta dimerización puede estar regulada por el dominio GGDEF o por la presencia de dominios accesorios adjuntos en el N-terminal de la proteína que captaran diversas señales citoplasmáticas, transmembranales y periplásmicas y cuyos dominios se localizan en los diversos compartimentos celulares (Römling, *et al.*, 2017).

Por medio de programas bioinformáticos, la estructura dimérica de la DGC DgcZ de *Escherichia coli* fue obtenida a partir de la estructura dimérica post-catálisis unida al

producto de di-GMPc (Figura 5-B). En esta estructura se logra identificar la parcial participación de ambos sitios catalíticos de cada monómero de DgcZ así como la formación de puentes de hidrógeno entre ciertos aminoácidos que permite la estabilización del dímero. En la Figura 5-A, se muestra el dímero de DgcZ, donde cada monómero contiene el GTP unido al motivo A, correspondiente, así como los iones de magnesio (Schirmer, 2016).



**Figura 5.- Estructura secundaria dimérica de la proteína DgcZ.** En la Figura A se observa la presencia de dos moléculas de sustrato GTP unidas respectivamente al motivo GGDEF (aza en color azul), y la formación de puentes de hidrogeno y los iones  $Mg^{+2}$  y/o  $Mn^{+2}$  que estabilizan la unión de la enzima con el sustrato. En la Figura B, se observa al dímero de las DgcZ unido a la molécula de di-GMPc (Modificado de Schirmer, 2016).

Dependiendo de la función que realicen las DGC, los aminoácidos conservados presentes en el motivo catalítico permiten clasificar a las enzimas en tres grupos: enzimáticamente activas, proteínas híbridas con dominio conservado o degenerado, o enzimáticamente inactivas (Römling, *et al.*, 2017).

Los dominios PAS de manera frecuente se localizan en el N-terminal de las proteínas (DGCs y/o PDEs), funcionando como un regulador o modulador universal de señalización en bacterias en respuesta a los estímulos ambientales (Schirmer, 2016).

### 1.5.- Dominios PAS

Los dominios PAS han sido identificados en todos los organismos. Sus siglas en inglés (Per/Arnt/Sim) provienen de aquellas proteínas inicialmente identificadas y descritas con

estos dominios en diferentes organismos. Per proviene de la proteína del periodo circadiano, Arnt de la proteína translocadora del receptor nuclear Ah y Sim de la proteína que regula el desarrollo del sistema nervioso central de *Drosophila* (Hefti, *et al.*, 2004).

El dominio PAS se encuentra constituido por aproximadamente 100 aminoácidos y es un dominio sensor importante para interacciones proteína-proteína, transferencia o captación de señales que involucra una gran variedad de moléculas e iones como lo son los grupos hemo, flavinas, ácidos di- y tricarbónicos, aminoácidos, cationes divalentes metálicos, ácido cumárico, ácidos grasos entre otros aún no identificados. De igual forma, es capaz de percibir cambios de concentración de varios gases O<sub>2</sub>, y CO, la intensidad de luz y voltaje. Sin embargo, no todos los dominios PAS cuentan con la misma secuencia de residuos de aminoácidos, pero presentan una estructura secundaria similar que les permite ser subclasificada dependiendo la molécula señal con la cual interaccionan. Además, algunas proteínas pueden incluir uno o más dominios PAS continuos (Qi, *et al.*, 2009; Henry, *et al.*, 2011).

## **2.- Antecedentes**

### **2.1.- *Azospirillum brasilense***

*Azospirillum* pertenece a la subclase alpha de las proteobacterias, Gram negativas, de vida libre, fijadoras de nitrógeno pertenecientes al grupo de rizobacterias promotoras del crecimiento vegetal o PGPRs (cuyas siglas provienen del inglés Plant Growth-Promoting Rhizobacteria). Estas bacterias cuentan con la capacidad de realizar quimiotaxis positiva en dirección a compuestos secretados por raíces como lo son ácidos orgánicos, azúcares, aminoácidos y compuestos aromáticos, por medio de mecanismos de motilidad, quimiotaxis o aerotaxis, la bacteria puede dirigirse hacia la raíz de las plantas y lograr la colonización de ésta de presentarse óptimas condiciones. La colonización de las PGPRs genera cambios en la arquitectura de la raíz, así como el desarrollo de ramificaciones de ésta, incrementando la captación de nutrientes. Este desarrollo es dado gracias a la secreción de fitohormonas producidas por la bacteria. Las fitohormonas (auxinas) que produce la bacteria se encuentran directamente involucradas en el crecimiento de la planta como lo son el ácido indol acético o el ácido indol butírico (Bashan, *et al.*, 2010).

Sin embargo, para que los beneficios por parte de las bacterias de *Azospirillum* puedan ser otorgados, debe presentarse una interacción prolongada entre la bacteria y la planta. Los microorganismos tienden a formar biopelículas posteriormente de haber colonizado el nicho de interés para poder establecer una interacción más íntima entre ambos organismos (Lin, *et al.*, 2009; Trabelsi, *et al.*, 2013).

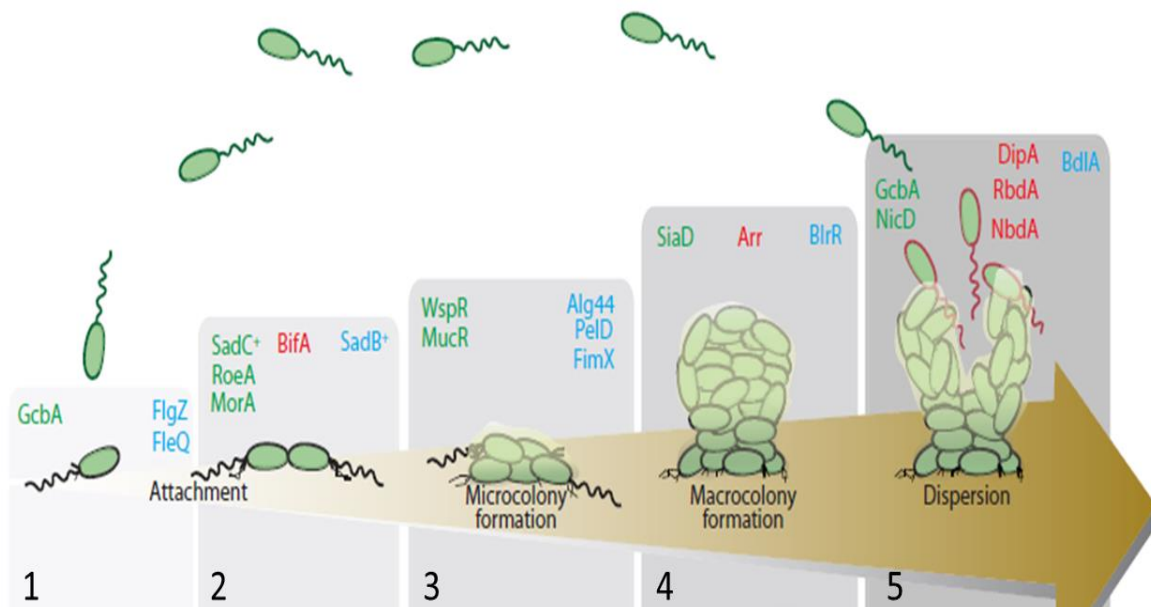
### **2.2- Biopelícula**

Las biopelículas son comunidades bacterianas inmersas en una matriz extracelular compuesta por sustancias poliméricas extracelulares como ADN, proteínas, exopolisacáridos y agua; estas comunidades se establecen en superficies que les permita desarrollarse. Las bacterias que crecen dentro de las biopelículas cuentan con la capacidad de modular su metabolismo para responder a la disponibilidad de nutrientes o gradientes de desechos y mantener una comunicación activa. La formación de estructuras comunitarias cumple con la función de proteger a los microorganismos del medio ambiente y proporcionar beneficios para su crecimiento y desarrollo. Factores ambientales como la

temperatura, concentración de electrolitos y disponibilidad de nutrientes juegan un papel importante en la integración de los microorganismos a estas comunidades (Ramey, *et al.*, 2004; Branda, *et al.*, 2005; Tallawi, *et al.*, 2017).

Para una óptima formación de biopelícula, se deben presentar dos etapas previas que forman parte de la colonización a un sitio o huésped específico. La primera etapa consiste en la adherencia a una superficie inerte o viva mediante interacciones fisicoquímicas y moleculares. Es una etapa reversible ya que, al no encontrar los nutrientes necesarios, el microorganismo puede desprenderse e ir en búsqueda de otros nichos. La segunda etapa es una adherencia irreversible específicamente para el desarrollo de la colonización, donde se presenta un incremento en producción de exopolisacáridos que refuerzan la adherencia celular (Loera Muro, *et al.*, 2012).

Posterior a los pasos de colonización, se presenta la formación de micro-colonias celulares estables y el inicio de intercambio de sustratos entre los organismos. Después, se genera la maduración de las micro-colonias, las células se comunican por medio de señales auto-inductoras y la secreción de exopolisacáridos se realiza para construir una matriz protectora que le otorga estructura tridimensional y permite el intercambio de comunicación, así como de nutrientes. Por último, los organismos de la biopelícula se reproducen y liberan células hijas, transitando de organismos sésiles a móviles, en búsqueda de otros nichos para establecerse y repetir el ciclo. (Jamal, *et al.*, 2017). En la Figura 6, Valentini, *et al.*, (2016) presentan el esquema de biopelícula integrando a DGCs, PDEs y proteínas efectoras (nombres en color verde, rojo y azul respectivamente) previamente estudiadas. Como se puede observar, la importancia de estas proteínas se presenta en diferentes etapas de desarrollo y formación de la biopelícula (Valentini, *et al.*, 2016).



**Figura 6.- Acción coordinada del mecanismo de señalización del di-GMPc en la formación de biopelícula en *P. aeruginosa*.** En color verde se presentan las DGCs, en color rojo las PDEs y en color azul los receptores o efectores de di-GMPc, así mismo las etapas de formación de biopelícula: 1) Anclaje reversible a la superficie de interés por medio del flagelo polar; 2) Anclaje irreversible por las células y la formación de una monocapa en superficie; 3) Formación de micro colonias; 4) Formación de macro colonias; 5) Dispersión (Modificado de Valentini, *et al.*, 2016).

### 2.3.- Interacciones proteína-proteína entre DGCs y efectores del di-GMPc

Se ha reportado que DGCs y PDEs intervienen de manera compleja en la red de señalización del di-GMPc.

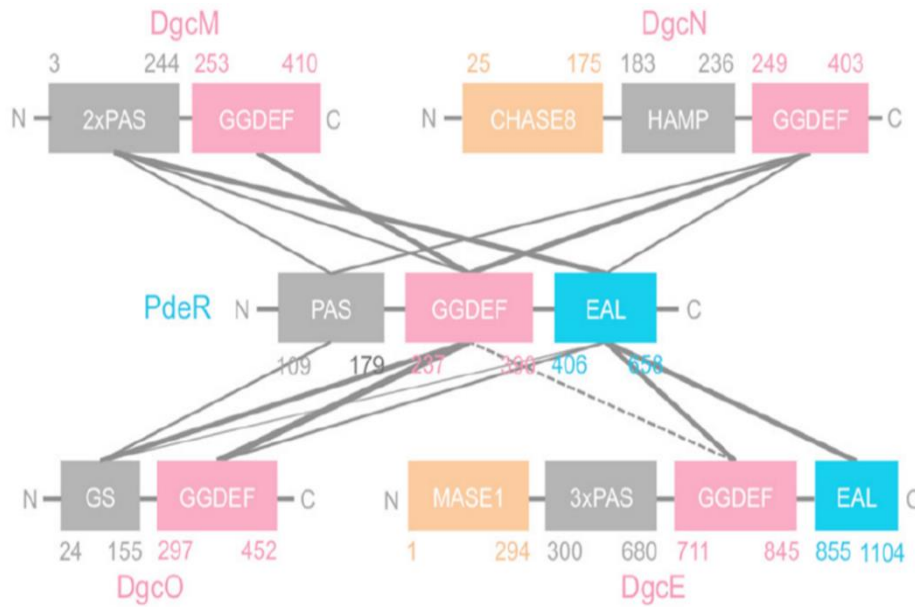
En *Xanthomonas campestris*, se describió la interacción física entre la PDE RpfG y dos DGCs (XC\_0249 y XC\_0420), de las cuales se sugiere en el estudio que participan en la regulación de factores de virulencia y/o movilidad. Dicho estudio apunta a la participación de las DGCs y PDEs, no solo en la síntesis y degradación del di-GMPc, sino también en una participación en la red de señalización del segundo mensajero de forma más compleja (Ryan, *et al*, 2010).

Dahlstrom, *et al.*, (2015) describieron la interacción física de una DGC GcbC y la proteína efectora LapD de *Pseudomonas fluorescens*. Al unirse el di-GMPc a LapD, esta interacción permite el secuestro de la proteasa periplásmica LapG y con ello, la acumulación de la

adhesina LapA en la membrana externa de *P. fluorescens* para la posterior formación de biopelícula bacteriana. Estos autores demuestran que la  $\alpha_5$ -hélice con la siguiente secuencia de aminoácidos <sup>477</sup>EQLLFAADK<sup>485</sup> de GcbC es importante para la interacción con la  $\alpha_2$ -hélice del dominio EAL de la proteína LapD con la secuencia <sup>462</sup>GRFLPWLER<sup>470</sup>. El di-GMPc producido por GcbC es conducido a LapD, proteína con el dominio EAL degenerado, pero que aún liga al sustrato di-GMPc, este complejo así formado permite la interacción con la proteasa para inactivar su función. Así mismo, señalan la importancia del ácido aspártico presente en la secuencia de la  $\alpha_5$ -hélice como uno de los aminoácidos más importantes que permite la interacción con otras proteínas (Dahlstrom, *et al.*, 2015).

Por otro lado, se demostró la importancia de un sitio inhibitorio constituido por los residuos de aminoácidos conservados: E360, R366 y E429 y el motivo RXXD de la proteína GcbC para ejercer interacción con la proteína LapD de *P. fluorescens*. Mutaciones en los aminoácidos anotados conlleva a la pérdida de la capacidad de unir al di-GMPc y la interacción con LapD (Dahlstrom, *et al.*, 2016).

En el estudio realizado en la cepa *E. coli* K-12 se presenta las interacciones proteína-proteína entre DGCs y PDEs (interacciones homodiméricas o heterodiméricas). Esta bacteria cuenta con 29 proteínas involucradas en la síntesis y degradación del di-GMPc, donde a pesar de la multiplicidad, estas proteínas generan funciones específicas y distintivas lo cual sugiere actividad localizada con efectores específicos. Los resultados corresponden a señalar una interacción no específica entre DGCs, no solo en los dominios característicos de las proteínas, sino también por los dominios accesorios con los cuales cuentan en su amino terminal como se puede observar en la Figura 7. Si bien se ha definido que el di-GMPc es sintetizado por homodímeros de DGCs, alguna otra función regulatoria podría tener la interacción de estas proteínas heterodimerizadas (Sarenko, *et al.*, 2017).



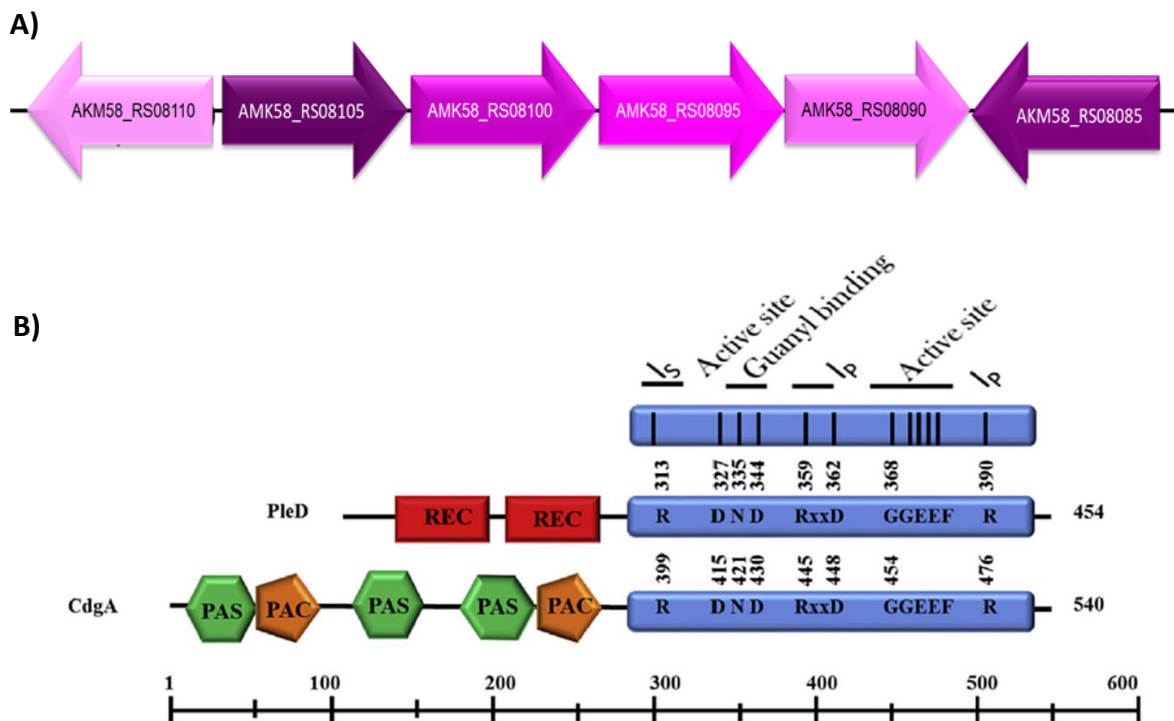
**Figura 7.- Interacción entre dominios catalíticos y dominios accesorios de una proteína fosfodiesterasa (PdeR) y cuatro diguanilato ciclasas (DgcM, DgcO, DgcN y DgcE) de *E. coli* K-12.** Las líneas gruesas, delgadas y punteadas indican la diferencia de fuerza presente de la interacción como resultado del ensayo de doble híbrido (Imagen tomada de Sarenko, *et al.*, 2017).

Entre mayor presencia de enzimas involucradas en la síntesis y degradación del di-GMPc, más complejo es el estudio del segundo mensajero y de la cascada de señalización que se establece en cada modelo bacteriano. Por ello, como resultado se proponen dos categorías para las proteínas con dominios GG(D/E)EF o EAL: (I) proteínas conectadas por múltiples interacciones involucradas en la red de señalización del di-GMPc y (II) enzimas solitarias que no intervienen con, o se asocian a otras DGCs o PDEs para modular su función. Las proteínas de esta última categoría podrían ya sea intervenir en modular los niveles de di-GMPc o interactuar con otros efectores o sistemas no reconocidos que unan al di-GMPc (Sarenko, *et al.*, 2017).

### 3.- Antecedentes directos

Ramírez-Mata, *et al.*, (2016), describieron la actividad de una DGC de *A. brasilense* Sp7, denominada CdgA, codificada por el gen *cdgA* (AMK58\_RS08100), cuyo contexto genético indicó que el gen *cdgA* comparte la misma dirección que otros 3 genes, como se observa en la Figura 8-A. El producto de traducción de la secuencia de los genes indicó que dos de las presuntas proteínas contienen dominios PAS. Además, la proteína CdgA cuenta con tres dominios accesorios PAS/PAC adicionales al dominio GGDEF, este último

conservado al ser comparado con la DGC PleD, ampliamente estudiada de *Caulobacter crescentus* (Figura 8-B).



**Figura 8.- Contexto genético en el cual se presenta el gen que codifica para la diguanilato ciclasa A y la comparación de dominios presentes entre la diguanilato ciclasa ya caracterizada PleD de *Caulobacter Crescentus* y CdgA de *Azospirillum brasilense* Sp7. A) AMK58\_RS08110 (Proteína de biosíntesis de cobalamina CobD), AMK58\_RS08105 (treonina-fosfato decarboxilasa), AMK58\_RS08100 (DgcA), AMK58\_RS08095 (PAS box), AMK58\_RS08090 (Proteína quimiotáctica aceptora de grupos metilos), AMK58\_RS08085 (Cinasa histidílica). B) Los motivos RXXD y GG(D/E)F se encuentran altamente conservados, así como la presencia de dominios accesorios (dominios PAS) (Ramírez-Mata, *et al.*, 2016).**

Para evaluar el efecto que tendría la ausencia del gen *cdgA* de la cepa *A. brasilense* Sp7, se realizó la mutación por delección del gen *cdgA* que codifica para la proteína CdgA. Una vez obtenida la mutante (1428-A) se prosiguió a realizar ensayos fenotípicos con la finalidad de analizar la producción de EPS y formación de biopelícula. Los resultados arrojaron una disminución en la formación de biopelícula y producción de EPS en la cepa mutante 1428-A en comparación a la cepa silvestre. La presencia de los genes que se localizan en la misma dirección del gen *cdgA* es sugerente de una participación en la red de señalización del di-GMPc (Ramírez-Mata, *et al.*, 2016).

## **4.- Hipótesis**

Los dominios que presenta la proteína CdgA de *A. brasilense* Sp7 (dominios accesorios PAS-PAC y dominio GGDEF) permiten la interacción con otras proteínas involucradas en la cascada de señalización del di-GMPc.

## **5.- Justificación**

Los niveles de di-GMPc tienen un papel importante en los diferentes estilos de vida de las bacterias, uno fundamental es la transición de célula planctónica a célula sésil y viceversa, por ello, entender el mecanismo por el cual este segundo mensajero interviene en las funciones de las células es relevante. El estudio del sistema de señalización del di-GMPc permitirá conocer y comprender las interacciones, adaptaciones metabólicas y fisiológicas del organismo en respuesta al entorno. El diseño de estrategias para determinar la participación de la proteína CdgA en la cascada de señalización del di-GMPc permitirá responder las siguientes preguntas, ¿Las diguanilato ciclasas intervienen en las vías regulatorias de transición de estilo de vida y no sólo participar en la catálisis del di-GMPc? ¿Habrá diferencias de interacción proteína-proteína entre dos estilos de vida bacteriana?

## **6.- Objetivos**

### **6.1.- Objetivo General**

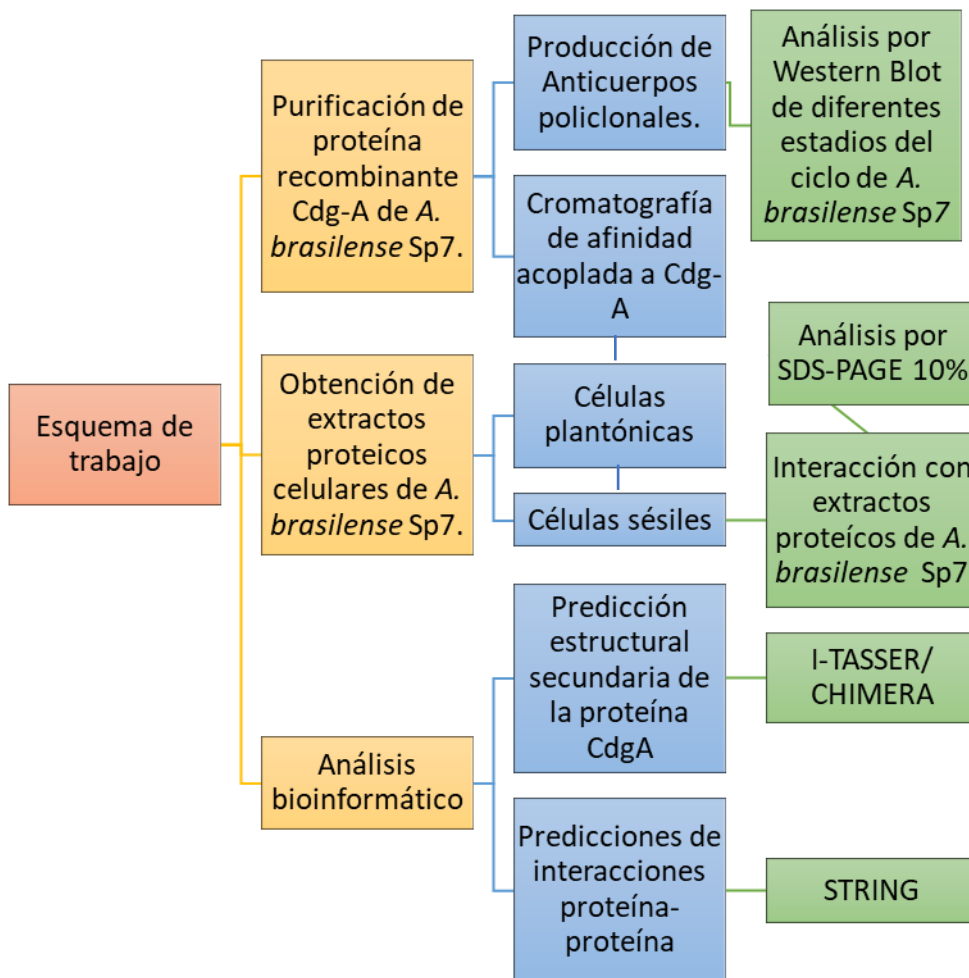
Diseñar estrategias que permitan detectar proteínas que interactúen con la proteína CdgA de *Azospirillum brasilense* Sp7.

### **6.2.- Objetivos Particulares**

1. Generar una cromatografía de afinidad acoplando la proteína CdgA recombinante de *E. coli* pLIL-2 a resina de afinidad Ni-NTA.
2. Realizar ensayos de interacción proteína-proteína, utilizando la resina de afinidad que tiene acoplada a la proteína CdgA y extractos proteicos totales de *A. brasilense* Sp7 a partir de cultivos en condiciones planctónicas y en formación de biopelícula.

3. Producir anticuerpos policlonales en contra de la proteína CdgA recombinante para identificar la expresión de la proteína en diferentes estadios del ciclo de vida de *A. brasilense* Sp7.
4. Generar un modelo estructural por homología por medio de programas bioinformáticos del dominio GGDEF de la diguanilato ciclasa A, con la finalidad de identificar los motivos A e I y la probable interface en la que la proteína pueda interactuar con otras DGCs y/o PDEs.
5. Realizar análisis bioinformáticos para predecir las interacciones proteína-proteína utilizando el programa STRING.

## 7.- Esquema de Trabajo



## 8.- Material y Métodos

### 8.1.- Material biológico

<i>Cepas</i>	<b>Fuente</b>
<i>Escherichia coli</i> (pLIL-2) derivado de pQE-31 que contiene un fragmento del alelo <i>cdgA</i> con etiqueta de polihistidinas	Laboratorio Interacción Bacteria-Planta
<i>Azospirillum brasilense</i> Sp7	Obtenida de Tarrand, <i>et al.</i> , 1987

### 8.2.- Inducción de la proteína recombinante CdgA

Se modificó el protocolo de inducción y extracción de la proteína CdgA recombinante presentado por Ramírez-Mata, *et al.*, (2016), para producir y extraer mayor cantidad de proteína CdgA. Se prosiguió a crecer a la cepa *E. coli* pLIL-2 (la cual cuenta con el gen que codifica para la proteína CdgA recombinante) en medio LB a 37°C con agitación de 150 rpm de 12-14 h. Posteriormente, se inoculó medio LB con el inóculo a una dilución de 1:50, se dejó crecer el cultivo a 37°C con agitación de 150 rpm durante 3 h aproximadamente hasta que la densidad óptica (DO) de 600<sub>nm</sub> alcanzara 0.5-0.7. Una vez obtenida la DO deseada, se prosiguió a agregar el inductor Isopropil-β-D-1-tiogalactopiranosido (IPTG) a una concentración final de 0.25 mM y se incubó durante 3 h a 30°C con agitación de 150 rpm. Finalmente se obtuvo la pastilla celular de crecimiento por centrifugación a 8,000 rpm a 4°C, durante 8 min y se guardó el sobrenadante a -80°C hasta su uso. La inducción se comprobó por medio de geles de poliacrilamida SDS-PAGE al 10% (Ramírez-Mata, *et al.*, 2016).

### 8.3.- Lisis celular y purificación de la proteína CdgA por cromatografía de afinidad

La lisis celular y purificación de la proteína CdgA se llevó a cabo siguiendo el protocolo establecido por Ramírez-Mata, *et al.*, (2016) con algunas adaptaciones. Se lisaron las

pastillas celulares utilizando la solución de lisis (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 30 mM Imidazol, 0.1 mM PMSF, 1mg/ml de lisozima y 0.5% de sarcosyl) Se utilizó para resuspender la pastilla celular un volumen del tampón correspondiente al 20% del cultivo. Posteriormente, se sonicó la pastilla así resuspendida hasta obtener una muestra translúcida (los parámetros de sonicación fueron: 1 minuto, 60 amplitud y 10 pulsos/segundo). Se centrifugó a 8,000 rpm a 4°C durante 10 min. Una vez fraccionada la muestra, se retiró el sobrenadante de la pastilla obtenida y se almacenó en alícuotas a -80°C. La lisis celular se comprobó por medio de geles de poliacrilamida SDS-PAGE al 10%.

