



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

---

CENTRO DE CIENCIAS GENÓMICAS

**Caracterización funcional de los reguladores hFixL y FikR en la cascada de regulación de las proteínas tipo FNR de *R. etli* CFN42**

T E S I S

QUE PARA OBTENER EL GRADO ACADÉMICO DE  
DOCTOR EN CIENCIAS

P R E S E N T A :

**L.I.B.B. DAVID SALVADOR ZAMORANO SÁNCHEZ**

TUTOR: DRA. MARÍA DE LOURDES GIRARD CUESY



Centro de Ciencias Genómicas

CUERNAVACA MOR.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Este trabajo fue realizado en el Programa de Genómica Funcional de Procariontes del Centro de Ciencias Genómicas de la Universidad Nacional Autónoma de México, bajo la dirección de la Dra. María de Lourdes Girard Cuesy.

# Índice

	Pág.
<b>Resumen</b>	1
<b>Abstract</b>	2
<b>Introducción General</b>	
a) La fijación biológica del nitrógeno	3
b) Proteínas detectoras de oxígeno en organismos fijadores de nitrógeno	6
c) Sistemas de dos componentes para la transducción de señales	10
d) Regulación genética de la fijación de nitrógeno	14
e) <i>Rhizobium etli</i> y <i>Rhizobium leguminosarum</i> bv. <i>viciae</i> como modelo de estudio para el análisis de la regulación de la fijación de nitrógeno	17
<b>Planteamiento del Problema</b>	21
<b>Objetivos</b>	22
<b>Diseño Experimental</b>	24

## **Materiales y Métodos**

1. Cepas bacterianas y plásmidos	26
2. Medios de cultivo y antibióticos	27
3. Condiciones de crecimiento microaeróbicas	27
4. Medición de la actividad de $\beta$ -glucuronidasa	28
5. Amplificación por PCR	28
6. 5' RACE (Amplificación Rápida de Extremos de cDNA)	28
7. Construcción de Plásmidos	29
8. Derivadas mutantes de <i>R. etli</i>	33
9. Purificación de His6-FxkR	33
10. Ensayo de cambio en la movilidad electroforética (“Electroforetic mobility shift analysis” EMSA)	34
11. Recursos bioinformáticos	34

## **Resultados**

Análisis de los elementos genéticos necesarios para la expresión de <i>fixKf</i>	36
Identificación de FxkR como el regulador de la respuesta cognado de la cinasa de histidinas hFixL en <i>R. etli</i>	37
El plásmido pRL9 de <i>R. leguminosarum</i> bv. <i>viciae</i> 3841 posee un gen ortólogo a <i>fxkR</i>	42
Identificación de proteínas tipo-FxkR en las $\alpha$ -proteobacterias	44
Análisis comparativo de la secuencia primaria de las proteínas tipo-FxkR y los reguladores de la respuesta OmpR y PhoB	47
La región reguladora de los genes tipo- <i>fixKf</i> comparte una caja reguladora	50
FxkR se une a la región intergénica <i>fixNf-fixKf</i> <i>in vitro</i>	54

Predicción de la interfaz de unión a DNA por métodos bioinformáticos	55
<b>Discusión y Perspectivas</b>	61
<b>Bibliografía</b>	66
<b>Anexo I</b>	76
<b>Anexo II</b>	78
<b>Anexo III</b>	80

## LISTA DE TABLAS

	Pág.
Tabla 1. Cepas bacterianas y plásmidos	26
Tabla 2. Oligos usados para la construcción de los plásmidos suicidas utilizados para la mutagénesis por inserción	30
Tabla 3. Oligos utilizados para la construcción de plásmidos	32
Tabla 4. Identificación del regulador de la respuesta que controla la expresión microaeróbica de <i>fixKf</i> en <i>R. etli</i> CFN42	39
Tabla 5. hFixL y FxkR son necesarios para la expresión microaeróbica de <i>fixKf</i> de <i>R. etli</i>	41
Tabla 6. <i>R. leguminosarum</i> bv. <i>viciae</i> tiene un ortólogo funcional de <i>fxkR</i> de <i>R. etli</i>	43

## LISTA DE FIGURAS

	Pág.
Figura 1. Estructura cuaternaria de la nitrogenasa	4
Figura 2. Arquitectura de dominios y estructura tridimensional de FixL y NifL	6
Figura 3. Interacción del dominio PAS con el cofactor hemo b en FixL	8
Figura 4. Química de fosfotransferencia de sistemas de dos componentes canónicos	11
Figura 5. Química de fosfotransferencia de sistemas de relevo de fosfatos	14
Figura 6. Cascada de regulación de los genes <i>fix</i> y <i>nif</i> de <i>Sinorhizobium meliloti</i> y <i>Bradyrhizobium japonicum</i>	16
Figura 7. Cascada de regulación de los genes <i>fix</i> , <i>nir</i> y <i>nor</i> en <i>Rhizobium etli</i> CFN42	18
Figura 8. Participación de los plásmidos pRet42d y pRet42f en la regulación de la expresión microaeróbica de <i>fixKf</i>	37
Figura 9. Arreglo genómico de los probables reguladores de la respuesta en pRet42f.	38
Figura 10. Análisis de la expresión de <i>fxkR</i> de <i>Rhizobium etli</i> en condiciones de vida libre	42
Figura 11. Representación esquemática del contexto genómico y análisis comparativo de genes tipo- <i>fixKf</i> , <i>hfixL</i> y <i>fxkR</i> de $\alpha$ -proteobacteria	45
Figura 12. Alineamiento múltiple de proteínas FxkR con los reguladores OmpR y PhoB de <i>E. coli</i>	47
Figura 13. Identificación de la caja-K en la región reguladora de genes tipo- <i>fixKf</i>	51
Figura 14. Análisis genético del probable sitio de unión de FxkR	53
Figura 15. La proteína recombinante FxkR de <i>R. etli</i> se une específicamente a la caja-K	55
Figura 16. Predicción de estructura secundaria de las proteínas PhoB y FxkR	56

Figura 17. Contactos en la interfaz de la proteína FxkR de <i>R. etli</i> y la caja-K de <i>fixKf</i>	57
Figura 18. Representación gráfica de los contactos en la interfaz Proteína-DNA de un modelo tridimensional de FxkR en el cual se sustituyó la Val210 por una Arginina	59
Figura 19. Estimación de la especificidad de unión de FxkR a la caja-K de <i>fixKf</i>	60



## Resumen

El control transcripcional del gen *fixK* en *Rhizobium etli* y *Rhizobium leguminosarum* bv. *viciae* está dado por un sistema de transducción de señales de dos componentes que difiere del sistema convencional, FixL-FixJ. *R. etli* posee un gen *fixL* el cual codifica para una cinasa de histidinas híbrida (hFixL), sin un homólogo a *fixJ* que funcione como su regulador de la respuesta cognado. En este trabajo presentamos evidencia de que el gen RHE\_PF00530 (*fxkR*) localizado en el plásmido pRet42f codifica para un regulador de la respuesta indispensable para la expresión microaeróbica de *fixKf*. Mediante ensayos de complementación demostramos que la activación de *fixKf* requiere de la presencia tanto de hFixL como de FxkR, y que el ortólogo de *fxkR* de *R. leguminosarum* bv. *viciae* es capaz de sustituir a FxkR en el fondo genético de *R. etli*. Con métodos bioinformáticos identificamos la presencia de proteínas tipo-hFixL y FxkR codificadas en la vecindad de un gen *fixK* en otras bacterias pertenecientes al grupo de las  $\alpha$ -proteobacterias. Usando fusiones transcripcionales con genes reporteros, mutaciones sitio-dirigidas y ensayos de EMSA, identificamos un sitio de unión de FxkR río arriba del inicio de transcripción de *fixKf*. Nuestros resultados muestran que una mutación nula en el gen *fxkR* no afecta la capacidad de fijar nitrógeno de la cepa de manera similar a lo observado previamente en mutantes en los genes *hfixL* y *fixKf*. En conclusión los resultados de este trabajo revelan que FxkR es el regulador de la respuesta que opera con hFixL para la transducción de la señal microaeróbica y la activación del regulón de FixKf en *R. etli* CFN42.

## Abstract

Transcriptional control of the *fixK* gene in *Rhizobium etli* and *Rhizobium leguminosarum* bv. *viciae* is governed by a two component signal transduction system that diverts from the conventional FixL-FixJ cascade that occurs in model rhizobia. Although a *fixL* gene, encoding a hybrid histidine kinase (hFixL), is present in *R. etli*, no *fixJ*, the cognate response regulator, has been identified. In this work, we present evidence that the pRet42f-located ORF RHE\_PF00530 (*fxkR*) encodes a novel response regulator indispensable for *fixKf* activation under microaerobic growth. Moreover, results from complementation assays demonstrate that the activation of *fixKf* expression requires the presence of both hFixL and FxkR, and that the *fxkR* ortholog from *R. leguminosarum* bv. *viciae* is able to substitute for FxkR transcriptional control in *R. etli*. Besides, in these two organisms, hFixL and FxkR-related proteins in other  $\alpha$ -proteobacteria could be identified located in close proximity to a *fixK*-related gene. Using reporter fusions, site directed mutagenesis and EMSA assays, we identified the FxkR binding site upstream from the transcriptional start site of *fixKf*. Similar to our previous observations for *fixL* and *fixKf* mutants, a null mutation in *fxkR* does not affect the symbiotic effectiveness of the strain. Thus, our findings reveal that FxkR is the long-standing missing key regulator that allows the transduction of the microaerobic signal for the activation of the FixKf regulon.

## Introducción General

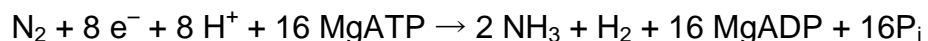
### *a) La fijación biológica del nitrógeno*

El nitrógeno es una molécula fundamental para la vida ya que forma parte de los ácidos nucleicos y las proteínas. El porcentaje de este gas en la atmósfera terrestre es del 78%; sin embargo para poder ser asimilado por los organismos debe de ser convertido en amoníaco o en nitrato. La fijación biológica de nitrógeno consiste en la reducción del dinitrógeno a amoníaco por bacterias especializadas que sintetizan la enzima que cataliza esta reacción, la nitrogenasa.

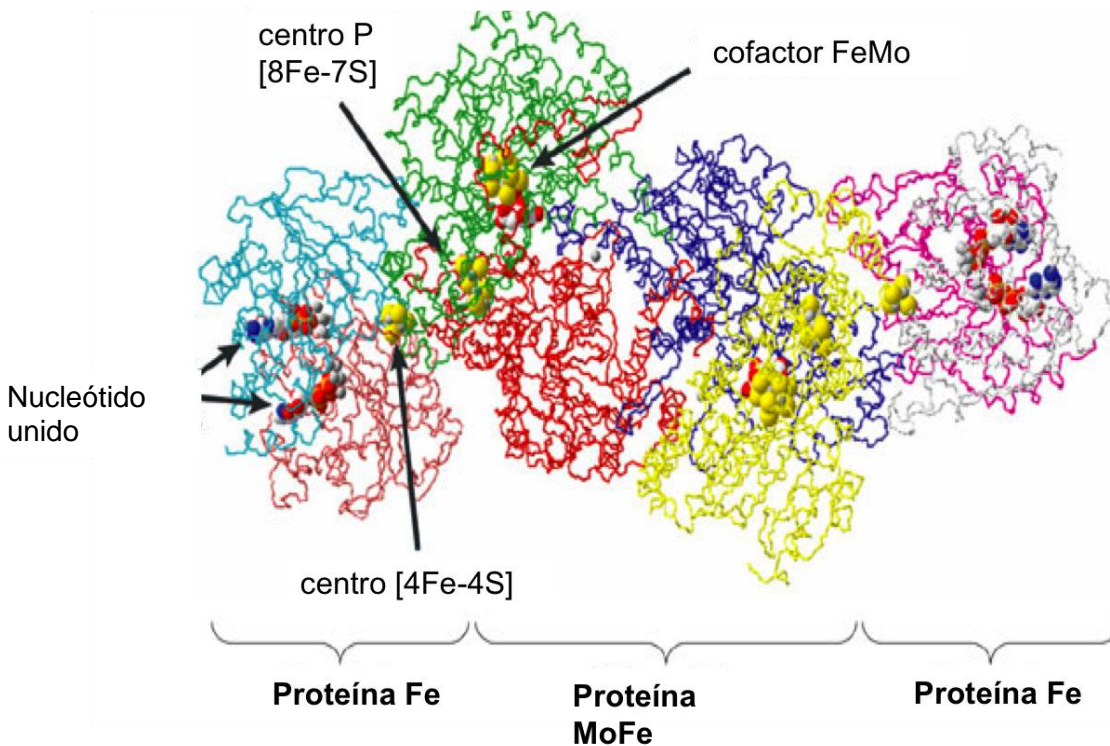
La nitrogenasa tiene una estructura y un mecanismo de acción ampliamente conservado (Rees DC. y Howard JB., 2000; Lawson DM. y Smith BE., 2002; Dixon R. y Kahn D., 2004). Es un complejo enzimático formado por dos componentes. Un componente dimérico, conocido como la proteína de hierro, el cual es el responsable de donar los electrones de forma dependiente de ATP. El otro componente es un heterotetrámero que es capaz de unir un cofactor de molibdeno-hierro (MoFe) y contiene el sitio activo de la enzima. Todos los organismos fijadores de nitrógeno poseen una nitrogenasa que incluye a la proteína molibdeno-hierro; sin embargo, en condiciones de escasez de molibdeno algunas bacterias producen un componente capaz de usar un cofactor de vanadio-hierro o de hierro-hierro para llevar a cabo la reducción de dinitrógeno (Eady RR., 1996; Dixon R y Kahn D., 2004).

La proteína de hierro de la nitrogenasa posee un centro de 4Fe-4S, mientras que la proteína molibdeno-hierro posee el centro P de 8Fe-7S y el cofactor FeMo. La presencia de estos centros metálicos hace que la nitrogenasa sea una enzima altamente sensible al oxígeno (Einsle O. y col. 2002; Seefeldt LC. y col. 2004; Dixon R y Kahn D., 2004).

La estequiometría de la reacción de fijación de nitrógeno es la siguiente:



Los electrones son donados a la proteína de hierro a través de donadores como la ferredoxina o la flavodoxina. La proteína de hierro transfiere uno a uno los electrones a la proteína de molibdeno-hierro, de manera dependiente de ATP, en cada ciclo catalítico el complejo se asocia y se disocia (Hageman RV. y Burris RH.,1978; Dixon R. y Kahn D., 2004).



**Figura 1.** Estructura cuaternaria de la nitrogenasa. Se muestra la estructura tridimensional del complejo oligomérico de la nitrogenasa de *Azotobacter vinelandii* compuesto por un par de dímeros de la proteína de hierro y las subunidades alfa y beta de la proteína MoFe. Con flechas se indica el sitio de unión del nucleótido que estabiliza la estructura (ADP-AIF<sub>4</sub>) y los centros metálicos que participan directamente en el transporte de electrones y en la catálisis de la reacción. Figura modificada de Dixon R. y Kahn D., 2004

Como se puede observar al analizar la estequiometría de la reacción, el proceso de fijación de nitrógeno es energéticamente costoso, requiere de un metabolismo activo capaz de suministrar el ATP necesario para que la catálisis se lleve a cabo. Como se mencionó anteriormente, la nitrogenasa es sensible al oxígeno, que al menos en el

caso de los organismos aerobios es el último aceptor de electrones en la cadena respiratoria. A esto se le conoce como la paradoja del oxígeno.

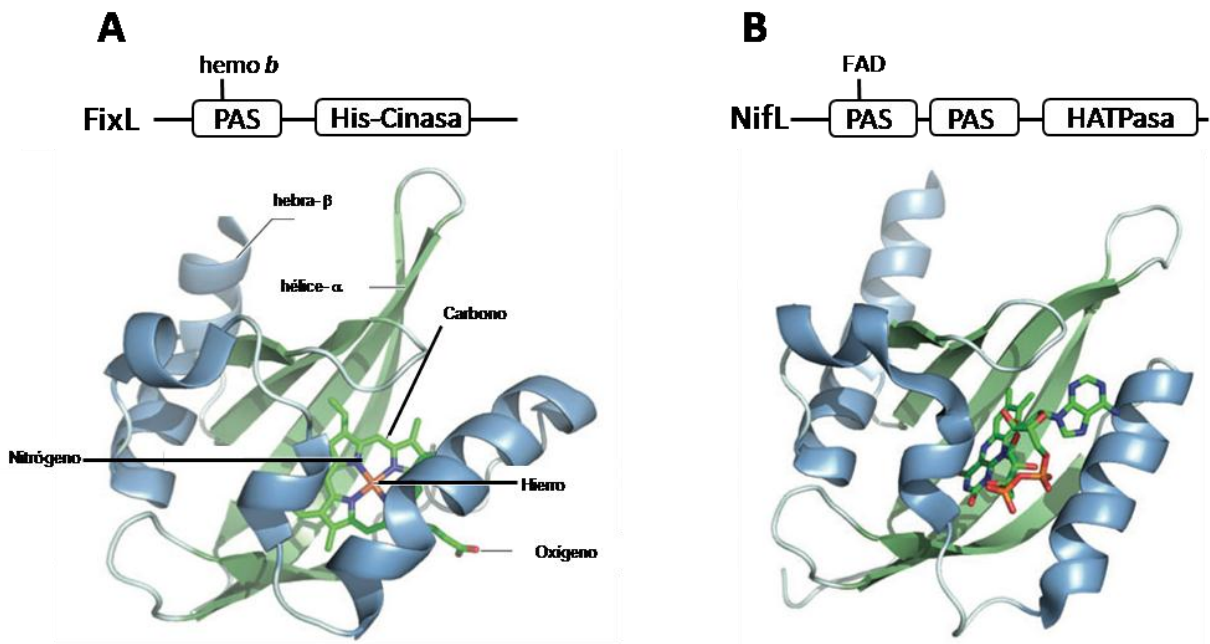
Existen diferentes mecanismos para que la fijación de nitrógeno se lleve a cabo a pesar de los factores negativos que la limitan. Algunas bacterias facultativas o capaces de respirar o crecer en ausencia de oxígeno, pueden fijar nitrógeno en ambientes anaerobios, algunos fijadores aerobios estrictos limitan el oxígeno capaz de reaccionar con la nitrogenasa optimizando su reducción en la cadena respiratoria a través de citocromo oxidasa de alta afinidad. Un ejemplo de los mecanismos empleados para que se lleve a cabo la fijación de nitrógeno, es el de la simbiosis leguminosa-rizobio. Los rizobios son capaces de colonizar las raíces de las plantas leguminosas e inducen la formación de estructuras especializadas en las cuales se lleva a cabo la fijación de nitrógeno. Estas estructuras se conocen como nódulos. En los nódulos, las células del córtex forman una barrera de difusión del oxígeno que permite que la concentración de esta molécula se mantenga en el orden nanomolar. En estas condiciones es fundamental un sistema de transporte intracelular de oxígeno que garantice la llegada del oxígeno a la bacteria, esta función la cumple una proteína de la planta llamada leghemoglobina. Por su parte, la bacteria requiere de una citocromo oxidasa capaz de detectar concentraciones nanomolares de oxígeno, para que su metabolismo se mantenga activo (Dixon R. y Kahn D., 2004).

Al ser la fijación de nitrógeno un proceso de alto costo energético, la célula cuenta con mecanismos para garantizar que este se lleve a cabo exclusivamente cuando las condiciones ambientales sean favorables y cuando el producto de la reacción sea necesario. Es por ello que la detección de la concentración de oxígeno favorable para la reacción y el monitoreo de la poza de nitrógeno intracelular son puntos clave para la regulación del proceso.

b) *Proteínas detectoras de oxígeno en organismos fijadores de nitrógeno*

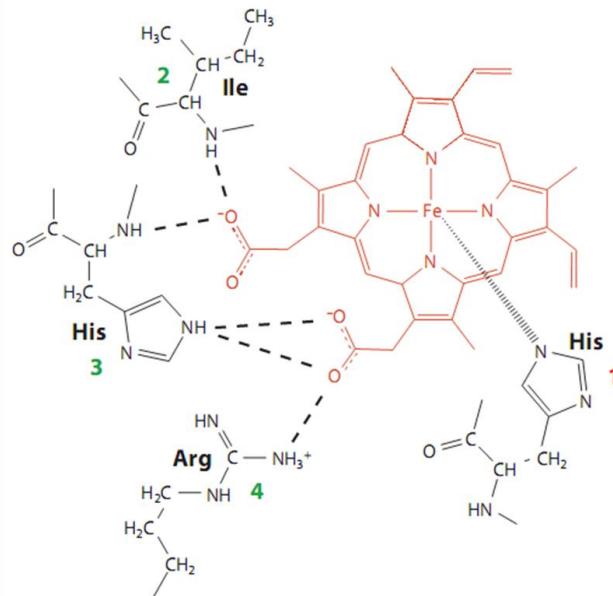
Las bacterias fijadoras de nitrógeno tienen la capacidad de responder a cambios en la concentración de oxígeno a través de diferentes tipos de sistemas de regulación. Se reconocen 4 proteínas detectoras de oxígeno: FixL, NifA, NifL y RegB.

Los dominios estructurales de FixL y NifL son análogos a excepción del dominio de cinasa de histidinas del cual carece NifL. Ambas proteínas poseen uno o dos dominios PAS en el extremo amino terminal, los cuales se han asociado a proteínas detectoras de oxígeno, de estado redox, de voltaje y de luz tanto en organismos procariontes como eucariontes (Taylor BL. y Zhulin IB., 1999). El mecanismo de detección es diferente para estas dos proteínas. Mientras que el dominio PAS de FixL posee un grupo prostético hemo *b*, el dominio PAS de NifL tienen como grupo prostético un grupo de flavina adenina dinucleótido (FAD) (Gilles-Gonzalez, MA. y col. 1991; Hill S. y col. 1996; Henry JT. y Crosson S. 2011).



**Figura 2.** Arquitectura de dominios y estructura tridimensional de FixL y NifL. Arquitectura de dominios de las proteínas FixL de *B. japonicum* y NifL de *Azotobacter vinelandii* y estructura tridimensional de los dominios PAS unidos a sus diferentes cofactores hemo *b* y FAD respectivamente. Figura modificada de Henry JT. y Crosson S. 2011.

La unión reversible de oxígeno al grupo hemo en FixL regula la actividad del dominio de cinasa de histidinas. En ausencia de oxígeno FixL se autofosforila y es capaz de transferir el fosfato al regulador de la respuesta FixJ (Tuckerman JR. y col. 2002). La estructura de FixL de *Sinorhizobium meliloti* y *Bradyrhizobium japonicum*, con y sin ligando ha sido reportada permitiendo sugerir un modelo estructural de detección de oxígeno. El residuo de histidina que une al dominio PAS de FixL con el hierro coordinado en el cofactor hemo *b* se encuentra conservado en otras proteínas con dominio PAS que unen este mismo tipo de cofactores, como la proteína DosP de *Escherichia coli*. Sin embargo otros residuos de histidina, arginina e isoleucina hacen contacto con las cadenas laterales del propionato. Estas interacciones parecen ser específicas de las proteínas tipo FixL (Henry JT. y Crosson S. 2011)(Figura 3). En FixL de *B. japonicum* se propone que la Arg<sup>220</sup> forma un puente salino con el propionato del grupo hemo, el cual se rompe en presencia de oxígeno (Hao B. y col. 2002). El mecanismo de regulación alostérica del dominio de cinasa de histidinas no es claro aún.



**Figura 3.** Interacción del dominio PAS con el cofactor hemo *b* en FixL. Representación gráfica de las interacciones de cadenas laterales y del esqueleto peptídico de residuos de aminoácidos conservados en las proteínas FixL con el cofactor hemo *b*. La interacción marcada con el número 1 corresponde a un residuo de histidina conservado en los dominios PAS que interacciona con cofactores hemo *b*. El resto de las interacciones son específicas de FixL. Figura modificada de Henry JT. y Crosson S. 2011.

NifL es un anti-activador del regulador NifA, su efecto regulador depende de la capacidad de interactuar con su blanco. En su estado oxidado el grupo FAD en el dominio PAS permite la interacción entre estas proteínas, sin embargo en el estado reducido FADH<sub>2</sub> NifL de *Azotobacter vinelandii* es incapaz de reprimir la actividad de NifA (Hill S. y col. 1996). A pesar de que NifL posee en el dominio carboxilo terminal un dominio de HATPasa, de la misma familia que el dominio HATPasa de las cinasas de histidina, es incapaz de hidrolizar ATP. Sin embargo, este dominio tiene la capacidad de unir nucleótidos de adenina y es fundamental para la interacción con NifA (Söderbäck E. y col. 1998). Aún se desconoce si cambios estructurales en el dominio PAS en respuesta al estado redox pueden ser transmitidos al dominio carboxilo terminal de NifL. En *Klebsiella pneumoniae* se ha propuesto al oxígeno como la molécula oxidante y a la poza reducida de quinonas como la fuente de transporte de electrones al grupo FAD. La reducción de NifL está asociada a su unión reversible a la



membrana, el hecho de que NifA permanezca siempre en el citoplasma, hace suponer que éste es un mecanismo adicional para evitar que NifA y NifL interactúen en condiciones anaeróbicas (Grabbe R. y Schmitz RA., 2003).

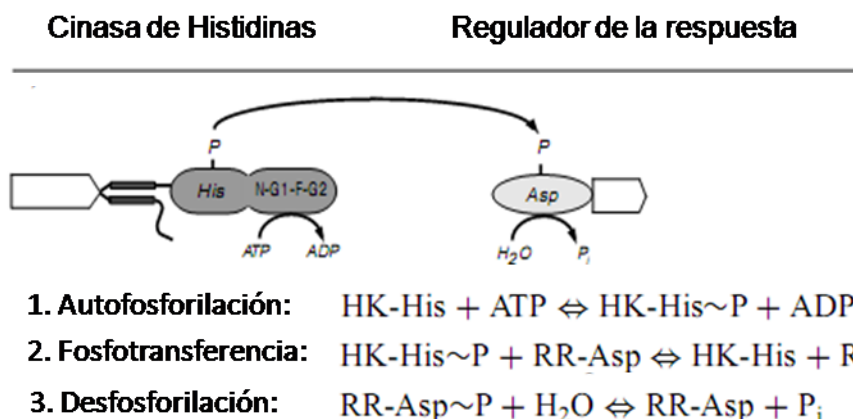
Los genes *nif* codifican para proteínas esenciales para el proceso de fijación de nitrógeno, entre ellas están las del complejo enzimático de la nitrogenasa (NifHDK) y NifA el regulador transcripcional encargado de activar la expresión de estos genes. Este regulador pertenece a la familia EBP (“enhancer binding proteins”) y funciona en conjunto con el factor sigma 54 de la RNA polimerasa. La estructura de NifA es similar a la de otros miembros de la familia. Posee un dominio central AAA<sup>+</sup> ATPasa flanqueado por un dominio regulador en el extremo amino terminal y un dominio de unión a DNA del tipo HTH en el extremo carboxilo terminal que le permite contactar las cajas UAS (Upstream-Activator-Sequences) típicas de este tipo de reguladores. En el extremo amino terminal se encuentra un dominio GAF capaz de interactuar con moléculas pequeñas como nucleótidos cíclicos (Ho YS. y col. 2000). El papel de este dominio en proteínas NifA de diferentes diazótrofos es variable, por ejemplo NifA de *A. vinelandi* une 2-oxoglutarato a través del dominio GAF y esto permite la interacción con el anti-activador NifL; mientras que la unión de 2-oxoglutarato no se ha observado en NifA de *K. pneumoniae* (Little R. y Dixon R. 2003). En organismos en los cuales no hay NifL, NifA posee una región de aproximadamente 100 residuos entre el dominio central AAA<sup>+</sup> y el dominio de unión a DNA en la que se localiza un motivo invariable del tipo Cys-X<sub>4</sub>-Cys. Adicionalmente, el dominio de unión a DNA tiene otra cisteína conservada (Fischer HM. 1994). La presencia de cada una de estas cisteínas es fundamental para la función de NifA. La presencia de metales en particular Fe<sup>2+</sup> es necesaria para la actividad de esta proteína lo cual sugiere un mecanismo de coordinación de hierro por estas cisteínas conservadas (Fischer HM. y col.1988).

