

03081

2
201



UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO

INSTITUTO DE BIOTECNOLOGIA

"ANALISIS MOLECULAR DE LOS SISTEMAS DE
REGULACION TRANSCRIPCIONAL QUE CONTROLAN LA
EXPRESION DEL OPERON *fixRnifA* DE *Bradyrhizobium*
japonicum"

T E S I S
QUE PARA OBTENER EL GRADO DE:
DOCTOR EN INVESTIGACION BIOMEDICA BASICA

PRESENTA:

HUMBERTO BARRIOS CAMACHO

CUERNAVACA, MORELOS

JUNIO DE 1999

TESIS CON
FALLA DE ORIGEN

274650



Universidad Nacional
Autónoma de México



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.

PASINACION

DISCONTINUA

Agradecimientos

A los Dres. Jesus Aguirre, Miguel Angel Cevallos, David Romero y Gabriel Guarneros por su asesoría durante el desarrollo de este trabajo de investigación.

A los Dres. Enrique Morett, Brenda Valderrama, Gabriel Guarneros, Carmen Gómez, Guadalupe Espín, Alicia González, Luis Servín, Laura Camarena, Miguel Angel Cevallos y Edmundo Calva por la revisión y comentarios en la escritura de esta tesis.

En especial a Mariana Peimbert, por los buenos momentos que compartimos y por su sinceridad y apoyo en esta etapa de mi vida. A ella un abrazo y un beso con cariño.

A mis mejores amigos, a los cuales menciono sin un orden prioritario: Juan Téllez, Rodrigo Arreola, Ramon Suárez, Ernesto Pérez, Filiberto Sánchez, Humberto Flores, Manuel Rivera, Gloria Saab, Víctor Gonzalez, Ismael Hernández, Gabriel Mercado y Maricela Olvera. A todos ellos, por haber soportado mi mal carácter durante estos años. Por todos esos momentos buenos y malos de los cuales nos reimos y seguiremos haciendolo hasta viejos, como de otras tantas anécdotas.

A Leticia Olvera, por esta larga amistad y por su apoyo incondicional en la realización del trabajo experimental. A Ricardo Grande, por fumarsela conmigo a la hora de discutir de ciencia y por esos 10 años de amistad. A Brenda Valderrama por muchas cosas que no podría enumerar en un espacio tan pequeño.

A toda mi queridísima familia por el apoyo incondicional que me han brindado durante mi corta vida, aunque todavía crean que estoy loco.

A mi jefe Enrique Morett, por su paciencia y todo ese tiempo utilizado en mi formación. Este trabajo se realizó entre 1993 y 1994 en el departamento de Ecología Molecular del CIFN-UNAM y de 1995 a 1999, en el departamento de Reconocimiento Molecular y Bioestructura del IBT-UNAM, bajo la asesoría del Dr J. Enrique Morett S.

A todo los compañeros del laboratorio, que sin su libre pensamiento los días serian aburridos.

Durante mis estudios de posgrado fui becario de CONACYT, durante el período de 1994 a 1998.

Indice

Resumen	1
Summary	2
1. INTRODUCCION	3
1.1. La RNA polimerasa en bacterias.	4
1.2. El mecanismo del inicio de la transcripción.	7
1.3. Múltiples factores sigma.	11
1.3.1. La familia del factor σ^{70}	12
1.3.2. La familia del factor σ^{54}	14
1.4. Las regiones promotoras σ^{70} y σ^{54}	16
1.4.1. Promotores múltiples	20
1.5. Regulación de la transcripción	23
1.5.1 La familia de los activadores a distancia (EBP)	24
1.5.2.El mecanismo de transcripción en promotores σ^{54} .	27
1.6. Cascada de regulación de los genes <i>fix</i> y <i>nif</i> en <i>B.japonicum</i> .	29
2. ANTECEDENTES Y OBJETIVO.	33
3. RESULTADOS (Discusión y conclusiones)	36
3.1. Apéndice I.	36
3.2. Apéndice II.	41
3.3. Apéndice III (Datos adicionales).	45
3.4. Apéndice IV.	53
4. REFERENCIAS.	55

RESUMEN

La información acumulada del estudio de diferentes genes bacterianos, muestra que la regulación de la expresión genética frecuentemente se lleva a cabo al nivel de la transcripción. El estudio de la regulación transcripcional a permitido entender de una manera más clara los diferentes mecanismos utilizados para la expresión específica de los genes.

La contribución de este trabajo, fue demostrar la existencia de un mecanismo de regulación complejo que controlan la expresión del operón *fixRnifA* en *Bradyrhizobium japonicum*. Este mecanismo consta de dos sistemas de regulación transcripcional distintos, en los cuales intervienen dos proteínas activadoras (NifA y RegR) y dos polimerasas diferentes ($E-\sigma^{54}$ y una RNA- polimerasa de la familia de σ^{70}), las cuales reconocen a dos promotores sobrelapados en la misma región de DNA.

Por otro lado, demostramos que el mecanismo que utilizan estas polimerasas para el reconocimiento del promotor es distinto, así como su capacidad para la formación del complejo abierto de la transcripción.

SUMMARY

The information in the study of many different bacterial genes shown that the regulation of genetic expression is most frequently become at level of transcription initiation. The documented transcriptional regulation allow to understand the different mechanism used for the specific gene expression.

The original contribution of this research was the description of the complex regulation mechanism of the *fixRnifA* operon expression in *Bradyrhizobium japonicum*. That mechanism use two different transcriptional regulation systems that involve two activator proteins (NifA and RegR) and two different RNA-polymerases (E- σ^{54} and an RNA- polymerase of the σ^{70} family). These different polymerases recognize two overlapping promoters in the same DNA region.

By other hand our dates show that the RNA-polymerase mechanism for the promoter recognition was different and the same differences happen with transcriptional open complex melting.

INTRODUCCION

La regulación de la expresión genética tiene como finalidad el controlar la cantidad relativa de proteínas específicas que se producen en las diferentes etapas de actividad celular en respuesta a diferentes estímulos dentro y fuera de la célula. Esta expresión se regula basicamente desde el inicio de la transcripción hasta la postraducción del RNAm. En eubacterias la regulación de la transcripción es un mecanismo importante para controlar finamente la expresión de los genes.

Normalmente involucra contactos directos entre las proteínas reguladoras con la RNA-polimerasa, desde sitios de unión que se encuentran aledaños al promotor. En el caso de la holoenzima RNA-polimerasa que contiene el factor σ^{54} , las proteínas reguladoras se unen a sitios más distales del promotor para activar la transcripción mediante la formación de un asa que permite la interacción de la proteína reguladora con la misma RNA-polimerasa- σ^{54} . En secciones posteriores describire los elemento involucrados en la transcripción, para finalmente recalcar la importancia del estudio de los mecanismos de regulación transcripcional involucrados en el control del operón *fixRnifA* en *Bradyrhizobium japonicum*,.

1.1 La RNA polimerasa bacteriana

La síntesis de RNA se lleva a cabo por la enzima RNA-polimerasa, la cual cataliza la formación de enlaces fosfodiéster en dirección 5'-3' entre los ribonucleótidos a partir de un molde de DNA. Esta RNA-polimerasa es una proteína oligomérica, cuya forma funcional se denomina comúnmente como holoenzima RNA polimerasa (E- σ ó RNA-pol- σ) (figura 1). La holoenzima se encuentra dividida funcionalmente en dos partes (15):

a) La RNA polimerasa "core" (E ó RNA-pol^c), que está constituida por cuatro subunidades que son $\alpha_2\beta\beta'$. Esta RNA-pol^c tiene toda la maquinaria catalítica para la síntesis de una nueva cadena de RNA. El gene *rpoA* codifica para subunidad α que tiene una masa de 155,613 daltons, el gene *rpoB* codifica para la subunidad β de 150,618 daltons y el gene *rpoC* codifica para la subunidad β' de 36,512 daltons (67, 109).

b) El factor σ , cuyo papel es reconocer específicamente las regiones promotoras en el DNA (15). En *E. coli* el factor σ "más común" es el σ^{70} , que se encuentra codificado por el gene *rpoD* y tiene una masa de 70,263 daltons (16).

RNAI-pol + FACTOR σ = HOLOENZIMA

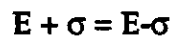
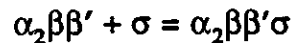


Figura 1. Esquema de la composición por subunidades de la holoenzima RNA polimerasa.

(Tomado de Doi, R.H. 1991)

Para que la holoenzima (RNA-pol- σ) se encuentre en su forma activa, es necesario que todas las subunidades se ensamblen de manera ordenada. Este comienza con la dimerización de las subunidades α , seguida por la integración de la subunidad β para conformar el complejo $\alpha_2\beta$ (46). Este primer paso es obligatorio para que se pueda asociar posteriormente la subunidad β' y ensamblar de esta manera a la RNA-pol^c (44). Una de las características particulares del ensamblaje, es la formación de una RNA-pol^c ($\alpha_2\beta\beta'$) inmadura, que enzimáticamente es inactiva y, aunque conformacionalmente es muy parecida a la forma nativa, presenta diferencias en sus propiedades fisicoquímicas (45). La maduración de esta enzima *in vivo* se establece cuando el factor σ se ensambla a la RNA-pol^c, dando lugar a una holoenzima madura y activa (RNA-pol- σ) (Figura 2) (47, 109).

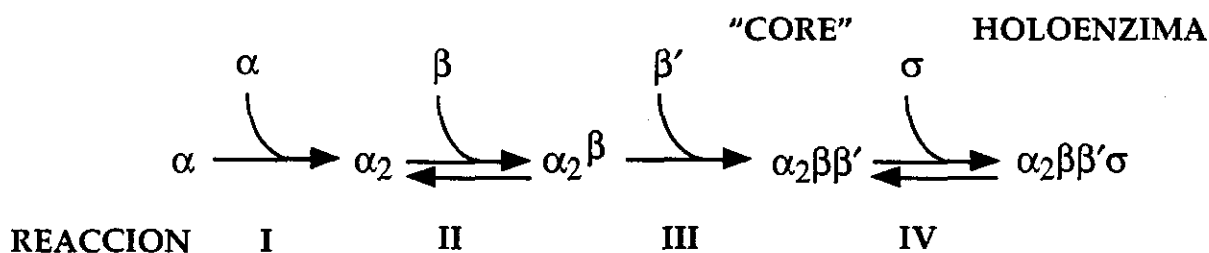


Figura 2. Orden de las diferentes etapas en la reacción de ensamblaje de la RNA-pol- σ .
(Tomado de Ishihama *et al.* 1979)

Algunas mutaciones en la subunidad β causan alteraciones en la especificidad de la transcripción; estas mutantes se encuentran probablemente alteradas en su capacidad de interactuar con el DNA o con algunos factores de regulación (109). La subunidad β cuenta con el sitio catalítico responsable de la formación de los enlaces fosfodiéster además de estar involucrada en unir el primer ribonucleótido durante la síntesis del RNA (109). Uno de los contactos de la RNA-pol- σ con el DNA se lleva a cabo por medio de la subunidad β' y se sabe que por sí sola se puede unir de manera estable al DNA (79). Diferentes evidencias experimentales postulan que las subunidades $\beta\beta'$ en conjunto dan lugar a la estructura llamada canal, la cual forma el sitio activo de la holoenzima. (21,80).

La región amino terminal de las subunidades α están involucrada en su dimerización, por otro lado su región carboxilo terminal está involucrada en el reconocimiento de la región promotora y con la interacción directa con los reguladores transcripcionales (41,53,75).

El factor σ determina la unión específica de la RNA-pol- σ al promotor, así como la determinación del inicio de transcripción, también llamado (+1). Es en esta posición donde se incorpora el extremo terminal 5' de la nueva cadena de RNA para iniciar su síntesis completa (24). Otra de las funciones que se proponen para el factor σ es la desnaturalización local del DNA en la región promotora, durante el primer paso del inicio de la transcripción (88).