Se utilizó la cromatografía de afinidad para purificar a la proteína CdgA recombinante, de acuerdo con Ramírez-Mata, *et al.*, (2016). El sobrenadante de la lisis celular y resina de afinidad Ni-NTA (previamente activada con solución de equilibrio basado en 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 30 mM Imidazol) interaccionaron por una h a temperatura ambiente con agitación suave. Se retiró el sobrenadante no absorbido y se guardó a -20°C como muestra control. La resina de interacción se lavó con solución de lavado (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 30 mM Imidazol, 0.1 mM PMSF), hasta no obtener DO a 280<sub>nm</sub>. Se continuaron los lavados con la solución de lavado con alta concentración de NaCl (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 1 M NaCl, 30 mM Imidazol, 0.1 mM PMSF), hasta no obtener DO a 280<sub>nm</sub>. Finalmente, se eluyeron las proteínas con solución de elusión (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 250 mM NaCl, 300 mM Imidazol, 0.1 mM PMSF) y obtener la proteína CdgA.

Como un control se tomaron muestras de resina, antes y después de los lavados y de las eluciones para verificar la interacción entre la resina-proteína y la pureza de la proteína, y posteriormente, comprobar las muestras por medio de geles de poliacrilamida SDS-PAGE al 10%. A partir de estas muestras, se prosiguió a inocular a los modelos animales para la producción de anticuerpos (Ramírez-Mata, *et al.*, 2016).

#### **8.4.- Crecimiento celular de *A. brasilense* Sp7 en estado plántónico y sésil**

Obtención de extractos proteicos celulares de condiciones de crecimiento plántónico y sésil de *A. brasilense* Sp7, para lo cual se crecieron las bacterias en medio NFB\* modificado, con una relación (Carbono 1: Nitrógeno 2), utilizando ácido málico como fuente de carbono

y  $\text{KNO}_3$  como fuente de nitrógeno. Se utilizaron tubos de ensayo en el crecimiento de células planctónicas y placas Petri en el crecimiento de células sésiles.

Se inició a partir de la cepa *A. brasilense* Sp7 crecida en medio Rojo Congo, se tomaron una o dos colonias crecidas en medio LB\* (LB modificado) hasta alcanzar una  $\text{DO}_{600\text{nm}}$  de 1.1-1.4. Los tubos de ensayo con medio NBF\* fueron inoculados con una relación de 1:100 de cultivo fresco de la cepa de *A. brasilense* Sp7 crecida en LB\* y se incubaron de 24 a 96 h a 30°C con agitación de 150 rpm, para los cultivos planctónicos. Las cajas Petri con medio NFB modificado fueron inoculadas con una relación de 1:100 de cultivo fresco de la cepa crecida en LB\* y se incubaron de 24 a 96 h a 30°C sin agitación, para los cultivos en formación de biopelícula (sésiles). Posteriormente, se realizó el mismo proceso para empaquetar las células en ambos casos, centrifugando el cultivo celular a 4°C con 8,000 rpm durante 10 minutos, se guardó el sobrenadante a -80°C para su posterior uso (Arruebarrena, *et al.*, 2012; Ramírez-Mata, *et al.*, 2016).

#### **8.5.- Lisis celular de extractos de *A. brasilense* Sp7**

Partiendo de las pastillas celulares tanto planctónicas como sésiles, se prosiguió a lisar las pastillas por medio de un tampón fosfato salino (PBS) con 0.1 mM de PMSF. Se sonicó el resuspendido celular utilizando: 60-80 de amplitud, 10 pulsos durante 60 seg, dejando reposar por igual de tiempo, hasta observar una solución traslúcida. Una vez finalizado este procedimiento, se cuantificó la cantidad de proteína obtenida por medio del método Bradford

#### **8.6.- Acoplamiento de la proteína CdgA a la resina de afinidad Ni-NTA para ensayos Pull Down**

La purificación de la proteína CdgA recombinante no se llevó a cabo hasta homogeneidad, ya que la finalidad de este protocolo es utilizar a la proteína CdgA recombinante como anzuelo que propicie la interacción con otras proteínas. Se realizaron las siguientes etapas: la interacción resina de afinidad Ni-NTA y sobrenadante celular, lavados con la solución de lavado y solución de lavado con altas concentraciones de cloruro de sodio. Finalmente, muestras de resina fueron tomadas antes y después de los lavados para ser analizadas por

medio de un gel de poliacrilamida SDS-PAGE al 10% para verificar la presencia de la proteína CdgA recombinante. Una vez comprobada la presencia única de la proteína CdgA en la resina de afinidad Ni-NTA, esta última se categorizaba como Ni-NTA-CdgA y sería utilizada posteriormente con los extractos celulares proteicos de las muestras obtenidas de células plantónicas y sésiles de *A. brasilense* Sp7.

### **8.7.- Ensayos Pull Down (interacción proteína-proteína)**

Se prosiguió a interaccionar la resina de afinidad Ni-NTA-CdgA y los extractos proteicos celulares de estados plantónicos y sésiles de *A. brasilense* Sp7 con una relación de 0.1 ml de resina por cada 1 mg de proteína, como control negativo se utilizó a la resina de afinidad Ni-NTA sin interacción con la proteína CdgA recombinante. La interacción se realizó a temperatura ambiente con agitación suave durante una h, se retiró el sobrenadante no absorbido y se guardó a -20°C como muestra control.

Posteriormente, las resinas (Ni-NTA-CdgA y Ni-NTA) fueron lavadas con solución de lavado (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 30 mM Imidazol, 0.1 mM PMSF) hasta no detectar proteínas por espectrofotometría 280<sub>nm</sub>, se tomaron muestras de la resina posterior a los lavados y se almacenó a 4°C. Una vez obteniendo una DO 280<sub>nm</sub>=0, se prosiguió a lavar las muestras con 3 soluciones de lavado con alta concentración de cloruro de sodio (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 0.5/0.75/1 M NaCl, 30 mM Imidazol, 0.1 mM PMSF) y se utilizaron dos volúmenes de resina por cada solución. Las muestras obtenidas de los lavados con la solución de lavado con alta concentración de cloruro de sodio se almacenaron a -20°C para su posterior análisis.

Una vez finalizados los lavados de las resinas, se analizaron las muestras obtenidas de los ensayos realizados por medio de geles de poliacrilamida SDS-PAGE al 10%.

### **8.8.- Producción de anticuerpos policlonales en contra de la proteína CdgA recombinante**

Siguiendo los protocolos de expresión y purificación de la proteína CdgA recombinante previamente establecidos en este trabajo, se obtuvo la proteína con el objetivo de emplearla como antígeno y obtener anticuerpos policlonales anti-CdgA.

Se prepararon emulsiones a base de adyuvante de Freud (completo e incompleto) más el antígeno (proteína CdgA-recombinante). El antígeno más el adyuvante se administraron de manera subcutánea al conejo (modelo animal otorgado y cuidado por parte del bioterio Claude Bernard perteneciente a la Benemérita Universidad Autónoma de Puebla) cada 3 a 4 semanas con un promedio de entre 10-30µg de proteína por reto y un total de 5 retos realizados.

Se tomaron muestras sanguíneas antes de realizar cada reto de inoculación como controles y al finalizar los retos de administración del antígeno, se realizó un Western Blot para comprobar la presencia de los anticuerpos anti-CdgA. Se utilizó el suero pre-inmune versus el suero obtenido del cuarto reto para comprobar la producción de anticuerpos. Una vez comprobado la presencia de los anticuerpos anti-CdgA, se prosiguió a sangrar al modelo animal y a preparar el suero final que contuviera los anticuerpos deseados.

### 8.9.- Geles de poliacrilamida desnaturalizantes o SDS-PAGE al 10%

Para preparar los geles de poliacrilamida SDS-PAGE al 10% se utilizaron las siguientes soluciones.

Material (gel de carga 4%)	3 mL	Material (gel de corrida 10%)	10 mL
30 % Acrilamida	399 µl	30 % Acrilamida	3.31 ml
1% Bis- Acrilamida	330 µl	1% Bis- Acrilamida	2.71 ml
2 M Tris pH 6.8	187.5 µl	2 M Tris pH 6.8	1.875 ml
10% SDS	30 µl	10% SDS	100 µl
Agua	2.020 ml	Agua	1.88 µl
10% APS	30 µl	10% APS	100 µl
Temed	3 µl	Temed	8 µl

Posteriormente, se prepararon las muestras con buffer de carga 2x el cual contuvo las siguientes características, 2M Tris-HCl pH 6.8, 10% SDS, 10% Mercaptoetanol, 20% Glicerol, 0.1% Azul Bromofenol y agua cbp. Se tomó la siguiente relación de buffer de carga con respecto a muestra es la siguiente 1:2, se mezcló la muestra con el buffer y finalmente se hirvió durante 5 min. Finalmente se cargó el gel y se corrió a 80 V y al finalizar, se tiñó con solución de Coomassie o por tinción de plata.

### **8.10.- Western Blot: Transferencia a membrana**

Primero se preparó el buffer de transferencia (25mM Tris-base, 192mM glicina y 20% metanol) y después se equilibraron los siguientes materiales con este buffer: dos esponjas Scotch brite, papel Wathman 3MM y el gel de poliacrilamida a transferir por 7 minutos aproximadamente. Al mismo tiempo, se activó la membrana de transferencia Immobilon-P PVDF (por sus siglas en inglés polyvinylidene difluoride) con metanol puro antes de usarlo. Se montó el experimento usando guantes, de la siguiente manera: en un vidrio se colocó el papel Wathman tratado y sin arrugas y burbujas, se colocó el gel a transferir a la membrana sobre el papel Wathman. Se colocó la membrana de transferencia Immobilon-P PVDF encima del gel con presión suave, se cubrió con papel filtro tratado, siempre cuidadosamente de no dejar burbujas y se transfirió al soporte de la cámara de transferencia. Se colocó previamente una esponja Scotch brite tratada sobre el soporte de la cámara y el conjunto del papel Wathman, gel de poliacrilamida, y membrana de transferencia Immobilon-P PVDF se colocaron encima de otra esponja Scotch brite (el gel deberá estar más próximo al electrodo negativo, mientras que la membrana deberá estar más próxima al electrodo positivo). Finalmente, se colocó el soporte dentro de la cámara de transferencia. De tal forma quedó en el soporte: soporte parte negativa (color negro, esponja Scotch brite, papel filtro, gel a transferir, membrana de transferencia Immobilon-P PVDF, papel filtro, esponja Scotch brite, soporte parte positiva (color transparente), el equipo utilizado para realizar la transferencia proviene del distribuidor BioRad. Para finalizar, se cubrió totalmente la cámara de transferencia con buffer de transferencia y se corrió a 200mA durante 1h, 20 min.

Una vez finalizada la transferencia, con la finalidad de verificar una transferencia exitosa se tiñó la membrana de transferencia con azul de Coomassie y al detectar la presencia de bandas proteicas, se destiñó totalmente la membrana con solución para desteñir o metanol puro.

Posteriormente al comprobar la transferencia, se prosiguió a bloquear la membrana para la inmunodetección (Buffer de bloqueo: 5% leche en polvo baja en grasa, 0.1% de Tween-20, y TBS 1x) durante toda una noche a 4°C.

### **8.11.- Western Blot: Inmunodetección anti-CdgA (anticuerpos policlonales de conejo)**

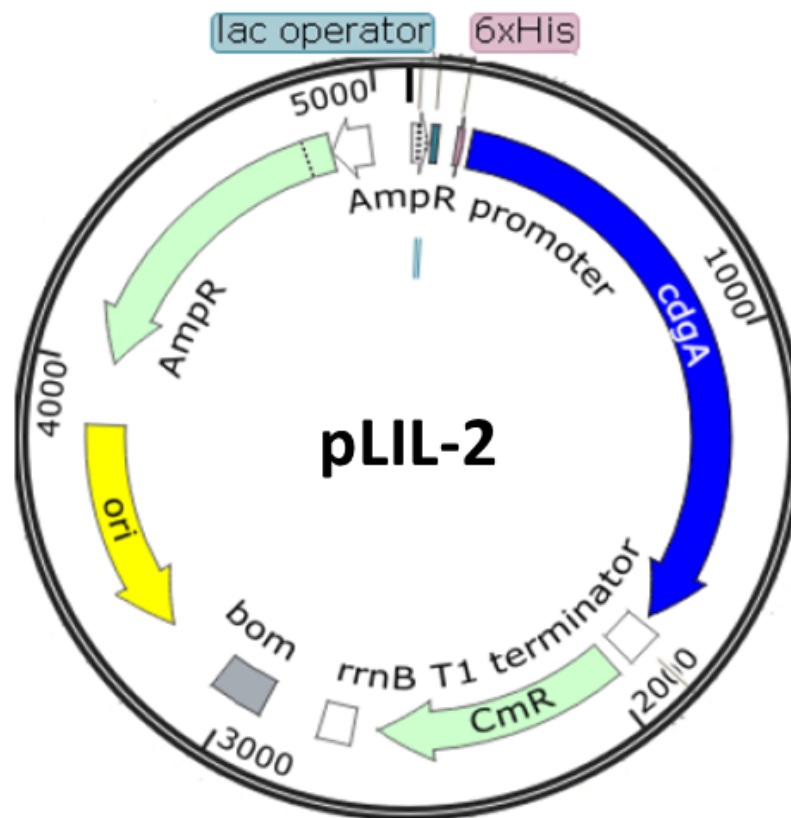
Posterior al bloqueo de la membrana de transferencia Immobilon-P PVDF previamente establecido, se prosiguió a realizar la inmunodetección utilizando el anticuerpo anti-CdgA como anticuerpo primario en el buffer de bloqueo. Previamente se estimó la dilución óptima para su uso (relación anticuerpo: buffer de bloqueo), se desechó el buffer de bloqueo, se agregó nuevo buffer de bloqueo más el anticuerpo anti-CdgA y se dejó interactuar durante 1 h a temperatura ambiente. Posteriormente, se retiró el buffer con el anticuerpo primario y se realizaron 3 o 4 lavados rápidos (de 30 seg utilizando agitación manual) y 3 o 4 lavados lentos (de 5 min utilizando agitador a 30 rpm) utilizando (1x TBS y 0.1% Tween-20). El TBS es un buffer salino a base de Tris y consiste en: 20 mM Tris-base y 150 mM NaCl.

Al finalizar los lavados, se preparó nuevo buffer de bloqueo con el anticuerpo secundario a utilizar. Se determinó la dilución a utilizar el anticuerpo secundario y éste se agregó a la membrana en el buffer de bloqueo. El anticuerpo secundario (antirabbit-IgG-HRP conjugado proporcionado por Promega o anti-6xHis proporcionado por Clontech) se dejó en agitación durante una h a temperatura ambiente. Se retiró el buffer de bloqueo con el anticuerpo secundario y se realizaron de 3 a 4 lavados rápidos y 3 a 4 lavados lentos (5 mi) con TBS y 0.1% Tween-20. Finalmente, se reveló la membrana con una solución a base de 100 µl de 3%  $\alpha$ -cloronaftol y 100 µl de 3% peróxido de hidrógeno por cada 10 ml de TBS 1x.

## 9.- Resultados

### 9.1.- Modificación de los protocolos de inducción, lisis y purificación de la proteína CdgA recombinante

Con el objetivo de utilizar la proteína CdgA recombinante (Figura 9) como anzuelo en ensayos Pull Down, los cuales consisten en presentar interacciones proteína-proteína, se estableció desarrollar una columna de afinidad acoplado la proteína CdgA recombinante de la cepa *E. coli* pLIL-2 a la resina de afinidad Ni-NTA (Figura 11) (Ramírez-Mata, *et al.*, 2016).



**Figura 9.- Plásmido pLIL-2 de la cepa *E. coli* pLIL-2.** Plásmido presente en *E. coli* que contiene el gen que codifica para la proteína CdgA y aporta una etiqueta de Histidinas al residuo N-terminal de la proteína, cuenta con un operador *lac* que permite la inducción con IPTG para obtener alta expresión del gen, y también cuenta con un casete de resistencia a Ampicilina que permite la selección de la cepa *E. coli* pLIL-2 (Ramírez-Mata, *et al.*, 2016).

Se realizaron modificaciones a los protocolos de inducción, lisis y purificación con el objetivo de mejorar el rendimiento y pureza de la proteína. En la Tabla 1 se muestran las modificaciones realizadas a los protocolos previamente mencionados para este trabajo.

**Tabla 1.- Modificaciones realizadas al protocolo de inducción, lisis y purificación de la proteína CdgA recombinante establecido por López, 2014.**

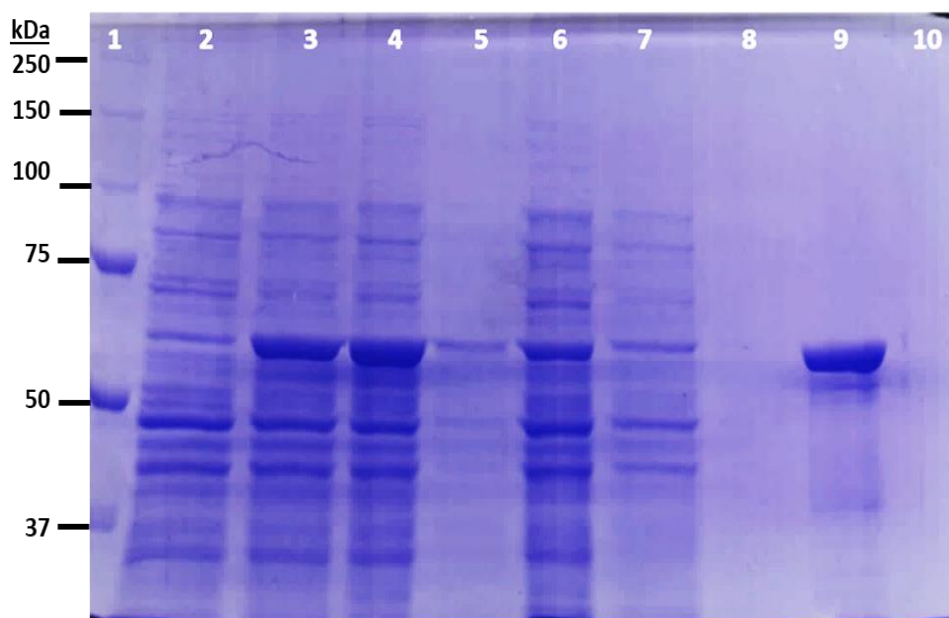
<b>Protocolo de Inducción</b>		
	Protocolo de inducción y purificación establecido por López, 2014.	Protocolo de inducción y purificación establecido en este trabajo
<i>Concentración de Inductor IPTG</i>	[1 mM] concentración final	[0.25 mM] concentración final
<i>Temperatura de incubación con inductor</i>	30°C	30°C
<i>Tiempo de incubación con el inductor</i>	2 horas	3 horas
<i>Agitación de incubación con el inductor</i>	150 rpm	150 rpm
<b>Protocolo de lisis</b>		
<i>Volumen de buffer de lisis</i>	10% del volumen del cultivo inducido utilizado	20% del volumen del cultivo inducido utilizado
<i>Concentración de Sarcocyl en buffer de lisis</i>	1 % concentración final	0.5% concentración final
<i>Concentración de Imidazol en buffer de lisis</i>	10 mM concentración final	30 mM concentración final
<i>Concentración de Lisozima</i>	0.5 mg/ml	0.5 mg/ml
<i>Tiempo de incubación de Lisozima</i>	10 min	30 min
<i>Temperatura de incubación de Lisozima</i>	37°C	Temperatura ambiente
<i>Sonicación</i>	Tiempo: 3 min	Tiempo: 1 min
	Amplitud: 50	Amplitud: 60
	Pulsos/seg: 6	Pulsos/seg: 10
	Repeticiones: 2 o hasta	Repeticiones: 2 o hasta

	observar fase transparente	observar fase transparente
<i>Centrifugación de muestra posterior a lisis</i>	8,000 rpm durante 20 min a 4°C	10,000 rpm durante 10 min a 4°C
<i>Almacenamiento de muestras</i>	-20°C	-20°C
<b>Protocolo de purificación</b>		
<i>Uso de cromatografía por exclusión de peso molecular</i>	Sí	No
<i>Concentración de Imidazol presente en solución de equilibrio de resina y solución de lavado para cromatografía de afinidad</i>	[20 mM] concentración final	[30 mM] concentración final
<i>Tiempo, temperatura y agitación de interacción resina-muestra</i>	2 horas a 4°C con agitación suave	1 hora a temperatura ambiente con agitación suave
<i>Volumen utilizado de soluciones de lavado</i>	En base al volumen de columna que se utilice	Lavar hasta obtener una DO= 0
<i>Uso de solución de lavado libre de NaCl</i>	No	Sí

Los nuevos parámetros establecidos y los resultados para incrementar la concentración de la proteína CdG A pueden ser observados en la Figura 10 donde se muestran los resultados obtenidos en las diferentes etapas analizadas mediante el empleo de electroforesis en geles de poliacrilamida en presencia del detergente dodecil sulfato de sodio (PAGE-SDS) condiciones desnaturizantes. La proteína CdG A recombinante tiene una masa molecular relativa (MMr) de 60 kDa. Se cuantificó la concentración de proteína por método Bradford para cargar 20 µg por carril, lo cual permitió observar diferencias entre las muestras y los ajustes realizados a los protocolos (Figura 10 y 11).

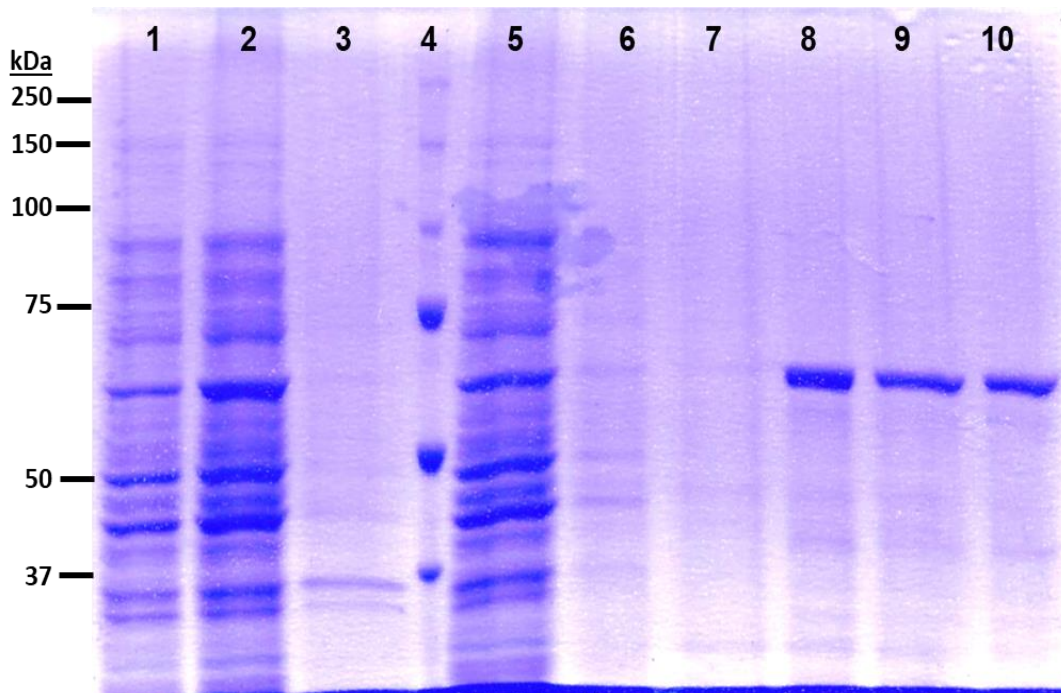
Tras haber estandarizado los cambios a los protocolos de inducción y lisis celular para obtener un mejor rendimiento de la proteína CdG A recombinante (CdG Ar), se verificó que

la cantidad de detergente no afectara el proceso de acoplamiento de la proteína a la resina de afinidad Ni-NTA con la finalidad de eliminar el excedente de detergente utilizado para la lisis celular (Ramírez-Mata, *et al.*, 2016). Como se observa en el electroforetograma se mejoró notablemente la obtención de la proteína CdgAr (Figura 10).



**Figura 10.- Modificaciones realizadas a los protocolos de inducción, lisis y purificación de la proteína CdgA recombinante (Tinción por Coomassie).** Carril 1) Marcador de peso molecular Carril 2) Muestra de extracto proteico total tiempo cero de cultivo celular de *E. coli* pLIL-2 previo a la inducción de [0.25 mM] de IPTG; Carril 3) Muestra de extracto proteico total final de *E. coli* pLIL-2 posterior a inducción con [0.25 mM] de IPTG y lisis celular; Carril 4) Muestra de sobrenadante del extracto proteico total final de *E. coli* pLIL-2 posterior a inducción [0.25mM] de IPTG y lisis celular; Carril 5) Muestra de pastilla del extracto proteico total final de *E. coli* pLIL-2 posterior a inducción [0.25mM] de IPTG y lisis celular; Carril 6) Muestra de sobrenadante no absorbido de *E. coli* pLIL-2 posterior a la interacción con resina de afinidad NTA; Carril 7) Primer lavado de resina con solución de lavado; Carril 8) Primer lavado de resina con solución de lavado con [↑NaCl]; Carril 9) Muestra de resina después del lavado con [↑NaCl] (10  $\mu$ l de resina); Carril 10) Vacío.

Sin embargo, en el carril 9 no sólo se muestra la presencia de la proteína CdgAr sino también un pequeño porcentaje de contaminantes. Por lo cual, se decide realizar modificaciones al protocolo de purificación que permitan la eliminación de estas impurezas retenidas. Aumentando la cantidad de imidazol presente en las soluciones de lavado y la cantidad de lavados realizados a las muestras, se pudo obtener la eliminación de impurezas retenidas en el acoplamiento de la resina de afinidad Ni-NTA y la proteína CdgAr. Esto se comprueba en la Figura 11.



**Figura 11.- Acoplamiento de la proteína a resina de afinidad a níquel mostrado en gel SDS-PAGE al 10% (Tinción por Coomassie). Carril 1)** Muestra de extracto proteico total tiempo cero de cultivo celular de *E. coli* pLIL-2previo a la inducción de [0.25mM] IPTG; **Carril 2)** Muestra de sobrenadante del extracto proteico total final de *E. coli* pLIL-2 posterior a inducción [0.25mM] IPTG y lisis celular; **Carril 5)** Muestra de pastilla del extracto proteico total final de *E. coli* pLIL-2 posterior a inducción [0.25mM] IPTG y lisis celular; **Carril 6)** Marcador de peso molecular; **Carril 7)** Muestra de sobrenadante no absorbido de *E. coli* pLIL-2 posterior a la interacción con resina de afinidad NTA; **Carril 8)** Muestra de resina después del último lavado con la solución de lavado; **Carril 9)** Muestra de resina después del último lavado con solución de lavado con [↑NaCl]; **Carril 10)** Muestra de resina después de lavado con solución de lavado sin NaCl.

Después de haber realizado cambios de concentración de imidazol en las soluciones de lavado y volumen utilizado para la purificación de la proteína recombinante como se muestra en la Tabla 1, se observa en la Figura 11 que la proteína CdGAr es mantenida posterior a los lavados (carriles 8, 9 y 10) con menor cantidad de impurezas. Las modificaciones realizadas al protocolo de purificación parecen no afectar la interacción entre la resina de afinidad NTA con la proteína CdGAr.