El sistema de dos componentes sensible al estado redox RegB-RegA inicialmente descrito en  $\alpha$ -proteobacterias fotosintéticas, es capaz de regular la expresión del regulador NifA en diferentes diazótrofos. En *Rhodobacter capsulatus* este sistema, en conjunto con NtrC, activa la expresión de *nifA2*. En *B. japonicum* los homólogos RegS-RegR controlan la expresión de *nifA* durante la simbiosis (Elsen S. y col. 2000; Emmerich R. y col. 2000; Bauer E. y col.1998).

La cinasa de histidinas RegB y sus homólogos, poseen una cisteína conservada en la región que une al dominio de autofosforilación y al dominio de unión de ATP, fundamental para detectar el estado redox. En su estado oxidado esta cisteína forma un puente disulfuro intermolecular que convierte a la proteína de un estado dimérico activo a un estado tetramérico inactivo. La formación de este puente disulfuro es dependiente de un metal, probablemente, en condiciones fisiológicas el cofactor sea cobre (Swem LR. y col. 2003). En PrrB de *R. sphaeroides*, un homólogo de RegB, el flujo de electrones a través del complejo *cbb<sub>3</sub>* regula su actividad de fosfatasa. Cuando la concentración de oxígeno es alta, la actividad de fosfatasa de PrrB se incrementa sin influir en la actividad de cinasa (Oh J. y col. 2004). Esto sugiere que la capacidad de detección de este tipo de proteínas depende tanto de la cisteína conservada como de la regulación por la *cbb<sub>3</sub>*.

### *c) Sistemas de dos componentes para transducción de señales*

La transducción de señales por sistemas de dos componentes ocurre en bacterias, arqueas, hongos y en algunas plantas. Los sistemas de dos componentes están constituidos por una proteína detectora de la señal y una proteína efectora a través de la cual se desencadenan una serie de cambios metabólicos o morfológicos en respuesta al estímulo detectado. La detección de la señal debe promover un cambio a nivel estructural que se traduzca en una nueva actividad o un cambio de un estado activo a un estado inactivo o viceversa. Este cambio tendrá como consecuencia modificaciones fisiológicas. En muchos sistemas de transducción de señales la fosforilación de residuos de aminoácidos como serina, treonina, tirosina, histidina y aspartato tienen un papel fundamental como modificadores post-traduccionales. En el caso particular de los sistemas de dos componentes la química de la transferencia del fosfato involucra tres reacciones y dos fosfo-proteínas intermediarias (Stock AM. y col. 2000; Stock JB., y col. 1995) (Figura 4).



**Figura 4.** Química de fosfotransferencia de sistemas de dos componentes canónicos. Representación gráfica de los dominios funcionales de un sistema de dos componentes canónico y la química de fosfotransferencia. Figura modificada de Stock AM. y col. 2000.

La primera reacción es la de autofosforilación, en ella participa la proteína detectora la cual tiene actividad de cinasa de histidinas. Las proteínas cinasa de histidina (HK) poseen al menos dos dominios, uno de ellos es el dominio detector. En las proteínas de este tipo que se encuentran ancladas a la membrana el dominio detector puede percibir señales externas, y señales que se producen en la región periplásmica o que no son capaces de atravesar la membrana plasmática. Sin embargo, no todas las HK poseen segmentos transmembranales lo cual hace pensar que estas proteínas detectan señales que se generan en el interior de la célula o que difunden a través de la membrana. Los cambios estructurales en el dominio detector en respuesta a un estímulo tienen como consecuencia la activación del dominio catalítico de cinasa de histidinas. En el dominio de cinasa de histidinas se lleva a cabo la transferencia del grupo  $\gamma$ -fosforilo del ATP a una histidina conservada que se localiza en un subdominio al que se le conoce como caja H. A pesar de que esta histidina se encuentra cercana al dominio catalítico, en la mayoría de estas proteínas la transferencia del fosfato se lleva a cabo de forma intermolecular. Las HK forman complejos diméricos, uno de los

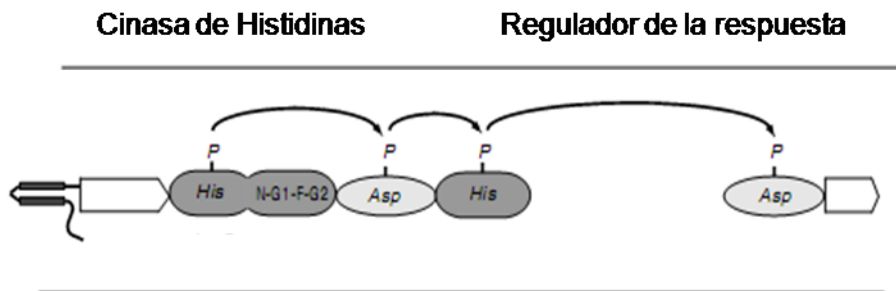
monómeros toma el grupo fosforilo del MgATP y cataliza la fosforilación del residuo de histidina en la caja H de otro monómero. Otro subdominio de las HK está compuesto por las cajas N, D, F, y G. Se presume que estos motivos altamente conservados forman en la estructura terciaria una superficie de unión para el nucleótido dentro del sitio activo (Stock AM. y col. 2000; Stock JB., y col. 1995).

La segunda reacción es la de fosfotransferencia, y es catalizada por el segundo componente del sistema al que se le conoce como regulador de la respuesta (RR). Esta proteína transfiere el grupo fosforilo del residuo fosfo-Histidina a un residuo conservado de aspartato dentro de sí misma en un dominio llamado receptor del fosfato (Rec). Además del dominio Rec estas proteínas poseen generalmente un dominio efector que en la mayoría de los casos le confiere la capacidad de unirse al DNA y de controlar la transcripción de genes. Existen tres principales familias de reguladores de la respuesta dependiendo del tipo de dominio de unión a DNA, la familia OmpR-PhoB, la familia NtrC y la familia NarL-FixJ. En estas proteínas el estado fosforilado controla su actividad como reguladores transcripcionales (Stock AM. y col. 2000; Stock JB., y col. 1995).

La tercera reacción es la desfosforilación, la cual se da por la hidrólisis del grupo fosforilo del residuo fosfo-Aspartato y que puede ser catalizada por el mismo RR, o en algunos casos por la HK que lo fosforila, u otras proteínas auxiliares que tienen actividad de fosfatasa (Stock AM. y col. 2000; Stock JB., y col. 1995).

Existen esquemas más complejos de fosfotransferencia dentro de los sistemas de dos componentes. Algunos autores han nombrado sistemas de relevo de fosfato a aquellos en los que participan más de un módulo receptor del fosfato (Rec1 y Rec2) y al menos un dominio de fosfotransferencia con una histidina conservada (HPt). Estos sistemas están compuestos por una cinasa de histidinas híbrida (hHK) que posee además del dominio detector y el dominio catalítico al dominio Rec1. Como se menciona anteriormente la química de la transferencia del fosfato en los sistemas de dos componentes exige que el grupo fosforilo vaya de un residuo fosfo-Histidina a un residuo de aspartato. Un regulador de la respuesta es incapaz de catalizar la transferencia de un grupo fosforilo en un residuo fosfo-Aspartato (Rec1) a su dominio receptor (Rec2). Por este motivo para que el grupo fosforilo en el dominio Rec1 de una

hHK pueda ser transferido a la proteína reguladora de la respuesta (proteína efectora) se requiere de un dominio adicional llamado HPt. Este dominio carece de actividad de cinasa o de fosfatasa y funciona exclusivamente como puente entre hHK y sus RR cognados. Los dominios HPt pueden estar dentro de la hHK como es el caso de ArcB o en proteínas independientes como ocurre en la mayoría de los sistemas de dos componentes descritos en organismos eucariontes. En los sistemas de relevo de fosfato además de la reacción de autofosforilación existen al menos tres reacciones de fosfotransferencia. La primera involucra el traspaso del grupo fosforilo en el residuo fosfo-Histidina en la hHK al residuo de aspartato en el dominio Rec1 dentro de la misma proteína. En la segunda reacción el grupo fosforilo es transferido del residuo fosfo-aspartato en Rec1 al dominio HPt en un residuo de histidina. La tercera reacción de fosfotransferencia es catalizada por el RR cognado y permite la transferencia del grupo fosforilo a su dominio Rec2 (Stock AM. y col. 2000) (Figura 5).



1. **Autofosforilación:**  $HK-His_I + ATP \rightleftharpoons HK-His_I \sim P + ADP$
2. **Fosfotransferencia 1:**  $HK-His_I \sim P + RR-Asp_I \rightleftharpoons HK-His_I + RR-Asp_I \sim P$
3. **Fosfotransferencia 2:**  $RR-Asp_I \sim P + HPt-His_{II} \rightleftharpoons RR-Asp_I + HPt-His_{II} \sim P$
4. **Fosfotransferencia 3:**  $HPt-His_{II} \sim P + RR-Asp_{II} \rightleftharpoons HPt-His_{II} + RR-Asp_{II} \sim P$
5. **Desfosforilación:**  $RR-Asp_{II} \sim P + H_2O \rightleftharpoons RR-Asp_{II} + P_i$

**Figura 5.** Química de fosfotransferencia de sistemas de relevo de fosfatos. Representación gráfica de los dominios funcionales de un sistema de relevo de fosfatos y la química de fosfotransferencia. Figura modificada de Stock AM. y col. 2000.

#### d) Regulación genética de la fijación de nitrógeno

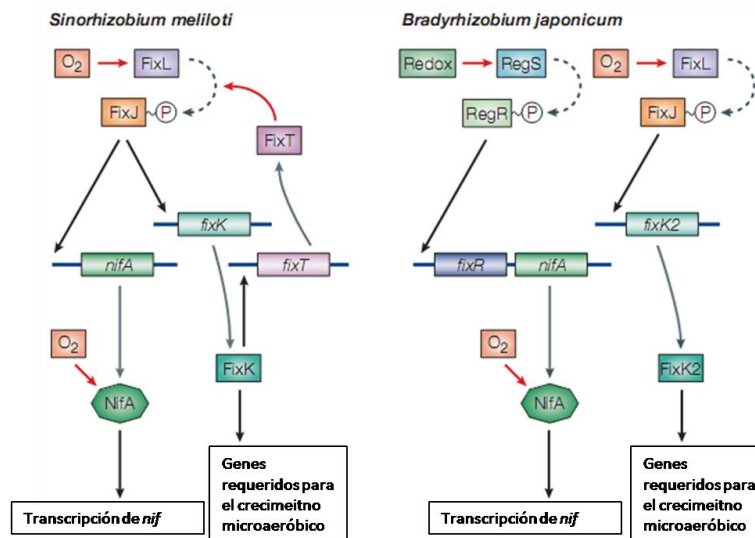
Como se mencionó anteriormente, el proceso de fijación de nitrógeno es finamente regulado de manera jerárquica. Las principales moléculas reguladoras son el nitrógeno y el oxígeno. La detección de los niveles de nitrógeno es particularmente determinante para la fijación de nitrógeno en diazótrofos de vida libre como *K. pneumoniae*, mientras que para diazótrofos simbióticos la señal clave la da el oxígeno.

En *K. pneumoniae*, en condiciones de limitación de nitrógeno, GlnB, una proteína tipo PII, es completamente uridilada lo que tiene como consecuencia la activación del sistema de dos componentes NtrB-NtrC. Este sistema de regulación activa la expresión de los genes *nifA*, *nifL*, *glnK* y *amtB*. GlnK codifica para una proteína tipo PII capaz de liberar la interacción entre NifL y NifA. En estado uridilado GlnK afecta su interacción con el transportador de amonio AmtB. Cuando la cantidad extracelular de amonio aumenta, GlnK se desuridila y AmtB la secuestra hacia la membrana y la represión de NifA por NifL se mantiene.

En diazótrofos simbióticos, como *S. meliloti* y *B. japonicum*, la principal señal que regula la fijación de nitrógeno es la concentración de oxígeno. Sin embargo, los mecanismos para regular la respuesta son un tanto diferentes en cada modelo. Los genes importantes para la fijación de nitrógeno son básicamente de dos tipos: los genes *nif* y los genes *fix*. Los genes *nif* codifican para la nitrogenasa y proteínas accesorias; mientras que los genes *fix* en su mayoría codifican para genes que permiten mantener un metabolismo activo en condiciones microaeróbicas. En *S. meliloti* los genes *nif* y los genes *fix* están bajo el control de un solo sistema de dos componentes, el sistema FixL-FixJ. Originalmente los genes *fixL* y *fixJ* se identificaron en *S. meliloti* gracias a su participación en la activación de los operones *fixNOQP*, *fixGHIS* y del gen *nifA* en condiciones de bajo oxígeno (David M. y col. 1988). El operón *fixNOQP* codifica para el complejo citocromo oxidasa de tipo *cbb<sub>3</sub>* el cual tiene un alta afinidad por oxígeno y es indispensable para respirar en el ambiente microaeróbico del nódulo (Preisig O. y col. 1993). Por su parte el operón *fixGHIS* participa en la internalización y metabolismo de cobre. Este metal y el cofactor hemo son fundamentales para la actividad de la oxidasa *cbb<sub>3</sub>*. Adicionalmente se ha sugerido que las proteínas FixGHIS participan en el ensamblaje y/o estabilidad de esta oxidasa (Preisig O. y col. 1993). En *S. meliloti* la regulación de los genes *fix* por parte de las proteínas FixL y FixJ depende del regulador transcripcional FixK (Batut J. y col. 1989; Hertig C. y col. 1989). FixK pertenece a la familia CRP/FNR y es homólogo del regulador Fnr de *Escherichia coli*, el cual es responsable de la activación de genes requeridos para la respiración anaeróbica en presencia de fumarato o nitrato como aceptor final de electrones (Batut J. y col. 1989). El regulador FixK conserva los aminoácidos importantes para la unión a DNA (EXXSR) presentes en el dominio HTH del regulador Fnr, en consecuencia este par de reguladores se unen a sitios muy similares en la región reguladora de sus genes blanco (Green J. y col. 2001). Una peculiaridad del sistema de *S. meliloti* es la presencia de la proteína FixT que regula negativamente la actividad de la cinasa de histidinas FixL. La expresión de este regulador está controlada positivamente por el sistema FixL-FixJ a través de FixK (Foussard M. y col 1997; Garnerone AM. y col. 1999). En *B. japonicum* se describieron dos proteínas homólogas a FixK (FixK<sub>1</sub> y FixK<sub>2</sub>) (Anthamatten D. y col. 1992; Nellen-

Anthamatten D. y col. 1998); sin embargo se demostró que la proteína FixK<sub>2</sub> es el regulador central de la cascada reguladora FixLJ, la cual en *B. japonicum* es responsable de la activación de los genes *fixNOQP*, *fixGHIS*, *rpoN<sub>1</sub>* y genes involucrados en la biosíntesis del grupo hemo y en la respiración anaeróbica en presencia de nitrato (Nellen-Anthamatten D. y col. 1998).

El regulador NifA como se mencionó anteriormente participa en la regulación de los genes *nif* los cuales incluyen al operón *nifHDKE* que codifica para las proteínas estructurales de la nitrogenasa (Fischer HM. 1994). En *S. meliloti* la expresión de *nifA* en condiciones microaeróbicas depende del sistema FixLJ, sin embargo a diferencia de lo que ocurre en la activación de los genes *fixNOQP*, FixJ controla de manera directa la expresión de *nifA* (Agron PG. y col. 1993; Reyrat JM. y col. 1993). Es importante hacer notar que la unión del regulador FixJ a sus blancos *fixK* y *nifA* ocurre en sitios con una secuencia nucleotídica diferente (Ferrières L. y col. 2002). En *B. japonicum* la expresión del gen *nifA* y en consecuencia del resto de los genes *nif* no está bajo el control de la cascada FixLJ como sucede en *S. meliloti* (Anthamatten D. y col. 1991). En esta bacteria el sistema de dos componentes compuesto por las proteínas RegS y RegR es responsable de la activación de este gen (Bauer E. y col. 1998).



**Figura 6.** Cascada de regulación de los genes *fix* y *nif* de *Sinorhizobium meliloti* y *Bradyrhizobium japonicum*.

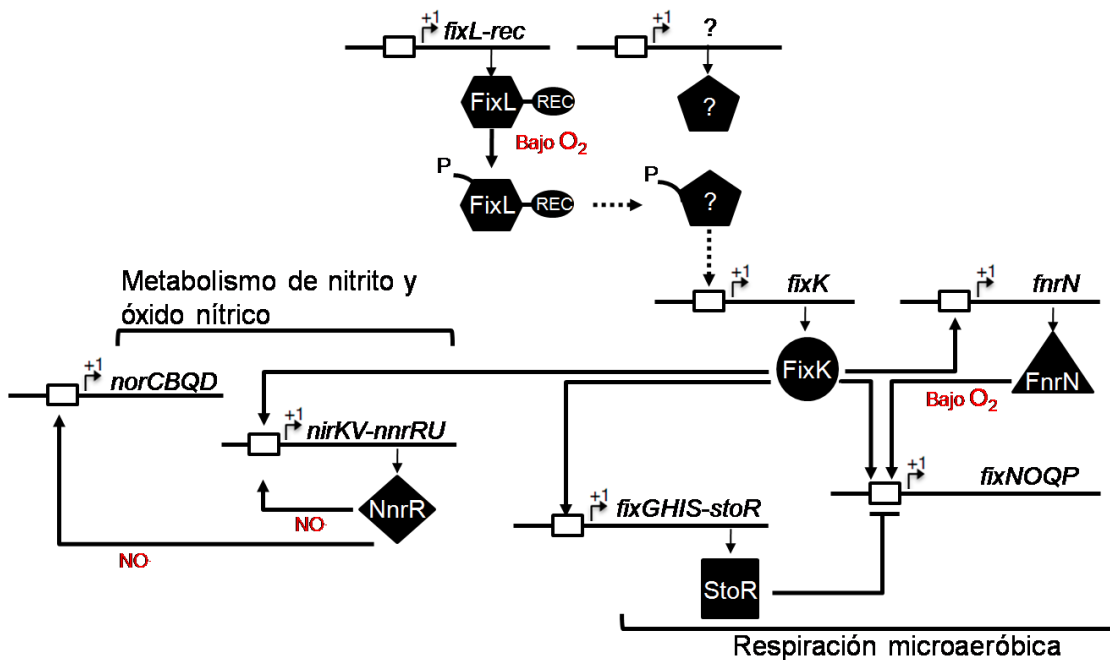


La estructura y la función de la cinasa de histidinas FixL y el regulador de la respuesta FixJ ha sido mayormente estudiada en *S. meliloti* y *B. japonicum*; muchas de las características que confieren la función como detector de oxígeno por parte de la cinasa de histidinas se comparten en estos dos organismos (Gilles-Gonzalez MA. y Gonzalez G. 2005). La actividad de cinasa de histidinas de las proteínas FixL se regula por la unión de oxígeno a un grupo hemo. Este grupo prostético se encuentra coordinado al dominio PAS en el extremo carboxilo terminal de la proteína. La saturación del dominio hemo-PAS con oxígeno ( $K_d= 50-140 \mu\text{M}$ ) inhibe la actividad de cinasa de FixL. Por otro lado en su estado no oxidado la actividad catalítica del dominio cinasa de histidinas permite transferir un grupo  $\gamma$ -fosforilo del ATP al regulador de la respuesta FixJ en un residuo conservado de aspartato. En presencia de oxígeno se ha reportado que FixL tiene actividad de fosfatasa la cual permite mantener a niveles muy bajos la cantidad de FixJ fosforilado en estas condiciones (Lois AF. y col. 1993a). La histidina que forma parte del enlace de coordinación entre el grupo hemo y el dominio PAS (histidina proximal) es el residuo más conservado en proteínas con dominio hemo-PAS como FixL (His<sup>194</sup> en FixL de *S. meliloti* e His<sup>200</sup> en FixL de *B. japonicum*). Otro residuo que se encuentra generalmente conservado en proteínas detectoras de oxígeno con dominios hemo-PAS, es una arginina situada a unos 20 residuos de la histidina proximal. Esta arginina distal interacciona directamente con el oxígeno (Gilles-Gonzalez MA. y Gonzalez G. 2005). El gen *fixL* de *S. meliloti* codifica para una proteína de membrana con cuatro cruces transmembranales (Lois AF. y col. 1993b), mientras que la proteína FixL de *B. japonicum* es una proteína citoplásmica (Gilles-Gonzalez MA. y col. 1994).

e) *Rhizobium etli* y *Rhizobium leguminosarum* bv. *viciae* como modelo de estudio para el análisis de la regulación de la fijación de nitrógeno

*Rhizobium etli* CFN42 es una bacteria del suelo que fija nitrógeno en asociación con plantas del frijol. Su genoma contiene un cromosoma y seis plásmidos (pRet42a a pRet42f) cuyo tamaño va desde 184.4 a 642.5 kb (González V. y col. 2006).

Previamente describimos un circuito de regulación complejo para controlar la expresión de los genes *fix* en *R. etli* CFN42. Este circuito integra varias características novedosas cuando se compara con los que son considerados los circuitos canónicos en *S. meliloti* y *B. japonicum*. Interesantemente elementos reguladores importantes como *fixL* y *fixK* que se encuentran en el plásmido simbiótico de *S. meliloti*, están localizados en el plásmido pRet42f en *R. etli*. En esta bacteria el gen *fixL* codifica para una cinasa de histidinas híbrida, por lo cual de aquí en adelante lo llamaremos *hfixL*. Este gen se localiza en una región genética que contiene copias de los genes *fixK*, *fixNOQP* y *fixGHIS-stoR*. La expresión de los operones *fixNOQP* y *fixGHIS-stoR* se controla por una cascada de regulación en la que participan los genes *hfixL* y *fixKf* y en la cual está ausente un gen *fixJ*. En esta cascada se desconoce la identidad del regulador de la respuesta que en coordinación con hFixL activa la expresión del regulador *fixKf* (Girard L. y col. 1996; Girard L. y col. 2000; González V. y col. 2003; González V. y col. 2006).



**Figura 7.** Cascada de regulación de los genes *fix*, *nir* y *nor* en *Rhizobium etli* CFN42. Las cajas blancas representan el sitio de unión conocido de las proteínas tipo-FNR.

Múltiples reguladores de la familia CRP/FNR participan en la regulación de genes cuya expresión se induce en condiciones microaeróbicas. Debido a la relación

filogenética que guardan las proteínas tipo-FNR, entre ellas y con los miembros de la familia CRP/FNR, se han agrupado en diferentes subfamilias: la subfamilia FixK, la subfamilia FnrN, la subfamilia NnrR y la subfamilia DNR (Körner H. y col. 2003).

Los miembros de la subfamilia FixK carecen de las cuatro cisteínas que coordinan el centro de hierro-azufre, conservadas en los homólogos más cercanos a FNR. Estos reguladores son capaces de unirse a cajas de anaerobiosis. El estado oligomérico de FixK<sub>2</sub> una proteína perteneciente a esta subfamilia en *B. japonicum*, se ha analizado *in vitro*. Esta proteína no posee centros de hierro-azufre y es capaz de activar la transcripción *in vitro* en presencia de oxígeno (Mesa S. y col. 2005).

Las proteínas tipo FnrN poseen 4 cisteínas conservadas en una localización similar a la de la proteína FNR. Se ha demostrado que FnrN de *R. leguminosarum* bv. *viciae* es capaz de activar la transcripción de algunos blancos de FNR en el fondo genético heterólogo de *E. coli* en condiciones anaeróbicas (Schlüter A. y col. 1992). Esto sugiere que este tipo de proteínas son capaces de detectar los niveles de oxígeno, quizás mediante un mecanismo similar al de FNR.

Las proteínas tipo NnrR y DNR controlan la transcripción de genes involucrados en la vía de desnitrificación. Se ha demostrado en diferentes modelos biológicos que esta clase de proteínas es capaz de detectar óxido nítrico cuyo mecanismo de detección se desconoce (Mesa S. y col. 2003; Hartsock A. y Shapleigh JP. 2010; Gómez-Hernández y col. 2011).

No obstante que la similitud a nivel de secuencia entre estas proteínas es baja, los miembros de las diferentes subfamilias son capaces de unirse a secuencias nucleotídicas muy similares y conservan el motivo de unión a DNA de FNR (5'-TTGAT-N<sub>4</sub>-ATCAA-3')

*R. etli* CFN42 posee seis proteínas tipo FNR: dos son de tipo FnrN, una está codificada en el plásmido simbiótico y la otra en el cromosoma y conservan una identidad del 82% (Lopez O. y col. 2001). Otra proteína tipo FNR está codificada en el plásmido pRet42f en orientación divergente al operón *fixNOQPf* y pertenece a la subfamilia FixK (Girard L. y col. 2000); una más es de la subfamilia NnrR y de igual forma se localiza en el plásmido pRet42f (Gómez-Hernández y col. 2011). Por último dos reguladores de la subfamilia DNR llamados StoR ("symbiotic terminal oxidase

regulator”) se localizan en el plásmido simbiótico (pRet42d) y en el plásmido pRet42f. Estos dos reguladores son 82% idénticos y forman parte de los operones *fixGHISd* y *fixGHISf*, respectivamente (Granados-Baeza y col 2007). Una reiteración del gen *fixKf* se encuentra codificada en el plásmido simbiótico (*fixKd*), sin embargo se trata de un pseudogen cuya expresión no se ha podido detectar en ninguna de las condiciones probadas para los genes *fix* (Girard y col. 2000).

En condiciones de vida libre a concentraciones bajas de oxígeno, la expresión de los genes *fix* es dependiente de la cascada *hfixL-fixKf*. La proteína FixKf, cuya expresión depende de la presencia de hFixL, funciona como el regulador primario de los genes *fixNOQPd*, *fixNOQPf*, *fixGHIS-stoRd*, *fixGHIS-stoRf*, *fnrNch* y *fnrNd*. En estas mismas condiciones las proteínas FnrNch y FnrNd también son capaces de activar los operones *fixNOQPd* y *fixNOQPf* aunque no a los mismos niveles a los que lo hace FixKf. Por otro lado las proteínas StoR regulan negativamente la expresión de estos operones. Por lo tanto, FixKf es capaz de regular directamente la expresión de la *cbb<sub>3</sub>* y de reprimirla indirectamente a través de las proteínas StoR.

En *R. leguminosarum* bv. *viciae* la cascada de regulación de los genes *fix* tiene similitudes con la de *R. etli*. En estas bacterias los genes *fix* están altamente conservados. FixL es una cinasa de histidinas híbrida y *fixJ* se encuentra ausente de su genoma. Sin embargo, en *R. leguminosarum* se ha reportado que el gen *fnrN* es el principal regulador de los genes *fixNOQP* reiterados (Patschkowski T. y col. 1996; Schlüter A. y col. 1997).

## Planteamiento del problema

Los sistemas de detección de señales ambientales son piezas fundamentales para entender el comportamiento celular en respuesta a cambios que ocurren en el exterior e interior de las células. En rizobios, el sistema de regulación de los cambios en la concentración de oxígeno que más ha sido estudiado, es el sistema de dos componentes FixL-FixJ. Sin embargo en los últimos años se ha demostrado que este sistema no opera en la totalidad de los rizobios.

En *R. etli* y en *R. leguminosarum* bv. *viciae* la cinasa de histidinas FixL posee un dominio receptor del fosfato y no tiene como regulador de la respuesta cognado a la proteína FixJ. El regulador de la respuesta cognado de FixL debe ser una proteína fundamental para la respuesta a cambios en la concentración de oxígeno en estas bacterias. En *R. etli* hay 67 reguladores de la respuesta codificados en su genoma, distribuidos en el cromosoma y los seis megaplásmidos que forman su genoma. Sabemos que el plásmido pRet42f es fundamental para la fijación de nitrógeno y la expresión de los genes *fix*, en gran medida porque codifica para el gen *fixL*, sin embargo 15 reguladores de la respuesta se encuentran codificados en este megaplásmido y desconocemos si alguno de estos pudiera ser el regulador de la respuesta cognado de esta cinasa de histidinas.

## Objetivos

### Objetivo general:

Identificar y hacer la caracterización inicial del regulador de la respuesta (F<sub>xkR</sub>) que activa la expresión de *fixKf*

### Objetivos particulares:

#### 1) Identificar al activador de *fixKf* (F<sub>xkR</sub>) en microaerobiosis.

- a. Análisis bioinformático de reguladores de la respuesta en el plásmido pRet42f.
- b. Generación de recombinantes sencillas de *R. etli* CFN42 en reguladores de la respuesta codificados en el pRet42f.
- c. Análisis de expresión de la fusión transcripcional del promotor de *fixKf* con el gen monitor *uidA*, en los fondos genéticos silvestre y mutante para cada uno reguladores de la respuesta seleccionados.

