

23088
6



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

UNIDAD ACADÉMICA DE LOS CICLOS PROFESIONAL Y DE
POSGRADO DEL COLEGIO DE CIENCIAS Y HUMANIDADES

INSTITUTO DE BIOTECNOLOGIA

Caracterización del mecanismo de regulación por oxígeno
de la actividad de NifA de *Bradyrhizobium japonicum*.

T E S I S

QUE PARA OBTENER EL GRADO DE
DOCTORA EN BIOTECNOLOGIA

PRESENTA:

M. en C. KATY JUAREZ LOPEZ

279052
250872

Cuernavaca, Morelos, Junio del 2000



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 SE DESARROLLO EN EL DEPARTAMENTO DE
BIOESTRUCTURA Y RECONOCIMIENTO MOLECULAR DEL
INSTITUTO DE BIOTECNOLOGIA, UNAM.

AGRADECIMIENTOS

Quiero agradecer a todas aquellas personas que de alguna manera contribuyeron y me apoyaron en la realización de este trabajo.

De manera especial a Enrique Morett, porque siempre conte con su apoyo tanto académica como personalmente

A los miembros de mi comité tutorial: Xavier Soberón y Alejandro Alagón.

A los revisores de este trabajo: Guillermo Dávila, Mario Soberón, Laura Camarena, Juan Carlos Almagro, Roberto Stock, Guadalupe Espín y Enrique Morett por sus comentarios.

A todos mis maestros y tutores, especialmente a Luis Servín, Alicia González y Rosa del Carmen Mateos.

A mis compañeros y amigos del laboratorio: Lorenzo, Fily, Manuel, Joel, Nely, Paty, Lidia, Gloria, Lety, Mary, Ricardo, Ismael, MariCarmen, Ale, Gaby, Gustavo, Luisa, Angel, Juanita, Luis, Karla, Sergio.

A mis amigos del IBT: Sonia, Pancho, Roberto, Lety, Paty, Gabo, Chayito, Moreno, Raunel, Oswaldo, Humbe.

A Karla Nevárez, Leticia Olvera, Maricela Olvera y por su apoyo técnico

DEDICATORIAS

A mis padres Carlos y Catalina por haberme dado su cariño, apoyo y ejemplo para lograr todo lo que he querido.

A Dany por su cariño y apoyo siempre y en todo momento

A Carlos, Lupita y Rafael

A Karla y Miguel

A Ana, Mony, Fallito y Laurita

A toda la Fam. Segura-González

A mis amigos

INDICE

INTRODUCCION	1
- TRANSCRIPCION EN BACTERIAS	1
- Componentes de la Transcripción	1
- RNA polimerasa	1
- Estructura del promotor	3
- Mecanismo del inicio de la transcripción	4
- Formación del complejo abierto	5
- Descripción de los diversos factores sigma	6
- Regulación del inicio de la transcripción en Bacterias	11
- Regulación Negativa	12
- Regulación Positiva	12
- Regulación a distancia	14
- Modelo de activación a distancia para $E\sigma^{54}$	14
ANTECEDENTES	16
- Estructura de NifA	16
- Dominios estructurales	17
- Motivo hélice-vuelta-hélice	19
- Relaciones filogenéticas de los dominios de las EBPs	20
- Regulación de la expresión de <i>nifA</i>	21
- Regulación de la Actividad de NifA	23
- Otros reguladores transcripcionales regulados por oxígeno	
- Centros hierro-azufre	28
- Grupos Hemo	28
- Flavín	28
- Formación de enlaces disulfuro	28

OBJETIVOS	29
MATERIAL Y METODOS	31
RESULTADOS Y DISCUSION	34
- Regulación de la proteína NifA de <i>S. meliloti</i>	
- Caracterización del dominio carboxilo terminal	34
- Regulación de la actividad de la NifA de <i>B. japonicum</i>	
- Expresión del dominio C-terminal	41
- Expresión de los dominios C-terminal y Central como proteínas de fusión	42
- Caracterización de las proteínas de fusión purificadas	48
- Caracterización de la fusión MBPCENCOOH	48
- Caracterización de la fusión MBPCYSCOOH2	52
- Obtención de mutantes en el motivo conservado de cisteínas	53
- Determinación del cofactor metálico	56
CONCLUSIONES	59
BIBLIOGRAFIA	61
ANEXO 1. ARTICULO PUBLICADO	71

RESUMEN

La proteína NifA es un activador transcripcional que interviene en la expresión de los genes de fijación de nitrógeno en organismos diazotófos. La NifA interactúa con la RNA polimerasa conteniendo el factor σ^{54} para transcribir los genes *nif*. Esta proteína se une a secuencias denominadas UAS (upstream activation sequences), localizadas a más de 100 pares de bases arriba del inicio de la transcripción de los promotores *nif* y contacta la $E\sigma^{54}$, generando un asa en el DNA. Para activar la transcripción, NifA cataliza la isomerización de un complejo cerrado a un complejo abierto productivo en un proceso que requiere de la hidrólisis de un nucleósido trifosfatado. La proteína NifA pertenece a la familia de proteínas denominadas Enhancer Binding Proteins (EBPs), las cuales son homólogas estructuralmente y presentan tres dominios independientes: el N-terminal no es esencial para la actividad de NifA; el dominio central está involucrado en la activación y contiene el sitio de unión al NTP y el C-terminal contiene el motivo de unión al DNA denominado hélice-vuelta-hélice.

En este trabajo mostramos evidencia de la coevolución de dominios en la NifA de *Sinorhizobium meliloti*. Encontramos que debido a la ausencia de una glicina conservada en el motivo de unión al DNA interactúa débilmente con las UAS (esta proteína presenta ácido glutámico). Este defecto es compensado con una mayor capacidad de activación del dominio central, y en contraste con la NifA de *B. japonicum*, esta proteína puede activar en trans. Al reemplazar la glicina en la NifA de *S. meliloti*, se observó una mayor actividad transcripcional y una fuerte unión al DNA. Para explorar el camino que llevó a esta mayor capacidad de activación, se mutagenizó la proteína de *B. japonicum* y después de tres rondas de mutagénesis y selección, se obtuvo una mutante más activa, interesantemente todas las mutaciones cambiaron a aminoácidos presentes en la NifA de *S. meliloti*, lo que indicaba que posiblemente se había recreado el camino seguido por esta proteína en su evolución.

Otro aspecto estudiado en este trabajo fue la regulación por oxígeno de la proteína NifA de *B. japonicum*. Esta proteína es sensible a oxígeno *per se*, contiene un motivo conservado de cisteínas involucrado posiblemente en la coordinación de un cofactor metálico y en la regulación por oxígeno. Debido

a que la proteína completa es inestable y difícil de purificar, se trabajó con los dominios aislados en fusión con la proteína de unión a maltosa (MBP), para facilitar su purificación y la identificación del posible cofactor metálico. Se obtuvieron varias proteínas de fusión, las cuales se purificaron por cromatografía de afinidad. La fusión que contiene los dominios central, C-terminal y el motivo de cisteínas (MBPCENCOOH) conserva el 30% de la capacidad de activación transcripcional *in vivo* y retiene la capacidad de unirse al DNA específicamente *in vitro*. El análisis de absorción atómica muestra que esta fusión contiene hierro en su estructura. La fusión que contiene sólo el C-terminal y el motivo de cisteínas (MBPCYSCOOH₂) mantiene la capacidad de unión al DNA *in vitro*. Sin embargo, no presenta ningún metal en su estructura. Estos resultados sugieren que la unión al DNA *in vitro* no es regulada directamente por oxígeno y que el oxígeno podría favorecer un cambio conformacional en el dominio central que afecta el estado oligomérico de la proteína disminuyendo su capacidad de unión al DNA, posiblemente mediante el estado redox del cofactor metálico que en base a nuestros datos podría tratarse del hierro.

INTRODUCCION

TRANSCRIPCION EN BACTERIAS

La transcripción es el primer paso en la expresión genética y uno de los principales blancos para su regulación. Los elementos básicos de la transcripción fueron elucidados mediante el estudio de los sistemas en procariontes. Básicamente, la RNA polimerasa en asociación a otros factores, localiza secuencias específicas en el DNA llamadas promotores y forma lo que se denomina un complejo cerrado, el cual se isomeriza a un complejo abierto al promover la separación de una porción de DNA cercana al inicio de la transcripción. Posteriormente se inicia la síntesis de una molécula de RNA, continuando su elongación hasta que se encuentra una señal específica de terminación y el RNA es liberado (Revisado en Reznikoff, W. et. al. 1985, von Hippel, P. et. al. 1984).

Componentes de la Transcripción

- RNA Polimerasa

Las RNA polimerasas son enzimas oligoméricas complejas, constituídas por diferentes subunidades. Llevan a cabo la síntesis DNA-dependiente de una molécula de RNA, mediante la formación de enlaces fosfodiéster entre ribonucleótidos trifosfatados.

En *Eubacteria* existe un solo tipo de RNA polimerasa, que es responsable de la transcripción de las diferentes clases de RNA: el RNA ribosomal (rRNA), el RNA mensajero (mRNA) y el RNA de transferencia, (tRNA). En los eucariontes hay tres tipos: la RNA polimerasa I y la RNA polimerasa II sintetizan el rRNA y el mRNA, respectivamente; mientras que la RNA polimerasa III se encarga de sintetizar el tRNA, el 5sRNA y algunos RNAs nucleares (Coulombe & Burton, 1999).

La RNA polimerasa de *Eubacteria* esta constituída por cuatro subunidades denominadas $\alpha(2)$, β y β' , las cuales componen el "core" de la enzima, también denominado **E**. El "core" contiene la maquinaria catalítica

necesaria para la polimerización de una molécula de RNA. Sin embargo, para que se lleve a cabo la iniciación específica de la transcripción se requiere del factor sigma (σ), que se asocia de manera reversible a **E**, constituyendo la holoenzima ó **E σ** . (Burgess, et al. 1987; Polyakov, et al. 1995; Coulombe & Burton, 1999).

Existen varios tipos de factores sigma, el σ^{70} es el que interviene en las funciones esenciales de la célula. Los otros tipos intervienen en funciones específicas celulares y se describirán mas adelante.

Los genes que codifican para la **E σ^{70}** de *E. coli* son: *rpoA*, *rpoB*, *rpoC* y *rpoD*, cuyas características se resumen en la Tabla 1.

TABLA 1. CARACTERISTICAS DE LAS SUBUNIDADES DE LA RNA POLIMERASA DE *E. coli* (Burgess, et al. 1987; Ishihama, 1988)

SUBUNIDAD	GENE	No. DE AA	PM (KD)	FUNCION
β	<i>rpoB</i>	1342	150.6	Catálisis (iniciación y elongación) Reconocimiento de terminadores Unión de ribonucleósidos trifosfatados Unión del RNA producido Regulación autogena de β β' Regulación por ppGpp y unión a rifampicina
β'	<i>rpoC</i>	1407	155.16	Unión al DNA templado y al factor σ
α	<i>rpoA</i>	329	36.5	Asociación de β y β' , unión al promotor (región UP), interacción con activadores
σ^{70}	<i>rpoD</i>	613	70.2	Reconocimiento de promotores de genes esenciales

El ensamblaje de las subunidades que constituyen la **E σ** de *E. coli* siguen el siguiente orden: $\alpha + \alpha \rightarrow \alpha 2\beta \rightarrow \alpha 2\beta\beta' \rightarrow \alpha 2\beta\beta' \sigma$ (Ishihama, 1988).

Cada subunidad α de la RNA polimerasa esta formada por dos dominios unidos por un interconector flexible. El α NTD, además de participar en el

ensamblaje de la polimerasa, interviene en la dimerización de esta subunidad (Gongyi, et. al. 1999). El dominio carboxilo (α CTD), interviene en la interacción con el elemento UP del promotor y con activadores transcripcionales (Ebright & Busby, 1994; Estream, et al. 1999; Ishihama, 1992).

La estructura de la E de *E. coli*, ha sido determinada a baja resolución (Darst, et al. 1998), y recientemente se determinó la estructura tridimensional del "core" de la RNAP de *Thermus aquaticus*, por cristalografía de rayos X a 3.3 Å de resolución, mostrando un gran parecido con los modelos existentes para la enzima de *E. coli* (Gongyi, et. al. 1999). Estos estudios revelan una estructura con forma de "tenaza de cangrejo" con un canal de 27 Å, donde al parecer se inserta una molécula de DNA de doble-cadena. Las subunidades β y β' de la RNAP constituyen dicho canal y presentan numerosos contactos con el dominio amino de la subunidad α (α NTD). Este α NTD interviene en el ensamblaje de β y β' , aunque no forma parte del sitio activo. En la pared trasera de este canal se encuentra un ion de Mg^{2+} requerido para la actividad catalítica, el cual está quelado por un motivo absolutamente invariante (NADFDGD) de la subunidad β' , presente en todas las RNA polimerasas (Gongyi, et. al. 1999).

- Estructura del Promotor

El proceso de inicio de la transcripción involucra el reconocimiento y la interacción de la $E\sigma$ con secuencias específicas presentes en el DNA denominadas promotores, localizadas antes de los genes u operones. El análisis de numerosas secuencias de promotores transcritos por la $E\sigma^{70}$ en *E. coli* permitió identificar dos secuencias hexanucleotídicas ubicadas a -35 y a -10 pares de bases arriba del inicio de la transcripción, con el consenso TTGACA y TATAAT respectivamente. Estos elementos juegan un papel determinante en la interacción con $E\sigma^{70}$, presentan un espaciamiento promedio de 17pb aunque la $E\sigma^{70}$ puede tolerar cambios en este espaciamiento de 15 a 20 pb. Existe una correlación entre la fuerza del promotor y su similitud con el consenso de las regiones -10 y -35 (Bujard, et al. 1987; Ishihama, 1988).

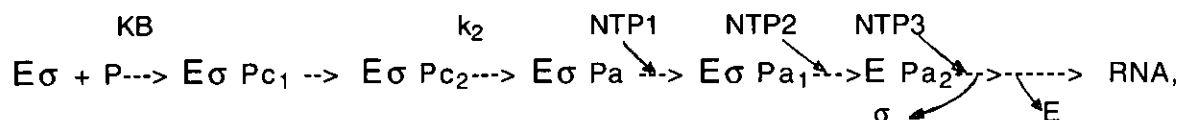
Existe un tercer elemento de reconocimiento para la $E\sigma^{70}$ en la región promotora de algunos genes denominado UP; se trata de una secuencia rica en A+T localizada de -40 a -60 pb del inicio de la transcripción. Inicialmente se describió para el promotor P1 del gen *rrnB* de *E. coli* y actualmente se ha caracterizado para numerosos genes bacterianos (Ross et al. 1993; Ross et al 1998; Tagami & Aiba, 1999). El dominio carboxilo de la subunidad α de la RNA polimerasa (α -CTD) interactúa directamente con el elemento UP, incrementando la frecuencia de inicio de la transcripción (Ross et al 1998; Tagami & Aiba, 1999).

Existe otra clase de promotores con una estructura diferente, reconocidos por el factor sigma alternativo σ^{54} . Estos promotores también presentan dos regiones consenso que se localizan a -12 (TTGCA/T) y -24 (TTGGCA) del inicio de la transcripción (Barrios, et al 1999; Thony & Hennecke, 1989). Los dinucleótidos GC y GG son prácticamente invariantes en 186 promotores de este tipo (Barrios, et al 1999). A diferencia de los promotores reconocidos por $E\sigma^{70}$, en los promotores reconocidos por $E\sigma^{54}$ existe una estricta dependencia del espaciamiento entre las regiones conservadas, y cambios de una sola base en el espaciador inactivan totalmente al promotor (Buck, 1986).

MECANISMO DEL INICIO DE LA TRANSCRIPCION

Existe una gran diversidad de mecanismos que contribuyen a la regulación del inicio de la transcripción, lo cual es consistente con la importancia que este proceso juega en el control de la expresión genética.

El mecanismo de iniciación de la síntesis de RNA involucra varios pasos que se describen en la siguiente reacción:



La holoenzima libre $E\sigma$ y el promotor P, se combinan reversiblemente para formar el complejo cerrado $E\sigma P_{c1}$, posteriormente se producen cambios

conformacionales que dan origen al intermediario $E\sigma^{Pc2}$ facilitando la separación de las cadenas del DNA. Subsecuentemente éste se isomeriza a un complejo abierto $E\sigma^{Pa}$, iniciando la síntesis del RNA. El proceso de inicio termina cuando se despega el factor σ del complejo (Gross, et.al.1992).

La afinidad de la unión de $E\sigma$ al promotor (K_B) y la eficiencia de isomerización del complejo cerrado a abierto (k_2) determinan la fuerza del promotor, estos parámetros se han determinado para varios promotores *in vitro* y el rango de valores en la mayoría es de 10^{-6} a 10^{-9} M^{-1} para K_B y de 10^{-3} a 10^{-1}sec^{-1} para k_2 (McClure, 1985; McClure & Hopes, 1987).

- Formación de complejo abierto.

La formación del complejo abierto involucra varios cambios conformacionales tanto en la $E\sigma$ como en el DNA. La molécula del DNA se abre de 12 a 14 pb en la región que va del -10 hasta el +2 respecto al inicio de la transcripción. Dadas las dimensiones de la $E\sigma$, y tomando en cuenta los cambios en las protecciones de la $E\sigma$ por DNAsa I en los diferentes pasos del inicio de la transcripción, se propone que el DNA se enrolla en la polimerasa, lo que facilita la apertura de las cadenas del DNA (Coulombe & Burton, 1999)

El factor sigma juega un papel determinante en el proceso de separación de las cadenas del DNA. Una vez que la nucleación ha ocurrido, procede la separación de las cadenas de una manera concertada a una o dos bases del sitio de inicio de la síntesis del RNA. La disposición de las cadenas en el complejo abierto maduro no está bien establecida aun, pero se propone que la $E\sigma^{70}$ estabiliza el complejo con la cadena no codificadora, previniendo el reapareamiento de las dos cadenas, mediante una barrera física (Hellman & deHaseh, 1999).

La iniciación continúa con una fase transitoria en la que se forman transcritos abortivos, pequeñas moléculas de RNA de 2-5 nucleótidos de largo que se presentan en la mayoría de los promotores transcritos por

$E\sigma^{70}$. Posteriormente se libera el factor sigma y empieza la elongación progresiva, hasta que se encuentra una señal de terminación (Gross, et al. 1992; Hellman & deHaseth, 1999)

DESCRIPCION DE LOS DIVERSOS FACTORES SIGMA

En las bacterias se sintetizan varios factores sigma, los cuales regulan la expresión genética, respondiendo a las necesidades metabólicas tanto para el mantenimiento celular, como para la adaptación a diversos estímulos ambientales. Estos factores se pueden agrupar en dos familias que presentan una estructura diferente y no están relacionadas filogenéticamente (Revisado en Gross, et al. 1992 y Wösten, 1998).

- Familia σ^{70}

En esta familia encontramos a los factores sigma relacionados en secuencia al factor σ^{70} de *E. coli*. Recientemente, la secuenciación de genomas completos de bacterias ha permitido inferir la presencia de varios factores σ pertenecientes a esta familia, los cuales según su función se pueden dividir en varios grupos; los factores σ primarios que intervienen en la transcripción de genes esenciales para el mantenimiento celular y los factores σ alternativos, que se requieren para la expresión de genes solo en condiciones especiales de crecimiento, descritos a continuación en la Tabla 2 (Wösten, 1998; Lonetto, et al 1992).

Recientemente se determinó la estructura cristalográfica de un fragmento obtenido por digestión de proteasas a 2.6 Å de resolución, revelando una forma de V. Los residuos involucrados en la interacción con el "core" caen en una cara de la estructura mientras que los residuos que interactúan con la región -10 y en la apertura del DNA se localizan en la cara opuesta (Malhotra, et al. 1996).

La comparación de estos factores a nivel de secuencia de aminoácidos ha permitido identificar cuatro regiones conservadas que se describen a continuación (figura 1A).

TABLA 2. DIVERSOS FACTORES SIGMA

GRUPO	ORGANISMO Y FACTOR σ	FUNCIONES	REFERENCIAS
<u>FACTORES PRIMARIOS</u>	<i>E. coli</i> y bacterias Gram (-): σ^{70} Bacterias Gram (+): σ^A <i>Mycobacteria</i> : MysA <i>Streptomyces</i> : HrdB	Funciones esenciales para el mantenimiento celular en fase exponencial del crecim.	Gruber & Bryant 1997
<u>FACTORES SIGMA NO-ESENCIALES</u> Factor σ de fase estacionaria	<i>Pseudomonas</i> y enterobacterias: RpoS o σ^S	Regulador global de la exp. genética en cels. expuestas a estrés hiperosmótico o ácido, y expresión en fase estacionaria.	Loewen & Hengge-Aronis 1994
Factores σ de cyanobacterias	<i>Synechococcus</i> , <i>Synechocystis</i> , <i>Anabaena</i> : SigB-E	Expresión genética durante las respuestas circadianas, y en la fase post-exponencial	Caslake, et al. 1997 Gruber & Bryant, 1997
Factores σ de bacterias Gram (+) de alto GC	<i>Mycobacteria</i> : MysB <i>Corynebacterium</i> : SigB <i>Streptomyces</i> : HrdA HrdC-E	Desconocida	Gruber & Bryant 1997
<u>FACTORES SIGMA-ALTERNATIVOS</u> Factor σ de flagelo	Enterobacteria: σ^{28} , <i>Streptomyces</i> : WhiG, <i>Bacillus subtilis</i> : σ^D	Expresión de quimiotaxis esporulación y flagelo	Helmann, 1991 Kelemen, et al 1996
Factor σ de ECF	<i>P. aeruginosa</i> : AlgU <i>E. coli</i> : Fecl <i>P. fluorescens</i> : PbrA <i>Myxococcus xanthus</i> : CarQ <i>S. coelicolor</i> , <i>E. coli</i> , <i>Mycobacteria</i> : σ^E <i>P. syringae</i> : HprL <i>Alcaligenes eutrophus</i> : CrnH <i>B. subtilis</i> : SigV-Z	Biosíntesis de alginato, utilización de fierro, biosíntesis de carotenoides, producción de antibióticos, inducción de factores de virulencia, resistencia a Co y Ni, síntesis de proteínas membranales o supervivencia a altas temperaturas	Lonetto, et al 1994
<u>FACTORES SIGMA DE ESTRES AL CALOR</u> Factor σ^{32} y factores σ -relacionados	Bacterias Gram (-): σ^{32} <i>M. xanthus</i> : SigB, <i>Streptomyces</i> : SigC.	Expresión de genes durante el estres. formación de cuerpos fructíferos maduración de mixosporas.	Yura, et al. 1993
Factores σ^B y factores relacionados	<i>B. subtilis</i> , <i>Staphylococcus</i> : σ^B , <i>Mycobacteria</i> , <i>Streptomyces</i> σ^B	Expresión durante el estres y esporulación	Hecker, et al. 1996 De Maio, et al. 1996
<u>FACTORES SIGMA DE ESPORULACION</u> Factores σ^{EFGKH}	<i>Bacillus</i> , <i>Clostridium</i> : σ^{EFGKH}	Expresión de genes de esporulación	Handelwang, 199

Región 1: Es la menos conservada, su función es modular la unión de la RNAP con el DNA (Dombroski, et al, 1992). La subregión 1.1 se encuentra sólo en los factores primarios y es importante en el inicio de la transcripción. La subregión 1.2 es importante para la formación del complejo abierto. Estudios recientes señalan que esta región está involucrada en la iniciación de la transcripción, ya que la remoción de este fragmento impide que ocurra la isomerización del complejo cerrado a abierto (Wilson & Dombroski, 1997).