Después de realizar las modificaciones y mejorar la producción de la proteína CdGAr (Tabla 1), se realizaron dos tipos de experimentos con la resina de afinidad Ni-NTA acoplada a la proteína CdGAr. El primero fue utilizar la resina acoplada a la proteína CdGAr para poder realizar ensayos Pull Down con extractos proteicos totales de *A. brasilense* Sp7 en estados platónicos y sésiles, mientras que el segundo consistió en la elución de la

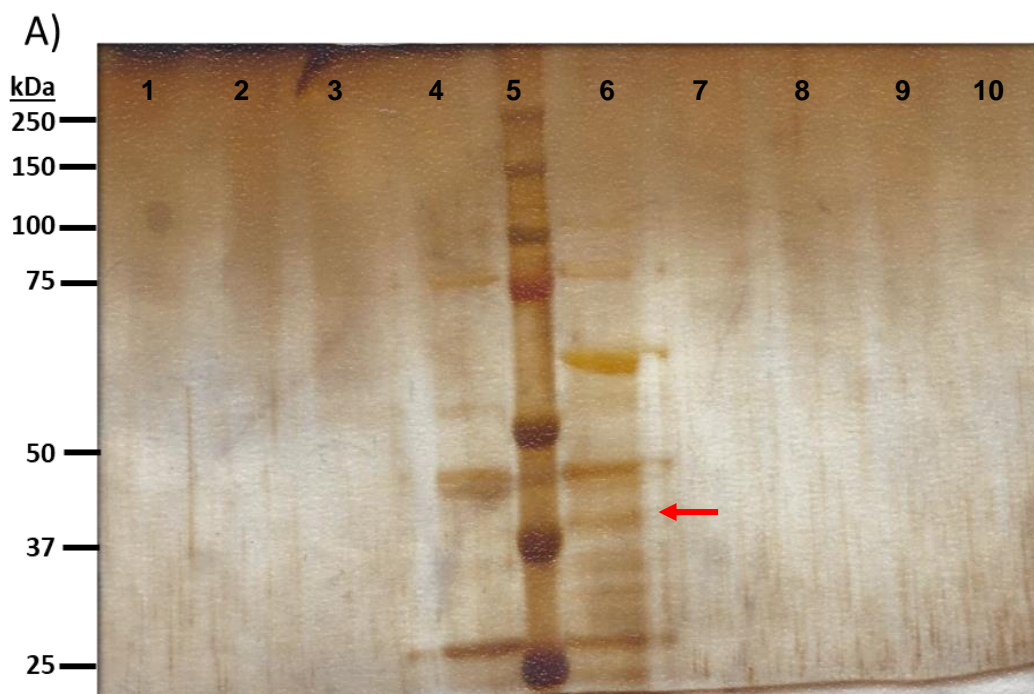
proteína CdgAr para utilizarla como antígeno y generar anticuerpos policlonales en contra de la proteína.

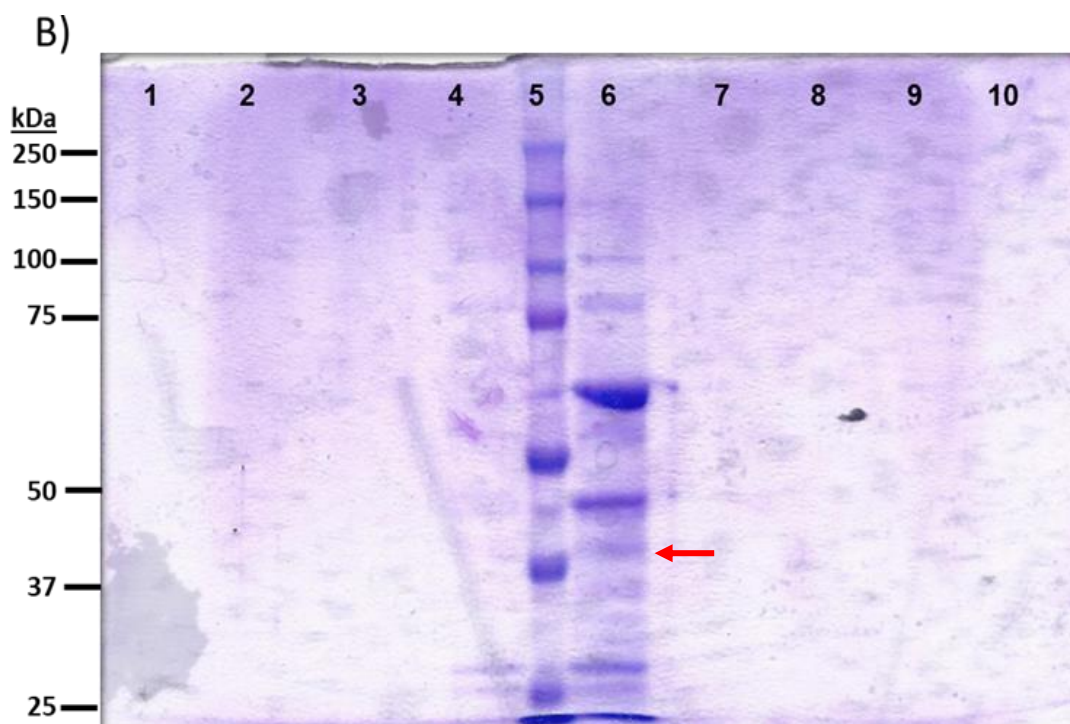
## 9.2.- Ensayos Pull Down

Una vez obtenido el acoplamiento de la proteína CdgAr a la resina de afinidad Ni-NTA (Ni-NTA-CdgAr), se prosiguió a utilizar a esta resina en ensayos Pull Down (interacción proteína-proteína). Para realizar estos ensayos, se utilizó extractos proteicos totales de células planctónicas o células sésiles de un día de crecimiento de *A. brasilense* Sp7.

La interacción proteína-proteína se llevó a cabo utilizando los extractos libres de células (1 mg de proteína) de *A. brasilense* Sp7, ya sea provenientes de células planctónicas o células sésiles de un día de crecimiento, por cada 100 µl de resina de afinidad Ni-NTA-CdgAr y como control negativo, se empleó la resina de afinidad Ni-NTA.

Se realizaron 3 ensayos independientes de interacción proteína-proteína por cada estilo de crecimiento de la bacteria *A. brasilense* Sp7. En las Figuras 12 y 13 se presentan los resultados obtenidos de los ensayos Pull Down utilizando extractos proteicos de células planctónicas y sésiles (respectivamente) de 1 día de crecimiento de *A. brasilense* Sp7.





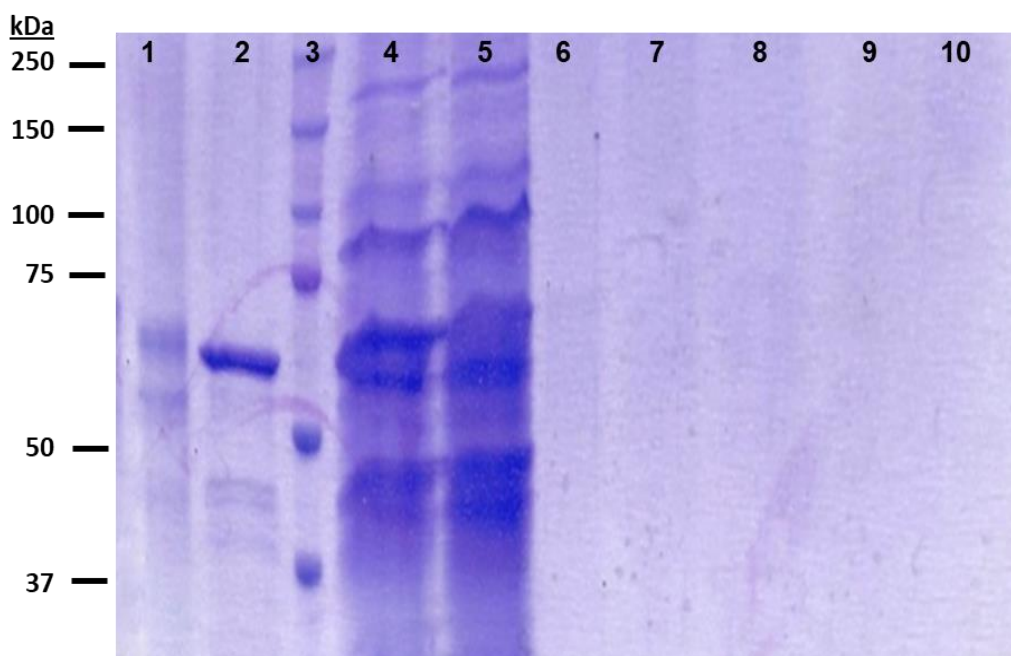
**Figura 12.- Geles de poliacrilamida SDS-PAGE 10% con muestras obtenidas posterior a la interacción de células plantónicas Día 1 de *A. brasilense* Sp7 y la resina acoplada a la proteína CdgA. A) Gel de poliacrilamida SDS-PAGE revelado con tinción de plata. B) Gel de poliacrilamida SDS-PAGE revelado con tinción de Coomassie. Ambos geles cuentan con el siguiente orden de carriles: **Carril 1)** Muestra de lavado posterior al uso de solución de lavado con 1M [NaCl] de la resina control; **Carril 2)** Muestra de lavado posterior al uso de solución de lavado con 750mM [NaCl] de la resina control; **Carril 3)** Muestra de lavado posterior al uso de solución de lavado con 500mM [NaCl] de la resina control; **Carril 4)** Muestra de la resina control final posterior a los lavados; **Carril 5)** Marcador de peso molecular; **Carril 6)** Muestra de resina de la afinidad Ni-NTA-CdgAr después de los lavados; **Carril 7)** Muestra de lavado posterior al uso de solución de lavado con 500 mM [NaCl] de la resina de afinidad Ni-NTA-CdgAr; **Carril 8)** Muestra de Lavado posterior al uso de solución de lavado con 750mM [NaCl] de la resina de afinidad Ni-NTA-CdgAr; **Carril 9)** Muestra de Lavado posterior al uso de solución de lavado con 1M [NaCl] de la resina de afinidad Ni-NTA-CdgAr; **Carril 10)** Vacío.**

En la Figura 12-A se presenta el gel SDS-PAGE revelado por tinción de plata (tinción que permite observar nanogramos de proteínas). En el gel se puede observar la presencia de diferentes bandas adheridas a la resina control (carril 4) y a la resina de afinidad Ni-NTA-CdgAr (Carril 6), por lo cual podemos determinar que la resina de afinidad Ni-NTA cuenta con la capacidad de tener proteínas de los extractos proteicos totales de células plantónicas de manera inespecífica, lo cual no se logra observar con la tinción de Coomassie (Figura 12-B carril 4). De este modo, aquellas bandas presentes en la resina control (carril 4) son descartadas como posibles proteínas que interactúan con CdgAr, y por ello, se logra detectar la presencia de una banda con un peso aproximado a  $>37$  kDa (de acuerdo con marcador de peso molecular) que se encuentra señalado con una flecha roja en la imagen.

La resina Ni-NTA-CdgAr retiene varias proteínas con tamaños alrededor de 25-50 kDa y aun de mayor MMr (carril 6); mientras que en el control (carril 4) solo se presentan dos proteínas en el mismo rango destacadas sólo en el gel teñido con AgNO<sub>3</sub> (Figura 12-A). Por otro lado, no se logra observar el desprendimiento de proteínas adheridas inespecíficamente a las resinas utilizando altas concentraciones de NaCl (carriles 1, 2 y 3 para muestras del control y carriles 7, 8 y 9 para muestras de resina de afinidad Ni-NTA-CdgAr).

Se realizó el mismo proceso de interacción entre la resina de afinidad Ni-NTA acoplada a la proteína CdgA recombinante y la resina control, con los extractos proteicos totales de células sésiles de 1 día de crecimiento de *A. brasilense* Sp7.

En la Figura 13 se muestran los datos obtenidos de SDS-PAGE al 10% teñido por Coomassie, de las muestras de extractos libres de células crecidas en cultivo estático (sésil), muestra de la resina de afinidad Ni-NTA-CdgAr final de los ensayos Pull Down con extractos proteicos de células planctónicas de 1 día de crecimiento de *A. brasilense* Sp7 con el objetivo de comparar los resultados obtenidos.



**Figura 13.- Gel de poliacrilamida SDS-PAGE 10% con muestras obtenidas posterior a la interacción de células de biopelícula Día 1 de *A. brasilense* Sp7 y la CdgA acoplada a la resina (tinción de Coomassie). Carril 1) Muestra de resina acoplada a CdgAr después de su interacción con extracto proteico planctónico de Día 1 y lavados; Carril 2) Muestra del sobrenadante de resina acoplada a CdgAr después de su interacción con extracto proteico planctónico de Día 1 y lavados; Carril 3) Marcador de peso molecular; Carril 4)**

Muestra de resina acoplada a CdGAr después de su interacción con extracto proteico biopelícula de Día 1 y lavados; **Carril 5**) Muestra del sobrenadante de resina acoplada a CdGAr después de su interacción con extracto proteico biopelícula de Día 1 y lavados; **Carril 6**) Muestra de resina control después de su interacción con extracto proteico de biopelícula de Día 1 y lavados; **Carril 7**) Muestra de sobrenadante de resina control después de su interacción con extracto proteico de biopelícula de Día 1 y lavados); **Carril 8, 9 y 10**) Vacío.

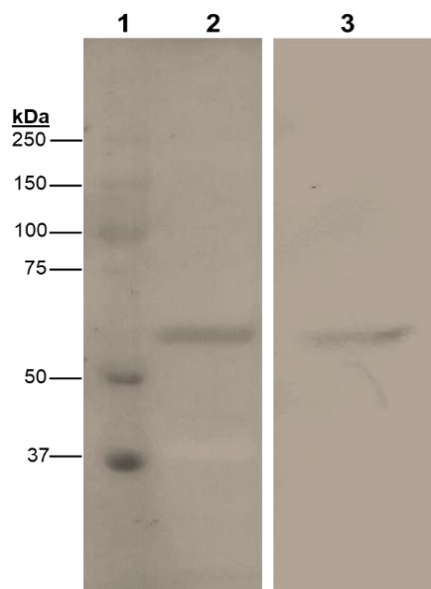
Se observa en la Figura 13 los resultados obtenidos de la interacción realizada con extractos proteicos totales de células sésiles de *A. brasilense* Sp7, donde la resina control no unió proteínas de manera inespecífica (carriles 6 y 7) como se observó en la Figura 12. Sin embargo, la resina Ni-NTA-CdGAr mostró la unión de proteínas en mayor cantidad como se indica en los carriles 4 y 5, las cuales deberán ser identificadas posteriormente por espectrometría de masas. Los resultados de las eluciones con altas concentraciones de NaCl de las muestras no se presentan ya que no mostraron nada al igual que en los geles presentados en la Figura 12).

Los resultados de estos ensayos fueron confirmados a partir de 3 ensayos independientes para cada interacción entre resinas y extractos proteicos (de células plantónicas y células sésiles) de *A. brasilense* Sp7.

### **9.3.- Anticuerpos policlonales Anti-CdGAr de *A. brasilense* Sp7**

Utilizando sueros obtenidos del primer reto previo a la inmunización y el último suero obtenido en la cuarta inmunización del animal, se realizó un Western Blot comparativo que permitiera determinar la producción o no de anticuerpos en el modelo animal (datos no mostrados), y la detección de las proteínas CdGAr de *E. coli* pLIL-2 y CdGAr en *A. brasilense* Sp7. Una vez verificado la producción de anticuerpos en contra de la proteína CdGAr, se sangró al animal y se determinó la cantidad mínima de suero inmune a utilizar para revelar la CdGAr por el ensayo Western Blot.

Como la proteína CdGAr contiene la etiqueta de histidinas el ensayo se realizó empleando ambos anticuerpos Anti-CdGAr y Anti-His respectivamente, dato que se muestra en Figura 14.

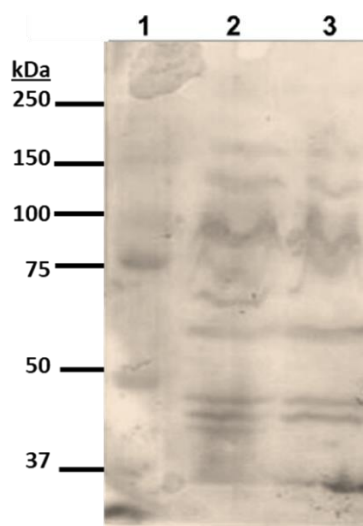


**Figura 14.- Western Blot utilizando anticuerpo comercial Anti-Histidinas y anticuerpo generado en el laboratorio Anti-CdgA de *A. brasilense* Sp7.** La membrana de transferencia Immobilon-P PVDF presenta el siguiente orden de carriles: **Carril 1)** Marcador de Peso molecular; **Carril 2)** Extracto proteico celular total de inducción a [0.25mM] de IPTG en *E. coli* pLIL-2; **Carril 3)** Extracto proteico celular total de inducción a [0.25mM] de IPTG en *E. coli* pLIL-2. Buffer de bloqueo a base de 5% de leche baja en grasa, 0.1% Tween-20 y TBS 1x. Anticuerpo primario Anti-CdgA dilución 1:1,500 utilizado en carril 1 y 2 y anticuerpo primario Anti-Histidinas dilución 1:10,000 utilizado en el carril 3 y ambas membranas de Western Blot ocuparon el anticuerpo secundario antirabbit-IgG-HRP conjugado a dilución 1:5,000.

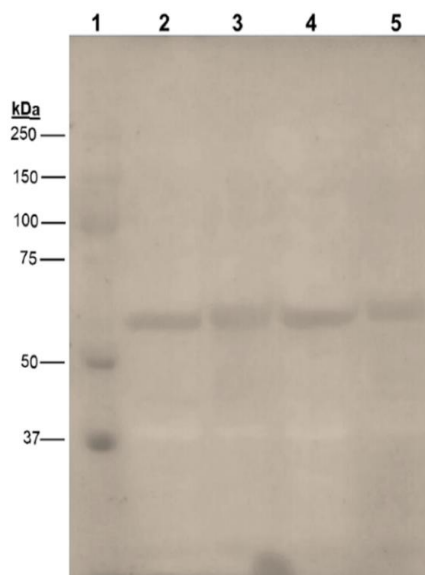
Se observa en la Figura 14, el resultado de la reacción antígeno-anticuerpo que reconoce a la proteína CdgAr en un extracto celular total de *E. coli* pLIL-2 (carril 2) de 20  $\mu$ g a la par del anticuerpo Anti-Histidinas (carril 3). Con estos datos, determinamos que el anticuerpo policlonal producido para este trabajo es funcional y permite el reconocimiento de la proteína CdgA recombinante. Se prosiguió a determinar la cantidad de Anti-CdgA necesaria para detectar a la proteína en *A. brasilense* Sp7.

En los ensayos para determinar la cantidad de anticuerpo Anti-CdgA a utilizar en los extractos proteicos de *A. brasilense* Sp7, se detectó la proteína CdgA (Figura 16). Sin embargo, varias bandas se revelaron en ambos extractos (Figura 15). Se sugiere la detección de otras proteínas diguanilato ciclasas, proteínas con dominios PAS o proteínas que reconozcan a la proteína CdgA. Cabe recordar que la proteína CdgA contiene 3 dominios PAS un dominio GGDEF altamente conservado, por lo que reacciones cruzadas estarían siendo detectadas por el anticuerpo policlonal obtenido. Por ello, fue necesario

determinar la cantidad exacta de anticuerpo primario que sólo permita observar exclusivamente la detección de la proteína CdgAr.



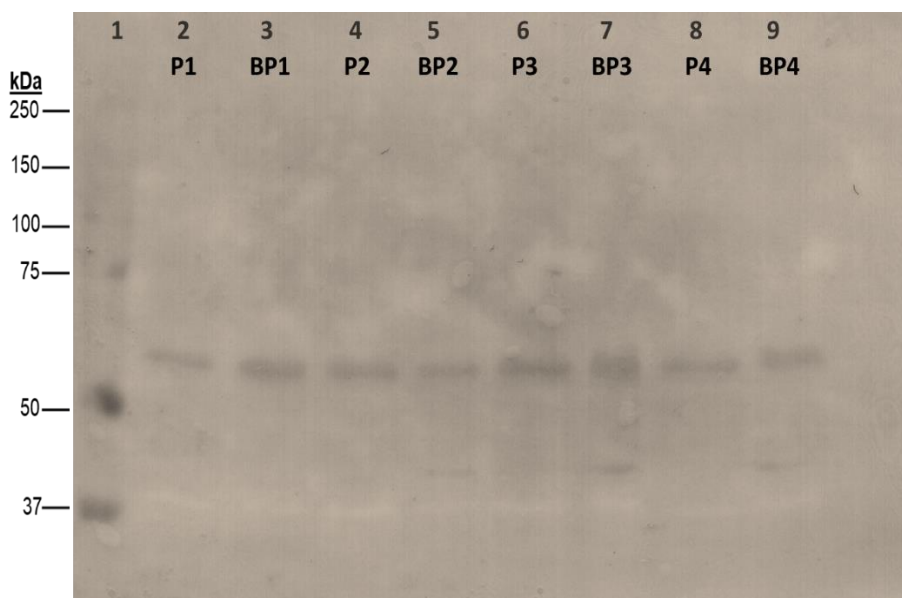
**Figura 15.- Detección de la proteína Cdg-A y otras proteínas en diferentes etapas de crecimiento de *A. brasilense* Sp7.** La membrana de transferencia Immobilon-P PVDF presenta el siguiente orden de carriles: **Carril 1)** Marcador; **Carril 2)** Extracto proteico total de células platónicas Día 3; **Carril 3)** Extracto proteico total de células de biopelícula Día 3. Anticuerpo primario Anti-CdgA dilución 1:1,000 y anticuerpo secundario antirabbit-IgG-HRP conjugado dilución 1:5,000.



**Figura 16.- Detección de la proteína CdgA y en diferentes etapas de crecimiento de *A. brasilense* Sp7.** La membrana de transferencia Immobilon-P PVDF presenta el siguiente orden de carriles: **Carril 1)** Marcador; **Carril 2)** *E. coli* pLIL-2 (Inducción utilizando 0.25 mM de IPTG por 3 h). **Carril 3)** Extracto proteico de células platónicas Día 1; **Carril 4)** Extracto proteico de células platónicas Día 3; **Carril 5)** Extracto proteico de células de biopelícula Día 3. Anticuerpo primario Anti-CdgA dilución 1:1,000 y anticuerpo secundario antirabbit-IgG-HRP dilución 1:5,000.

Ahora bien, debido a la detección de la proteína CdgA en 2 días diferentes con dos estilos de vida diferentes (células planctónicas y células sésiles) se prosiguió a realizar el ensayo que permitiera analizar la detección de la proteína CdgA en ambos estilos de vida, pero a 4 días de crecimiento.

En la Figura 17 se muestra la presencia de la proteína CdgA en los cuatro días analizados tanto de células planctónicas como de células sésiles. En el primer carril se presenta el marcador de peso molecular, en los carriles 2 y 3 se presentan los extractos proteicos de células planctónicas y sésiles (respectivamente) de 1 día de crecimiento, en los carriles 4 y 5 se presentan los extractos proteicos de células planctónicas y sésiles (respectivamente) de 2 días de crecimiento, en los carriles 6 y 7 se presentan los extractos proteicos de células planctónicas y sésiles (respectivamente) de 3 días de crecimiento, y finalmente, en los carriles 8 y 9 se presentan los extractos proteicos de células planctónicas y sésiles (respectivamente) de 4 días de crecimiento.



**Figura 17.- Detección de la proteína Cdg-A en dos estilos de vida diferente de *A. brasilense* Sp7.** La membrana de transferencia Immobilon-P PVDF presenta el siguiente orden de carriles: **Carril 1)** Marcador de Peso Molecular; **Carril 2)** Extracto proteico total de células planctónicas Día 1; **Carril 3)** Extracto proteico de células de biopelícula Día 1; **Carril 4)** Extracto proteico de células planctónicas Día 2; **Carril 5)** Extracto proteico de células de biopelícula Día 2; **Carril 6)** Extracto proteico de células planctónicas Día 3; **Carril 7)** Extracto proteico de células de biopelícula Día 3; **Carril 8)** Extracto proteico de células planctónicas Día 4; **Carril 9)** Extracto proteico de células de biopelícula Día 4. Anticuerpo primario Anti-CdgA dilución 1:1,500 y anticuerpo secundario antirabbit-IgG-HRP homogeniza el nombre y define que tiene acoplado HRP dilución 1:5,000.

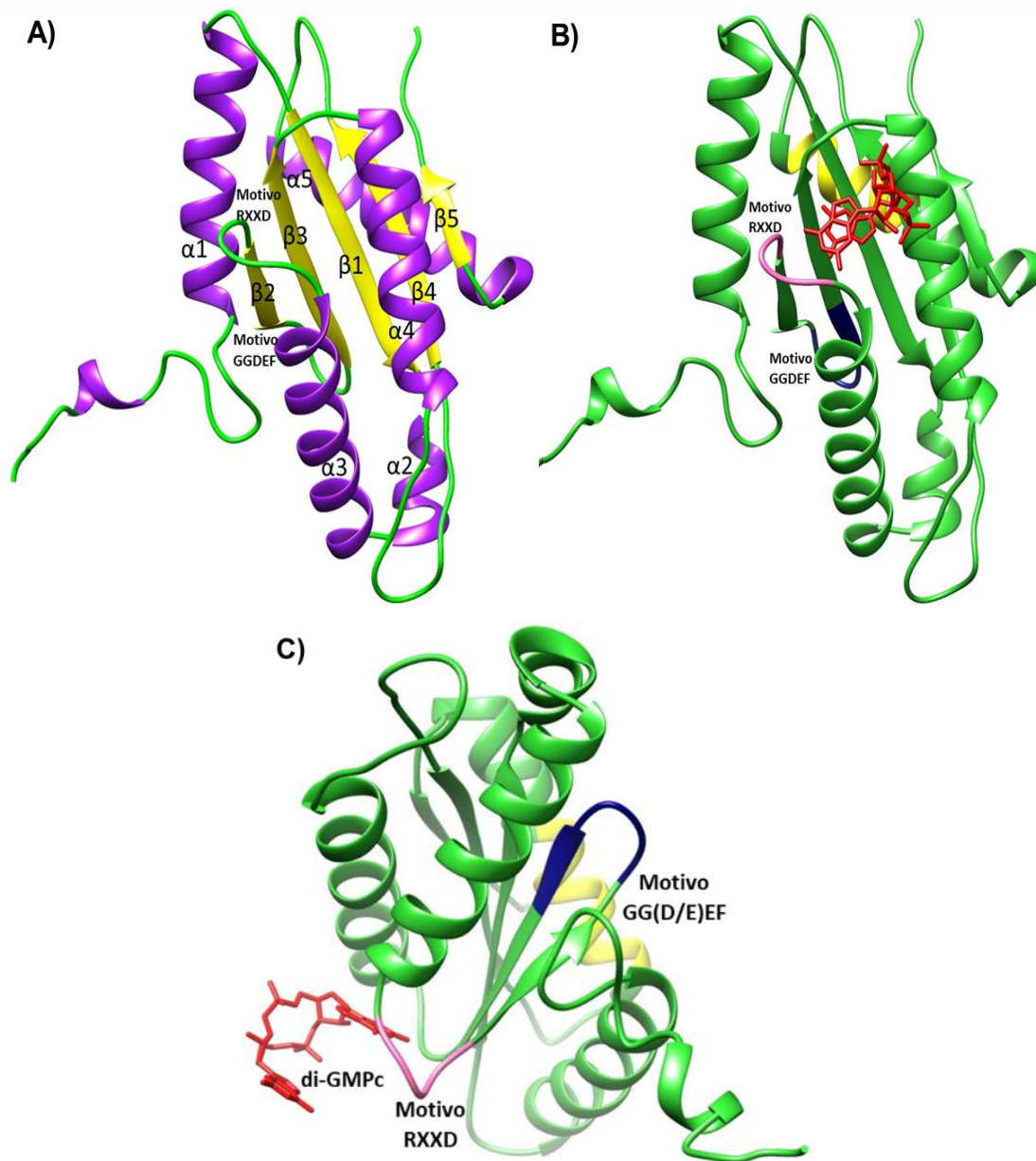
En base a lo obtenido en la Figura 17, la proteína CdgA se presenta en los cuatro días diferentes de crecimiento de los estados planctónico y sésil respectivamente. Este resultado confirma los estudios realizados por Ramírez-Mata, *et al.*, (2016) que muestran la presencia de la proteína tanto para movilidad, como para la formación de biopelícula de la bacteria.

#### **9.4.- Análisis bioinformático estructural de la proteína CdgA de *A. brasilense* Sp7**

##### **9.4.1- Modelo estructural del dominio GGDEF de la proteína CdgA**

Se analizó la estructura terciaria del dominio GGDEF de la proteína CdgA de *A. brasilense* Sp7, con un análisis *in silico* empleando el servidor bioinformático I-TASSER el cual consiste en generar un modelo estructural por homología a partir de secuencias aminoacídicas y estructuras cristalográficas de proteínas ya registradas en el banco de datos de proteínas o PDB (por sus siglas en inglés de Protein Data Bank). La exactitud del modelo otorgado por el programa (C-score) es dado por un rango de -5 a 2. Entre más próximo a 2, el modelo comparte mayor similitud a la base estructural del cual es generado y, por lo tanto, la similitud entre ambas estructuras (estructura registrada versus estructura generada) es alto.