La estructura tridimensional de la RNA-pol- σ de *E. coli* ha sido determinada por microscopía de fuerza atómica. La baja resolución de esta estructura, comparada con la obtenida para la DNA polimerasa, sirvió para describir algunas características estructurales de la holoenzima; siendo de las más sobresalientes la identificación de un canal de aproximadamente 25 Å de diámetro por 55 Å de longitud, con las dimensiones apropiadas para unir un segmento de 16 pares de bases de DNA (80). Este canal es muy similar al observado en la estructura obtenida por rayos X del sitio

de pegado al DNA de un fragmento de la DNA-polimerasa de *E. coli* (Klenow) (21). Además, se ha determinado la estructura de la RNA-pol^c por microscopía de fuerza atómica a una resolución de aproximadamente 23 Å. Esta estructura es muy parecida a la holoenzima RNA polimerasa II de levadura y muestra notables cambios conformacionales cuando se compara con la estructura de la RNA-pol- σ (Figura 3) (80).

Figura 3. Comparación de las diferentes estructuras: a) Modelo de la RNA polimerasa II de Levadura, b) Modelo de la RNA-pol^c "core" de *E.coli* y c) Modelo de la RNA-pol- σ "holoenzima" de *E.coli*.
(Tomado de Polyakov, A. *Et al.*1995)

1.2. Mecanismo del inicio de la transcripción: formación de los complejos cerrado y abierto de la transcripción.

Se ha considerado el mecanismo del inicio de transcripción como un proceso de tres pasos. Actualmente, utilizando evidencias cinéticas y estructurales, se ha propuesto como un mecanismo de cuatro pasos que se pueden ejemplificar con la siguiente formula:

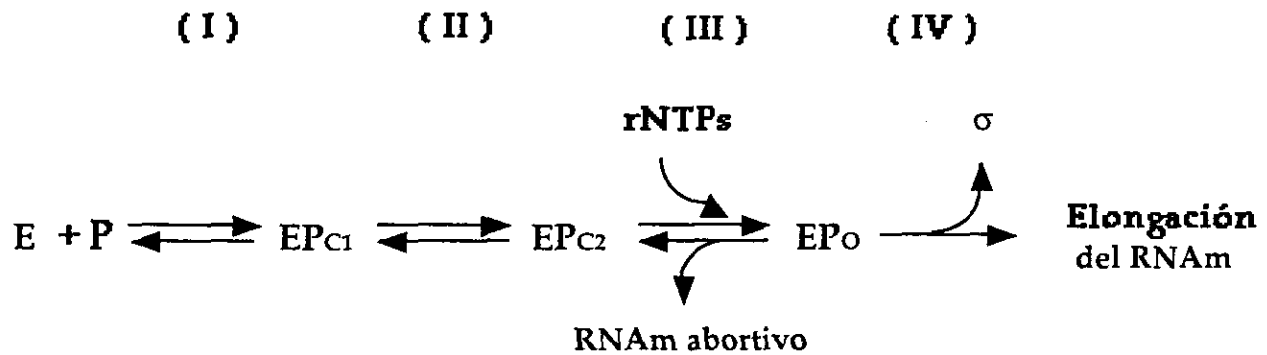


Figura 4. Mecanismo del inicio de la transcripción. E=holoenzima RNA-polimerasa, P=promotor, EP_{c1} =complejo cerrado de la transcripción, EP_{c2} =complejo intermediario, EP_o =complejo abierto de la transcripción. (Tomado de Record *et al.* 1996)

Con la comparación de las estructuras entre RNA-pol- σ y la RNA-pol^c de *E.coli* y basándose en evidencias experimentales, se ha propuesto un modelo que explica los diferentes cambios conformacionales que sufre la RNA polimerasa de *E.coli*, así como el mecanismo de transcripción (Figuras 4 y 5) (80,82).

En el primer paso (I) la RNA-pol^c se encuentra en solución y se considera que predomina bajo una conformación de elongación con la estructura del canal cerrada (Figura 3b). Posteriormente se da la unión del factor σ para formar a la holoenzima y favorecer la apertura de la estructura del canal. Durante este paso, se da la formación de un complejo reversible y específico entre la RNA-pol- σ y el DNA denominado "EP" (67,80). Posteriormente la holoenzima reconoce específicamente la secuencia promotora sobre la doble cadena de DNA formando el complejo cerrado inicial de la transcripción (EP_{c1}). Datos experimentales señalan que la formación del complejo (EP_{c1}) y la polimerasa libre llegan rápidamente a un equilibrio (98). Este modelo

corresponde al llamado de difusión lineal, en donde la holoenzima va recorriendo a todo lo largo el DNA celular hasta que localiza los sitios específicos de pegado que corresponden a las regiones promotoras. Wu y colaboradores han reportado evidencias del modelo de difusión lineal (67).

En el paso (II), durante la formación del complejo cerrado intermedio de la transcripción (Epc2), la holoenzima sufre cambios conformacionales. Estos cambios se dan cuando la estructura del canal se cierra completamente alrededor del DNA, abarcando desde la posición -12 a la +2 de la región promotora. Sobre esta región, la holoenzima puede abrir específicamente de 10 a 15 pares de bases que se encuentran en la vecindad del inicio de la transcripción de manera reversible (22,67). En el paso (III) se forma el complejo abierto de la transcripción (EP_o), comenzando con la unión del primer ribonucleótido complementario a la cadena molde de DNA en el inicio de la transcripción (+1); iniciando la polimerización de la molécula de RNA mensajero (RNAm). Este complejo (EP_o) funciona como un proceso reversible dando lugar a productos abortivos de la transcripción. El último paso de esta reacción (IV) comienza una vez que se han unido de 7 a 12 nucleótidos del RNAm, es entonces cuando el factor σ se despega de la región promotora y continúa la síntesis completa del RNAm (80).

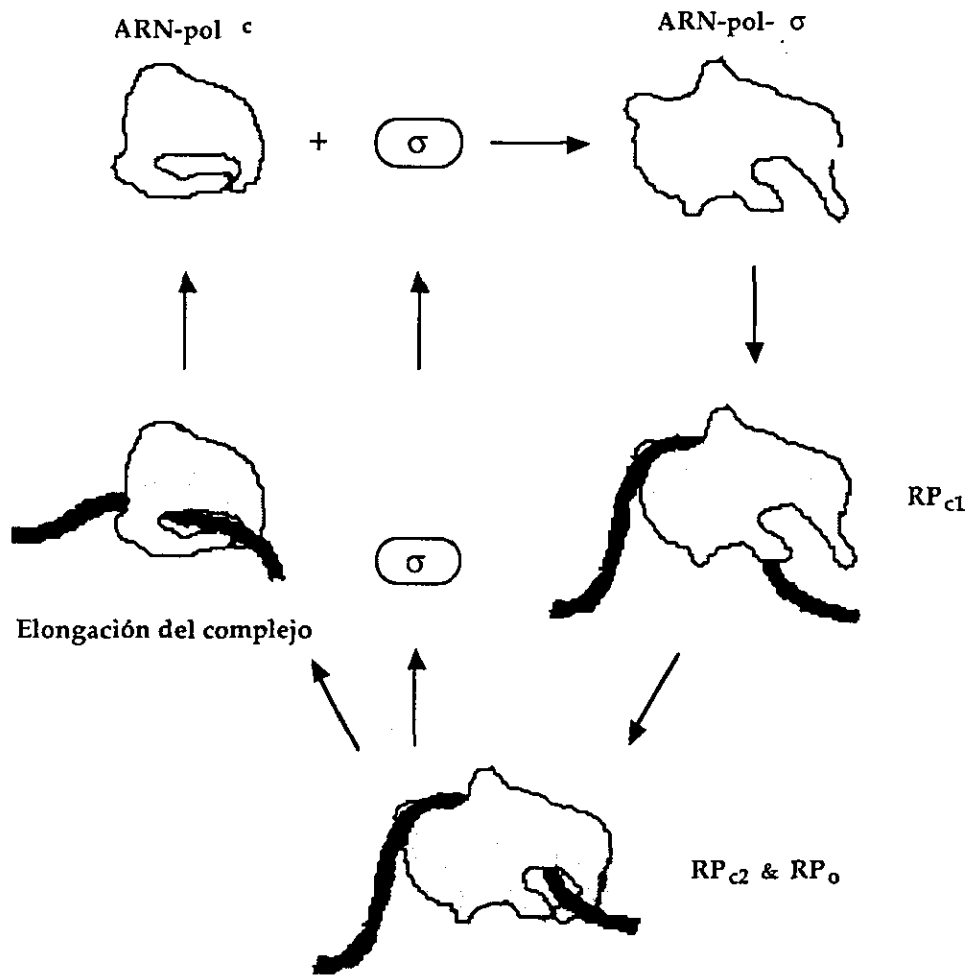


Figura 5. Modelo de los cambios estructurales que sufre la RNA polimerasa durante la transcripción.
 (Tomado de Polyakov *et al.* 1995)

1.3. Múltiples factores sigma

Por medio del análisis y comparación de las secuencias protéicas de los diferentes factores sigma así como de sus mecanismos de activación, éstos se han dividido en dos familias: aquéllos relacionados con la secuencia protéica de σ^{70} y aquellos relacionados con σ^{54} (35,38).

Dentro de la familia σ^{70} se encuentra el σ primario de mantenimiento en *E.coli* σ^{70} , que se encargan de transcribir todos los genes involucrados en funciones basicas de la bacteria. Por otro lado estan los sigmas alternativos que se encuentran involucrados en la expresión de grupos de genes específicos no esenciales para la sobrevivencia de la célula. En la figura 6 se muestra un ejemplo de los diferentes factores σ existentes en *E.coli*, así como algunas de las funciones en la que se encuentran involucrados

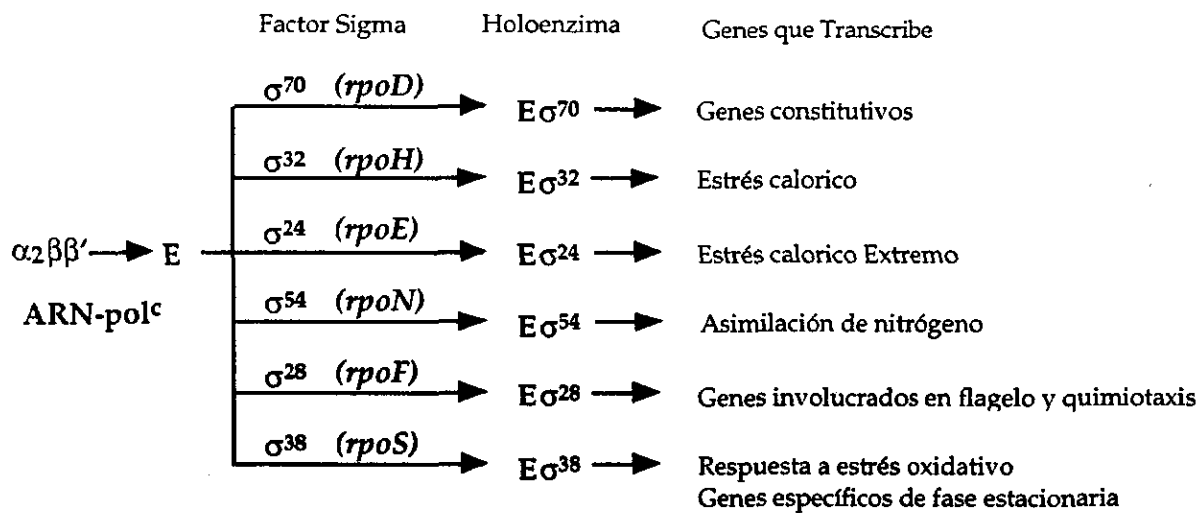


Figura 6. Diferentes holoenzimas presentes en *E.coli*. Los diferentes factores σ dan especificidad a la holoenzima para transcribir grupos de genes involucrados en funciones particulares.

(Tomado de Ishihama *et al.* 1993)

1.3.1 La familia del factor σ^{70}

Los promotores reconocidos por la holoenzima asociada a algún factor de la familia de σ^{70} , se componen de secuencias conservadas a -10 y -35 pares de bases por arriba del inicio de la transcripción (+1) (35).