#### 2) Caracterizar funcionalmente la proteína F<sub>xkR</sub> de *R. etli* CFN42 y *R. leguminosarum* bv. *viciae*.

- a. Analizar la expresión de *fixKf*, en la mutante F<sub>xkR</sub><sup>-</sup> complementada y en la mutante hFixL<sup>-</sup> que expresa el gen *fxkR* bajo el control de un promotor constitutivo (*plac*).
- b. Estudiar el patrón de expresión de *fxkR* mediado por las proteínas hFixL y F<sub>xkR</sub>.
- c. Identificar al homólogo de *fxkR* en *R. leguminosarum* bv. *viciae* 3841 por métodos bioinformáticos.
- d. Clonar y secuenciar el gen *fxkR* de *R. leguminosarum* bv. *viciae* (3841 y VF39).
- e. Caracterizar la actividad del regulador homólogo mediante ensayos de complementación heteróloga.

#### 3) Identificar proteínas relacionadas a hFixL, F<sub>xkR</sub> y FixKf en diferentes

**bacterias.**

- a. Búsqueda de proteínas con alta similitud a hFixL, FxkR y FixKf.
- b. Análisis del contexto genómico de estos genes.

**4) Caracterizar la región reguladora de *fixKf* y determinar si FxkR es el regulador directo de este gen.**

- a. Análisis bioinformático de la secuencia río arriba de los genes relacionados a *fixKf* para identificar secuencias conservadas.
- b. Determinar la región mínima necesaria para la activación de *fixKf* en condiciones microaeróbicas.
- c. Mutagénesis puntual de los nucleótidos del motivo conservado.
- d. Ensayos de unión *in vitro* de FxkR a la región reguladora de *fixKf*.
- e. Hacer un análisis funcional del motivo conservado en la capacidad de unión del regulador.

**5) Predicción *in silico* de los residuos importantes para la función del regulador FxkR.**

- a. Análisis de la conservación de residuos importantes para la función de reguladores caracterizados de la misma familia .
- b. Modelaje de la interfaz de interacción Proteína-Dna, de la estructura de FxkR unido al DNA usando como templado el cristal de un homólogo.

## Diseño experimental

El diseño experimental de este proyecto se basa principalmente en un enfoque genético, sin embargo también se incluirá evidencia *in vitro* para la caracterización de la región reguladora de *fixKf*.

Con el fin de identificar al regulador de *fixKf* en los megaplásmidos de *R. etli* se evaluará la expresión de una fusión reportera del promotor de este gen con el gen estructural *uidA*, en diferentes fondos genéticos con curaciones de estos replicones.

Para predecir la presencia de reguladores de la respuesta en los diferentes replicones del genoma de *R. etli* se utilizará el servidor web P2CS. Con la finalidad de evaluar el papel de algunos de los reguladores de la respuesta predichos, se analizará la expresión de la fusión reportera *fixKf::uidA* en fondos genéticos mutantes en estos reguladores.

Una vez que se haya identificado al regulador (*fxkR*) se determinará si se requiere la expresión simultánea de *hfixL* para la expresión del gen *fixKf* a los niveles observados en la cepa silvestre. Este análisis se hará cuantificando la expresión de *fixKf* en un fondo genético con una mutación nula en *hfixL* y que contenga *fxkR* bajo un promotor constitutivo. Adicionalmente para determinar si *fxkR* está regulado a nivel transcripcional en condiciones microaeróbicas se analizará la expresión del gen *uidA* bajo el promotor de *fxkR* en diferentes fondos genéticos (*hFixL<sup>-</sup>* y *FxkR<sup>-</sup>*).

Con el objetivo de identificar un posible homólogo a *fxkR* en las cepas VF39 y 3841 de *R. leguminosarum* bv. *viciae* se analizará la secuencia completa de esta última con la herramienta bioinformática BLAST. Una vez identificado el homólogo se diseñarán oligonucleótidos para amplificar al gen tanto en la cepa 3841 como en la VF39 considerando que en esta última es en la que se ha hecho la caracterización de la regulación de los genes *fix* en *R. leguminosarum* bv. *viciae*. Los productos de PCR se clonarán en el vector TOPO\_TA y serán secuenciados. Construiremos plásmidos para ensayos de complementación heteróloga y para determinar si los homólogos de *fxkR* identificados son funcionales.



Por medio de alineamientos tipo BLAST y análisis del contexto genómico determinaremos la presencia de genes que codifican proteínas relacionadas con la cascada hFixL-FxkR en otras bacterias.

Como parte del análisis de la región reguladora de *fixKf* se buscarán motivos conservados entre homólogos de este gen utilizando el servidor web MEME. Una vez que se haya identificado algún motivo conservado, analizaremos su funcionalidad *in vivo*, utilizando fragmentos recortados de la región reguladora clonados en una fusión reportera, con el fin de delimitar la región mínima necesaria para la expresión de *fixKf*. Mediante mutagénesis sitio dirigida determinaremos la importancia del motivo conservado en el reconocimiento del activador de la transcripción de *fixKf*. Por otra parte, se analizará *in vitro* la capacidad del regulador de la respuesta de unirse directamente a la región reguladora de *fixKf* mediante ensayos de retardo.

Para cumplir con el objetivo de predecir residuos importantes para la actividad del nuevo regulador, se buscará una proteína de la misma familia con un porcentaje alto de identidad y cuya función haya sido estudiada a nivel bioquímico. Se hará un análisis comparativo mediante alineamientos múltiples para determinar si los aminoácidos importantes para la función del regulador homólogo están conservados en los reguladores tipo FxkR. Para predecir la interfaz de interacción Proteína-Dna se hará uso del servidor web 3d-footprint. Este programa utiliza un algoritmo que permite predecir los aminoácidos que hacen contacto con el DNA, mediante un análisis comparativo de la conservación de residuos en la región de unión al DNA, de la proteína de interés y la interfaz determinada experimentalmente (a nivel de estructura terciaria) para una serie de proteínas que fueron cristalizadas en presencia de su DNA blanco. Si la estructura secundaria de la proteína de interés es muy similar a aquella con la cual se comparó para predecir las interacciones Proteína-Dna, se puede también modelar su estructura tridimensional utilizando el algoritmo de TF-modeller. El modelo servirá para hacer un gráfico que muestre el tipo de interacción que tienen los aminoácidos del modelo con el DNA encajado en la estructura, utilizando una vez más el programa 3d-footprint.

## Materiales y Métodos

### 1. Cepas bacterianas, plásmidos y condiciones de cultivo.

Las cepas bacterianas y los plásmidos utilizados se enlistan en la siguiente tabla:

**Tabla 1. Cepas bacterianas y plásmidos**

Cepa o plásmido	Características relevantes	Referencia
<b>Rhizobiales</b>		
CE3	Derivada Sm <sup>r</sup> de la CFN42	Noel KD y col. 1984
CFNX89	Derivada de la CE3 pRet42d <sup>r</sup>	Brom S y col. 1992
CFNX186	Derivada de la CE3 pRet42f <sup>r</sup>	Brom S y col. 1992
CFNX218	Derivada de la CE3 pRet42a <sup>r</sup> ,b <sup>r</sup> ,c <sup>r</sup> ,d <sup>r</sup> , e <sub>Δ</sub> ,f <sup>r</sup>	Brom S y col. 2000
CFNX636	Derivada de la CE3, <i>fixL::loxP</i> Sp	Girard L y col. 2000
<i>R. leguminosarum</i> bv. <i>viciae</i> VF39	Silvestre, Sm <sup>r</sup>	Hynes MF y col. 1988
<i>R. leguminosarum</i> bv. <i>viciae</i> 3841	Silvestre	Johnston AW y Beringer JE 1975
<i>S. meliloti</i> 1021	Silvestre	Meade HM y col. 1982
CFN-LG1	Derivada de la CE3, <i>fxkR::loxP</i> Sp	Este trabajo
<b>Escherichia coli</b>		
DH5α	<i>supE44ΔlacU169(φ80lacZΔM15)</i> <i>hsdR17recA1endA1gyrA96 thi-1relA1</i> F – <i>ompT hsdSB(rB– mB–) gal dcm</i>	Sambrook J y col. 1989
BL21 (DE3)	(DE3)	Invitrogen
<b>Plásmidos</b>		
pCR2.1 TOPO	Vector de clonación de productos de PCR, Ap <sup>r</sup> Km <sup>r</sup>	Invitrogen
pK* <i>mobsacB</i>	Vector de clonación suicida, Km <sup>r</sup>	Schäfer A y col. 1994
pRK415	Plásmido de amplio espectro, Tc <sup>r</sup>	Keen NT y col. 1988
pET28a	Vector de expresión, His-Tag, Km <sup>r</sup>	Novagen
pBBMCS53	pBBR1MCS-5 Δ <i>placZ</i> que lleva el gen <i>uidA</i> sin promotor, Gm <sup>r</sup>	Girard L y col. 2000
pRK2073	Plásmido ayudador, Sm/Sp <sup>r</sup>	Better M y Helinski D 1983
pRK2013	Plásmido ayudador, Km <sup>r</sup>	Figurski DH y Helinski DR 1979
pLG1	pBBMCS53::R <i>fixKf</i>	Girard L y col. 2000
pARG1	pK* <i>mobsacB::fxkR-loxP</i> Sp	Este trabajo
p <i>fxkR</i> -Ret	pRK415:: <i>fxkR</i> ( <i>R. etli</i> CE3).	Este trabajo
p <i>fxkR</i> -VF	pRK415:: <i>fxkR</i> ( <i>R. leguminosarum</i> bv. <i>viciae</i> VF39)	Este trabajo
p <i>fixJ</i> -Sm	pRK415:: <i>fixJ</i> ( <i>S. Meliloti</i> )	Este trabajo
pARG2	pBBMCS53::RR <i>fxkR</i>	Este trabajo
pDZS1	pBBMCS53::RRΔ <i>fixKf</i> (Δ K-box)	Este trabajo
pDZS2	pBBMCS53:: RRΔ <i>fixKf</i> (K-box wt).	Este trabajo
pDZS3	pBBMCS53:: RRΔ <i>fixKf</i> (Ka ΔGTT )	Este trabajo
pDZS4	pBBMCS53::RRΔ <i>fixKf</i> (Δ Ka)	Este trabajo
pDZS5	pBBMCS53:: RR <i>fixKf</i> -Ka* (mutación sitio dirigida en subsitio Ka)	Este trabajo
pDZS6	pBBMCS53:: <i>fixKf</i> -Kb* (mutación sitio dirigida en subsitio Kb)	Este trabajo
pARG3	pET28a:: <i>fxkR</i> ( <i>R. etli</i> )	Este trabajo

## 2. Medios de cultivo y antibióticos.

Las cepas de *Escherichia coli* se crecieron a 37°C in medio Luria-Bertani (LB). Las concentraciones de los antibióticos utilizados fueron: carbenicilina (Cb), 100  $\mu\text{g ml}^{-1}$ ; gentamicina (Gm), 15  $\mu\text{g ml}^{-1}$ ; kanamicina (Km) 30  $\mu\text{g ml}^{-1}$ ; espectinomicina (Sp), 100  $\mu\text{g ml}^{-1}$ . Las cepas de *Rhizobium* se crecieron a 30°C en medio rico PY suplementado con  $\text{Ca}^+$  (Noel KD. y col. 1984) o en medio mínimo Y (MMY) con succinato (10 mM) y cloruro de amonio (10 mM) como fuentes de carbono y nitrógeno respectivamente (Bravo A. y Mora J., 1988). Los antibióticos se utilizaron en las siguientes concentraciones: estreptomycin (Sm), 200  $\mu\text{g ml}^{-1}$ ; gentamicina (Gm), 15  $\mu\text{g ml}^{-1}$ ; kanamicina (Km) 30  $\mu\text{g ml}^{-1}$ ; ácido nalidíxico (Nal) 20  $\mu\text{g ml}^{-1}$ ; neomicina (Neo), 60  $\mu\text{g ml}^{-1}$ ; espectinomicina (Sp), 100  $\mu\text{g ml}^{-1}$ . Cuando se requirió se agregó sacarosa al 12.5% (w/v). Para la detección de la actividad de  $\beta$ -galactosidasa y  $\beta$ -glucuronidasa se utilizaron los sustratos 5-bromo-4-cloro-3-indolil- $\beta$ -D-galactosido (X-gal) y 5-bromo-4-cloro-3-indolil- $\beta$ -D-glucuronido (X-gluc) respectivamente a una concentración de 20  $\mu\text{g ml}^{-1}$  en cajas de cultivo con medio sólido.

## 3. Condiciones de crecimiento microaeróbicas.

Los crecimientos en condiciones microaeróbicas de *R. etli*, silvestre y mutantes, se hicieron de acuerdo con el protocolo estandarizado y reportado por nuestro laboratorio (Girard L. y col. 2000). Se inocularon 3 ml de PYCa+, con las cepas recién sembradas. Estos cultivos se incubaron 12 horas a 30°C y 200 rpm. Se toma 1.0 ml de estos precultivos y se inoculan 50 ml de PYCa+. Los cultivos se crecieron hasta fase exponencial media. Las células se colectaron por centrifugación, se lavaron con MMY estéril y se concentraron 100 veces. 20 ml de MMY con succinato y amonio a una concentración final de 10 mM, se inocularon con la suspensión celular a una densidad óptica inicial a 540 nm ( $\text{D.O.}_{540}$ ) de 0.05 en botellas cerradas con tapón aislante. Para crear un ambiente microaeróbico, estas fueron gaseadas con una mezcla oxígeno/argón 1 : 99 v/v. Los cultivos se crecieron con agitación (200 rpm) por 10 horas a 30°C. En los experimentos que incluyen cepas curadas del plásmido pRet42f (CFNX186 o CFNX218), los cultivos se inocularon a una  $\text{D.O.}_{540}$  de 0.25 y las botellas

se incubaron con la mezcla de oxígeno/argón (1:99, v/v) por 10 horas a 30°C para la inducción microaeróbica.

#### **4. Medición de la actividad de $\beta$ -glucuronidasa.**

Los cultivos aeróbicos y microaeróbicos de las cepas de *R. etli* que llevan fusiones transcripcionales se crecieron como se describió previamente (Girard L. y col. 1996). La cuantificación de la actividad de GUS se determinó por un método colorimétrico a partir de muestras de 1.0 ml usando como sustrato 4-nitrofenil  $\beta$ -D-glucuronido el cual al ser hidrolizado por la  $\beta$ -glucuronidasa desarrolla un color amarillo, como se describió previamente (Girard L. y col. 2000). Los datos se normalizaron con respecto a la concentración total de proteína, calculada con el método de Lowry utilizando un segundo juego de muestras de 1.0 ml. La actividad específica está expresada en nanomoles de producto por minuto por miligramos de proteína. El perfil de plásmidos de las cepas utilizadas para los ensayos de actividad GUS, se determinó usando un protocolo modificado de la técnica de Eckhardt (Hynes MF. y McGregor NF., 1990).

#### **5. Amplificación por PCR.**

Los oligonucleótidos específicos para llevar a cabo la Reacción en Cadena de la Polimerasa (PCR) se diseñaron utilizando el programa Oligo 6.0 y fueron sintetizados por la Unidad de Síntesis Química IBt-UNAM. Para las reacciones de amplificación se utilizó una polimerasa de alta fidelidad (Accuprime Taq DNA Polymerase High Fidelity de Invitrogen) con programas de síntesis de ciclos de temperatura que incluyen un paso de desnaturalización a 94°C por 1 min seguido de 30 ciclos de 94°C por 30 seg, 45-60°C (de acuerdo con la temperatura de “melting” de los oligonucleótidos) por 30 seg y 72°C por 90 seg y una extensión final a 72°C de 10 minutos.

#### **6. 5' RACE (Amplificación Rápida de Extremos de cDNA)**

Los sitios de inicio de la transcripción de los genes *fixKf* y *fixMf* se identificaron mediante ensayos de 5' RACE (Rapid amplification of cDNA ends 2.0 de Invitrogen) como se describió previamente (Gómez-Hernández N. y col. 2011). El scDNA se

sintetizó a partir de RNA purificado de cultivos microaeróbicos de cepas derivadas de la silvestre que se conjugaron con fusiones transcripcionales de *fixKf* y *fixNf* en el plásmido pBBMCS53 (pLG1 y pLG2 respectivamente) (Girard L. y col. 2000), utilizando el oligo complementario gusLw (5'-ACAGGACGTAACATAAGGGACT-3'). Después de la síntesis de la primera hebra de cDNA, el mRNA se eliminó mediante el tratamiento con RNAsa y el scDNA se purificó usando una columna S. N. A. P. (Invitrogen). Las colas homopoliméricas al extremo 3' del cDNA se añadieron usando TdT (Terminal deoxynucleotidyl transferase) y dCTP. Posteriormente se hizo un PCR anidado usando el oligo de anclaje que provee el kit (AAP, GGCCACGCGTCGACTAGTACGGG||GGG||GGG||G) y los oligonucleótidos específicos para el gen de interés *fixKf*-RT (5'-GGCGAAATGAAATGCGGTTAT-3', en las posición 214 a 234) o *fixNf*-RT (5'-CGCTGCACCACGTAGAAC-3', en la posición 403 a 420). Los productos de amplificación fueron clonados en el vector pCR2.1 TOPO con extremos cohesivos T-A y secuenciados.

## 7. Construcción de Plásmidos.

Para introducir una mutación en el gen *fxkR*, se utilizó un producto de amplificación de PCR obtenido a partir de DNA total de la cepa CE3 con los oligonucleótidos específicos UpDR530 y LwDR530. Este producto de amplificación se clonó en el vector pCR 2.1-TOPO. Para construir un plásmido suicida útil para intercambiar el alelo silvestre de *fxkR* por el alelo mutado *fxkR::loxP*Sp (pARG-1), el fragmento de 1517 pb *BglII-HindIII* que contiene el gen *fxkR* se clonó en sitios *BamHI-HindIII* del vector pK<sup>\*</sup>mobsacB (Schäfer A. y col. 1994). El gen *fxkR* se interrumpió clonando en el sitio *SaII* (nucleótido 274) un interposón *loxP*Sp (Martínez-Salazar JM. y Romero D., 2000). Los plásmidos utilizados para la mutagénesis por inserción (mutación por recombinación sencilla) se construyeron clonando un producto específico de amplificación de la región intragénica del gen de interés en el vector pK<sup>\*</sup>mobsacB, el cual es un plásmido suicida en bacterias del género *Rhizobium*. Las regiones intragénicas se amplificaron por PCR usando DNA total de la cepa CE3 y oligonucleótidos específicos (Las regiones intragénicas y la secuencia de los oligonucleótidos se muestran en la Tabla 2). Estos productos de amplificación se

clonaron en el vector pCR2.1-TOPO y se analizaron por secuenciación. Para construir el plásmido suicida, las regiones intragénicas se clonaron usando sitios *EcoRI*, con la excepción de los fragmentos clonados en los plásmidos pPF530 y pPF72 que contienen un sitio *EcoRI* interno.

**Tabla 2. Oligonucleótidos usados para la construcción de los plásmidos suicidas utilizados para la mutagénesis por inserción.**

Oligo / Posición	Secuencia (5' → 3')	Fragmento intragénico utilizado para clonar en el vector pK*mobsacB
UpPF72 / 238 - 258	AAGCGCAATATTCCTTCTTC	540 pb del nucleótido
LwPF72 / 949 - 969	GAGCTTCTCGCCCTCAAGGAC	238 al 773
UpPF224 / 71 - 90	TCTGGAATCGGCGGGCTATG	244 pb del nucleótido
LwPF224 / 296 - 315	GAATGGCTTGC GGAAGAATC	71 al 315
UpPF225 / 10 - 29	GACCAGCCGTTTCGTTCTAG	560 pb del nucleótido
LwPF225 / 551 - 570	GCTGATGCCAAGGTCGAAGG	10 al 570
UpPF479 / 58 - 77	GCTGTCAACGCTGCCCTGCC	600 pb del nucleótido
LwPF479 / 638 - 657	ACGCAAGAAAGCAATCGTGA	58 al 657
UpPF530 / 69 - 88	GACGCGCTATTTTCGAGGATG	437 pb del nucleótido
LwPF530 / 597 - 616	CGATCGTCCGGTCAAAGGCT	69 al 506
UpPF533 / 62 - 81	CACCGGTCGTATGCCTCGTG	558 pb del nucleótido
LwPF533 / 600 - 619	CGGCCATATTCACCAGCTCC	62 al 619
UpPF547 / 33 - 52	GAGCCGGCTGTTCTCCTGTG	182 pb del nucleótido
LwPF547 / 195 - 214	GCGTGCGCTCATCGTCGTAG	33 al 214
UpPF548 / 121 - 140	GAAGCGTTTCGAAAGCACCTC	478 pb del nucleótido
LwPF548 / 579 - 598	CGACAGACCGGACTTCCATC	121 al 598

Los plásmidos de complementación que contienen los genes *fixkR* de *R. etli* CFN42 (*pfixkR*-Ret) y de *R. leguminosarum* bv. *viciae* VF39 (*pfixkR*-VF) así como el plásmido que lleva el gen *fixJ* de *S. meliloti* (*pfixJ*-Sm) expresados bajo el promotor constitutivo de *lacZ* se obtuvieron mediante la clonación de las respectivas regiones codificantes en el plásmido conjugativo pRK415 (Keen NT. y col. 1988). Para construir *pfixkR*-Ret, se

amplificó una región de 981 pb utilizando DNA total de la cepa CE3 y los oligonucleótidos específicos Up530-*Hind*III y Lw530-*Xho*I. Para generar el plásmido p*fxkR*-VF se amplificó una región de 1230 pb utilizando DNA total de la cepa VF39 y los oligonucleótidos específicos UpRL9-26 y LwRL9-26 que se diseñaron utilizando la secuencia del ORF pRL90026 de *R. leguminosarum* bv. *viciae* 3841. El gen *fixJ* de *S. meliloti* 1021 se amplificó usando los oligonucleótidos específicos Up-*fxJ*Sm y Lw-*fxJ*Sm. Para construir el plásmido utilizado para el análisis de expresión del gen *fxkR* de *R. etli*, la región reguladora correspondiente se sintetizó por PCR utilizando los oligonucleótidos Up-530gus y Lw-PF530. El fragmento *Eco*RI de 767 pb (que contiene 262 pb río arriba del codón de inicio de *FxkR*) se clonó en el plásmido de amplio espectro pBBRMCS53 que contiene al gen *uidA* el cual carece de promotor (Girard L. y col. 2000) generando así el plásmido pARG-2. La orientación correcta de los fragmentos clonados en este plásmido se analizó por PCR usando el oligo *gus*Lw y el oligo 5' correspondiente.

Los plásmidos pDZS1, pDZS2, pDZS3 y pDZS4 son variantes recortadas del plásmido pLG1. Cuatro diferentes productos de amplificación se obtuvieron usando como templado el plásmido pLG1 y los oligonucleótidos específicos de la hebra directa Up-NA, Up-GTT, Up-ACA, Up-CAC y el mismo oligonucleótido específico de la hebra complementaria Lw-RRKf. Los productos de amplificación se clonaron en el vector pCR2.1 TOPO. Los fragmentos de 526 pb (que contiene 67 pb río arriba del codón de inicio de *FixKf*), 559 pb (que contiene 100 pb río arriba del codón de inicio de *FixKf*), 556 pb (que contiene 97 pb río arriba del codón de inicio de *FixKf*) y 553 pb (que contiene 67 pb río arriba del codón de inicio de *FixKf*) se clonaron por separado en el vector pBBMCS53 en los sitios de *Xba*I-*Kpn*I para el plásmido pDZS1 y *Eco*RI-*Kpn*I para los plásmidos pDZS2, pDZS3 y pDZS4. Las mutaciones de la probable caja K en la región reguladora de *fixKf* se introdujeron usando el QuickChange Site-Directed Mutagenesis Kit de Stratagene. Para ello se utilizó como templado el plásmido pCR2.1-TOPO que tiene clonada la misma secuencia que el plásmido pLG1. Cada subsitio de la caja se modificó cambiando la secuencia GTTACA por la correspondiente al sitio de restricción de *Eco*RI (GAATTC). Los oligonucleótidos utilizados para mutar el subsitio Ka fueron UpKa-mut y LwKa-mut, mientras que el subsitio Kb se mutó usando

los oligonucleótidos UpKb-mut y LwKb-mut. La región reguladora con mutaciones en el subsitio Ka o en el Kb fue clonada en el plásmido pBBMCS53 para obtener los plásmidos pDZS5 y pDZS6.

Para generar una proteína FxkR recombinante con una etiqueta de 6 histidinas, la región codificante del gen se amplificó usando DNA total de la cepa CE3 como templado, con los oligonucleótidos UpFvkR-NdeI y Lw530-XhoI. Los sitios de restricción introducidos en los oligonucleótidos se utilizaron para clonar el fragmento en el vector pET28a para obtener el plásmido pARG3.

Los oligonucleótidos utilizados en la construcción de los plásmidos reportados en este trabajo se enlistan en la Tabla 3 (los nucleótidos subrayados indican el sitio de restricción).

**Tabla 3. Oligonucleótidos utilizados para la construcción de plásmidos.**

Nombre del Oligo	Secuencia (5'→3')
gusLw	ACAGGACGTAACATAAGGGACT
UpDR530	CCCTCG CCTCTTCAGTTGATCGAA
LwDR530	CTTCCTGCTGATTCAGTCCAGT
Up530- <i>Hind</i> III	<u>AAGCTT</u> CTGGCAAATCTCTGAAAG
Lw530- <i>Xho</i> I	<u>CTCGAG</u> CAAGCGAGCCGAGGG
UpRL9-26	GGCGCTGTAGAGATAGTGGTTCTT
LwRL9-26	TCT <u>GGATCC</u> GAGGCAACTGCTTCA
Up-fxJSm	AGAAGTCGCCGGTCAGCTCTT
Lw-fxJSm	GCAACGAGCCGCCATTAA GCC
UP-530gus	CACCAATGGCGACACCCTTC
Lw-PF530	CGATCGTCCGGTCAAAGGCT
Up-NA	<u>TCTAGA</u> ACCGTCCGCTCATTC
Up-GTT	CCGGAATTCGTTACACACTGTTA
Up-ACA	CCGGAATTCACACACTGTTACAA
Up-CAC	CCGGAATTCCACTGTTACAACT
Lw-RRKf	GGTACCAGAGGCGCTCCTCGC
UpKa-mut	AACCGTCGCT <u>GAATTC</u> CACTGTTACA
LwKa-mut	TGTAACAGTGG <u>GAATTC</u> AGCGACGGTT
UpKb-mut	GTTACACACT <u>GAATTC</u> AACTTTCCACC
LwKb-mut	GGTAAAAGTT <u>GAATTC</u> AGTGTGTAAC
UpFvkR- <i>Nde</i> I	<u>CATATG</u> GCAGCACCTCAGCA
Lw530- <i>Xho</i> I	<u>CTCGAG</u> CAAGCGAGCCGAGGG
Up-RRKf	ACCACCATCGTTTCC
Lw2-RRKf	CATTGTCGAACGCGTGGTTGTTCC

Los sitios de restricción introducidos se encuentran subrayados.



## 8. Derivadas mutantes de *R. etli*

La mutagénesis por inserción se hizo usando a los plásmidos pPF72, pPF224, pPF225, pPF479, pPF530, pPF533, pPF547 y pPF547. Cada plásmido transferido por conjugación a *R. etli*, se co-integró por recombinación sencilla con el blanco de homología, produciendo la mutación del gen de interés (Landeta C. y col. 2011). Las mutantes se seleccionaron como Sm<sup>r</sup> Km<sup>r</sup> y la integración correcta del plásmido se verificó por hibridación de tipo “Southern blot”.

El intercambio del alelo silvestre *fxkR* de la cepa CE3 de *R. etli* por el alelo mutante *fxkR::loxP*Sp del plásmido pARG1 se llevó a cabo por homogenización usando el marcador *sacB* presente en el vector pK\*mobsacB; este marcador permite una fácil selección de eventos de homogenización mediante el análisis de la resistencia a sacarosa. Las derivadas donde ocurrió la doble recombinación se seleccionaron por sus resistencias a sacarosa, Sm Sp y sensibilidad a Km<sup>s</sup>. Para verificar que el reemplazo alélico ocurrió, se analizó el DNA total por la técnica de “Southern Blot” con sondas apropiadas para los genes y para el cassette de resistencia.