Al ingresar la secuencia de aminoácidos del dominio GGDEF de la proteína CdgA de *A. brasilense* Sp7 al programa, nos proporcionó varios modelos estructurales que tomamos como base para el modelado por homología. Los tres primeros modelos estructurales obtenidos fueron: 2V0N de PleD de *Caulobacter vibrioides* con una resolución de 2.71 Å, 5XGB de la proteína híbrida (PAS-GGDEF-EAL) de *P. aeruginosa* PA0861 y 3BRE de WspR de *P. aeruginosa* PA3702. En base a la estructura secundaria de los dominios GGDEF de estos modelos, el servidor arrojó la predicción del dominio GGDEF de la proteína CdgA de *A. brasilense* Sp7 como se puede observar en la Figura 18. Se decidió trabajar con el modelo otorgado por parte del modelo PleD de *C. vibrioides* ya que el C-score que otorgó el programa es igual a 1.10 y el RMSD de 2.9+/-2.1 Å, lo cual proporciona una alta confiabilidad al modelo que el servidor bioinformático otorga.



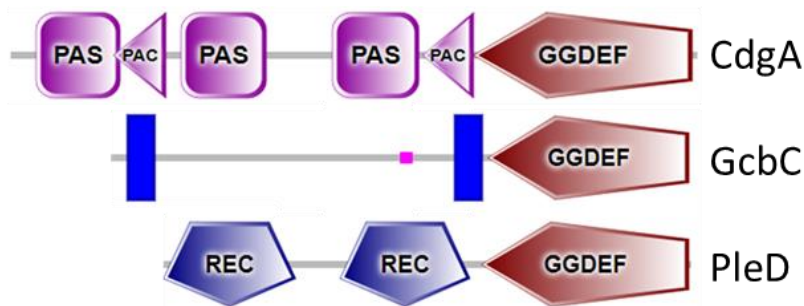
**Figura 18.- Modelo estructural por homología del dominio GG(D/E)EF de la DGC CdgA de *A. brasilense* sp7 otorgado por el programa bioinformático I-TASSER.** A) Estructura terciaria otorgada por el programa bioinformático I-TASSER con un C-score de 1.10, en color morado se presentan las  $\alpha$ -hélices, en color amarillo se presentan las hojas  $\beta$ -plegadas y las azas u horquillas se presentan en color verde. B) En verde se presenta la estructura secundaria del modelo, en rojo se presenta el dinucleótido interaccionando con el motivo RXXD cuyos aminoácidos se presentan en color rosa y en azul se presenta el motivo GGDEF de la proteína CdgA de *A. brasilense* Sp7. C) Acercamiento en el motivo RXXD (rosa) y motivo GG(D/E)EF (azul). La estructura que arroja el I-TASSER cuenta con un C-score de 1.10, resultando muy confiable, con una TM de  $0.86\pm 0.07$  y RMSD de  $2.9\pm 2.1$  Å. Los óvalos en color naranja indican el amino terminal de la secuencia, mientras que los óvalos en color verde indican el carboxilo terminal de la secuencia.

En la Figura 18 se presenta el modelo del dominio GGDEF que proporciona el programa I-TASSER para la proteína CdgA. El dominio presenta 5 hojas  $\beta$ -plegadas y 5  $\alpha$ -hélices

intercaladas, estructura base similar a lo reportado por Schirmer (2016) (Figura 18-A). Así mismo, se presenta el aza con el motivo GG(D/E)EF entre las hojas  $\beta_2$  y  $\beta_3$  -plegadas y 5 aminoácidos río arriba se localiza el motivo RXXD formado por el aza entre la  $\alpha_3$ -hélices y la hoja  $\beta_2$ -plegada (Figura 18-B). En la misma imagen podemos observar la interacción del di-GMPc con el motivo RXXD (sitio inhibitorio), esto se logró al sobreponer el dominio GGDEF de PleD de *C. crescentus* con el dominio GGDEF de CdgA de *A. brasilense* Sp7.

#### 9.4.2.- Análisis estructural comparativo entre los dominios GGDEF de las Diguanilato ciclasas GcbC de *Pseudomonas fluorescens*, PleD de *C. vibrioides* y CdgA de *A. brasilense* Sp7

En la Figura 19 se presenta una imagen que representa los dominios presentes en las DGCs CdgA de *A. brasilense* Sp7, GcbC de *P. fluorescens* y PleD de *C. vibrioides*. Como se observa, las tres proteínas cuentan con dominios accesorios, elementos importantes para regular la actividad catalítica y, solo comparten la similitud de presentar el dominio GGDEF en su carboxilo terminal.

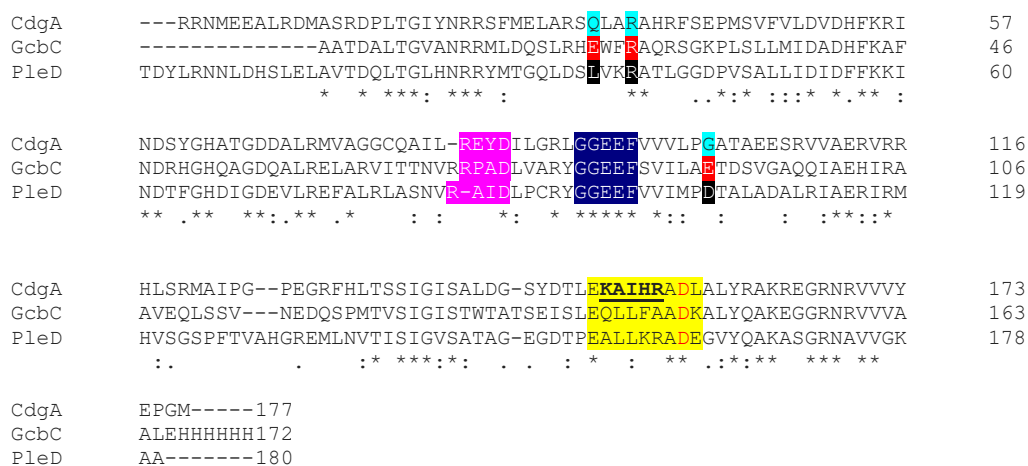


**Figura 19.- Alineamiento de los dominios presentes en las diguanilato ciclasas GcbC, PleD y CdgA otorgadas por el programa bioinformático SMART.** La proteína CdgA de *A. brasilense* Sp7 presenta 3 dominios accesorios (dominios PAS) y el dominio GGDEF. La proteína GcbC de *P. fluorescens* presenta dominios transmembranales más el dominio GGDEF. La proteína PleD de *C. vibrioides* presenta dos dominios REC y el dominio GGDEF.

Una vez comparados los dominios presentes en las DGC CdgA, GcbC y PleD, se prosiguió a realizar la comparación aminoacídica de los dominios GGDEF de las proteínas citadas. De acuerdo con los antecedentes, no solo se buscaron los motivos RXXD y GG(D/E)EF, también se identificaron los aminoácidos importantes para la formación del sitio I y la

secuencia aminoacídica ‘EQLLFAADK’ de la  $\alpha_5$ -hélice, de acuerdo con lo registrado de la proteína GcbC.

El alineamiento de aminoácidos se presenta en la Figura 20, donde se presenta la identificación de los motivos RXXD y GG(D/E)EF subrayados en los colores rosa y azul respectivamente, mientras que los aminoácidos importantes para la formación del sitio I se identifican con colores azul cielo, rojo y negro para CdgA, GcbC y PleD, respectivamente. La secuencia aminoacídica ‘EQLLFAADK’ de la  $\alpha_5$ -hélice se presenta subrayada en color amarillo, y el ácido aspártico en color rojo por su importancia.

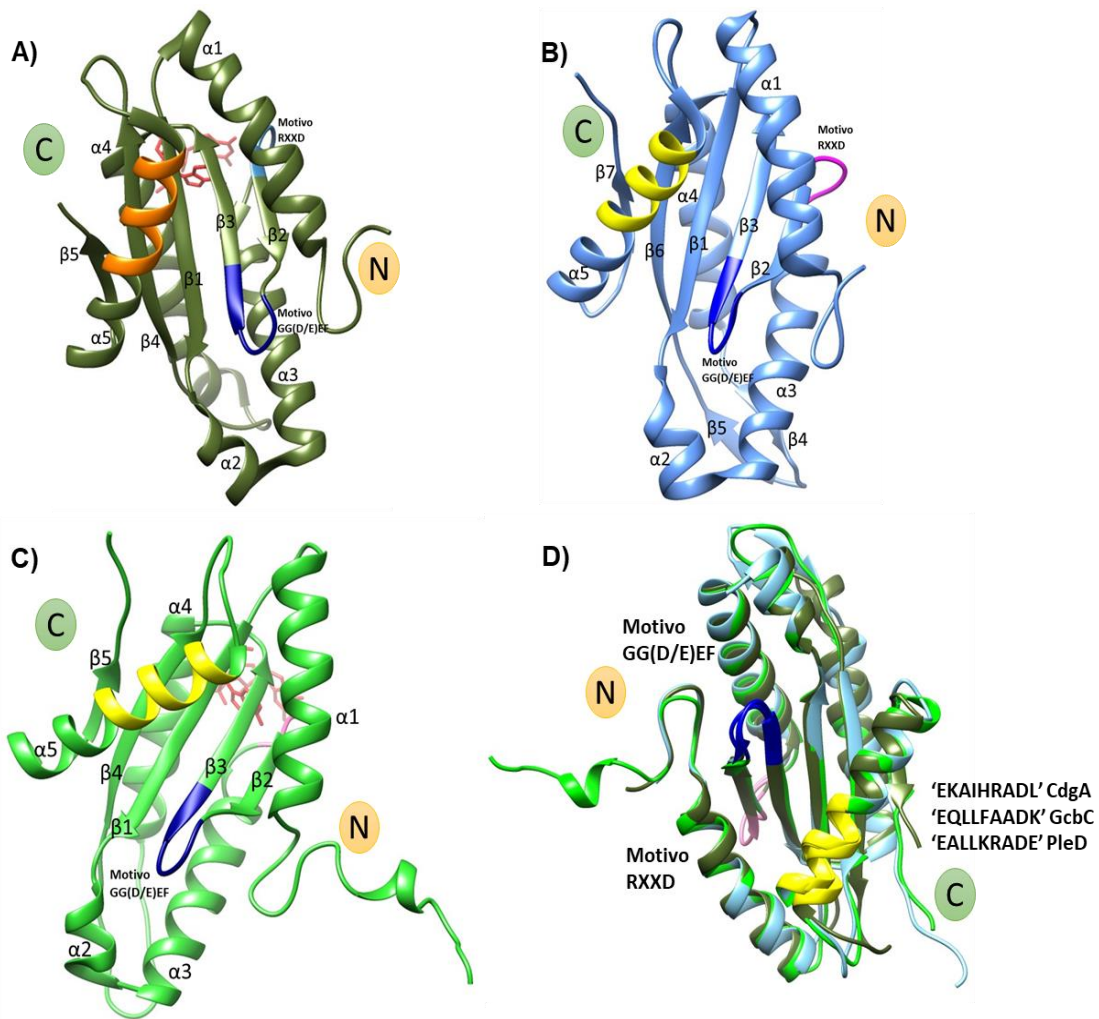


**Figura 20.- Alineamiento aminoacídico entre los dominios GGDEF de las proteínas CdgA de *A. brasilense* Sp7, GcbC de *P. fluorescens* y PleD de *C. vibrioides* por medio del programa CLUSTAL O.** En color rosa se encuentra marcado el motivo RXXD, en color azul se encuentra marcado el motivo GG(D/E)EF y en amarillo se encuentra subrayado la secuencia aminoacídica que Dahlstrom, *et al.*, (2015) proponen como secuencia participante en la interacción DGC-PDE entre GcbC y el dominio alterado EAL de LapD en *P. fluorescens*.

La Figura 20, muestra los motivos GG(D/E)EF y RXXD altamente conservados, lo cual sugiere una actividad catalítica regulada. Por otro lado, la presencia de la secuencia de interacción ‘EQLLFAADK’ de la proteína GcbC de *P. fluorescens*, PleD y CdgA de *C. vibrioides* y *A. brasilense* Sp7 respectivamente, presentan cierta similitud. El ácido aspártico se encuentra conservado; si bien no están conservados algunos de estos pertenecen a la misma familia lo que sugiere que la alfa hélice mantendrá su conformación estructural. El sitio inhibitorio conservado es indispensable también para la interacción con posibles proteínas efectoras, por lo cual no sólo el motivo RXXD debe mantenerse

conservado, sino también la presencia de aminoácidos que intervienen en la formación de la cavidad del sitio I como se muestra en el alineamiento subrayado en azul cielo, rojo y negro (Dahlstrom, *et al.*, 2015 y 2016).

En base a los datos expuestos se realizó un análisis estructural comparativo empleando el programa CHIMERA de los dominios GGDEF de las proteínas CdgA de *A. brasilense* Sp7 otorgado por I-TASSER, PleD de *C. vibrioides* y GcbC de *P. fluorescens* provenientes de la base de datos PDB.



**Figura 21.- Estructura secundaria de los dominios GGDEF de GcbC de *P. fluorescens*, PleD de *C. vibrioides* con CdgA de *A. brasilense* Sp7 y la presencia del di-GMPc. A)** Estructura secundaria del dominio GGDEF de GcbC de *P. fluorescens*. En color verde olivo se presenta la estructura secundaria del dominio GGDEF, en azul cielo se presenta el motivo RXXD, en azul se presenta el motivo GG(D/E)EF, y en color dorado se encuentran señalados los aminoácidos pertenecientes a la secuencia de interacción 'EQLLFAADK' con una resolución de 2.989 Å; **B)** Estructura secundaria del dominio GGDEF de PleD de *C. vibrioides*. En azul claro se presenta la estructura del dominio GGDEF, en rosa, el motivo RXXD, y en azul fuerte se presenta el motivo GG(D/E)EF; en amarillo se indican los aminoácidos 'EALLKRADE' de la

secuencia de interacción de GcbC. **C)** Estructura secundaria del dominio GGDEF de CdgA de *A. brasilense* Sp7. En verde la estructura secundaria del modelo, en rojo se presenta el dinucleótido interaccionando con el motivo RXXD cuyos aminoácidos se presentan en color rosa y en azul el motivo GGEEF. En amarillo la secuencia presenta los siguientes 'EKAIHRADL' aminoácidos alineados por el programa Clustal Omega parcialmente homólogos a la secuencia de interacción de GcbC. **D)** Alineamiento de los dominios GGDEF de CdgA de *A. brasilense* Sp7, GcbC de *P. fluorescens* y PleD de *C. vibrioides*. En color verde olivo se presenta la estructura del dominio GGDEF de GcbC de *P. fluorescens*, en verde la estructura del dominio GGDEF de CdgA de *A. brasilense* Sp7, en color azul cielo/rosa se presenta el motivo RXXD, en color azul rey el motivo GG(D/E)EF, en dorado/amarillo las secuencias de interacción 'EKAIHRADL', 'EQLLFAADK' y 'EALLKRADE', respectivamente para cada proteína. Los óvalos en color naranja indican el amino terminal de la secuencia, mientras que los óvalos en color verde indican el carboxilo terminal de la secuencia.

El análisis de la estructura terciaria mostrado en la Figura 21 muestra que poco difieren los modelos de los dominios GGDEF de las tres proteínas analizadas como se había mostrado previamente en el alineamiento de la Figura 20. Esto sugiere una potencial interacción de estas DGCs con otras proteínas que intervendrían en una probable vía de señalización del di-GMPc.

De acuerdo con el alineamiento de la Figura 20 y a los datos mostrados en la Figura 21, de los tres aminoácidos sugeridos a participar en la formación de la cavidad del sitio I, el segundo aminoácido (arginina o R) se encuentra altamente conservado, sin embargo, los otros dos aminoácidos difieren parcialmente, aunque pertenecen a la misma familia. A pesar de esto, el modelo de la estructura secundaria con la sobreposición de los tres dominios GGDEF parece no diferir mucho entre las tres proteínas, sugiriendo que los cambios o modificaciones de estas proteínas podría deberse a una adaptación del sitio I.

#### **9.5.- Análisis bioinformático de interacción proteína-proteína de la diguanilato ciclasas CdgA con otras proteínas de *A. brasilense* Sp7 por medio del programa bioinformático STRING**

El programa STRING predice a través de su base de datos la posibilidad de interacción entre proteínas, considerando siete evidencias, las cuales son: presencia de fusión (línea roja), contexto genético (línea verde), concurrencia (línea azul), experimental (línea morada), búsqueda de datos (línea amarilla), banco de datos (línea azul cielo) y co-expresión (línea negra), y con un grado de confianza dependiente del grosor de la línea.

Por otro lado, se otorga un grado de puntuación entre cada nódulo (proteína) que se ingresa para interactuar con otras. Esta puntuación va de un rango de entre 0.15 a 0.9, siendo ésta última la de mayor confianza en la predicción que proporciona el programa.

### 9.5.1.- Predicción de interacción proteína-proteína sometiendo el contexto genético de CdgA de *A. brasilense* Sp7

El contexto genético del gen *cdgA* de *A. brasilense* Sp7 (Figura 22-A) muestra cuatro genes con la misma dirección de transcripción, es decir el gen que codifica la treonina-fosfato descaboxilasa (AMK58\_RS08105), ubicada río arriba del gen *cdgA* (AMK58\_RS08100); dos genes río abajo que codifican para la ORF con dominio PAS (AMK58\_RS08095); y la proteína de función desconocida (AMK58\_RS08090), datos registrados en el NCBI. Sin embargo, Rodríguez-Vázquez (2016) realizó el estudio del gen AZOBR\_140291 de *A. brasilense* Sp245, gen homólogo a AMK58\_RS08090 de *A. brasilense* Sp7 (Anexo1), y determinó su función como una MCP (proteína aceptora de grupos metilo) importante para la movilidad quimiotáctica de bacterias.

De acuerdo con las características que presentan los genes de la unidad transcripcional donde se encuentra el gen *cdgA*, se decidió realizar una búsqueda bioinformática por la posibilidad de interacción proteína-proteína en la que podrían participar estas dos proteínas (Figura 22-B).

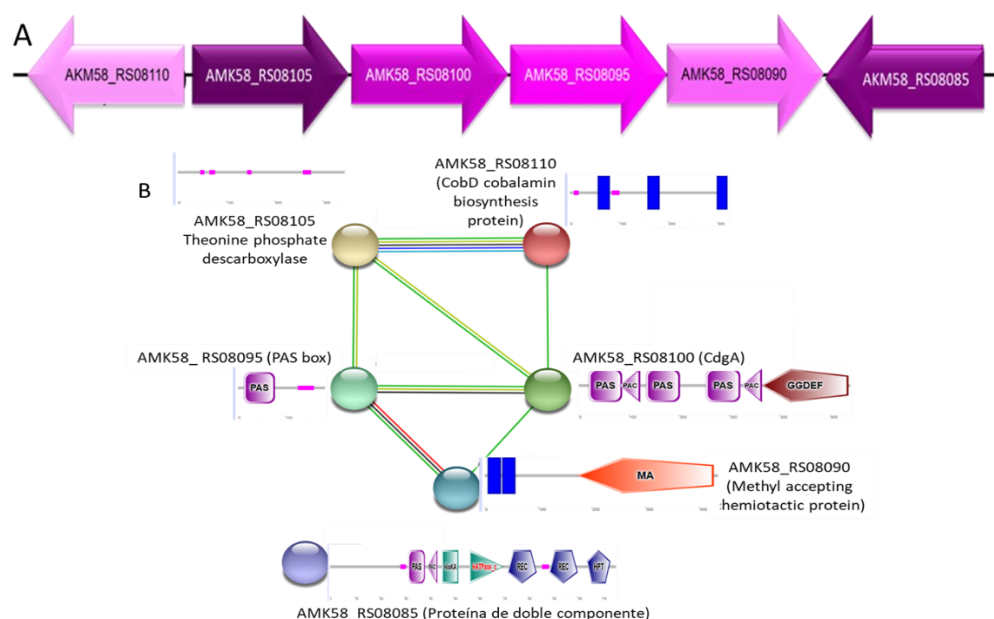


Tabla 2.- Puntaje dado entre la predicción de interacción de proteínas para el contexto genético de la unidad de transcripción de la proteínas CdgA de <i>A. brasilense</i> Sp7 y dos genes aledaños		
Proteína 1	Proteína 2	Score de Interacción
AMK58_RS08105	AMK58_RS08110	0.996
AMK58_RS08105	AMK58_RS08095	0.403
AMK58_RS08105	AMK58_RS08100 (CdgA)	0.588
AMK58_RS08110	AMK58_RS08100 (CdgA)	0.434
AMK58_RS08095	AMK58_RS08100 (CdgA)	0.630
AMK58_RS08095	AMK58_RS08090	0.906
AMK58_RS08090	AMK58_RS08100 (CdgA)	0.538

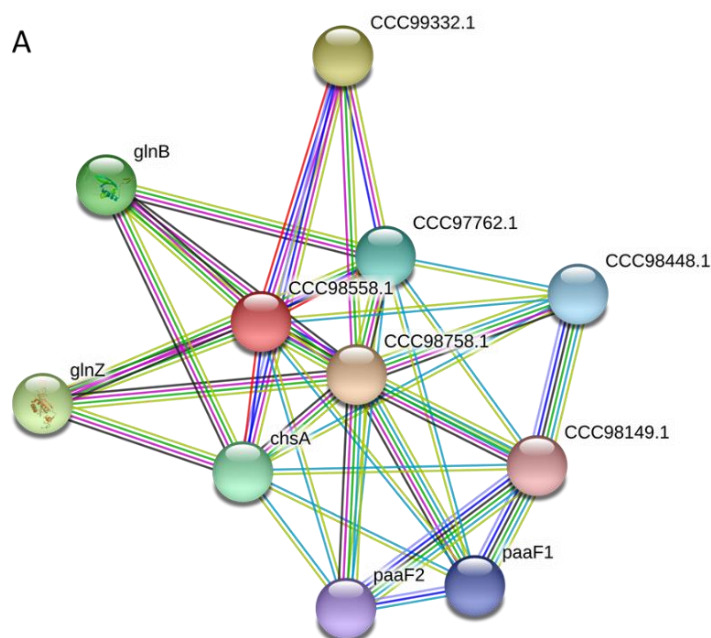
**Figura 22.- Contexto genético de la diguanilato ciclasa A de *Azospirillum brasilense* Sp7 y análisis *in silico* de interacciones entre proteínas utilizando el programa bioinformático STRING.** A) Contexto genético en donde se presenta el gen que codifica para la proteína CdgA en conjunto con otros genes con el siguiente orden: AMK58\_RS08110 (Proteína de biosíntesis de cobalamina CobD), AMK58\_RS08105 (Treonina-fosfato decarboxilasa), AMK58\_RS08100 (Diguanilato ciclasa A), AMK58\_RS08095 (PAS box), AMK58\_RS08090 (Proteína quimiotáctica aceptora de grupos metilos), AMK58\_RS08085 (Cinasa histidínica). B) Predicción de interacción proteica entre la proteína CdgA y otras proteínas presentes en el mismo contexto genético en *A. brasilense* Sp7 utilizando el programa STRING. **Tabla 2-** Puntaje de interacciones proteína-proteína otorgado por el programa STRING. La arquitectura de los dominios de las proteínas se realizó utilizando el programa bioinformático SMART.

En la Figura 22-B se presenta la imagen obtenida por STRING donde se comparten líneas verdes y amarillas entre CdgA y las otras proteínas, que comparten el mismo contexto genético. Además, CdgA comparte una línea negra con el ORF de dominio PAS indicando que pudiera presentarse una co-expresión de los genes que codifican ambas proteínas. Así mismo, (Tabla 2) se presenta la puntuación dada para cada interacción entre nodos, con rango de 0.434-0.630 de confiabilidad de interacción entre la proteína CdgA y ORF de dominio PAS, la cual a su vez podría interaccionar con la MCP, función establecida por Rodríguez-Vázquez (2016).

### 9.5.2.- Predicción de interacción proteína-proteína utilizando el banco de datos del programa STRING

El programa bioinformático STRING tiene acceso a la base de datos del EMBL (por sus siglas en inglés European Molecular Biology Laboratory), CRP (por sus siglas en inglés NNF Center for Protein Research) y SIB (por sus siglas en inglés Swiss Institute of Bioinformatics). Esto nos permitió analizar el contexto genético de una cepa de *A. brasilense*. De acuerdo con Ramírez-Mata, *et al.*, (2018), CdgA está presente tanto en la cepa de *A. brasilense* Sp7 como *A. brasilense* Sp245.

El genoma de la cepa de *A. brasilense* Sp7 no se localizó en la base de datos del STRING, por lo cual se empleó la información registrada de *A. brasilense* Sp245 y, así realizar la búsqueda de interacciones proteína-proteína con CdgA, posteriormente empleando el algoritmo BLAST ubicar las secuencias aminoacídicas de *A. brasilense* Sp245 en *A. brasilense* Sp7. En la Figura 23-A, se muestra la predicción CdgA (CCC98558.1) con el grupo de proteínas siguientes: una cinasa histidínica, 3 fosfodiesterasas, 2 reguladores transcripcionales y 4 enoyl-CoA hidratasas. Una vez obtenidos estos datos se procedió a encontrar las proteínas homólogas en el genoma de *A. brasilense* Sp7 disponible en la base de datos del NCBI. Los datos se muestran en la Figura 23 (Tabla 3), los números de acceso de las secuencias de ambas cepas, su peso molecular y el tipo de proteína.



**Tabla 3.- Números de acceso de las proteínas que posiblemente interaccionen con la proteína CdgA de *A. brasilense* Sp245 y las proteínas homólogas en *A. brasilense* Sp7.**

<i>A. brasilense</i> Sp245	<i>A. brasilense</i> Sp7	Peso Molecular	Proteína
AZOBR_140288 (CCC98558.1)	AMK58_RS08100	60 kDa	Diguanilato ciclasa CdgA
AZOBR_150176 (CCC98758.1)	AMK58_RS10050	91.2 kDa	Cinasa Histidínica
AZOBR_200037 (CCC99332.1)	AMK58_RS03020	44.05 kDa	Fosfodiesterasa
AZOBR_100148 (CCC97762.1)	AMK58_RS05685	28 kDa	Fosfodiesterasa
AZOBR_100230 ( <i>glnB</i> )	AMK58_RS06055	12.3 kDa	Regulador Transcripcional

AZOBR_10503 ( <i>glnZ</i> )	AMK58_RS13700	12.27 kDa	Regulador Transcripcional
AZOBR_100013 ( <i>chsA</i> )	AMK58_RS05115	61.94 kDa	Fosfodiesterasa
AZOBR_100375 ( <i>paaF2</i> )	AMK58_RS06715	28.56 kDa	Enoyl-CoA hidratasa
AZOBR_10280 ( <i>paaF1</i> )	AMK58_RS00900	27.8 kDa	Enoyl-CoA hidratasa
AZOBR_110123 (CCC98149.1)	AMK58_RS07475	38.82 kDa	Enoyl-CoA hidratasa isomerasa
AZOBR_140178 (CCC98448.1)	AMK58_RS08575	26.87 kDa	Enoyl-CoA hidratasa

**Tabla 4.- Puntaje de predicción de interacción dado entre proteínas del banco de datos de STRING y CdgA de *A. brasilense* Sp7**

Proteína 1	Score	Proteína 2	Dominios de las proteínas
AMK58_RS08100 (CdgA)	0.980	AMK58_RS10050	
AMK58_RS08100 (CdgA)	0.803	AMK58_RS03020	
AMK58_RS08100 (CdgA)	0.756	AMK58_RS05685	
AMK58_RS08100 (CdgA)	0.801	AMK58_RS06055	
AMK58_RS08100 (CdgA)	0.801	AMK58_RS13700	
AMK58_RS08100 (CdgA)	0.778	AMK58_RS05115	
AMK58_RS08100 (CdgA)	0.752	AMK58_RS06715	
AMK58_RS08100 (CdgA)	0.752	AMK58_RS00900	
AMK58_RS08100 (CdgA)	0.742	AMK58_RS07475	
AMK58_RS08100 (CdgA)	0.752	AMK58_RS08575	

**Tabla 5.- Tipo de evidencia que permite la predicción de interacción dada entre proteínas del banco de datos de STRING y CdgA de *A. brasilense* sp7**

Proteína 1	Proteína 2	Tipo de interacción de acuerdo al programa bioinformático STRING
AMK58_RS08100 (CdgA)	AMK58_RS10050	<ul style="list-style-type: none"> <li>• Contexto genético</li> <li>• Búsqueda de datos</li> </ul>
AMK58_RS08100 (CdgA)	AMK58_RS03020	<ul style="list-style-type: none"> <li>• Búsqueda de datos</li> <li>• Experimental</li> <li>• Concurrencia</li> <li>• Presencia de fusión</li> </ul>
AMK58_RS08100 (CdgA)	AMK58_RS05685	<ul style="list-style-type: none"> <li>• Concurrencia</li> <li>• Presencia de fusión</li> </ul>
AMK58_RS08100 (CdgA)	AMK58_RS06055	<ul style="list-style-type: none"> <li>• Contexto genético</li> </ul>

		<ul style="list-style-type: none"> <li>• Búsqueda de datos</li> <li>• Presencia de fusión</li> <li>• Co-expresión</li> </ul>
AMK58_RS08100 (CdG A)	AMK58_RS13700	<ul style="list-style-type: none"> <li>• Contexto genético</li> <li>• Búsqueda de datos</li> <li>• Presencia de Fusión</li> <li>• Concurrencia</li> </ul>
AMK58_RS08100 (CdG A)	AMK58_RS05115	<ul style="list-style-type: none"> <li>• Presencia de Fusión</li> <li>• Concurrencia</li> </ul>
AMK58_RS08100 (CdG A)	AMK58_RS06715	<ul style="list-style-type: none"> <li>• Contexto genético</li> <li>• Banco de datos</li> </ul>
AMK58_RS08100 (CdG A)	AMK58_RS00900	<ul style="list-style-type: none"> <li>• Contexto genético</li> <li>• Banco de datos</li> </ul>
AMK58_RS08100 (CdG A)	AMK58_RS07475	<ul style="list-style-type: none"> <li>• Contexto genético</li> <li>• Banco de datos</li> </ul>
AMK58_RS08100 (CdG A)	AMK58_RS08575	<ul style="list-style-type: none"> <li>• Contexto genético</li> <li>• Banco de datos</li> </ul>

**Figura 23.- Predicción de interacción proteína-proteína utilizando el banco de datos del programa STRING.** A) El programa STRING cuenta con la base de datos del genoma de *A. brasilense* Sp245 y CdG A de *A. brasilense* Sp7. **Tabla 3** de análisis bioinformático en búsqueda de secuencias homologas de proteínas identificadas en el genoma de *A. brasilense* Sp245 en el genoma de *A. brasilense* Sp7 y su número de acceso de acuerdo a BLAST. **Tabla 4** de puntaje de interacciones proteína-proteína otorgado por el programa STRING. La arquitectura de los dominios de las proteínas se realizó utilizando el programa bioinformático SMART. La arquitectura de los dominios de las proteínas se realizó utilizando el programa bioinformático SMART. **Tabla 5** muestra el tipo de evidencias para determinar la posible interacción proteína-proteína otorgado por el programa STRING.