Región 2: Es la región con mayor conservación entre los distintos factores sigma. Se puede dividir en cuatro subregiones; la subregión 2.1 es importante para la interacción con el "core" de la RNAP; la subregión 2.3 está involucrada en la apertura de las cadenas del DNA; la subregión 2.4 intervienen en el reconocimiento del elemento -10 del promotor (Waldburger. et al. 1990). Recientemente se identificó la región 2.5, la cual está involucrada en el reconocimiento de los nucleótidos -14 y -15 en aquellos promotores que contienen al dinucleótido TG, homólogos a la secuencia consenso -16 de los promotores de *Bacillus subtilis*, como se observa en la figura 1A (Voskuil, et al. 1995)

Región 3: Esta región se divide en dos subregiones, la 3.1 contiene un motivo de unión al DNA del tipo hélice-vuelta-hélice (HTH), ampliamente encontrado en proteínas de unión al DNA y descrito detalladamente mas adelante. La región 3.2 es menos conservada y probablemente intervenga en la unión al "core", ya que mutantes en σ^{32} que carecen de esta región presentan una deficiencia en la unión a la polimerasa, pero mantienen intactas las otras funciones de la proteína (Zhou, et al. 1992).

Región 4: Esta región participa en el reconocimiento de la región -35 del promotor, se subdivide en dos; la subregión 4.1 se propone que se une a ciertos activadores transcripcionales durante el inicio de la transcripción. El reconocimiento del elemento -35 se lleva a cabo por la subregión 4.2, lo cual se deriva de estudios con mutaciones en este elemento (Siegele, et al. 1989).

- Familia σ^{54}

El factor σ^{54} fue descrito inicialmente tanto en *E. coli* como en *Salmonella thyphimurium* por su participación en la expresión de genes que intervienen en el metabolismo nitrogenado (Hirschman, et al, 1985; Hunt & Magasanik, 1985). Actualmente se han encontrado homólogos en un gran número de especies bacterianas y se sabe que intervienen procesos celulares, por ejemplo: en la fijación de nitrógeno (en *Klebsiella pneumoniae* y *Rhizobia* entre otros), la biosíntesis del flagelo (en *Caulobacter crescentus*) y el catabolismo de xileno y tolueno (en *Pseudomonas*) (Kustu, et al. 1989).

Análisis mutacionales y la comparación de secuencias de σ^{54} , han permitido dividir esta proteína en tres regiones (figura 1B).

Región I o región de activación: Esta región está constituida por 50 aminoácidos del extremo amino terminal, es la región más conservada, rica en glutaminas y leucinas, juega un papel importante en mediar la respuesta a proteínas activadoras. Mutantes que carecen de la región I son capaces de activar en ausencia del activador, en condiciones que favorecen la separación de cadenas en el DNA (Syed & Gralla, 1998; Gallegos, et al. 1999) Actualmente se sugiere que presenta una función dual inhibiendo la isomerización en ausencia de un activador y estimulando la iniciación mediante un cambio conformacional promovido por las EBPs (Enhancer Binding Proteins)(Casaz, et al. 1999).

Región II: Esta región es variable y consta de 60 a 110 aminoácidos, presentando numerosos residuos ácidos

Región III: Comprende alrededor de 400 aminoácidos del carboxilo-terminal de la proteína, presenta dos motivos altamente conservados: el HTH y la caja RpoN, involucrados en el reconocimiento de las regiones -24 y -12, respectivamente en promotores σ^{54} dependientes (Taylor, et al. 1996)

Recientemente se obtuvo la estructura a baja resolución del factor σ^{54} por

dispersión de rayos-X en solución. Presenta una forma muy alargada, una forma de V, similar a la observada en el modelo atómico de σ^{70} , indicando que aunque contienen una secuencia primaria diferente, muestran una estructura similar (Svergum, et al. 2000).

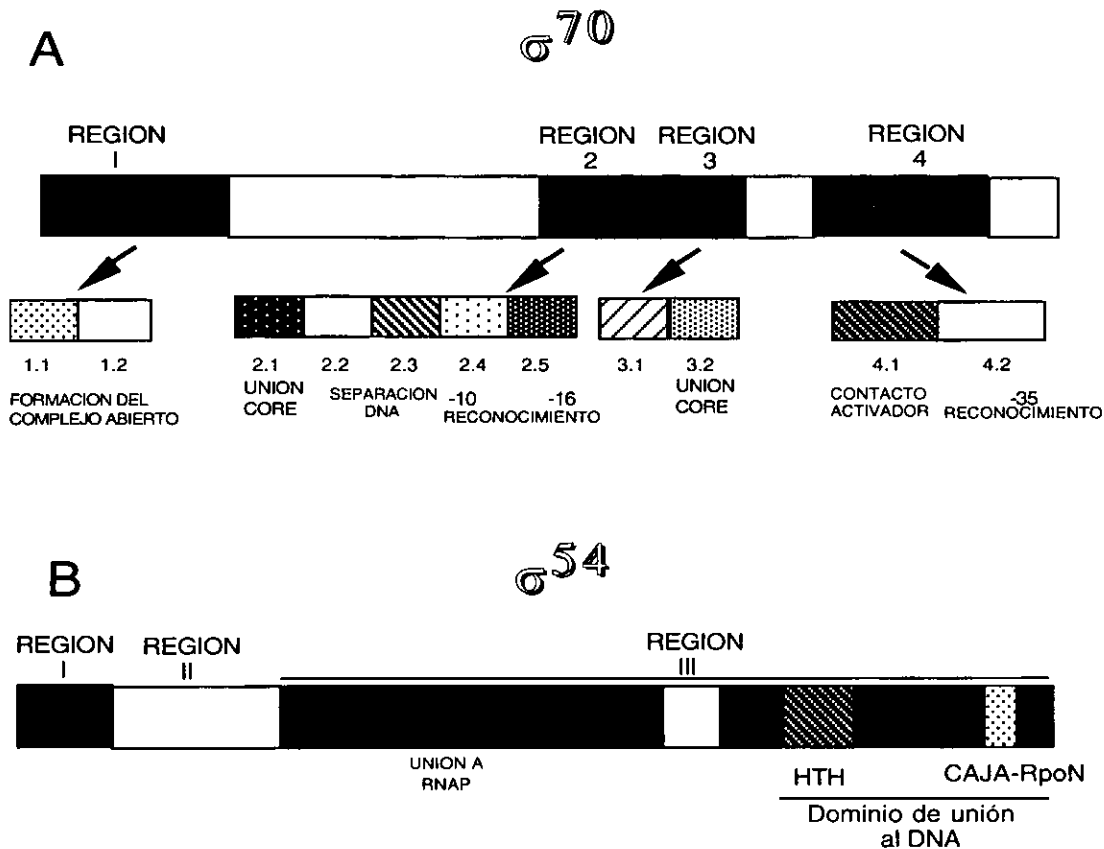


FIGURA 1. ESTRUCTURA DE FACTORES SIGMA. A) σ^{70} , B) σ^{54}

- Formación de complejo abierto en $E\sigma^{54}$

La formación del complejo abierto en promotores reconocidos por la $E\sigma^{54}$ difiere significativamente del proceso para la $E\sigma^{70}$.

La principal diferencia radica en que $E\sigma^{54}$ es capaz de formar un complejo

estable transcripcionalmente inactivo, que requiere de la hidrólisis de un nucleósido trifosfatado (NTP), para se lleve a cabo la isomerización a un complejo abierto productivo. La hidrólisis del NTP es efectuada por una proteína activadora EBP (Enhancer Binding Protein), (Weiss, et al 1991; Austin & Dixon, 1992).

Los modelos actuales proponen la formación de un complejo pre-inicio, en el que la $E\sigma^{54}$ se encuentra establemente unida al DNA abierto, pero es incapaz de iniciar la transcripción. Se requiere de un cambio conformacional en la holoenzima, el cual es promovido por la EBP (Wang & Gralla, 1996). Actualmente se ha determinado que el amino terminal de σ^{54} (Región I), contiene los determinantes que permiten la interacción con las EBPs e inhiben la iniciación en ausencia de éstos (Casaz, et al. 1999; Syed & Gralla, 1999).

REGULACION DEL INICIO DE LA TRANSCRIPCION EN BACTERIAS

Los genes son expresados solamente cuando se necesitan y en el nivel requerido por las células. Para que se expresen de manera adecuada existen numerosos mecanismos que le permiten a la célula mantener la armonía y enfrentar los diversos cambios ambientales. Estos mecanismos son parte de una red regulatoria que asegura la máxima eficiencia con el mínimo gasto de energía para mantener la homeostasis en la célula.

Existen varios elementos que intervienen en la regulación del inicio de la transcripción. En promotores constitutivos, la eficiencia de la transcripción depende únicamente de la concentración de $E\sigma$ libre, la estructura y secuencia del promotor, la topología del DNA y factores que se unen directamente a la $E\sigma^{70}$, como el nucleótido guanosa tetrafosfatada (ppGpp), que en niveles altos lleva a una reducción dramática de la $E\sigma$ activa. (Liang, et al 1999; Glass, et al. 1996). En promotores no constitutivos, la regulación de la transcripción involucra el contacto directo de la RNAP con proteínas que se unen a sitios adyacentes o sobrelapados a los reconocidos por la polimerasa para el inicio de la transcripción y que pueden activar o reprimir la expresión de genes, proporcionando el nivel requerido para las diferentes condiciones metabólicas celulares (McClure, 1985)

El estudio de 150 promotores de *E. coli*, reconocidos por el factor σ^{70} , muestra las siguientes características; la mayoría involucra en su regulación una sola proteína reguladora, generalmente un represor. Los represores se unen principalmente a regiones que se sobrelapan con los sitios de unión a la $E\sigma^{70}$. Los activadores se unen generalmente a la región localizada entre -30 y -80 pb con respecto al inicio de la transcripción. Muchos de estos activadores también pueden comportarse como represores sobre ciertos promotores, si la unión se da en regiones abajo del -30, interfiriendo con la unión de $E\sigma^{70}$. Por otro lado, se ha observado que cuando los represores o activadores se unen en regiones distales al promotor, generalmente actúan en conjunto con sitios adicionales localizados en regiones adyacentes a las regiones de unión a $E\sigma^{70}$ (Gralla & Collado-Vides, 1996)

REGULACION NEGATIVA (Represión)

En este caso una proteína con la capacidad de unirse al DNA, bloquea el sitio de unión de la $E\sigma$ (promotor), impidiendo que se lleve a cabo el inicio de la transcripción. La actividad de unión al DNA del represor puede ser modulada por la unión de una pequeña molécula denominada efector.

Actualmente contamos con numerosos ejemplos que describen este tipo de regulación. Sin embargo, se mencionará la represión del operón de lactosa por ser el primer modelo de regulación descrito de este tipo (Majors, 1975).

El operón de lactosa (*lac*) consta de tres genes que se expresan a un nivel alto solo cuando las células están creciendo en un medio que contiene lactosa. El gene *lacI* codifica para el represor el cual inhibe el inicio de la transcripción más de 1000 veces. Este operón contiene tres operadores, siendo el más importante el O_1 u operador primario (localizado a +11), los otros dos pseudo-operadores son O_2 (+401) y O_3 (-82). El operador O_1 , actúa cooperativamente con los otros dos operadores. El complejo *LacI*-operador es desestabilizado por la unión de compuestos galactósidos a *LacI*, permitiendo que la polimerasa inicie la transcripción de este operón en presencia de estos compuestos. Los detalles moleculares del sistema aun

continúan siendo estudiados (Choy & Adhya, 1996; Gralla, 1992).

REGULACION POSITIVA (Activación)

Numerosos genes bacterianos son regulados positivamente por proteínas activadoras que se unen a regiones adyacentes al promotor. Se pueden distinguir dos tipos; los que interactúan con la subunidad α de la E (α -CTD) y los que lo hacen con el factor σ . La activación tiene como consecuencia la modificación de las constantes cinéticas de la transcripción (k_2 y K_B , ver mecanismo), incrementando su eficiencia (Busby & Ebright, 1994)

Dentro de el primer tipo, el activador mejor caracterizado que interactúa con α -CTD en *E. coli* es CRP (proteína receptora de AMP cíclico), esta proteína dimérica activa la transcripción de más de 100 promotores diferentes, uniéndose a los operadores en presencia de AMPc (Busby, 1999).

Los promotores dependientes de CRP se han dividido en dos clases, dependiendo de su sitio de unión con respecto al inicio de la transcripción. En la clase I, el sitio de unión de CRP se localiza aproximadamente a -61.5 pb del inicio de la transcripción, un ejemplo es la activación de *lac*. En la clase II, el sitio de reconocimiento para CRP se localiza aproximadamente en la posición -41.5 pb, sobrelapando la región -35 y un ejemplo es la regulación del promotor *gal* (Busby & Ebright, 1997; Savery, et al. 1998) La combinación de los mecanismos de activación de la transcripción presentes en la clase I y II origina un sinergismo dado por dos activadores, por ejemplo en la regulación del promotor *ansB*, intervienen CRP y FNR (Regulador de los genes de anaerobiosis). FNR se une a la posición -41.5 y contacta σ y CRP se une a la posición -91.5 contactando al α -CTD (Scott, et al. 1995; Busby & Ebright, 1999)

El segundo tipo de activadores interactúa con σ^{70} . Este tipo de activadores se unen generalmente a sitios sobrelapados al elemento -35 del promotor y el ejemplo mejor estudiado es la proteína cI del bacteriófago λ , la cual se une al sitio -42 del promotor λP_{RM} e interactúa con el factor σ^{70} .

Mutantes supresoras en σ^{70} han permitido identificar a la región 4 de σ^{70} como la región de contacto (Li, et al 1997).

REGULACION A DISTANCIA

Los mecanismos de activación descritos anteriormente implican la unión del activador a regiones adyacentes al promotor. Sin embargo, existen ejemplos de interacciones con proteínas que se unen a regiones alejadas del promotor, tanto para represión como para activación. La regulación del operón de arabinosa por AraC es el paradigma de este tipo de regulación. AraC es una proteína bifuncional que regula entre otros, al operon *araBAD*, que contiene a los genes cuyos productos son necesarios para la utilización de arabinosa. AraC se encuentra en equilibrio conformacional entre la forma de activador o represor. En ausencia de arabinosa, el equilibrio se desplaza a la forma de represor en la cual se favorece la unión a los sitios *araO2* y *araI1* formando una asa en el DNA. Esta configuración previene la iniciación de la transcripción en el promotor. Cuando el ligando (arabinosa) esta presente, el equilibrio se desplaza hacia la activación y entonces se une a sitios adyacentes *araI1* y *araI2*. (Gallegos, et al. 1997)

Otro mecanismo de regulación de la transcripción a distancia es el caso de los genes transcritos por $E\sigma^{54}$ y las proteínas activadoras EBPs. Las EBPs se unen a sitios localizados entre 80 y 140 pb arriba del inicio de la transcripción denominados UAS (Upstream Activation Sequences) (Morett & Buck,1988). Las UAS presentan características similares a los "enhancers" descritos en eucariontes, ya que pueden moverse varios miles de nucleótidos arriba del promotor independientemente de su orientación y seguir siendo activos (Buck et al,1986; Kustu et el, 1989).

MODELO DE ACTIVACION A DISTANCIA para $E\sigma^{54}$

Este modelo propone que las EBP's se unen a las UAS y hacen contacto simultáneamente con $E\sigma^{54}$ que se encuentra unida al promotor, formando un "loop" en el DNA, el cual es favorecido por el factor de integración IHF. Este factor se une a una región del DNA entre las UAS y el promotor

generando un doblado de 140°, lo que favorece los contactos entre la proteína activadora y la polimerasa, como se observa en la figura 2 (Kustu, et al, 1991; Su et al, 1990). La $E\sigma^{54}$ forma un complejo cerrado que es isomerizado a un complejo abierto, mediante la hidrólisis de ATP que lleva a cabo la EBP correspondiente.

Como se mencionó anteriormente muchas especies bacterianas, incluyendo a *E. coli*, tienen múltiples EBP's que intervienen en la regulación de diferentes procesos celulares; por ejemplo, en la asimilación de nitrógeno, la proteína NtrC; en el transporte de ácidos dicarboxílicos DctD; en la oxidación de hidrógeno HoxA; en la degradación de compuestos aromáticos XylR, por citar algunos ejemplos. (Kustu et al, 1989). Estas proteínas constituyen una familia con características estructurales similares, las cuales serán discutidas más adelante.

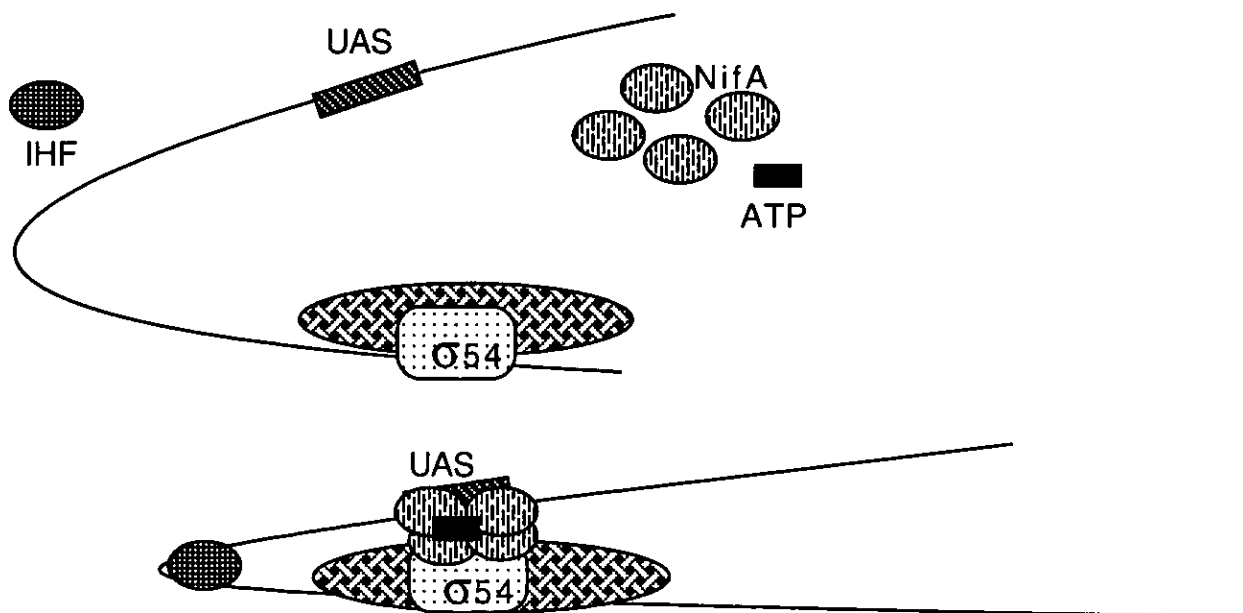


FIGURA 2. MODELO DE ACTIVACION A DISTANCIA DEL PROMOTOR *nifH* DE *K. pneumoniae*

ANTECEDENTES

La fijación de nitrógeno es un proceso que demanda un alto gasto de energía; la nitrogenasa es inactivada irreversiblemente por oxígeno por lo que los organismos diazótrofos han desarrollado mecanismos de control estrictos que garantizan su expresión sólo en las condiciones ambientales apropiadas (Merrick, 1992).

El regulador positivo NifA interviene en la expresión de los genes *nif* y *fix*, responsables de la fijación de nitrógeno. Estudios comparativos entre los genes *nif* de diferentes especies sugieren que hay un grupo común de 14 genes, necesarios para la biosíntesis de la nitrogenasa (*nifH, D, K, E, N, X, U, S, V, Z, W, M, B, Q*). En los *Rhizobia*, bacterias que interactúan con las raíces de plantas leguminosas, existe otro grupo de genes asociados a la fijación de nitrógeno denominados *fix* que no tienen homólogos en los diazótrofos de vida libre (Revisado en Merrick, 1997).

Los genes *nif* y *fix* se encuentran agrupados en varios operones cuyos promotores generalmente son reconocidos por $E\sigma^{54}$ y activados por NifA. En el caso de los organismos diazótrofos de vida libre, la cantidad de compuestos nitrogenados y oxígeno en el medio, son las señales que regulan la actividad de NifA, a diferencia de los organismos simbiotes, en los que la concentración de oxígeno celular es la principal señal de regulación para la actividad de esta proteína (Revisado en Fisher, 1994).

La proteína NifA pertenece a la familia de activadores transcripcionales que actúan a distancia denominados EBPs. Esta familia de proteínas presenta una estructura modular característica, que se describirá a continuación. Entre las EBPs de las que se tiene mayor información se encuentran NifA y NtrC (Buck et al, 1986; Kustu et al, 1989).

ESTRUCTURA DE NifA

La proteína NifA y en general todas las EBPs, están formadas por tres dominios: un dominio amino terminal (N-terminal), cuya actividad se relaciona con la recepción de señales; un dominio central que interviene

directamente en el control positivo y en la interacción con σ^{54} y un dominio carboxilo terminal (C-terminal) que presenta la actividad de unión al DNA. Existen evidencias bioquímicas y genéticas que indican que estos dominios son funcional y estructuralmente independientes (Drummond, et al., 1986; Drummond et al 1990). En el caso de NifA y DctD de *S. meliloti*, el dominio central es suficiente para activar la transcripción (Huala & Ausubel, 1989; Huala et al, 1992), mientras que el dominio C-terminal aislado de *K. pneumoniae*, retiene la capacidad de unirse específicamente al DNA (Morett, et al. 1988, Lee, et al. 1994).

- **Dominio N-terminal:** Es el que presenta mayor variabilidad, tanto en tamaño como en función. Un subgrupo formado por proteínas como NtrC, DctD, HydG, AlgB, HoxA y FlbD, este dominio es característico de las proteínas reguladoras de la familia de “dos componentes”, que intervienen en la transducción de señales (Morett & Segovia, 1993).

El dominio N-terminal de la proteína NifA es de tamaño variable (164-216 aa) y al parecer es dispensable para la función de activación, como lo sugieren estudios con una mutante de *S. meliloti* que carece de este dominio y conserva la capacidad de activación, además del hecho de que no se presenta naturalmente en NifA de *R. leguminosarum* bv *trifolii* (Huala & Ausubel, 1989; Iismaa & Watson, 1989). El N-terminal de NifA de *K. pneumoniae*, *A. vinelandii* interacciona con la proteína sensora NifL en respuesta a las condiciones de nitrógeno fijado y oxígeno celular (Berger et al 1994; Blanco, et al 1993). Actualmente se estudia el papel que tiene en la regulación de la actividad por compuestos nitrogenados en la NifA de *Azospirillum brasilense* y *Herbaspirillum seropididae* (Monteiro, et al. 1999, Arsene, et al. 1999).

- **Dominio central:** El dominio N-terminal y el central se encuentran separados por un interdominio conector constituido de numerosos residuos de glutamina. El dominio central está constituido por 240 aminoácidos y es común en todas las proteínas de la familia de las EBPs. Es el dominio que muestra la mayor similitud a nivel de aminoácidos entre las EBPs, por ejemplo: entre las NifAs de *S. meliloti*, *B. japonicum*, *A. caulinodans* y *K. pneumoniae* existe una identidad del 53-72%. Este dominio aislado, es capaz de promover la activación (Huala y Ausubel, 1989). Presenta el

motivo (GXXXGK) característico de unión a NTP's, conocido como "Walker A", que se presenta en los plegamientos típicos de unión a monucleótido. La hidrólisis de un NTP aporta energía necesaria para isomerizar el complejo de promotor cerrado a un complejo abierto productivo, en coordinación con la $E\sigma^{54}$ (Weiss et al,1991)

Recientemente se propuso un modelo para este dominio basado diversos parámetros como son; la predicción de estructura secundaria, mutaciones correlativas descritas en diferentes EBPs y algoritmos que permiten la relación de plegamientos con estructuras comunes (Osuna, et al 1997). El modelo propuesto por Osuna y cols. sugiere que el dominio central adopta un plegamiento similar al dominio G del factor de elongación bacteriano EF-Tu y a la proteína de señalización p21^{ras}, estas proteínas pertenecen a la familia de las GTPasas que llevan a cabo la hidrólisis de mononucleótidos en sus función. El modelo se basó en la estructura cristalográfica de EF-Tu que consiste en un polipéptido que adopta una topología α/β . Actualmente existen datos experimentales que apoyan este modelo (Rombel, et al 1999).