Las proteínas pertenecientes a esta familia presentan una gran similitud entre ellas. Cada una puede ser subdividida funcionalmente en cuatro regiones distintas. La región 1 se encuentra localizada en la parte amino terminal de la proteína y se presenta casi exclusivamente en sigmas primarios (65). Esta región está constituida aproximadamente por 120 aminoácidos que presentan alrededor de un 30% de identidad entre miembros de la misma familia. A su vez, esta región se ha subdividido en regiones 1a y 1b (38). La región 1a parece estar relacionada con la capacidad de iniciar la transcripción, mientras que proteínas a las cuales se les a removido fragmentos correspondientes a la región 1b son incapaces de isomerizar el complejo cerrado a un complejo abierto (104). Entre las regiones 1 y 2 se encuentra un fragmento de aproximadamente 245 aminoácidos que posiblemente esté relacionado con la unión a la RNA-pol^c (38)

La región 2 se encuentra dividida en cuatro subregiones de la 2.1 a la 2.4. Esta se compone de aproximadamente 70 aminoácidos, que suelen ser los más conservados en las secuencias de los factores de esta familia (38). Diferentes evidencias sugieren que la región está relacionada en el reconocimiento de la RNA-pol^c (61). Se ha visto que fragmentos de esta región se unen a la RNA-pol^c. Por otro lado, mutaciones en la cara expuesta de la hélice alfa de la subregión 2.4, sugieren que ésta puede estar involucrada en el reconocimiento de la región conservada -10 del promotor (35). La región 2.3 parece estar involucrada en el pegado a la cadena sencilla de DNA durante la formación del complejo abierto de la transcripción (38).

De esta manera, la región 2 juega un papel central en la unión del factor σ a la RNA-pol, en el reconocimiento de la secuencia promotora -10 y en el mecanismo de apertura del DNA. La propuesta es congruente con la identificación de mutantes en

la misma región del gen homólogo de σ^{70} en *Bacillus subtilis* (σ^A) que son incapaces de abrir el DNA (38).

La región 3 se encuentra dividida en las subregiones 3.1 y 3.2. (38). Presuntamente la región 3.2 está involucrada en el reconocimiento de la RNA-pol^c, así como en el reconocimiento de la región promotora, aunque actualmente no hay evidencias contundentes al respecto. Mientras que la región 3.1 posiblemente presentan una estructura de helice-vuelta-helice que puede estar relacionada con el reconocimiento del promotor. La poca conservación y la carencia de su existencia en la mayoría de los factores σ de la familia, hacen suponer que su función ha sido sustituida por otras regiones en los factores sigma (65).

La región 4 se encuentra dividida en subregiones 4.1 y 4.2. En varios factores sigma el extremo carboxilo terminal de la subregión 4.2 presenta una estructura hélice-vuelta-hélice (HTH) que se encuentra en una subregión conservada de aproximadamente 30 aminoácidos (Figura 7). Por diferentes tipos de evidencias se sabe que la hélice-vuelta-hélice de la subregión 4.2 está involucrada en el reconocimiento específico de la región -35 del promotor (25,38,89). La subregión 4.1 presenta una pequeña región conservada donde se encuentra una posible hélice alfa anfipática, constituida de aminoácidos aromáticos que se presenta en la mayoría de los factores σ (Figura 7) (65).

Se ha visto que σ^{70} es incapaz de interaccionar específicamente con el DNA en ausencia de la RNA-pol^c, aunque péptidos que contienen las regiones 2 y 4 se unen debilmente. Las evidencias sugieren que muy probablemente la RNA-pol^c contribuye directamente en estabilizar el pegado de σ^{70} con el DNA e indirectamente en dar especificidad para reconocer a la región promotora.

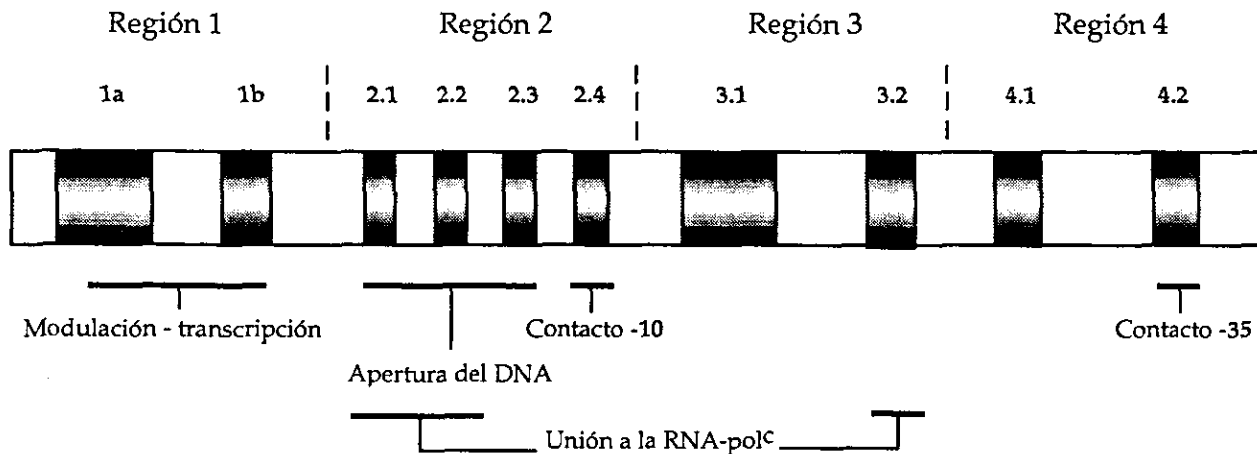


Figura 7. Regiones conservadas en la familia de los factores σ^{70} (en oscuro) y sus posibles funciones.
(Tomado de Record *et al.* 1996)

1.3.2 La familia del factor σ^{54}

El gen que codifica para σ^{54} se ha denominado *ntrA*, *glnF* y recientemente *rpoN* (96). Normalmente este gen está presente en una sola copia, aunque hay algunas excepciones como *B. japonicum*, *Rhizobium etli*, *Azorhizobium caulinodans* y *Rhodobacter sphaeroides* que tienen dos copias (48,56). Hasta ahora hay 36 secuencias de genes *rpoN* reportadas en 33 especies de bacterias distintas, pero tenemos evidencia de que por lo menos otras 24 especies de bacterias posiblemente contienen dicho gen (7).

La holoenzima que contiene el factor σ^{54} ($E-\sigma^{54}$), reconoce promotores con dos motivos conservados en las posiciones -24 (GG) y -12 (GC), que se encuentran por arriba del inicio de la transcripción (+1). Ambos motivos están separados por 10 pares de bases menos conservadas (96). En general, los factores sigma reportados hasta ahora que pertenecen a esta familia son secundarios o alternativos, ya que no se encuentran involucrados en funciones básicas de la célula, excepto en *Myxococcus*

xanthus, donde σ^{54} es esencial para la sobrevivencia de la célula (52). Las proteínas de la familia σ^{54} tienen una gran similitud entre ellas y pueden ser divididas en tres regiones estructurales (Figura 8) (68).

El extremo amino terminal de esta proteína, designado como región I está constituido por aproximadamente 50 aminoácidos. La predicción de estructura secundaria de esta región indica su similitud a una hélice alfa rica en leucinas y glutamínas. La remoción de fragmentos pequeños de aminoácidos de esta región, da como resultado que RpoN sea incapaz de unirse a la secuencia promotora, además de no isomerizar el complejo cerrado a uno abierto. Se ha sugerido que heptámeros repetidos constituidos por residuos hidrofóbicos se encuentran flanqueando la región II y que estos motivos pueden estar interaccionando cooperativamente para formar una estructura de cierre de leucinas entre el extremo amino terminal de la proteína y la región III; proteínas con mutaciones puntuales sobre estos residuos hidrofóbicos pierden la capacidad de unirse a la RNA-pol^c (17,35,40,88,97).

La región II de σ^{54} se encuentra poco conservada, su tamaño es variable y puede estar constituida de 60 a 110 aminoácidos. Esta región contiene una considerable cantidad de residuos ácidos. Se ha sugerido que la región amino terminal de RpoN está involucrada en el reconocimiento del motivo -12 del promotor pero no en la unión con el motivo -24 del promotor. Datos recientes indican que la región II está involucrada tanto en la desnaturalización del DNA como en el reconocimiento del motivo -12 (88,106).

La región III que se encuentra en el extremo carboxilo terminal constituida aproximadamente por 400 aminoácidos altamente conservados, teniendo dos motivos bastante notables. Una hélice-vuelta-hélice (HTH) potencial y un motivo llamado caja rpoN, que está constituido por 10 aminoácidos altamente conservados (93). Se ha visto que la región carboxilo terminal, que conforma más de la tercera parte de la proteína, puede unirse por sí sola al DNA (17). La función del motivo hélice-vuelta-hélice es el reconocimiento del motivo -12 por otro lado, la caja RpoN está involucrada directamente en el reconocimiento del motivo -24 (18,68).

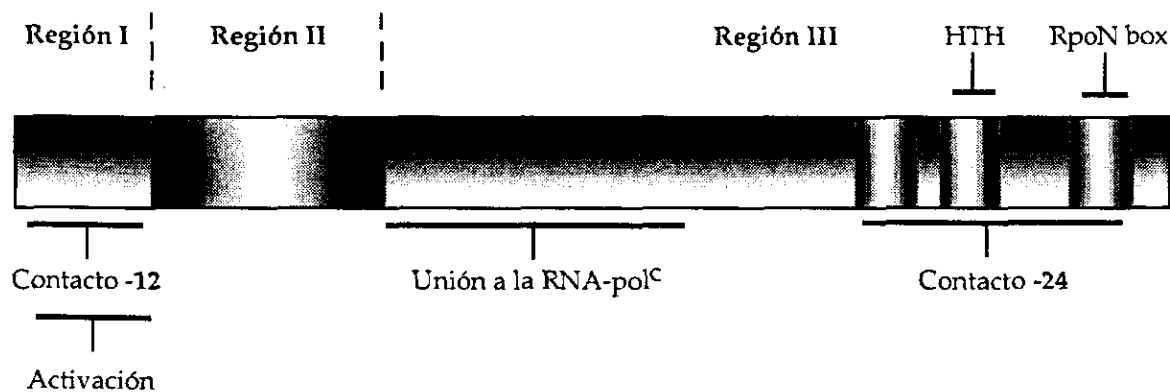


Figura 8. Regiones conservadas en los factores σ^{54} y sus posibles funciones.
(Modificado de Merrick *et al.* 1993)

1.4. Las regiones promotoras σ^{70} y σ^{54}

Las secuencias promotoras reconocidas por la $E-\sigma^{70}$ (Figura 9) están constituidas por dos hexanucleótidos bastante conservados que se localizan en las regiones -35 y -10 por arriba del inicio de la transcripción. Estos motivos presentan las secuencias consenso TTGACA y TATAAT para σ^{70} . La distancia de pares de bases no conservados entre estos motivos es de 17 ± 1 , aunque ejemplos naturales de 15 a 20 pares de bases entre estos dos motivos conservan la función (37). En este tipo de promotores, el inicio de la transcripción comienza en una purina, aunque preferencialmente es usada adenina.

La distancia que existe entre la última base conservada de la región -10 y el inicio de la transcripción, normalmente es de 7 ± 1 pares de bases, existiendo ejemplos de promotores totalmente funcionales con distancias entre 4 y 10 pares de bases (36). Recientemente, se ha reportado la existencia de una tercera región en este

tipo de promotores denominado elemento UP, que se localiza por arriba de la región -35. Este elemento UP presenta secuencias ricas en A y T, que se localizan aproximadamente entre -60 y -40 pares de bases por arriba del inicio de la transcripción (28).

Se sabe que el extremo carboxilo terminal de la subunidad α de la RNA-pol interacciona con la secuencia del elemento UP, incrementando los niveles de transcripción (32,86). Este mismo extremo interactúa con ciertos activadores cuando la región -35 del promotor es débil, y se sugiere que esta misma interacción pasa con el elemento UP estabilizando el complejo cerrado de la transcripción en ambos casos (43).