La Figura 23-C presenta el score de interacción de la proteína CdG A con las proteínas encontradas. El score oscila desde 0.742-0.980, los cuales son puntuaciones de confiabilidad muy altos que el programa otorga y esto a la vez nos sugiere una probabilidad de interacción alta.

Asimismo, la Figura 23 muestra la Tabla 5, con el tipo de evidencias que permite la predicción de interacción. Las más destacadas son las líneas verdes y amarillas (de contexto genético y búsqueda de datos respectivamente) que se presentan en la mayor parte de interacciones entre proteínas que arroja el programa. Cuatro de las once proteínas que se presentan como posibles proteínas de interacción con la proteína CdG A cuentan con 4 o más tipos de evidencias por parte del programa mientras que las siete proteínas restantes cuentan con dos tipos de evidencias que permite la predicción de interacción proteína-proteína.

### 9.5.3.- Predicción de interacción proteína-proteína de CdgA y las presuntas DGCs *A. brasilense* Sp7

En base a los resultados obtenidos por el programa STRING para los dos últimos apartados del contexto genético donde se encuentra el gen *cdgA* que codifica para la proteína CdgA y el banco de datos del genoma de *A. brasilense* Sp245, se decidió retar contra otras Diguanilato ciclasas de *A. brasilense* Sp7. Se conoce que *A. brasilense* Sp7 cuenta con la presencia de 20 DGC, 5 PDE con dominio EAL y 9 proteínas híbridas, datos obtenidos del análisis bioinformático de genes que codifican para presuntas DGCs y PDE de dominio EAL en diferentes genomas de cepas de *Azospirillum* (Ramírez-Mata, *et al.*, 2018).

Previamente, se separaron en subgrupos a las DGC dependiendo los dominios accesorios que presentaran para la mejor evaluación de la presunta interacción proteína-proteína. Aunque algunas proteínas cuentan con más de un tipo de dominio accesorio, sólo se analizan en un subgrupo y no se realizó repetición de éstas en los demás subgrupos. Como primer subgrupo se evaluó a las DGC que presentaran dominios PAS (Ramírez-Mata, *et al.*, 2016). Este subgrupo de cinco DGC de *A. brasilense* Sp7, cuyo número de acceso es: AMK58\_RS21735 (DCG con dos dominios PAS), AMK58\_RS17510 (DGC con un dominio transmembranal y uno PAS), AMK58\_RS00200 (DGC con un dominio PAS), AKM58\_RS12258 (DGC con dominio PAS y GAF) y CdgA (DGC con tres dominios PAS). Estas proteínas fueron analizadas por STRING y se obtuvieron los siguientes resultados que se presentan en la Figura 25.

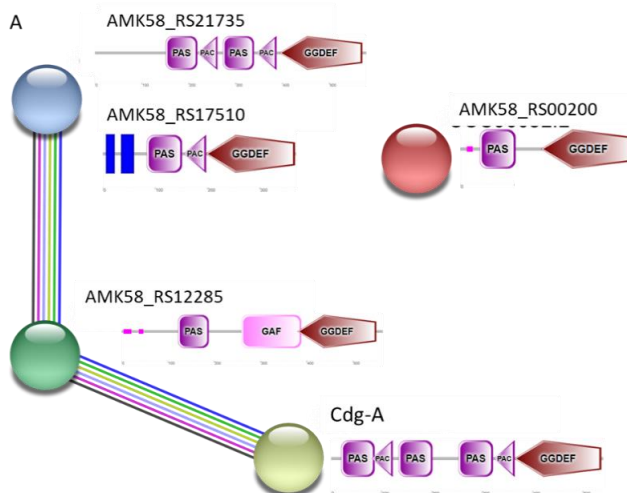


Tabla 6.- Puntaje de predicción de interacción dado entre DGC con dominios accesorios PAS y CdgA de *A. brasilense* Sp7

Proteína 1	Puntaje de Interacción	Proteína 2
AMK58_RS08100 (CdgA)	0.425	AMK58_RS12285
AMK58_RS12285	0.433	AMK58_RS17510 AMK58_RS21735

**Figura 24.- Predicción de interacciones proteína-proteína entre la proteína CdgA y otras DGC con dominios PAS en *A. brasilense* Sp7 utilizando el programa STRING. Imagen A)** Análisis de interacción proteína-proteína ingresando las secuencias aminoacídicas de diguanilato ciclasas con dominio PAS versus la secuencia aminoacídica de la proteína CdgA de *A. brasilense* Sp7. Tabla 6. Puntaje de interacciones proteína-proteína otorgado por el programa STRING. La arquitectura de los dominios de las proteínas se realizó utilizando el programa bioinformático SMART.

Únicamente la proteína AMK58\_RS12285 presentó una presunta interacción con CdgA y un puntaje de 0.425, con diferentes tipos de evidencia, como se indica en la Figura 24.

Las proteínas AMK58\_RS17510 Y AKM58\_RS21735 tendrían una interacción con la proteína AMK58\_RS12285 con un puntaje de 0.433, mientras que AMK58\_RS00200 no presenta ninguna interacción.

Posteriormente se prosiguió a realizar un alineamiento aminoacídico de los dominios GGDEF de las DGC con dominios accesorios PAS como se observa en la Figura 25 para identificar sus respectivos motivos RXXD (color rosa) y GG(D/E)EF (color azul), así como los aminoácidos importantes para la formación del sitio I (color morado) y finalmente la secuencia ‘EQLLFAADK’ (color amarillo). Como se puede observar los motivos RXXD y GGDEF se encuentran totalmente conservados, los aminoácidos importantes para la formación del sitio I se encuentran parcialmente conservados y el ácido aspártico esencial dentro de la secuencia ‘EQLLFAADK’ se encuentra conservado.

AMK58_RS21735	---LRRSEAEKDHLAHHDALTGLPNRRLLFDRLLDMLARSRRSRGRSLALLFIDLDDGFKAV	57
AMK58_RS17510	---QHHLEERLRYLAGHDALTGLCNRIREFDRLEHAIARAKREDKSFALLFLDDGFKEV	57
AMK58_RS12285	---HLAYQERLRRLSERDGLTGLFNRRFFERLEEVLGRSD--SGPSALLYVDLDFKAV	55
GcbC	-----AATDALTGVANRRMLDQSLRHWFQAQRSGKPLSLLMIDADHFKAF	46
PleD	TDYLRNNLDHSLELAVTDQLTGLHNRRYMTGQLDSLVKRAATLGGDPVSALLIDIDFFKKI	60
CdgA	---RRNMEALRDMASRDPLTGTYNRRSFMELARSCLAAHHRFSEPMSEVFLVDVDFKRI	57
AMK58_RS00200	---LKESEERLRRLSITDSLTVSNRRHFVEAAEQELARARRHGRPVTLMLDLDFKSI	57
	: * ***: ** : * : . : : * * * .	
AMK58_RS21735	NDRMGHDCGDEVLRIAAARIAGCL-RNSDITVARVGGDEFITILEDVSSAADAGVAAKIV	116
AMK58_RS17510	NDTHGHQKGDEVLVTTAQRSLGAL-RNSDITVARVGGDEFVAVLLECDISVENARAVAEKVI	116
AMK58_RS12285	NDLHGHQQGDAVLMAIGLLTNGV-RPGDLPGRVGGDEFVWLWGRDDE-AQAVAVAERLL	113
GcbC	NDRHGHQAGDQALRELARVITTNVRRPADLVARYGGEEFVSVILATDS-VGAQQIAEHIR	105
PleD	NDTFGHDI GDEVLRREFALRLASNV-RATDLPCTRYGGEEFVIVIMPETAL-ADALRIAERIR	118
CdgA	NDSYGHATGDDALRMVAGGCQAIL-REYDILGRVGGEEFVIVVLPQATA-EESRVVAERVR	115
AMK58_RS00200	NDTHGHAVGDEALRTFTAACRALL-RENDLLGRVGGEEFAILLPEITDI-AGARMAERIR	115
	** ** * * . * : * * * ** * : : : * : :	
AMK58_RS21735	EAVALPIPVQGPVQNTNAVIGASVGIALSPA-DGSDARSLLVADGAMYRAKQAGRGTHI	175
AMK58_RS17510	DSIDRPLPGL-----DLSVSASVGMAFYPS-DGADAGVLLRCADQAMYRAKRLGKNRLE	169

```

AMK58_RS12285      AGMAA-LRHLSAGPDK--PLGLSIGIAVHQQV-GSETLQELTDRADSAM YDAKKRGKGHYA      169
GcbC               AAVEQ-LS---SVNEDQSPMTVSIIGISTWTATSEISLEQLLFAADKALYQAKEGGRNRVV      161
PleD               MHVSG-SPFTVAHGREMLNVTISIGVSATAG-EGDTP EALLKRADEGVYQAKASGRNAVV      176
CdG A              RHLSR-M--AIPGPEGRFHLTSSIGISALDG-SYDTLEKAIHRA DLALYRAKREGNRNVV      171
AMK58_RS00200     RRTAE-L--AVPAGDETVRFTVSI GVACCAA-GTRD V D A M L S S A D E A L Y R A K A A G R N R V V      171
                   .  * : * : :                               * * . : * * * * : .

AMK58_RS21735     FY-----              177
AMK58_RS17510     LG-----              171
AMK58_RS12285     VA-----              171
GcbC              VAALEHHHHHHH          172
PleD              GKAA-----            180
CdG A             VYEPGM-----         177
AMK58_RS00200     CA-----              173

```

**Figura 25.- Alineamiento de los dominios GGDEF de las secuencias de DGC con dominios PAS y los dominios GGDEF de las DGC CdG A de *A. brasilense* Sp7 con GcbC de *P. fluorescens*, y PleD de *C. vibrioide*.** En rosa se encuentra señalado el motivo RXXD, en azul se presenta el motivo GG(D/E)EF y en amarillo se presentan los 9 aminoácidos de la secuencia de interacción identificada en GcbC, donde se destaca el aminoácido ácido aspártico más cercano al carboxilo terminal del dominio. Subrayados en morado se presentan los aminoácidos importantes para la formación del sitio I. Si bien no se comparte una identidad del 100% en estos aminoácidos, la mayor parte de ellos pertenecen a la misma familia y se conservan las argininas entre las secuencias para la formación del sitio I.

El siguiente grupo por analizar fue DGC con dominios accesorios GAF que correspondió únicamente a AMK58\_RS25025. Los resultados que arroja el programa para la interacción de la proteína AMK58\_RS25025 con CdG A son 6 de 7 evidencias que se indican en la Figura 26-A y el puntaje es de 0.425 de confianza de un rango de 0.1-0.9 (Figura 26-Tabla 7).

A)

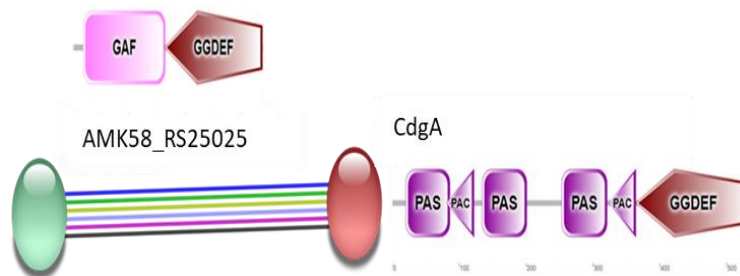


Tabla 7.- Puntaje de predicción de interacción dado entre DGC con dominio accesorio GAF y CdG A de <i>A. brasilense</i> Sp7		
Proteína 1	Puntaje de Interacción	Proteína 2
AMK58_RS08100 (CdG A)	0.425	AMK58_RS25025

**Figura 26.- Predicción de interacción proteína-proteína entre la proteína CdG A y la DGC con dominio GAF en *A. brasilense* Sp7 utilizando el programa STRING.** Imagen A) Análisis de interacción proteína-proteína ingresando la secuencia aminoacídica de la diguanilato ciclase con dominio GAF versus la secuencia aminoacídica de la proteína CdG A de *A. brasilense* Sp7. Tabla 7. Score de interacciones proteína-proteína otorgado por el programa STRING. La arquitectura de los dominios de las proteínas generadas utilizando el programa bioinformático SMART.

Además, en el alineamiento presentado en la Figura 27 los motivos RXXD y GG(D/E)EF se encuentran conservados, la secuencia de aminoácidos para la formación del sitio I muestra la conservación de 2 de 3 de los aminoácidos, pero presenta una alteración la arginina que previamente se había conservado en los alineamientos cambiándose por una metionina. Finalmente, el ácido aspártico de la secuencia ‘EQLLFAADK’ que permite la interacción con otras proteínas se mantiene conservado.

AMK58_RS21735	---FRDLNAALESYSYTDALDTGLPNRRAIVAEHLR <small>LF</small> MAERAGQSVLIGFIDLDDGFKAI	57
CdgA	---RRNMEEALRDMASRDPLTGIYNRRSFMELARS <small>CLAF</small> AHRFSEPMSEVFLVDVDFKRI	57
GcbC	-----AATDALTGVANRRMLDQSLRH <small>WFF</small> AQRSGKPLSLLMIDADHFKAF	46
PleD	TDYLRNLDHSLELAVTDQLTGLHNRRYMTGQLDS <small>IVK</small> EATLGGDPVSALLIDIDFFKKI	60
	: * ***: *** :	* .. : .:* * * :
AMK58_RS21735	NDTHGHEAGDAFLTKVGQRIGAGL- <small>RAGD</small> TLGR <small>LGDEF</small> IIIVGM <small>CA</small> TPGAEGEDAADALR	116
CdgA	NDSYGHATGDDALRMVAGGCQAIL- <small>REYD</small> ILGR <small>GGEEF</small> VVVLP <small>ATAE</small> -ESRVVAERV	115
GcbC	NDRHGHQAGDQALRELARVITTNVRR <small>PAD</small> LVARY <small>GGEEF</small> SVILA <small>TDSV</small> -GAQQIAEHIR	105
PleD	NDTFGHDIGDEVLREFALRLASNVR- <small>AIDL</small> PCRY <small>GGEEF</small> VVIM <small>P</small> TALA-DALRIAERIR	118
	** .** * * .. : :	* * **:* * : : . * : *
AMK58_RS21735	ERVEPLIRGVYRLDSCAFDYPGASVGVVSADP-GTTTT <small>DDALRAADA</small> AMYVQKKIRRAAR	175
CdgA	RHLSRMAIPG----PEGRFHLTSSIGISALDG-SYDTLE <small>KAIHRADL</small> ALYRAKREGGRNV	170
GcbC	AAVEQLSSV-----NEDQSPMTVISIGISTWTATSEISLE <small>EQLLFAADK</small> ALYQAKEGGRNV	160
PleD	MHVSGPSPTVAH--GREMLNVTISIGVSATAG-EGDTP <small>EALLKRADE</small> GVYQAKASGRNAV	175
	: .	*:* : : : : * * .:* * *
AMK58_RS21735	-----	175
CdgA	VVYEPMG-----	177
GcbC	VVALEHHHHHH	172
PleD	VGKAA-----	180

**Figura 27.- Alineamiento de los dominios GGDEF de la secuencia DGC con dominio GAF y los dominios GGDEF de las DGC CdgA de *A. brasilense* Sp7, GcbC de *P. fluorescens* y PleD de *C. vibrioide*.** En rosa se encuentra señalado el motivo RXXD, en azul se presenta el motivo GG(D/E)EF y en amarillo se presentan los 9 aminoácidos de la secuencia de interacción identificada en GcbC, donde se destaca el aminoácido ácido aspártico más cercano al carboxilo terminal del dominio. Subrayados en morado están los aminoácidos importantes para la formación del sitio I. Si bien no se comparte una homología en estos aminoácidos, la mayor parte de ellos pertenecen a la misma familia, sin embargo, como se puede observar no existe la conservación de argininas entre las secuencias para la formación del sitio I (cambio de una arginina por una metionina).

El tercer grupo de DGC consistió en aquellas proteínas con dominios REC. Este grupo lo integran 5 presuntas DGCs con las siguientes características: AMK58\_RS13115, AMK58\_RS12695, AMK58\_RS05940 y AMK58\_RS16780 (contienen dominio REC); mientras que AMK58\_RS16070 (presenta dos dominios REC).

El análisis mostró que AMK58\_RS12695 y AMK58\_RS16070 no interaccionan con otras proteínas. En tanto que CdgA tendría una posible interacción con AMK58\_RS13115 ya que el programa otorga 6 de 7 evidencias, con un puntaje de 0.505 (Figura 28, Tabla 8).

A)

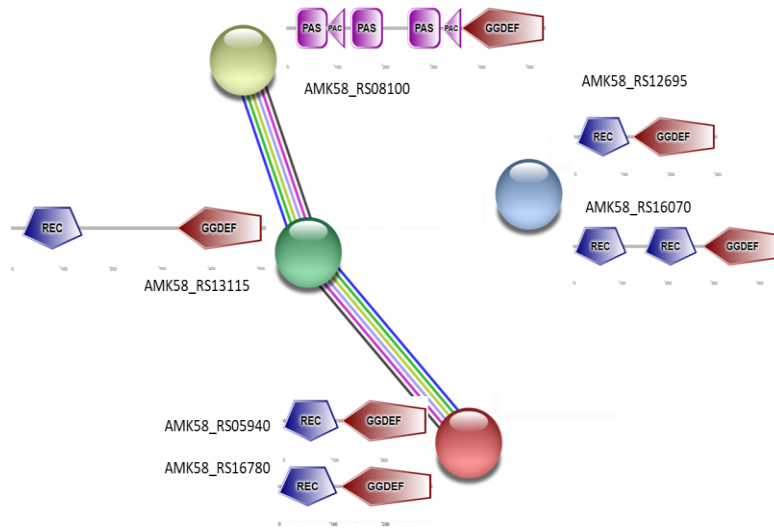


Tabla 8.- Puntaje de predicción de interacción dado entre DGC con dominios accesorios REC y CdgA de <i>A. brasilense</i> Sp7		
Proteína 1	Score de Interacción	Proteína 2
AMK58_RS08100 (CdgA)	0.505	AMK58_RS13115
AMK58_RS13115	0.413	AMK58_RS05940
AMK58_RS13115	0.413	AMK58_RS16780

**Figura 28.- Predicción de interacciones proteína-proteína entre la proteína CdgA y DGC con dominios accesorios REC en *A. brasilense* Sp7 utilizando el programa STRING. A) Análisis de interacción proteína-proteína ingresando las secuencias aminoacídicas de DGC con dominios accesorios REC y la secuencia de la proteína CdgA de *A. brasilense* Sp7. Tabla 8. Puntuación de interacciones proteína-proteína otorgado por el programa STRING. La arquitectura de los dominios de las proteínas se realizó utilizando el programa bioinformático SMART.**

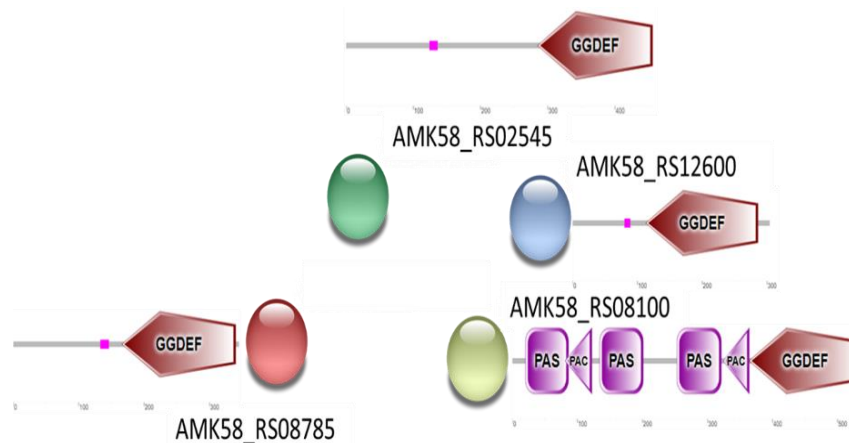
El análisis con el Clustal Omega que se muestra en la Figura 29, indica la presencia de los motivos GG(D/E)EF y RXXD conservados. De igual forma, se presentan los aminoácidos importantes para la formación del sitio I. Si bien no están todos conservados pertenecen al mismo grupo, a excepción de los aminoácidos presentes en CdgA Q33 y G101 mientras que en PleD y AMK58\_RS16070 presentan L36 y L33 respectivamente, y el grupo de argininas se mantiene conservado.

PleD	TDYLRNNLDHSLELAVTDQLTGLHNRRYMTGQLDS	VK	ATLGGDPVSALLIDIDFFKKI	60		
AMK58_RS16070	---LRANYEQSLSMALTDSDLTGVEFNRRYINAHLPRLLE	RAIDN	NHKPVAILLFDIDHFKV	57		
AMK58_RS12695	---LKRQTDLLRTLSTFNGLTGIANRRRFDEVLLREWR	CGRAQL	PPLSLIMLDVDQFKPY	57		
GcbC	-----AATDALTGVANRRMLDQSLRH	WF	AQRSGKPLSLLMIDADHFKAF	46		
CdgA	---RRNMEALRDMASRDPLTGIYNRRSFMELARS	CLAH	AHRFSEPMVSFVLDVDHFKRI	57		
AMK58_RS05940	---ITEMHRLLVQAQTDPLTGVANRRALLDRADQ	EIR	RALRSRQPLSLLIADIDHFKGI	57		
AMK58_RS16780	---VVDMQRQLERLATDTSLTGALNRRGFARAAD	EVAR	AQRSGQPLSAIMLDIDHFKKV	57		
AMK58_RS13115	---LEAANAKLERLATVDPLTAVWNRHFLLDLAA	ELGR	AHRHGRRLSVILLDLDNFKAV	57		
	: * ** . ** :	*	:: : * * **			
PleD	NDFGHGHDIGDEVLREFALRLAS-NV	RAID	LPCRYGGEEFV	VIMP	TALADALRIAERIRM	119
AMK58_RS16070	NDSYGHIVGDEVLKEVSSRASR-NL	RTFD	LVARLGGEEF	VVILP	TDADAALTVARLRT	116
AMK58_RS12695	NDHYGHQAGDECLRAVAQLLAEQMM	RPGL	IARYGGEEF	VCLLE	ETDEDGAVQVAERLRR	117
GcbC	NDRHGHQAGDQALRELARVITTNVR	RPAD	LVARVGGEEF	SVILA	TDSVGAQQIAEHIRA	106

CdgA	NDSYGHATGDDALRMVAGGCA-ILREYDILGRLLGEEFVVVLPATAEESRVAERVRR	116
AMK58_RS05940	NDSRGHAAGDEALQRFVRTLGD-ALRPGDLIGRYGEEFVVVLENTALADASVVAERLRR	116
AMK58_RS16780	NDTYGHATGDEAIRTVVRVCRS-MVRGADVLGRLLGEEFVAILLPDTPQQGAALLAERLRR	116
AMK58_RS13115	NDGHGHAAGDEALRTVVVRARE-ALRPSDQLAREGEEFVALLPEVTDLAGARIVAERVRA	116
	** ** ** : : . * * * * * : : : : * : *	
PleD	HVSGSPFTVAH-GREMLNVTISIGVSATA-GEGDTEALLKRADEGVYQAKASGRNAVVG	177
AMK58_RS16070	RIADTLFKVSA-DVGEIPVTVSIGVAVGG-RLGDTAEGLIRRADEALYEAKRAGRNCCTVA	174
AMK58_RS12695	TVADRRLPHAV-SHVADHVTISLGVATAREMLDDTPERLTQLADGLLYEAKRAGRNRVCS	176
GcbC	AVEQLSSV----NEDQSPMTVSIIGISTWTATSEISLEQLLFAADKALYQAKEGGRNRVVV	162
CdgA	HLSRMAIPG---PEGRFHLTSSIGISALD-GSYDTLEKAIHRADLALYRAKREGRNRVVV	172
AMK58_RS05940	FVAEQEMGP---NGDSFRMTASFGVAPIALDAPNGLDVALRVADAALYRAKAEGRNRVVP	173
AMK58_RS16780	ALAEETDVRTANGAGSVLFTFTVSIIGVSALK-PGENGVAAVLARADEALYRAKNSGRNRVMC	175
AMK58_RS13115	AIAGSPVVT---DGRAIPITASLGVAEWN-TAEASTELTLRRADAALYEAKLAGRDRVCV	172
	: . * * : * : * * : * * : * * : * * : *	
PleD	KAA-----	180
AMK58_RS16070	D-----	175
AMK58_RS12695	G-----	177
GcbC	AALEHHHHHH	172
CdgA	YEPGM-----	177
AMK58_RS05940	S-----	174
AMK58_RS16780	D-----	176
AMK58_RS13115	A-----	173

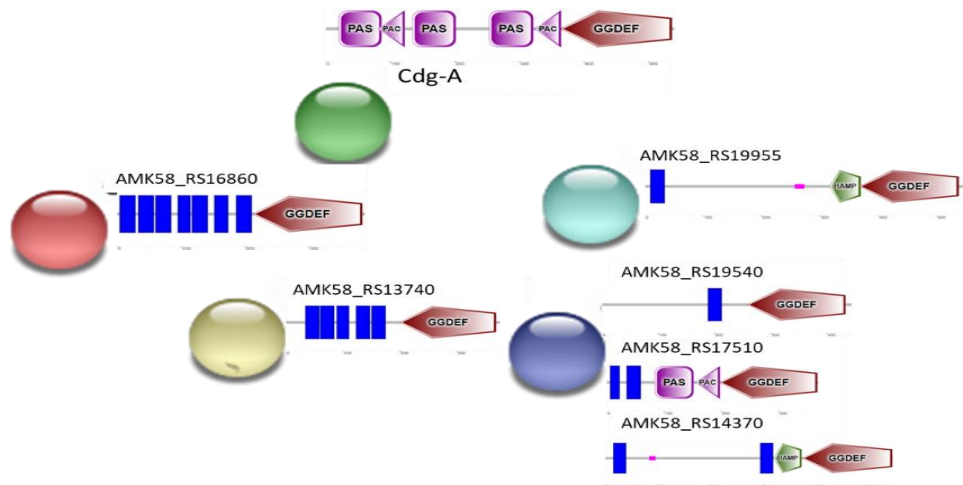
**Figura 29.- Alineamiento de los dominios GGDEF de las DGC con dominios accesorios REC y los dominios GGDEF de las DGC CdgA de *A. brasilense* Sp7, GcbC de *P. fluorescens* y PleD de *C. vibrioide*.** En rosa se encuentra señalado el motivo RXXD, en azul se presenta el motivo GG(D/E)EF y en amarillo se presentan los 9 aminoácidos de la secuencia de interacción identificada en GcbC, donde se destaca el ácido aspártico más cercano al carboxilo terminal del dominio. Subrayados en morado se presentan los aminoácidos importantes para la formación del sitio I, si bien no se comparte una homología en estos aminoácidos, la mayor parte de ellos o son igual o pertenecen a las mismas familias de los aminoácidos presentes en GcbC, PleD y CdgA para este sitio. En el primer grupo tenemos la presencia de E, L y Q para GcbC, PleD y CdgA respectivamente y la presencia de E y L en las demás secuencias, en el segundo grupo se presente la conservación de argininas, y por último se presenta E, D, y G para GcbC, PleD y CdgA respectivamente en conjunto con D, E, N en las demás secuencias.