Entre el dominio central y el dominio C-terminal se presenta un conector de tamaño variable en las NifAs. En las NifAs de *K. pneumoniae* y *A. vinelandii* el dominio central se encuentra directamente unido al dominio C-terminal, a diferencia de las NifAs de *Azospirillum brasilense*, *Rhodobacter capsulatus* y *Rhizobia* (sensibles *per se* a oxígeno) en las que estos dos dominios se encuentran separados por un conector de 32-44 aa. En este conector se pueden distinguir dos cisteínas totalmente conservadas, que junto con otras dos presentes al final del dominio central, conforman el motivo C-X₁₁-C-X₁₉-C-X₄-C propuesto como un probable motivo de unión a algún cofactor metálico.

- **Dominio C-terminal:** Consta de 65 a 130 residuos de aminoácidos, presenta un motivo altamente conservado de unión al DNA con la estructura de hélice-vuelta-hélice (HTH) (Morett & Segovia, 1993). Este dominio puede ser expresado independientemente, reteniendo la capacidad de unirse específicamente a las UAS (Lee et al 1994).

MOTIVO DE HELICE-VUELTA-HELICE

Como ya se mencionó en la introducción, la activación de la transcripción requiere de la formación de un complejo entre la proteína regulatoria y una secuencia específica de DNA. Existen varios motivos estructurales que cumplen con esta función y se agrupan en diversas familias. Entre los motivos mejor estudiados se pueden mencionar los “dedos” de zinc, cierre de leucinas, los homeodominios, la hélice-vuelta-hélice y la hélice-loop-hélice. (Revisadas en Pabo & Sauer, 1992). La existencia de distintas familias demuestra que hay múltiples soluciones para el problema estructural de unirse al DNA (Brennan & Mathews, 1995). Sin embargo, cada motivo que ha sido caracterizado involucra una estructura secundaria simple (generalmente una α -hélice), que presenta cierta complementaridad con la estructura de DNA-B.

Reconocimiento DNA-proteína.

La mayoría de las proteínas de unión al DNA usan α -hélices para hacer contacto con las bases del DNA en el surco mayor (Pabo & Sauer, 1992).

Con base en las estructuras disponibles, varios autores han tratado de elaborar un código de reconocimiento. Sin embargo, aunque hay algunas reglas claras, es importante tomar en cuenta las interacciones que el motivo de unión al DNA tiene con el resto de la proteína

El primer motivo de unión al DNA descrito fue el HTH. Numerosas proteínas que lo contienen han sido cristalizadas incluso como complejos en interacción con el DNA. Las primeras estructuras del HTH fueron deducidas a partir de cristales de la proteína CRP de *E.coli*, el represor del fago λ y la proteína Cro del mismo fago, donde se observó que la HTH está formada por dos α -hélices interrumpidas por una vuelta (Revisado en Pabo & Sauer, 1992). El residuo más conservado en este motivo es una glicina en la posición 10 (G10) al inicio de la vuelta que, como se observa en la figura 3, se presenta invariante en las proteínas de las que se ha determinado la estructura en interacción con el DNA (Brennan & Mathews, 1988; Suzuki, et al, 1995a). Cabe mencionar que en las EBPs este residuo también se encuentra muy conservado. La presencia de esta glicina se ha tratado de explicar en términos estructurales y se piensa que es indispensable para

romper la conformación de la α -hélice (con giro a la izquierda) presente en el motivo HTH (Baumeister, et al. 1992).

RCRO_LAMBD	GQTKTAKDL@VYQSAINKAIH	6CRO
RCRO_BP434	TQTELATKA@VKQOSIQLIEA	3CRO
RPC1_BP434	SQESVADKM@MGQSGVGALFN	1LMB
TRPR_ECOLI	SQRELKNEL@AGIATITRGSN	1TRO
LACT_ECOLI	TLYDVAEYA@VSYQTVSRVFN	1LCC
CRP_ECOLI	TRQEIGQIV@CSRETVGRILK	1CGP
RPC1_BP434	NQAELAQKV@TTQQSIEQLEN	2ORI
HMEN_DROME	RRQQLSSEL@LNEAQIKIWFQ	1HDD
DNIV_SALTY	PRQQLAIIIF@IGVSTLYRYFP	1HCR

10

FIGURA 3. Secuencia de HTH de proteínas cocristalizadas con el DNA. Las abreviaturas de lado izquierdo corresponden a la base de datos Swiss Prot y las abreviaturas del lado derecho corresponden al PDB (Protein Data Bank). 6CRO (lambda Cro), 3CRO (434 Cro), 1LMB (Represor de Lambda), 1TRO (represor *trp*), 1LCC (represor *lac*), 1CGP (CAP), 2ORI (represor 434), 1HDD (Homeoproteína *Drosophila*), 1HCR (recombinasa Hin).

De las dos hélices que componen el motivo HTH, una es de reconocimiento que da la especificidad y la otra es parte de la estructura. La llamada hélice de reconocimiento está formada por varios residuos entre los que encontramos los que contactan las bases del DNA. Algunos de estos residuos contactan los fosfatos y otros no hacen contacto directo con el DNA, pero interactúan con el resto de la proteína, limitando el ángulo de rotación de la hélice. (Suzuki, et al, 1995b). En el caso de las EBPs la hélice de reconocimiento interactúa específicamente con las UAS y es la más variable, esto corresponde al hecho de que las diferentes EBPs se unen a diferentes UAS (Jacob, J. & Drummond, M. 1993).

RELACIONES FILOGENETICAS DE LOS DOMINIOS DE LAS EBP'S

El análisis de las secuencias que componen a esta familia permitió hacer un estudio filogenético de sus dominios (Morett & Segovia, 1993). El estudio incluyó 13 diferentes proteínas activadoras, incluyendo a 20 especies distintas de organismos. Este análisis reveló que el dominio central y el C-terminal tuvieron un origen común para toda la familia, a diferencia del dominio N-terminal. Solo en un subgrupo de EBPs, las que

pertencen a su vez a la familia de proteínas activadoras de dos componentes, se observa un origen común. En general no se observa un intercambio de dominios dentro de las EBPs, al parecer estos dominios han evolucionado juntos, y muestran la divergencia natural de las especies y grupos bacterianos a los que pertenecen. El hecho de que existan EBPs en proteobacteria, firmibacteria y aún en un organismo que divergió muy temprano en la evolución como es *Aquifex aeolicus*, indican que el sistema σ^{54} es parte de un sistema regulatorio ancestral (Morett & Segovia, 1993)

REGULACION DE LA EXPRESION DE *nifA*

La expresión de *nifA* fue descrita inicialmente en *K. pneumoniae*, diazótrofo facultativo de vida libre. Actualmente se cuenta con la secuencia de numerosos genes *nifA* pertenecientes a diferentes diazótrofos. Entre los que encontramos varias especies tanto de organismos endosimbiontes como de organismos de vida libre.

En *K. pneumoniae*, la expresión de *nifA* está regulada a dos niveles; tanto por la concentración de oxígeno celular, como por los niveles de nitrógeno combinado en la célula. La regulación de la concentración de nitrógeno celular en *K. pneumoniae* forma parte del sistema *ntr* y controla la expresión de genes que intervienen en el metabolismo de nitrógeno. Este sistema está conformado por los productos de los genes *glnD*, *glnB*, *glnK*, *ntrB* y *ntrC*. Los tres primeros codifican para la uridiltransferasa (UTasa), la proteínas efectoras PII y la proteína GlnK, respectivamente, mientras que los dos últimos codifican para el par NtrB y NtrC, proteínas del sistema de dos componentes. En esta bacteria, el gene *nifA* es transcrito coordinadamente con *nifL*, en el operón *nifLA* el cual presenta un promotor σ^{54} -dependiente y es activado por NtrC (Merrick & Stewart, 1985; Govantes, et al. 1996). En condiciones limitantes de nitrógeno combinado, la Utasa promueve la uridilación de PII, la cual a su vez promueve la autofosforilación de NtrB y posteriormente NtrC es fosforilado por NtrB, produciendo una forma activa necesaria para la transcripción. En exceso de amonio, la cascada regulatoria se revierte y NtrB promueve la defosforilación de NtrC, disminuyendo su actividad e impidiendo la isomerización del complejo del promotor *nifLA* cerrado a abierto.

Recientemente se sabe que la proteína GlnK también promueve la fosforilación de NtrB, lo que añade complejidad al sistema de regulación en estas bacterias. El promotor de *nifLA* contiene dos sitios de unión a NtrC, localizados -120 a -140 pb arriba del promotor -24 /-12. Estos sitios no son completamente idénticos a la secuencia consenso de unión a NtrC y presentan una baja afinidad, por lo que sólo en condiciones severas de limitación de nitrógeno se expresan los genes *nif* y se puede llevar a cabo fijación de nitrógeno (Austin et al, 1990).

Como se mencionó anteriormente, la expresión de *nifLA* en *K. pneumoniae* también responde a la concentración de oxígeno celular. La regulación está mediada por cambios en la topología del DNA. En anaerobiosis, condición en la que el DNA se encuentra superenrollado negativamente, la expresión del promotor *nifLA* se incrementa 100 veces (Whitehall et al, 1992; Hill et al, 1985).

En *S. meliloti*, *B. japonicum* y *A. caulinodans*, organismos diazótrofos endosimbiontes, tanto la expresión de *nifA* como la actividad de la proteína, están reguladas principalmente por la concentración de oxígeno celular. En *S. meliloti*, *nifA* se encuentra abajo del operón *fixABCX*, presentando dos promotores; uno σ^{54} dependiente, arriba de *fixABCX* y el otro entre *fixABCX* y *nifA* que es reconocido por σ^{54} . El primero se autorregula por NifA y el segundo se regula mediante un mecanismo que involucra los productos de los genes *fixL* y *fixJ*. FixL es una hemoproteína de membrana que tiene la capacidad de autofosforilarse en condiciones de baja concentración de oxígeno, posteriormente transduce la señal a FixJ fosforilándola, lo que permite activar transcripcionalmente la expresión del segundo promotor *pnifA*, expresándose NifA en mayor cantidad y autorregulando así a todo el operón (Reyrat et al, 1992). Además de esta autoregulación positiva, la expresión de *nifA* está regulada negativamente por FixK, previniendo la sobreexpresión de *nifA* (Fisher, 1994).

En *B. japonicum* la regulación de la expresión de *nifA* es también compleja. El operón *fixRnifA* es transcrito desde promotores sobrelapados reconocidos por dos diferentes $E\sigma$, que actúan dependiendo de las condiciones celulares de oxígeno. El operón *fixRnifA*, se expresa a un nivel

basal en condiciones aeróbicas a partir del promotor *fixRp2*, el cual es reconocido por una $E\sigma$ aun no determinada; actualmente se sabe que la expresión en estas condiciones depende de otro regulador identificado como RegR (Bauer, et al. 1998). En condiciones anaeróbicas la expresión de *fixRnifA* se incrementa 5 veces y lo hace a partir del promotor *fixRp1*, el cual es reconocido por $E\sigma^{54}$ y autorregulado por NifA (Barrios, et al, 1995). En este organismo también intervienen las proteínas FixLJ, sensando y transduciendo la señal dependiendo de la concentración de oxígeno celular; a bajas concentraciones FixJ activa la expresión de *fixK*. FixK a su vez interviene en la expresión de varios genes, entre ellos una copia de σ^{54} , el cual interactúa con NifA para transcribir los genes *nif* y *fix* e incluso autorregulando su expresión (Revisado en Fisher,1994).

REGULACION DE LA ACTIVIDAD DE NifA

La actividad de NifA también está sujeta a la regulación por oxígeno, en base al mecanismo regulatorio se pueden diferenciar tres tipos de NifA: Las NifAs en las que la regulación por oxígeno y amonio depende de la proteína sensora NifL; las que son intrínsecamente sensibles a oxígeno y que contienen un posible motivo de unión a metales (C-X₁₁-C-X₁₉-C-X₄-C) y las que además de contener este motivo de cisteínas, responden a la concentración de amonio a través de su dominio amino terminal, descritas con más detalle a continuación.

En *K. pneumoniae*, *A. vinelandii* y *Enterobacter agglomerans* la actividad de NifA es regulada mediante la proteína sensora NifL, la cual inhibe la actividad de NifA en respuesta a altas concentraciones de nitrógeno fijado y oxígeno celular. Como ya se mencionó, los genes *nifA* y *nifL* forman un operón cuya expresión se encuentra acoplada traduccionalmente, asegurando la presencia de cantidades estequiométricas de las dos proteínas (Govantes, et al. 1996). Aunque el mecanismo de inhibición no ha sido elucidado completamente, se sabe que NifL interacciona con NifA formando el complejo NifLA, que interfiere en el contacto de NifA con $E\sigma^{54}$. El mecanismo que controla la actividad de NifL se explicará más adelante, en esta misma sección (Austin, et al., 1994; Dixon, 1998)

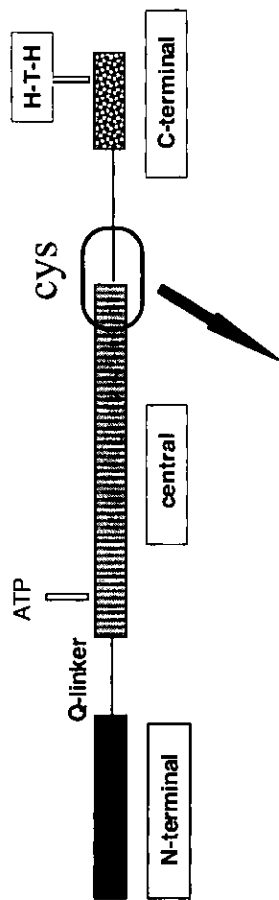
La actividad de la proteína NifA de *Herbaspirillum seropedicae* y

Azospirillum brasilense responde directamente a la concentración celular de oxígeno y compuestos nitrogenados, mediante un mecanismo en el que no participa NifL, al parecer son dos mecanismos independientes, que actúan concertadamente regulando la actividad de estas proteínas (Souza, et al., 1991; Arsene, et al., 1999).

En la regulación por amonio intracelular y glutamina de este tipo de NifA, interviene la proteína sensora PII, componente de una cascada que sensa los niveles de nitrógeno celular (Bennelli, et al 1997) y el dominio N-terminal de la proteína. Las proteínas que carecen del N-terminal se vuelven insensibles a la regulación por amonio y estudios recientes muestran que al expresar el dominio N-terminal de *H. seropedicae* en *E. coli*, se reestablece el control negativo de amonio en mutantes de NifA truncadas que carecen del N-terminal (Monteiro, et al. 1999, Arsene, et al. 1999). En la regulación por oxígeno de este tipo de proteínas interviene el motivo de cisteínas también presente en las NifAs de los *Rhizobia*, descrito con más detalle a continuación.

Sinorhizobium, *Bradyrhizobium* y *Azorhizobium* entre otras, presentan el tercer tipo de NifA y no tienen NifL, estas proteínas son sensibles intrínsecamente a las concentraciones de oxígeno celular. El alineamiento de las secuencias de aminoácidos de los diferentes tipos de NifAs, revela una diferencia en el dominio interconector (central-C-terminal). Solamente las proteínas sensibles a oxígeno *per se* presentan de 36-44 aminoácidos extras, entre los que se pueden identificar dos de las cuatro cisteínas propuestas como un posible motivo de unión a metales, con el arreglo C-X₁₁-C-X₁₉-C-X₄-C, como se observa en la figura 4 (Fischer et al, 1988).

Los datos anteriores concuerdan con el hecho de que la actividad de las NifAs de *B. japonicum*, *R. trifolii* y *S. meliloti*, a diferencia de la NifA de *K. pneumoniae*, son dependientes de iones metálicos. En estudios *in vivo*, la adición de agentes quelantes al medio dió como resultado la inhibición de la actividad, este efecto puede revertirse con la adición de algunos iones metálicos divalentes (Fischer et al, 1988; Screen, et al 1994).



Rc	RCQFPGNERELENCVNRAAALSDGAIVLAEELACRQGA CL SAEL----FRLQDGTSPIGGLAVGRVIT
Ac	RCYFPGNVRELENCIRRTATLAHDAVITPHDFACDSGQCL SAM LWKSAPKPVMPNVPAPTPLTPLS
Hs	NCYWPGNVRELENCVERTATMMRGDLITEVHFS CQ QNKCLTKVL-----HEPGQQQPVV
Bj	SCKPFGNVRELENCIERATLSAGT SIV RSDFAC SQ QQLSTTLWKSTSYGKTDPAAPMQPVPAKSII
Rm	KCKFPGNVRELENCVRRRTATLARSKTITSSDFACQTDQCFSSRLWKGVHCSHGHIIEI-DAPAGTTPLL
Rl	KCAFPGNVRELDNCVQRTATLASSNFTSSDFACQDQ CS SALLRKADGDGIGNDAM-NGLNSRD TMS
Rt	NCEFPGNIRELQ NC TQRTATLARSDVIVPQVLACEQGR CYS PILKKAVAEQVKGAIHGLARGETESM
Av	SHRWPGNVRELENC LE RSAINSE DT ITR-----DVVS
Kp	EYSWPGNVRELENC LE RSVLS ES GLIDR-----DVIL

Figura 4. Alineamiento del domini COOH terminal de diferentes NifA's. Los organismos dentro del recuadro presentan NifA's sensibles perse. (Rc, Ac, Hs, Bj, Rm y Rt), presentan las 4 cisteinas involucradas en la regulacion por oxigeno, a diferencia de *Klebsiella pneumoniae* (Kp) y *Azotobacter vinelandii* (Av).

La importancia del motivo de cisteínas en la actividad de NifA, se demostró en estudios realizados con la NifA de *B.japonicum*. El reemplazo de cada una de las cisteínas por otro aminoácido da como resultado proteínas inactivas. Además, se determinó que la distancia entre éstas, más no la naturaleza de los aminoácidos que las separan también es fundamental para la actividad de la proteína (Fischer, et al. 1987; Fischer, et al. 1988).

Por otro lado, en estudios *in vivo* con la NifA de *B. japonicum*, se observó que la concentración de oxígeno celular juega un papel importante en la capacidad de unión al DNA. En estos estudios al elevar la concentración de oxígeno en el medio, la NifA deja de unirse al DNA y la transcripción del gene *nifD* se abate, poco después la proteína se degrada rápidamente. Estos resultados se han interpretado como un cambio en el estado redox de un cofactor metálico, lo que afecta la conformación de la proteína favoreciendo su degradación (Morett, et al. 1991).

Los estudios descritos anteriormente demuestran que hay un claro efecto del oxígeno sobre la actividad de NifA, sin embargo, actualmente se desconoce el mecanismo que regula la actividad de este tipo de proteínas.

El estudio de los mecanismos regulatorios capaces de responder a diferentes concentraciones de oxígeno en otras proteínas, nos puede dar indicios del mecanismo empleado en este tipo de NifA, por lo que se describirán a continuación los más relevantes.

OTROS REGULADORES TRANSCRIPCIONALES REGULADOS POR OXIGENO

Las bacterias presentan respuestas adaptativas a las transiciones entre el crecimiento aeróbico al anaeróbico. Estas respuestas son mediadas por la regulación coordinada de grupos de genes, controladas generalmente por un regulador global, que interviene en procesos metabólicos diversos como la respiración, la fotosíntesis o la fijación de nitrógeno.

Las proteínas regulatorias que responden a cambios en la concentración de oxígeno celular o agentes oxidativos son denominados "reguladores redox".

Estos reguladores muestran diferentes sensibilidades al oxígeno o a las especies reactivas de oxígeno, lo cual se regula mediante un control diferencial en respuesta a las diferentes necesidades metabólicas (O'Halloran, 1993). Existen varios mecanismos moleculares empleados por estos reguladores para sensar la concentración celular de oxígeno; la aparición de especies reactivas de oxígeno o cambios en el estado redox de la célula.

A) Centros Fierro-Azufre

El regulón SoxR/SoxS interviene en la expresión de genes que se requieren durante el estrés oxidativo, causado por especies reactivas de oxígeno como el superóxido ($O_2\cdot^-$). La inducción de los genes para la defensa oxidativa involucra la activación en dos pasos. El sensor SoxR es un homodímero en solución que contiene un centro fierro-azufre [2Fe-2S], el cual en condiciones oxidadas promueve la expresión de *soxS*. SoxS activa directamente los genes de respuesta al estrés-oxidativo, por ejemplo la superóxido dismutasa y la catalasa (Revisado en Hidalgo, et al, 1995).

La oxido-reducción del centro fierro-azufre es reversible y solo en condiciones oxidadas, estimula a la $E\sigma^{70}$ para formar complejos de promotor abiertos y productivos, aunque en condiciones reducidas e incluso como apo-SoxR, mantiene la capacidad de unirse al DNA, sin promover la transcripción de *soxS*, como lo indican estudios de "footprinting" (Hidalgo & Demple, 1994).

El regulador FNR, también posee un centro fierro-azufre del tipo [4Fe-4S], que sensa el estado redox de la célula. En *E. coli*, promueve la expresión de 120 genes necesarios para el metabolismo anaeróbico. Por ejemplo, activa la expresión de enzimas respiratorias, que emplean aceptores de electrones alternativos y actúa como represor en la expresión de enzimas de la respiración aeróbica (Spiro & Guest, 1994).

Los análisis *in vitro* indican que el centro [4Fe-4S] es muy lábil y que rápidamente se desensambla cuando es expuesto al oxígeno. Estudios *in vivo* demuestran que el cambio de condiciones aeróbicas a anaeróbicas

provocan un cambio del centro [4Fe-4S] a uno del tipo [2Fe-2S]. La estructura también cambia de un homodímero con alta afinidad de unión al DNA, a un monómero con una baja afinidad. Esta conversión oxidativa es reversible (Khoroshilova, et al. 1995; Koroshilova, et al. 1997).

B) Grupos Hemo

FixL es una proteína de membrana que tiene un grupo hemo, pertenece a los sistemas regulatorios de dos componentes con el par FixJ. FixL es la proteína sensora histidina-cinasa y FixJ es el regulador con actividad de unión al DNA. Interviene en la regulación de FixK y NifA, que a su vez activan los genes *nif* y *fix* para la fijación de nitrógeno (Gilles-González, et al. 1994).

C) Flavin

NifL es una flavoproteína compuesta de dos dominios, el dominio carboxilo terminal, presenta cierta similitud con los dominios transmisores de las proteínas histidina-cinasas, inhibe la actividad de NifA mediante una interacción proteína-proteína (Dixon, 1998). En el dominio amino contiene FAD como grupo protético (Hill, et al. 1996). Aunque NifL aparentemente no contiene hierro, se ha visto que el hierro participa en la respuesta redox de NifL *in vivo*. Esta proteína también responde a la concentración de nitrógeno fijado en la célula y a la concentración de adenosin nucleósidos, pero al parecer son eventos independientes de la función redox (Söderbäck, et al. 1998).

D) Formación de enlaces disulfuro

Otro ejemplo de la regulación por oxígeno es OxyR, que regula positivamente los genes del regulón de estrés oxidativo. Esta proteína se activa directamente por oxidación y aparentemente no requiere de un ion metálico. Recientemente mediante estudios de espectrometría de masas, de muestras de OxyR en estado oxidado, indican que el resultado de la oxidación lleva a la formación de un puente disulfuro intramolecular entre la cisteína 199 y la cisteína 208, favoreciendo un cambio conformacional poco favorable a la unión al DNA (Zheng, et al. 1998; Farr, 1991).