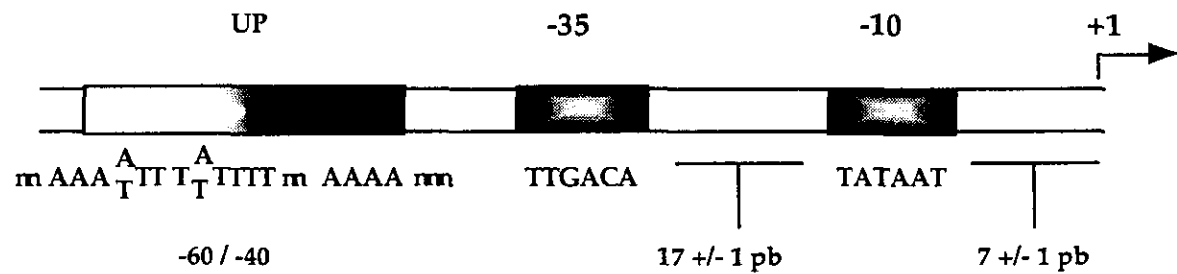


Figura 9. Estructura básica de los promotores reconocidos por sigma 70.

En eubacterias se ha descrito la coexistencia de múltiples factores sigma con la misma capacidad de unirse al único tipo de RNA-pol^c presente. Esto da como resultado un número tan grande de holoenzimas distintas como factores sigma existan en la célula. En la tabla 1 se muestran algunos de los factores sigma más estudiados de la familia así como la secuencia promotora que reconocen (38,43).

Sigma	Organismo	Región -35	Región "-10"
σ^{70}	<i>E. coli</i>	TTGACA	TATAAT
σ^{32}	<i>E. coli</i>	TCTC-CCCTTGAA	CCCAT-TA
σ^A	<i>B. subtilis</i>	TTGACA	TATAAT
σ^B	<i>B. subtilis</i>	AGGTTTAA	GGGTAT
σ^D	<i>B. subtilis</i>	CTAAA	CCGATAT
σ^E	<i>B. subtilis</i>	ATATT	ATACA
σ^G	<i>B. subtilis</i>	TGAATA	CATACTA
σ^K	<i>B. subtilis</i>	AC	CATA--T
σ^H	<i>B. subtilis</i>	CAGGA	GAATT-T
σ^{gp28}	phage SP01	AGGAGA	TTT-TTT ^a
σ^{gp55}	phage T4		TATAAATA

Tabla 1. Secuencia consenso de los promotores reconocidos por la RNA-polimerasa que contiene diferentes factores sigma. (Tomado de Gross *et al.* 1992)

Las secuencias promotoras reconocidas por la holoenzima E- σ^{54} (Figura 10) presentan dos regiones conservadas en -24 y -12 pares de bases por arriba del inicio de la transcripción. Por cuestiones obvias, se les denomina regiones -24 (GG) y -12 (GC), ya que es en esas posiciones donde se encuentran los dinucleótidos más conservados y característicos de estos promotores. La secuencia consenso extendido para este tipo de promotores es mrNrYTGGCACG en la región -24 y TTGCAWNNw en la región -12 (Apéndice IV) (7,96). La remoción de uno o más nucleótidos no conservados entre las regiones -24/-12, han demostrado que la única distancia que puede existir entre estos elementos es de 10 pares de bases para que el promotor sea funcional (11).

Se ha sugerido que el posicionamiento de las dos regiones del promotor -24/-12, en la misma cara de la hélice del DNA, puede ser una condición necesaria para

el reconocimiento y pegado específico de σ^{54} a la región promotora (69).

La distancia que existe entre la C conservada de la región -12 y el inicio de transcripción es normalmente entre 11 y 13 pares de base, aunque hay ejemplos funcionales entre 8 y 21 pares de bases (7). En este tipo de promotores, la holoenzima E- σ^{54} es incapaz de iniciar la transcripción por sí sola, y necesita de la intervención de las proteínas activadoras que pertenecen a la familia de las "enhancer binding proteins" EBP (60,74). Estas proteínas tienen la capacidad de unirse con el DNA a más de 100 pares de bases por arriba del inicio de la transcripción. Esto permite el contacto a distancia con la E- σ^{54} , para isomerizar el complejo cerrado de la transcripción a un complejo abierto (59). Actualmente, se ha sugerido la existencia de los elementos UP en este tipo de promotores, con una función similar a la de los elementos correspondientes en los promotores σ^{70} (Figura 10) (10).

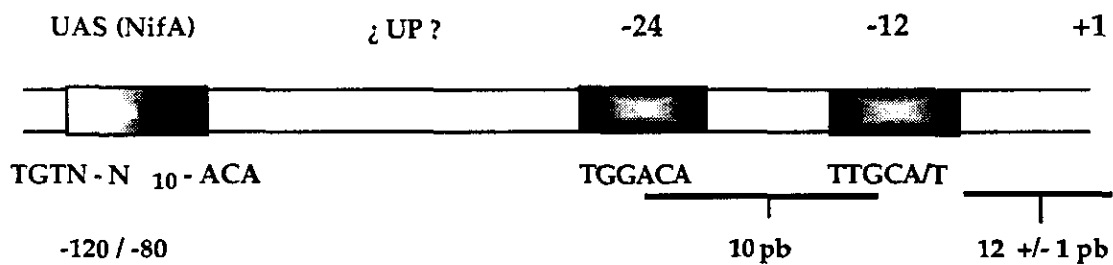


Figura 10. Estructura básica de los promotores reconocidos por sigma 54.

1.4.1. Promotores múltiples.

La expresión tanto de genes como de operones puede estar bajo el control transcripcional de uno o más promotores, los cuales a su vez pueden estar regulados de diferentes maneras. Este tipo de promotores, además son reconocidos por uno o más tipos distintos de holoenzimas. Por sí sola la existencia de las regiones promotoras tiene un papel importante en la regulación de la transcripción. Por otro lado la presencia de promotores múltiples da lugar a sistemas complejos y a niveles más finos en la regulación transcripcional. De los diferentes tipos de promotores múltiples que existen en la naturaleza (Figura 11), podemos decir que cada uno de ellos ha sido seleccionado evolutivamente para una función específica (24).

Los promotores divergentes normalmente tienen secuencias que facilitan la regulación de ambos genes, por ejemplo reprimiendo a uno y activando otro. En *Helicobacter pylori*, los genes *cagA* y *cagB* se transcriben por promotores divergentes y ambos se encuentran involucrados en patogenicidad; además *cagB* tiene dos promotores sobrelapados y ambos funcionales. Estos promotores divergentes son reconocidos por la misma forma de holoenzima E- σ^{70} , aunque la máxima expresión de *cagA* solo puede llevarse a cabo cuando es utilizado el elemento UP que se encuentra sobre *cagB*, disminuyendo su expresión (90).

Otro tipo de promotores son los que se encuentran en tándem. En este caso, ambos promotores son reconocidos por el mismo tipo de holoenzima para aumentar los niveles de expresión, aunque hay ejemplos como el del gene *fliA* en *E.coli*, el cual es controlado por E- σ^{28} y E- σ^{70} ; la holoenzima E- σ^{70} inicia la transcripción y posteriormente la holoenzima E- σ^{28} aumenta los niveles de expresión (62).

El tipo menos común de promotores son los convergentes, un ejemplo de este tipo de promotores lo encontramos en *B. subtilis* donde el promotor Px del gen *antE* se transcribe de manera convergente con el promotor P3 de *sigA*. En este caso, ambos promotores dejan de transcribir cuando la célula entra en la etapa temprana

de esporulación y ambos utilizan dos tipos distintos de holoenzima (99).

En el caso de los promotores sobrelapados estos pueden funcionar como moduladores de la transcripción dependiendo de las condiciones de crecimiento en las que se encuentra la célula. Un ejemplo de este tipo de promotores es el operón *groESL* de *Neisseria gonorrhoeae* donde se encuentran secuencias promotoras que son reconocidas por las holoenzimas E- σ^{32} y E- σ^{70} . Se ha sugerido que durante la exposición de *N.gonorrhoeae* a un estrés calórico, el operón es completamente transcrito a partir del promotor reconocido por la holoenzima E- σ^{32} por otro lado, durante condiciones normales de crecimiento, se transcribe constitutivamente a partir del promotor reconocido por la holoenzima E- σ^{70} (92). Se considera que la manera en que este tipo de promotores se organiza puede estar funcionando como un elemento adicional en la regulación. Si un tipo de holoenzima reconoce a uno de los promotores sobrelapados, esta puede impedir que el segundo promotor sea utilizado. Así tenemos que el primer promotor puede estar regulando de manera negativa la utilización del segundo promotor en otras palabras, se genera una competencia basada en la diferente afinidad de cada holoenzima por la regiones promotoras. Un ejemplo de esto es el operón *fixRnifA* de *B.japonicum* estudiado en este trabajo (6).

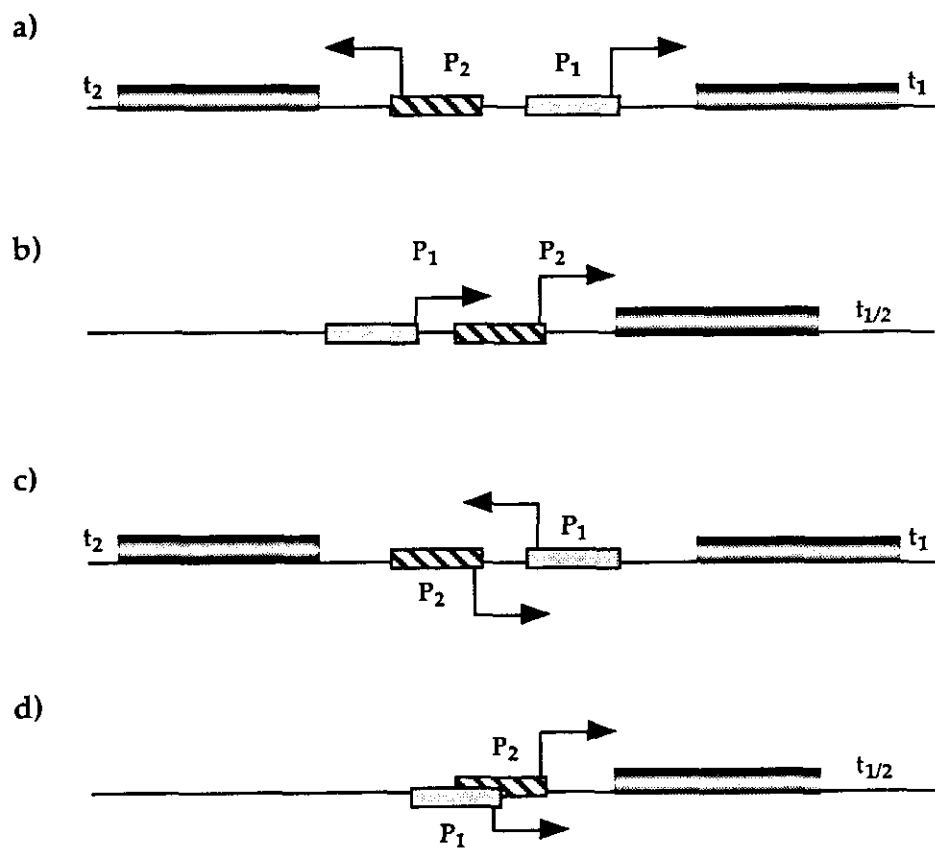


Figura 11. Esquema de promotores múltiple más comunes: a) sobrelapados, b) divergentes, c) tándem y d) convergentes. (Tomado de Doi, R.H. 1991)

1.5. Regulación de la transcripción.

Previamente se han descrito de manera general los elementos básicos de la transcripción, los cuales están involucrados en la funcionalidad de los mecanismos de regulación genética reportados en bacterias.