## 9. Purificación de His<sub>6</sub>-F<sub>xkR</sub>

El plásmido pARG3 se utilizó para sobreexpresar la proteína recombinante His<sub>6</sub>-F<sub>xkR</sub> en el fondo genético de *E. coli* BL21 (DE3). La sobreexpresión se logró induciendo cultivos crecidos en medio LB (O.D.<sub>540</sub> 0.6) con 0.3 mM de IPTG por dos horas a 30°C. Las células se colectaron por centrifugación a 4°C y se rompieron en amortiguador de lisis nativo (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl pH 8.0) después de dos pases a través de la prensa de French. La separación de la fracción soluble de la insoluble se logró después de centrifugar a 10000xg por 30 minutos a 4°C. Muestras de los cultivos inducidos y no inducidos así como de las fracciones soluble e insoluble, se analizaron por SDS-PAGE. La proteína His<sub>6</sub>-F<sub>xkR</sub> se purificó a partir de 5 ml de fracción soluble usando la agarosa comercial Ni-NTA (Invitrogen) en columnas de plástico (Bio-Rad) siguiendo las instrucciones de los fabricantes. Brevemente, la columna equilibrada por la cual se pasó la fracción soluble, se lavó 3 veces con amortiguador de lavado (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 40 mM Imidazol pH 8.0) y 16 fracciones de 200 µl se colectaron después de adicionar a la columna el amortiguador

de elución (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 500 mM Imidazol pH 8.0). Se tomaron muestras de cada paso de purificación y se analizaron por SDS-PAGE.

## **10. Ensayo de cambio en la movilidad electroforética (“Electroforetic mobility shift analysis” EMSA)**

Una secuencia de 300 pb que comprende la región intergencia de *fixNf-fixKf* se amplificó por PCR usando como templado los plásmidos pLG1, pDZS5 y pDZS6 y los oligonucleótidos Up-RRKf y Lw2-RRKf. Los productos de amplificación se purificaron con el kit “PureLink PCR Purification” (Invitrogen) y se cuantificaron con un espectrofotómetro del tipo Nanodrop. Se utilizaron 100 ng de DNA en la reacción de marcaje con 3 µl de [γ-<sup>32</sup>P]-ATP (6000 Ci/mmol) en un volumen final de 20 µl. La marca que no fue incorporada se eliminó utilizando columnas “illustra Microspin G-25” (Amersham) siguiendo las instrucciones del fabricante.

Para las reacciones de unión, el DNA marcado (1 ng) se incubó en presencia de concentraciones crecientes de la proteína recombinante pura (de 0.3 a 1.4 nM) en un amortiguador de unión que contiene 100 mM Tris-HCl (pH 7.4), 100 mM KCl, 10 mM MgCl<sub>2</sub>, 10% glicerol, 2 mM dithiothreitol y 1 mM de acetil fosfato (SIGMA), en un volumen final de 20 µl por 20 minutos a temperatura ambiente. Cuando fue necesario se utilizaron 100 ng de DNA de esperma de salmón como competidor inespecífico; mientras que se utilizaron 100 ng de DNA sin marcar como competidor específico. La electroforésis se llevó a cabo en geles de poliacrilamida (79:1) al 5% en amortiguador 0.5X TBE por una hora a 300 V a temperatura ambiente.

## **11. Recursos bioinformáticos**

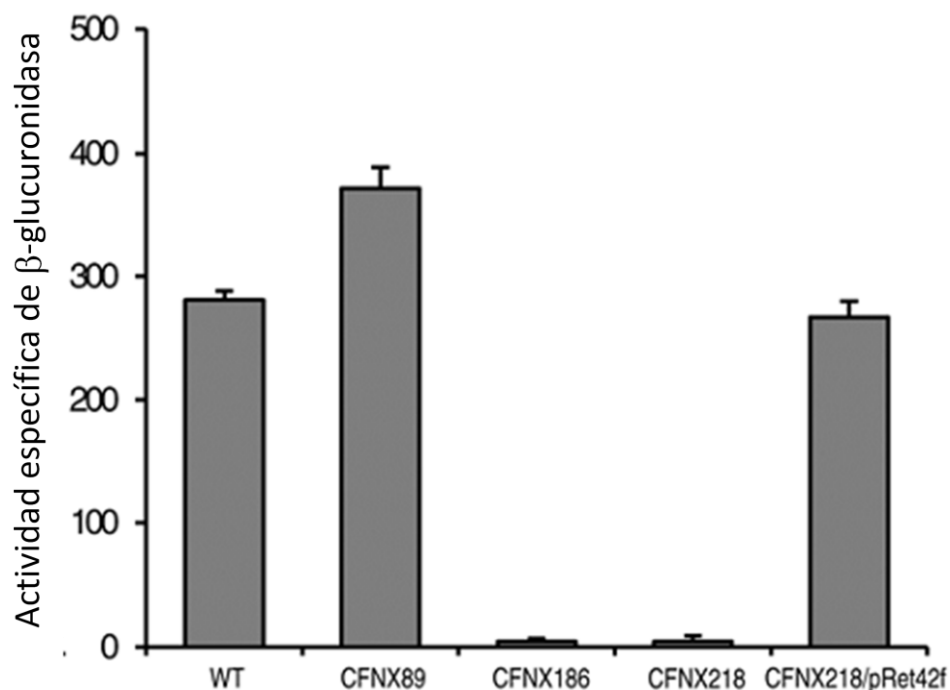
Las proteínas similares a FxkR y hFixL en diferentes organismos se identificaron con BLASTP. El contexto genómico de los genes *fxkR* se analizó con el visor de información del gen de la página web NCBI. El alineamiento de secuencias para el análisis de residuos conservados y el porcentaje de identidad se realizó con los parámetros por omisión del servidor web ClustalW2 de las herramientas bioinformáticas de EMBL-EBI (Larkin MA. y col. 2007). La conservación de dominios se analizó usando

diferentes librerías como INTERPRO, PFAM, PROSITE, SMART y SUPERFAMILY (Finn RD. y col. 2010; Gasteiger E. y col. 2003; Letunic I. y col. 2009; Sigrist CJ. y col. 2010). Para identificar motivos conservados en la región reguladora de los genes tipo *fixKf* se utilizó el servidor web MEME con los parámetros por omisión (Bailey TL. y Elkan C., 1994). La predicción de estructura secundaria se hizo con el método de predicción PSIPRED (Jones DT. 1999) a través del servidor web PRALINE (Simossis V.A. and Heringa J. 2005). Para modelar la estructura cuaternaria de FxkR se utilizó como templado la estructura cristalográfica de las cadenas A y B del dominio efector de PhoB unido a la caja pho (1GXP) (Blanco AG. y col. 2002). Las coordenadas atómicas de este complejo se obtuvieron del “Protein Data Bank” (PDB) en un archivo con terminación pdb. Las cadenas A y B de los dominios efectores de PhoB y FxkR se alinearon con ClustalW y se generó un archivo con terminación fas. La secuencia de las cadenas A y B del dominio efector de FxkR, así como el archivo del alineamiento (.fas) y del cristal de PhoB-pho (.pdb) se utilizaron para generar el modelo cuaternario de FxkR-pho utilizando el servidor web TFmodeller (Contreras-Moreira B. et al. 2007). Con base en información obtenida del modelaje tridimensional del regulador FxkR de *R. etli* se hizo una predicción de la probable interfaz de reconocimiento Proteína-DNA. La caja pho se intercambió por la caja K para generar los modelos 1GXP-Kab y FxkR-Kab. La modificación de la caja pho se logró usando un “script” de python escrito por Bruno Contreras Moreira. El servidor web 3d-footprint se utilizó para analizar la probable interfaz de reconocimiento (Contreras-Moreira B. et al. 2010). Mediante la aplicación “footprint your structure” el servidor predice los contactos entre los aminoácidos y los nucleótidos del complejo y una matriz de posición y peso del probable sitio de unión a DNA. Se utilizó la misma aplicación para hacer la estimación de la especificidad de unión usando la opción “relax hydrophobics” y “sample rotamers” con los parámetros por omisión. El Logo de la estimación de la especificidad de unión se generó con la combinación de las matrices de posición y peso (PWM) “readout” y “contact”. Las mutaciones *in silico* de los aminoácidos de la interfaz Proteína-Dna se generaron con la herramienta MUTATE del visor de archivos .pdb “Swiss-PdbViewer Deep View”.

## Resultados

### **Análisis de los elementos genéticos necesarios para la expresión de *fixKf*.**

El análisis de la secuencia completa del genoma de *R. etli* CFN42 reveló que en esta bacteria no hay proteínas con un alto porcentaje de identidad al regulador de la respuesta FixJ de *Sinorhizobium meliloti* o de *Bradyrhizobium japonicum*. Dado que los plásmidos pRet42d y pRet42f codifican a la mayoría de los genes *fix* y sus reguladores conocidos (Girard L. y col. 1996; Girard L. y col. 2000; González V. y col. 2003; González V. y col. 2006) decidimos evaluar la participación de estos megaplásmidos en el control de la expresión de *fixKf*. El plásmido pLG1 que contiene la fusión transcripcional *fixKf::uidA* (Girard L. y col. 2000) se introdujo en diferentes derivadas de la cepa CE3 con curaciones en al menos uno de los megaplásmidos nativos de *R. etli* CFN42. Los resultados en la Figura 8 muestran que la presencia del pRet42f es indispensable para la expresión de *fixKf* ya que el promotor de este gen se expresa en todos los fondos genéticos probados a excepción de aquellos en los que el plásmido pRet42f se curó como en el caso de la cepa CFNX186, una derivada de la cepa CE3 curada del pRet42f y la cepa CFNX218 que carece de todos los megaplásmidos a excepción de un producto de eliminación parcial de pRet42e que no se puede curar por ser indispensable para la viabilidad de la cepa (pRet42e $\Delta$ ) (Brom S. y col. 1992). Dado que el gen *hfixL*, el cual es fundamental para la expresión de *fixKf* se encuentra codificado en pRet42f, este resultado era el esperado. Sin embargo, en una derivada de la cepa CE3 con solo el plásmido pRet42f y el (pRet42e $\Delta$ ) (CFNX218/pRet42f), la expresión de *fixKf* ocurre a los mismos niveles que en la cepa silvestre (Figura 8). Por lo anterior podemos concluir que el pRet42f es el único plásmido indispensable para la expresión de *fixKf*. El ligero incremento en la expresión de *fixKf* observado en la cepa derivada que carece del plásmido simbiótico (CFNX89) podría explicarse por la pérdida del represor StoRd el cual se ha reportado que modula negativamente la expresión de este gen (Granados-Baeza M. y col. 2007).

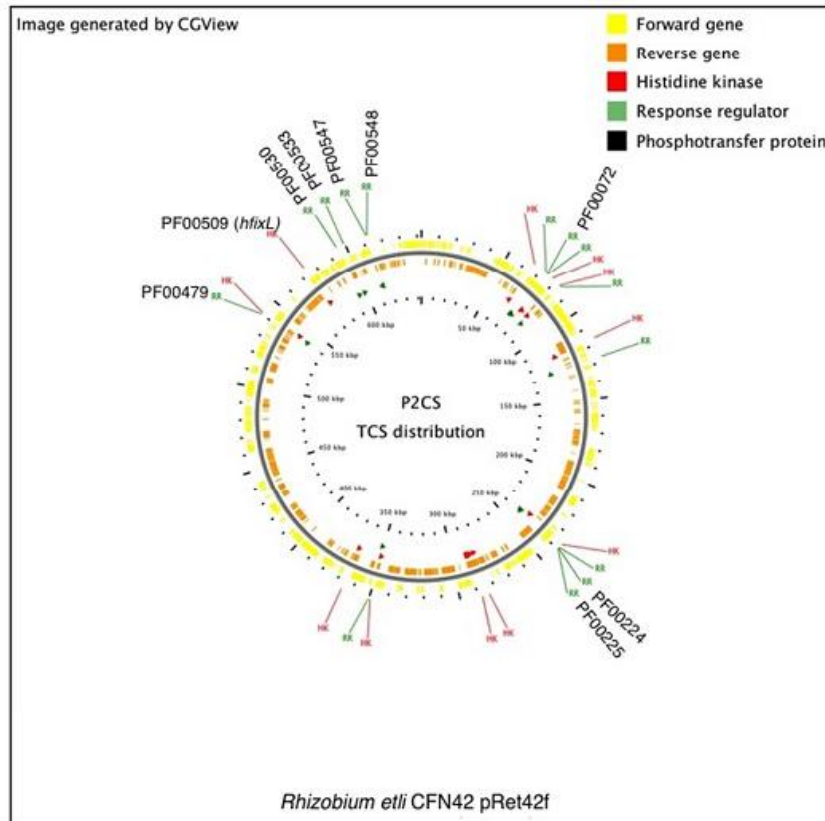


**Figura 8.** Participación de los plásmidos pRet42d y pRet42f en la regulación de la expresión microaeróbica de *fixKf*. La expresión de *fixKf* se analizó en bajas concentraciones de oxígeno (1%) en diferentes derivadas: CE3 (silvestre-WT), CFNX89 (curada del plásmido Ret42d), CFNX186 (curada del plásmido Ret42f), CFNX218 (que solo contiene el pRet42e $\Delta$ ) y la CFNX218/pRet42f. Los cultivos se indujeron en microaerobiósis por 10 horas. La actividad específica de  $\beta$ -glucuronidasa se expresó como nmol min<sup>-1</sup> mg<sup>-1</sup> de proteína. Los datos corresponden al promedio de dos réplicas de tres experimentos independientes.

### Identificación de FxkR como el regulador de la respuesta cognado de la cinasa de histidinas hFixL en *R. etli*.

Para predecir la presencia de reguladores de la respuesta en el genoma de *R. etli* utilizamos la base de datos P2CS (Barakat M. y col. 2011). De acuerdo con esta base de datos, se predijeron un total de 68 marcos abiertos de lectura que codifican para reguladores de la respuesta en el genoma de esta bacteria: de ellos, 23 están codificados en diferentes megaplásmidos: 14 en el pRet42f; mientras que solo 2 en el plásmido pRet42e $\Delta$ . El arreglo genómico de estos probables reguladores de la respuesta en el pRet42f se ilustra en la Figura 9. El análisis de similitud con FixJ de *S.*

*meliloti* mostró que las proteínas RHE\_PF00225 y RHE\_PF00548 son las que tienen el mayor porcentaje de identidad con esta proteína, 40% y 37%, respectivamente.



**Figura 9.** Arreglo genómico de los probables reguladores de la respuesta en pRet42f. La posición de los 8 reguladores de la respuesta que se analizaron en este trabajo se indica con el identificador del gen (“Gene ID”).

Dado que en *R. etli* CFN42 la mayoría de las proteínas reguladoras de los genes *fix* se encuentran codificadas en pRet42f, hipotetizamos que el regulador cognado de hFixL podría estar codificado en este replicón. Mediante una estrategia de recombinación sencilla generamos mutantes nulas en los dos reguladores más conservados a nivel de secuencia con respecto a FixJ de *S. meliloti* y en otros seis reguladores de la respuesta predichos en la secuencia de pRet42f. Los genes RHE\_PF00530 y RHE\_PF00533, que pertenecen a la familia OmpR/PhoB y NarL/FixJ

respectivamente, los cuales no se encuentran en operón con una cinasa de histidinas. Por otro lado los genes RHE\_PF00225, RHE\_PF00479 y RHE\_PF00548, que pertenecen a la familia NarL/FixJ parecen ser parte de un operón que incluye una cinasa de histidinas. Mientras que los genes RHE\_PF00072, RHE\_PF00547 y RHE\_PF00224 no tienen un sitio de unión a DNA identificable por la base de datos SMART (Letunic I. y col. 2009), y el primero de ellos posee un dominio de unión de nucleótidos cíclicos de monofosfato (cNMP). Estos genes también parecen formar parte de un probable operón que incluye a una cinasa de histidinas (Rhizobase Database).

Estudiamos la expresión microaeróbica de *fixKf* en la cepa silvestre y en los fondos genéticos con mutaciones nulas en los reguladores de la respuesta mencionados anteriormente (SR72/pLG1, SR24/pLG1, SR225/pLG1, SR479/pLG1, SR530/pLG1, SR533/pLG1, SR547/pLG1 y SR548/pLG1). Nuestros resultados demuestran claramente que la expresión microaeróbica de *fixKf* es completamente dependiente de una proteína codificada por el marco abierto de lectura RHE\_PF00530 (Tabla 4). Ninguna de las mutaciones en cualquiera de los otros reguladores de la respuesta, incluyendo a las proteínas tipo-FixJ, tuvo un efecto en la expresión de *fixKf*.

**Tabla 4. Identificación del regulador de la respuesta que controla la expresión microaeróbica de *fixKf* en *R. etli* CFN42.**

Cepa	Gen Mutado	Arquitectura de dominios y organización genética	Expresión microaeróbica de <i>fixKf</i> <sup>1</sup>
CE3		Silvestre	381 ± 12
SR 72	RHE_PF00072	Sin dominio HTH. Con un dominio cNMP en el extremo N-terminal. Huérfano.	330 ± 15
SR 224	RHE_PF00224	Sin dominio HTH. Probablemente en operón con una HK y dos RR.	292 ± 22
SR 225	RHE_PF00225	De la familia NarL/FixJ. Probablemente en operón con RHE_PF00224.	282 ± 17
SR 479	RHE_PF00479	De la familia NarL/FixJ. Probablemente en operón con RHE_PF00480 (HK).	289 ± 18

SR 530	RHE_PF00530	De la familia OmpR/PhoB. Huérfano.	4 ± 4
SR 533	RHE_PF00533	De la familia NarL/FixJ. Huérfano.	315 ± 17
SR 547	RHE_PF00547	Sin dominio HTH. Probablemente en operón con RHE_PF00548.	299 ± 24
SR 548	RHE_PF00548	De la familia NarL/FixJ. Probablemente en operón con RHE_PF00547.	302 ± 17

<sup>1</sup>La actividad específica de  $\beta$ -glucuronidasa ( $\text{nmol min}^{-1} \text{mg protein}^{-1}$ ) se determinó después de 10 horas de incubación en condiciones microaeróbicas a 30°C. Los datos son el promedio de dos réplicas y tres experimentos independientes.

El gene RHE\_PF00530 codifica un polipéptido de 241 aminoácidos y está localizado a 19.5 kb río abajo del gen *hfixL*. Este gen se nombró *fxkR* por “*fixK* regulator”. Interesantemente este regulador no pertenece a la familia NarL/FixJ si no a la familia OmpR/PhoB.

Para determinar si la actividad de FxkR depende de la presencia de hFixL, se estudió la expresión microaeróbica de *fixKf* en el fondo genético silvestre y en el de las mutantes CFNX636 (*fixL::loxSp*) y CFN-LG1 (*fxkR::loxSp*) (Tabla 5). Para ello se conjugó en estas cepas el plásmido *pfxkR-Ret*, el cual expresa *fxkR* bajo el promotor del operón *lac*. Nuestros resultados muestran que la expresión microaeróbica de *fixKf* depende completamente de la presencia de hFixL y de FxkR. El plásmido *pfxkR-Ret* complementa el fenotipo de la cepa CFN-LG1, obteniéndose un nivel de expresión de *fixKf* de prácticamente el doble del nivel observado en la cepa silvestre ( $697 \pm 70$  vs  $367 \pm 20$ ). Sin embargo, en la cepa CFNX636/*pfxkR-Ret*, el nivel de expresión de *fixKf* es apenas del 16% del observado en la cepa silvestre ( $60 \pm 12$  vs  $367 \pm 20$ ). Este resultado demuestra claramente que para que FxkR active la expresión de *fixKf* es necesaria la presencia de hFixL y sugiere fuertemente que la dependencia por hFixL no es a través del control de la transcripción de *fxkR*.

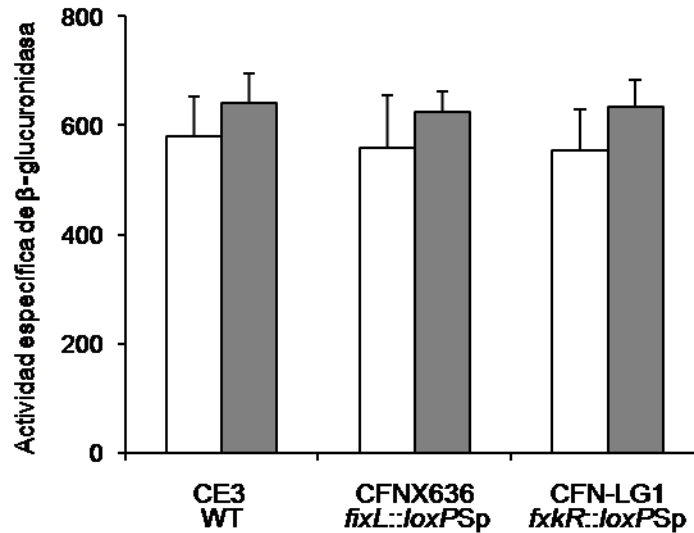


**Tabla 5. hFixL y FxkR son necesarios para la expresión microaeróbica de *fixKf* de *R. etli*.**

Cepa	Genotipo	Plásmidos	Nivel de Expresión <sup>1</sup>
CE3/pLG1	Wild type	pBBMCS53:: <i>fixKf</i>	367 ± 20
CFNX636/pLG1	<i>fixL::loxP</i> Sp	pBBMCS53:: <i>fixKf</i>	2 ± 2
CFN-LG1/pLG1	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i>	1 ± 4
CFNX636/pLG1/ pfxkR-Ret	<i>fixL::loxP</i> Sp	pBBMCS53:: <i>fixKf</i> pRK415:: <i>fxkR-Ret</i>	60 ± 12
CFN-LG1/pLG1/ pfxkR-Ret	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i> pRK415:: <i>fxkR-Ret</i>	697 ± 70
CFN-LG1/pLG1/ pRK415	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i> pRK415	2 ± 2

<sup>1</sup> La actividad específica de β-glucuronidasa (nmol min<sup>-1</sup> mg protein<sup>-1</sup>) se determinó después de 10 horas de incubación en condiciones microaeróbicas a 30°C. Los datos son el promedio de dos réplicas y tres experimentos independientes.

Para corroborar lo anterior, se analizó la expresión del gen *uidA* codificado en el plásmido pARG2, que se encuentra bajo el control del promotor de *fxkR*. La regulación transcripcional del promotor de *fxkR* se estudió en las cepas CE3 (silvestre), CFNX636 (*fixL::loxSp*) y CFN-LG1 (*fxkR::loxSp*), en condiciones aeróbicas y microaeróbicas. El patrón de expresión de *fxkR* muestra que este gen se expresa de manera constitutiva independientemente de la concentración de oxígeno en el cultivo y que el gen no está sujeto a auto-regulación (Figura 10). Considerando que la expresión de *fixKf* es altamente inducida en condiciones microaeróbicas (25 veces) (Girard L. y col. 2000) podemos proponer que hFixL activa la expresión de FxkR post-traduccionalmente en esta condición.



**Figura 10.** Análisis de la expresión de *fvkR* de *Rhizobium etli* en condiciones de vida libre. Las células se crecieron aeróbicamente (barras blancas) o microaeróbicamente (barras grises) por 10 horas. La actividad específica de  $\beta$ -glucuronidasa está expresada como  $\text{nmol min}^{-1} \text{mg protein}^{-1}$ . Los datos son el promedio de tres experimentos independientes por duplicado cada uno.

### El plásmido pRL9 de *R. leguminosarum* bv. *viciae* 3841 posee un gen ortólogo a *fvkR*.

Como se mencionó anteriormente *R. etli* CFN42 y *R. leguminosarum* bv. *viciae* poseen una hFixL y carecen de un gen *fixJ* anotado (Patschkowski T. y col. 1996; Girard L y col. 2000). Dadas estas similitudes consideramos altamente probable la existencia de un gen ortólogo a *fvkR* en *R. leguminosarum* bv. *viciae*. El análisis de la secuencia completa de la cepa 3841 por blastp (NCBI GenBank Database) reveló que el gen pRL90026, codifica para un polipéptido de 220 aminoácidos con un porcentaje de identidad del 88% con respecto a FvkR de *R. etli*. Por “Bi-directional Best-Hit” pudimos determinar que ambas proteínas son ortólogas. En virtud de que en la cepa VF39 se describió el mecanismo de regulación de los genes *fix* de *R. leguminosarum* bv. *viciae* pero no ha sido secuenciada, nos basamos en la secuencia del gen pRL90026 para diseñar oligonucleótidos para generar un fragmento por PCR tanto de la cepa 3841 como de la cepa VF39. El análisis de la secuencia completa del producto de PCR de la cepa VF39 reveló un marco de lectura abierto completo con una alta

identidad a pRL90026 y RHE\_PF00530 (98% y 88% respectivamente). Para determinar si el ortólogo presente en la cepa VF39 es funcional en el fondo genético de *R. etli*, se realizó un experimento de complementación heteróloga, utilizando el plásmido pfxkR-VF el cual expresa el gen *fxkR*-VF bajo el promotor del operón *lac*. El plásmido pfxkR-VF se transfirió por conjugación a la cepa CFN-LG1 y se evaluó la expresión de la fusión *fixKf::uidA* comparando con los niveles de expresión en la cepa silvestre y en la cepa no complementada. Como se observa en la Tabla 6 el regulador heterólogo FxkR-VF complementa a la cepa mutante en FxkR-Ret (CFN-LG1).

**Tabla 6. *R. leguminosarum* bv. *viciae* tiene un ortólogo funcional de *fxkR* de *R. etli*.**

Cepa	Genotipo	Plásmidos	Nivel de Expresión <sup>1</sup>
CE3/pLG1	Wild type	pBBMCS53:: <i>fixKf</i>	367 ± 20
CFN-LG1/pLG1	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i>	1 ± 4
CFN-LG1/pLG1/ pfxkR-Ret	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i> pRK415:: <i>fxkR-Ret</i>	697 ± 70
CFN-LG1/pLG1/ pfxkR-VF	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i> pRK415:: <i>fxkR-Rleg</i>	510 ± 71
CFN-LG1/pLG1/ pfixJ-Sm	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i> pRK415:: <i>fixJ-Smel</i>	6 ± 5
CFN-LG1/pLG1/ pRK415	<i>fxkR::loxP</i> Sp	pBBMCS53:: <i>fixKf</i> pRK415	2 ± 2

<sup>1</sup> La actividad específica de  $\beta$ -glucuronidasa ( $\text{nmol min}^{-1} \text{mg protein}^{-1}$ ) se determinó después de 10 horas de incubación en condiciones microaeróbicas a 30°C. Los datos son el promedio de dos réplicas y tres experimentos independientes.

Esto sugiere que ambos reguladores podrían cumplir funciones similares en sus propios fondos genéticos. Como se mencionó anteriormente, en *S. meliloti* y en *B. japonicum*, FixJ es el regulador de la respuesta cognado de FixL que activa la expresión de *fixK*. Sin embargo, el análisis de la región reguladora de *fixKf* muestra que el sitio consenso de unión de FixJ (Ferrière L. y Kahn D., 2002) no está conservado en

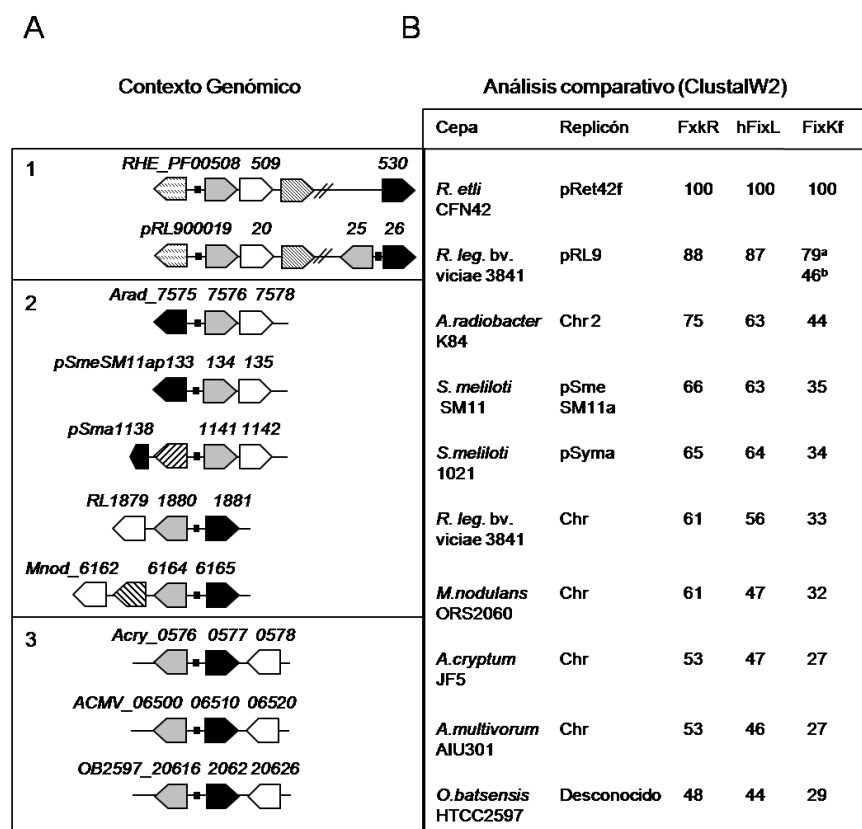
este gen (datos no mostrados). Por lo tanto, no es de sorprender que *fixJ* de *S. meliloti* 1021 sea incapaz de sustituir al gen *fxkR* en *R. etli* CFN42 (Tabla 6).