OBJETIVOS

Dada la relevancia que presenta el estudio de la actividad de NifA como uno de los reguladores globales en el proceso de fijación de nitrógeno y la importancia de entender el mecanismo molecular que permite a esta proteína sensar el oxígeno y modular su actividad, nos llevaron a plantear los objetivos que a continuación se describen.

La evidencia de que el motivo de cisteínas presente en la NifA de *B. japonicum* es fundamental para la actividad de la proteína completa y el hecho de que exista una dependencia de iones divalentes, apoyan la idea de que estas cisteínas constituyen un motivo de unión a un cofactor metálico. Lo anterior, nos llevó a proponer que el estado redox de un metal involucrado, podría ser el control molecular que permite regular la actividad de esta proteína, por lo que la comprobación de la existencia e identificación del metal involucrado, fue uno de nuestros principales objetivos.

La proteína NifA es una proteína modular con actividades específicas en cada dominio que la constituye. Estudios anteriores habían demostrado que el dominio central de la NifA de *S. meliloti* conteniendo el motivo de cisteínas, era suficiente para activar la transcripción (Huala & Ausubel, 1988). Existían también antecedentes en los que el dominio C-terminal de *K. pneumoniae*, retiene la capacidad de unirse específicamente a las UAS, por lo que otro de nuestros objetivos fue determinar si el efecto regulatorio por oxígeno actuaba modulando la actividad de los dominios de manera independiente o concertadamente. Para contestar estas preguntas, y tomando en cuenta la dificultad que presenta la purificación de la proteína completa, la estrategia planteada fue expresar los dominios aislados de la proteína o como péptidos de fusión y determinar si los dominios central y carboxilo terminal, con el motivo de unión a metales, seguían siendo activos o perdían la regulación por oxígeno. La purificación de las proteínas truncadas permitiría la identificación del posible cofactor metálico y la caracterización del efecto del oxígeno sobre estas proteínas.

La modularidad de este tipo de proteínas nos hizo plantearnos preguntas en cuanto a la regulación de la actividad de los dominios y el balance que debe existir entre estas actividades, lo cual nos llevó a proponer que en este tipo de proteínas se presenta una evolución recíproca de sus dominios. Lo que permite compensar mutaciones desfavorables para un dominio y llevando al "óptimo" de actividad al otro dominio, permitiendo la existencia de estas variantes.

MATERIAL Y METODOS

- Cepas bacterianas y plásmidos

Las cepas empleadas se resumen como sigue:

CEPA	REFERENCIA
<i>Escherichia coli</i> JM101	Messing 1981
<i>Sinorhizobium meliloti</i> 1021	Simon, et al. 1983
<i>S. meliloti</i> 2	Segovia, et al. 1991
<i>S. meliloti</i> 7	Martínez, et al. 1990
<i>Sinorhizobium sp</i> BRL816	Martínez, et al. 1990
<i>Sinorhizobium fredii</i>	Martínez, et al. 1990

PLASMIDO	CARACTERISTICAS
pSUPG1	Derivado de pSUP202 (Simon, et al 1983) lleva el gen de <i>nifA G10E</i> de <i>S. meliloti</i>
pMAL-C2	NEW ENGLAND BIOLABS

- Purificación de las proteínas de fusión

Se creció un preinoculo con la cepa de *E. coli* llevando el plásmido pMALC-NifA a 37°C, posteriormente se inocularon botellas de 500 ml conteniendo el medio NFDM modificado como se describe en Morett, et al. 1991, con tapones con septo con el fin de eliminar el aire al cerrarlos, estos se crecieron por aproximadamente 10-14 hrs. Al alcanzar la D.O. a 600nm de 0.4 a 0.6, se adicionó el IPTG a una concentración final de 0.3mM. Se continuó el cultivo por 3-5 horas más, se recuperaron las células por centrifugación y se extrajo de dos formas, en la cámara anaeróbica o en condiciones normales.

Se lavó con amortiguador 1 y se resuspendió en 1/20 del volumen original, adicionando inhibidores de proteasas y ditionita.

Se sonicó (3 pulsos de 20 segundos cada uno), y se centrifugó a 8000 rpm.

Se diluyó 5 veces en amortiguador 2 y se metió a la columna de afinidad conteniendo la resina de amilosa (New England Biolabs).

Se hicieron varios lavados antes de eluir la proteína de fusión:

- 1) 5 volúmenes de columna con amortiguador 2
- 2) 3 volúmenes con amortiguador 3
- 3) elución con maltosa a una concentración de 10 mM.

-Amortiguador 1

10mM Fosfatos pH 7.2 ($\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ y $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$), 30 mM NaCl, 0.25% Tween 20, 10 mM β -mercaptoetanol. Cuando se agregó ditionita la concentración final fue de 1mM.

-Amortiguador 2

10mM Fosfatos pH 7.2 ($\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ y $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$), 30 mM NaCl, Cuando se agregó ditionita la concentración final fue de 1mM.

-Amortiguador 3

10mM Fosfatos pH 7.2 ($\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ y $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$), 300 mM NaCl
En este último se disolvió la maltosa para la elución final.

- Inmunodetección y obtención de anticuerpos policlonales de conejo.

Los anticuerpos empleados se hicieron como se describe en el artículo anexo. Las inmunodetecciones se hicieron como en González, V. (1997).

- Electroforesis SDS-PAGE (PolyAcrylamide Gel Electrophoresis) se realizó según el método de Laemmli (1970)

- PCR

Esta técnica se empleó para la generación de las fusiones, amplificando los dominios que posteriormente se clonaron en el plásmido pMalC y también se empleó para la generación de las mutantes empleando oligonucleótidos con las mutaciones que se introdujeron en los genes. Para la amplificación del extremo C-terminal de las diferentes cepas de *Sinorhizobium* se emplearon oligonucleótidos degenerados, diseñados a partir de alineamientos tomando en cuenta las regiones más conservadas del gene de *nifA*. Se emplearon las condiciones estándar de amplificación, variando solamente la temperatura de alineamiento del oligonucleótido, esto se decidió con base en la temperatura de "melting" del oligonucleótido.

- Ensayos de retardamiento.

Los ensayos se hicieron con las proteínas en estado anaeróbico, excepto

cuando se expusieron al aire para ver el efecto del oxígeno.

La reacción de unión para formar el complejo DNA-proteína, se realizó a 30 C por 15 minutos.

Los extremos 5' de un oligo correspondiente a la región del promotor de *nifD* de *B. japonicum* se marcaron radiactivamente con la kinasa de T4, y se extendió un fragmento de 250 pb, que se empleó para los ensayos de unión.

Amortiguador empleado para la reacción de unión.

50 mM KCl, 3 mM MgCl₂, 0.5 mM ditiotritol, 10% de glicerol v/v.

- **Determinacion de metal**

Con la proteína purificada, se determinó cualitativamente el contenido de Fe, empleando ensayos colorimetricos, con diferentes agentes quelantes.

Se empleo ferrozina segun Stookey (1970).

y sulfonato de bathofenantrolina según Woodland & Dalton (1984).

Para la determinación por absorción atómica, las proteínas purificadas se difirieron con HCN al 0.5% y se enviaron para su análisis a la compañía Difaza, se empleó el método de espectrometría de emisión (ICP).

- **Crecimiento de las plantas**

Semillas de alfalfa fueron esterilizadas y germinadas en placas de agar, posteriormente se inocularon con cultivos de *S. meliloti* 1021. Las plantas se crecieron a 25°C con un foto-periodo de 16h y regadas según fuera necesario. El número de nódulos y la actividad de nitrogenasa fueron determinados en grupos de cuatro plantas, por el método de reducción de acetileno

RESULTADOS Y DISCUSION

Regulación de la proteína NifA de *S. meliloti*.

- Caracterización del dominio carboxilo terminal.

Debido a la inestabilidad y sensibilidad al oxígeno que presenta la proteína NifA, los intentos de purificación de la proteína completa realizados por diferentes grupos han sido insatisfactorios. Como una alternativa se planteó que los dominios aislados podrían ser purificables con mayor facilidad, lo que nos permitiría además determinar si conservan su actividad y regulación con respecto al oxígeno de manera independiente.

Inicialmente se trabajó con la caracterización de la proteína NifA de *S. meliloti*, intentando expresar independientemente el dominio C-terminal o dominio de unión al DNA. Se eligió este dominio debido a que la unión al DNA es una actividad simple y fácilmente detectable. Contábamos con varias construcciones de péptidos truncados conteniendo el dominio de unión al DNA y el motivo conservado de cisteínas completo o solo con dos de las cisteínas.

Con la finalidad de determinar si estos péptidos truncados mantenían la actividad de unión al DNA, se realizaron ensayos de "footprinting" *in vivo*, empleando como control la proteína completa de *S. meliloti*. Los resultados obtenidos de estos ensayos fueron; una débil y poco reproducible protección de la guanina (G136) de las UAS con la proteína completa y una falta de unión con los péptidos truncados. Para comprobar que el ensayo permitía detectar la unión de NifA a las UAS, se hicieron ensayos de "footprinting" con proteínas en las que anteriormente se había determinado dicha unión (Morett, et al 1988). En el ensayo, se emplearon las proteínas silvestres de *K. pneumoniae* y *B. japonicum*, además de la NifA de *S. meliloti*. Los resultados mostraron una notable protección de la G136 de las UAS con las NifAs de *K. pneumoniae* y *B. japonicum*, como se

observa en la figura 1A del artículo anexo, a diferencia del ensayo con la NifA *S. meliloti*, en el que no se observa tal protección. Los ensayos de β -galactosidasa con la fusión *nifH::lacZ*, muestran que las tres proteínas empleadas para el "footprinting" son capaces de activar a muy altos niveles el promotor *nifH*.

Para determinar que la proteína de *S. meliloti* se estuviera expresando en cantidades similares a las observadas para las proteínas de *K. pneumoniae* y *B. japonicum*, se hicieron extractos celulares de los cultivos empleados en el "footprinting" y se analizaron mediante inmunodetección utilizando anticuerpos contra el dominio central de la proteína. Efectivamente, todas las proteínas probadas en este ensayo se expresaron en cantidades similares (datos no presentados).

Además, para comprobar la expresión de las construcciones que llevan solamente dominio C-terminal, se diseñaron dos péptidos sintéticos conteniendo el motivo HTH de las NifAs de *S. meliloti* y *B. japonicum*. Estos péptidos se acoplaron a albúmina sérica bovina y se inyectaron a conejos para la obtención de anticuerpos policlonales. Se hicieron inmunodetecciones con los extractos de cultivos de las construcciones del C-terminal de la NifA de *S. meliloti* y no se detectó proteína, lo que indicaba que estos no se estaban expresando o que eran demasiado inestables.

Se intentó encontrar una explicación que nos permitiera entender a qué se debía la débil unión al DNA presentada en la NifA de *S. meliloti*. Para esto se analizó la secuencia del motivo HTH de esta proteína y se comparó con las NifAs disponibles en las bases de datos. La diferencia más significativa fue que la NifA de *S. meliloti* carece de una glicina (G) altamente conservada y en su lugar presenta un residuo de ácido glutámico (E).

La glicina G10 es uno de los residuos más conservados en el motivo HTH de numerosas proteínas, incluso en las que se ha comprobado funcionalmente el motivo y de las que se tiene estructura cristalográfica en complejo con

el DNA (figura 1C, del artículo anexo) (Mathews & Brenan 1992). En estudios de mutaciones en la G10 del motivo HTH de la proteína HU de *Bacillus stearothermophilus*, muestran que la glicina es el aminoácido energéticamente mas favorable para la estabilidad de la HTH. En estos estudios se hicieron mutantes cambiando la G10E, que es el aminoácido presente en el motivo HTH de *B. subtilis* y se demostró que este cambio causa un decremento en la termoestabilidad de la proteína. Además, estudios de NMR con la mutante G10E, muestran que este cambio provoca una conformación mas abierta del motivo HTH, lo cual desestabiliza los contactos entre las hélices (Kawamura, et al. 1996).

Los datos anteriores sugerían que el cambio de G por E presente en la NifA de *S. meliloti* podría afectar la actividad de unión al DNA, debido posiblemente a una incorrecta estructuración del motivo HTH, dando como resultado la débil protección de las UAS observada en los ensayos de "footprinting" *in vivo*, lo cual pensamos que es compensado con una mayor capacidad de activación transcripcional llevada a cabo por el dominio central. Para probar esta hipótesis, se construyeron mutantes sencillas en las que se reestableció la glicina conservada y mutantes dobles, en las que además de la glicina se cambió el aminoácido contiguo por el más conservado en las NifAs (Dávila, S. 1997), a estas mutantes se les determinó la capacidad de activación mediante la fusión *nifH::lacZ*, cuantificando la actividad de β -galactosidasa. Como se esperaba, la activación del promotor aumentó dos veces, en el caso de la mutante sencilla y una vez en el caso de la mutante doble (Dávila, S. 1997) y se reestableció la unión al DNA (figura 2A, artículo anexo).

Adicionalmente se creó una mutante en el motivo HTH de la NifA de *B. japonicum*, esta proteína presenta G y se cambió por E, lo cual la hizo deficiente en la unión al DNA (figura 2, artículo anexo). Para comprobar si la NifA de *S. meliloti*, es menos dependiente de la unión al DNA, se midió la activación del promotor *nifH::lacZ* con y sin UAS y se comparó con la mutante G10E de *B. japonicum*. Como se esperaba, observamos que la

dependencia de la unión al DNA es menor en la NifA de *S. meliloti*, la cual en solución es capaz de retener el 25% de la actividad total a diferencia de la NifA de *B. japonicum* en la que se abate la capacidad de activación, (figura 2D). Estos resultados forman parte del artículo publicado como aportación original en el anexo 1.

Por otro lado, en el laboratorio se produjeron mutantes más activas de la NifA de *B. japonicum*. Lo cual se logró alterando el dominio central, después de tres rondas de mutagénesis, los cambios observados se localizan en posiciones conservadas en la familia, mutando hacia los aminoácidos presentes en la NifA de *S. meliloti*. Estos datos apoyaban la idea de que la NifA de *S. meliloti* presenta un dominio central más activo, que le permite compensar la débil actividad de unión al DNA. Además, posiblemente la NifA se encuentra en un espacio de secuencia restringido con soluciones limitadas para aumentar la actividad mediante mutaciones puntuales. La explicación de cómo se presentó este cambio en la proteína y cómo fue compensado nos llevó a proponer un modelo de lo que denominamos “Evolución recíproca de dominios”, descrito en el artículo anexo y esquematizado en la figura 5. Este modelo describe la coevolución entre los dominios de la NifA de *S. meliloti*, de manera que la mutación en la HTH que provoca la débil unión al DNA del dominio C-terminal es compensada mediante mutaciones en el dominio central que permiten aumentar su capacidad de activación.

La siguiente pregunta con relación al motivo HTH de la NifA de *S. meliloti* fue ¿qué tan recientemente ocurrió el cambio en la HTH de esta proteína? Para responder esta pregunta se obtuvo la secuencia del motivo HTH de varias especies cercanas filogenéticamente a *S. meliloti* (con base en comparaciones de RNA 16S ribosomal, (Martínez, et al. 1997). Como se mencionó anteriormente, la conservación a nivel de aminoácidos del motivo HTH de las NifAs es alta. De las secuencias disponibles en las

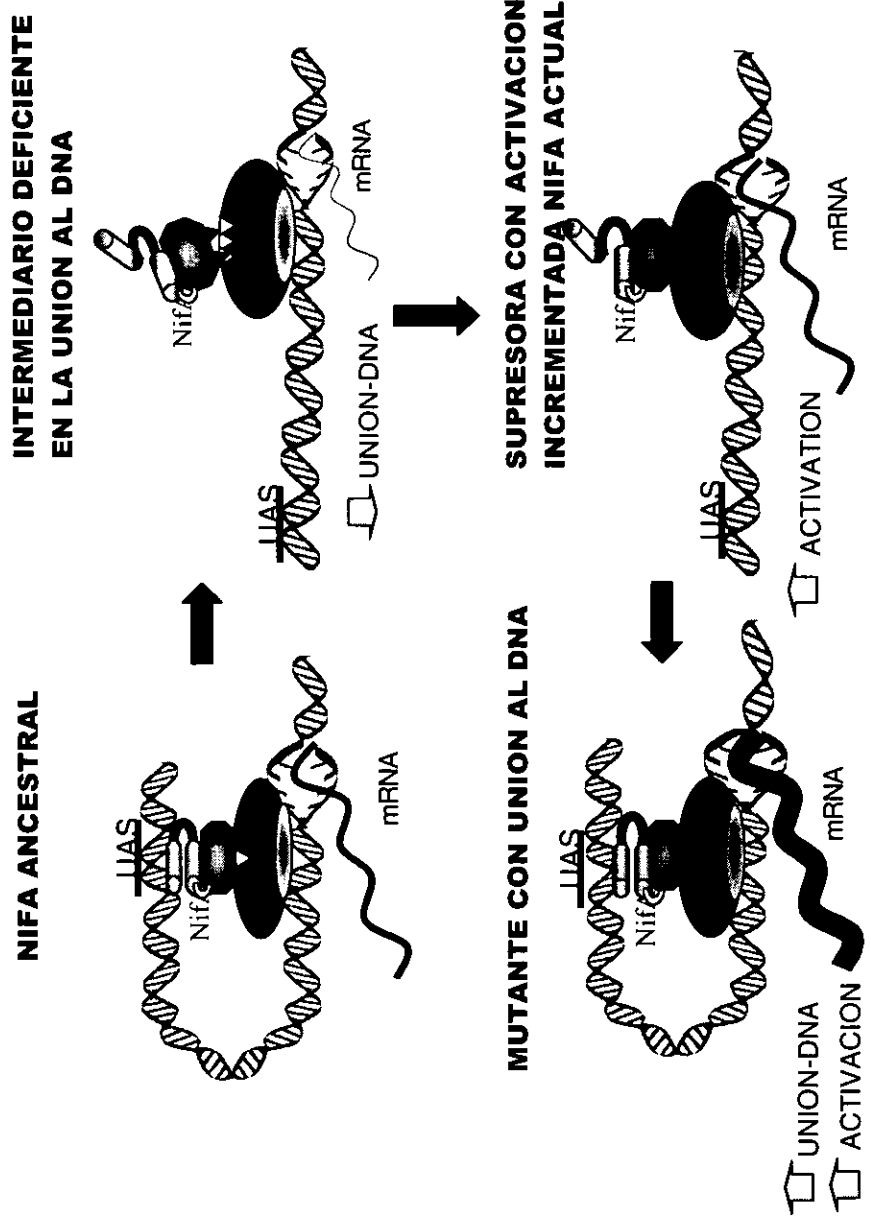


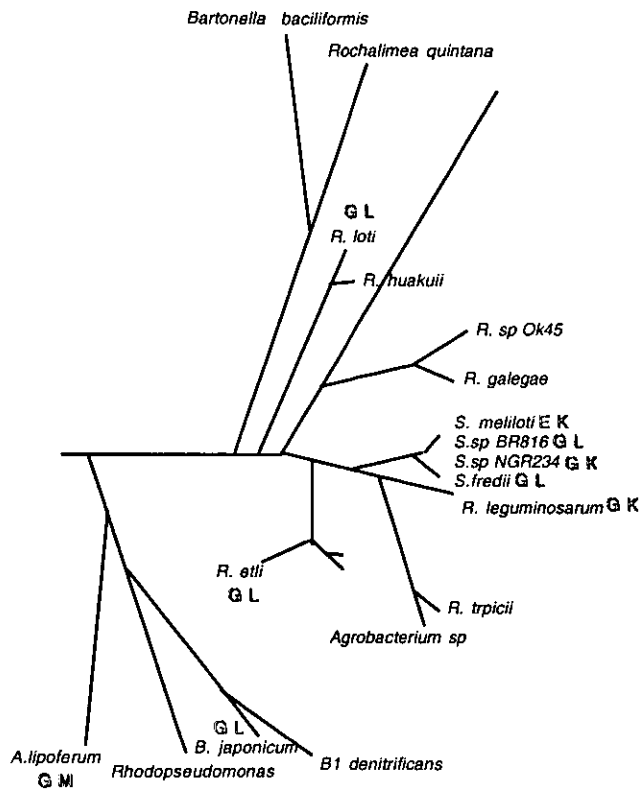
FIGURA 5. MODELO DE LA EVOLUCION RECIPROCA DE DOMINIOS

bases de datos, ninguna proteína de la familia presenta E en la posición diez, como en el caso de *S. meliloti* 1021. Para el análisis se emplearon varias especies de *Sinorhizobium*, varias cepas de *S. meliloti*, así como otras especies cercanas. Se diseñaron oligonucleótidos degenerados de las regiones más conservadas del gene de *nifA* y se amplificó por PCR la región que codifica para el extremo C-terminal de las NifAs de distintas especies.

La secuencia de los fragmentos de DNA amplificados reveló que todas las cepas analizadas de *S. meliloti* presentan E en esta posición y que las especies más cercanas filogenéticamente presentan G, figura 6A. Es interesante mencionar que una de las cepas analizadas de *S. meliloti*, presenta un cambio contiguo al E10. Se trata de una arginina (R) que es un aminoácido cargado positivamente, al igual que en el caso de la NifA de *S. meliloti* 1021, la cual presenta lisina (K) en esta posición y que pensamos puedan tener la función de compensar la carga negativa de E. A nivel de DNA también observamos cambios que en algunos casos son conservativos.

Los resultados anteriores nos llevaron a concluir que el cambio de G por E, ocurrió recientemente en *S. meliloti* y después de su divergencia de otras especies de *Sinorhizobium*. Actualmente estamos analizando otras especies y cepas del grupo de los *Sinorhizobia* para compararlos con las secuencias ya existentes, lo que nos permitirá rastrear precisamente cuándo sucedió el cambio y qué cambios compensatorios se presentan en el dominio central.

Como se mencionó anteriormente, la mutante E10G en el HTH de la NifA de *S. meliloti* presenta una mayor capacidad de activación, por lo que decidimos analizar su efecto en la fijación de nitrógeno. Para esto, se integró el plásmido pSUPNifAE10G (que lleva la NifA mutante) por recombinación homóloga en el cromosoma de *S. meliloti*. La recombinación nos permitiría tener dos distintas posibilidades de integración, en las que quedarían dos copias del gene *nifA* (la silvestre y la mutante) bajo diferentes promotores. La integración se comprobó por resistencia a

A**B**

Sm 1021	SALEEAGWNQAKAARILE <u>EK</u>
Sm 2	SAMEEAGWNQAKAARILE <u>ER</u>
Sm 7	SALEEAGWNQAKAARILE <u>EK</u>
S. C1	SALEEAGWNQAKAAAILE <u>EK</u>
S.sp.BR816	DAMEKSGWVQAKAARIL <u>GL</u>
S.loti	DAMEKAGWVQAKAARFL <u>GL</u>
S.fredii	LGMEKAGWDQAKAARIL <u>GL</u>

C

<i>S. meliloti</i> 1021	AGTGCCTGGAGGAGGCCGGTTGGAATCAGGCAAAGGCAGCTCGCATCCTCGAAAAAACG
<i>S. meliloti</i> 2	AGTGCATGGAGGAGGCCGGTTGGAATCAGGCAAAGGCAGCTCGCATCCTCGAAAGAACG
<i>S. melilotii</i> 7	ATCAGTGCCTGGAGGAGGCCGGTTGGAATCAGGCAAAGGCAGCTCGCATCCTCGAAAAAACG
<i>Sinorhizobium</i> C1	AGTGCCTGGAGGAGGCCGGTTGGAATCAGGCAAAGGCAGCTGCCATCCTCGAAAAAACG
<i>S. sp.</i> BR816	GACGCGATGGAAAAGTCCGGGTGGGTTTCAGGCCAAGGCGGCTCGTATTCTCGGCCTCAGC
<i>S. loti</i>	GATGCGATGGAGAAAGCCGGCTGGGTTTCAGGCCAAGGCGGCTCGTTTGCTCGGTCTCAGC
<i>S. fredii</i>	CTCGGATGGAGAAGGCGGGCTGGGATCAGGCCAAGGCAGCTCGGATCCTCGGCCTCAGC

Figura 6. Motivo HTH en especies cercanas filogenéticamente a *S. meliloti* 1021. A) Arbol filogenético basado en las secuencias de 16S ribosomal, (Martínez, et.al. 1997), B) Secuencias de proteína deducidas a partir de la secuencia nucleotídica obtenida por PCR. C) secuencia nucleotídica de los fragmentos de PCR obtenidos de especies cercanas a *R. meliloti*

tetraciclina (codificada en el plásmido). Además, dado que nos interesaba la integración en la que hubiera quedado la NifAE10G bajo el promotor silvestre, ya que el otro promotor del vector no se expresa en *S. meliloti*; se comprobó mediante PCR el tipo de integración que nos interesaba, empleando un oligonucleótido que hibrida con el promotor del vector en contrasentido a un oligonucleótido que hibridiza con el gene de NifA.