La regulación global de la expresión genética, responde a estímulos externos entre los que se encuentran; los mecanismos de respuesta a la disposición de fuentes de carbono, nitrógeno, oxígeno, a la tensión ejercida por calor o al daño sobre el DNA. Por otro lado, el mecanismo de respuesta puede ser por estímulos específicos, funcionando para ciertas vías metabólicas las cuales se inducen por las mismas moléculas utilizadas como sustrato. Por último el mecanismo de respuesta puede ser temporal, dentro del cual se incluyen la división y diferenciación celular.

La expresión genética comúnmente se encuentra controlada al nivel de la transcripción, la cual puede dividirse en dos mecanismos de regulación principales. Uno es el control directo de la RNA polimerasa en el reconocimiento del promotor. Este mecanismo depende de las características intrínsecas de la secuencia promotora, básicamente de cuánta información estructural contenga. En otras palabras mientras la secuencia promotora esté más cercana o alejada del consenso para un tipo específico de holoenzima, dependerá la afinidad de la polimerasa por el promotor.

El otro mecanismo involucra proteínas reguladoras las cuales pueden ser represoras (regulación negativa) o activadoras (regulación positiva). Estas proteínas tienen características muy particulares dependiendo de su función. Un ejemplo de regulación negativa lo encontramos en *E.coli*, cuando crece en ausencia de lactosa, el represor LacI se une a la región promotora inhibiendo la transcripción del operón *lacZYA*. Por otro lado cuando hay lactosa en el medio, esta actúa como un efector alostérico sobre LacI impidiendo que se una a la región promotora. En secciones posteriores se describirá una familia de reguladores positivos y su función en la regulación de la transcripción.

1.5.1. La familia de los activadores a distancia (EBP) .

Los promotores dependientes de E- σ^{54} forman complejos cerrados de transcripción estables y para que este complejo pueda isomerizarse a un complejo abierto e iniciar la transcripción, es necesaria la intervención de alguna de las proteínas activadoras de la familia de las EBPs, las cuales tienen la capacidad de unirse al DNA a distancias que varían entre 100 y 200 pares de bases por arriba del +1 (81,91).

Múltiples evidencias físicas y genéticas han demostrado que estas proteínas presentan tres diferentes dominios estructurales (74). El dominio amino terminal de esta familia presenta un tamaño bastante variable; por ejemplo, en NifA de *R. leguminosarum* *bv. trifolii* sólo tiene 12 aminoácidos, mientras que en FhlA de *E. coli* tiene aproximadamente 400 residuos. Por otro lado se puede hacer un subgrupo con algunas proteínas como NtrC, DctD, HydG, AlgB, HoxA y FlbD, las cuales llegan a tener un 38% de similitud en este dominio. Estas proteínas pertenecen a la familia transductora de señales denominada de dos componentes (74).

Dentro de este subgrupo las proteínas histidin cinasa (HPK) funcionan como detectora para transducir las señales hacia el citoplasma, sabiéndose que su forma activa es cuando se encuentra fosforilada. Se ha visto que la contraparte de este sistema de dos componentes es la proteína reguladora de la respuesta (RR), la cual es fosforilada en un residuo de ácido aspártico en el dominio amino terminal por la histidina cinasa (103). Este aminoácido se encuentra conservado en todas las proteínas de este subgrupo, exceptuando FlbD. La fosforilación de este residuo en NtrC estimula la unión cooperativa al DNA y aumenta su afinidad por el ATP (51).

Por otro lado, la actividad de NifA en *K. pneumoniae* se encuentra regulada tanto por oxígeno como por la concentración intracelular de nitrógeno, ambas regulaciones están mediadas por la proteína NifL. Actualmente se sabe que en bajas concentraciones de oxígeno y en condiciones de fijación de nitrógeno el dominio amino terminal protege a NifA de la inactivación por NifL (26). También se sabe

miembros de esta familia de proteínas como NifA de *R. trifolii* y PspF de *E. coli* carecen de este dominio (42,49). La función de dicho dominio es variable y puede ir desde dispensable hasta tener un papel clave en la función y regulación de estas proteínas.

Entre el dominio amino terminal y el dominio central se encuentra un interdominio pequeño, el cual los conecta y da flexibilidad. Este conector es comúnmente encontrado en proteínas modulares y se le denomina como conector-Q, este nombre se debe a que este conector de aproximadamente 20 aminoácidos es rico en glutaminas y otros aminoácidos hidrofóbicos (107).

El dominio central está constituido de aproximadamente 240 aminoácidos altamente conservados en todos los miembros de la familia, y ha sido dividido en 7 regiones numeradas de C1 a C7 (74). La región C1 tiene un motivo de unión a nucleótidos parecido al de las adenilato ciclasas y al de otras proteínas que unen ATP. La hidrólisis de ATP es un paso importante en la formación del complejo abierto de la transcripción, esto se ha establecido con evidencias bioquímicas y genéticas en estudios realizados con la proteína NtrC (5). La región C2 está constituida por residuos hidrofóbicos y contiene la única cisteína conservada en la familia. La región C3 se caracteriza por estar formada por dos hélices alfa y es en esta región donde se encuentran los motivos conservados GAFTGA y ESELFGHEK. El motivo GAFTGA pudiera estar involucrado en el reconocimiento entre las EBPs y la holoenzima E- σ^{54} . Sin embargo, se sugieren que el motivo ESELFGHEK entra en contacto directo con la polimerasa E- σ^{54} (34). Se ha propuesto que esta región funciona como transductora de señales al llevarse a cabo la hidrólisis de ATP (60). La región C4 es rica en glicinas y se encuentra cargada negativamente; en cambio la región C6 está cargada positivamente y se encuentra conformada por residuos aromáticos y prolinas. Finalmente, la región C7 tiene ocho residuos bastante conservados, aunque no se les ha encontrado una función específica (74).

Entre el dominio central y el carboxilo existe una región denominada interdominio de unión, de aproximadamente 32 a 44 residuos. Este interdominio presenta la secuencia Cys-X₁₁-Cys-X₁₉-Cys-X₄-Cys entre las proteínas NifA. Se ha propuesto que las cisteínas están involucradas en la unión de un ion metálico el cual funciona como cofactor; en presencia de oxígeno, este ion se oxida dando como resultado una proteína NifA inactiva (30). Las proteínas correspondientes de *K.pneumoniae* y *A.vinelandii* carecen de estas dos cisteínas y característicamente no son sensibles a oxígeno por sí solas (29).

El dominio carboxilo terminal es el más corto en la familia de las EBPs, su tamaño puede variar entre los 65 y 130 aminoácidos. Este dominio cuenta con un motivo "HTH" el cual reconoce y se une al DNA. Este dominio también es variable ya que cada una de las proteínas de esta familia reconoce secuencias de DNA específicas. Esta variabilidad tan sutil en el motivo HTH puede estar funcionando como un elemento adicional en el mecanismo de regulación transcripcional (Figura 12) (72,74).

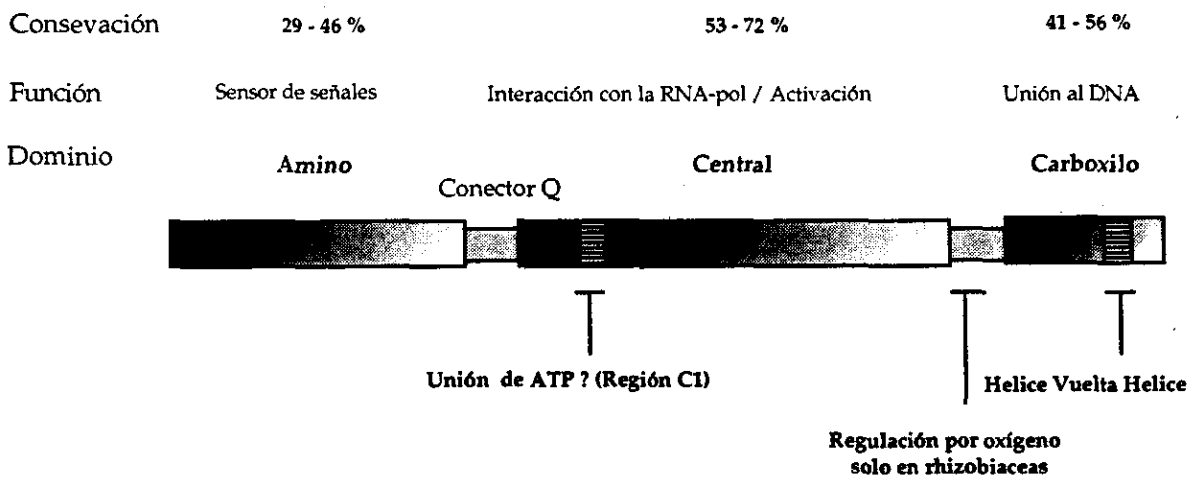


Figura 12. Esquema de los dominios y elementos estructuralmente funcionales de la EBPs.

1.5.2. El mecanismo de transcripción en promotores dependientes de σ^{54} .

El mecanismo de transcripción de todos los promotores dependientes de σ^{54} es bastante similar, siendo el uso de las EBPs para activar la transcripción una de las características (101). Las EBPs se unen a sitios localizados alrededor de 100 pares de bases por arriba de la región promotora -24/-12 (77). A los sitios potenciadores de la transcripción que reconocen estas EBPs se les denomina UAS "upstream activating sequences", y contienen secuencias específicas dependiendo del activador que las reconozca (58). Uno de los ejemplos más documentados son los sitio de pegado para la proteína NifA, los cuales presenta la secuencia TGT N₁₀ ACA y se pueden presentar en una o más copias (19,71).

Para que se lleve a cabo la transcripción es necesario que el activador transcripcional, en este caso NifA, este unido a la secuencia de las UAS que se deben encontrar en la orientación correcta con respecto al promotor, por otro lado la holoenzima debe de estar formando el complejo cerrado de la transcripción (13). La interacción entre la proteína activadora y la holoenzima E- σ^{54} es consecuencia de la formación de un asa entre el activador y el promotor (83,91).

Existen ejemplos de la intervención de una proteína accesoria llamada IHF "integration host factor", que tiene la capacidad de doblar el DNA facilitando la formación del asa entre la EBP y la holoenzima E- σ^{54} (39). Por último, se da la isomerización del complejo cerrado a un complejo abierto de la transcripción formado entre la E- σ^{54} y el promotor (54,78). Esto involucra la hidrólisis de una molécula de ATP por la EBP, lo que se da al momento en que la proteína multimeriza contactando y transduciendo la señal a la holoenzima E- σ^{54} , para que ésta comience la síntesis de RNA (Figura 13) (81,100,108).

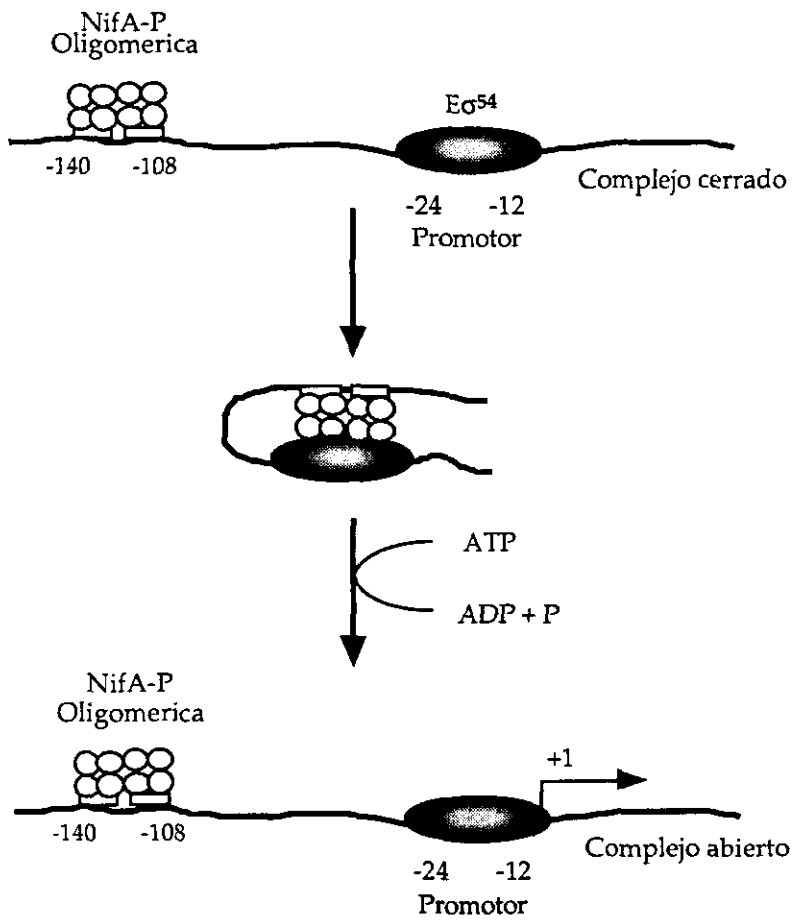


Figura 13. Modelo de activación de la transcripción por NifA a distancia.
 (Modificado de Wyman *et al.* 1997)

1.6. Cascada de regulación de los genes *fix* y *nif* en *B.japonicum*

B.japonicum utiliza una cascada de regulación para activar la expresión de los genes involucrados en la fijación de nitrógeno. La cascada constituida por las proteínas FixLJ-FixK₂, detectan la baja concentración de oxígeno por medio del sistema de dos componentes FixLJ; este se expresa tanto en condiciones aeróbicas como anaeróbicas. FixLJ activa la transcripción del sistema NifA y sigma 54 el cual regula la expresión de los genes *fix* y *nif* entre otros (29).