### **Identificación de proteínas tipo-FxkR en las $\alpha$ -proteobacterias**

Para identificar la presencia de proteínas tipo FxkR y hFixL en otras bacterias, se hizo una búsqueda con blastp usando las proteínas de *R. etli* CFN42 como referencia. Encontramos que las proteínas con un mayor porcentaje de identidad a FxkR y hFixL de *R. etli* CFN42 pertenecen a bacterias del grupo de los *Rhizobiales* (Anexos I y II). Continuamos nuestro análisis con aquellas bacterias que tuvieran tanto una proteína hFixL como una proteína FxkR altamente conservadas con respecto a las de *R. etli* CFN42 y que sus genomas estuvieran completamente secuenciados. Las proteínas hFixL se seleccionaron con base en la conservación de un dominio receptor del fosfato (Rec1) en el extremo C-terminal y la presencia de los residuos His<sup>200</sup> y Arg<sup>220</sup> los cuales son importantes para la unión de oxígeno en FixL de *B. japonicum* (Dunham CM. y col. 2003; Gong W. y col. 1998). Las proteínas tipo-FxkR se seleccionaron considerando una calificación de alineamiento  $\geq 48$  definido por ClustalW2, la presencia de aminoácidos clave para la función del regulador OmpR y su contexto genómico. Este análisis reveló la presencia de proteínas tipo-hFixL y tipo-FxkR en los genomas de *Rhizobium rhizogenes* K84 (antes *Agrobacterium radiobacter* K84), *Sinorhizobium meliloti* 1021, *Sinorhizobium meliloti* SM11, *Methylobacterium nodulans* ORS 2060, *Acidiphilium cryptum* JF-5, *Acidiphilium multivorum* AIU301 y *Oceanicola batsensis* HTCC2597. Interesantemente, este análisis reveló la presencia de genes homólogos a *hfixL* y *fxkR* en el cromosoma de *R. leguminosarum* bv. *viciae* 3841.

Como es característico de las proteínas tipo-hFixL de *R. etli* CFN42, *R. leguminosarum* bv. *viciae* VF39 y *S. meliloti* 1021 (Boesten B. y Priefer UB. 2004), el resto de las proteínas tipo-hFixL encontradas poseen una caja H no ortodoxa con el motivo HDFNLL (Apéndice II). El análisis del contexto genómico en el que se encuentran los genes tipo-*hfixL* y *fxkR* en los genomas analizados reveló la presencia de un gen que codifica para una proteína que forma parte de la familia de reguladores CRP/FNR. En la mayoría de los casos esta proteína tiene un alto porcentaje de identidad cuando se compara con FixKf de *R. etli* CFN42. El análisis del contexto

genómico y el análisis comparativo con respecto a las proteínas FxkR, hFixL y FixKf de *R. etli* CFN42 se presenta en la Figura 11.



**Figura 11.** Representación esquemática del contexto genómico y análisis comparativo de genes tipo-*fixKf*, *hfixL* y *fxkR* de  $\alpha$ -proteobacteria. (A) Representación esquemática del contexto genómico. Las cajas 1, 2 y 3 corresponden a diferentes tipos de arreglos identificados. Los arreglos similares se encuentran en la misma caja. Los genes estructurales están representados por flechas que indican la direccionalidad de la transcripción. Los genes tipo-*fixKf* están en color gris, los genes tipo-*hfixL* en blanco y los tipo-*fxkR* en negro. La línea continua denota que los genes están codificados de forma contigua, cuando se rompe la continuidad de la línea se indica que los genes se encuentran separados. Los genes que se encuentran dentro de la región *hfixL*-*fixKf* están representados con un patrón diferente que representa su identidad (genes con el mismo patrón pertenecen a la misma familia proteica). La probable caja-K identificada río arriba de los genes estructurales tipo-*fixK* se indica con cuadros negros. (B) Calificación para las proteínas tipo-FxkR, hFixL and FixKf comparadas con las de *R. etli* CFN42 obtenida por ClustalW2. <sup>a</sup> corresponde a pRL90019 y <sup>b</sup> a pRL90025.

Cuando se analizó el contexto genómico de los genes *hfixL* y *fxkR* se observaron tres tipos de arreglos (Figura 11A). En el arreglo tipo 1, presente en pRet42f y pRL9 de *R. etli* CFN42 y *R. leguminosarum* bv. *viciae* 3841 la región de los genes *hfixL* y *fxkR* incluye al operon *fixNOQP* orientado en dirección opuesta a los genes *fixK*, *hfixL* y

*azuP* que codifica para una pseudoazurina regulada a nivel transcripcional por FixKf y NnrR en *R. etli* CFN42 (Gómez-Hernández N., no publicado). Es importante hacer notar que en *R. etli* CFN42 *fxkR* y *hfixL* están separados por un grupo de 20 genes mientras que en *R. leguminosarum* bv. *viciae* 3841 la región entre ambos genes contiene solamente 6. El gen *fxkR* en *R. etli* se encuentra en vecindad con genes blanco del regulón FixKf-NnrR como son el operón *nirKV-nnrRU* y *norCBQD*. Estos genes importantes para el metabolismo de nitrito y óxido nítrico (Gómez-Hernández N. y col. 2012) se encuentran ausentes en *R. leguminosarum* bv. *viciae*. Mientras que en esta región en ambos genomas está presente un gen *hemN* que codifica para una oxidasa de coproporfirinógeno III importante para el metabolismo de grupos hemo. Una observación interesante es que en orientación divergente con respecto al ortólogo de *fxkR* de *R. leguminosarum* (Figura 11A) se encuentra un gen que codifica para una proteína de la familia CRP/FNR. Esta probable proteína FNR tiene una calificación de 46 cuando se compara con FixKf usando ClustalW2 (Figura 11B) y conserva el dominio de unión EXXSR presente en el HTH de proteínas tipo-FNR (Green J. y col. 2001). Esto nos hace sugerir que en el plásmido pRL9 están codificadas dos proteínas tipo-FixKf.

Se encontró un arreglo diferente (tipo 2) en los genomas de *R. rhizogenes* K84, *S. meliloti* 1021, *S. meliloti* SM11, *M. nodulans* ORS 2060 y en los genes tipo-*hfixL* y *fxkR* cromosomales de *R. leguminosarum* bv. *viciae* 3841. En este tipo de arreglo, se encontró que los genes *fxkR* y *fixK* están orientados divergentemente, en el caso de *S. meliloti* 1021 parece haber una proteína hipotética codificada entre estos dos genes. En *M. nodulans* una proteína hipotética podría estar codificada entre los genes *fixK* y *hfixL*. Como ocurre en el arreglo de tipo 1, los genes *fixK* y *hfixL* son codireccionales. Sin embargo no aparecen en esta región el operón *fixNOQP* y el gen *azuP*.

El arreglo genómico en las bacterias *A. cryptum* JF-5, *A. multivorum* AIU301, y *O. batsensis* HTCC2597 muestra que *fxkR* está codificado entre *hfixL* y *fixKf* en una orientación convergente con respecto a *hfixL* (tipo 3). Las proteínas hFixL, FxkR y FixK de estas bacterias son las menos conservadas de las analizadas en este trabajo (Figura 11B).

## Análisis comparativo de la secuencia primaria de las proteínas tipo-FxkR y los reguladores de la respuesta OmpR y PhoB.

El análisis de dominios de la secuencia de aminoácidos de FxkR usando varias librerías, clasifica a este regulador de la respuesta como parte de la familia OmpR/PhoB. Las proteínas OmpR y PhoB de *Escherichia coli* son las mejor caracterizadas dentro de la familia. En *E. coli* OmpR, es el regulador de la respuesta cognado de EnvZ, y en conjunto son importantes para el control de la respuesta a cambios osmóticos (Mizuno, T. y Mizushima S., 1987). El residuo de aspartato en la posición 55 es el sitio primario de fosforilación del regulador, sin embargo el Asp<sup>11</sup> puede ser fosforilado particularmente cuando el Asp<sup>55</sup> está ausente (Delgado J. y col. 1993). La Val<sup>203</sup>, la Arg<sup>207</sup> y la Arg<sup>209</sup> son residuos importantes para que el regulador reconozca de manera específica el DNA (Rhee JE. y col. 2008). El análisis comparativo de la secuencia primaria de las proteínas tipo-FxkR reveló que todas conservan los dos sitios de fosforilación de OmpR con la excepción del ortólogo FxkR de *R. leguminosarum* bv. *viciae* 3841 como aparece anotado en la secuencia reportada (Figura 12). Sin embargo, si tomamos como sitio de inicio de la traducción el codón de la metionina que se localiza a 63 pb río arriba del inicio de la traducción predicho, se genera un marco abierto de lectura ininterrumpido que cubre la región completa de *fxkR* de *R. etli* CFN42 y que conservara el Asp<sup>11</sup> de OmpR. Adicionalmente, seis bases río arriba de este ATG encontramos una secuencia parecida al sitio de unión del ribosoma (RBS) (GACGGA); mientras que en la cercanía del sitio de inicio de la traducción anotado no pudimos identificar un RBS.

Acry_0577	-----MDAGTILIVDDDPQIRTMLRRYLAGEGYGTREADSEAGTHAALLGGDVTLVLLDL	▲	55
ACMV_06510	-----MDAGTILIVDDDPQIRTMLRRYLAGEGYGTREADSEAGTHAALLGGDVTLVLLDL		55
RHE_PF00530	-MAAPQHDAHLLVDDDPRIQMLTRYFEDEGYTVSSAADGTEMRVRLRQQNFDAVLLDL		59
pRL90026	-----MLTRYFEDEGYAVSSASDGIEMRVHMRQQSFDAIILLDL		38
Arad_7575	MAPDQQRREHVLVDDDPRIQMLTRYFEDEGYAVSTASNGLEMRAQLQENFDAIILLDW		60
RL1881	MTSGTVRRTRILVDDDVRIQRMLAQYLGEEGFDVDVAGGSSQMRARMAGAKFDAVLLDL		60
pSmeSM11ap133	MANAPTLPHEHILVDDDSRIRQMLSRYLESEGFVSTAESGTAMRKALASEKIDLVLLDL		60
Sma1138	-----		
Mnod_6165	----MSEPPHILITDDDPRIQMLSRYLESEGFVSTAESGTAMRKALASEKIDLVLLDL		56
OB2597_20621	---MDVTAKRILVDDDAKIRTLRLRCFESDGYEVVDAEAAESVRAAFAEGDFDLVTLDL		57
ompR_Ec	----MQENYKILVDDDMRLRALLERYLTEQGFQVRSVANAEQMDRLLTRESFHLMLVLLDL		56

```

phoB_Ec          -----MARRILVVEDEAPIREMVCVLEQNGFQPVEAEDYDSAVNQLNEPWPDLILLDW 54

Acry_0577       MLGGE-DGLALARAIRARSD---VPIIMLTGKDDVIDRVAGLEAGADDYIAKPFHLREVL 111
ACMV_06510      MLGGE-DGLALARAIRARSD---VPIIMLTGKDDVIDRVAGLEAGADDYIAKPFHLREVL 111
RHE_PF00530     VLPGGTDGLDLAREIRAQSD---VPIMMLTGRDDVVDRIVGLVGGADDYIAKPFHLREVL 116
pRL90026        VLPGGADGLDLAREIRAQSD---VPIMMLTGRDDVVDRIVGLVGGADDYIAKPFHLREVL 95
Arad_7575       VLPGGKDGDLAREIRAQSD---VPIIMLTGRDDVVDRIVGLVGGADDYIAKPFHLREVL 117
RL1881          VLPGDCDGLQLAREIRSGSD---VPIIMLSGRDDVMDRVVGLVGGADDYIGKPFHLREVL 117
pSmeSM11ap133  MLPGE-DGLTLARDIRARSD---VPIIMLTGRSDVVDVGVGLVGGADDYIAKPFHLREVL 116
SMa1138         -----MLTGRSDVVDVGVGLVGGADDYIAKPFHLREVL 33
Mnod_6165       VLAGE-DGLELAREIRASSR---VGIIMVTGRSDIIDMVVGLVGGADDYIAKPFQIREVL 112
OB2597_20621   NLGAE-DGLDIARELRRDHD---VPIFMVTGKDDVIDRVVGLVGGADDYLTKEPFHVREVL 113
ompR_Ec         MLPGE-DGLSICRRLRSQSNP--MPIIMVTAKGEEVDRIVGLVGGADDYIPKPFNPPELL 113
phoB_Ec         MLPGG-SGIQFIKHLKRESMTRDI PVVMLTARGEEDRVGRLETGADDYITKPFSPKELV 113

                *:::..: * : *:* *****: *** :*:

```

```

Acry_0577       ARIRALKRRMTGRGTPEPAEPDPP---EAVIRFDGHALDPLGRTLGRGDGADIPLTTAEF 168
ACMV_06510      ARIRALKRRMTGRGTPEPAEPDPP---EAVIRFDGHALDPLGRTLGRGDGADIPLTTAEF 168
RHE_PF00530     ARLKSI LRRRQPPARNPEAAG-----EEIIGFEDWKLNL SRRQLLDPEGTEIELTTGEF 170
pRL90026        ARLKSI LRRRQSSGPN AEPVS-----EEIVGF EYWKLNL SRRQLVDPQGA EIELTTGEF 149
Arad_7575       ARLKSI LRRRQPAIRSPDKNP-----DDVIRFEGWQLDTGKRQLLAAEGHEIELTTGEF 171
RL1881          ARVKT V LRRRSPAG-DPVRIE-----GDVFRFEGWKLDVERRAVTNPAGDEVAMSTGEF 170
pSmeSM11ap133  ARVRGVLRRRQPRRSPEVGDQQ-----TEIYCFEGLRLDVGSRQLLSDDEREVPLTTGEF 171
SMa1138         ARVRGVLRRRQPRRSPEVGDQQ-----AEVYSFEGLRLDVGSRQLLSDDEREVPLTTGEF 88
Mnod_6165       ARTRSVLRRLQPSASTETAKAVGPPSRAEVIGFEGWRLDISRRELMSP EGQEVPLTTGEF 172
OB2597_20621   ARVRSVLRRTARREAPPAASRETA---APWQGLDGLRINLDGMQLLDREGRDCALTTADF 170
ompR_Ec         ARIRAVLRRQANELPGAPSQE-----EAVIAFGKFKLNLGTREMFR-EDEPMP L TSGEF 166
phoB_Ec         ARIKAVMRRISPMA-----VEEVIEMQGLSLDPTSHRVMAGEEP-LEMGPTEF 160

** : : **                : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Acry_0577       ELLAVFLRHAGRVLDRDQLLNLLKGRDFEAFDRL VDIQVRLRRKIEQDPGQPQLIKTVR 228
ACMV_06510      ELLAVFLRHAGRVLDRDQLLNLLKGRDFEAFDRL VDIQVRLRRKIEQDPGQPQLIKTVR 228
RHE_PF00530     DMLTTFVRHAGRVLTRDVLMDLTRGRNLEAFDRT IDAQIVRLRRKIETDPKKPQFIKAVR 230
pRL90026        DMLATFVRHAGRVLTRDVLMDLTRGRNLEAFDRT IDAQIVRLRRKIETDPKRPQFIKAVR 209
Arad_7575       DILVVFARHAGRVLSRELLMDLTRGRNLEAFDRT IDAQIVRLRRKIETDPARPQLIKSIR 231
RL1881          DMLHVFAAHAGRVLTRDMLMDLTRKGRDGFDR IDALIVRLRRKIETDPASPRLIKSVR 230
pSmeSM11ap133  DMLCVLVKHAARVLQREFLMDLTRGRNLEAFDRS IDAQIVRLRRKIERDHTRPALIKSVR 231
SMa1138         DMLCVLVKHAGRVLQREFLMDLTRGRNLEAFDRS IDAQIVRLRRKIERDHTRPALIKSVR 148
Mnod_6165       DLLAAMAKHPGRVLSRDM LMDLTRGREWDALDRS IDAQVRLRRKLEADPKNP KLIKSVR 232
OB2597_20621   KLLAAFVGNPQRQLSRDLMDLVNGSDWTPLDRT IDNQVRLRRKLERDPARPQLIKTVR 230
ompR_Ec         AVLKALVSHPREPLSRDKLMNLARGREYSAMERS IDVQISRLRRMVEEDPAHPRYIQTVW 226

```



```

phoB_Ec      KLLHFFMTHPERVYSREQLLNHVWGTNVYVEDRT VDVHIRRLLRKALEPG-GHDRMVQTVR 219
              :*  :  :. .  * : *::          :* :*  : ***: :* .      :::

Acry_0577    GAGYVFAGKLER----- 240
ACMV_06510   GAGYVFAGKLER----- 240
RHE_PF00530  GVGYVFTARLD----- 241
pRL90026     GVGYVFTARLD----- 220
Arad_7575    GVGYVFTGKV----- 241
RL1881       GVGYLFTARLETASGQP 247
pSmeSM11ap133 GVGYVFTARTTRQRS-- 246
SMa1138      GVGYVFTARTTRQRS-- 163
Mnod_6165    GAGYVFAGRIERM---- 245
OB2597_20621 GLGYMLTEKPVDLA-- 245
ompR_Ec      GLGYVFVPDGSKA---- 239
phoB_Ec      GTGYRFSTRF----- 229
              * * * :

```

**Figura 12.** Alineamiento múltiple de proteínas FxkR con los reguladores OmpR y PhoB de *E. coli*. La secuencia primaria de las proteínas FxkR de *R. etli* CFN42 (RHE\_PF00530), *R. leguminosarum* bv. *viciae* (pRL90026 and RL1879), *R. rhizogenes* (Arad\_7578), *S. meliloti* 1021 (SMa1142), *S. meliloti* SM11 (pSmeSM11ap133), *A. cryptum* (Acry\_0578), *A. multivorans* (ACMV\_06520), *O. batsensis* (OB2597\_20626) y *M. nodulans* (Mnod\_6162) se alinearon con las secuencias de OmpR y PhoB de *E. coli* para analizar la conservación de aminoácidos importantes para la función de estas últimas. El triángulo marca las posiciones en las que OmpR puede ser fosforilado (D<sup>11</sup> y D<sup>55</sup>). Los aminoácidos que contactan al DNA están indicados por un punto. Los residuos que contactan a la RNA polimerasa están doblemente subrayados en PhoB y subrayados en OmpR. Los aminoácidos que participan en la interfaz de interacción Proteína-Dna en el modelo tridimensional de FxkR y la estructura tridimensional de PhoB se encuentran coloreados cada uno con un color diferente.

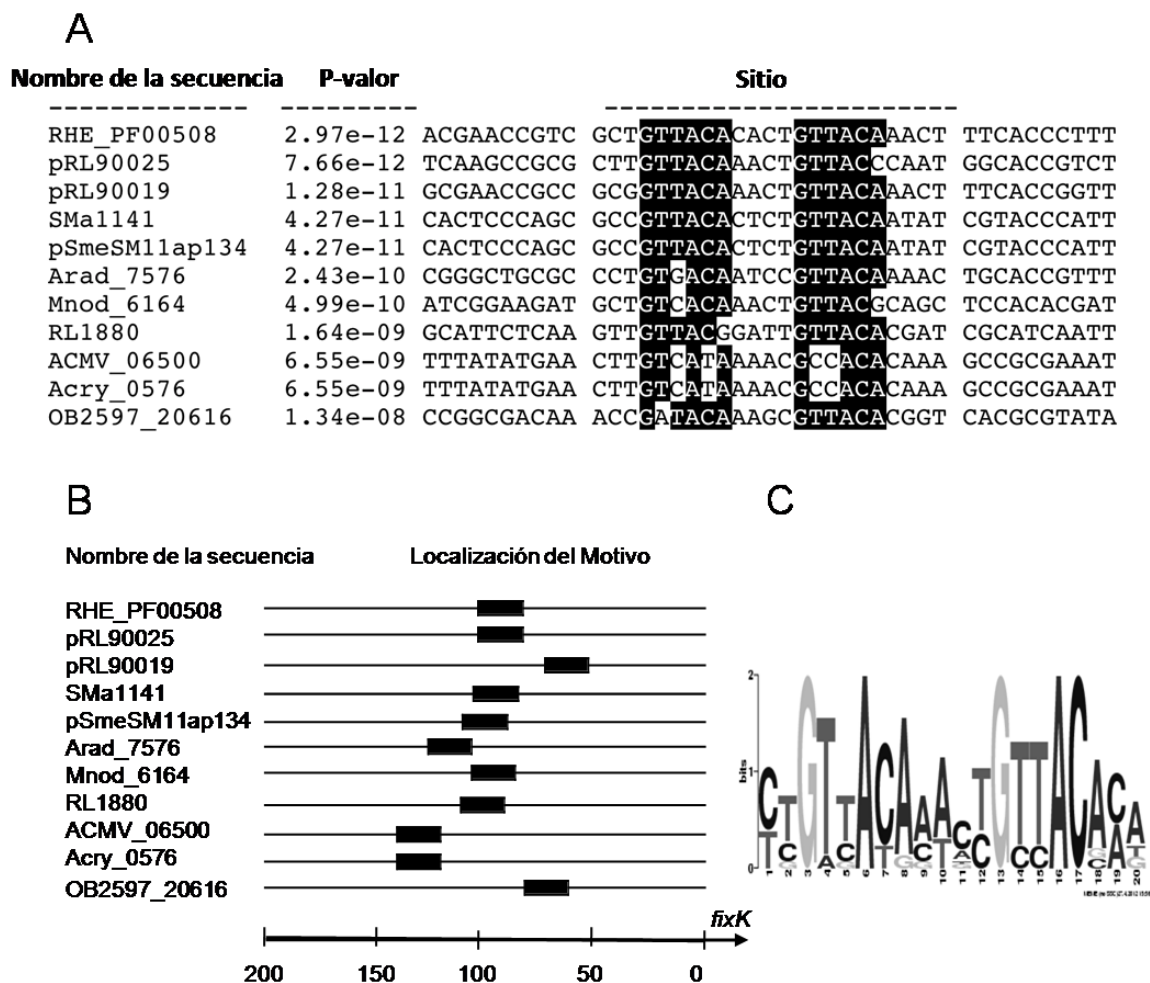
Nuestro análisis también reveló que en la región de reconocimiento del DNA todas las proteínas FxkR estudiadas conservan la Arg<sup>207</sup> y la Arg<sup>209</sup>, mientras que el residuo Val<sup>203</sup> está sustituido por una alanina. Se ha propuesto que los aminoácidos que contactan a la subunidad  $\sigma^{70}$  de la RNA polimerasa en PhoB son Trp<sup>184</sup>, Gly<sup>185</sup>, Val<sup>190</sup> y el Asp<sup>192</sup>, mientras que los aminoácidos que contactan a las subunidades  $\alpha$  de la RNA polimerasa en OmpR son Arg<sup>42</sup>, Pro<sup>179</sup>, Glu<sup>193</sup>, Aln<sup>196</sup> and Glu<sup>198</sup> (Makino K. y col. 1996; Pratt LA. y col. 1994). La mayoría de las proteínas FxkR examinadas conservan únicamente el residuo Aln<sup>196</sup> de OmpR y Asp<sup>192</sup> y Gly<sup>185</sup> de PhoB (Figura 12). Estas

observaciones nos permiten sugerir que la interacción con la RNA-polimerasa podría ser considerablemente diferente en los reguladores FxkR con respecto tanto a OmpR como a PhoB, mientras que la región de reconocimiento del DNA podría parecerse a las presentes en los blancos de regulación de estos.

### **La región reguladora de los genes tipo-*fixKf* comparte una caja reguladora**

La conservación del contexto genómico de los genes tipo-*fixKf* nos llevó a proponer que podrían estar regulados de manera similar. Para identificar elementos reguladores en la secuencia río arriba de estos genes usamos el servidor web MEME (Bailey TL. y Elkan C. 1994). Interesantemente, en la región reguladora de los 11 genes tipo-*fixKf* estudiados identificamos un motivo de 20 pb con un E-valor de  $1.6(10)^{-29}$ . En la Figura 13 se muestra una representación gráfica del alineamiento del motivo, su posición con respecto al inicio de traducción y la secuencia consenso (Logo).

El motivo conservado está compuesto por dos sub-sitios de 6 pb separados por 4 nucleótidos variables (GTTACA-N<sub>4</sub>-GTTACA). La tercer posición del primer sub-sitio es la menos conservada y los motivos presentes en los genes tipo-*fixKf* de *A. cryptum* y *A. multivorans* tienen variaciones considerables con respecto al consenso GTTACA-N<sub>4</sub>-GTTACA (Figure 13). Esta secuencia se asemeja al sitio F1 de reconocimiento presente en *ompF*, al cual se une el regulador OmpR (Yoshida T. y col. 2006). A esta secuencia conservada se le nombró caja-K y a los dos sub-sitios Ka y Kb respectivamente.



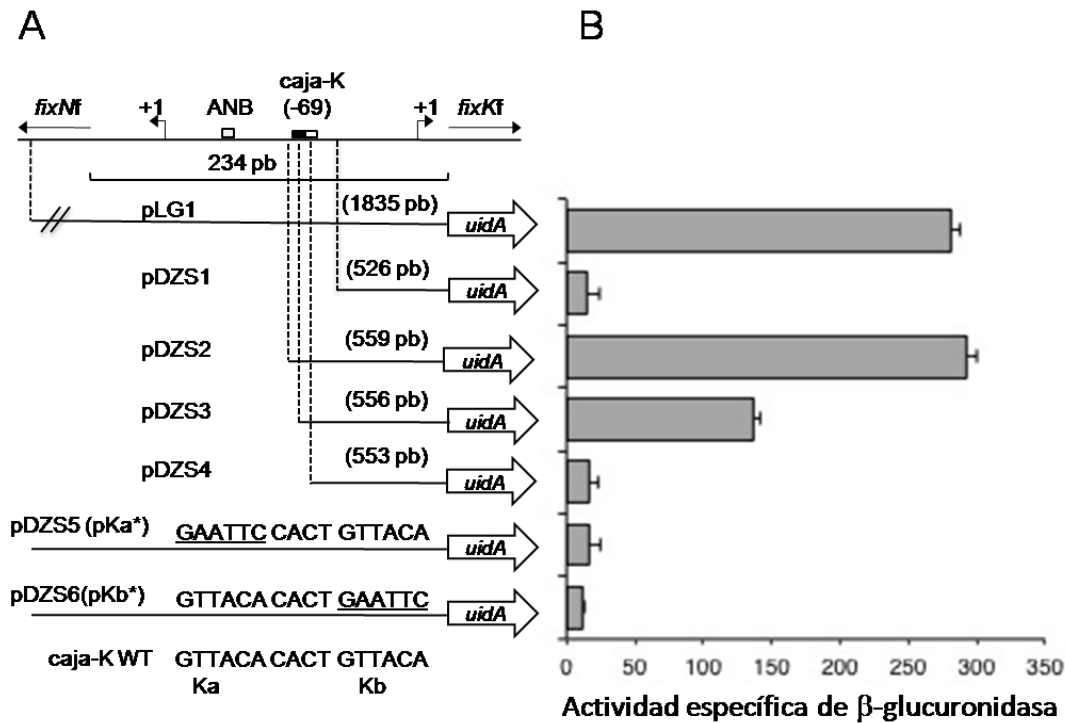
**Figura 13.** Identificación de la caja-K en la región reguladora de genes tipo-*fixKf*. (A) Motivo conservado y *p*-valor encontrados por MEME mediante el análisis de los 200 pb río arriba de los genes *fixK*. Los nucleótidos que están dentro de las cajas negras son los más conservados. (B) Representación gráfica de la localización de la secuencia conservada con respecto al sitio de inicio de la traducción de los genes *fixK*. (C) Una representación tipo Logo que muestra el grado de conservación de los nucleótidos a lo largo del motivo descubierto.