Se inocularon raíces de alfalfa con cultivos de las cepas resultantes. Interesantemente, la cepa con la mutación E10G a los 18 días después de la inoculación, indujo un gran número de pequeños nódulos. A esta cepa se le determinó la actividad de nitrogenasa, obteniéndose solo un décimo de la actividad alcanzada en la cepa silvestre. Posteriormente se observó que las plantas dejaron de crecer y a los 28 días después de la inoculación, no se detectó actividad de nitrogenasa.

Los resultados anteriores indican que la copia extra del gen que codifica para la NifA E10G fue desfavorable para la planta, posiblemente por un desbalance metabólico en los nódulos provocado por la mayor actividad de esta proteína. No se hicieron experimentos posteriores para explicar estos resultados y pensamos que es necesario repetirlos e incluir un control con la proteína silvestre integrada de la misma forma que se integró la mutante o hacer un reemplazo del gen de NifA, asegurando que exista una sola copia expresada a partir del promotor silvestre. Además, sería interesante determinar a que nivel afecta esta mayor capacidad de activación de NifA y si se pueden encontrar las condiciones en las que se alcance una mayor capacidad de fijar nitrógeno en los nódulos de la planta.

Regulación de la actividad de la NifA de *B.japonicum*

Con base en los resultados descritos anteriormente con respecto a la actividad de unión al DNA de la NifA de *S. meliloti*, y tomando en cuenta que para cumplir nuestros objetivos con respecto a la regulación por oxígeno era indispensable contar con una proteína en la que las actividades de los dominios aislados pudieran detectarse y purificarse individualmente, se decidió cambiar de proteína y se eligió la NifA de *B. japonicum*, la cual presenta una unión al DNA conspicua. Esto nos permitiría caracterizar las actividades del dominio de unión al DNA (C-terminal) y del dominio de activación (central), de manera independiente con respecto al oxígeno. Se planteó trabajar sólo con estos dos dominios de la proteína debido a que el dominio N-terminal en este tipo de NifAs no es esencial para la activación y no se le ha asociado alguna función regulatoria; incluso no se presenta en la NifA de *R. leguminosarum* bv *trifolii* (Huala & Ausubel, 1989; Iismaa & Watson, 1989)

- Expresión del dominio C-terminal

Con la finalidad de sobreexpresar el dominio C-terminal de la NifA de *B. japonicum* y purificarlo fácilmente, se diseñaron construcciones que llevaban una fusión de 6 histidinas (6H) en el extremo carboxilo del dominio C-terminal, introducidas por PCR y clonadas bajo un promotor fuerte e inducible por IPTG, lo que permitiría su purificación empleando una columna de afinidad a las 6H por medio de níquel. Se comprobó por secuencia nucleotídica que dichas construcciones estuvieran en fase traduccional y que llevaran las 6H.

Las construcciones del C-terminal conteniendo además el motivo de las cuatro cisteínas y la fusión de 6H se intentaron expresar en varias cepas de *E. coli*, incluso en una cepa deficiente de proteasas (BL21) y no se logró detectar la expresión, tanto en geles de poliacrilamida (PAGE) como por inmunodetección, por lo que se concluyó que estas construcciones no se

expresan en niveles detectables o son muy inestables y se degradan rápidamente. Con el propósito de saber si las 6H afectaban de alguna manera la actividad de la proteína, se fusionaron las 6H a la proteína silvestre de *B. japonicum*. En ensayos *in vivo* se determinó la capacidad de esta proteína para activar la fusión *nifH::lacZ*, se comprobó que la fusión de 6H en el C-terminal afecta la actividad de esta NifA y al igual que las construcciones no se observó su expresión en PAGE.

- Expresión de los dominios C-terminal y Central como proteínas de fusión.

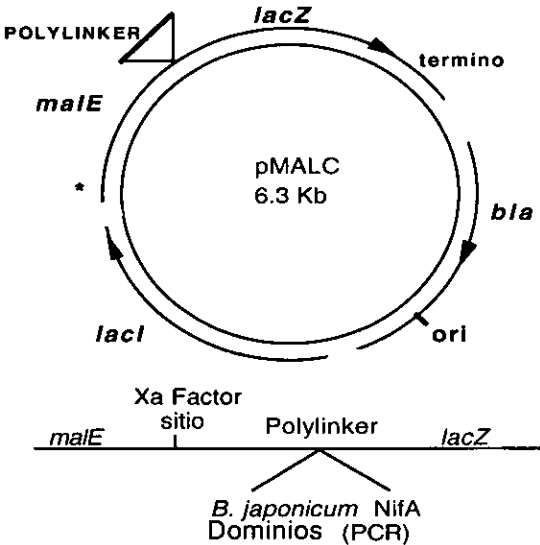
Para contender con el problema de inestabilidad presentado en las construcciones descritas anteriormente, se decidió cambiar de sistema de expresión, basándose en varios reportes de la literatura en los que se ha conseguido la expresión de proteínas inestables y poco solubles mediante una fusión con otra proteína (Bossé, et al.1993; La Vallie & McCoy,1995). Existen reportes de proteínas homólogas que se han logrado expresar de esta manera como son: la NifA de *K. pneumoniae* (Berger, et al. 1994) y recientemente la NifA de *Herbaspirillum brasilense* (Chubatsu, et al. 1999).

Con base en lo anterior, elegimos un sistema de fusión que emplea la proteína de unión a maltosa de *E. coli* o MBP (Maltose Binding Protein), codificada en el plásmido pMAL-C. Este plásmido permitió clonar la proteína o dominio de interés en fase con el gene *malE*, el cual se transcribe a partir de un promotor fuerte y es inducible por IPTG (figura 7A)

Este sistema permite una purificación sencilla de los péptidos expresados, empleando una columna de afinidad (amilosa-agarosa) a la que se une la parte del MBP. Se introdujo entre los dos péptidos una secuencia específica de corte del factor Xa, lo que permitiría eliminar el MBP, mediante otro paso de purificación. Se diseñaron y produjeron varios péptidos de fusión conteniendo la proteína de unión a maltosa y los

dominios de NifA, cuyas características se resumen en la tabla 3 y se esquematizan en la figura 7B.

A



B

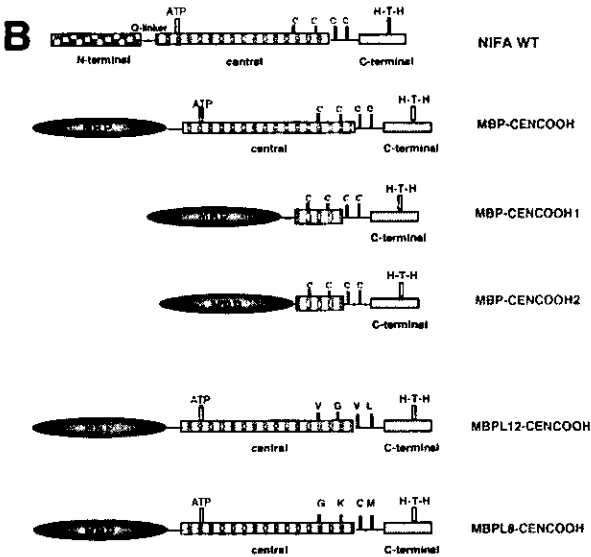


TABLA 3. DESCRIPCION DE LOS PEPTIDOS DE FUSION

NOMBRE DE LA FUSION	PESO MOL. ESPERADO (Da)	CYS	ESTABILIDAD	INMUNOBLOT ANTI-MBP	INMUNOBLOT ANTI-CEN	INMUNOBLOT ANTI-COOH
MBPCYSCOOH1	58500	4C	INESTABLE	si	nd	si
MBPCYSCOOH2	58480	4C	ESTABLE	si	nd	si
MBPCENCOOH	83630	4C	ESTABLE	si	si	si
MBPL8CENCOOH	83630	3C	ESTABLE	nd	si	nd
MBPL12CENCOOH	83630	----	ESTABLE	nd	si	nd

Las primeras construcciones empleando este sistema fueron inestables. La proteína de fusión MBPCYSCOOH1 se expresó del tamaño esperado. Sin embargo, se degradó rápidamente dando como resultado un fragmento de proteína que contenía solo la porción de MBP, lo cual se determinó por inmunoblot con un anticuerpo anti-MBP. En la figura 8, se muestra la sobreexpresión de la proteína de fusión (a) y el producto de degradación (b) en un PAGE. Este problema de inestabilidad se solucionó acortando el conector entre la MBP y los dominios de NifA, dando como resultado proteínas de fusión más estables.

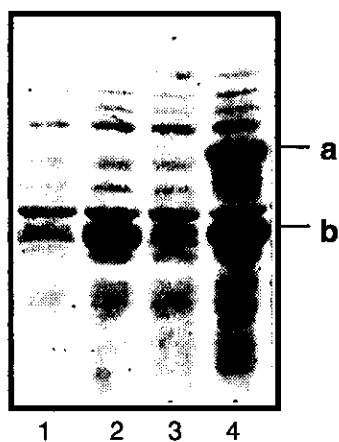


Figura 8. Expresión de fusión MBPCYSCOOH1. 1 y 3) sin IPTG, 2 y 4) con IPTG. a) MBPCYSCOOH1 completa, b) MBPCYSCOOH1 degradada

La expresión de algunas de estas construcciones se puede observar en la figura 9, así como su inducción con IPTG. Se comprobó la identidad de estas construcciones empleando dos anticuerpos; uno contra el C-terminal de NifA y otro contra MBP. En la figura 9, solo se muestra la inmunodetección empleando el anticuerpo anti-C-terminal ya que ambas construcciones comparten este dominio.

La purificación de las diferentes proteínas de fusión generadas se realizó empleando una columna de amilosa. En la fracción soluble de los extractos celulares de las proteínas de fusión, se presenta aproximadamente el 50% de éstas, el resto se agrega y se precipita, posiblemente formando cuerpos de inclusión. Los amortiguadores empleados para la purificación se describen en métodos. Cabe mencionar que se modificó la fuerza iónica con respecto al protocolo recomendado por la casa comercial que vende el plásmido ya que no se pegaba la proteína a la resina. Se disminuyó la concentración de sales en el pegado inicial y se incrementó su concentración en la elución de las proteínas de fusión. En la figura 10, se pueden observar los diferentes pasos de purificación de una de las construcciones (MBPCENCOOH) tanto en PAGE como por inmunodetección. La purificación de las otras construcciones se realizó siguiendo la misma estrategia.

La purificación se realizó en dos condiciones: aeróbicas y anaeróbicas. La purificación en condiciones aeróbicas se realizó manteniendo las proteínas en condiciones reductoras empleando ditionita como agente reductor, además de que los amortiguadores se desgasaron y saturaron con argón. En la purificación anaeróbica se empleó una cámara conteniendo una mezcla de gases (CO₂-5%, Hidrógeno-10% y Nitrógeno-85%) empleada para crecimiento anaeróbico.

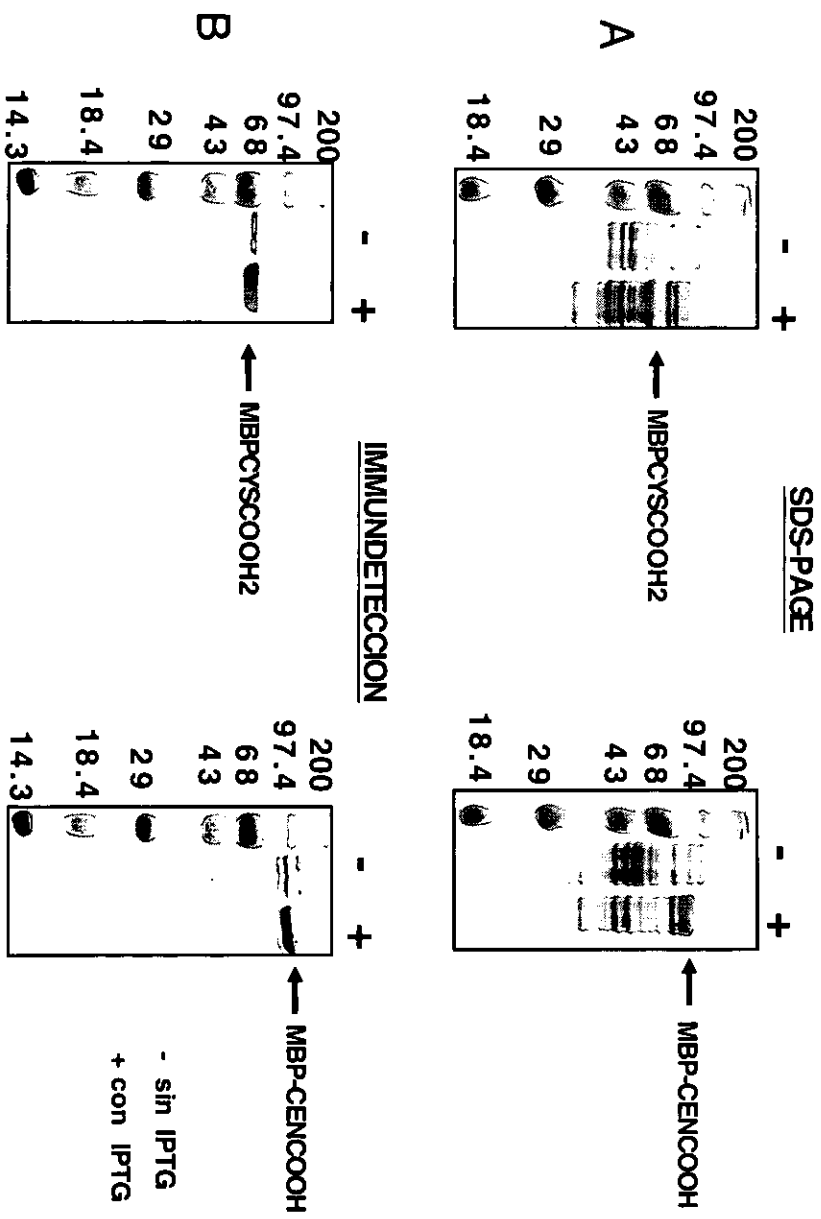


FIGURA 9. EXPRESION DE DIFERENTES PROTEINAS DE FUSION. A) SDS-PAGE, B) Immunoblot empleando anticuerpo anti-COOH de NifA de *B. japonicum*

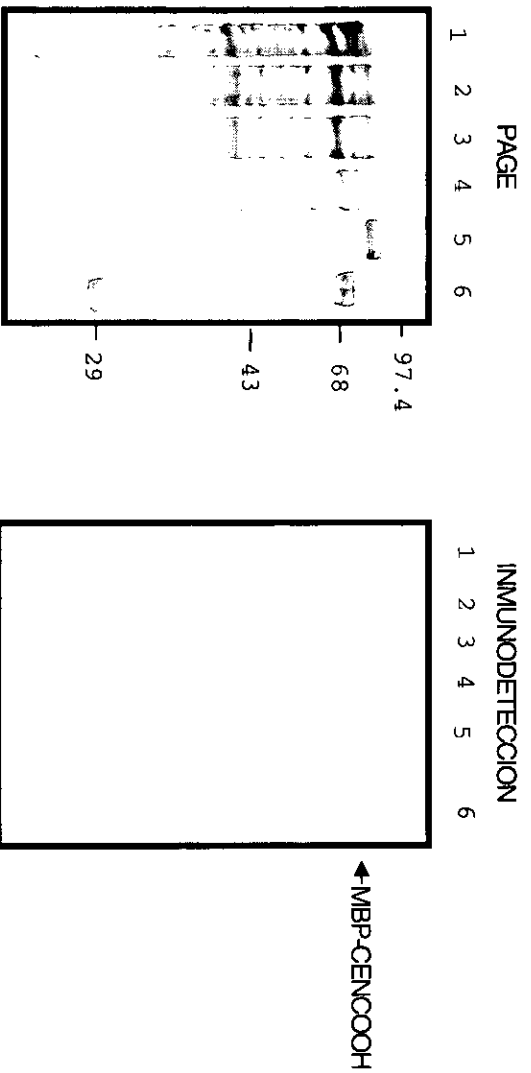


Figura 10. PURIFICACION DE LA CONSTRUCCION MIBP-CENCOOH

PAGE-1) EXTRACTO CRUDO, 2) COLUMNA, 3) 1er LAVADO, 4) 2do LAVADO,
 5) ELUCION CON MALTOSA, 6) MPM
 INMUNODETECCION- 1) EXTRACTO CRUDO, 2) COLUMNA, 3) 1er LAVADO,
 4) 2do LAVADO, 5) ELUCION CON MALTOSA, 6) MPM
 * AB (antiniFA)

Como se puede observar en la figura 11, se obtuvieron las proteínas puras de las construcciones MBP-CENCOOH y MBPCYSCOOH2 en condiciones anaeróbicas.

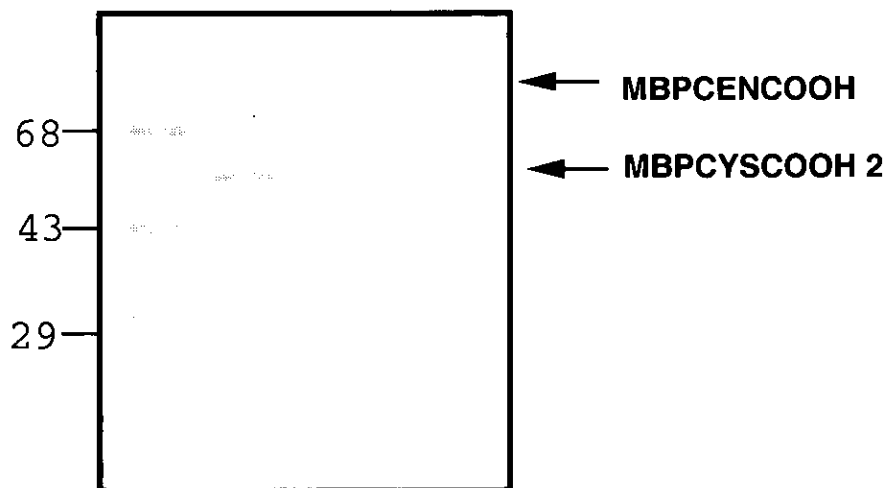


Figura 11. PAGE de las construcciones puras

- CARACTERIZACION DE LAS PROTEINAS DE FUSION PURIFICADAS

La caracterización de las proteínas de fusión purificadas se realizó tomando en cuenta varios aspectos como son: la actividad de unión al DNA y la capacidad de activación *in vivo* del promotor *nifH::lacZ*, en el caso de las que contenían el dominio central.

- Caracterización de la proteína de fusión MBPCENCOOH

La fusión MBPCENCOOH conserva una capacidad de activación del 30%, en comparación a la activación obtenida con la proteína silvestre. La actividad de unión al DNA de esta proteína se determinó por ensayos de retardamiento empleando un fragmento de 250 pb sintetizado por PCR el cual lleva las UAS correspondientes al promotor *nifD* de *B. japonicum*.

Es importante mencionar que en los ensayos de retardamiento se emplearon dos técnicas de separación del complejo formado entre la proteína y el DNA, ésto con la finalidad de obtener una mejor definición del complejo y separación con el fragmento libre:

- El primer tipo de ensayo se realizó marcando el extremo 5' del oligonucleótido con γ -[³²P] empleado en la amplificación por PCR. Las reacciones de unión al DNA se incubaron 15 min a 30°C y se hicieron con la proteína pura. Para la separación del complejo DNA-proteína (D-P) del fragmento de DNA libre, se empleó un gel de acrilamida al 6%. Fue difícil encontrar las condiciones en las que el complejo D-P pudiera entrar a la matriz del gel, posiblemente porque se forman agregados o multímeros de alto peso molecular que no entran a la matriz del gel y se quedan retenidos en el pozo (ver figura 12 B).

- En el segundo tipo de ensayo, se empleó un fragmento de DNA de 250 pb sin marcar radiactivamente en las reacciones de unión. Los complejos formados se separaron en geles de agarosa de bajo punto de fusión al 2.5-3% y se tiñeron con bromuro de etidio. En este caso la mayor parte del complejo D-P entró a la matriz del gel de agarosa y se observó una mayor definición del complejo (ver figura 12 A).

Se analizó el efecto de la concentración de proteína en la formación del complejo D-P con la fusión MBPCENCOOH, como se puede observar en la figura 12A. En el ensayo se probaron cantidades crecientes de esta proteína y el retardamiento observado por la formación de dicho complejo es dependiente de la concentración de proteína. Para comprobar que la unión al DNA fuera específica, se empleó un fragmento del mismo tamaño pero sin UAS, el cual no se retardó, aún empleando la mayor concentración de proteína. Por último se agregó DNA inespecífico (DNA de esperma de salmón) a una concentración 10 veces mayor a la empleada en el fragmento con UAS, lo cual no impidió la formación del complejo específico D-P, como se observa en la figura 12A. Es importante hacer notar que la

cantidad de fragmento libre no corresponde a la cantidad del fragmento retardado en el complejo D-P a concentraciones mayores de proteína. Esto se debe a que con altas concentraciones de proteína una parte del complejo se quedó retenido en la parte superior del gel.