En *B.japonicum* mutaciones en el operón *fixLJ* muestran una pérdida del 90% en la actividad de fijación de nitrógeno durante la simbiosis en comparación con la cepa silvestre. La cepa mutante también es incapaz de crecer con nitrato como aceptor terminal de electrones en anaerobiosis, lo que sugiere un papel regulador de FixLJ en la respiración microaeróbica (2, 4).

La proteína FixL es una hemoproteína transmembranal que se regula por oxígeno, y cuenta con las actividades enzimáticas de cinasa y fosfatasa. Funcionalmente puede ser dividida en tres dominios (23): a) El dominio amino terminal que contiene cuatro elementos transmembranales, los cuales son esenciales para la actividad de esta proteína tanto *in vivo* como *in vitro* (63). b) El dominio central que contiene un grupo hemo por monómero tiene la capacidad de unir oxígeno. Dos de las tres histidinas conservadas entre las proteínas detectoras se encuentran en este dominio; actualmente se ha sugerido que estos aminoácidos juegan un papel importante en el pegado del grupo hemo (1,70). c) El dominio carboxilo terminal que corresponde a la región más conservada entre las proteínas detectoras en estos sistemas. Se ha observado que este dominio tiene la capacidad de autofosforilación *in vitro*. La autofosforilación de FixL *in vivo* se estimula específicamente por bajas concentraciones de oxígeno en el medio (64,70).

La proteína FixJ contiene dos dominios estructuralmente funcionales. El dominio amino terminal tiene cuatro aminoácidos altamente conservados los cuales son característicos de las proteínas reguladoras encargadas de la respuesta (1).

Debido a la alta similitud que existe entre las proteínas de esta familia se ha propuesto que el residuo Aspártico 54 de FixJ es fosforilado por FixL, aunque no se ha demostrado directamente (1,29). El dominio carboxilo terminal tiene una alta similitud con un buen número de activadores transcripcionales, además de su homología con una parte del dominio carboxilo terminal de los factores σ^{70} , los cuales interactúan con la región -35 de los promotores. En la región más conservada se encuentra un motivo de HTH, el cual puede estar involucrado en el reconocimiento específico de las secuencias promotoras (50).

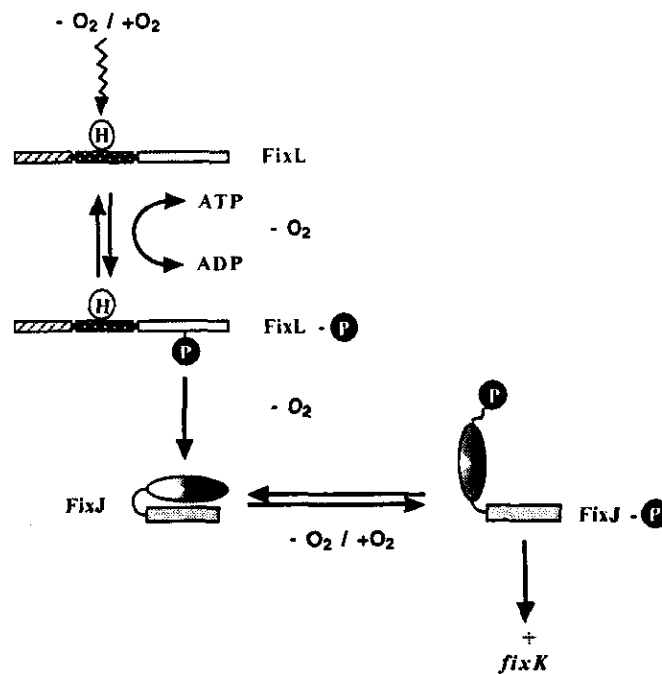


Figura 14. Mecanismo propuesto para la fosforilación de FixL-FixJ.

(Tomado de Fischer, H.M. 1994)

En el modelo del mecanismo de transducción de señales por medio del sistema de dos componentes FixLJ, se induce cuando FixL detecta la concentración de oxígeno en el medio a través de su grupo hemo, el cual está unido al dominio central y se encuentra orientado hacia el citoplasma de la célula. Este dominio modula tanto la actividad de cinasa como de fosfatasa del dominio carboxilo terminal. A bajas concentraciones de oxígeno favorece la actividad de cinasa que permite el incremento de FixL-P. En este estado FixL-P transfiere su grupo fosfato a la proteína FixJ dando como resultado la un incremento de FixJ-P, la cual se encuentra en su forma activa para inducir la transcripción de los genes blanco. Por otro lado, cuando la concentración de oxígeno aumenta, los niveles de FixJ-P decrecen por medio del aumento de la actividad de fosfatasa de FixL (Figura 14) (64).

Bajo condiciones de microaerobiosis, FixJ-P activa la transcripción del gen *fixK₂*. La proteína FixK₂ es un regulador positivo de los genes involucrados en la respiración en condiciones microaeróbicas (por ejemplo *fixNOQP*, *fixGHIS*), incluyendo a una de las dos copias que codifican para el factor σ^{54} (*rpoN₁*); la otra copia de σ^{54} (*rpoN₂*) se autorregula negativamente (56). Mutaciones sobre *fixK₂* reducen la capacidad de respirar en microaerobiosis, presentando el mismo fenotipo que las mutaciones en *fixLJ* (76). La regulación del factor sigma σ^{54} por FixK₂, es la conexión para que se active el sistema de regulación involucrado en la fijación simbiótica de nitrógeno y otros procesos en *B.japonicum*. Bajo condiciones donde la concentración de oxígeno es baja, la proteína NifA activa la transcripción de un grupo específico de genes que presentan promotores -24/-12 (29). En *B.japonicum*, *nifA* es el gen distal del operón *fixRnifA*. Este operón no solamente se expresa en condiciones microaerobicas, si no además en condiciones aeróbicas donde está bien documentado que la proteína NifA se inactiva rápidamente (Figura 15) (73,94,95).

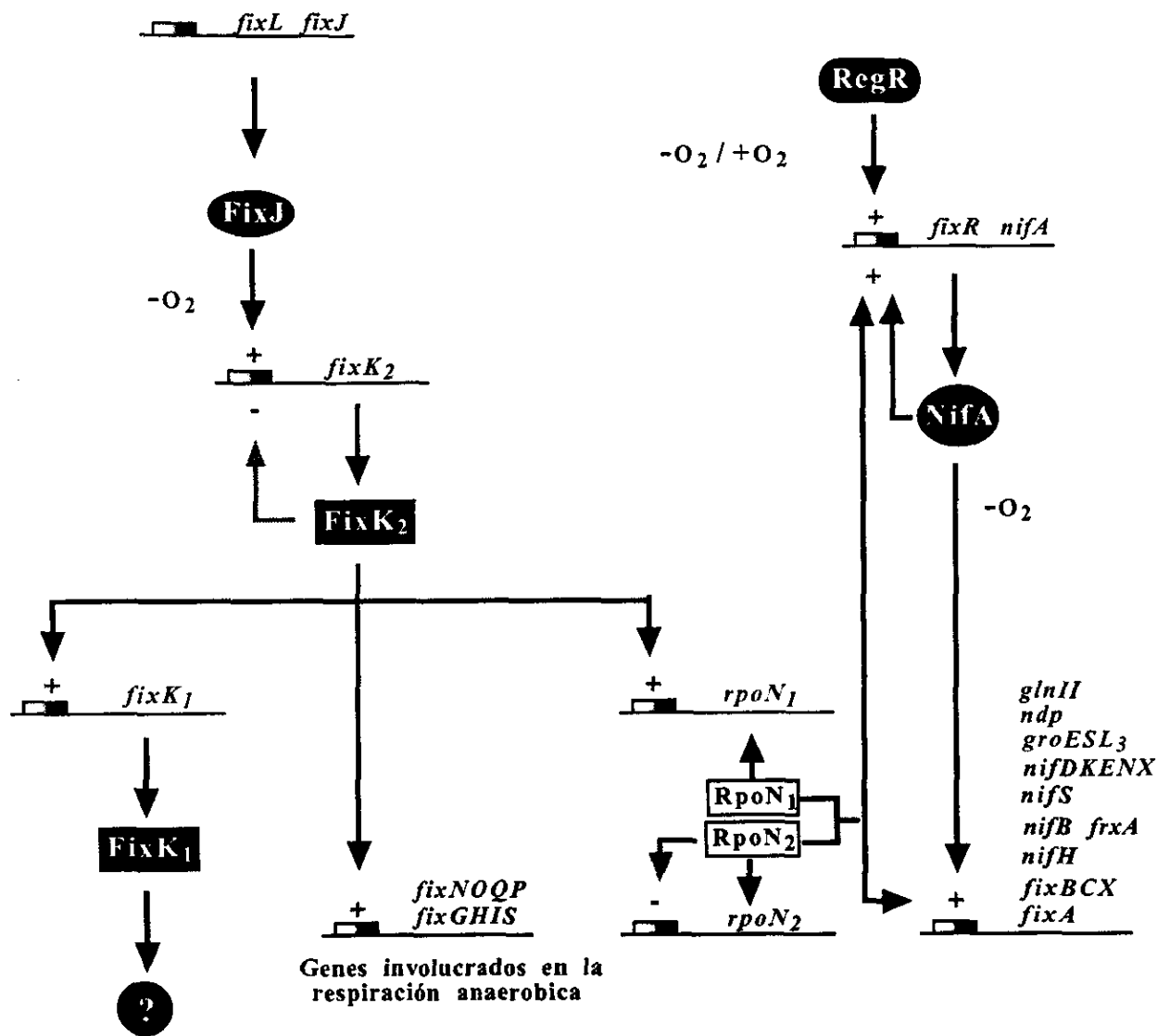


Figura 15. Esquema de la cascada de regulación transcripcional en *Bradyrhizobium japonicum*. FixLJ-FixK₂ controla principalmente la respiración aerobica y activa la expresión del sistema NifA y sigma 54 que regulan los genes involucrados en la fijación de nitrógeno principalmente.

(Tomado de Fischer, H.M. 1994)

2. ANTECEDENTES Y OBJETIVO

La fijación biológica del nitrógeno es el proceso enzimático de la reducción del nitrógeno molecular (N_2) a amonio. Para que se lleve a cabo esta reacción enzimática es necesaria la expresión de por lo menos 20 genes (*fix* y *nif*), los cuales son regulados positivamente por el activador transcripcional NifA en condiciones de microaerobiosis. La enzima encargada de la reducción del nitrógeno diatómico se conoce como nitrogenasa; se ha calculado que esta enzima necesita 16 moles de ATP por cada mol de amonio producido. La nitrogenasa es funcional sólo en condiciones de microaerobiosis. El alto consumo energético de este proceso hace necesario un control muy fino, el cual se consigue mediante una regulación genética muy compleja para evitar un gasto excesivo de energía (69,85).