Para determinar la relevancia de la caja-K en la activación de la expresión microaeróbica de *fixKf*, realizamos el siguiente análisis de expresión. Primero identificamos el sitio de inicio de la transcripción de *fixKf* por la técnica de 5'-RACE (amplificación rápida de los extremos 5' de cDNA), a partir de RNA total purificado de cultivos de la cepa silvestre en condiciones microaeróbicas. Nuestros resultados muestran que el inicio de la transcripción de *fixKf* es una A localizada 21 nucleótidos río

arriba del probable codón de inicio. El análisis de la secuencia de DNA río arriba del gen *fixKf* reveló la presencia de una región rica en purinas que podría ser un RBS (AGGGGA) a 7 bases del codón ATG. El motivo conservado identificado en este trabajo (GTTACA-N<sub>4</sub>-GTTACA), está centrado en la posición -69 bp con respecto al inicio de la transcripción de *fixKf* (Figura 14A). Se construyeron diferentes fusiones transcripcionales con el gen monitor *uidA*, clonando diferentes fragmentos de la región reguladora de *fixKf* (Figura 7A). El plásmido pDZS1 (RR $\Delta$ *fixKf*- $\Delta$ K-box) contiene una región de 528-pb que no incluye la caja-K, el plásmido pDZS2 (RR $\Delta$ *fixKf*-K-box wt) contiene una versión recortada de la región reguladora (560-bp) que comienza con la caja-K, en pDZS3 (RR $\Delta$ *fixKf*-Ka $\Delta$ GTT) se eliminaron los primeros tres nucleótidos del sub-sitio Ka mientras que en pDZS4 (RR $\Delta$ *fixKf*- $\Delta$ Ka) el sub-sitio Ka se eliminó completamente. Estos plásmidos se introdujeron por separado en la cepa CE3 de *R. etli* y su actividad transcripcional se monitoreó bajo condiciones microaeróbicas y se comparó con el nivel de expresión obtenido con el plásmido pLG1 el cual contiene completa la región intergénica *fixNf* – *fixKf* (RR*fixKf*-wt) (Figura 14A).

La expresión microaeróbica de *fixKf* se vio drásticamente afectada cuando se eliminó la caja K, indicando que este motivo o parte de la secuencia río arriba de él se requiere para la expresión (Figura 14B comparar el nivel de expresión de pLG1 vs. pDZS1). El nivel de expresión microaeróbica de los plásmidos pLG1 ( $281 \pm 7$ ) y pDZS2 ( $292 \pm 8$ ) es muy similar, lo cual sugiere que la región reguladora presente en el pDZS2 es suficiente para la expresión de *fixKf* y que la secuencia río arriba de la caja-K parece ser dispensable para su activación. La eliminación de los primeros 3 nucleótidos de la caja-K afecta el nivel de expresión de *fixKf*, el cual disminuye a niveles cercanos al 50% (pDZS3,  $136 \pm 6$ ); mientras que la eliminación completa del sub-sitio abate en su totalidad la transcripción (pDZS4,  $17 \pm 6$ ) (Figura 14B). Este resultado sugiere que la presencia del sub-sitio Ka completo es esencial para la activación y máxima expresión de *fixKf*. Para evaluar la importancia de los sub-sitios Ka y Kb de manera independiente, se generaron mutaciones sitio-específicas en cada uno de ellos. En el pLG1 el motivo GTTACA se sustituyó por un sitio de restricción *EcoRI* de manera individual en pKa\* o en pKb\* (de GTTACA a GAATTC, ver Figura 14A). Nuestro resultados revelan que ambas mutaciones afectan drásticamente la expresión

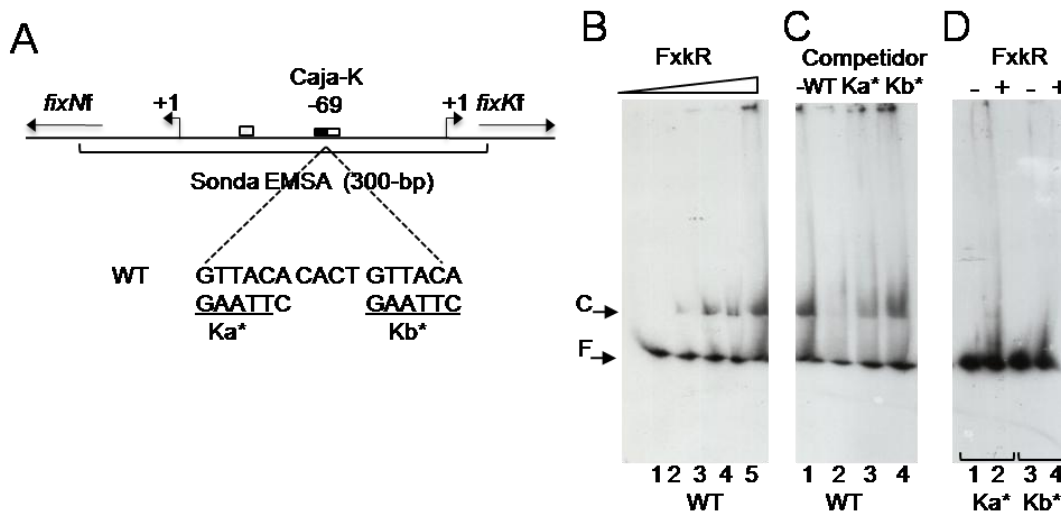
microaeróbica de *fixKf* ( $17 \pm 8$  y  $11 \pm 2$  respectivamente) (Figura 14B), confirmando que la caja-K es crucial para la activación de este gen y que ambos sub-sitios son necesarios para su correcto funcionamiento.



**Figura 14.** Análisis genético del probable sitio de unión de FxkR (A) Representación esquemática de la región reguladora de *fixKf* y los fragmentos clonados en el plásmido pBBMCS53 para generar las fusiones transcripcionales *fixKf-gus*. Los genes estructurales *fixNf* y *fixKf* están representados por flechas que indican la dirección de la transcripción. El inicio de la transcripción de cada gen está indicado con un +1 y una flecha doblada. Los sub-sitios de la caja-K están representados por cuadros negros y blancos. La caja de anaerobiosis de *fixNf* está representada por una caja sin relleno llamada ANB. La posición relativa de los elementos identificados y el tamaño de los fragmentos clonados en el plásmido pBBMCS53 se indica. Una representación a escala del tamaño de las fusiones se muestra mediante líneas punteadas. Las flechas al final de cada fusión transcripcional representan al gen *uidA*. Las leyendas pKa\* y pKb\* denotan la presencia de mutaciones sitio dirigidas en los sub-sitios Ka y Kb respectivamente. Los nucleótidos subrayados indican el sitio *EcoRI* introducido. (B) Análisis de expresión microaeróbica. Cada barra de actividad está posicionada de acuerdo con la fusión transcripcional correspondiente mostrada en el panel A. La actividad específica de  $\beta$ -glucuronidasa se expresa en  $\text{nmol min}^{-1} \text{mg protein}^{-1}$ . Los datos son el promedio de dos réplicas de tres experimentos independientes.

### **FxkR se une a la región intergénica *fixNf-fixKf* *in vitro*.**

Por medio de ensayos de movilidad eletroforética (EMSA) estudiamos la posible interacción de FxkR con la región reguladora de *fixKf*. Para ello generamos una proteína recombinante FxkR de *R. etli* con cola de histidinas (His<sub>6</sub>-FxkR) y amplificamos por PCR fragmentos de 300 pb de la región reguladora de *fixKf* que incluyeran la caja-K silvestre o los sub-sitios mutados pKa\* y pKb\* (Figura 15A). La incubación de la sonda silvestre marcada con concentraciones crecientes de la proteína His<sub>6</sub>-FxkR resultó en un incremento en la cantidad de complejo Proteína-DNA (Figura 15B). La unión de la proteína recombinante fue específica a juzgar por el hecho de que el DNA inespecífico no marcado (DNA de espermatozoos de salmón) a una concentración 100 veces superior al de la sonda no eliminó la formación del complejo (Figura 15C comparar carriles 1 y 2). Adicionalmente los ensayos de competencia revelaron que las regiones mutadas no son capaces de competir por la unión de la proteína mientras que la región silvestre sí lo es. Esto sugiere que FxkR es incapaz de unirse a una región que contenga mutaciones específicas en los sub-sitios Ka o Kb (Ka\* or Kb\* respectivamente) (Figura 15C, carriles 3 y 4). Además cuando se sustituyeron los sub-sitios de la caja-K por sitios de restricción *EcoR1* en la sonda marcada (Ka\* and Kb\*) la unión de la proteína His<sub>6</sub>-FxkR se vio severamente comprometida (Figura 15D). Estos resultados nos permiten concluir que FxkR es el activador directo de la expresión de *fixKf*, que la secuencia GTTACA-N<sub>4</sub>-GTTACA es el sitio de unión que reconoce esta proteína y que ambos sub-sitios son necesarios para que la unión ocurra.



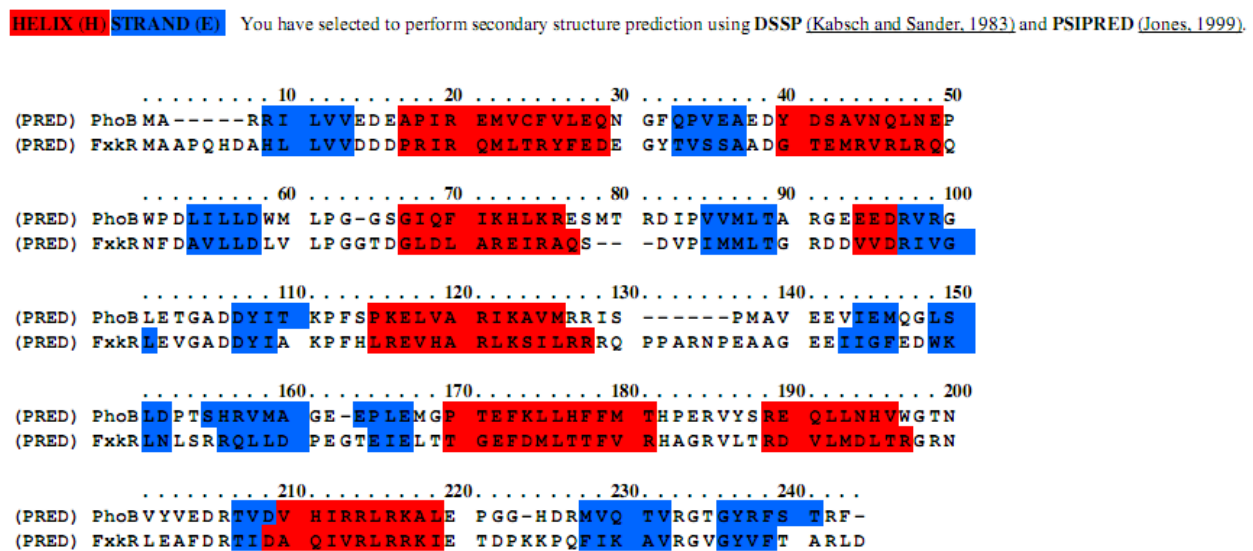
**Figura 15.** La proteína recombinante FxkR de *R. etli* se une específicamente a la caja-K. (A) Representación gráfica de la región utilizada como sonda para los ensayos de EMSA. Se muestra la sonda silvestre y las sustituciones hechas por mutación sitio dirigida (WT, Ka\* y Kb\* respectivamente). (B) El ensayo de EMSA se realizó sin proteína His<sub>6</sub>-FxkR (carril 1) o con 0.3, 0.7, 1.4 y 2.9 nM de proteína His<sub>6</sub>-FxkR (carriles 2, 3, 4, y 5 respectivamente) más sonda silvestre marcada. (C) Se hizo un ensayo de competencia con 2.9 nM de His<sub>6</sub>-FxkR más sonda silvestre marcada sin competidor específico no marcado (carril 1) o con un exceso de 100 veces de competidor específico silvestre (carril 2), o mutado en Ka\* (carril 3) o en Kb\* (carril 4) sin marcar. D) Ensayo de unión utilizando como sonda marcada fragmentos con mutaciones ya sea en el sub-sitio Ka (carriles 1, 2) o Kb (carriles 3, 4) con o sin 2.9 nM His<sub>6</sub>-FxkR. El complejo FxkR-DNA (C→) y el DNA libre (F→) están indicados.

### Predicción de la interfaz de unión a DNA por métodos bioinformáticos

Existen métodos bioinformáticos que permiten predecir a partir de modelos tridimensionales el tipo de aminoácidos que participan en la interacción con el DNA y los tipos de contacto que ocurren. Para poder modelar la estructura de FxkR unido a su DNA blanco se requiere utilizar como templado a las coordenadas atómicas de un complejo Proteína-dna cristalizado que pertenezca a la familia OmpR/PhoB. El único complejo de este tipo que se tiene corresponde a 1GXP (PhoB-pho\_box) y está depositado en la base de datos "Protein Data Bank". A pesar de que la similitud de secuencia entre las proteínas FxkR y PhoB es relativamente baja, la predicción de estructura secundaria muestra una alta conservación dentro de estas proteínas (Figura 16). Lo anterior hace posible modelar la estructura cuaternaria de FxkR de *R. etli*

usando como templado la estructura cristalográfica del complejo PhoB-pho\_box (1gxpAB). Utilizando el servidor web 3d-footprint pudimos determinar que una gran parte de los aminoácidos identificados dentro de la interfaz de unión Proteína-Dna en 1gxp se encuentran conservados en FxkR (Figura 12).

Para generar el modelo dimérico de FxkR en complejo con la caja pho (FxkR\_AB-pho\_box) se utilizó el servidor web de modelaje Tformodeller como se menciona en materiales y métodos. La energía de los contactos intramoleculares se minimizó utilizando el programa Swiss-Pdbviewer.



**Figura 16.** Predicción de estructura secundaria de las proteínas PhoB y FxkR. Con rojo se marcaron los aminoácidos que se predice forman estructuras tipo helice y con azul aquellos que forman láminas.

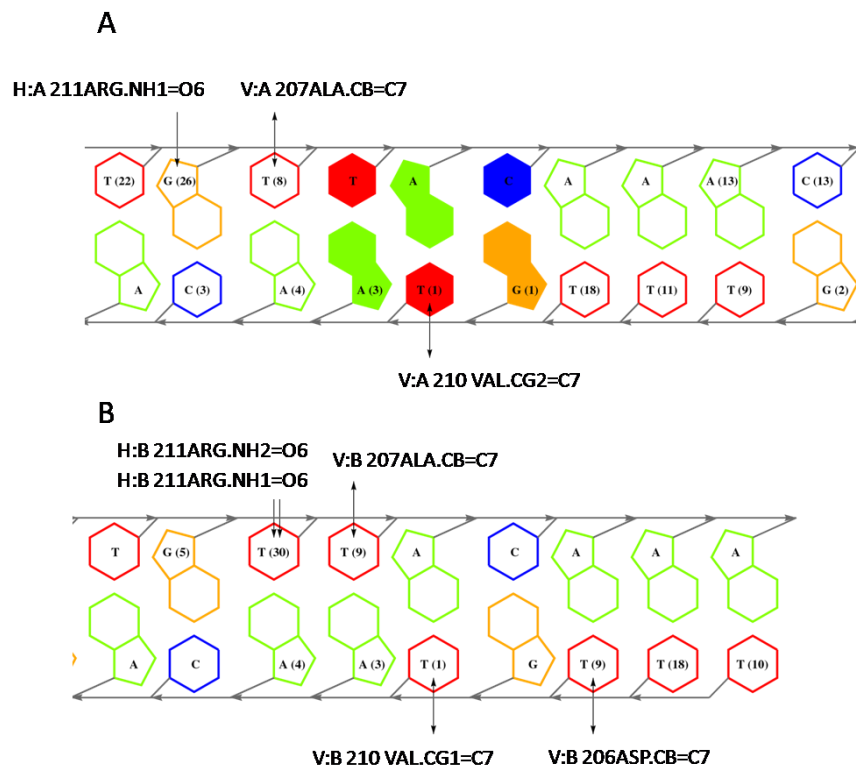
La caja-pho del complejo original (GAGCTGTCATAAAGTTGTCACGG) se modificó para generar una secuencia similar a la caja-K (GAGCTGTTACAAACTGTTACAAA). Esto se logró usando un programa (“script”) generado con el lenguaje de programación python.

Con base en información obtenida del modelaje tridimensional se hizo una predicción de la probable interfaz de reconocimiento Proteína-Dna. El servidor web 3d-footprint se utilizó para analizar la probable interfaz de reconocimiento. Mediante la aplicación “footprint your structure” el servidor predice los contactos entre los



aminoácidos y los nucleótidos del complejo, el Logo y una matriz de posición y peso del probable sitio de unión a DNA.

Como se observa en la Figura 17 basado en el modelo estructural los contactos entre la cadena A y el sub-sitio Ka son muy similares pero no idénticos a los de la cadena B con el sub-sitio Kb. Las diferencias podrían estar dadas por estimaciones no necesariamente correctas debidas al modelaje de la estructura cuaternaria. En los geles de retardo no tenemos la resolución suficiente para aseverar que existan diferencias en la unión de la proteína recombinantes con las regiones de DNA mutadas en los sub-sitios Ka y Kb.

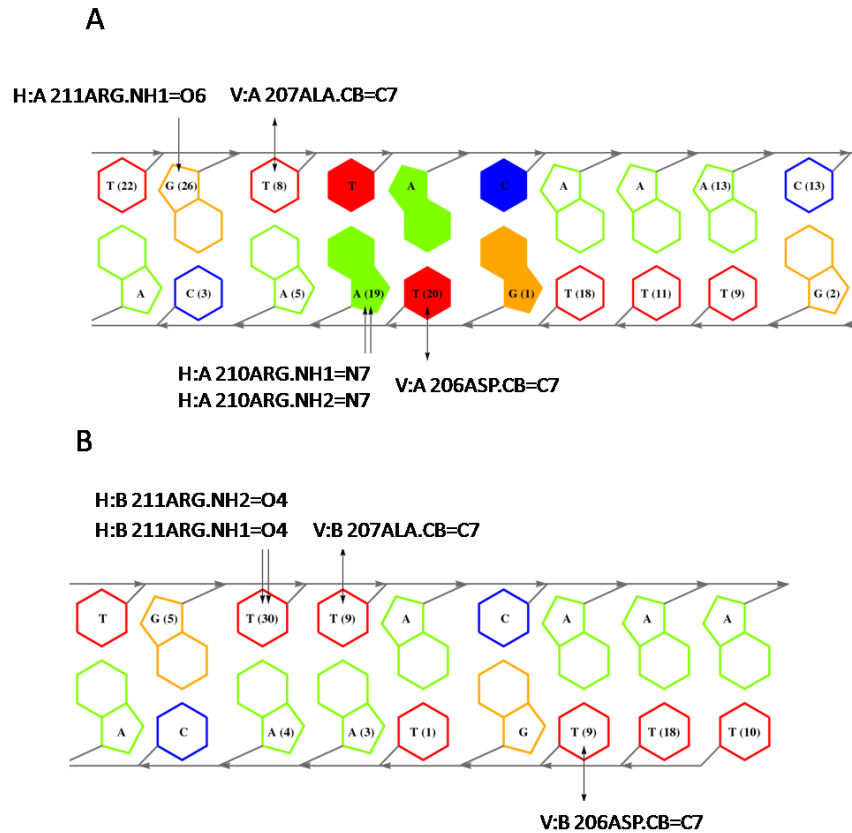


**Figura 17.** Contactos en la interfaz de la proteína FxkR de *R. etli* y la caja-K de *fixKf*. A) Contactos de la cadena A con el sub-sitio Ka. B) Contactos de la cadena B con el sub-sitio Kb. Las flechas con una sola punta indican contactos por puente de hidrógeno, las flechas con doble punta indican interacciones hidrofóbicas. Los números encerrados en paréntesis indican la cantidad de contactos de la base nitrogenada con cadenas laterales. La leyenda superior de las flechas indica el tipo de contacto (H- puente de hidrógeno; V-interacción hidrofóbica), la cadena (A o B), el aminoácido que contacta la base nitrogenada y los grupos que forman la interacción. Los nucleótidos rellenos tienen una deformación en la geometría del DNA generada por el algoritmo DNAPROT.

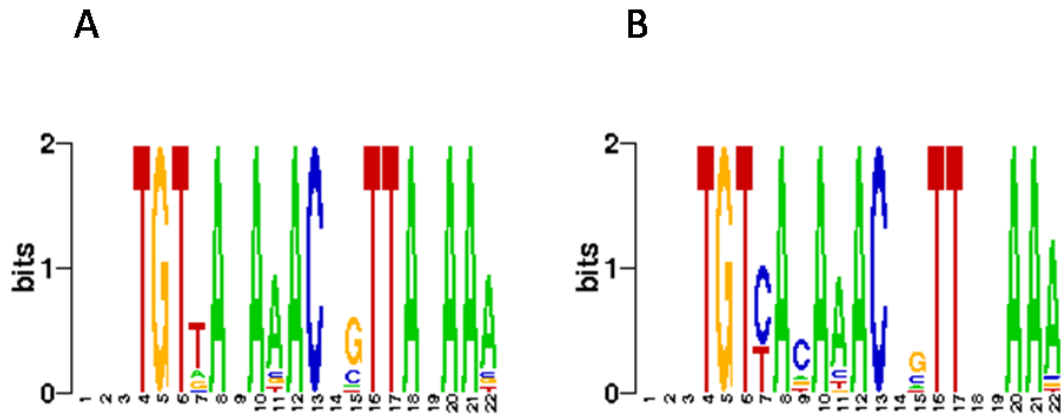
Cuatro aminoácidos parecen ser los que generan la mayoría de los contactos con la caja-K, la Arg<sup>211</sup> que forma puentes de hidrógeno con la primera guanina del sub-sitio Ka (GTTACA) y la primer timina del sub-sitio Kb (GTTACA), la Ala<sup>207</sup> que tiene interacciones hidrofóbicas con la primer timina del sub-sitio Ka (GTTACA) y la segunda timina del sub-sitio Kb (GTTACA), la Val<sup>210</sup> forma interacciones hidrofóbicas con la timina complementaria de la primera adenina del sub-sitio Ka y del sub-sitio Kb (GTTACA), mientras que el Asp<sup>206</sup> forma interacciones hidrofóbicas con la timina complementaria de la segunda adenina del sub-sitio Kb (GTTACA). La Arg<sup>211</sup> y el Asp<sup>206</sup> están conservados con respecto a la Arg<sup>201</sup> y el Asp<sup>196</sup> de PhoB, la Ala<sup>207</sup> esta semi-conservada con respecto a la Val<sup>197</sup> de PhoB mientras que la Val<sup>210</sup> sustituye a la Arg<sup>200</sup> de PhoB. El análisis de esta sustitución podría dar pistas sobre las diferencias en el reconocimiento de blancos de regulación por estas dos proteínas. Como una aproximación inicial hicimos una mutación *in silico* para sustituir la Val<sup>210</sup> por una arginina en el modelo tridimensional (Val<sup>210</sup>Arg) (Figura 18). Las interacciones son similares sin embargo esta modificación genera interacciones por puentes de hidrógeno de la Arg<sup>210</sup> con la adenina complementaria a la segunda timina del sub-sitio Ka y no así en Kb (Figura 18).

El servidor 3d-footprint usa dos clases de matrices de peso y posición (PWM) para estimar las especificidades de unión. La PWM llamada "Contact", es calculada sumando los contactos entre las cadenas laterales y las bases nitrogenadas. La PWM llamada "Readout" se deriva tanto del arreglo del "score" de interacciones atómicas en la interfaz y la serie de deformaciones dependientes de secuencia a partir de las coordenadas de DNA utilizando el algoritmo DNAPROT. El programa genera una representación gráfica (Logo) de las estimaciones hechas con base en un modelo o estructura tridimensional (Figura 19). El Logo de la especificidad de unión del modelo FxkR-caja\_K se asemeja al consenso generado por MEME (Figura 13C), sin embargo los contactos con la citosina conservada de ambos sub-sitios no se observan. De igual forma la guanina del sub-sitio Kb que se encuentra altamente conservada en la región reguladora de genes tipo-*fixKf* parece no ser reconocida con alta especificidad en el modelo (Figura 19A). El Logo generado con el modelo mutado en la Val<sup>210</sup>Arg tiene diferencias interesantes. El contacto con el tercer nucleótido del sub-sitio Ka (GTTACA)

podría ocurrir tanto con una C cómo con una T (GTCACA) que da como resultado un sitio muy similar al que se une PhoB (TGTC A) (Figura 19B). Además la primera adenina del sub-sitio Kb parece no ser reconocida (GTTACA).



**Figura 18.** Representación gráfica de los contactos en la interfaz Proteína-DNA de un modelo tridimensional de FxkR en el cual se sustituyó la Val<sup>210</sup> por una Arginina. A) Contactos de la cadena A con el sub-sitio Ka. B) Contactos de la cadena B con el sub-sitio Kb. Las flechas con una sola punta indican contactos por puente de hidrógeno, las flechas con doble punta indican interacciones hidrofóbicas. Los números encerrados en paréntesis indican la cantidad de contactos de la base nitrogenada con cadenas laterales. La leyenda superior de las flechas indica el tipo de contacto (H- puente de hidrógeno; V-interacción hidrofóbica), la cadena (A o B), el aminoácido que contacta la base nitrogenada y los grupos que forman la interacción. Los nucleótidos rellenos tienen una deformación en la geometría del DNA generada por el algoritmo DNAPROT.



**Figura 19.** Estimación de la especificidad de unión de FxkR a la caja-K de *fixKf*. A) Representación gráfica de la especificidad de unión de la proteína silvestre FxkR a la caja-K. B) Representación gráfica de la especificidad de unión a la caja-K de la proteína FxkR donde se sustituyó la Val<sup>210</sup> por una Arginina.

La mayoría de estos resultados se incluyeron en el artículo (Anexo III)

**FxkR provides the missing link in the *fixL-fixK* signal transduction cascade in *Rhizobium etli* CFN42.** David Zamorano-Sánchez, Alma Reyes-González, Nicolás Gómez-Hernández, Patricia Rivera, Dimitris Georgellis and Lourdes Girard. *Molecular Plant-Microbe Interactions*. doi.org/10.1094/MPMI-05-12-0136-R. En prensa.

## Discusión y Perspectivas

Hemos demostrado que un regulador de la respuesta de la familia OmpR/PhoB aquí llamado *fxkR* (por “*fixK* regulator”), es el elemento regulador clave que completa la cascada reguladora dependiente de hFixL. Nuestros resultados demuestran que FxkR es indispensable para activar la expresión de *fixKf* en respuesta a una baja concentración de oxígeno en *R. etli* CFN42. Esta activación la logra gracias a su unión a una secuencia específica (GTTACA-N<sub>4</sub>-GTTACA) en la región reguladora de *fixKf* (caja-K). Dado que este es el primer regulador reportado que activa la expresión microaeróbica de *fixKf*, este trabajo ha servido para dilucidar una pregunta que permaneció sin respuesta por mucho tiempo y que es de relevancia en el campo de la regulación genética de la fijación de nitrógeno.

La adaptación co-evolutiva ha dado forma a los circuitos genéticos que controlan la simbiosis y la fijación de nitrógeno, llevando a diferencias significativas entre parejas simbióticas no relacionadas. La intrincada cascada de regulación de los genes *fix* en *Rhizobium etli* revela a un nuevo participante, el regulador de la respuesta FxkR. Es digno de destacar que este nuevo regulador no pertenece a la familia de reguladores de FixJ los cuales son ampliamente aceptados como proteínas cognado de las proteínas detectoras FixL y los principales distribuidores de la señal de oxígeno en varias  $\alpha$ -proteobacterias. El descubrimiento de este regulador abre varias posibilidades de líneas de investigación interesantes, como son la caracterización bioquímica de sus mecanismos de regulación, su especificidad por hFixL como la cinasa de histidinas cognado y la caracterización a nivel genómico de los circuitos reguladores de la fijación de nitrógeno u otros procesos microaeróbicos en los que podría estar involucrado no solo en *R. etli* si no en otros organismos modelo.