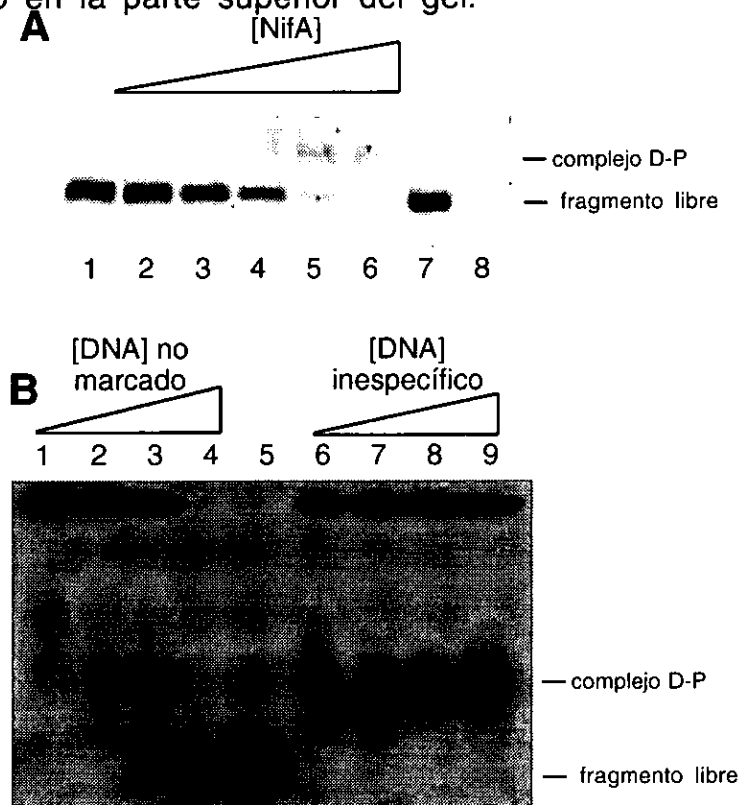


Figura 12. Ensayos de retardamiento con la fusión MBPCENCOOH. A) Gel de agarosa teñido con bromuro de etidio, 1, fragmento con UAS sin proteína; 2-6, fragmento con UAS y concentraciones crecientes de proteína 30-120 ug/ul de proteína, 7, fragmento sin UAS (120 ug/ul de proteína), 8, fragmento con UAS y DNA de esperma de salmón B) Gel de retardamiento empleando un fragmento marcado radiactivamente; 1-3, concentraciones decrecientes de fragmento no marcado, 4-5 fragmento sin proteína; 6-9, competencia con DNA de esperma de salmón.

La unión específica de la fusión MBPCENCOOH al DNA, también se comprobó mediante ensayos de competencia empleando un fragmento con UAS marcado radiactivamente. En la figura 12B, se puede observar el retardamiento del complejo D-P en competencia con un fragmento no-marcado de las mismas características pero en cantidades crecientes. En

este ensayo se observó que concentraciones crecientes del fragmento no-marcado compiten por la unión específica de la proteína desplazando la unión con el fragmento marcado radiactivamente. También se hicieron ensayos de retardamiento con el fragmento de DNA con UAS en competencia con cantidades crecientes de DNA inespecífico; para esto se empleó DNA de esperma de salmón. Como se puede observar en este último caso el complejo D-P sigue formándose y no es desplazado por el DNA inespecífico.

Con el fin de elucidar el efecto del oxígeno sobre la actividad de unión al DNA en la fusión MBPCENCOOH, se hicieron los siguientes ensayos: La proteína pura se dejó 12-14 hrs expuesta al aire y antes del ensayo se agitó en el vortex por 3 min. Como se observa en la figura 13, la unión al DNA disminuyó después de la exposición al aire, aunque cabe hacer notar que el efecto no es total. Se comprobó que la cantidad de proteína fuera igual en ambas condiciones (figura 13B) y que el efecto del oxígeno sobre la actividad de unión al DNA no se debió a una degradación de la proteína sino posiblemente a un cambio en la estructura que afecta la actividad de unión al DNA.

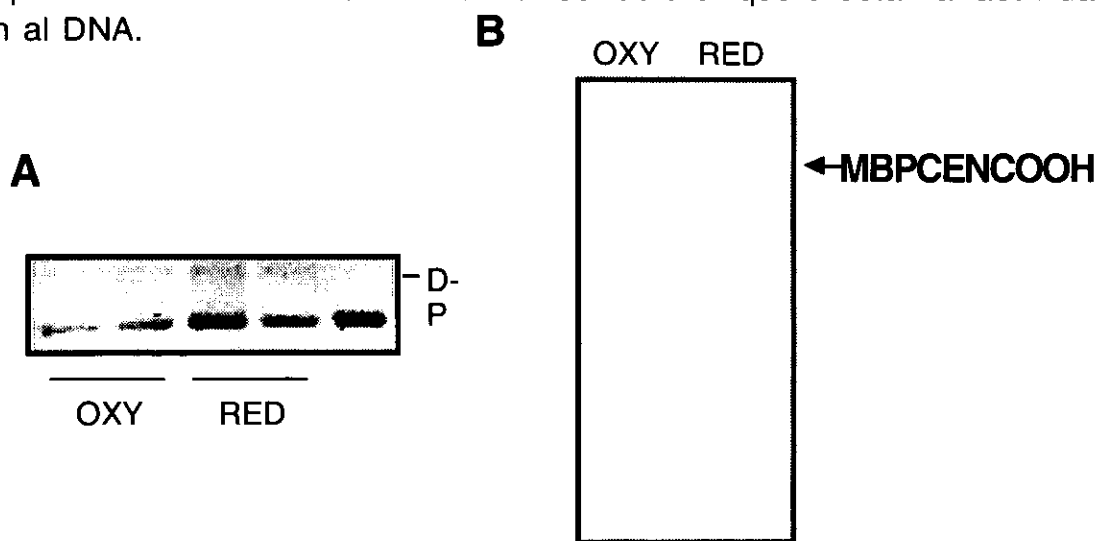


Figura 13. Efecto del oxígeno en el retardamiento de DNA con la proteína de fusión MBPCENCOOH. A) Gel de reatrdamiento 1-2) Oxidadas, 3-4) Reducidas, 5)sin proteína B) PAGE de las proteínas; 1) oxidada 2) reducida.

- Caracterización de la proteína de fusión MBPCYSCOOH2

Como ya se ha mencionado, uno de nuestros objetivos era determinar si el control molecular que le permite a la proteína NifA sensor el oxígeno afecta a los dominios de manera independiente. Contando con la proteína de fusión MBPCYSCOOH2, que lleva solo el C-terminal y el motivo de cisteínas se analizó la actividad de unión al DNA. Como se puede observar en la figura 14, la fusión MBPCYSCOOH2 retuvo la capacidad de unión al DNA y también es dependiente de la concentración de proteína. La unión al DNA es específica, lo cual se comprobó empleando un fragmento sin UAS y éste no se retardó. Como se describió anteriormente, se hicieron ensayos de retardamiento con proteínas expuestas al aire 12 hrs y no se observó ninguna diferencia con respecto a la actividad de unión al DNA (no se muestran estos datos).

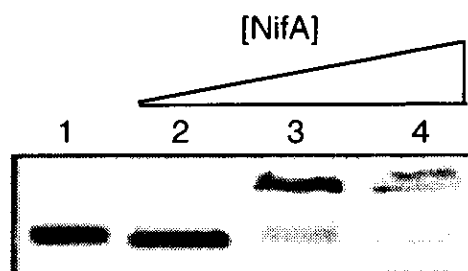


Figura 14. Retardamiento de DNA-MBPCYSCOOH2, 1) sin proteína, 2-4) concentraciones crecientes de la proteína de fusión MBPCYSCOOH2, 10-80 ug/ml.

Los resultados anteriores con respecto a la regulación por oxígeno de actividad de unión al DNA *in vitro* con las fusiones MBPCENCOOH y MBPCYSCOOH2, nos indicaban que el efecto del oxígeno sobre esta actividad no es total y que posiblemente se requiere de la regulación de los dominios concertadamente para observar un efecto mayor, la discusión de este punto se describirá mas adelante integrando todos los resultados.

- Obtención de mutantes en el motivo conservado de cisteínas.

Paralelamente a la caracterización de las proteínas de fusión con respecto a la regulación por oxígeno, estábamos interesados en obtener una proteína completa insensible a oxígeno, purificable y que nos aportara datos estructurales. Propusimos llevar a cabo la mutagénesis de las cuatro cisteínas que probablemente coordinan un cofactor metálico, mediante la introducción de aminoácidos con ciertas características, que permitieran sustituir al cofactor metálico y obtener una proteína estable. Existen ejemplos en los que se ha podido llevar a cabo este tipo de sustitución, reemplazando la unión de Zn con aminoácidos hidrofóbicos (referencia).

En la sustitución de las cisteínas pensamos que sería importante introducir aminoácidos hidrofóbicos, así como aminoácidos que favorecen la formación de puentes salinos, éste último con el fin de proporcionarle estabilidad a la proteína.

Debido a que la combinatoria de cuatro sustituciones en una mutagénesis del tipo NNG/C eleva el tamaño del banco necesario para obtener todas las posibles combinaciones, además de la sobrerrepresentación de algunos aminoácidos en el banco (Del Río, et al. 1994), y debido a que no todos los aminoácidos nos interesaban, se decidió seguir otra estrategia de mutagénesis sitio-dirigida que consiste en emplear mezclas de oligonucleótidos con una proporción determinada de bases en la primera, segunda y tercera posición del codón (Arkin & Youvan, 1992). Basándose en esta última, diseñamos una mezcla que nos permitía la posibilidad de introducir 16 aminoácidos (14 representados por un solo codón y 2 aminoácidos representados por 2 codones y solo un codón de paro). Con esto el tamaño del banco calculado para obtener todas las posibles variantes disminuyó considerablemente.

La selección de las mutantes se realizó con base en la capacidad que presentaron de activar una fusión *nifH::lacZ* en cajas conteniendo X-gal (selección por color). Se obtuvieron varias mutantes que se estructuran

correctamente en cuanto a tamaño y estabilidad, como lo demuestran ensayos de inmunodetección, pero con una actividad residual. Actualmente se están analizando algunas mutantes que fueron seleccionadas por presentar cierta actividad.

En estudios anteriores se había descrito que las cuatro cisteínas que conforman el probable motivo de unión a metales son esenciales para la actividad de la proteína completa, determinado por ensayos de activación *in vivo* (Fischer, et al. 1988). Además, en estudios con una mutante de la NifA en una de las cisteínas (Cys472), se comprobó que ésta pierde la capacidad de unión al DNA en ensayos de "footprinting" *in vivo* (Morett et al. 1991). Para esclarecer este punto decidimos reemplazar el motivo de cisteínas de dos de las mutantes inactivas descritas anteriormente en la fusión MBPCENCOOH, lo que nos permitió purificarlas y caracterizar su actividad de unión al DNA.

Las fusiones resultantes fueron MBP-L12CENCOOH y MBP-L8CENCOOH, las cuales se purificaron y se les determinó la actividad de unión al DNA. A diferencia de lo esperado, la construcción MBP-L12CENCOOH que lleva los cambios (V-X₁₁-G-X₁₉-V-X₄-L) conserva la actividad de unión al DNA, como se observa en la figura 15A. Igualmente, la construcción MBP-L8CENCOOH que presenta una de las cuatro cisteínas del motivo con los cambios (G-X₁₁-K-X₁₉-C-X₄-M) también sigue uniéndose al DNA como se ve en la figura 15B (ver discusión adelante)

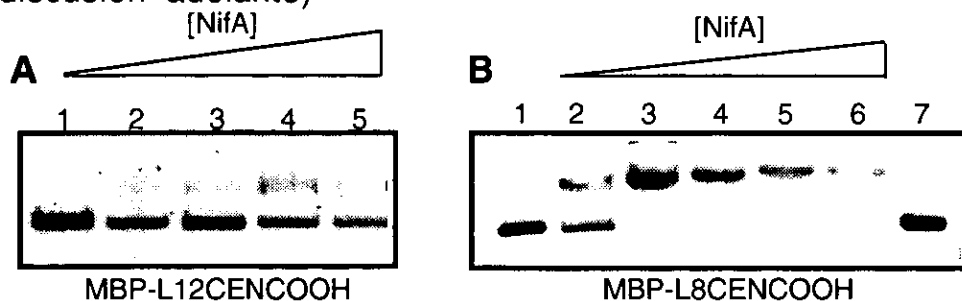


Figura 15. Retardamiento con mutantes en el motivo de cisteínas. A)MBP-L12CENCOOH B)MBP-L8CENCOOH, en ambos casos se probó la concentración de proteína.

Los resultados anteriores con respecto a la actividad de unión al DNA, sugieren que el dominio C-terminal de la NifA de *B. japonicum* es independiente del control por oxígeno en ensayos *in vitro*. Sin embargo, en estudios anteriores en los que se analizó el efecto del oxígeno sobre la actividad de unión al DNA mediante ensayos de "footprinting" *in vivo* con la proteína completa de *B. japonicum*, se observó un efecto inmediato del oxígeno sobre la unión al DNA (Morett et al. 1991). Una posible explicación a esta diferencia entre los ensayos *in vitro* e *in vivo* podría ser que el oxígeno no afecta directamente la actividad de unión al DNA, sino que posiblemente actúa de manera concertada afectando alguna de las actividades del dominio central, lo cual posteriormente se ve reflejado en la unión al DNA.

Los estudios con otros reguladores transcripcionales que son capaces de regular su actividad en respuesta a la concentración de oxígeno, nos permiten entender cómo podría funcionar esta regulación en la proteína NifA. En el caso de FNR la unión al DNA se lleva a cabo principalmente por homodímeros de la proteína, que en condiciones anaeróbicas se ensamblan conteniendo un centro [4Fe-4S] el cual es muy lábil y en presencia de oxígeno cambia a [2Fe-2S], favoreciendo la formación de monómeros que disminuyen la afinidad por el DNA notablemente (Lazazzera et al. 1993, Ziegelhoffer & Kiley, 1995).

En el caso de la NifA de *K. pneumoniae*, el dominio C-terminal aislado se comporta como monómero, aunque mantiene cierta capacidad de unión al DNA, sugiriendo que los determinantes para la oligomerización residen en el dominio central. (Lee et al 1993; Morett, et al. 1988). La proteína de fusión MBPCYSCOOH₂, que lleva el C-terminal y el motivo de cisteínas, conserva la actividad de unión al DNA y al parecer no es sensible a oxígeno, a diferencia de proteínas truncadas de la NifA de *S. meliloti* que contienen el dominio central solamente y el motivo de cisteínas y conservan la regulación por oxígeno sobre la capacidad de activación de promotores dependientes de NifA (Huala et al 1992).

Se ha demostrado que otras EBP's también forman oligómeros al unirse al DNA, y que la unión de NTP's en el dominio central provoca el cambio conformacional requerido para ensamblarse, favoreciendo la unión a las UAS (Pérez & De Lorenzo, 1996). Por otro lado cabe mencionar que existe una mutante tolerante a oxígeno de la NifA de *S. meliloti*, en la que el cambio se localiza en la zona de unión a NTP (Krey et al. 1992), lo que sugiere que posiblemente esta mutante adopta una conformación que le permite interactuar con el NTP y favorecer la oligomerización aun en presencia de oxígeno.

Con los datos anteriores proponemos que en la NifA podría suceder algo similar a lo observado en FNR, donde el estado reducido del cofactor metálico favorece una conformación de la proteína en la que puede entrar el nucleótido y promover su oligomerización, incrementando la unión al DNA en forma activa. En presencia de oxígeno el cofactor promueve cierta conformación de la proteína que impide la unión al nucleótido y aunque sigue habiendo unión al DNA, ésta ocurre con una afinidad menor y en forma monomérica inactiva.

DETERMINACION DEL COFACTOR METALICO

Aunque el mecanismo mediante el cual el oxígeno inactiva a la proteína NifA, aun es desconocido, existen antecedentes de que las NifA's de *B. japonicum* y *R. trifolii* son dependientes de iones metálicos divalentes, principalmente fierro y sensibles a agentes quelantes (Fischer et al, 1988; Fischer, et al. 1989; Screen, et al. 1994). Existen varios ejemplos de metaloproteínas que intervienen en la expresión genética, en este sentido los metales de transición juegan un papel muy importante, SoxR y Fnr sensan cambios en la presión parcial de O₂ mediante centros fierro-azufre. Existen otros ejemplos en los que el metal que interviene, por ejemplo Zn, mantiene la conformación tridimensional de la proteína en forma activa (revisado en Karlin, 1993)

Contando con las contrucciones antes descritas y purificadas, se llevó a

cabo mediante métodos bioquímicos la determinación cualitativa del metal (Stookey, 1970; Woodland & Dalton, 1984). Para lo cual fue necesario hidrolizar la proteína y por medio de un ensayo colorimétrico a cada metal, identificarlo espectrofotométricamente.

Se hicieron determinaciones para hierro y para cobre, que por las características que presenta la proteína y el tipo de regulación que lleva a cabo son los más factibles. La proteína de fusión MBP-CENCOOH contiene hierro, como se observa en la figura 16. Posteriormente este resultado se

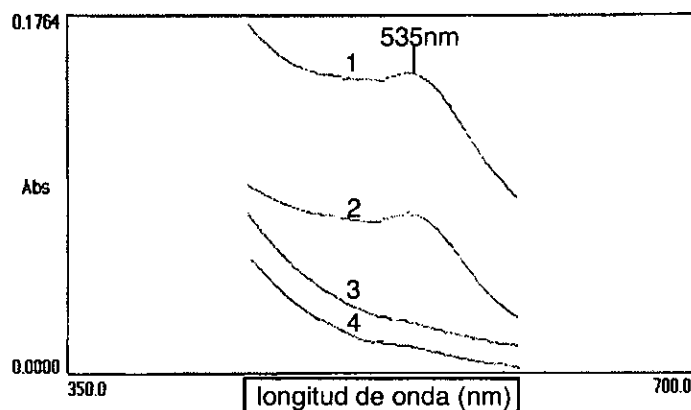


Figura 16. Análisis cualitativo del contenido de Fe en la proteína de fusión MBPCENCOOH. 1) Sulfato de hierro amoniacal, 2) MBPCENCOOH, 3) BSA, 4) Control sin proteína

corroboró mediante un espectro de absorción atómica para determinar el contenido de Fe.

Como se observa en la tabla 4, MBPCENCOOH presenta Fe en su estructura, a diferencia de la fusión de MBPCYSCOOH2 la cual también presenta el

motivo de cisteínas. Una explicación a este resultado podría ser que debido a que la fusión MBPCYSCOOH2 presenta solamente el C-terminal, requiera de algún determinante del dominio central para que se estructure de manera adecuada el cofactor metálico.

TABLA 4. DETERMINACION DEL CONTENIDO DE FIERRO EN LAS PROTEINAS DE FUSION PURIFICADAS.

Proteína de fusión	DETERMINACION DE Fe
MBPCYSCOOH2	0.002 mg/L
MBPCENCOOH	0.091 mg/L
CONTROL (RNASA)	0.006 mg/L

La estructura del cofactor involucrado no es posible determinarla aún, ya que faltan datos del contenido de sulfidrilos que pudieran estar involucrados en la unión de Fe y su comprobación por EPR (espectro de resonancia paramagnética) lo cual se realizará próximamente.

CONCLUSIONES

Los resultados obtenidos con respecto a la caracterización de la NifA de *S. meliloti* nos llevaron a las siguientes conclusiones:

- La NifA de *S. meliloti* presenta una deficiente actividad de unión al DNA, provocada probablemente por una incorrecta estructuración del motivo HTH.
- La deficiente actividad de unión al DNA en esta proteína, es compensada con una mayor capacidad de activación.
- El cambio en la HTH que provoca la baja actividad de unión al DNA ocurrió recientemente, después de su divergencia con otras especies de *Sinorhizobium*.
- Estudios preliminares muestran que las mutantes que restablecen la capacidad de unión al DNA y que presentan una mayor actividad transcripcional *in vitro*, al integrarse al cromosoma de *S. meliloti* e inocular plantas de alfalfa con estas cepas provocan un efecto negativo en el crecimiento de la planta. Además, la actividad de nitrogenasa disminuye un 90% con respecto a la cepa silvestre.

Los resultados obtenidos con respecto a la regulación por oxígeno de la NifA de *B. japonicum* nos permitieron concluir lo siguiente:

- La fusión de 6 histidinas al extremo carboxilo terminal de la NifA de *B. japonicum* produce proteínas inactivas.
- Las proteínas de fusión con MBP son estables y conservan la actividad de los dominios *in vitro*.

- La proteína de fusión MPBCENCOOH que contiene los dominios central, C-terminal y el motivo de cisteínas conserva un 30% de la capacidad de activación, con respecto a la proteína silvestre. Esta fusión conserva la actividad de unión al DNA, la cual es sensible a oxígeno.
- La fusión MPBCENCOOH contiene hierro en su estructura y posiblemente conforma un cofactor metálico que responde a las concentraciones celulares de oxígeno.
- La proteína de fusión MBPCYSCOOH que contiene el C-terminal aislado y el motivo de cisteínas, es capaz de unirse al DNA específicamente. Sin embargo, aunque contiene el motivo de cisteínas, no se detectó hierro en su estructura y no responde a la concentración de oxígeno.
- Las fusiones MBPL12CENCOOH y MBPL8-CENCOOH que contienen mutaciones en las cuatro cisteínas o en tres de ellas respectivamente, pierden la actividad de activación transcripcional *in vivo* pero conservan la actividad de unión al DNA *in vitro*.
- Con base en las conclusiones anteriores proponemos que el dominio de unión al DNA no responde directamente al oxígeno y que la regulación se da manera concertada con alguna de las actividades presentes en el dominio central.

BIBLIOGRAFIA

- Arsene, F., Kaminski, A & Elmerich, C. (1999) FEMS. Microbiol. lett. **179**: 339-343.
- Arkin, A. & Youvan, C. (1992) Bio/Technology **10**:297-300.
- Austin, S., Hendersonm N. & Dixon, R. (1990) Eur. J. Biochem. **187**: 353-360.
- Austin, S. & Dixon, R. (1992) EMBO J.**11**:2219-2228.
- Austin, S., Buck, M., Cannon, W., Eydmann, T. & Dixon, R. (1994). J. Bacteriol. **176**(12):3460-3465
- Barrios, H., Fischer, H. M., Hennecke, H. & Morett, E. (1995) J. Bacterio. **177**: 1760-1765.
- Barrios, H., Grande, R., Olvera, L. & Morett, E. (1998) Proc. Natl. Acad. Sci. USA **95**: 1014-1019.
- Barne, K. A., Bown, J. A., Busby, S. J., & Minchin, S. D. (1997) EMBO k., **16**:4034-4040.
- Baumeister, R, Müller, G., Hecht, B and Hillen W. (1992). Proteins. **14**:168-177.
- Bauer, E., Kaspar, T., Fischer, H.M. & Hennecke, H. (1998) J. Bacteriol **180**(15):3853-3863
- Berger, D. K., Narberhaus, F. & Kustu, S. (1994) Proc. Natl. Acad. Sci. USA **91**:103-107
- Blanco, G., Drummond, M., Woodley, P. & Kennedy, C. (1993) Mol. Micobiol. **9**:869-880.