El cuarto grupo de DGC analizado en el programa STRING consiste en aquellas proteínas con dominio único: AMK58\_RS02545, AMK58\_RS12600, AMK58\_RS08785 y CdgA (AMK58\_RS08100). En la Figura 30 se muestra que el programa indica que las proteínas no interaccionan.



**Figura 30.- Predicción de interacción proteína-proteína entre la proteína CdgA y DGC con dominio GGDEF único en *A. brasilense* Sp7 utilizando el programa STRING.** Las diguanilato ciclasas de dominio único de *A. brasilense* Sp7 no presentan interacción alguna por medio del programa STRING. La arquitectura de los dominios de las proteínas se realizó utilizando el programa bioinformático SMART.

El último grupo de DGC analizado consiste en aquellas que contengan dominios transmembranales. Este grupo se compone por: AMK58\_RS16860, AMK58\_RS19955, AMK58\_RS13740, AMK58\_RS19540, AMK58\_17510, AMK58\_RS14370. Como se puede observar en la Figura 31, el programa STRING no muestra la predicción de interacción para las secuencias de estas proteínas.

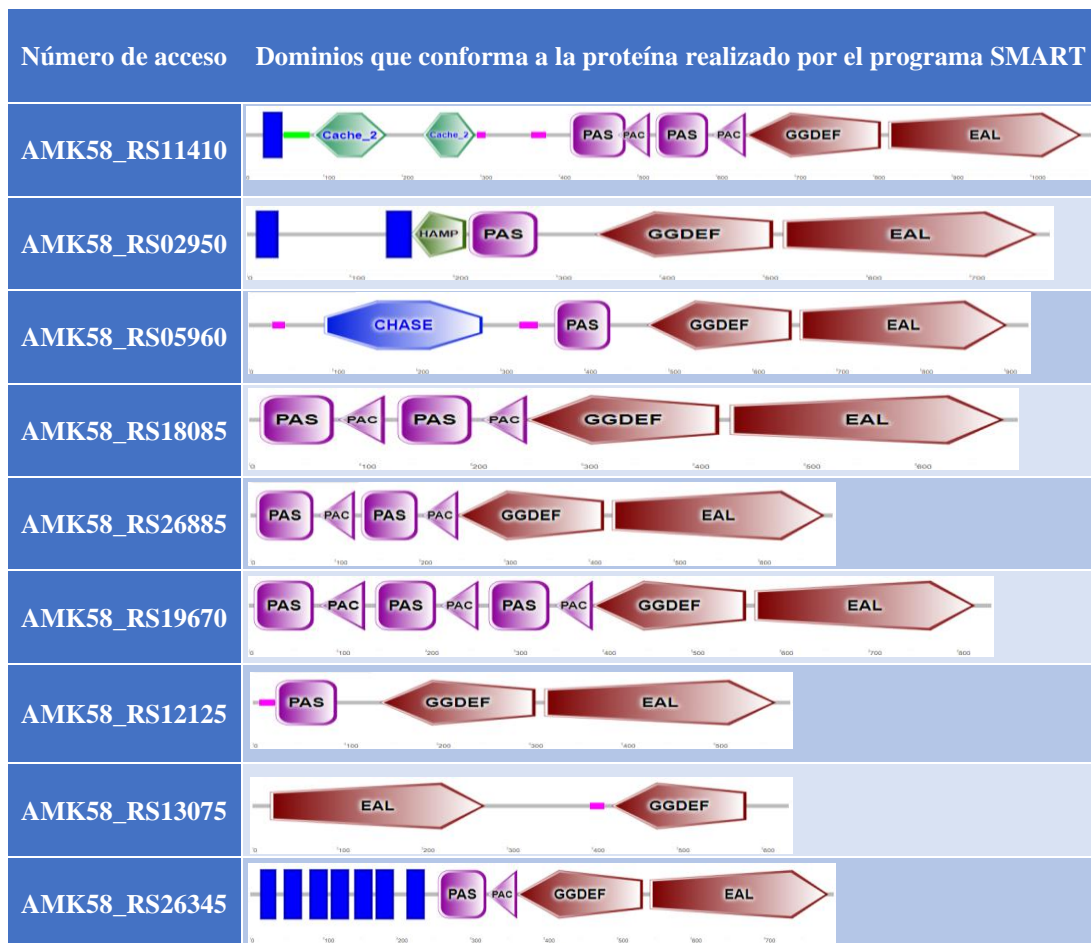


**Figura 31.- Predicción de interacciones proteína-proteína entre la proteína CdgA y las DGC con dominios transmembranales en *A. brasilense* Sp7 utilizando el programa STRING.** Análisis de interacción proteína-proteína ingresando las secuencias aminoacídicas de diguanilato ciclasas con dominio transmembranales versus la secuencia aminoacídica de la proteína CdgA de *A. brasilense* Sp7. La arquitectura de los dominios de las proteínas se realizó utilizando el programa bioinformático SMART.

#### 9.5.4.- Predicción de interacción proteína-proteína retando el contexto genético para las proteínas de dominio híbrido GGDEF-EAL de *A. brasilense* Sp7

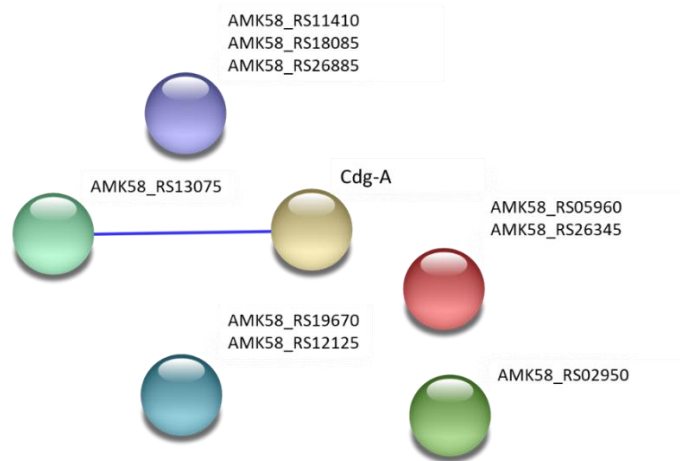
Por los antecedentes de interacción entre GcbC y LapD de *P. fluorescens* (DGC y LapD proteína híbrida, respectivamente), se prosiguió a realizar un análisis bioinformático de interacción entre las proteínas híbridas (con dominios GGDEF-EAL) y la DGC CdgA.

Están presentes en el genoma de *A. brasilense* Sp7 9 genes que codifican de manera presuntiva para proteínas híbridas GGDEF-EAL (Figura 33) Ramírez-Mata, *et al.*, (2018). Estas proteínas presentan diversos dominios accesorios en su amino terminal (dominios PAS, REC, transmembranales, CHASE, CACHE y HAMP).



**Figura 32.-** Tabla de las proteínas híbridas de *A. brasilense* Sp7 y los dominios que conforma a estas proteínas. El genoma de *A. brasilense* Sp7 presenta 9 proteínas híbridas las cuales son: AMK58\_RS11410 presenta dominios transmembranales, CACHE, PAS y GGDEF-EAL; AMK58\_RS02950 presenta dominios transmembranales, HAMP, PAS y GGDEF-EAL; AMK58\_RS05960 presenta dominios CHASE, PAS y GGDEF-EAL; AMK58\_RS18085 presenta dominios PAS y GGDEF-EAL; AMK58\_RS26885 presenta dominios PAS y GGDEF-EAL; AMK58\_RS19670 presenta dominios PAS y GGDEF-EAL; AMK58\_RS12125 presenta dominio PAS y GGDEF-EAL; AMK58\_RS13075 no presenta dominios accesorios, solo los dominios GGDEF-EAL; y AMK58\_RS26345 presenta dominios transmembranales, PAS y GGDEF-EAL. La arquitectura de los dominios de las proteínas se realizó utilizando al programa bioinformático SMART.

Al analizar este grupo de proteínas en el programa de STRING, se presentó la siguiente información (Figura 33). Como se puede observar, solo AMK58\_RS13075 (proteína sin dominios accesorios) presenta la evidencia de concurrencia (línea azul) con CdgA y el puntaje otorgado por el programa es de 0.726.



**Figura 33.- Predicción de interacciones proteína-proteína entre la proteína CdgA y las nueve proteínas híbridas GGDEF-EAL presentes en *A. brasilense* Sp7 utilizando el programa STRING.** Análisis de interacción proteína-proteína ingresando las secuencias aminoacídicas de proteínas híbridas versus la secuencia aminoacídica de las proteínas CdgA de *A. brasilense* Sp7.

En base a los datos presentados por Dahlstrom, *et al.*, (2015), se prosiguió a analizar *in silico* los dominios EAL de las proteínas híbridas y compararlas con el dominio EAL de LapD de *P. fluorescens*.

AMK58_RS13075	RKAALPAQWTDLL---AGLSVAFQPIVQMTGR---CHGMEALRGLERS-PFASPNDLI	53
EAL LapD	DHHAWHRLLDQALNQ-RRFELFFQFVVAQDTQ--LVLHYKVL SRLLDEQGQTIPAGRFL	57
AMK58_RS18085	RRMFLTMHLAHAIER-NELHMAYQPIVEARSGR---IVGAEALLRWNQHTLGAVSPGEFI	56
AMK58_RS11410	ERMWLESNLRHALAD-NELELHFQPFQFAVDGRR---LVGAEGLLRWRRRDGMTIPPQQFI	56
AMK58_RS05960	ERLDLEAALRRALQN-RELFVYQFQLRISDDL---VVGVEALLRWRHPEAGLIMPNRFI	56
AMK58_RS26345	ARRGLERDLRQALVD-GALNVHYQPLEDCRTHR---ILGFEALVRWKHPERGMIPPEFV	56
AMK58_RS02950	RHLALDRDLRAALAR-GEFQLHYQPKVSLVDQS---LEGFEALLRWDKPGFGMIPPEGFI	56
AMK58_RS26885	FTLGLESDLRRAAATGEEFELHFQPKVDMPHRGYHRVIGAEALIRWRHPVRGLVSPAEFI	60
AMK58_RS19670	AQMRTETDLRMALEQ-NQLVLYYQFIVALSSGQ---IAGFEALMRWKHPERGLVPPGEFI	56
AMK58_RS12125	EQVRLEHDLRRALERGEEELWVAYQPIVEMVTGG---LAGFEALVRWNHPERGNIPPEVFI	57
	: : * * : * * : :	
AMK58_RS13075	DAA MQGLLAEVECALHAKAVAAFRALDHPNAKL-----FLNIDARVVGVAGSPWL	105
EAL LapD	PWL RFGWTARLDRLMLER---VLEQMAGHEE-----SLALNLSATLA-----	98
AMK58_RS18085	PVALESGLIVPITDWWLNA---VVEQMAVWEPAGTL---PNQIFINISGPFQL-----	103
AMK58_RS11410	PVAIDTGLIQPIGDWVLGE---ACRRAADLLA-RGH---PLLVAINLSAVQLR-----	102
AMK58_RS05960	PVALETGLIQAIGDWVLDE---VCAQIRWRDQIGL---TVPIAVNVSAQQFR-----	103
AMK58_RS26345	PLATE SGLVMQLGGEFVLR---ACRDARNWPDVAVK-----VAVNLSPIQFR-----	99
AMK58_RS02950	PVAIDTGFIVPLGDWVIDE---ACRQLREWIDRGLE---PVPVAVNISPRHLR-----	103
AMK58_RS26885	PLATE TGLIWEIGAWTLRE---ACGRLAGWLAAGLD---IASVSVNLSPRQFQ-----	107
AMK58_RS19670	PLATE SGLIVPMGRWALRE---AARQLGRWQT-SFPRPTPLFMSVNVSSRQFS-----	105
AMK58_RS12125	PIATS TGLIVTLGAWVLRE---ACRQIAEWQDKRLPGSANLFMSINLSTRQLN-----	107
	* : . : * : .	
AMK58_RS13075	APFAPAPFAPAPFALDGGAPDPRLTISLEISERRELRPDHSIEQAIDRYRQAGIGVALDD	165
EAL LapD	----DPQALNKVFEILRAHSNLGARLTLEIGEEQLPEQA-VLEQLTRRLRELGFSLSLQR	153
AMK58_RS18085	----RGNLSARLEELLGRHPELRGLGLEITEQAAVRDLKAAVQTLTELEAIDVQVAIDD	159
AMK58_RS11410	----RPGLADRVAFWLDAYKVPARCLELLETESVLMDESDMVTNTFARLREMGVALAIID	158
AMK58_RS05960	----DGQLPAKVATALDRNGLQGWLEIEVTEGTLIDDVPSAIATLRALKQRGCLIALDD	159
AMK58_RS26345	----HSDLSATILGILEEGLPPHRLEIEVTEGVMIDDERALAALTALKAAGVRIALDD	155
AMK58_RS02950	----QRS AED-FRRIIDRHRLSPDLVELEITEGAVMQMDHALSVLAALKAMGIRVAVDD	158
AMK58_RS26885	----DARLVNFVRDVVDRSGVPPERLELELETEGAMIGDIEKAVTILHGLKIGIRLSIDD	163
AMK58_RS19670	----DDD LIGLVSDVLLSEGVPPSSLKLEITESLLMKDPAKCRVLMQGIRDMDVRLSIDD	161
AMK58_RS12125	----DPDMVSLVREVLRGTGAEP SWIKLEITESAVMEKAEQSI RLLQTLRGLGKMSIDD	163
	: : * : * . . : : :	

AMK58_RS13075	FGVGFAGLRLLYESKPDYLKIDRYFVAGIDQDTRKRAIAHALVGYAHAHGILIIAEGVET	225
EAL_LapD	FGGRFSGMIGNLARLGLAYLKIDGSYIRAIQDESCKRLFIEAIQRAAHSIDLPLIAERVET	213
AMK58_RS18085	FGSGYSSLSYVQQLPVSKLIDRAFMAIPDNLKNGALVRAAVGMAHGLGLVTVAEVET	219
AMK58_RS11410	FGTGYSSLSYLKRFVVDKLIKIDRSFVTSITDGS DGGAI AEAI IGMARSLRMQTLAEVET	218
AMK58_RS05960	FGTGYSSLNYLHRFPIDKLIKIDRSFIHDLQDQAGDSIPRAIVGLGRSLGLSVVAEVET	219
AMK58_RS26345	FGTGYSSLSYLQKFTFDKLIKIDRSFVQLMAENRESLSIIRTIILAKSLDIAVTAEVET	215
AMK58_RS02950	FGTGHSSLSYLKRLPVTTLKIDRSFVNGVPSEREDNGIVSTIIAMADMLGLDVVAEVET	218
AMK58_RS26885	FGTGYSSLSYLKRFPIINTLKIDRSFVRDIVQSATDPAIVNTIVNLADSLGFDTLAEVET	223
AMK58_RS19670	FGTGYSSLSYLHKFPADTLKIDRSFVQAISSGEGNAIVQVIATLAAIILGMEVAEVET	221
AMK58_RS12125	FGTGYSSLSYLHKLPIDSLKVDRSFVMAHQSEENRAIVRIIMDLARLLGFDVIAEGVET	223
	** . : : . ** : : . : . : : ** : *	
AMK58_RS13075	EKEFHTCRDLGCDFAGQGYLIARPSLDLR	253
EAL_Lap	EGELSVIREMGLYGVQQGLFGEKPK---	238
AMK58_RS18085	EEQRDFLVSVCDDLQGYLFGRPAAPDA	247
AMK58_RS11410	EEQFGGLARLGCDQIQGYLLGRPMFYDD	246
AMK58_RS05960	EAQLQLLRSLKCESFQGYLFSRPVLAEE	247
AMK58_RS26345	EEQRTLKKNESCQLQGYLIGRPVPASE	243
AMK58_RS02950	QEQAFLRHHNCTLVQGWLTGRVPVAGQ	246
AMK58_RS26885	EEQADMLRRQRCTRIQGFLLSRPLDVDA	251
AMK58_RS19670	EMESEFLRDIMCKYAQGYLYARPAPADA	249
AMK58_RS12125	SADANLLRALACDYQGYHFARMPAAD	251
	. : . ** . *	

**Figura 34.- Alineamiento de las secuencias de los dominios EAL de las proteínas híbridas de *A. brasilense* Sp7 y el dominio EAL de LapD de *P. fluorescens*.** Alineamiento de las secuencias de los dominios EAL de las proteínas híbridas de *A. brasilense* Sp7 y el dominio EAL de LapD de *P. fluorescens*. En rosa se encuentra señalado el motivo EAL, en verde se presenta la secuencia de interacción identificada en el dominio EAL de LapD donde se destacan los aminoácidos de triptófano y ácido glutámico más cercano al carboxilo terminal del dominio. En azul se presentan los aminoácidos que conforman al loop 6, secuencia dada por los siguientes aminoácidos ‘DGF(A/S/T)(G/A)(Y/F)(S/A/T)(S/A/GV/T)’.

En la Figura 33 se presenta el alineamiento de las secuencias proteicas de los dominios EAL de las proteínas híbridas de *A. brasilense* Sp7 y el dominio EAL de LapD de *P. fluorescens*. Como se puede observar el motivo EAL se encuentra altamente conservado en los dominios EAL de las proteínas híbridas a excepción de AMK58\_RS11410 donde se cambia A por G.

#### 9.5.5.- Predicción de interacción proteína-proteína de CdG con PDEs de dominio EAL de *A. brasilense* Sp7

El programa STRING obtuvo como resultado la posible interacción de CdG con 3 PDEs de dominio EAL en búsqueda de proteínas que puedan interaccionar con la proteína CdG de *A. brasilense* Sp7, al analizar el genoma de *A. brasilense* Sp 245 (Figura 23).

Por lo tanto, que se decidió realizar un análisis bioinformático de interacción proteína-proteína entre CdgA y las cinco PDEs de dominio EAL que presenta *A. brasilense* Sp7. Las 5 PDEs son: AMK58\_RS05685, AMK58\_RS05115, AMK58\_RS03020, AMK58\_RS26825/CAJ18244 (estas últimas agrupadas debido a que sólo existe la diferencia de 20 aminoácidos entre los dominios que presenta) (Ramírez-Mata, *et al.*, 2018).

Como podemos observar en la Figura 35-A, estas proteínas cuentan con dominios accesorios en su amino terminal (dominios PAS y transmembranales). De acuerdo con el programa STRING se presentan evidencias para la interacción de CdgA con AMK58\_RS05685, AMK58\_RS05115/CAJ18244, y AMK58\_RS03020.

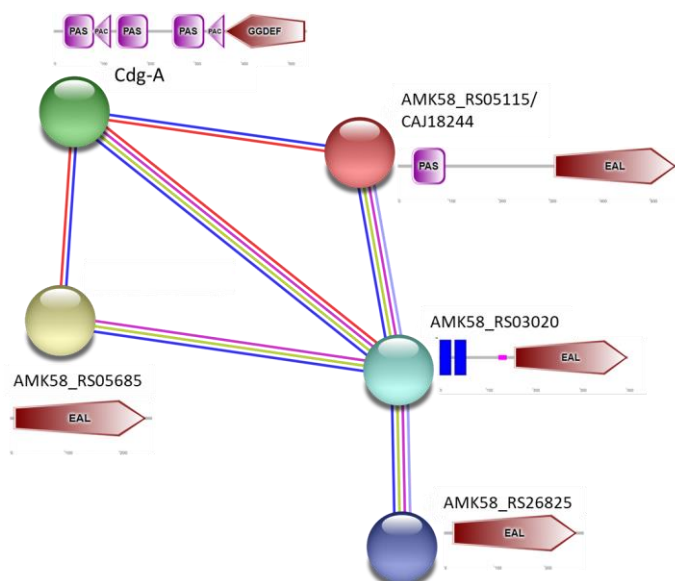


Tabla 9.- Puntaje de predicción de interacción dado entre DGC con dominios accesorios PAS y CdgA de <i>A. brasilense</i> Sp7			
Proteína 1	Puntaje de Interacción	Proteína 2	Tipo de Evidencia
AMK58_RS08100 (CdgA)	0.778	AMK58_RS05115 CAJ18244	<ul style="list-style-type: none"> <li>• Presencia de Fusión</li> <li>• Concurrencia</li> </ul>
AMK58_RS08100 (CdgA)	0.803	AMK58_RS03020	<ul style="list-style-type: none"> <li>• Presencia de Fusión</li> <li>• Concurrencia</li> <li>• Experimental</li> <li>• Búsqueda de datos</li> </ul>
AMK58_RS08100 (CdgA)	0.756	AMK58_RS05685	<ul style="list-style-type: none"> <li>• Presencia de Fusión</li> <li>• Concurrencia</li> </ul>
AMK58_RS05115/ CAJ18244	0.415	AMK58_RS03020	<ul style="list-style-type: none"> <li>• Concurrencia</li> <li>• Experimental</li> <li>• Búsqueda de datos</li> <li>• Banco de datos</li> </ul>
AMK58_RS05685	0.802	AMK58_RS03020	<ul style="list-style-type: none"> <li>• Concurrencia</li> </ul>

			<ul style="list-style-type: none"> <li>• Experimental</li> <li>• Búsqueda de datos</li> </ul>
AMK58_RS03020	0.415	AMK58_RS26825	<ul style="list-style-type: none"> <li>• Concurrencia</li> <li>• Experimental</li> <li>• Búsqueda de datos</li> <li>• Banco de datos</li> </ul>

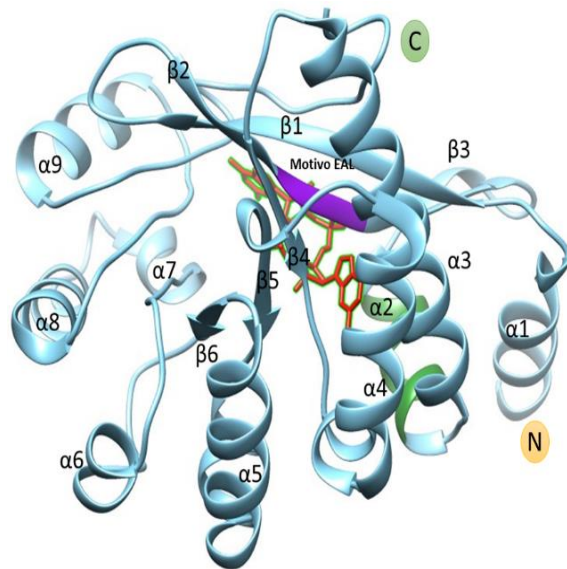
**Figura 35.- Predicción de interacciones proteína-proteína entre la proteína CdGA y las cinco fosfodiesterasas presentes en *A. brasilense* Sp7 utilizando el programa STRING.** A) Análisis de interacción proteína-proteína ingresando la secuencia aminoacídica de las fosfodiesterasas con dominio EAL versus la secuencia aminoacídica de las proteínas CdGA de *A. brasilense* Sp7. Tabla 9. Score de interacciones proteína-proteína otorgado por el programa STRING. La arquitectura de los dominios de las proteínas fue realizada utilizando el programa bioinformático SMART.

En la Tabla 9 se presenta el puntaje de interacción entre las proteínas y el tipo de evidencia que se da por parte del programa STRING. Como se puede observar, el puntaje que otorga el programa es de 0.756-0.803 lo cual indica una confiabilidad alta para una posible interacción entre estas proteínas.

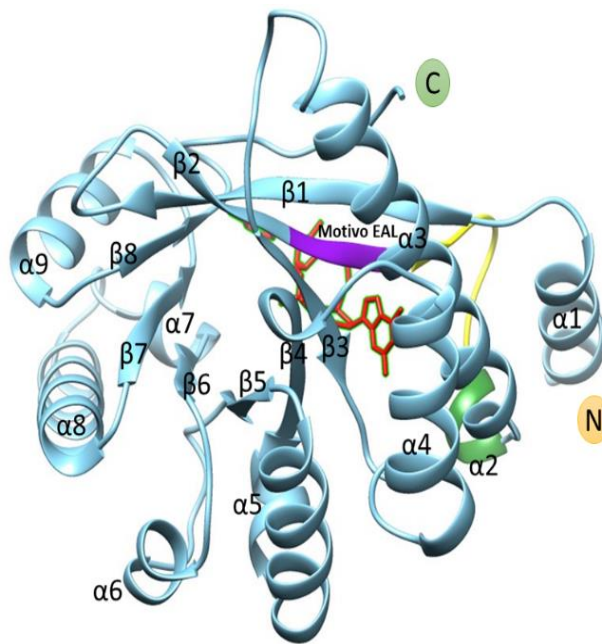
En base a los resultados previamente obtenidos, se decidió analizar la estructura secundaria de las cuatro PDEs (ya que aunque no todas presenten más de tres evidencias de interacción con la DGC CdGA, presentan interacciones directas) y compararlas con el dominio EAL de LapD de *P. fluorescens* con el objetivo de identificar homología en secuencia y estructura por medio de los programas I-TASSER y Clustal Omega.

Para el análisis de estructuras terciarias de las PDEs con dominio EAL de *A. brasilense* Sp7 se estudiaron las secuencias aminoacídicas de las proteínas AMK58\_RS03020, AMK58\_RS05685, AMK58\_RS05115 y CAJ18244, con el servidor I-TASSER.

En la Figuras 36 se presenta la estructura del dominio EAL de la PDE AMK58\_RS03020 con un C-Score de 1.48. En la Figura 37 se presenta la estructura del dominio EAL de la PDE AMK58\_RS05685 con un C-Score de 1.15.



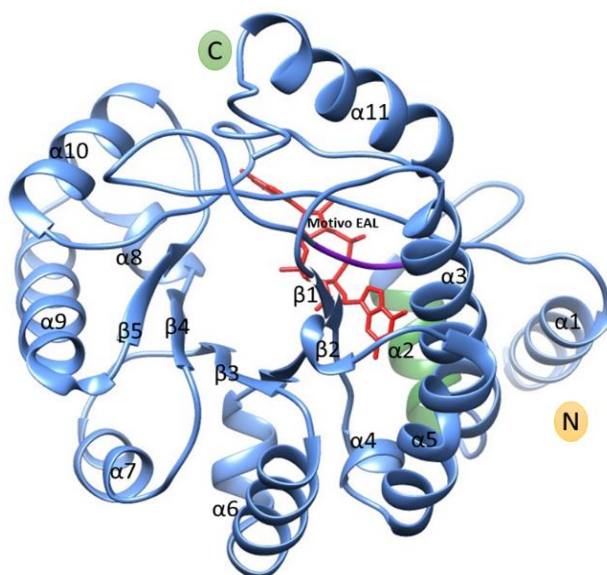
**Figura 36.- Modelo estructural por homología del dominio EAL de la PDE AMK58\_RS03020 de *A. brasilense* Sp7.** En azul cielo se presenta el modelo del dominio EAL de la PDE AMK\_58\_RS03020 y en morado se presenta el motivo EAL de *A. brasilense* Sp7. La estructura que arroja el I-TASSER cuenta con un C-score de 1.48, lo que es confiable en la escala de -5 a 2, con una TM de 0.92+/-0.06 y RMSD de 2.8+/-2.1Å.