Previamente se había reportado que mutaciones sencillas en *fixL* o *fixKf* no afectan la fijación de nitrógeno. Con este antecedente no era inesperado que la mutante nula en *fxkR* de *R. etli* mostrara un fenotipo Fix<sup>+</sup> (datos no mostrados). Las proteínas FnrNch y FnrNd pueden sustituir a FixKf en condiciones simbióticas (Lopez O. y col. 2001). Los genes *fnrNch* y *fnrNd* no poseen la caja consenso K en su región reguladora, esta observación sugiere fuertemente que estos reguladores no están bajo el control directo

de FxkR en el nódulo. El hecho de que los reguladores tipo-FnrN se caracterizaron como detectores de oxígeno con capacidad de autoactivación podría explicar su papel como una rama independiente de control de los genes involucrados en la fijación de nitrógeno.

Uno de los objetivos de este trabajo era sentar las bases para trabajos futuros en otras bacterias que pudieran compartir este novedoso elemento regulador. Con este propósito hicimos una búsqueda *in silico* para encontrar tractos genéticos conservados que incluyeran al gen *fxkR* en  $\alpha$ -proteobacterias. Una cascada de regulación de los genes *fix* muy similar a la de *R. etli* había sido reportada previamente en *R. leguminosarum* bv. *viciae* VF39. Esta bacteria induce simbiosis efectiva con planta leguminosas de la tribu *Fabeae* (Patschkowski T. y cols. 1996). En este trabajo identificamos la presencia de un ortólogo de *fxkR* en *R. leguminosarum* bv. *viciae* VF39 y demostramos que este gen codifica para una proteína FxkR capaz de restaurar la expresión de *fixKf* en una mutante nula de *fxkR* en *R. etli* CFN42. Una perspectiva interesante sería probar si este regulador controla la expresión de *fixK* o *fnrN* en su propio fondo genético. Por otro lado el análisis funcional de los genes *fixK*, *fnrN* y *hfixL* en la cepa 3841, no ha sido reportado. En este trabajo encontramos la caja-K en la región reguladora de tres genes tipo-*fixK* pero no en la del gene *fnrN*. Por lo tanto podemos sugerir que uno o los tres genes *fixK* podrían controlar la expresión de la oxidasa terminal simbiótica de manera directa o indirecta a través de *fnrN*. El escenario regulador se vuelve más complejo si integramos las proteínas tipo-hFixL y FxkR codificadas en el cromosoma de esta bacteria. No tenemos evidencia de que ambos juegos de proteínas sean funcionales. La demostración de que FxkR pueda unirse con diferente especificidad a variante de la caja-K presente río arriba de los diferentes genes *fixK* y la especificidad de las diferentes cinasas hFixL por diferentes reguladores FxkR, podría servir para revelar si estos sistemas de dos componentes forman parte de circuitos independientes o con entrecruzamientos (“cross-talk”).

Usando la información genómica disponible, pudimos identificar proteínas hFixL, FxkR y FixK tanto en organismos fijadores de nitrógeno como en  $\alpha$ -proteobacterias no fijadoras. Las bacterias *M. nodulans* ORS2060 y *S. meliloti* SM11 tienen un juego de genes muy interesante en lo que se refiere a la respuesta microaeróbica, dado que

adicionalmente a los elementos de la cascada FixL-FixJ-FixK que podrían controlar la expresión de los operones *fixNOQP* y *fixGHIS*, tienen codificados los elementos de la cascada hFixL-FxkR-FixK. Será importante determinar la participación de cada regulador en el control de la expresión dependiente de la concentración de oxígeno en estas bacterias. En la misma línea de pensamiento, sería crucial determinar cuál de las dos cascadas controla la expresión de diferentes reguladores FNR/CRP. Además el hecho de que ambas proteínas FixL (canónica e híbrida) coexistan en la misma célula, nos hace preguntarnos si detectan el mismo estímulo y de ser éste el caso ¿habrá diferencias en el umbral de detección? Interesantemente la proteína FxkR de la cepa 1021 está truncada y no tiene un dominio receptor del fosfato completo. Esto sugiere que esta cascada en particular no está operando o que la proteína hFixL usa otro regulador de la respuesta cognado en esta cepa en particular. Todas estas preguntas deberán abordarse para poder entender de una mejor manera cómo circuitos reguladores similares pueden ser orquestados dentro de una célula.

La presencia de sistemas de regulación FixL-FixJ y hFixL-FxkR en  $\alpha$ -proteobacterias que no son capaces de fijar nitrógeno, implica un diseño modular conservado para detectar el oxígeno. Análisis previos han descrito una correlación entre la presencia de estos sistemas de detección y el operón *ccoNOQP*. En *C. crescentus* el sistema FixL-FixJ también controla la expresión de citocromo oxidasa alternativas (Crosson S. y cols. 2005; Cosseau C. y Batut J. 2004). En este trabajo reportamos al menos tres diferentes  $\alpha$ -proteobacterias que carecen de *cbb<sub>3</sub>* y poseen al sistema hFixL-FxkR (*A. cryptum*, *A. multivorum* and *R. rhizogenes*). Esto podría sugerir que la utilización de estos sistemas de respuesta a oxígeno no está limitada a bacterias que se adaptaron evolutivamente a ambientes microaeróbicos mediante la presencia de un complejo *cbb<sub>3</sub>* de alta afinidad por oxígeno. Interesantemente en *A. cryptum* y *A. multivorum* identificamos motivos de unión de FixK (TTGAT-N<sub>4</sub>-ATCAA) en la región reguladora de genes implicados en diferentes tipos de respiración, como *cydA* (involucrado en la respiración aeróbica), *narG* (importante para la respiración dependiente de nitrato) o *hemN* y *bchE* que forman parte de la ruta de biosíntesis del tetrapirrol. Estos dos últimos genes están involucrados en la respiración microaeróbica y la fotosíntesis de *Rubrivivax gelatinosus* (Ouchane S. y cols. 2007). Un gen que

codifica para un regulador de la familia CRP/FNR posee esta caja consenso. Ambos reguladores tipo FNR podrían estar regulando coordinadamente los genes antes mencionados.

Estudios previos enfocados en la caracterización de la distribución de marcadores genéticos de plásmidos en cepas noduladoras de frijol, revelaron una diversidad apreciable intra-especies. Sin embargo la mayoría de las cepas de *R. etli* probadas mostraron un homólogo de *hfixL* de la cepa CFN42 codificado en un plásmido diferente al simbiótico y la ausencia de un gen *fixJ* (Brom S. y col. 2002). Por ello para clarificar la presencia de genes *hfixL* y *fxkR*, 18 cepas de *R. etli* se analizaron para identificar la presencia de ambos genes. Usando oligonucleótidos específicos para *fxkR* y *hfixL* de la cepa CFN42 obtuvimos los productos esperados de "PCR" (548 pb para *fxkR* y 184 pb para *hfixL*) en cada cepa. El análisis parcial de la secuencia de 12 productos de "PCR" incluyendo los de la cepa CFN42 y de *R. leguminosarum* bv. *viciae* 3841 como controles, demostró que corresponden a *fxkR* (Girard datos sin publicar). Estos resultados dan soporte a la idea de que la presencia de hFixL y FxkR es predominante en las cepas de *R. etli* y sugiere que este circuito de regulación es común dentro de estas bacterias estrechamente relacionadas. Sin embargo estas cepas se aislaron en España y en México, sería interesante saber que tan conservada está la presencia de estos genes en una colección más rica de cepas de *R. etli* de diferentes suelos. El genoma de *R. etli* CIAT652 (ahora *Rhizobium phaseoli* CIAT652 López-Guerrero MG. y cols. 2012) ha sido recientemente secuenciado (González V. y col. 2010). Mediante el análisis de la secuencia total del genoma no pudimos identificar proteínas hFixL, FxkR o FixKf conservadas. Sin embargo cuando usamos como referencia FixL, FixJ y FixK de *S. meliloti* pudimos identificar proteínas conservadas. Interesantemente los genes *fixL*, *fixJ* y *fixK* se encontraron en la misma región genética en el cromosoma (datos no mostrados).

El análisis comparativo de la secuencia de los reguladores FxkR reveló que estas proteínas conservan el sitio de fosforilación primario de OmpR. Con respecto a los residuos de reconocimiento del DNA todas las proteínas FxkR conservan la Arg<sup>207</sup> y Arg<sup>209</sup>, pero la mayoría tienen una alanina en lugar de la Val<sup>203</sup>. Estas observaciones nos hacen sugerir que la interfaz de reconocimiento Proteína-Dna debe estar



conservada en estos reguladores. Mediante modelaje de la estructura tridimensional pudimos predecir algunos otros aminoácidos que pudieran jugar un papel importante en el reconocimiento de la caja-K, varios de estos residuos están conservados en los reguladores OmpR y PhoB.

A pesar de la evidente variabilidad entre las regiones intergénicas de los genes *fixK* analizados en este estudio, la caja-K está muy conservada. Aquellas cajas-K con el menor *p*-valor (Figura 6A) fueron las de *A. cryptum*, *A. multivorans* y *O. batsensis*. Las proteínas FxkR de estas bacterias poseen un dominio HTH con la identidad mas baja cuando se compara con el dominio de FxkR de *R. etli*. Es probable que estas proteínas se unan a una caja de regulación diferente. Los sitios de unión de PhoB (TGTCa y TTACA) han sido identificados en la región reguladora del gen *pstS* de *E. coli* (Makino K. y col. 1996). Interesantemente la caja TTACA también forma parte de la caja-K pero la región TGTCa no, aunque esta secuencia se puede encontrar en la región intergénica de los genes *fixK* de *R. rhizogenes*, *M. nodulans*, *A. cryptum*, *A. multivorans* y en el gen *fixKf* codificado en el cromosoma de *R. leguminosarum* (datos no mostrados). Hay que señalar que la importancia de estos nucleótidos aún debe ser demostrada.

Tomando en cuenta las evidencias presentadas en este trabajo, proponemos que hFixL y FxkR forman parte de un nuevo sistema de relevo de fosfatos que activa la expresión de genes tipo *fixK* en respuesta a la concentración de oxígeno, no solo en en el género *Rhizobium* sino también en otros grupos de las  $\alpha$ -proteobacterias. Dado que los sistemas de dos componentes que involucran cinasas de histidina híbridas están caracterizados por una cascada de reacciones de transferencia del fosfato donde la actividad de los dominios HK, HPt y Rec son esenciales, hemos planeado experimentos encaminados a demostrar si se requieren elementos adicionales para el correcto funcionamiento de la cascada de regulación hFixL-FxkR.

## Bibliografía

- Agron, P.G., Ditta, G.S., Helinski, D.R. 1993. Oxygen regulation of *nifA* transcription *in vitro*. Proc. Natl. Acad. Sci. U S A. 90: 3506-3510.
- Anthamatten, D., and Hennecke, H. 1991. The regulatory status of the *fixL*-like and *fixJ*-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*. Mol. Gen. Genet. 225: 38-48.
- Anthamatten, D., Scherb, B., Hennecke, H. 1992. Characterization of a *fixLJ*-regulated *Bradyrhizobium japonicum* gene sharing similarity with the *Escherichia coli* *fnr* and *Rhizobium meliloti* *fixK* genes. J. Bacteriol. 174: 2111-2120.
- Bailey, T.L., Elkan, C. 1994 Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. 2: 28-36
- Barakat, M., Ortet, P., Whitworth, D.E. 2011. P2CS: a database of prokaryotic two-component systems. Nucleic. Acids. Res. 39: D771-D776.
- Batut, J., Daveran-Mingot, M.L., David, M., Jacobs, J., Garnerone, A.M., Kahn, D. 1989. *fixK*, a gene homologous with *fnr* and *crp* from *Escherichia coli*, regulates nitrogen fixation genes both positively and negatively in *Rhizobium meliloti*. EMBO J. 8: 1279-1286.
- Bauer, E., Kaspar, T., Fischer, H.M., Hennecke, H. 1998. Expression of the *fixR-nifA* operon in *Bradyrhizobium japonicum* depends on a new response regulator, RegR. J. Bacteriol. 180: 3853-3863.
- Better, M., and Helinski, D. P. 1983. Isolation and characterization of a *recA* gene of *Rhizobium meliloti*. J. Bacteriol. 115: 311-316.
- Blanco, A.G., Sola, M., Gomis-Rüth, F.X., Coll, M. 2002 Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator. Structure. 10: 701-713.
- Boesten, B., and Priefer, U.B. 2004. The C-terminal receiver domain of the *Rhizobium leguminosarum* bv. *viciae* FixL protein is required for free-living microaerobic induction of the *fnrN* promoter. Microbiology. 150: 3703-3713.
- Bravo, A., Mora, J. 1988. Ammonium assimilation in *Rhizobium phaseoli* by the glutamine synthetase-glutamate synthase pathway. J. Bacteriol. 170: 980-984.
- Brom, S., García de los Santos, A., Stepkowsky, T., Flores, M., Dávila, G., Romero, D., Palacios, R. 1992. Different plasmids of *Rhizobium leguminosarum* bv. *phaseoli* are required for optimal symbiotic performance. J. Bacteriol. 174: 5183-5189.

- Brom, S., García-de los Santos, A., Cervantes, L., Palacios, R., Romero, D. 2000. In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitiveness and cellular growth require interaction among different replicons. *Plasmid*. 44: 34-43.
- Brom, S., Girard, L., García-de los Santos, A., Sanjuan-Pinilla, J.M., Olivares, J., Sanjuan, J. 2002. Conservation of plasmid-encoded traits among bean-nodulating *Rhizobium* species. *Appl. Environ. Microbiol.* 68: 2555-2561.
- Contreras-Moreira, B., Branger, P.A., Collado-Vides, J. 2007. TFmodeller: comparative modelling of protein-DNA complexes. *Bioinformatics*. 23: 1694-1696.
- Contreras-Moreira, B. 2010. 3D-footprint: a database for the structural analysis of protein-DNA complexes. *Nucleic Acids Res.* 38(Database issue): D91-97.
- Crack, J.C., Jervis, A.J., Gaskell, A.A., White, G.F., Green, J., Thomson, A.J., Le Brun, N.E. 2008. Signal perception by FNR: the role of the iron-sulfur cluster. *Biochem. Soc. Trans.* 36(Pt 6):1144-1148.
- Cosseau, C., Batut, J. 2004. Genomics of the *ccoNOQP*-encoded *cbb<sub>3</sub>* oxidase complex in bacteria. *Arch. Microbiol.* 181: 89-96.
- Crosson, S., McGrath, P.T., Stephens, C., McAdams, H.H., Shapiro, L. 2005. Conserved modular design of an oxygen sensory/signaling network with species-specific output. *Proc. Natl. Acad. Sci. U S A.* 102: 8018-8023.
- David, M., Daveran, M.L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., Kahn, D. 1988. Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. *Cell*. 54: 671-683.
- Delgado, J., Forst, S., Harlocker, S., Inouye, M. 1993. Identification of a phosphorylation site and functional analysis of conserved aspartic acid residues of OmpR, a transcriptional activator for *ompF* and *ompC* in *Escherichia coli*. *Mol. Microbiol.* 10: 1037-1047.
- Dixon, R., Kahn, D. 2004. Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* 2: 621-631.
- Dunham, C.M., Dioum, E.M., Tuckerman, J.R., Gonzalez, G., Scott, W.G., Gilles-Gonzalez, M.A. 2003. A distal arginine in oxygen-sensing heme-PAS domains is essential to ligand binding, signal transduction, and structure. *Biochemistry*. 42: 7701-7708.
- Eady, R.R. 1996. Structure–function relationships of alternative nitrogenases. *Chem. Rev.* 96: 3013–3030.

- Einsle, O., Tezcan, F.A., Andrade, S.L., Schmid, B., Yoshida, M., Howard, J.B., Rees D.C. 2002. Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. *Science*. 297: 1696-1700.
- Elsen, S., Dischert, W., Colbeau, A., Bauer, C.E. 2000. Expression of uptake hydrogenase and molybdenum nitrogenase in *Rhodobacter capsulatus* is coregulated by the RegB–RegA two-component regulatory system. *J. Bacteriol.* 182: 2831–2837.
- Emmerich, R., Hennecke, H., Fischer, H.M. 2000. Evidence for a functional similarity between the two-component regulatory systems RegSR, ActSR, and RegBA (PrrBA) in  $\alpha$ -Proteobacteria. *Arch. Microbiol.* 174: 307–313.
- Ferrières, L., Kahn, D. 2002. Two distinct classes of FixJ binding sites defined by *in vitro* selection. *FEBS Lett.* 517: 185-189.
- Figurski, D. H., and Helinski, D. R. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc. Natl. Acad. Sci. U.S.A.* 76:1648-1652.
- Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A. 2010. The Pfam protein families database. *Nucleic Acids Res.* 38: D211-222.
- Fischer, H.M., Bruderer, T., Hennecke, H. 1988. Essential and non-essential domains in the *Bradyrhizobium japonicum* NifA protein: identification of indispensable cysteine residues potentially involved in redox reactivity and/or metal binding. *Nucleic Acids Res.* 16: 2207–2224.
- Fischer, H.M. 1994. Genetic regulation of nitrogen fixation in rhizobia. *Microbiol. Rev.* 58: 352-386.
- Foussard, M., Garnerone, A.M., Ni, F., Soupène, E., Boistard, P., Batut, J. 1997. Negative autoregulation of the *Rhizobium meliloti* *fixK* gene is indirect and requires a newly identified regulator, FixT. *Mol. Microbiol.* 25: 27-37.
- Garnerone, A.M., Cabanes, D., Foussard, M., Boistard, P., Batut, J. 1999. Inhibition of the FixL sensor kinase by the FixT protein in *Sinorhizobium meliloti*. *J. Biol. Chem.* 274: 32500-32506.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31: 3784-3788.

- Gilles-Gonzalez, M.A., Ditta, G.S., Helinski, D.R. 1991. A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. *Nature*. 350: 170–172.
- Gilles-Gonzalez, M.A., Gonzalez, G. 1993. Regulation of the kinase activity of heme protein FixL from the two-component system FixL/FixJ of *Rhizobium meliloti*. *J. Biol. Chem.* 268: 16293-16297.
- Gilles-Gonzalez, M.A., Gonzalez, G. 2005. Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses. *J. Inorg. Biochem.* 99: 1-22.
- Girard, L., Valderrama, B., Palacios, R., Romero, D., Dávila, G. 1996. Transcriptional activity of the symbiotic plasmid of *Rhizobium etli* is affected by different environmental conditions. *Microbiology*. 142: 2647-2856.
- Girard, L., Brom, S., Dávalos, A., López, O., Soberón, M., Romero, D. 2000. Differential regulation of *fixN*-reiterated genes in *Rhizobium etli* by a novel *fixL*-*fixK* cascade. *Mol. Plant-Microbe Interact.* 13:1283-1292.
- Gómez-Hernández, N., Reyes-González, A., Sánchez, C., Mora, Y., Delgado, M.J., Girard, L. 2011. Regulation and symbiotic role of *nirK* and *norC* expression in *Rhizobium etli*. *Mol. Plant-Microbe Interact.* 24: 233-245.
- Gong, W., Hao, B., Mansy, S.S., Gonzalez, G., Gilles-Gonzalez, M.A., Chan, M.K. 1998. Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proc. Natl. Acad. Sci. U S A.* 95: 15177-15182.
- González, V., Bustos, P., Ramírez-Romero, M.A., Medrano-Soto, A., Salgado, H., Hernández-González, I., Hernández-Celis, J.C., Quintero, V., Moreno-Hagelsieb, G., Girard, L., Rodríguez, O., Flores, M., Cevallos, M.A., Collado-Vides, J., Romero, D., Dávila, G. 2003. The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42 and its relation to other symbiotic genome compartments. *Genome Biol.* 4: R36.
- González, V., Santamaría, R.I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S.C., Ramírez, M.A., Jiménez-Jacinto, V., Collado-Vides, J., Dávila, G. 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc. Natl. Acad. Sci. U S A.* 103: 3834-3839.
- González, V., Acosta, J.L., Santamaría, R.I., Bustos, P., Fernández, J.L., Hernández-González, I.L., Díaz, R., Flores, M., Palacios, R., Mora, J., Dávila, G. 2010. Conserved symbiotic plasmid DNA sequences in the multireplicon pangenomic structure of *Rhizobium etli*. *Appl. Environ. Microbiol.* 76:1604-1614.
- Grabbe, R., Schmitz, R.A. 2003. Oxygen control of *nif* gene expression in *Klebsiella pneumoniae* depends on NifL reduction at the cytoplasmic membrane by

- electrons derived from the reduced quinone pool. *Eur. J. Biochem.* 270: 1555–1566.
- Granados-Baeza, M.J., Gómez-Hernández, N., Mora, Y., Delgado, M.J., Romero, D., Girard, L. 2007. Novel reiterated Fnr-type proteins control the production of the symbiotic terminal oxidase *cbb<sub>3</sub>* in *Rhizobium etli* CFN42. *Mol. Plant-Microbe Interact.* 20: 1241-1249.
- Green, J., Scott, C., Guest, J.R. 2001. Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Adv. Microb. Physiol.* 44: 1-34.
- Guex, N. and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis.* 18: 2714-2723.
- Hageman, R.V., Burris, R.H. 1978. Nitrogenase and nitrogenase reductase associate and dissociate with each catalytic cycle. *Proc. Natl. Acad. Sci. USA* 75: 2699–2702.
- Hao, B., Isaza, C., Arndt, J., Soltis, M., Chan MK. 2002. Structure-based mechanism of O<sub>2</sub> sensing and ligand discrimination by the FixL heme domain of *Bradyrhizobium japonicum*. *Biochemistry* 41: 12952–12958.
- Hartsock, A., Shapleigh, J.P. 2010. Identification, functional studies, and genomic comparisons of new members of the NnrR regulon in *Rhodobacter sphaeroides*. *J. Bacteriol.* 192: 903-911.
- Henry, J.T., Crosson, S. 2011. Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu. Rev. Microbiol.* 65: 261-286.
- Hertig, C., Li, R.Y., Louarn, A.M., Garnerone, A.M., David, M., Batut, J., Kahn, D., Boistard, P. 1989. *Rhizobium meliloti* regulatory gene *fixJ* activates transcription of *R. meliloti nifA* and *fixK* genes in *Escherichia coli*. *J. Bacteriol.* 171: 1736-1738.
- Hill, S., Austin, S., Eydmann, T., Jones, T., Dixon, R. 1996. *Azotobacter vinelandii* NifL is a flavoprotein that modulates transcriptional activation of nitrogen-fixation genes via a redox-sensitive switch. *Proc. Natl. Acad. Sci. USA.* 93: 2143–2148.
- Ho, Y.S., Burden, L.M., Hurley, J.H. 2000. Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *EMBO. J.* 19: 5288–5299.
- Hynes, M.F., Brucksch, K., Priefer, U. 1988. Melanin production encoded by a cryptic plasmid in a *Rhizobium leguminosarum* strain. *Arch. Microbiol.* 150: 326-332.

- Hynes, M.F., McGregor, N.F. 1990 Two plasmids other than the nodulation plasmid are necessary for formation of nitrogen-fixing nodules by *Rhizobium leguminosarum*. Mol. Microbiol. 4: 567-574.
- Johnston, A.W., and Beringer, J.E. 1975. Identification of the *Rhizobium* 749 strains in pea root nodules using genetic markers. J. Gen. Microbiol 87:343-350.
- Jones, D.T. 1999. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. 292: 195-202.
- Kaminski, P. A., Batut, J., and Boistard, P. 1998. A survey of symbiotic nitrogen fixation by Rhizobia. Pages 431-460 in: The Rhizobiaceae: Molecular Biology of Model Plant-Associated Bacteria. H. P. Spaink, A. Kondorosi, and P. J. J. Hooykaas, eds. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Keen, N.T., Tamaki, S., Kobayashi, D., Trollinger, D. 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. Gene. 70:191-197.
- Körner, H., Sofia, H.J., Zumft, W.G. 2003. Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. FEMS Microbiol. Rev. 27: 559-592.
- Landeta, C., Dávalos, A., Cevallos, M.A., Geiger, O., Brom, S., Romero, D. 2011. Plasmids with a chromosome-like role in rhizobia. J. Bacteriol. 193: 1317-1326.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G. 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23: 2947-2948.
- Lawson, D.M., Smith, B.E. in Metal Ions in Biological Systems Vol. 39 (eds Sigel, A. & Sigel, H.) 75–119 (Marcel Dekker, New York, 2002).
- Letunic, I., Doerks, T., Bork, P. 2009. SMART 6: recent updates and new developments. Nucleic Acids Res. 37: D229-232.
- Little, R., Dixon, R. 2003. The amino-terminal GAF domain of *Azotobacter vinelandii* NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions. J. Biol. Chem. 278: 28711–28718.
- Lois, A.F., Weinstein, M., Ditta, G.S., Helinski, D.R. 1993a. Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of *Rhizobium meliloti* are coordinately regulated by oxygen. J. Biol. Chem. 268: 4370-4375.

- Lois, A.F., Ditta, G.S., Helinski, D.R. 1993b. The oxygen sensor FixL of *Rhizobium meliloti* is a membrane protein containing four possible transmembrane segments. *J. Bacteriol.* 175: 1103-1109.
- Lopez, O., Morera, C., Miranda-Rios, J., Girard, L., Romero, D., Soberón, M. 2001. Regulation of gene expression in response to oxygen in *Rhizobium etli*: role of FnrN in *fixNOQP* expression and in symbiotic nitrogen fixation. *J. Bacteriol.* 183: 6999-7006.
- López-Guerrero, M.G., Ormeño-Orrillo, E., Velázquez, E., Rogel, M.A., Acosta, J.L., González, V., Martínez, J., Martínez-Romero, E. 2012. *Rhizobium etli* taxonomy revised with novel genomic data and analyses. *Syst. Appl. Microbiol.* En prensa.
- Makino, K., Amemura, M., Kawamoto, T., Kimura, S., Shinagawa, H., Nakata, A., Suzuki, M. 1996. DNA binding of PhoB and its interaction with RNA polymerase. *J. Mol. Biol.* 259: 15-26.
- Martínez-Salazar, J.M., Romero, D. 2000. Role of the *ruvB* gene in homologous and homeologous recombination in *Rhizobium etli*. *Gene* 243: 125-131.
- Meade, H.M., Long, S.R., Ruvkun, G.B., Brown, S.E., Ausubel, F.M. 1982. Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J. Bacteriol.* 149: 114–122.
- Mesa, S., Bedmar, E.J., Chanfon, A., Hennecke, H., Fischer, H.M. 2003. *Bradyrhizobium japonicum* NnrR, a denitrification regulator, expands the FixLJ-FixK2 regulatory cascade. *J. Bacteriol.* 185: 3978-3982.
- Mesa, S., Ucurum, Z., Hennecke H, Fischer HM. 2005. Transcription activation *in vitro* by the *Bradyrhizobium japonicum* regulatory protein FixK2. *J. Bacteriol.* 187(10):3329-38.
- Mizuno, T., Mizushima, S. 1987. Isolation and characterization of deletion mutants *ompR* and *envZ*, regulatory genes for expression of outer membrane proteins OmpC and OmpE. *J. Biochem.* 101: 387-396.
- Nellen-Anthamatten, D., Rossi, P., Preisig, O., Kullik, I., Babst, M., Fischer, H.M., Hennecke, H. 1998. *Bradyrhizobium japonicum* FixK2, a crucial distributor in the FixLJ-dependent regulatory cascade for control of genes inducible by low oxygen levels. *J. Bacteriol.* 180: 5251-5255.
- Noel, K.D., Sanchez, A., Fernandez, L., Leemans, J., Cevallos, M.A. 1984. *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. *J. Bacteriol.* 158:148-155.