- Bossé, M., Handl, C., Lortié, L.A, Harel, J. & Dubreuil, D. (1993). *Journal of General Microbiology*. **139**:631-638.
- Brennan R.G. & Matthews, B W. 1989 *J. Biol Chem* 264(4) 1903-1906.
- Buck, M., Miller, S., Drummond, M. & Dixon, R. (1986) *Nature (London)* **320**:374-378.
- Buck, M., Cannon, W. y Woodcock, J. (1987) *Mol. Microbiol* **1**:243-249.
- Buck, M. (1986). *J. Bacteriol.* **166**:545-551.
- Bujard, H., Brunner, M., Deuschle, U., Kammerer, W. & Rainer, K. (1987). En Reznicoff, WS, Burgues, R., Dahlberg. J. Gross. C. Record, T. and Wickens (eds). *RNA Polymerase and the Regulation of Transcription*. Elsevier, N.Y. 95-113.
- Burgess, R., Erickson, D., Gentry, K., Gribskov, D., Hager, S., Lesley, M. Strickland, M. & Thompson, N. (1987) En Reznicoff, WS, Burgues, R., Dahlberg. J. Gross. C. Record, T. and Wickens (eds). *RNA Polymerase and the Regulation of Transcription*. Elsevier, N.Y. 3-16.
- Busby, S & Ebright, R.H. (1997) *Mol. Microbiol.* **23**(5):853-859.
- Busby, S & Ebright, R.H. (1999) *J. Mol. Biol.* **293**:199-213
- Cannon, W., Charlton, W. y Buck, M. (1991) *J. Mol. Biol.* **220**:915-921.
- Cannon, W., Gallegos, M. T., Casaz, P. & Buck, M. (1999) *Genes Dev.* **13**(3):357-370
- Casaz, P. & Buck,M. (1999) *J. Mol. Biol.* **285**:507-514.
- Casaz, P., Gallegos, M. T. & Buck,M. (1999) *J. Mol. Biol.* **292**:229-239.
- Caslake, L.F., Gruber, T. M. & Bryant, D. A. (1997) *Microbiology.* **143**:3807-

3818

- Choy, H. & Adhya, S. (1996) En Neidhart, F. C. et al. *Escherichia coli* and *Salmonella*. Cellular and Molecular Biology, ASM Press USA. pag. 1287-1297.
- Chubatsu 1999
- Coulombe, B. & Burton, Z.F. (1999). Microbiol. Mol.Biol.Rev. **63**(2):457-478.
- Darst, S. A., Polyakov, A., Ritcher, C., & Zhang, G. (1998). J. Struct. Biol **124**:115-122.
- David, M., Daveran, M., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P. & Kahn, D. (1988) Cell **54**:671-683.
- Dávila, S. (1998). Análisis de la interacción de NifA de *R. meliloti* con las regiones promotoras UAS de los genes *nif* mediante mutagénesis dirigida, Tesis de Maestría en Ciencias Bioquímicas, IUNAM.
- Del Río, G., Osuna, J. & Soberón, X. (1994) Biotechniques **17**:1132-1139.
- DeMaio, J. Zhang, Y., KO, C., Young, D. B. & Bishai, W. R. (1996) Proc. Natl. Acad. Sci. USA **93**: 2790-2794.
- Dombroski, A. J., Walter, W., Record, M. T. Siegele, D. A. & Gross, C. (1992) Cell **70**:501-512.
- Estream, S. T., Ross, W., Gaal, T., Chen, Z. W., Niu, W., Ebrigh, R. & Gourse, R. (1999) Genes & Development. **13**:2134-2147.
- Ebrigh, R. H. & Busby, S. (1994) Current Op. Gen. Genet. **5**: 197-203.
- Dixon, R. 1998. Arch.Microbiol. **169**:371-380.
- Drummond, M. , Contreras, A. & Mitchenall, L. A. (1990) Mol Microbiol

4(1):29-30

- Drummond, M. , Whitty, P. & Wooton J. (1986) EMBO J. **5**: 441-447.
- Farr,S. B. & Koguma, T. (1991) Microbiol. Rev. **55**(4):561-585.
- Fischer, H.M. & Hennecke, H. (1987) Mol. Gen. Genet. **209**:621-626.
- Fischer, H. M. (1994) Microbiology Rev. **58**:352-386
- Fischer, H.M., Bruderer, T. & Hennecke, H.(1988) Nucleic Acids Res. **16**:2207-2224.
- Fischer, H.M., Acuna, G., Anthametten, D., Arigoni, F., Babst, M., Brawer, P. Kaspar, T., Kullik, I., Preisig, O., Scherb, B., Weidenhaupt, M., & Hennecke, H. (1992) En Palacios, R., Mora, J. y Newton, E. (eds). New Horizons in Nitrogen Fixation Kluwer Acad. Pub. London.
- Gallegos, M.T., Cannon, W. & Buck, M. (1999) J. Biol. Chem. **274**(36):25285-25290.
- Gilles-González, M. A., González, G. Perutz, M. F., Kiger, L., Marden , M. C. Poyart, C. (1994) Biochemistry **33**:8067-8073.
- Glass, R., Jones, S. T. & Ishihama , A. (1987) En Reznicoff, W. S., Burgues, R., Dahlberg. J. Gross. C. Record, T. and Wickens (eds). RNA Polymerase and the Regulation of Transcription. Elsevier, N.Y. pag. 393-399.
- González, V., Olvera, L., Soberón, X. & Morett. E. (1997). J. Bact **27**:
- Govantes, F. Molina-Lopez, J.A. & Santero, E. (1996) J. Bacteriol. **178**:6817-6823.
- Gralla, J.D. (1992) En McKnight, S & Yamamoto K.R. Transcriptional Regulation II. CSHL Press, N Y.629-642.

- Gross, C.A., Lonetto, M. & Losick, R. (1992). En McKnight, S & Yamamoto K.R. Transcriptional Regulation I. CSHL Press, N Y. 129-175.
- Gruber, T. & Bryant, D. (1997). J. Bacteriol. **179**:1734-1747.
- Guo, Y. & Gralla, J. (1997) J. Bacteriol **179**:1239-1245.
- Jacob, J. & Drummond, M. (1993). Mol.Microbiol. **10**(4): 813-821.
- Haldenwang, W. G. (1995) Microbiol. Rev. **59**:1-30
- Hecker, M., Schumann, W. & Volker, U. (1996) Mol. Microbiol. **19**:417-428.
- Hellman, J. D. (1991). Mol. Microbiol. **5**:2875-2882.
- Helmann, J. & deHaseth, P. (1999) Biochemistry **38**(19): 5959-5967.
- Henderson, N., Austin, S. & Dixon, R. (1989) Mol. Gen. Genet. **216**:484-491.
- Hennecke, H. (1990) Mol. Microbiol. **4**(10):1621-1628.
- Hidalgo, E. & Demple, B. (1994) EMBO J. **13**: 138-146.
- Hidalgo, E. Bollinger, J. M., Bradley, T. M., Walsh, C. T. & Demple, B. (1995) J. Biol. Chem. **270**: 20980-20914.
- Hirschman, J., Wong, K., Sei, J., Keener, J. & Kustu, S. (1985) Proc. Natl. Acad. Sci. USA. **82**:7525-7529.
- Hoover, T.R., Santero, E., Porter, S. & Kustu, S. (1990) Cell **63**:11-22.
- Hochschild, A. & Dove, S. (1998) Cell **92**:597-600.
- Huala, E. & Ausubel, F.M. (1989) J. Bacteriol. **171**:3354-3365.
- Huala, E., Stigter, J. & Ausubel, F.M. (1992) J. Bacteriol. **174**:1428-1431.
- Iismaa, I. & Watson, M. (1989) Mol. Microbiol. **3**:943-945.
- Ishihama, A. (1988) TIG. **4**(10):282-286.

- Ishihama, A. (1992) *Mol. Microbiol.* **6**(22):3283-3288
- Kawamura, S., Kakuta, Y., Tanaka, I. Hikichi, K., Kuhara, S, Yamasaki, N. & Kimura, M. (1996) *Biochemistry* **35**:1195-11200.
- Khoroshilova, N., Beinert, H. & Kiley, P. J. (1995) *Proc. Natl. Acad. Sci. USA* **92**: 2499-2503.
- Khoroshilova, N., Popescu, C. Münck, E, Beinert, H. & Kiley, P. J. (1997) *Proc. Natl. Acad. Sci. USA* **94**: 6087-6092.
- Kullik, I., Hennecke, H. y Fischer, M. (1989) *Arch. Microbiol.* **151**:191-197.
- Kustu, S., Santero, E., Keener, J., Popham, D. & Weiss, D. (1989) *Microbiol. Rev.* **53**:367-376.
- Kustu, S., Northm A.K. & Weiss, D. S. (1991) *Trends Biochem. Sci.* **16**:397-402.
- Langdon, R.C. & Hochschild, A. (1999) *Proc. Natl. Acad. Sci. USA* **96**:12673-12678
- La Vallie, E. & McCoy, J. (1995). *Curren Opinion in Biotechnology* **6**:501-506
- Laemmli, U. K. (1970) *Nature. London.* **227**:680-685.
- Lee, J.H. & Hoover, T. R. (1995). *Proc. Natl. Acad. Sci. USA* **92**:9702-9706.
- Lee, H. S., Berger, D.K. & Kustu, S.(1994) *Proc. Natl. Acad. Sci. USA* **90**: 2266-2270
- Liang, S., Bipatnath, M., Xu, Y., Chen, S., Dennis, P., Ehrenberg, M. & Bremer, H. (1999) *J. Mol. Biol.* **292**(1):19-37.
- Loewen, P.C. & Hengge-Aronis, R. (1994) *Annu. Rev. Microbiol.* **48**:53-80.

- Lonetto, M. A., Brown, K. L., Rudd, K. E. & Buttner, J. M. (1994) Proc. Natl. Acad. Sci. USA **91**: 7573-7577.
- Lonetto, M. A., Gribskov, M. & Gross, C. A. (1992) J. Bacteriol. **174**:3843-3849.
- Malhotra, A., Severinova, E. & Darst, S. A. (1996) Cell **87**:127-136.
- Majors, J. (1975) Proc. Natl. Acad. Sci. USA **72**: 4394-4398.
- Martínez, E., Romero, D. & Palacios, R. (1997) Crit. Rev. Plant.Sci.**9**:59-63
- McClure, W. (1985). Ann. Rev. Biochem, **54**: 171-204.
- McClure & Hopes, (1987) En Reznicoff, WS, Burgues, R., Dahlberg. J. Gross. C. Record, T. and Wickens. RNA Polymerase and the Regulation of Transcription. Elsevier, N.Y.
- Meade, H. M., Long, R., Ruvkun, G., Brown, E & Ausubel, M. (1982) J.Bacteriol. **149**:114-122.
- Merrick, M. J. (1992) En Stacy, G., Burris, R. y Evans, H, (eds), Biological nitrogen fixation. Chapman Hall, New York. p.835-847.
- Merrick, M.J. & Stewart, D,P. (1985) Gene **35**:297-303.
- Messing, M. J. (1981). Nucleic Acids Res **9**:259.
- Monteiro, R.A., Sousa, E. M., Yates, G.M., Pedrosa, F.O. & Chubatsu, L.S. (1999) FEMS Microbiol. Lett. **180**:157-161.
- Morett, E. y Buck, M. (1989) J. Mol. Biol. **210**:65-77.
- Morett, E., Cannon, W. & Buck, M. (1988) Nucleic Acids Res. **16**:11469-11488.
- Morett, E., Fischer, H.M. & Hennecke, H. (1991) J. Bacteriol. **173**:3478-

3487.

- Morett, E. & Segovia, L. (1993) *J. Bacteriol.* 175(19):6067-6074.
- Morris, C.E., McGraw, N.J. Joho, K., Brown, J.E., Kement, Ling, M.L. and McAllister, W.T. (1987) En Reznicoff, WS, Burgues, R., Dahlberg. J. Gross. C. Record, T. and Wickens. RNA Polymerase and the Regulation of Transcription. Elsevier, N.Y.
- O'Halloran, T. (1993) *Science* **261**:715-725.
- Osuna, J., Soberon, X. & Morett, E. (1997). *Protein Science* **6**:543-555
- Pabo, C.O. & Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61**:1053-1095.
- Reyrat, J. M., David, M., DePhilip, P., Garnerone, A.M., Babut, J. y Boistard, P. (1993) en Palacios, R., Mora, J. y Newton, E. (eds). *New Horizons in Nitrogen Fixation* Kluwer Acad. Pub. London.
- Reznicoff, W.S, Deborah, S. , Debora, W. & Gross, C.A (1985). *Ann. Rev. Genet.* **19**: 355-387.
- Ross, W., Gosink, K., Salomon, K., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. & Gourse, R.L. (1993) *Science* **262**:1407-1413.
- Ross, W., Aiyar, S. E., Salomon, J. & Gourse, R.L. (1998) *J. Bacteriol.* **180**(20): 5375-5383.
- Savery, N., Lloyd, G. S., Kainz, Gaal, T. Ross, W., Ebright, R., Gourse, R. & Busby, S. (1998) *EMBO J.* **17**(12):3439-3447.
- Schleif, R. (1996) En Neidhart, F. C. et al. *Escherichia coli* and *Salmonella*. Cellular and Molecular Biology, ASM Press USA. pag. 1300-1308.
- Scott, S., Busby, S. & Beachman, I. (1995). *Mol. Microbiol* **18**:521-531.

Monsalve y Hoschild 98

- Screen, S. Watson, J. y Dixon. R. (1994). Mol. Gen. Genet. **245**:313-322.
- Segovia, L., Piñero, D. Palacios, R. & Martínez, E. (1991) Appl. Environ. Microbiol. **57**: 426-433.
- Siegele, D. A. Hu, J.C., Walter, W. A. & Gross, C.A. (1989) J. Mol. Biol. **206**: 591-603.
- Sied, A. & Gralla, J. (1998). J. Bacteriol. **180**(21):5619-5625.
- Simon, R., Priefer, U. & Puhler, A. (1983). En Puhler, A (eds) Molecular Genetics of the bacteria-plant interaction. Springer-Verlag, KG. Berlin.
- Souza, E. M., Funuyama, S., Rigo, L. U. Yates, M. G. & Pedrosa, F. O. (1991) J. Gen. Microbiol. **137**: 1511-1522.
- Spiro, S. & Guest, J. (1994) Antonie Van Leeuwenhoek. **66**:23-36.
- Stookey. L (1970). Anal. Chem. **42**(7): 779-782.
- Struthers, M. D. Cheng, R. & Imperiali, B. (1996) Science. **271**: 342-344.
- Su, W., Porter, S., Kustu, S. y Echols, H. (1990) Proc. Natl. Acad. Sci. USA. **87**:5504-5508.
- Suzuki, M., Yagi, N, & Gerstein M (1995) Prot. Engineering **8**(4):329-338.
- Suzuki, M., Brenner, S., Gerstein, M. & Yagi, N,(1995) Prot. Engineering **8**(4):319-328.
- Svergun, D. I., Malfois, M., Koch, M., Wigneshweraraj, S.R. & Buck, M. (2000) J. Biol. Chem. **275**(6): 4210-4214.
- Syed, A. & Gralla, J. (1997) Mol. Microbiol. **23**: 987-995.
- Tagami, H. & Aiba, H.(1999) Proc. Natl. Acad. Sci. USA. **96** (13):7202-7207

- Taylor, M. Butler, R., Chambers, S. Casimiro, M., Badii, M. & Merrick, F. (1996) Mol. Microbiol. **22**:1045-1054.
- Thöny, B. y Hennecke, H. (1989) FEMS Microbiol. Rev. **63**:341-358.
- von Hippel, P. H., Bear, D, Morgan, W.D. & McSwiggen, J. (1984). Ann. Rev. Biochem, **53**:389-446
- von Hippel, P., Yager, T. & Gill, S.C. (1992) En McKnight, S & Yamamoto K.R. Transcriptional Regulation I. CSHL Press, N Y. 179-201.
- Voskuil, M.I., Voepel, K. & Chambliss, G. H. (1995) Mol. Microbiol. **17**:271-279.
- Waldburger, C., Gardella, T., Wong, R, & Susskind, M.M. (1990) J. Mol. Biol. **215**:267-276.
- Wang, W. & Gralla, J. D. (1996) J. Mol. Biol. **271**:32707-32713.
- Weiss, D.S., Batut, J., Klose, K.E., Keener, J. y Kustu, S. (1991) Cell **67**: 155-167.
- Whitehall, S., Austin, S. y Dixon, R. (1992) J. Mol. Biol. **225**:591-607.
- Wilson, C. & Dombroski, A. J. (1997). J. Mol. Biol. **267**:60-74.
- Woodland, M & Dalton H. (1984). J. Biol Chem. **259**(1):53-59.
- Wösten, M. M. (1998) FEMS Microbial. Rev. **22**(3):127-50
- Yura, T., Nagai, H. & Mori, H. (1993). Annu. Rev. Microbiol. **47**:321-350.
- Zhou, Y. N. & Gross, C. A. (1992) J. Bacteriol. **174**:7128-7137.

Reciprocal domain evolution within a transactivator in a restricted sequence space

Katy Juárez^{*†}, Humberto Flores^{*†}, Sonia Dávila^{*‡}, Leticia Olvera^{*}, Víctor González^{*§}, and Enrique Morett^{*¶||}

^{*}Instituto de Biotecnología, Universidad Nacional Autónoma de México, AP 510-3, Cuernavaca, Morelos, 62250, Mexico; and [†]European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117, Heidelberg, Germany

Edited by Sydney Kustu, University of California, Berkeley, CA, and approved January 18, 2000 (received for review October 15, 1999)

Although the concept of domain merging and shuffling as a major force in protein evolution is well established, it has been difficult to demonstrate how domains coadapt. Here we show evidence of coevolution of the *Sinorhizobium meliloti* NifA (SmNifA) domains. We found that, because of the lack of a conserved glycine in its DNA-binding domain, this transactivator protein interacts weakly with the enhancers. This defect, however, was compensated by evolving a highly efficient activation domain that, contrasting to *Bradyrhizobium japonicum* NifA (BjNifA), can activate in trans. To explore paths that lead to this enhanced activity, we mutagenized BjNifA. After three cycles of mutagenesis and selection, a highly active derivative was obtained. Strikingly, all mutations changed to amino acids already present in SmNifA. Our artificial process thus recreated the natural evolution followed by this protein and suggests that NifA is trapped in a restricted sequence space with very limited solutions for higher activity by point mutation.

Contemporary proteins are often assemblages of functionally and evolutionarily independent domains (1, 2). This modular architecture has conferred great flexibility for new specificities, altered recognition properties, and modified functions to flourish with a strikingly limited set of structurally different domains (3–6). But how the domains coadapt to achieve an optimal fitness is poorly understood. It is anticipated that appropriate domain interfaces and balanced activities need to be attained by a complex combinatorial optimization process (7) that is constantly in operation but limited by the evolutionary constraints inherent to the protein folds. Moreover, the optimal fitness needs to be dynamically maintained by compensatory mutations in a changing environment.

Earlier work showed that the enhancer-binding proteins (EBP), as the majority of transactivator proteins, are modular regulators with evolutionarily distinct DNA-binding, transcriptional activation, and regulatory domains (8–10). The DNA-binding and positive control functions of NifA, a member of the EBP family that controls nitrogen fixation gene expression in eubacteria, have been separated (11). The DNA-binding function resides at the C-terminal domain, whereas the activation domain is located at the center of the protein.

The EBPs bind to remote DNA sites, functionally similar to the eukaryotic enhancers, and activate transcription by contacting the σ^{54} form of the RNA polymerase, bound at the promoter, in a process that requires nucleoside triphosphate hydrolysis (12–15). The binding of the EBP at the enhancers may help to increase the local concentration of the activator in the vicinity of the promoter and to direct the central domain to interact in the correct orientation with $E\sigma^{54}$. Thus, the level of expression of a given promoter results from both the DNA-binding affinity and the intrinsic activation activity of the EBP.

The C-terminal region of several EBPs is predicted to form a helix–turn–helix (hth) supersecondary structure (8). Mutagenesis (16), spectroscopic (17), and NMR (P. Ray, K. J. Smith, R. A. Dixon, and E. I. Hyde, personal communication) studies of NifA support this interpretation. This motif is present in a wide range of site-specific DNA-binding proteins (18). When we compared the hth of several different NifA proteins we observed

that the protein of *Sinorhizobium meliloti* NifA (SmNifA) has a glutamic acid (Fig. 1B) in a position where glycine (G10) is highly conserved (18) (Fig. 1C). In this study, we report the detrimental effect of the lack of the glycine on DNA binding, which, in SmNifA, was compensated by developing a highly efficient activation domain, in a process we named reciprocal domain evolution.

Materials and Methods

Bacterial Strains and Plasmids. *Escherichia coli* JM101 strain was used for all experiments except for activation in trans, where ET8894 Δ glnA Δ nifBC was used to avoid cross-activation of the *nifH* promoter by NtrC, as described (19). Plasmid pRJ7511 (20) carries the *Bradyrhizobium japonicum* *nifA* gene or its derivatives. Plasmid pACYCNifA carries the *B. japonicum* *nifA* gene constitutively expressed from the *cat* promoter. This plasmid was constructed by inserting a *Bam*HI–*Pst*I fragment from pRJ7511 into pACYC177. Plasmids pKK*nifH* and pVB007 are derivatives of pKK232–8 (21) and carry the *cat* gene under the control of the *Klebsiella pneumoniae* and *S. meliloti* (deleted of the enhancer) *nifH* promoters, respectively. Plasmids pRT22 (22) and pMB210 (23) carry the *K. pneumoniae* and *S. meliloti* *nifH* promoter regions, respectively, fused to *lacZ* gene. Plasmid pUCNifA was constructed by subcloning the entire *S. meliloti* *nifA* gene, or its mutant derivatives, into pUC19. Plasmid pCU101 (19) carries the *S. meliloti* *nifH* promoter fused to the *lacZ* gene, and pSU003 is a derivative from pCU101 without the enhancer.

β -Galactosidase Assays. Strains carrying the different plasmids were grown in modified NFDM medium as described (20), at 30°C in aerobic or microaerobic conditions until they reached an optical density of 0.4–0.6 at 600 nm, as described (19).

In Vivo Dimethyl Sulfate (DMS) Footprinting. The accessibility of *K. pneumoniae* *nifH* promoter DNA to DMS was performed *in vivo* as described (19). A ³²P-5'-labeled synthetic oligonucleotide priming upstream of the *nifH* promoter was extended with 0.5 unit of the Klenow fragment of DNA polymerase for 10 min at 50°C, and the products were analyzed on sequencing gels.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: EBP, enhancer-binding proteins; SmNifA and BjNifA, NifA from *Sinorhizobium meliloti* and *Bradyrhizobium japonicum*, respectively; MIC, minimal inhibitory concentration; DMS, dimethyl sulfate; hth, helix–turn–helix; Cm, chloramphenicol.

[†]K.J. and H.F. contributed equally to this work.

[‡]Present address: Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Mexico, D.F.

[§]Present address: Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, 62210, Mexico.

[¶]To whom reprint requests should be addressed at the * address. E-mail: emorett@ibt.unam.mx.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.060444897. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.060444897

Immunoblotting Techniques. *E. coli* cells, expressing the *nifA* mutant derivatives, were cultured as described for β -galactosidase. Cells were pelleted, suspended in SDS sample buffer, and incubated for 10 min at 90°C. Detection was carried out as described previously (19). Two peptides, corresponding to the sequences of the hth of BjNifA (QAKAARLLGLTPROVGY) and SmNifA (QAKAARILEKTPROVGY), were synthesized and conjugated to BSA. New Zealand rabbits were immunized with these products to produce the specific polyclonal antibodies.

Site-Directed and PCR Mutagenesis. When required, single mutant oligonucleotides were synthesized to generate the specific mutants by PCR site-directed mutagenesis or by replacing the appropriated restriction fragment, as described (19). PCR mutagenesis was performed essentially as described (24). The PCR products were ligated into pACYCNifA, pRJ7511, or pUCNifA plasmids, which were restricted with the appropriate enzymes to replace the wild type with mutant fragment.

Selection and Screening Systems. *E. coli* JM101 cells harboring pKK*nifH* plasmid were electrotransformed with pools of mutant *nifA* genes in pACYCNifA. Expression of the *cat* gene in the former plasmid is activated by the BjNifA derivatives coded in the latter plasmid. The minimal inhibitory concentration (MIC) of chloramphenicol (Cm) for cells that carried the parental gene was determined, and a higher concentration was used in plates to select for mutants with enhanced transcriptional activity, as indicated below.

Sensitivity of NifA Mutant Proteins to Oxygen. *E. coli* cultures with the appropriate plasmids were grown in NFDM under microaerobic conditions. When cultures reached an OD of 0.4–0.6 at 600 nm, they were shifted to aerobic conditions, as described (20). Every 20 min a 2-ml aliquot was taken to measure β -galactosidase activity and the amount of the NifA derivatives by immunoblotting.