B. japonicum es una proteobacteria de la subdivisión alfa que habita en el suelo y tiene la capacidad de establecer una relación simbiótica con las raíces de la soya (*Glycine max*). En esta simbiosis se da la formación de nódulos en las raíces donde se lleva a cabo el proceso de la fijación de nitrógeno. Este proceso ocurre en un estadio diferenciado de las bacterias denominado bacteroide, el cual se encuentra dentro de los nódulos. *B. japonicum* tiene la capacidad de fijar nitrógeno tanto en vida libre como en simbiosis (3,105).

En los organismos que fijan nitrógeno este proceso se encuentra regulado por la proteína NifA, que pertenece a la familia de las EBPs. Esta proteína activa la transcripción de los genes *nif* y *fix* concertadamente con la RNA-pol- σ^{54} , al interaccionar con secuencias UAS, las cuales presentan los motivos TGT - N₁₀ - ACA que son reconocidas por NifA (14,74).

El estudio del mecanismo de regulación del operón *fixRnifA*, comenzó en el grupo del Prof. Hauke Hennecken; a continuación describo la información con la que contábamos al iniciar este proyecto de investigación.

En *B. japonicum*, la proteína NifA está codificada dentro del operón *fixRnifA*. Se sabe que este operón es autorregulado positivamente por NifA tanto en simbiosis

como en microaerobiosis. La secuencia nucleotídica de la región promotora de este operón, presenta una alta similitud al consenso de los promotores -24/-12 que son reconocidos por la holoenzima E- σ^{54} (95). En la región que se encuentra por arriba del promotor -24/-12, se han descrito dos posibles motivos UAS para NifA, los cuales se localizan a 100 y 121 pares de bases por arriba del inicio de transcripción. Estos motivos son imperfectos y presentan las secuencias TGT N₁₀ CCA y TGT N₁₀ ACC, respectivamente, en vez de TGT N₁₀ ACA (94).

Está bien documentado que la proteína NifA de *B.japonicum* es sensible por sí sola a oxígeno, y bajo condiciones aeróbicas queda inactiva (73). En estas condiciones no existe expresión del gen *rpoN₁*, en cambio la expresión de *rpoN₂* se ve regulada negativamente. Aun así, se ha visto que el operón *fixRnifA* continúa activo en condiciones aeróbicas (56,94). Por otro lado, al evaluar la expresión de este operón en un fondo genético *nifA*⁻ o en una doble mutante de los genes *rpoN_{1/2}*, se ha visto una pérdida total en el proceso de fijación de nitrógeno.

De esta manera, queda descartada la posibilidad de que el sistema dependiente de NifA estuviera involucrado en la expresión aeróbica del operón *fixRnifA*. Por medio de la remoción de fragmentos de DNA que abarcan la región promotora, se encontró que la región entre -80 y -50 pares de bases, por arriba del inicio de transcripción es necesaria para la expresión aeróbica del operón *fixRnifA*. Al analizar los mismos fragmentos removidos del promotor, en experimentos de retardamiento electroforético, se observó que sólo aquellos fragmentos que contienen secuencias de DNA por arriba de la posición -50 tienen la capacidad de ser retardados, al incubarse con extractos de crecimientos aeróbicos de *B.japonicum*. Utilizando el mismo análisis, se ha visto que el cambio puntual de A por C en la posición -68 de un fragmento sintético de DNA, que abarca desde la posición -80 a -50 por arriba del inicio de la transcripción, es suficiente para evitar el retardo de este fragmento. Este mismo cambio en la posición -68 impide la expresión del operón en condiciones aeróbicas. Se ha propuesto que la región -68 es el sitio de pegado de una proteína que regula aeróbicamente al operón *fixRnifA*. Con los datos anteriores, se

llego a proponer la existencia de un doble sistema de regulación que controla la expresión del operón *fixRnifA* de manera diferencial (94).

En base a estos antecedentes se propuso como objetivo general de este trabajo el estudio de los mecanismos de regulación transcripcional que controlan la expresión del operón *fixRnifA* en *B. japonicum*. El objetivo particular de este proyecto fue, comprobar la existencia del doble sistema de regulación transcripcional que controla la expresión de *fixRnifA*, mediante el mapeo del inicio de la transcripción, en diferentes fondos genéticos y simbiosis.

APENDICE I

Barrios H, Fischer H M, Hennecke H and Morett E. (1995). Overlapping promoters for two different RNA polymerase holoenzymes control *Bradyrhizobium japonicum nifA* expression. *J. Bacteriol.* **177**: 1760- 1765.

Overlapping Promoters for Two Different RNA Polymerase Holoenzymes Control *Bradyrhizobium japonicum* *nifA* Expression

HUMBERTO BARRIOS,^{1†} HANS-MARTIN FISCHER,² HAUKE HENNECKE,² AND ENRIQUE MORETT^{1*}

Departamento de Ecología Molecular, Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62271, México,¹ and Mikrobiologisches Institut, Eidgenössische Technische Hochschule, CH-8092 Zürich, Switzerland²

Received 21 November 1994/Accepted 27 January 1995

The *Bradyrhizobium japonicum* NifA protein, the central regulator for nitrogen fixation gene expression, is encoded in the *fixRnifA* operon. This operon is activated during free-living anaerobic growth and in the symbiotic root nodule bacteroid state. In addition, it is expressed in aerobic conditions, albeit at a low level. Here, we report that this pattern of expression is due to the presence of two overlapping promoters: *fixRp*₁, which is of the -24/-12 class recognized by the RNA polymerase σ^{54} , and *fixRp*₂, which shares homology with the -35 and -10 regions found in other putative *B. japonicum* housekeeping promoters. Primer extension analyses showed that *fixRp*₁ directed the synthesis of a transcript, P1, that starts 12 nucleotides downstream of the -12 region. In addition to σ^{54} , P1 was dependent on NifA and low oxygen tension. Transcripts originating from *fixRp*₂ started at two sites: one coincided with P1, while the most abundant, P2, initiated just two nucleotides further downstream of P1. Expression from *fixRp*₂ was dependent on the upstream -68 promoter region, a region known to bind a putative activator protein, but it was independent of σ^{54} and NifA. This promoter was expressed in aerobic and anaerobic conditions but was not expressed in 30-day-old bacteroids. Mutations in the conserved -12 region for the σ^{54} promoter did not show any transcript, because these mutations also disrupted the overlapping -10 region of the *fixRp*₂ promoter. Conversely, mutations at the -24 region only affected the σ^{54} -dependent P1 transcript, having no effect on the expression of P2. In the absence of σ^{54} , anaerobic expression from the *fixRp*₂ promoter was enhanced threefold, suggesting that in the wild-type strain, the two RNA polymerase holoenzymes must compete for binding to the same promoter region.

In the soybean root nodule endosymbiont *Bradyrhizobium japonicum*, as in most of the nitrogen-fixing bacteria, expression of the genes involved in free-living and symbiotic nitrogen fixation (*nif* and *fix*) is controlled by the oxygen status of the cell. This control is exerted by regulating both the activity and the synthesis of the central regulatory protein NifA (13, 24, 29). NifA belongs to the bacterial enhancer-binding protein family of transcriptional regulators that activate promoters recognized by the RNA polymerase holoenzyme containing the alternative sigma factor σ^{54} ($E\sigma^{54}$) (15, 21). The σ^{54} promoters are unique among the prokaryotic promoters in having two closely spaced conserved sequences, centered at -12 and -24 bp from the transcription start, instead of the more common -10 and -35 boxes found in most of the prokaryotic promoters (23, 27). To activate *nif* and *fix* gene expression, NifA binds to DNA sites, named upstream activator sequences, which function as transcriptional enhancers, and causes the isomerization of an $E\sigma^{54}$ closed promoter complex to the transcriptionally active open promoter form, a process that is dependent on ATP hydrolysis (2, 6, 22, 26, 27).

The in vivo activity of the NifA protein of *B. japonicum* is sensitive to oxygen (17). It can only bind to the upstream activator sequences and activates *nif* and *fix* gene expression at

low oxygen concentrations. Upon a shift to aerobic growth, NifA is immediately inactivated and subject to degradation (28). The mechanism of oxygen sensing by this protein is not clear. However, it has been shown that both metal ions and a cluster of four conserved cysteine residues not present in the oxygen-insensitive NifA protein of *Klebsiella pneumoniae* are required for its activity (16, 28). It has been proposed that the cysteine residues coordinate a metal cofactor whose redox state modulates NifA activity (16, 28).

The *B. japonicum* NifA protein is encoded in the *fixRnifA* operon, which, in contrast to the majority of the nitrogen-fixing bacteria, is expressed in aerobic and anaerobic cultures (13, 40). The aerobic expression is about one-fifth of the level observed in anaerobiosis and is completely dependent on the integrity of a region located around 66 bp upstream of the transcription start site (39). This upstream region is the binding site for a protein proposed to be an activator. The enhanced expression observed in anaerobically growing cells is dependent on NifA (39).

B. japonicum has two copies of the *rpoN* gene, *rpoN*₁ and *rpoN*₂, that code for σ^{54} (20). Mutational analyses of these genes have shown that both are functional: single mutations did not have any symbiotic effect, while an *rpoN*₁ *rpoN*₂ double-mutant strain was unable to activate the expression of the nitrogenase genes and, therefore, to fix nitrogen (20). Expression of the *rpoN*₁ gene is regulated by the oxygen conditions of the cell via the two-component pair of proteins FixL and FixJ, while the *rpoN*₂ gene is negatively autoregulated (20). The promoter region of the *fixRnifA* operon has a sequence that closely resembles the consensus -24/-12 promoter and, because it is also autoactivated by NifA, it has been proposed to

* Corresponding author. Present address: Instituto de Biotecnología, Universidad Nacional Autónoma de México, A. Postal 510-3, Cuernavaca, Morelos 62271, México. Phone: 52 (73) 114900. Fax: 52 (73) 172388. Electronic mail address: emorett@pbr322.ceingebi.unam.mx.

† Present address: Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62271, México.

be recognized by σ^{54} (39, 40). However, in an *rpoN₁ rpoN₂* double-mutant strain devoid of σ^{54} , or in a *fixR* promoter mutant in which the conserved TG dinucleotide of the -24 promoter region was replaced by CT, the expression of the *fixRnifA* operon was only marginally reduced (20, 40). These observations suggested that *fixRnifA* expression might be independent of the apparent -24/-12 promoter and therefore must originate from an alternative, yet unidentified promoter (20).

In this report, we show by transcription mapping of *fixRnifA* that the σ^{54} promoter, here denominated *fixRp₁*, is transcriptionally active and that there is a second overlapping promoter, *fixRp₂*, recognized by an RNA polymerase with a sigma factor other than σ^{54} , likely to be the housekeeping factor. We also show that the previously observed marginal reduction in *fixRnifA* expression in an *rpoN₁ rpoN₂* double mutant (20) was due to the enhanced expression from the *fixRp₂* promoter, probably caused by a relief of competition for RNA polymerase binding to the promoter region in the absence of σ^{54} .

MATERIALS AND METHODS

Bacterial strains and plasmids. All of the strains used in this work were constructed by chromosomal integration by homologous recombination of the appropriate *fixR-lacZ* fusion as described by Kullik et al. (20). Strain 8085 contains the wild-type *fixR-lacZ* promoter fusion in the wild-type *B. japonicum* 110*spc4* strain. Strain 8091 harbors a *fixR-lacZ* promoter fusion with an A-to-C [A(-68)-to-C] transversion at position -68 (previously denominated -66), integrated in the chromosome of the 110*spc4* strain. Strains A8085 and A8091 are the same as strains 8085 and 8091, except that the *fixR-lacZ* promoter fusions were integrated into the *nifA* mutant strain A9 (14). Strains N8085 and N8091 are the same as strains 8085 and 8091, except that the *fixR-lacZ* promoter fusions were integrated into the *rpoN₁ rpoN₂* double-mutant strain N50-97 (20). Strain 7292R has a *fixR-lacZ* promoter fusion in which the GC dinucleotide of the -12 promoter region was replaced by AG in the 110*spc4* background. Strain 7291R has a *fixR-lacZ* promoter fusion in which the TG dinucleotide of the -24 promoter region was replaced by CT in the 110*spc4* background.