**Figura 37.- Modelo de estructura secundaria del dominio EAL de la PDE AMK58\_RS05685 de *A. brasilense* Sp7.** B) En azul cielo se presenta la estructura secundaria del dominio EAL de la PDE AMK\_58\_RS05685 y en morado se presenta el motivo EAL de *A. brasilense* Sp7. La estructura que arroja el I-TASSER cuenta con un C-score de 1.15, lo que es confiable en la escala de -5 a 2, con una TM de 0.78+/-0.10 y RMSD de 4.9+/-3.2Å.

Debido a que 2 de las 4 proteína a analizar presentan dominios parecidos (un dominio PAS y un dominio EAL), se realizó un alineamiento utilizando Clustal Omega para determinar

las diferencias del dominio EAL (Figura 38-A). El alineamiento muestra una homología del 100% en cuanto a la comparación de las secuencias de los dominios EAL de AMK58\_RS05115 y CAJ18244 (Anexo 2) por lo cual, sólo se estudió una secuencia con I-TASSER para generar la estructura secundaria (Figura 38-B) con un C-Score otorgado es de 0.48. En la Figura 38-B se presenta la cantidad de  $\alpha$ -hélices (color morado) y hojas  $\beta$ -plegadas (color amarillo), así como su amino y carboxilo terminal. En la Figura 38-C se presenta el motivo EAL en color morado interaccionando con di-GMPc.



**Figura 38.- Modelo estructural por homología del dominio EAL de las PDE AMK58\_RS05115/CAJ18244 de *A. brasilense* Sp7.** En azul cielo se presenta la estructura secundaria del dominio EAL de la PDE AMK58\_RS05115/CAJ18244 y en morado se presenta el motivo EAL de *A. brasilense* Sp7. La estructura que arroja el I-TASSER cuenta con un C-score de 0.48, lo que es confiable en la escala de -5 a 2, con una TM de 0.87 $\pm$ 0.07 y RMSD de 3.4 $\pm$ 2.4Å

Los principales modelos de los cuales I-TASSER se basó para poder generar las estructuras secundarias de AMK58\_RS03020, AMK58\_RS05685 y AMK58\_RS05115/CAJ18244 fueron MucR y PA0575 de *P. aeruginosa* (número de acceso de PDB: 5m1t y 5m3c respectivamente) con la resolución del cristal de 2.27 Å para MucR y 2.8 Å PA0575. AMK58\_RS03020 y AMK58\_RS05685 provinieron de MucR mientras AMK58\_RS05115/CAJ18244 de PA0575.

Posteriormente a obtener las estructuras terciarias de las PDEs de *A. brasilense* Sp7, se prosigió a realizar un alineamiento de las secuencias para determinar la conservación del

motivo EAL el loop 6 y la secuencia de aminoácidos que conforman la segunda  $\alpha$ -hélice que permite la interacción con otras proteínas (delineado en rosa, azul cielo y verde respectivamente en Figura 39).

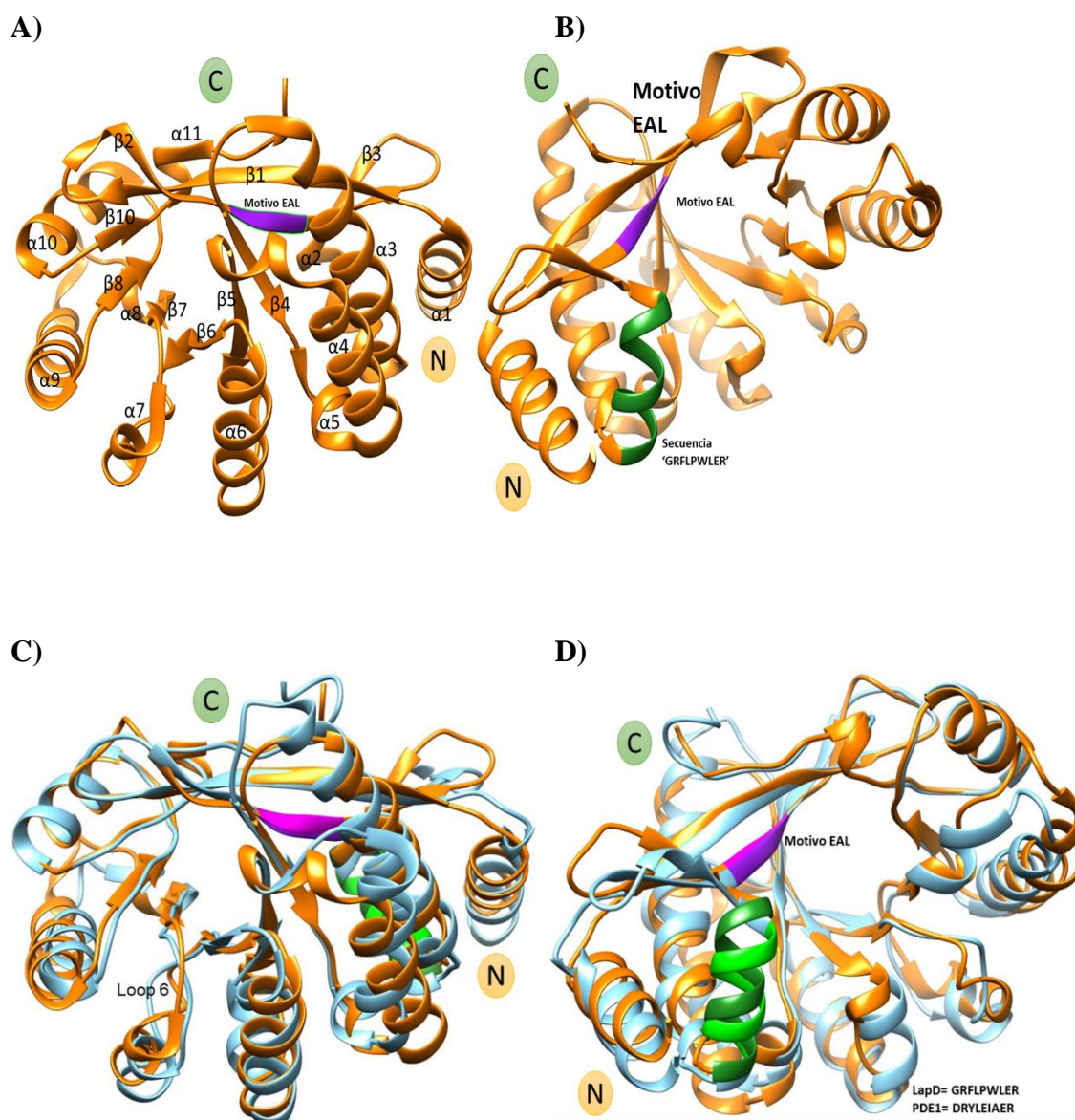
El motivo EAL de LapD es parcialmente degenerado ya que E cambia por K y A por V. En AMK58\_RS03020 se presentan cambios parciales en los aminoácidos del motivo EAL, se cambia A por una V y L por F. Mientras que las secuencias de AMK58\_RS05685 y AMK58\_RS05115/CAJ18244 mantienen al motivo EAL conservado.

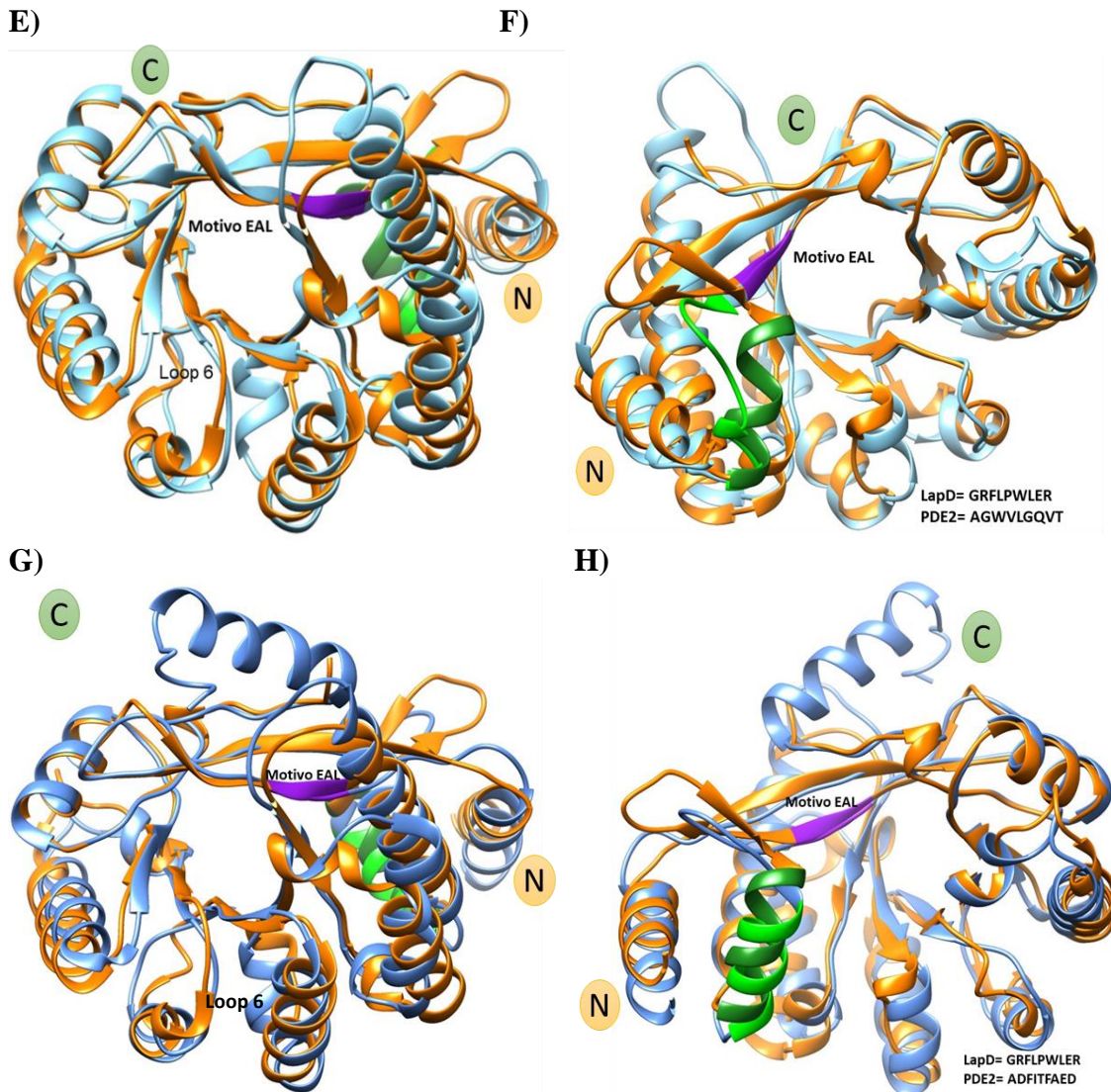
Para la secuencia de interacción de la segunda  $\alpha$ -hélice ‘GRFLPWLER’ destacando a los aminoácidos W y E, donde AMK58\_RS05115/CAJ18244 incluye F, AMK58\_RS03020 contiene I, mientras que, AMK58\_RS05685 incluye G en lugar de W respectivamente. Por otro lado, AMK58\_RS05685 presenta V en lugar de E y AMK58\_RS05115/CAJ18244 y AMK58\_RS03020 mantienen conservado a E.

AMK58_RS05685	-KGGGCGGQCTEGLGFDFMTMAFQPIVDVGGGQPW-AH	EAT	VRGLDGGQ----	A-GWVLGQ	53
AMK58_RS05115/CAJ18244	DAVDRIQRRLRATIEKRDFVYQPIVDIGTRTVH-HL	EAT	TRVEGTLS----	PADFTVA	55
EAL LapD	DHHAWHTLLDQALNHQRFELYQFVVAQDTQVVLHY	KVI	SRLLEHGQTI	PAGRFLPVI	60
AMK58_RS03020	DDAAVLEAVRDALAADRIDIHLPQIVSLPQR-KHRFF	VVF	SRVRAADGSLILP	DRYLEIA	59
	: : *:*	:: *	.	::	
AMK58_RS05685	EANRYAFDQACRVKAIELAASLGLGQQPD-TRLSINFLPNAVYQPEACIRATLAAAKR				112
AMK58_RS05115/CAJ18244	VGLIYDFDLLLTQSVLDT--MHAHRKEPTLPDVAINLSAKTLM-SPIFLRQFLAVTEP				112
EAL LapD	FGWTARLDRMLLEQVLRQ--MAG-----HEESLALNLSSATLA-DPQALNKIFEILRA				112
AMK58_RS03020	AGLMATIDNLLLVRCIQL--IRETERRQHAIGFFCNMSAATLS-DAAFMRQFLDLLAR				116
	. :*	:	. *:	::	::
AMK58_RS05685	TGFPPERIIFEVTENERVVDGAHLKSI FTEYKRQGFHTAIDDFGSGYSGLNLLAEFPDI				172
AMK58_RS05115/CAJ18244	YGELARKLLIEVTETVVVTDIAKLNEVLQKLRAGFRICLDDVAGAGS TSFQSLYGIQADF				172
EAL LapD	HSNLGSRLTLEIGEEQLPEQ-AVLEQLTRRLRELGFSLSLQRFGRFMSI GNLSRLGLAY				171
AMK58_RS03020	NQTLVPKLVFELSQQELRAGGAVTMGILSLSRIGFRFSMDRWSDDLIDVDALLRHEIRY				176
	:: :* : :	*	:	**	:: : . *
AMK58_RS05685	IKLDMELTRSIDTDRARRSIVSAILTVCRD LGITPVAEGIETPGEAKALRDLGITLMQGY				232
AMK58_RS05115/CAJ18244	AKIDGHFVRGAVHNGRDMAMLRSMVDVCRQLGLGLIGE QVEGPEHAQLLTELGVSLAQGY				232
EAL LapD	LKIDGSYIRAIDQESDKRLFIEATQRAAHSIDLPLIAERVETEGELAVIREMGLFGVQGG				231
AMK58_RS03020	LKLDCAALLD----SALAPRVEALRRRLDGTGIDLIAEKIETE AQLDAILRNGIDFGQGY				232
	*:*	::	::	::* :*	. : *:
AMK58_RS05685	LFARPAVETL-----				242
AMK58_RS05115/CAJ18244	LYSRPSRDFAYFAKDWTKMAGTGKVLKRG				262
EAL LapD	LVGE-----				235
AMK58_RS03020	LFGEPRPARKPP-----				244

**Figura 39.- Alineamiento de aminoácidos del dominio EAL de LapD de *P. fluorescens* con cuatro dominios EAL de PDEs de *A. brasilense* Sp7.** Alineamiento aminoacídico del dominio EAL de LapD de *P. fluorescens* y cuatro dominios EAL de cuatro fosfodiesterasas de *A. brasilense* sp7. En rosa se encuentra señalado el motivo EAL, en verde se presenta la secuencia de interacción identificada en el dominio EAL de LapD donde se destacan los aminoácidos de triptófano y ácido glutámico más cercano al carboxilo terminal del dominio. En azul cielo se resaltan los aminoácidos que conforman al loop 6, secuencia dada por los siguientes aminoácidos ‘DGF(A/S/T)(G/A)(Y/F)(S/A/T)(S/A/GV/T)’ y los cambios se presentan en azul oscuro, secuencia importante para la interacción proteína-proteína entre PDEs de dominio EAL.

En base a los resultados obtenidos por parte del alineamiento de los dominios EAL de las PDE de *A. brasilense* Sp7 y el dominio EAL de LapD de *P. fluorescens*, se prosiguió a realizar el empalme de estructuras secundarias por medio del programa bioinformático CHIMERA (Figura 40). En la Figura 40-A y 40-B se presenta la estructura en color dorado del dominio EAL de LapD obtenido del PDB (por sus signos en ingles Protein Data Bank), en morado se presenta el motivo EAL y en verde a la secuencia de interacción.





**Figura 40.- Alineamiento de la estructura secundaria del dominio EAL de LapD de *P. fluorescens* con el EAL de tres PDEs de *A. brasilense* Sp7 usando el programa CHIMERA. A) Estructura secundaria del dominio EAL de LapD de *P. fluorescens*. En color naranja se presenta la estructura secundaria del dominio EAL, en morado se presenta el motivo EAL. B) En color verde olivo se indican los aminoácidos pertenecientes a la secuencia de interacción ‘GRFLPWLER’ del dominio EAL de LapD de *P. fluorescens*. C) Alineamiento de los dominios EAL de LapD de *P. fluorescens* y PDE1 (AMK58\_RS03020) de *A. brasilense* sp7. D) Alineamiento de las secuencias de interacción ‘DRYLEIAER’ de PDE1 (AMK58\_RS03020) de *A. brasilense* sp7 con ‘GRFLPWLER’ del dominio EAL de LapD de *P. fluorescens*. E) Alineamiento de los dominios EAL de LapD de *P. fluorescens* y PDE2 (AMK58\_RS05685) de *A. brasilense* sp7. F) Alineamiento de las secuencias de interacción ‘AGWVLGQVT’ de PDE2 (AMK58\_RS05685) de *A. brasilense* sp7 con ‘GRFLPWLER’ del dominio EAL de LapD de *P. fluorescens*. G) Alineamiento de los dominios EAL de LapD de *P. fluorescens* y PDE3 (AMK58\_RS05115 y CAJ18244) de *A. brasilense* sp7. H) Alineamiento de las secuencias de interacción ‘ADFITFAED’ de PDE3 (AMK58\_RS05115 y CAJ18244) de *A. brasilense* sp7 con ‘GRFLPWLER’ del dominio EAL de LapD de *P. fluorescens*.**

En la Figura 40-C/D se presenta el empalme del dominio EAL de AMK58\_RS03020 de *A. brasilense* Sp7 y el dominio EAL de LapD de *P. fluorescens*. Como se puede observar, el motivo EAL se encuentra conservado, y las secuencias para el loop 6 y la secuencia de

interacción con otra proteína se encuentran parcialmente modificadas por lo cual ambas estructuras no coinciden en cierto porcentaje de acuerdo con el alineamiento de la Figura 40.

Por otra parte, en la sección E y F de la misma figura, se presenta el empalme del dominio AMK58\_RS05685 y el dominio EAL de LapD de *P. fluorescens*. Se observa la conservación del motivo EAL y el loop 6, sin embargo, se presenta gráficamente las alteraciones en la secuencia que permite la interacción con otras proteínas como se había predicho en el alineamiento de la Figura 40.

Finalmente, en la Figura 40-G/H se presenta el motivo EAL conservado, aunque en lugar de presentarse en forma de hoja- $\beta$ -plegada se presenta en forma de aza. Las secuencias del loop 6 y la secuencia de interacción con otras proteínas se encuentran parcialmente conservadas de acuerdo con el alineamiento de la Figura 39.

## 10.- Discusión

El di-GMPc es una molécula señal importante para el desarrollo de las bacterias ante diversos ambientes. Recientemente, el estudio del dinucleótido no sólo abarca a las enzimas encargadas de producir y degradar al segundo mensajero, si no también integra a aquellas proteínas que efectúan función de proteínas efectoras al ligar al dinucleótido como molécula señal. Gran parte de las proteínas que participan en esta red cuentan con dominios accesorios que permite captar señales ambientales que regulen su actividad y dependiendo de estos dominios su ubicación en el citoplasma o membrana celular será determinado. Se conoce que las enzimas las DGC y PDE presentan dominios característicos que le otorgan participación en la red de señalización del di-GMPc, y de misma forma las proteínas efectoras deben contar con dominios receptores del di-GMPc. Algunos ejemplos son: el dominio PilZ el cual contiene los motivos RxxxR y DxSxxG comprometidos en la unión al dinucleótido y caracterizado en enterobacterias y en otras proteobacterias como: *C. crescentus*, *P. aeruginosa*, y *Vibrio cholerae*, con dominio GGDEF degenerado y con sitio inhibitorio conservado (i.e. PelD de *P. aeruginosa*), proteínas híbridas con dominios parcial o totalmente degenerados (LapD de *P. fluorescens*), factores de transcripción (tipo FleQ de *P. aeruginosa* o Bcam1349 de *Burkholderia cenocepacia*) o con motivo W[F/L/M][T/S]R (VpsT de *V. cholerae*). Sin embargo, no todas las macromoléculas efectoras que reciben a este dinucleótido son proteínas también pueden presentarse riboswitches los cuales ejercen el control de expresión de ciertos genes. Los riboswitches cuentan con el motivo GEMM que les permite interactuar con algunos marcos de lectura abiertos de DGCs, PDEs y genes regulados por el di-GMPc para su regulación transcripcional o transduccional. (Smith & Strobel., 2011; Sondermann, *et al.*, 2012; Abel S., *et al.*, 2013; Dahlstrom, *et al.*, 2017). A pesar de contar con cierta información respecto a esta red de señales, mucho queda por estudiar y descubrir de estas vías.

Un modelo bacteriano ampliamente estudiado ha sido *C. crescentus*, el cual ha demostrado que alteraciones a la concentración intracelular del di-GMPc (nulas o sobreexpresión) afectan la división y morfología celular (tanto en células individuales como en formación de colonias), la expresión de EPSs y la formación de flagelos, pilis tipo IV y adhesinas. Así

mismo, se ha demostrado que tanto DGC como PDE participan de diversas formas en diferentes etapas de crecimiento y desarrollo bacteriano de otras proteínas.

A pesar de los recientes avances sobre el mecanismo y función de la red de señalización del di-GMPc existe el paradigma de la complejidad que representa estudiar estas redes. En el Laboratorio de la Interacción Bacteria-Planta se continuó el estudio de la DGC CdgA de *A. brasilense* Sp7 Ramírez-Mata, *et al.*, (2016). Con el objetivo de profundizar el conocimiento que se tiene de ésta DGC y facilitar el posible estudio de la red de señalización en la que participa CdgA, surgió la necesidad de diseñar estrategias de estudio que permitan abarcar otras funciones que pueda presentar la proteína. Al obtener los resultados positivos (Figuras 12 y 13) y en donde se observa unión de proteínas con la DGC, se prosiguió a realizar estudios bioinformáticos tanto de estructuras secundarias como de su posible interacción con otras proteínas como se describió en los estudios de Dalhstrom, *et al.*, 2015; Dalhstrom, *et al.*, 2016.

Si bien diversos autores emplean el sistemas de doble híbrido (ya sea en levadura o *E. coli*) para realizar interacciones de proteína-proteína *in vivo* como lo presenta Sarenko, *et al.*, (2017) en su estudio con *E. coli* K-12, algunos otros ocupan la metodología de ensayos Pull Down que permite detectar interacciones proteína-proteína o sustrato-proteína *in vitro* como Lindenberg, *et al.*, (2013) en el estudio de la DGC Ydam y PDE YciR de *E. coli*, así como Düvel, *et al.*, (2012) que utilizó al di-GMPc para identificar proteínas que unen al sustrato en *P. aeruginosa*. En este estudio empleamos éste último ensayo; sin embargo, para que la proteína “anzuelo” pueda ser utilizada debe contener alguna etiqueta que le permita adherirse a la superficie de la resina que se emplea en la purificación aspecto que fue considerado en este estudio (para eliminar inespecificidades) sin perder la interacción de la proteína en el proceso.

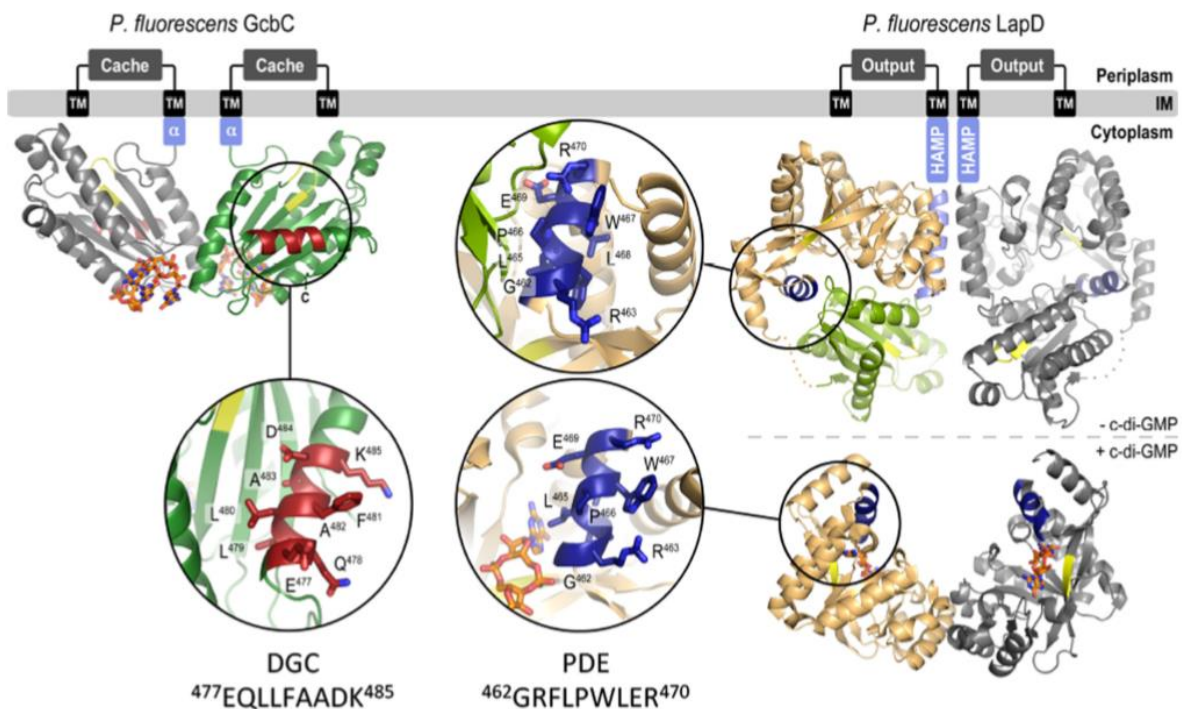
En las Figuras 16 y 17 observamos que la proteína CdgA está presente en ambos estadios de vida es decir células platónicas como sésiles (formación de biopelícula), lo que sugiere un comportamiento no canónico; ya que los estudios efectuados en varios modelos bacterianos indican que la concentración del di-GMPc es baja en células móviles, por tanto, sería de esperar que CdgA no estuviera presente en extractos libres de células en vida libre en condiciones de formación de biopelícula. Es posible que CdgA no esté activa, o bien que

inhiba o active a través de la interacción proteína-proteína efectores que participen en la cascada de señalización como fue descrito en los modelos de *P. fluorescens*. Es tentativo especular esta interacción ya que estudios de nuestro grupo de trabajo indican la participación de un gen río abajo del gen *cdgA*, particularmente el gen AMK58\_RS08090 que codifica para la presunta MCP, el cual está implicado en quimiotaxis; y como se predice por el estudio bioinformático (Figura 22), los genes AMK58\_RS08095 y AMK58\_RS08090 tendrían una posible interacción con CdgA (Dahlstrom, *et al.*, 2015; Rodríguez-Vázquez, 2016; Jenal, *et al.*, 2017). A través de qué mecanismo esta interacción física funcionaría es un aspecto que debe ser analizado de manera experimental, en futuros trabajos.