Results and Discussion

G10 of the hth Is Critical for DNA Binding. We compared the DNA-binding properties of several NifA proteins by *in vivo* footprinting. We observed that the protein of *S. meliloti* interacts very weakly with the enhancer of the *nifH* gene, compared with *B. japonicum* NifA (BjNifA) and *K. pneumoniae* NifA (KpNifA), such that it protected very weakly guanine-136 of the enhancer from methylation by DMS (Fig. 1A). To assess whether the lack of the conserved glycine G10 in the hth causes the weak binding, we replaced the glutamic acid with glycine by site-directed mutagenesis, resulting in SmNifA_{E10G} (Fig. 1B). Additionally, a second derivative with a lysine-to-leucine substitution at position 11 (SmNifA_{E10G/K11L}) was constructed to make the hth motif more similar to that of the other NifA proteins (Fig. 1B). *In vivo* footprinting analysis showed that both mutant proteins protected the *nifH* enhancer from methylation, indicating that they now bind strongly to the DNA (Fig. 2A) and that, indeed, the lack of G10 impairs this function. In the cocrystal structure of several hth proteins with their DNA-binding sites, G10, the first amino acid of the turn, is facing opposite to the DNA (25) (Fig. 1C). Thus, it is unlikely that residue 10 of the hth motif of NifA could make a direct contact with the DNA; the weak binding of SmNifA perhaps is due to an indirect effect, such as a more rigid or improperly oriented motif.

Transcriptional Activation Activity of BjNifA and SmNifA. Notwithstanding the weak DNA binding, SmNifA activates efficiently *nif* gene expression (23) (it is a natural protein after all), and the SmNifA_{E10G} and SmNifA_{E10G/K11L} mutants showed a very clear but moderate increase in transcriptional activation when strongly bound at the enhancer (Fig. 2D). These results suggest that

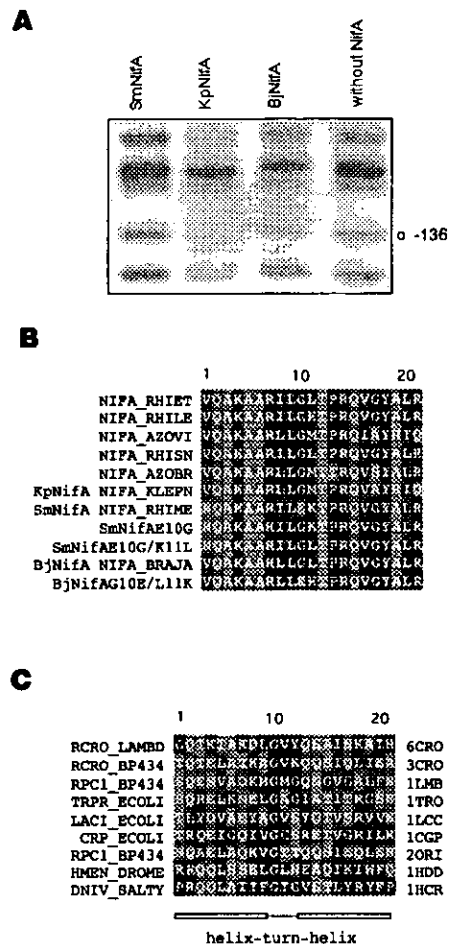


Fig. 1. DNA-binding properties of different NifA proteins and amino acid sequence of various hth motifs. (A) *In vivo* DMS footprinting of the *K. pneumoniae nifH* enhancer with different NifA proteins, as indicated. Protection from methylation of guanine-136 by KpNifA and BjNifA is indicated. This residue is part of the TGT-N₁₀-ACA *nifH* enhancer. (B) Sequence alignment of the hth motif of several NifA and the SmNifA and BjNifA mutant proteins constructed in this work. The SwissProt name for each protein is indicated. (C) Alignment of the hth motifs of proteins whose structures have been solved as cocrystals with their DNA-binding sites. SwissProt (first column) and Protein Data Bank (last column) names are specified. 6CRO, lambda cro; 3CRO, 434 cro; 1LMB, lambda repressor; 1TRO, *trp* repressor; 1LCC, *lac* repressor; 1CGP, catabolite activator protein (cap); 2ZOR, 434 repressor; 1HDD, *Drosophila* homeodomain protein; 1HCR, Hin recombinase. A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

SmNifA has a highly efficient transcriptional activation function that allows *nif* gene expression even when weakly bound to the enhancers. To test this hypothesis we carried out two different experiments.

In the first experiment we altered the hth motif of BjNifA to make it as that of SmNifA by constructing a G10-to-E and L11-to-K (BjNifA_{G10E/L11K}) derivative (Fig. 1B). We rationalized that if BjNifA has a similar transcriptional activation activity as SmNifA, impairing its DNA-binding properties would result only in a moderate reduction of *nifH* expression, whereas if it is less active, then it should be more dependent on a strong interaction with the enhancers. As anticipated, the mutations strongly impaired the binding to the *nifH* enhancer (Fig. 2A) and, in contrast to SmNifA, the mutant protein failed to activate *nifH* gene expression (Fig. 2E). Immunodetection of the NifA deriv-

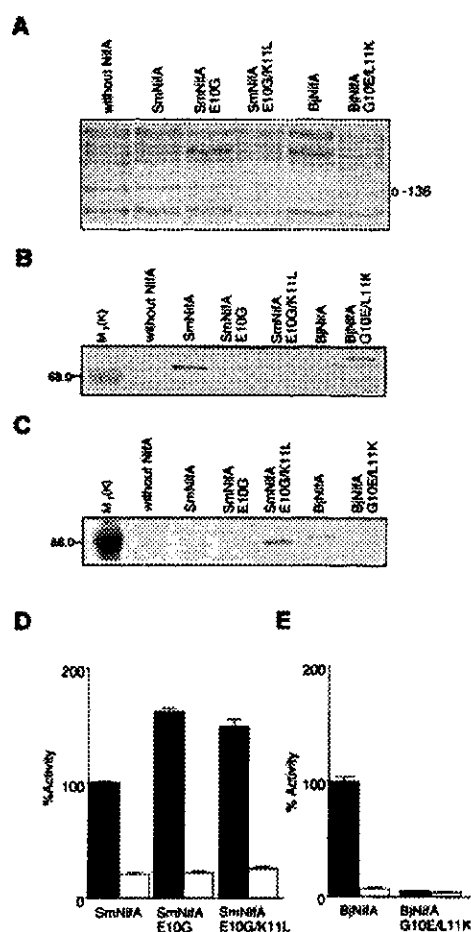


Fig. 2. Analysis of DNA binding, protein stability, and transcriptional activation of SmNifA, BjNifA, and their mutant derivatives. (A) Footprinting *in vivo* of the *K. pneumoniae nifH* enhancer with different NifA proteins and mutant derivatives, as indicated. Guanine-136 of the *nifH* enhancer is marked. (B and C) Immunodetection of the SmNifA and BjNifA (see *Materials and Methods*) proteins and their mutant derivatives. Antibodies raised against synthetic peptides corresponding to the hth of SmNifA or BjNifA were used for the immunodetection shown in B and C, respectively. M, values of control proteins are denoted. Note that the anti-SmNifA hth peptide antibodies did not react against either BjNifA or SmNifA_{E10G/K11L} and only weakly against SmNifA_{E10G}, but efficiently recognized BjNifA_{G10E/L11K}. Conversely, antibodies raised against BjNifA hth peptide did not recognize either SmNifA or BjNifA_{G10E/L11K} and reacted weakly against SmNifA_{E10G}, but efficiently recognized SmNifA_{E10G/K11L}. (D and E) Transcriptional activation of a *nifH-lacZ* fusion carrying (solid bars) or deleted of (open bars) the enhancer by different NifA proteins and their mutant derivatives, as indicated. Because SmNifA and BjNifA are coded in different vectors and expressed from different promoters, it is not feasible to correlate actual β -galactosidase activities. Values represent the mean \pm SD of three independent experiments and are expressed as percentage of β -galactosidase activity in Miller units.

atives showed that the mutations in the hth motif did not significantly affect the stability of any of the proteins (Fig. 2 B and C). Interestingly, the mutations selectively inverted the immunoreactivity of the SmNifA and BjNifA derivatives to antibodies raised against synthetic peptides comprising the hth of SmNifA (Fig. 2B) and BjNifA (Fig. 2C).

In the second experiment we directly compared the activation function of each protein by assaying their ability to activate transcription from solution, that is, to activate promoters devoid of the enhancer. Promoters that form strong, closed complexes with the RNA polymerase σ^{54} , such as the *S. meliloti nifH*

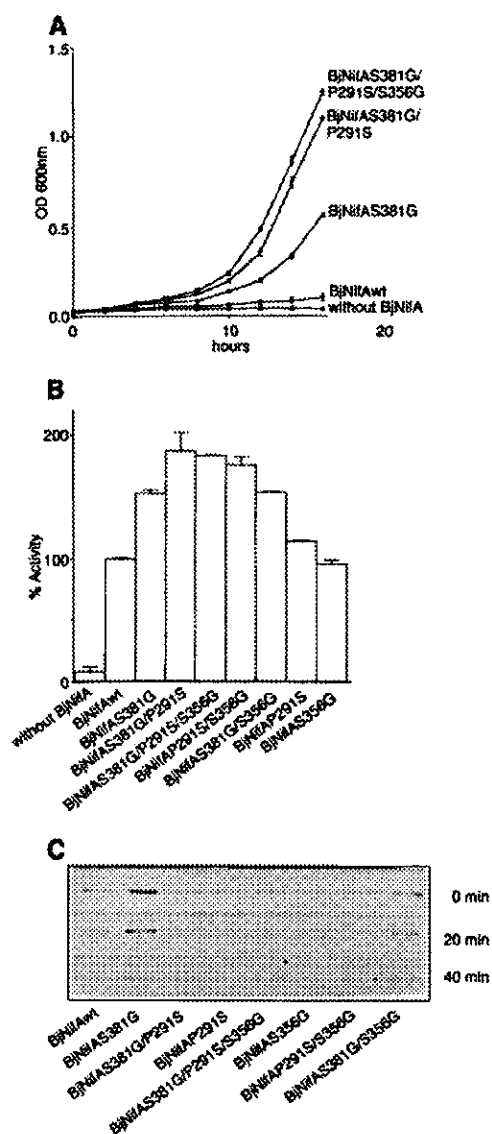


Fig. 3. Transcriptional activity, stability, and sensitivity to oxygen of the BjNifA derivatives. (A) Growth curves of strains carrying the wild type or the directly evolved BjNifA derivatives, as indicated. Cells were grown on NFD medium with Cm (40 μ g/ml) under aerobic conditions. The level of expression of the *cat* gene in pVB007, under the *S. meliloti nifH* promoter (deleted of the enhancer), and, therefore, the resistance to Cm is dependent on the activity of each BjNifA derivative. Values represent the mean \pm SD of three independent experiments. (B) Transcriptional activation of a *nifH-lacZ* fusion by the BjNifA mutant derivatives. (C) Stability and sensitivity of BjNifA mutant proteins to oxygen. To analyze whether the mutations affected the stability of the protein or any intrinsic function leading to a higher transcriptional activation, we grew strains carrying each of the mutant BjNifA proteins in microaerobic cultures, and the amount of the protein was detected after being shifted to heavily aerated flasks. Samples were taken at time 0 min, 20 min, and 40 min, and the BjNifA derivatives were immunodetected in soluble cell extracts with an antibody raised against a polypeptide comprising part of the central domain. All mutant proteins had similar stability and oxygen sensitivity because they decreased at the same rate as the wild type after being shifted to air.

(*SmnifH*), can be activated partially by NifA when deleted from the enhancers (26). If SmNifA is less dependent on the binding to the DNA, we rationalized that it would be able to activate at higher levels than BjNifA, a *SmnifH* promoter lacking the

enhancer. Fig. 2D shows that SmNifA significantly activates this promoter (at about 25% of the intact promoter), in sharp contrast to BjNifA, which was almost totally dependent on the enhancer (Fig. 2E). This result indicates that, although SmNifA is still dependent on the binding to the enhancers for full activation [required for fidelity of activation (27)], indeed it has a higher transcriptional activation activity than BjNifA.

Directed Evolution of the Central Domain of BjNifA. How did SmNifA achieve its higher transcriptional activity? Perhaps this was selected to compensate for the weak DNA binding caused by the G10-to-E mutation, although it is also possible that a lower DNA-binding activity was selected in response to an enhanced activation activity detrimental to the cell. Reconstructing the sequence of ancestral proteins by using sequence information from closely related extant proteins has helped to identify residues responsible for enhanced activity (28). Alternatively, directed evolution (29) can provide information about relevant positions critical for enhanced activity. To explore possible paths that lead SmNifA to its higher transcriptional activity, we subjected BjNifA to a directed evolution process. We developed a positive selection scheme for NifA mutants displaying enhanced ability to activate the expression of a Cm resistance gene under the control of a *KpnI*H promoter. The activation domain of NifA was mutagenized by random, error-prone PCR, and the derivatives conferring increasing resistance to Cm were selected. About 1×10^5 clones were analyzed in each cycle. The majority of the mutations impaired the function, as observed for other proteins (30), about 30% retained some activity (as assayed for their ability to activate a *nifH-lacZ* fusion), and only very few grew on Cm. In the first cycle three clones were selected on 10- μ g/ml Cm (the MIC of the BjNifA wild type is 8 μ g/ml), and the DNA fragment corresponding to their activation domains sequenced. Remarkably, all had an amino acid replacement in common, S381 to G. Two of them had other synonymous substitutions (BjNifA_{S381G/A377A/L342L}). Because the activity of the mutant derivatives was very similar, the clone with the single mutation (BjNifA_{S381G}) was chosen and subjected to a second cycle of mutagenesis and selection on 15- μ g/ml Cm plates. Only one colony grew faster than the parental strain. This clone, BjNifA_{S381G/P291S}, which had an additional P291-to-S substitution, was mutagenized further, and only a single clone of the third generation (BjNifA_{S381G/P291S/S356G}) grew on plates containing 20 μ g/ml Cm. Sequencing of their entire activation domain showed that it had S356 to G as the only additional substitution. Fig. 3A shows growth curves of the strains with the wild type and each of the mutant BjNifA proteins, which successively attained a faster growth in aerobic liquid cultures containing Cm, as higher expression of a *nifH-lacZ* fusion (Fig. 3B).

To assess the contribution of each mutation to the enhanced activity we segregated them and constructed the combination of all double mutants by different molecular genetic tech-

niques. Their ability to activate the expression of a *nifH* promoter fused to *lacZ* was determined as β -galactosidase activity in microaerobic cultures in NFD. In this assay the second- and the third-generation mutants (BjNifA_{S381G/P291S} and BjNifA_{S381G/P291S/S356G}) showed similar activities (about twice the wild type), whereas the mutant of the first generation (BjNifA_{S381G}) was more active than any of the other two single mutants (BjNifA_{P291S} and BjNifA_{S356G}), which themselves had a similar activity as the wild type. The double mutant, BjNifA_{S381G/P291S}, was slightly more active than BjNifA_{P291S/S356G} and BjNifA_{S381G/S356G}. These results are consistent with the additive effect of the mutations (Fig. 3A) and with the fact that the S381-to-G mutation was the only one selected in the first generation appearing in three independent clones.

Oxygen Stability of the Mutant BjNifA Derivatives. To test whether the mutations affected the stability or a catalytic property of the mutant proteins, the relative concentration of the NifA derivatives was determined by immunoblotting. As shown in Fig. 3C, the mutant proteins were present at levels similar to the wild-type NifA. Our previous studies on the regulation of NifA activity showed that BjNifA is oxygen-labile (20). To determine the oxygen sensitivity of the mutant proteins, cultures harboring each of the mutant BjNifA derivatives were grown overnight microaerobically and shifted to heavily aerated conditions. Samples were taken at 20-min intervals and assayed for the fraction of protein remaining. Fig. 3C shows that the mutant BjNifA derivatives decreased at about the same rate as the wild type. Thus, this experiment indicates that the mutations did not significantly affect the oxygen stability of the mutant BjNifA derivatives.

NifA Is Trapped in a Restricted Sequence Space. Despite the inexorable limitation of any attempt to explore the sequence space of a protein, the mutagenesis strategy and the number of independent BjNifA derivatives analyzed here allowed us to screen a great number of mutants, including all possible single and a considerable proportion of double mutants in the central domain. Therefore, we were surprised to find that, from the many possible substitutions, all those with higher activity changed to amino acids that are already present in the highly active SmNifA (Fig. 4) and normally are the most frequent amino acids at these positions in other EBP (9, 10). Thus, our directed evolution process led to the same solutions for enhanced activity as SmNifA.

Our attempts to further increase the activity of the central domain by the same mutagenic procedure were unsuccessful. Segregation by *in vitro* recombination of all individual mutations [gene shuffling (31)], with further mutagenesis, resulted only in the regeneration of the triple-mutant clone (with one additional mutation that did not affect the activity) after selecting on 25- μ g/ml Cm plates. Thus, it is likely that with these three mutations the protein attained a local optimum in a sequence

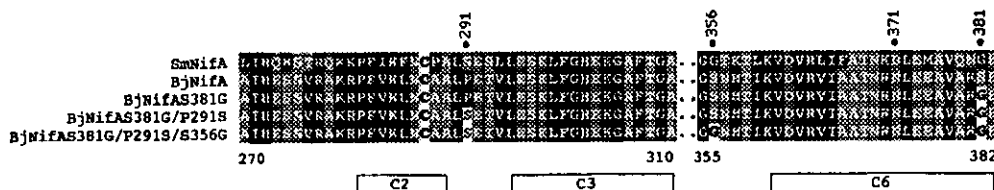


Fig. 4. Multiple sequence alignment of SmNifA, BjNifA, and its mutant derivatives. Only regions near the mutated residues on BjNifA are shown. Dots indicate discontinuity of the alignment. Numbers above the sequences indicate amino acid positions mutated in BjNifA. Numbers below the alignment show amino acid positions of BjNifA. Boxes indicate the highly conserved regions in the EBP proteins (9, 10). The central domains of BjNifA and SmNifA are 68% identical, whereas the degree of identity of the whole family is around 40%.

space accessible by point mutation [one-mutant distance neighborhood (32)].

That we found the same solutions for higher activity as natural proteins suggests that NifA is trapped in a very narrow activity "peak" on its fitness landscape. Is it possible to move NifA to a different activity "peak"? It has been proposed that multiple mutations can make a protein "jump beyond" to a new activity "peak" in its fitness landscape (32). We currently are searching for suppressor derivatives of stable mutants of BjNifA with residual activity in an attempt to explore new solutions on its sequence space or, following the metaphor of the fitness landscape (32), to move BjNifA to different activity "peaks."

Conclusions

The results presented here show that SmNifA has found a novel optimum fitness, perhaps by increasing its activation function in response to a mutation that impaired DNA binding. Alternatively, a weak DNA binding could have been selected in response to an unfavorable high transcriptional activation function. Whatever the order of the events was, it is clear that neither this protein nor BjNifA is at its maximum ability to

activate gene expression. Taken together, our results show that a modular protein can stay around its optimal fitness by dynamically balancing the activities of its different domains, a process we called reciprocal domain evolution. We propose that continuous compensatory mutations allow the exploration of the sequence space, which occasionally might result in novel activities. Thus, it is likely that reciprocal domain evolution has been a strong force in the sharpening of the contemporary proteins.

We are grateful to M. Olvera, R. Hernández, and S. Caro for technical assistance, P. Gaytan and E. López for oligonucleotides synthesis, T. Olamendi for polypeptides synthesis, E. Mata and G. Cabeza for assistance in rabbit immunizations, B. Valderrama for pVB007 plasmid, and X. Soberón, A. Alagón, A. Garcíarrubio, M. Huynen, and P. Bork for fruitful discussions. E.M. thanks the Consejo Nacional de Ciencia y Tecnología and Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México, for support for a sabbatical leave. K.J. was a recipient of a Consejo Nacional de Ciencia y Tecnología scholarship. E.M. was an Alexander von Humboldt fellow. This research was supported in part by grants from Consejo Nacional de Ciencia y Tecnología (G0030-N9607) and The European Commission (C11*-CT94-0060).

1. Henikoff, E. S., Greene, A., Pietrokovski, S., Bork, P., Attwood, T. K. & Hood, L. (1997) *Science* **278**, 609–614.
2. Doolittle, R. F. (1995) *Annu. Rev. Biochem.* **64**, 287–314.
3. Doolittle, R. F. & Bork, P. (1993) *Sci. Am.* **269**, 50–56.
4. Netzer, W. J. & Hart, F. U. (1997) *Nature (London)* **388**, 343–349.
5. Chothia, C. (1992) *Nature (London)* **357**, 543–544.
6. Orengo, C. A., Jones, D. T. & Thornton, J. M. (1994) *Nature (London)* **372**, 631–634.
7. Kauffman, S. A. & Weinberger, E. D. J. (1989) *Theor. Biol.* **141**, 211–245.
8. Drummond, M., Whitty, P. & Wootton, J. (1986) *EMBO J.* **5**, 441–447.
9. Morett, E. & Segovia, L. (1993) *J. Bacteriol.* **175**, 6067–6074.
10. Osuna, J., Soberón, X. & Morett, E. (1997) *Protein Sci.* **6**, 543–555.
11. Lee, H. S., Berger, D. K. & Kustu, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2266–2270.
12. Kustu, S., Santero, E., Keener, J., Popham, D. & Weiss, D. (1989) *Microbiol. Rev.* **53**, 367–376.
13. Merrick, M. J. (1993) *Mol. Microbiol.* **10**, 903–909.
14. Rippe, K., von Hippel, P. H. & Langowski, J. (1995) *Trends Biochem. Sci.* **20**, 500–506.
15. Wyman, C., Rombel, J., North, A. K., Bustamante, C. & Kustu, S. (1997) *Science* **275**, 1658–1661.
16. Morett, E. & Buck, M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9401–9405.
17. Missailidis, S., Jaseja, M., Ray, P., Chittock, R., Wharton, C. W., Drake, A. F., Buck, M. & Hyde, E. I. (1999) *Arch. Biochem. Biophys.* **361**, 173–182.
18. Wintjens, R. & Rooman, M. (1996) *J. Mol. Biol.* **262**, 294–313.
19. González, V., Olvera, L., Soberón, X. & Morett, E. (1998) *Mol. Microbiol.* **28**, 55–67.
20. Morett, E., Fischer, H. M. & Hennecke, H. (1991) *J. Bacteriol.* **173**, 3478–3487.
21. Brosius, J. (1984) *Gene* **27**, 151–160.
22. Tuli, R. & Merrick, M. J. (1988) *J. Gen. Microbiol.* **134**, 425–432.
23. Better, M., Ditta, G. & Helinski, D. (1985) *EMBO J.* **4**, 2419–2424.
24. Tintut, Y. & Gralla, J. D. (1995) *J. Bacteriol.* **177**, 5818–5825.
25. Susuki, M., Brenner, S., Gerstein, M. & Yagi, N. (1995) *Protein Eng.* **8**, 319–328.
26. Buck, M. & Cannon, W. (1992) *Mol. Microbiol.* **6**, 1625–1630.
27. Morett, E. & Buck, M. (1989) *J. Mol. Biol.* **210**, 65–77.
28. Opitz, J. G., Ciglic, M. I., Haugg, M., Trautwein-Fritz, K., Raillard, S. A., Jermann, T. M. & Benner, S. A. (1998) *Biochemistry* **37**, 4023–4033.
29. Kushner, O. & Arnold, F. H. (1997) *Trends Biotechnol.* **12**, 523–530.
30. Suckow, J., Markiewicz, P., Kleina, L. G., Miller, J., Kisters-Woike, B. & Muller-Hill, B. (1996) *J. Mol. Biol.* **261**, 509–523.
31. Stemmer, W. P. (1994) *Nature (London)* **370**, 389–391.
32. Kauffman, S. & Levin, S. (1987) *J. Theor. Biol.* **128**, 11–45.