- Oh, J., Ko, I., Kaplan, S. 2004. Reconstitution of the *Rhodobacter sphaeroides* *cbb<sub>3</sub>*-PrrBA signal transduction pathway *in vitro*. *Biochemistry*. 43: 7915-7923
- Ouchane, S., Picaud, M., Therizols, P., Reiss-Husson, F., Astier, C. 2007. Global regulation of photosynthesis and respiration by FnrL: the first two targets in the tetrapyrrole pathway. *J. Biol. Chem.* 282: 7690-7699.
- Patschkowski, T., Schlüter, A., Priefer, U.B. 1996. *Rhizobium leguminosarum* bv. *viciae* contains a second *fnr/fixK*-like gene and an unusual *fixL* homologue. *Mol. Microbiol.* 21: 267-280.
- Pratt, L.A., Silhavy, T.J. 1994. OmpR mutants specifically defective for transcriptional activation. *J. Mol. Biol.* 243: 579-594.
- Preisig, O., Anthamatten, D., Hennecke, H. 1993. Genes for a microaerobically induced oxidase complex in *Bradyrhizobium japonicum* are essential for a nitrogen-fixing endosymbiosis. *Proc. Natl. Acad. Sci. U S A.* 90: 3309-3313.
- Rees, D.C., Howard, J.B. 2000. Nitrogenase: standing at the crossroads. *Curr Opin Chem Biol.* 4: 559-566.
- Reyrat, J.M., David, M., Blonski, C., Boistard, P., Batut, J. 1993. Oxygen-regulated *in vitro* transcription of *Rhizobium meliloti* *nifA* and *fixK* genes. *J. Bacteriol.* 175: 6867- 6872.
- Rhee, J.E., Sheng, W., Morgan, L.K., Nolet, R., Liao, X., Kenney, L.J. 2008. Amino acids important for DNA recognition by the response regulator OmpR. *J. Biol. Chem.* 283: 8664-8677.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. *Molecular Cloning. A Laboratory Manual*. 2nd ed. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY, U.S.A.
- Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., Pühler, A. 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene.* 145: 69-73.
- Schlüter, A., Patschkowski, T., Uden, G., Priefer, U.B. 1992. The *Rhizobium leguminosarum* FnrN protein is functionally similar to *Escherichia coli* Fnr and promotes heterologous oxygen-dependent activation of transcription. *Mol. Microbiol.* 6: 3395-3404.
- Schlüter, A., Patschkowski, T., Quandt, J., Selinger, L.B., Weidner, S., Krämer, M., Zhou, L., Hynes, M.F., Priefer, U.B. 1997. Functional and regulatory analysis of

- the two copies of the *fixNOQP* operon of *Rhizobium leguminosarum* strain VF39. *Mol. Plant-Microbe Interact.* 10: 605-616.
- Seefeldt, L. C., Dance, I. G. & Dean, D. R. Substrate interactions with nitrogenase: Fe versus Mo. *Biochemistry* (2004) 43: 1401–1409.
- Sigrist, C.J., Cerutti, L., de Castro, E., Langendijk-Genevaux, P.S., Bulliard, V., Bairoch, A., Hulo, N. 2010. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res.* 38: D161-166.
- Simossis, V.A., Heringa, J. 2005. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Res.* 33(Web Server issue):W289-294.
- Söderbäck, E. Reyes-Ramirez, F., Eydmann, T., Austin, S., Hill, S., Dixon, R. 1998. The redox-and fixed nitrogen-responsive regulatory protein NIFL from *Azotobacter vinelandii* comprises discrete flavin and nucleotide-binding domains. *Mol. Microbiol.* 28: 179–192.
- Stock, A.M., Robinson, V.L., Goudreau, P.N. 2000. Two-component signal transduction. *Annu. Rev. Biochem.* 69: 183-215.
- Stock, J.B., Surette, M.G., Levit, M., Park, P. Two-component signal transduction systems: structure-function relationships and mechanisms of catalysis. In: Hoch JA, Silhavy TJ, eds. *Two-component Signal Transduction*. American Society for Microbiology, Washington, DC 2005, 1995: 25-51.
- Swem, L.R., Kraft, B.J., Swem, D.L., Setterdahl, A.T., Masuda, S., Knaff, D.B., Zaleski, J.M., Bauer, C.E. 2003. Signal transduction by the global regulator RegB is mediated by a redox-active cysteine. *EMBO J.* 22: 4699–4708.
- Taylor, B.L., Zhulin, I.B. 1999. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* 63: 479–506.
- Tuckerman, J.R., Gonzalez, G., Dioum, E.M., Gilles-Gonzalez, M.A. 2002. Ligand and oxidation-state specific regulation of the heme-based oxygen sensor FixL from *Sinorhizobium meliloti*. *Biochemistry* 41: 6170–6177.
- Yoshida, T., Qin, L., Egger, L.A., Inouye, M. 2006. Transcription regulation of *ompF* and *ompC* by a single transcription factor, OmpR. *J. Biol. Chem.* 281: 17114-17123.
- Young, J.P., Crossman, L.C., Johnston, A.W., Thomson, N.R., Ghazoui, Z.F., Hull, K.H., Wexler, M., Curson, A.R., Todd, J.D., Poole, P.S., Mauchline T.H., East, A.K., Quail, M.A., Churcher, C., Arrowsmith, C., Cherevach, I., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabinowitsch, E., Sanders, M.,

Simmonds, M., Whitehead, S., Parkhill, J. 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol.* 7(4):R34.

## Alineamiento tipo BLASTP

## FkxR Sequence Alignment

Accession_List	GI_List	Organism name	ClustalW2 alignment score vs FxkR_RetCFN4
YP_473146.1	86361259	Rhizobium etli CFN 42]	
ZP_03510313.1	218513473	Rhizobium etli 8C-3]	s (1:2) Aligned. Score: 98
YP_765323.1	116249485	Rhizobium leguminosarum bv. viciae 3841]	s (1:3) Aligned. Score: 88
YP_002540636.1	222081273	Agrobacterium radiobacter K84]	s (1:4) Aligned. Score: 75
ZP_03524328.1	218674659	Rhizobium etli GR56]	s (1:5) Aligned. Score: 97
EHJ96356.1	358004027	Agrobacterium tumefaciens 5A]	s (1:6) Aligned. Score: 66
CCE97252.1	365180397	Sinorhizobium fredii HH103]	s (1:7) Aligned. Score: 66
EHK78932.1	359506418	Sinorhizobium meliloti CCNWSX0020]	s (1:8) Aligned. Score: 66
CCE99124.1	365182274	Sinorhizobium fredii HH103]	s (1:9) Aligned. Score: 66
YP_003329431.1	270208660	Sinorhizobium meliloti]	s (1:10) Aligned. Score: 66
YP_767484.1	116251646	Rhizobium leguminosarum bv. viciae 3841]	s (1:11) Aligned. Score: 61
YP_002501284.1	220925982	Methylobacterium nodulans ORS 2060]	s (1:12) Aligned. Score: 61
CCE99068.1	365182218	Sinorhizobium fredii HH103]	s (1:13) Aligned. Score: 68
AAA74221.1	152259	Rhizobium sp.]	s (1:14) Aligned. Score: 55
ZP_00056502.1	23016749	Magnetospirillum magnetotacticum MS-1]	s (1:15) Aligned. Score: 50
YP_001233720.1	148259593	Acidiphilium cryptum JF-5]	s (1:16) Aligned. Score: 53
YP_858400.1	117621451	Aeromonas hydrophila subsp. hydrophila ATCC 7966]	s (1:17) Aligned. Score: 47
CCC96325.1	356875587	Azospirillum brasilense Sp245]	s (1:18) Aligned. Score: 49
YP_001143701.1	145300860	Aeromonas salmonicida subsp. salmonicida A449]	s (1:19) Aligned. Score: 47
YP_004390864.1	330827912	Aeromonas veronii B565]	s (1:20) Aligned. Score: 47
YP_002894193.1	237809753	Tolomonas auensis DSM 9187]	s (1:21) Aligned. Score: 47
YP_003449844.1	288959503	Azospirillum sp. B510]	s (1:22) Aligned. Score: 47
YP_002297237.1	209964322	Rhodospirillum centenum SW]	s (1:23) Aligned. Score: 48
ZP_01000594.1	84502458	Oceanicola batsensis HTCC2597]	s (1:24) Aligned. Score: 48
ZP_08831279.1	345879666	endosymbiont of Riftia pachyptila (vent Ph05)]	s (1:25) Aligned. Score: 48
YP_003447640.1	288957299	Azospirillum sp. B510]	s (1:26) Aligned. Score: 48
ZP_08817587.1	345865402	endosymbiont of Tevnia jerichonana (vent Tica)]	s (1:27) Aligned. Score: 48
ZP_01215498.1	90407312	Psychromonas sp. CNPT3]	s (1:28) Aligned. Score: 43
YP_005082460.1	374332276	Pseudovibrio sp. FO-BEG1]	s (1:29) Aligned. Score: 51
YP_004108130.1	316933148	Rhodopseudomonas palustris DX-1]	s (1:30) Aligned. Score: 45
YP_004677807.1	338740845	Hyphomicrobium sp. MC1]	s (1:31) Aligned. Score: 48
YP_585618.1	94312408	Cupriavidus metallidurans CH34]	s (1:32) Aligned. Score: 46
ZP_05083678.1	254470274	Pseudovibrio sp. JE062]	s (1:33) Aligned. Score: 51
NP_948908.1	39936632	Rhodopseudomonas palustris CGA009]	s (1:34) Aligned. Score: 45
ZP_08522011.1	334706145	Aeromonas caviae Ae398]	s (1:35) Aligned. Score: 47
YP_004011747.1	312114151	Rhodomicrobium vannielii ATCC 17100]	s (1:36) Aligned. Score: 46
YP_113275.1	53804897	Methylococcus capsulatus str. Bath]	s (1:37) Aligned. Score: 45
CCD03382.1	356882372	Azospirillum brasilense Sp245]	s (1:38) Aligned. Score: 47
ADP99725.1	311696852	Marinobacter adhaerens HP15]	s (1:39) Aligned. Score: 44
YP_004610813.1	337266758	Mesorhizobium opportunistum WSM2075]	s (1:40) Aligned. Score: 46
YP_005091216.1	374334529	Oceanimonas sp. GK1]	s (1:41) Aligned. Score: 47
YP_001425412.1	154706185	Coxiella burnetii Dugway 5J108-111]	s (1:42) Aligned. Score: 46
YP_426370.1	83592618	Rhodospirillum rubrum ATCC 11170]	s (1:43) Aligned. Score: 47
AEO47797.1	346717782	Rhodospirillum rubrum F11]	s (1:44) Aligned. Score: 47
ZP_09162141.1	358451708	Marinobacter sp. Mnl7-9]	s (1:45) Aligned. Score: 43
ZP_08867802.1	347734814	Azospirillum amazonense Y2]	s (1:46) Aligned. Score: 49
YP_001355016.1	152981322	Janthinobacterium sp. Marseille]	s (1:47) Aligned. Score: 46
YP_004873548.1	348030862	Glaciecola nitratireducens FR1064]	s (1:48) Aligned. Score: 43
NP_104011.1	13472444	Mesorhizobium loti MAFF303099]	s (1:49) Aligned. Score: 47
YP_004130132.1	319779219	Taylorella equigenitalis MCE9]	s (1:50) Aligned. Score: 45
CAM74227.1	144897363	Magnetospirillum gryphiswaldense MSR-1]	s (1:51) Aligned. Score: 48
YP_783556.1	115526645	Rhodopseudomonas palustris BisA53]	s (1:52) Aligned. Score: 44
YP_004874302.1	348589840	Taylorella asinigenitalis MCE3]	s (1:53) Aligned. Score: 44
ZP_00054307.2	46201806	Magnetospirillum magnetotacticum MS-1]	s (1:54) Aligned. Score: 46
YP_002290061.1	209886205	Oligotropha carboxidovorans OM5]	s (1:55) Aligned. Score: 44
ZP_09426576.1	365887756	Bradyrhizobium sp. STM 3809]	s (1:56) Aligned. Score: 43
ZP_09432436.1	365894283	Bradyrhizobium sp. STM 3843]	s (1:57) Aligned. Score: 44
ZP_09418224.1	365878767	Bradyrhizobium sp. ORS 375]	s (1:58) Aligned. Score: 43
YP_001207589.1	146342541	Bradyrhizobium sp. ORS 278]	s (1:59) Aligned. Score: 43
NP_820981.1	29655289	Coxiella burnetii RSA 493]	s (1:60) Aligned. Score: 46
YP_001418289.1	154247331	Xanthobacter autotrophicus Py2]	s (1:61) Aligned. Score: 45
NP_435861.1	16263068	Sinorhizobium meliloti 1021]	s (1:62) Aligned. Score: 65

ZP_09092601.1	357030657	Gluconobacter morbifer G707]	s (1:63) Aligned. Score: 46
ZP_09037784.1	354723569	Enterobacter mori LMG 25706]	s (1:64) Aligned. Score: 45
ZP_02218375.1	165918289	Coxiella burnetii RSA 334]	s (1:65) Aligned. Score: 46
YP_659811.1	109896556	Pseudoalteromonas atlantica T6c]	s (1:66) Aligned. Score: 44
YP_485580.1	86749084	Rhodopseudomonas palustris HaA2]	s (1:67) Aligned. Score: 44
YP_297528.1	73543008	Ralstonia eutropha JMP134]	s (1:68) Aligned. Score: 46
ZP_08831159.1	345879522	glutamine-hydrolyzing] 1	s (1:69) Aligned. Score: 45
YP_744558.1	114327401	Granulibacter bethesdensis CGDNIH1]	s (1:70) Aligned. Score: 46
ZP_08825435.1	345873528	Thiorhodococcus drewsii AZ1]	s (1:71) Aligned. Score: 46
YP_004675592.	338738630	Hyphomicrobium sp. MC1]	s (1:72) Aligned. Score: 44
ZP_01945737.1	153206919	Coxiella burnetii 'MSU Goat Q177']	s (1:73) Aligned. Score: 46
YP_191997.1	58040033	Gluconobacter oxydans 621H]	s (1:74) Aligned. Score: 46
YP_002304402.	212213466	Coxiella burnetii CbuG_Q212]	s (1:75) Aligned. Score: 46
ZP_09474441.1	367474952	Bradyrhizobium sp. ORS 285]	s (1:76) Aligned. Score: 43
YP_001242045.	148257460	Bradyrhizobium sp. BTAi1]	s (1:77) Aligned. Score: 43
ZP_08828256.1	345876488	endosymbiont of Riftia pachyptila (vent Ph05)]	s (1:78) Aligned. Score: 44
ZP_08990181.1	354571041	Mesorhizobium australicum WSM2073]	s (1:79) Aligned. Score: 46
YP_004141323.	319781847	Mesorhizobium ciceri biovar biserrulae WSM1271]	s (1:80) Aligned. Score: 46
YP_003693902.	298291963	Starkeya novella DSM 506]	s (1:81) Aligned. Score: 45
CAZ89725.1	294341311	Thiomonas sp. 3As]	s (1:82) Aligned. Score: 44
ZP_07661410.1	307946075	Roseibium sp. TrichSKD4]	s (1:83) Aligned. Score: 43
ZP_04716156.1	239995632	Alteromonas macleodii ATCC 27126]	s (1:84) Aligned. Score: 44
YP_002947437.	239820252	Variovorax paradoxus S110]	s (1:85) Aligned. Score: 44
YP_001951325.	189424148	Geobacter lovleyi SZ]	s (1:86) Aligned. Score: 44
ZP_06835893.1	296117303	Gluconacetobacter hansenii ATCC 23769]	s (1:87) Aligned. Score: 46
ZP_03561287.1	221134984	Glaciecola sp. HTCC2999]	s (1:88) Aligned. Score: 42
YP_005037694.	374290659	Azospirillum lipoferum 4B]	s (1:89) Aligned. Score: 48
EHB77624.1	353212224	Hyphomicrobium denitrificans 1NES1]	s (1:90) Aligned. Score: 48
AET93970.1	357940414	Burkholderia sp. YI23]	s (1:91) Aligned. Score: 44
YP_004429076.	332143338	Alteromonas macleodii str. 'Deep ecotype']	s (1:92) Aligned. Score: 44
YP_564536.1	91794885	Shewanella denitrificans OS217]	s (1:93) Aligned. Score: 43
YP_002281561.	209549644	Rhizobium leguminosarum bv. trifolii WSM2304]	s (1:94) Aligned. Score: 46
NP_773291.1	27381762	Bradyrhizobium japonicum USDA 110]	s (1:95) Aligned. Score: 43
YP_271244.1	71277926	Colwellia psychrerythraea 34H]	s (1:96) Aligned. Score: 43
ZP_08387658.1	332185911	Sphingomonas sp. S17]	s (1:97) Aligned. Score: 46
ZP_08868183.1	347735276	Azospirillum amazonense Y2]	s (1:98) Aligned. Score: 46
YP_004952051.	365970490	Enterobacter cloacae EcWSU1]	s (1:99) Aligned. Score: 45
YP_004230723.	323528571	Burkholderia sp. CCGE1001]	s (1:100) Aligned. Score: 45

## Anexo II. Proteínas tipo-hFixL

## Alineamiento tipo BLASTP

Accession_List	GI_List	Organism name	ClustalW2 alignment score vs hFixL_RetCFN4
AAG00949.1	9857983	Rhizobium etli CFN 42]	
YP_473125.1	86361238	Rhizobium etli CFN 42]	s (1:2) Aligned. Score: 99
ZP_03513251.1	218516411	Rhizobium etli 8C-3]	s (1:3) Aligned. Score: 94
ZP_03522871.1	218673202	Rhizobium etli GR56]	s (1:4) Aligned. Score: 95
ZP_03523665.1	218673996	Rhizobium etli GR56]	s (1:5) Aligned. Score: 93
CAA94319.1	1240049	Rhizobium leguminosarum]	s (1:6) Aligned. Score: 88
YP_765317.1	116249479	Rhizobium leguminosarum bv. viciae 3841]	s (1:7) Aligned. Score: 87
CCE97254.1	365180399	Sinorhizobium fredii HH103]	s (1:8) Aligned. Score: 64
NP_435864.1	16263071	Sinorhizobium meliloti 1021]	s (1:9) Aligned. Score: 64
YP_003329433.	270208662	Sinorhizobium meliloti]	s (1:10) Aligned. Score: 63
YP_002540638.	222081275	Agrobacterium radiobacter K84]	s (1:11) Aligned. Score: 63
CCE99126.1	365182276	Sinorhizobium fredii HH103]	s (1:12) Aligned. Score: 62
EHK78934.1	359506420	Sinorhizobium meliloti CCNWSX0020]	s (1:13) Aligned. Score: 61
EHJ96354.1	358004025	Agrobacterium tumefaciens 5A]	s (1:14) Aligned. Score: 58
YP_767482.1	116251644	Rhizobium leguminosarum bv. viciae 3841]	s (1:15) Aligned. Score: 56
YP_002501281.	220925979	Methylobacterium nodulans ORS 2060]	s (1:16) Aligned. Score: 47
ZP_01465884.1	115378738	Stigmatella aurantiaca DW4/3-1]	s (1:17) Aligned. Score: 37
YP_632059.1	108758219	Myxococcus xanthus DK 1622]	s (1:18) Aligned. Score: 34
YP_004668486.	338535152	Myxococcus fulvus HW-1]	s (1:19) Aligned. Score: 35
NP_945847.1	39933571	Rhodopseudomonas palustris CGA009]	s (1:20) Aligned. Score: 36
YP_001989530.	192288925	Rhodopseudomonas palustris TIE-1]	s (1:21) Aligned. Score: 36
YP_001922895.	188579450	Methylobacterium populi BJ001]	s (1:22) Aligned. Score: 34
YP_771342.1	116255509	Rhizobium leguminosarum bv. viciae 3841]	s (1:23) Aligned. Score: 35
YP_001208913.	146343865	Bradyrhizobium sp. ORS 278]	s (1:24) Aligned. Score: 37
ZP_09429869.1	365891452	Bradyrhizobium sp. STM 3809]	s (1:25) Aligned. Score: 40
ZP_09439484.1	365901652	Bradyrhizobium sp. STM 3843]	s (1:26) Aligned. Score: 36
ZP_09472489.1	367472917	Bradyrhizobium sp. ORS 285]	s (1:27) Aligned. Score: 36
YP_001203997.	146338949	Bradyrhizobium sp. ORS 278]	s (1:28) Aligned. Score: 37
YP_002984779.	241666695	Rhizobium leguminosarum bv. trifolii WSM1325]	s (1:29) Aligned. Score: 35
ZP_09421066.1	365881777	Bradyrhizobium sp. ORS 375]	s (1:30) Aligned. Score: 41
BAL07285.1	354954606	Bradyrhizobium japonicum USDA 6]	s (1:31) Aligned. Score: 31
YP_001260126.	148550687	Sphingomonas wittichii RW1]	s (1:32) Aligned. Score: 34
YP_001243575.	148258990	Bradyrhizobium sp. BTAi1]	s (1:33) Aligned. Score: 36
YP_001233721.	148259594	Acidiphilium cryptum JF-5]	s (1:34) Aligned. Score: 47
ZP_08632972.1	338983819	Acidiphilium sp. PM]	s (1:35) Aligned. Score: 47
YP_002280721.	209548804	Rhizobium leguminosarum bv. trifolii WSM2304]	s (1:36) Aligned. Score: 33
YP_004282881.	326402800	Acidiphilium multivorum AIU301]	s (1:37) Aligned. Score: 46
YP_002975121.	241204025	Rhizobium leguminosarum bv. trifolii WSM1325]	s (1:38) Aligned. Score: 32
EGE56939.1	327189798	Rhizobium etli CNPAF512]	s (1:39) Aligned. Score: 33
YP_003544106.	294010646	Sphingobium japonicum UT26S]	s (1:40) Aligned. Score: 33
YP_469066.1	86357174	Rhizobium etli CFN 42]	s (1:41) Aligned. Score: 33
ZP_01000595.1	84502459	Oceanicola batsensis HTCC2597]	s (1:42) Aligned. Score: 44
NP_772546.1	27381017	Bradyrhizobium japonicum USDA 110]	s (1:43) Aligned. Score: 36
ZP_02153691.1	163746333	Oceanibulbus indolifex HEL-45]	s (1:44) Aligned. Score: 44
YP_001977763.	190891221	Rhizobium etli CIAT 652]	s (1:45) Aligned. Score: 33
ZP_05102698.1	254489495	Roseobacter sp. GAI101]	s (1:46) Aligned. Score: 38
ZP_02730300.1	168698023	Gemmata obscuriglobus UQM 2246]	s (1:47) Aligned. Score: 33
YP_570402.1	91977743	Rhodopseudomonas palustris BisB5]	s (1:48) Aligned. Score: 41
YP_530436.1	90422066	Rhodopseudomonas palustris BisB18]	s (1:49) Aligned. Score: 35
ZP_09398096.1	365858140	Acetobacteraceae bacterium AT-5844]	s (1:50) Aligned. Score: 35
YP_003544047.	294010587	Sphingobium japonicum UT26S]	s (1:51) Aligned. Score: 37
YP_567425.1	91974766	Rhodopseudomonas palustris BisB5]	s (1:52) Aligned. Score: 34
YP_484167.1	86747671	Rhodopseudomonas palustris HaA2]	s (1:53) Aligned. Score: 35
YP_001922824.	188579379	Methylobacterium populi BJ001]	s (1:54) Aligned. Score: 34
YP_485765.1	86749269	Rhodopseudomonas palustris HaA2]	s (1:55) Aligned. Score: 41
YP_004106507.	316931525	Rhodopseudomonas palustris DX-1]	s (1:56) Aligned. Score: 36
YP_779071.1	115522160	Rhodopseudomonas palustris BisA53]	s (1:57) Aligned. Score: 36
ZP_05102691.1	254489488	Roseobacter sp. GAI101]	s (1:58) Aligned. Score: 40
BAL23049.1	358635752	Azoarcus sp. KH32C]	s (1:59) Aligned. Score: 39
ZP_08627679.1	338972304	Bradyrhizobiaceae bacterium SG-6C]	s (1:60) Aligned. Score: 31
YP_532993.1	90424623	Rhodopseudomonas palustris BisB18]	s (1:61) Aligned. Score: 38
YP_001414747.	154253923	Parvibaculum lavamentivorans DS-1]	s (1:62) Aligned. Score: 34
ZP_08773565.1	344342694	Marichromatium purpuratum 984]	s (1:63) Aligned. Score: 34
ZP_01459219.1	115371906	Stigmatella aurantiaca DW4/3-1]	s (1:64) Aligned. Score: 35

YP_003956599.1	310824241	Stigmatella aurantiaca DW4/3-1]	s (1:65) Aligned. Score: 35
ZP_02151528.1	163744168	Oceanibulbus indolifex HEL-45]	s (1:66) Aligned. Score: 36
ZP_09474559.1	367475084	Bradyrhizobium sp. ORS 285]	s (1:67) Aligned. Score: 36
YP_422676.1	83312412	Magnetospirillum magneticum AMB-1]	s (1:68) Aligned. Score: 31
YP_001860904.1	186473562	Burkholderia phymatum STM815]	s (1:69) Aligned. Score: 36
ZP_02734442.1	168702165	Gemmata obscuriglobus UQM 2246]	s (1:70) Aligned. Score: 37
AET93694.1	357940138	Burkholderia sp. YI23]	s (1:71) Aligned. Score: 35
ZP_09433679.1	365895573	Bradyrhizobium sp. STM 3843]	s (1:72) Aligned. Score: 35
AET94958.1	357941402	Burkholderia sp. YI23]	s (1:73) Aligned. Score: 36
YP_001236718.1	148252133	Bradyrhizobium sp. BTAi1]	s (1:74) Aligned. Score: 39
ZP_09439544.1	365901714	Bradyrhizobium sp. STM 3843]	s (1:75) Aligned. Score: 39
BAL09330.1	354956651	Bradyrhizobium japonicum USDA 6]	s (1:76) Aligned. Score: 36
YP_559180.1	91783974	Burkholderia xenovorans LB400]	s (1:77) Aligned. Score: 35
YP_556487.1	91781280	Burkholderia xenovorans LB400]	s (1:78) Aligned. Score: 36
YP_001863001.1	186471683	Burkholderia phymatum STM815]	s (1:79) Aligned. Score: 37
ZP_06843785.1	296160973	Burkholderia sp. Ch1-1]	s (1:80) Aligned. Score: 34
YP_004231830.1	323529678	Burkholderia sp. CCGE1001]	s (1:81) Aligned. Score: 36
NP_772317.1	27380788	Bradyrhizobium japonicum USDA 110]	s (1:82) Aligned. Score: 36
YP_558384.1	91783178	Burkholderia xenovorans LB400]	s (1:83) Aligned. Score: 36
YP_002418936.1	218528120	Methylobacterium chloromethanicum CM4]	s (1:84) Aligned. Score: 36
ZP_09426307.1	365887465	Bradyrhizobium sp. STM 3809]	s (1:85) Aligned. Score: 36
ZP_02883999.1	170692837	Burkholderia graminis C4D1M]	s (1:86) Aligned. Score: 36
YP_003609372.1	295680798	Burkholderia sp. CCGE1002]	s (1:87) Aligned. Score: 34
ZP_09474283.1	367474789	Bradyrhizobium sp. ORS 285]	s (1:88) Aligned. Score: 38
ZP_09645180.1	374572084	Bradyrhizobium sp. WSM471]	s (1:89) Aligned. Score: 36
YP_003909260.1	307726047	Burkholderia sp. CCGE1003]	s (1:90) Aligned. Score: 35
NP_767025.1	27375496	Bradyrhizobium japonicum USDA 110]	s (1:91) Aligned. Score: 37
ZP_09421056.1	365881766	Bradyrhizobium sp. ORS 375]	s (1:92) Aligned. Score: 36
ZP_07673969.1	309779206	Ralstonia sp. 5_7_47FAA]	s (1:93) Aligned. Score: 36
YP_001241215.1	148256630	Bradyrhizobium sp. BTAi1]	s (1:94) Aligned. Score: 35
ZP_09435246.1	365897229	Bradyrhizobium sp. STM 3843]	s (1:95) Aligned. Score: 35
YP_004677808.1	338740846	Hyphomicrobium sp. MC1]	s (1:96) Aligned. Score: 33
BAL09539.1	354956860	Bradyrhizobium japonicum USDA 6]	s (1:97) Aligned. Score: 35
ZP_05101771.1	254488566	Roseobacter sp. GAI101]	s (1:98) Aligned. Score: 31
YP_004350561.1	330821699	Burkholderia gladioli BSR3]	s (1:99) Aligned. Score: 33
ZP_09644752.1	374571656	Bradyrhizobium sp. WSM471]	s (1:100) Aligned. Score: 34