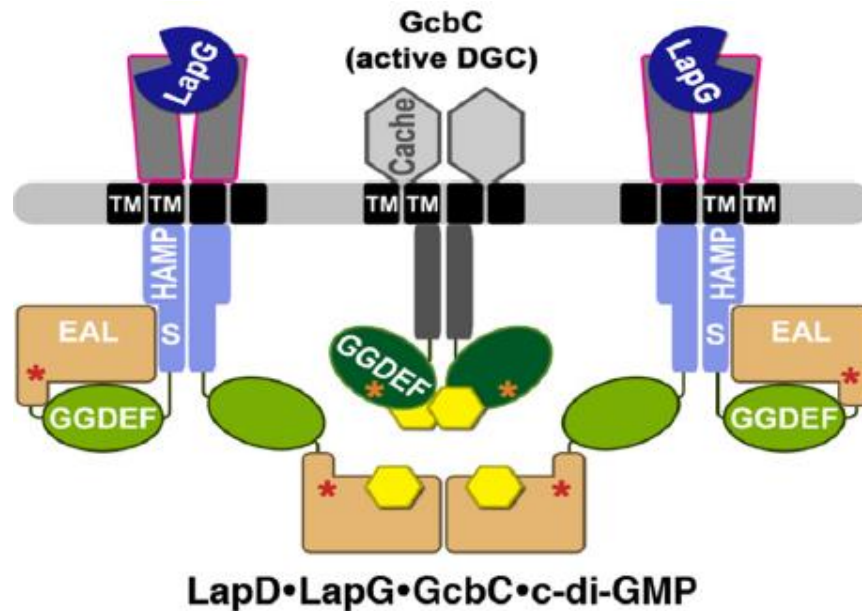
En relación con los datos obtenidos por medio de la interacción de CdgAr con extractos libres de células obtenidas en cultivos en ambas condiciones, se observó la presencia de bandas con pesos moleculares de entre 37 a 50 kDa de donde se destaca una banda en especial (~37kDa) a partir de los cultivos creciendo en condiciones plantónicas; sin embargo, ya que las demás también son parcialmente visibles en el carril control (carril 6 vs carril 4 respectivamente) de las Figuras 12-A y 12-B éstas deben ser consideradas como artefactos. Los datos obtenidos con extractos proteicos celulares en formación de biopelícula a un día de crecimiento, se observó la presencia de varias proteínas con diferentes MMr (de 37-50 kDa, 50-75 kDa, 75-100 kDa, 100 kDa y 150-250 kDa) como se muestra en la Figura 13, estos datos fueron confirmados a partir de tres ensayos independientes realizados. Los datos apuntan a que CdgA está presente en ambos estilos de vida de *A. brasilense* Sp7 y que bajo formación de biopelícula CdgAr interacciona con un número mayor de proteínas; sin embargo, es necesario identificar a estas proteínas por espectrometría de masas. Por otro lado, la formación de anticuerpos policlonales anti-CdgAr permitió detectar la presencia de la proteína en ambos estilos de vida (plantónico y sésil) a cuatro días diferentes de crecimiento, confirmando su posible participación en ambos estilos de vida. Ha sido descrito por el grupo de G. Alexandre adaptaciones metabólicas durante la transición de la célula móvil (plantónica) a sésil (biopelícula) por la presencia del oxígeno (cultivo en condiciones de agitación), en donde se observa la formación de flóculos o agregados de bacterias; bajo estas condiciones la concentración celular del di-GMPc es similar a la de biopelícula por lo tanto, la señalización podría ser

similar en ambas condiciones a los tres días de crecimiento, bajo las condiciones que se emplearon en el presente estudio (Bible, *et al.*, 2015; O’Neal, *et al.*, 2017).

Ahora bien, Dahlstrom, *et al.*, (2015) publica la detección de interacción proteína-proteína entre la quinta alfa hélice de la DGC GcbC (secuencia ‘EQLLFAADK’) con la segunda alfa hélice de la proteína híbrida LapD (secuencia ‘GRFLPWLER’) de *P. fluorescens* y en el (2016) menciona la importancia del sitio I para que la interacción se realice (Figura 42). Y Giacalone, *et al.*, (2018), presentan que GcbC no puede producir la misma cantidad de di-GMPc que se produce al interactuar con LapD de *P. fluorescens* y que la interacción entre GcbC y LapD, permite que el di-GMPc producido por la DGC sea dirigido a LapD. Por otro lado, la presencia de citrato o isocitrato aumenta la producción del dinucleótido ya que los dominios accesorios CACHE pueden detectar a estos compuestos.<sup>7</sup>



**Figura 42.- Modelos de las proteínas transmembranales GcbC de y LapD de *P. fluorescens* y acercamiento a las secuencias importantes para la interacción de estas proteínas.** Cuando LapD no cuenta con di-GMPc se encuentra en un estado no activo, sin embargo, cuando LapD recibe di-GMPc la hélice se expone y posteriormente se realiza la interacción entre ambas proteínas (Tomado de Dahlstrom, *et al.*, 2015).



**Figura 43.- Modelo del complejo formado entre GcbC-LapD en *P. fluorescens*.** La proteína híbrida LapD unirá su dominio EAL degenerado al dominio GGDEF de GcbC a través del sitio inhibitorio y las alfa-hélices 2 y 5 (del dominio EAL de LapD y dominio GGDEF de GcbC respectivamente). La producción del di-GMPc aumentará con dicha interacción más aparte la presencia de citrato e isocitrato captado por los dominios CACHE de GcbC, y finalmente el dinucleótido será dirigido al dominio EAL de LapD que ayudará a un cambio estructural de esta proteína que permita captar y retener a la proteasa LapG, la cual esta encargada de regular la producción de la adhesina LapA. Este pasó da inicio a la producción de LapA y con ello el inicio de la formación de la matriz extracelular para la formación de biopelícula (Imagen tomada de Giacalone, *et al.*, 2018).

En base a la posibilidad de que la proteína CdgA esté participando en la red de señalización del di-GMPc de acuerdo con los ensayos Pull Down y cuente con la capacidad de interactuar con otras proteínas, se decidió analizar la estructura secundaria de la proteína en base a los datos de Dahlstrom, *et al.*, (2015 y 2016). En la Figura 18 se presenta el modelo del dominio GGDEF de CdgA por el servidor I-TASSER, con un C-score de 1.10 de un rango de -5 a 2 (siendo 2 la puntuación más alta), el servidor predijo un modelo de alta confiabilidad el cual se identifica a los motivos RXXD y GG(D/E)EF en color rosa y azul respectivamente, y al di-GMPc interactuando con el motivo RXXD. Ya que Dahlstrom, *et al.*, (2016) discuten sobre la importancia de los aminoácidos que conforman al sitio I para la interacción proteína-proteína entre GcbC y LapD, se realizó un alineamiento entre los dominios GGDEF de las proteínas CdgA de *A. brasilense*, GcbC de *P. fluorescens* y PleD de *C. crescentus*. En la Figuras 19 se presenta el alineamiento de dominio, mientras que en la Figura 20 se presenta el alineamiento del motivo RXXD y GG(D/E)EF conservado de acuerdo con Chan, *et al.*, (2004). Así mismo, los aminoácidos importantes para la quinta

alfa hélice como para el sitio I se presentan parcialmente conservados. Por lo que se prosigue a realizar un alineamiento a nivel estructural, y debido a que GcbC y PleD ya son proteínas cristalizadas, comparar los datos de Dahlstrom, *et al.*, (2015 y 2016) para determinar si los aminoácidos de la quinta alfa hélice, aunque estén parcialmente degenerados mantengan estructura similar a la presente en GcbC. Como se puede observar en la Figura 21 la quinta alfa hélice de CdgA permanece con una estructura similar a la de GcbC lo cual podría ser importante para que CdgA pueda llevar a cabo interacción con otras proteínas. Así mismo, se analiza la estructura secundaria de 3 PDE como se puede observar en las Figuras 36, 37 y 38 otorgados por I-TASSER para AMK58\_RS03020, AMK58\_RS05685 y AMK05115/CAJ18244 respectivamente. Como podemos observar, solo AMK58\_RS05685 no presenta la formación de la segunda alfa hélice mientras que las otras dos la presentan de forma parcial. Además, las 3 estructuras presentan la estructura base tipo barril conformada por 8 alfa hélices y 8 beta plegadas intercaladas, con una aza flexible (denominado loop 6) que permite la dimerización y control del ion metálico para la actividad enzimática (Chan, *et al.*, 2004; Dahlstrom, *et al.*, 2015; Dahlstrom, *et al.*, 2016; Römling, *et al.*, 2017; Giacalone, *et al.*, 2018).

En base a los resultados obtenidos de los ensayos Pull Down y a la presencia parcialmente conservada de la quinta alfa hélice, se prosiguió a predecir por medios bioinformáticos la posible interacción de CdgA con otras proteínas. Los grupos por analizar consistían en: proteínas de los genes que están en la misma unidad transcripcional que CdgA (AMK58\_RS08105, AMK58\_RS08095 y AMK58\_RS08090), DGC separadas por dominios accesorios, proteínas híbridas y PDE con dominios EAL de *A. brasilense* Sp7, así como la base de datos del mismo programa que contiene el genoma de *A. brasilense* Sp245. De los análisis realizados indicamos las proteínas que sobresalieron (con >3 evidencias indicadas por el programa STRING) en la Tabla 10.

Tabla 10. Proteínas de interés posterior al análisis realizado por el programa STRING.			
No. Acceso	Dominios de Proteína	Peso Molecular	Grupo de análisis
AMK58_RS06055	P-II	12.30 kDa	Base de datos STRING

AMK58_RS13700	P-II	12.27 kDa	Base de datos STRING
AMK58_RS08095	Dominio PAS	20.24 kDa	Unidad transcripcional
AMK58_RS12285	PAS-GAF-GGDEF	60.75 kDa	DGC con dominios PAS
AMK58_RS25025	GAF-GGDEF	36.33 kDa	DGC con dominios GAF
AMK58_RS13115	REC-GGDEF	56.05 kDa	DGC con dominios REC
AMK58_RS03020	Transmembranal-EAL	44.05 kDa	Fosfodiesterasa

De la Tabla 10, se presentan a las proteínas AMK58\_RS06055 (GlnB) y AMK58\_RS13700 (GlnZ) que son proteínas involucradas en el sistema de fijación o asimilación de nitrógeno en *A. brasilense*. Por otro lado, el dominio PAS es un dominio accesorio que capta señales ambientales o intracelulares para regular la actividad de proteínas para una gran cantidad de microorganismos, y finalmente a 3 DGC y 1 PDE con dominio EAL (Taylor & Zhulin, 1999; Inaba, *et al.*, 2015; Ramírez-Mata, *et al.*, 2018).

En base a los datos obtenidos por vías bioinformáticas, las bandas detectadas en el SDS-PAGE (Figura 12) del ensayo realizado con extractos proteicos de células platónicas podrían corresponder a AMK58\_RS25025 y AMK58\_RS03020 (una DGC y PDE respectivamente), mientras que en las proteínas detectadas en el SDS-PAGE (Figura 13) del ensayo realizado con extractos proteicos de células en formación de biopelícula podrían corresponder a AMK58\_RS12285, AMK58\_RS25025, AMK58\_RS13115, y AMK58\_RS03020. Sin embargo, estas especulaciones quedan por ser confirmadas por espectrometría de masas.

## 11.- Conclusión

En conclusión, de acuerdo con los datos obtenidos de los ensayos de Pull Down y el análisis bioinformático realizado, permite determinar que la estrategia diseñada para la obtención de información de interacción proteína-proteína es viable. Se predice que la proteína CdgA es capaz de interactuar con otras proteínas tanto en crecimiento plantónico como en la formación de biopelícula, sin embargo, se desconoce si las posibles proteínas presentadas en los resultados de los ensayos Pull Down son igual o semejantes a las arrojadas por el análisis bioinformático, por lo cual, esto último será determinado por análisis de espectrometría de masas posteriormente. Así mismo, la formación de anticuerpos policlonales permitió verificar la presencia de la proteína CdgA en ambos estilos de vida analizados, y aunque ya se cuenta con previos reportes de la participación de la proteína en formación de biopelícula (Ramírez-Mata, *et al.*, 2016), cabe identificar la participación de la proteína en estado plantónico.

## 12.- Perspectivas

- Secuenciar las bandas obtenidas de los ensayos Pull Down para determinar qué tipo de proteínas son obtenidas de los ensayos.
- Realizar ensayos de microscopia de fluorescencia etiquetando con un fluoroforo al anticuerpo Anti-CdgA para localizar a la proteína en *A. brasilense* Sp7 en ambos estilos de vida (células plantónicas y células sésiles).
- Realizar un análisis doble híbrido que permita confirmar los datos obtenidos en este estudio utilizando a las proteínas presentadas en la Tabla 10 más CdgA.

### 13.- Referencias

1. Abel S, Bucher T, Nicollier M, Hug I, Kaever V, Abel Zur Wiesch P, Jenal U. (2013). Bi-modal distribution of the second messenger c-di-GMP controls cell fate and asymmetry during the *Caulobacter* cell cycle. *PLoS GENET*. 9(9): e1003744. doi: 10.1371/journal.pgen.1003744
2. Arruebarrena D-P. A, Pereyña. C., Moreno Ramírez L., Xiqui Vázquez ML., Baca B.E., Pereyña M. A., Lamattina L., & Creus C.M. (2012). Denitrification-derived nitric oxide modulates (or enhances) biofilm formation in *Azospirillum brasilense*. *FEMS Microbiol. Lett.* 22. P. 77-85. doi: 10.1111/1574-6968.12030
3. Bashan, Y., & de-Bashan, L. (2010). How the plant growth-promoting bacterium azospirillum promotes plant growth - a critical assessment. *In Advances in Agronomy*. P. 77-136. doi:10.1016/s0065-2113(10)08002-8.
4. Bible AN, Khalsa-Moyers GK, Mukherjee T, Green CS, Mishra P, Purcell A, Aksenova A, Hurst GB, Alexandre G. (2015). Metabolic adaptations of *Azospirillum brasilense* to oxygen stress by cell-to-cell clumping and flocculation. *Appl Environ Microbiol* 81:8346–8357. doi: 10.1128/AEM.02782-15.
5. Branda, S., Vik, A., Friedman, L., & Kolter, R. (2005). Biofilms: the matrix revisited. *TRENDS in Microbiology*. P. 20-26. doi: 10.1016/j.tim.2004.11.006
6. Chan, C., Paul, R., Samoray, D., Amiot, N., Giese, B., Jenal, U., & Schirmer, T. (2004). Structural basis of the activity and allosteric control of diguanylate cyclase. *PNAS*. P. 17084-17089. doi: 10.1073/pnas.0406134101
7. Dahlstrom, K. M., Giglio, K. M., Collins, A. J., Sondermann, H., & O'Toole, G. A. (2015). Contribution of physical interactions to signaling specificity between a diguanylate cyclase and its effector. *mBio*. 6(6):e01978-15. doi: 10.1128/mBio.01978-15
8. Dahlstrom, K. M., Giglio, K. M., Collins, A. J., Sondermann, H., & O'Toole, G. A. (2016). The inhibitory site of a diguanylate cyclase is a necessary element for interaction and signaling with an effector protein. *Journal of Bacteriology*. 198(11):1595-603. doi: 10.1128/JB.00090-16.
9. Dahlstrom, K. M., & O'Toole, G. A. (2017). A symphony of cyclases: specificity in diguanylate cyclases signaling. *Annu. Rev. Microbiol.* P. 179-195. doi: 10.1146/annurev-micro-090816-093325
10. D'Argenio, D., & Miller, S. (2004). Cyclic di-GMP as a bacterial second messenger. *Microbiology*. P. 218-228. doi: 10.1016/j.mib.2006.02.010
11. Deepthi, A., Liew, C., Liang, Z.-X., & Swaminathan, K. (2014). Structure of a diguanylate cyclase from *Thermotoga maritima*: insights into activation, feedback inhibition and thermostability. *PLOS ONE*. 9(10):e110912. doi: 10.1371/journal.pone.0110912
12. Düvel J., Bertinetti D., Möller S., Schwede F., Morr M., Wissing J., Radamm L., Zimmermann B., Genieser H.G., Jansch L., Herberg F.W., Häussler S. (2012). A chemical proteomics approach to identify c-di-GMP binding proteins in *Pseudomonas aeruginosa*. *Journal of Microbiological Methods*. P.229-236. doi: 10.1016/j.mimet.2011.11.015
13. Feirer, N., Xu, J., Allen, K. D., Koestler, B. J., Bruger, E. L., Waters, C. M., White, R. H., & Fuqua, C. (2015). A pterin-dependent signaling pathway regulates a dual-

- function diguanylate cyclase-phosphodiesterase controlling surface attachment in *Agrobacterium tumefaciens*. *mBio*. 6(4): e00156-15. doi: 10.1128/mBio.00156-15
14. Giacalone D, Smith TJ, Collins AJ, Sondermann H, Koziol LJ, O'Toole GA. (2018). Ligand-mediated biofilm formation via enhanced physical interaction between a diguanylate cyclase and its receptor. *mBio*. 9:e01254-18. doi: 10.1128/mBio01254-18
  15. Hefti, M., Francoijs, K.-J., de Vries, S., Dixon, R., & Vervoort, J. (2004). The PAS fold A redefinition of the PAS domain based upon structural prediction. *Eur. J. Biochem*. 271(6):1198-208. doi: 10.1111/i.1432-1033.2004.04023.x
  16. Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nature Review: Microbiology*. 7(4):263-73. doi: 10.1038/nrmicro2109.
  17. Hengge, R., Grundling, A., Jenal, U., Ryan, R., & Yildiz, F. (2016). Bacterial signal transduction by cyclic di-GMP and other nucleotide second messengers. *Journal of Bacteriology*. 198(1): 15–26. doi: 10.1128/JB.00331-15
  18. Henry, J.T., & Crosson, S. (2011). Ligand binding PAS domains in a genomic, cellular and structural context. *Annu Rev Microbiol*. 65: 261–286. doi: 10.1146/annurev-micro-121809-151631
  19. Inaba J., Thornton J., Fernandes Huergo L., Monteiro R.A., Klassen G., Oliveira Pedrosa F., Merrick M., & Maltempi de Souza E. (2014). Mutational analysis of GlnB residues critical for NifA activation in *Azospirillum brasilense*. *Microbiological Research*. P. 65-72. doi: 10.1016/j.micres.2014.12.005
  20. Jamal, M., Ahmad, W., Andleeb, S., Jalil, F., Imran, M., Asif Nawaz, M. Hussain, T., Ali, M., Rafiq, M., & Atif Kamil, M. (2017). Bacterial biofilm and associated infections. *Journal of the Chinese Medical Association*. P.7-11. doi: 10.1016/j.jcma.2017.07.012
  21. Jenal U., Reinders A., & Lori C. (2017). Cyclic di-GMP: second messenger extraordinaire. *Nature Review*. Vol.15, p. 271-284. doi: 10.1038/nrmicro.2016.190
  22. Krasteva, P.V., Giglio, K.M., & Sondermann, H. (2012). Sensing the messenger: The diverse ways that bacteria signal through c-di-GMP. *Protein Science*. (7):929-48. doi: 10.1002/pro.2093.
  23. Lin, S.-Y., Young, C. C., Hupfer, H., Siering, C., Arun, A. B., Chen, W.-M., Lai, W.-A., Shen, F.-T., Rekha, P. D., & Yassin, A. (2009). *Azospirillum picis* sp. nov., isolated from discarded tar. *International Journal of Systematic and Evolutionary Microbiology*. 761-765. doi: 10.1099/ij.s.0.65837-0
  24. Lindenberg S., Klauck G., Pesavento C., Klauck E., & Hengge R. (2013). The EAL domain protein YciR acts as a trigger enzyme in a c-di-GMP signaling cascade in *E. coli* biofilm control. *The EMBO Journal*. 32(14): 2001–2014. doi: 10.1038/emboj.2013.120
  25. Loera Muro, A., Ramírez Castillo, F., Avelar González, F., & Guerrero Barrera, A. (2012). Biopelícula: asociarse para sobrevivir. *Investigación y Ciencia de la Universidad Autónoma de Aguascalientes*. Número 54,(49-56). doi: 10.13140/2.1.4508.2567
  26. Lopez, L. (2014). Determinación de la actividad enzimática de diguanilato ciclase de la proteína CdgA de *Azospirillum brasilense* Sp7. *Puebla: Posgrado en Microbiología CICM ICUAP BUAP*.
  27. O'Neal L, Ryu M-H, Gomelsky M, Alexandre G. (2017). Optogenetic manipulation of cyclic di-GMP (c-di-GMP) levels reveals the role of c-di-GMP in regulating

- aerotaxis receptor activity in *Azospirillum brasilense*. *J Bacteriol.* 199:e00020-17. doi: 10.1128/jb.00020-17
28. Qi, Y., Rao, F., Luo, Z., & Liang, Z. (2009). A flavin cofactor-binding PAS domain regulates c-di-GMP synthesis in AxDGC2 from *Acetobacter xylinum*. *Biochemistry.* 48(43):10275-85. doi: 10.1021/bi901121w
  29. Ramey, B., Koutsoudis, M., Bodman, S., & Fuqua, C. (2004). Biofilm formation in plant-microbe associations. *Current Opinion in Microbiology.* P. 602-609. doi: 10.1016/j.mib.2004.10.014
  30. Ramírez-Mata, A., López-Lara, L., Xiqui-Vázquez, M., Jijón-Moreno, S., Romero-Osorio, A., & Baca, B. (2016). The cyclic-di-GMP diguanylate cyclase CdgA has a role in biofilm formation and exopolysaccharide production in *Azospirillum brasilense*. *Research in Microbiology.* Vol. 167, p. 190-201. doi: 10.1016/j.resmic.2015.12.004
  31. Ramírez Mata, A., Millán Pacheco, C., Cruz Pérez, J.F., Minjárez Sáenz, M., & Baca B.E. (2018). In silico comparative analysis of GGDEF and EAL domain signaling proteins from the *Azospirillum* genomes. *BMC Microbiology.* 6: 23769. doi: 10.1038/srep23769
  32. Rodríguez Vázquez, S. (2016). Estudio de la participación del gen AZOBR\_140291 (proteína aceptora de grupos metilo putativo) en la quimiotaxis de *Azospirillum brasilense*. *Puebla, Puebla: Tesis de Posgrado en Microbiología, BUAP.*
  33. Römling, U., Galperin, M., & Gomelsky, M. (2013). Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiology and Molecular Biology Reviews.* 77(1): 1–52. doi: 10.1128/MMBR.00043-12
  34. Römling, U., Liang, Z.-X., & Dow, M. (2017). Progress in understanding the molecular basis underlying functional diversification of cyclic dinucleotide turnover proteins. *Journal of Bacteriology.* 199(5):JB.00790-16 doi: 10.1128/JB.00790-16
  35. Ross P., Weinhouse H., Aloni Y., Michaeli D., Weinberger-Ohana P., Mayer R., Braun S., de Vroom E., van der Marel G.A., van Boom J.H., Benziman M. (1987). Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature.* 325(6101):279-81. doi: 10.1038/325279a0
  36. Ryan, Robert P. & Dow, J. Maxwell. (2010). Cell-cell signal-dependent dynamic interactions between HD-GYP and GGDEF domain proteins mediate virulence in *Xanthomonas campestris*. *Virulence.* 107(13): 5989–5994. doi: 10.1073/pnas.0912839107
  37. Sarenko, O., Klauck, G., Wilke, F. M., Pfiffer, V., Richter, A. M., Herbst, S., Kaefer, V., & Hengge, R. (2017). More than enzymes that make or break cyclic Di-GMP-local signaling in the interactome of GGDEF/EAL domain proteins of *Escherichia coli*. *American Society for Microbiology.* 8(5):e01639-17 doi: 10.1128/mBio.01639-17
  38. Schirmer, T., & Jenal, U. (2009). Structural and mechanistic determinants of c-di-GMP signaling. *Nature Reviews Microbiology.* 7(10):724-35. doi: 10.1038/nrmicro2203
  39. Schirmer, T. (2016). C-di-GMP synthesis: structural aspects of evolution, catalysis and regulation. *Journal of Molecular Biology.* Vol. 248, p. 3683-3701. doi: 10.1016/j.jmb.2016.07.023

40. Smith K.D., & Strobel S.A. (2011). Interactions of the c-di-GMP riboswitch with its second messenger ligand. *Biochem Soc. Trans.* 39(2): 647–651. doi: 10.1042/BST0390647
41. Sonderman H., Shikuma N.J., Yildiz F.H., (2012). You've come a long way: c-di-GMP signaling. *Curr Opin Microbiol.* 15(2):140-6 doi: 10.1016/j.mib.2011.12.008
42. Tallawi, M., Opitz, M., & Lieleg, O. (2017). Modulation of the mechanical properties of bacterial biofilms in response to environmental challenges. *Biomaterials Science.* 5(5), 887–900. doi:10.1039/c6bm00832a
43. Taylor B.L., & Zhulin I.B. (1999). PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiology and Molecular Biology Reviews.* 63(2): 479–506. doi: 1092-2172/99/\$04.0010
44. Trabelsi, D., & Mhamdi, R. (2013). Microbial inoculants and their impact on soil microbial communities: a review. *BioMed Research International.* doi: 10.1155/2013/863240
45. Valentini, M., & Filloux, A. (2016). Biofilms and cyclic di-GMP (c-di-GMP) signaling: lessons from *Pseudomonas aeruginosa* and other bacteria. *The Journal of Biological Chemistry.* 291(24): 12547–12555. doi: 10.1074/jbc.R115.711507
46. Whiteley, C., & Lee, D.-J. (2015). Bacterial diguanylate cyclases: structure, function and mechanism in exopolysaccharide biofilm development. *Elsevier: Biotechnology Advances.* Vol. 33, p. 124-141. doi: 10.1016/j.biotechadv.2014.11.010

## 14.- ANEXOS

### ANEXO 1.- Alineamiento entre las proteínas MCP de *A. brasilense* Sp 245 y *A. brasilense* Sp7

CLUSTAL O(1.2.4) multiple sequence alignment

```

AZOBR_140291      -----MAALMSGAVYGLATGDATAALLGLLAVLPCLAAIRWLVLTNRT 45
AMK58_RS08090    MTSLSSVSKSLAAAASAALLSGAVYGFMTGDAVAAALGLLAILPCLAAIRWLVLTNRT 60
                  ***:*****:***:** *****:*****:*****
AZOBR_140291      IRKASTVIAAAEKGDLPRIILNIHGASSIADMLRTINRLDRVESFGKEANAAMQHAAEG 105
AMK58_RS08090    IHKAIAVIAAAEKGDLPRIILNIQGASPIADMLRTINRLDRVESFGKEANAAMQHAAEG 120
                  *:* :*****:*** *****:*****
AZOBR_140291      QYYRRIVMTGMVGEFGAYARQINDGLAAMDGKSREFVESATRIGANIKEVAQSLASAAQ 165
AMK58_RS08090    QYYRRIVMTGMVGEFGAYAHQINDGLAAMDGKSREFVESATRIGANIKEVAQSLASAAQ 180
                  *****:*****
AZOBR_140291      LEASSTAMTATAATASEQSFSAASAAEQVSSNVDGVAATGEVSGAIGEVAQGVSRADL 225
AMK58_RS08090    LEASSTAMTATAATASEQSFSAASAAEQVSSNVDGVAATGEVSGAIGEVAQGVSRADL 240
                  *****
AZOBR_140291      ARSSVEKVREADATIRSLLAASDQIGAVVQLINDIASQTNLLALNATIEAARAGEAGKGF 285
AMK58_RS08090    ARSSVEKVREADATIRSLLAASDQIGAVVQLINDIASQTNLLALNATIEAARAGEAGKGF 300
                  *****
AZOBR_140291      AVVATEVKNLANQTAKATEDITAQITQVQSVTRDTAAVIQHVGGMIHDIQIAVGIAGAA 345
AMK58_RS08090    AVVATEVKNLANQTAKATEDISAQISQVQSVTRDTAAVIQHVGGMIHDIQIAVGIAGAA 360
                  *****:***:*****:*****
AZOBR_140291      EQQSAAIDEISRSIREASAGVRTVADAVTSVSSGQDASAAASQVLSSAGELARRAVTLN 405
AMK58_RS08090    EQQSAAIDEISRSIREASAGVRTVADAVTSVSSGQDASAAASQVLSSAGELARRAETLN 420
                  ***** **
AZOBR_140291      GDIDNFVARVCGGQR 420
AMK58_RS08090    GDIDKFVARVCGGQR 435
                  ***:*****

```

### ANEXO 2.- Alineamiento entre los dominios EAL de las proteínas AMK58\_RS05115 y CAJ18244

```

AMK58_RS05115    DAVDRIGRLRATIEKRDFVYQPIVDIGTRTVHHLALTRVEGTLSPADFITFAEDVGL 60
CAJ18244         DAVDRIGRLRATIEKRDFVYQPIVDIGTRTVHHLALTRVEGTLSPADFITFAEDVGL 60
                  *****
AMK58_RS05115    IYDFDLLLTQSVLDTMHAHRKEPTLPDVAINLSAKTLMSPIFLRQFLAVTEPYGELARKL 120
CAJ18244         IYDFDLLLTQSVLDTMHAHRKEPTLPDVAINLSAKTLMSPIFLRQFLAVTEPYGELARKL 120
                  *****
AMK58_RS05115    LIEVTETVVVTDIAKLNEVLQKLRAAGFRICLDDVGAGSTSFQSLYGIQADFAKIDGHFV 180
CAJ18244         LIEVTETVVVTDIAKLNEVLQKLRAAGFRICLDDVGAGSTSFQSLYGIQADFAKIDGHFV 180
                  *****
AMK58_RS05115    RGAVHNGRDMAMLRSMVDVCRQLGLGLIGEVEGPEHAQLLTELGVSLAQGYLYSRPSRD 240
CAJ18244         RGAVHNGRDMAMLRSMVDVCRQLGLGLIGEVEGPEHAQLLTELGVSLAQGYLYSRPSRD 240
                  *****
AMK58_RS05115    FAYFAKDWTKMAGTGGKVLKRG 262
CAJ18244         FAYFAKDWTKMAGTGGKVLKRG 262
                  *****

```