Growth conditions. *B. japonicum* strains were normally grown aerobically in PSY medium (34) with the following antibiotics: kanamycin, 100 μ g/ml; tetracycline, 50 μ g/ml; streptomycin, 100 μ g/ml; and spectinomycin, 100 μ g/ml. Anaerobic cultures were grown in yeast extract-mannitol (YEM) medium (11) with 10 mM KNO_3 as the electron acceptor.

β -Galactosidase assays. Determination of β -galactosidase activity of the *fixR-lacZ* promoter fusion strains was carried out according to the standard procedure of Miller (25).

RNA isolation. Total RNA was prepared from 100-ml cultures by the hot acidic phenol procedure. Briefly, cells grown to mid-exponential phase (60 h) were collected by centrifugation and washed with saline phosphate solution (150 mM NaCl, 40 mM K_2HPO_4 , 22 mM KH_2PO_4 [pH 7.2]). Lysis was carried out in RNA isolation buffer (20 mM sodium acetate, 0.5% sodium dodecyl sulfate, 1 mM EDTA [pH 5.5]) with phenol, which was previously equilibrated with sodium acetate (20 mM [pH 4.5]) and incubated for 5 min at 65°C. RNA was purified with the RNAid RNA purification kit from Bio 101, Inc. (La Jolla, Calif.) according to the manufacturer's instructions. To isolate total RNA from root nodule bacteroids, soybean *Glycine max* (Promase, Mexico D. F., Mexico) plants were grown in a greenhouse, and 30 days after inoculation, the nodules were harvested and collected in dry ice. Bacteroids were isolated by a Percoll gradient preparation, as reported previously (35), and the RNA was isolated by the same procedure as that used for free-living cultures.

Transcriptional mapping. The transcription initiation sites were determined by primer extension analysis. For each extension reaction, RNA isolated from 20-ml cultures or from bacteroids isolated from nodules collected from five plants was used as a template for reverse transcription. The following synthetic oligonucleotides were used as primers: lac4 (5'-AGTTGGGTAACGCCAGGG-3') and Pbj16S (5'-CGGGCCGATCTTCGGCG-3'). These oligonucleotides were complementary to the 5' *lacZ* part of the *fixR-lacZ* mRNA and to positions 180 to 162 of the 16S rRNA (42), respectively. Five picomoles of oligonucleotides lac4 and Pbj16S was 5' end labelled with 27 μ Ci (ca. 4.7 pmol) of [γ - 32 P]ATP (>6,000 Ci/mmol) as previously described (26). Oligonucleotides (0.25 pmol of lac4 and approximately 0.001 pmol of Pbj16S) were annealed in 0.12 M NaCl-0.25 M Tris-HCl (pH 7.0) with the total RNA by being heated for 3 min at 95°C and then slowly cooling to 42°C. Primer extensions were carried out at 42°C for 1 h with avian myeloblastosis virus (Boehringer Mannheim) or RAV-2 (Amersham) reverse transcriptase enzyme in transcription buffer (50 mM Tris-HCl, 50 mM KCl, 10 mM MgCl_2 , 5 mM dithiothreitol [pH 7.5]), 1 mM deoxyribonucleoside triphosphates, 100 mM dithiothreitol, and 108 U RNasin (Amersham). The extension products were precipitated with 2.5 M ammonium acetate and ethanol

TABLE 1. Expression of different *fixR* promoter derivatives fused to *lacZ* in wild-type, *nifA*, and *rpoN₁ rpoN₂* backgrounds

Strain	Background	Relevant characteristic of <i>fixR</i> promoter	β -Galactosidase activity (Miller units)	
			Aerobic	Anaerobic
8085	Wild type	Wild type	278 \pm 14	1,891 \pm 398
A8085	<i>nifA</i>	Wild type	254 \pm 33	173 \pm 20
N8085	<i>rpoN₁ rpoN₂</i>	Wild type	173 \pm 17	1,150 \pm 110
8091	Wild type	A(-68) to C	5 \pm 1	1,078 \pm 197
A8091	<i>nifA</i>	A(-68) to C	5 \pm 1	2 \pm 1
N8091	<i>rpoN₁ rpoN₂</i>	A(-68) to C	2 \pm 1	12 \pm 4
7291R	Wild type	TG(-25) to CT	135 \pm 25	1,256 \pm 21
7292R	Wild type	GC(-12) to AG	42 \pm 2	120 \pm 49

and analyzed by electrophoresis on 6% polyacrylamide gels. The gels were exposed on Hyperfilm- β -max films (Amersham) or scanned in a Molecular Dynamics PhosphorImager device. The sequence ladder of the *fixR* promoter region used to determine the transcription start sites was obtained from plasmid pRJ211 (39) with oligonucleotide lac4 and Sequenase 2.0 (U.S. Biochemical Corp.) by the standard procedure.

RESULTS AND DISCUSSION

Mapping of the *fixR* transcription start. In previous studies, we had shown that expression of the *B. japonicum fixRnifA* operon is subject to a dual positive control (39, 40). In aerobic cultures, expression from this operon is dependent on an upstream region centered around 66 bp upstream of the transcription start, which is the binding site for a protein proposed to be a positive regulator (39). During anaerobiosis, the expression is increased fivefold and is dependent on NifA (Table 1 [39]). However, the dependence on NifA did not correlate with the almost complete independence from σ^{54} or the integrity of the -24 promoter region for anaerobic expression (Table 1 [20, 39]). Thus, it was postulated that perhaps the putative -24/-12 promoter is not functional and that the observed effect of NifA could be indirect and could occur in a σ^{54} -independent mode (20). To assess whether the putative -24/-12 promoter is functional and to identify the promoter that is independent of σ^{54} , we carried out an analysis of the transcription initiation sites by primer extension with several strains and *fixR* promoter mutants.

Figure 1 shows a transcription mapping experiment of the *fixRnifA* operon in the 8085 (wild type), A8085 (*nifA*), and N8085 (*rpoN₁ rpoN₂*) strains grown anaerobically. In strain 8085 (lane 1), two major primer extension products were detected; the longest one, P1, was the most abundant transcript and was initiated at a guanine residue located exactly 12 bp downstream of the presumed -24/-12 promoter. The second transcript, P2, was also initiated at a guanine residue located two nucleotides downstream from the start point of P1. In the N8085 strain (lane 5), the P1 transcript was considerably diminished, while the P2 transcript was not affected. The A8085 strain (lane 3) showed the same pattern as the N8085 strain; i.e., the P2 transcript was more abundant than P1, but the total amount of *fixR-lacZ* transcripts was reduced (see below).

Our interpretation of these results is that P1 originates from a functional -24/-12 promoter, *fixRp₁*, activated by NifA and that this is the major promoter in anaerobic conditions. The transcripts observed in the *nifA* and in the *rpoN₁ rpoN₂* backgrounds must stem from a promoter recognized by a form of the RNA polymerase with a sigma factor other than σ^{54} (see below). This promoter, *fixRp₂*, can initiate at two sites. The preferred site, P2, is two nucleotides downstream from the

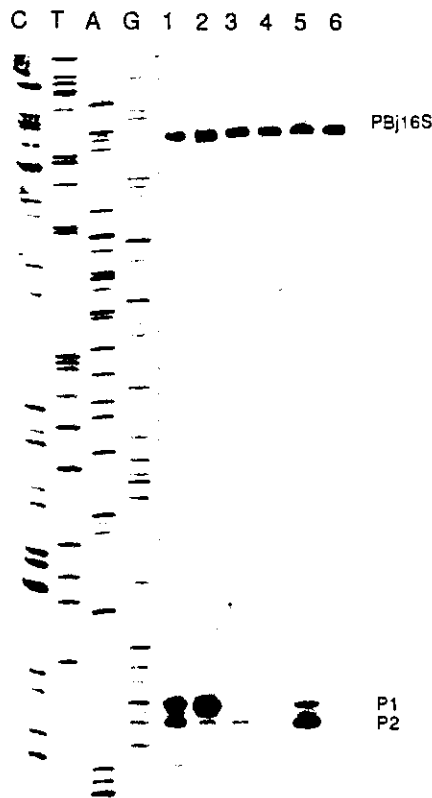


FIG. 1. Transcript mapping of the *fixRnifA* mRNA in wild-type, *nifA*, and *rpoN₁ rpoN₂* strains and effect of the A(-68)-to-C mutation. Total RNA was purified from the following strains grown anaerobically with KNO₃: lane 1, 8085 (wild type); lane 2, 8091 [A(-68)-to-C mutation in wild type]; lane 3, A8085 (*nifA*); lane 4, A8091 [A(-68)-to-C mutation in *nifA* strain]; lane 5, N8085 (*rpoN₁ rpoN₂*); lane 6, N8091 [A(-68)-to-C mutation in *rpoN₁ rpoN₂* strain]. The RNA was annealed to γ -³²P 5'-end-labelled oligonucleotides lac4 and Pbj16S to detect the *fixRnifA* and the 16S rRNA transcripts, respectively, and extended with reverse transcriptase as described in Materials and Methods. The products were separated by electrophoresis in a 6% polyacrylamide gel next to a sequence ladder of plasmid pRJ7211 made with the same oligonucleotide. Transcripts P1, P2, and Pbj16S (control) are indicated.

σ^{54} -dependent P1, although a second transcription site coincides with that used by $E\sigma^{54}$.

It is remarkable that in the *nifA* background, both the *fixR-lacZ* transcripts and the β -galactosidase activity of the fusion were reduced compared with those in the *rpoN₁ rpoN₂* background (Fig. 1, lanes 3 and 5). This is probably due to the formation of a strong $E\sigma^{54}$ -*fixR* promoter complex, which, in the absence of NifA, cannot escape from the promoter and then interferes with the binding of the other holoenzyme to the same DNA region. It has previously been shown that $E\sigma^{54}$ can form stable closed complexes with certain promoters (9, 27, 32, 38). Moreover, Magasanik and collaborators (36) have shown that $E\sigma^{54}$ can inhibit transcription of an artificially created σ^{70} promoter that overlaps the *Escherichia coli glnAp₂* σ^{54} promoter in the absence of the NRI (NtrC) activator protein. In the *rpoN₁ rpoN₂* background, there is no interfering $E\sigma^{54}$, and therefore expression from the *fixRp₂* promoter is enhanced.

In a previous transcription mapping analysis of the *fixRnifA* operon, it was proposed that the +1 position corresponds to a C residue located two nucleotides upstream of P1 (40). The discrepancy with our results is due to the low-resolution nature of the previous mapping. Now, we have denominated the P1 start site as position +1 and renumbered the upstream positions accordingly.

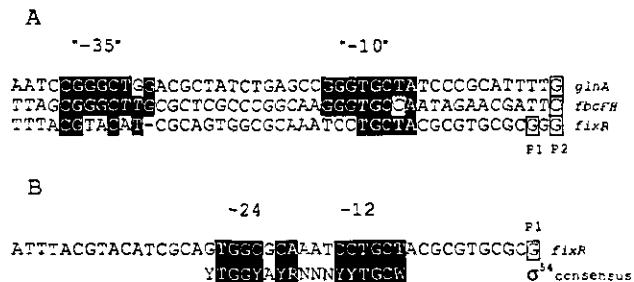


FIG. 2. Comparison of the *glnA*, *fbcFH*, and *fixR* promoter regions. (A) The sequences of the *glnA* (7) and *fbcFH* (40a, 41) promoter regions were aligned with the +1 transcription start site as a reference. The *fixR* promoter region was aligned with the 5' end of P2 as +1. (B) The *fixR* promoter region was aligned with the consensus σ^{54} promoter sequence (27). Identical nucleotides are shown in reverse type. The -12 region of the *fixRp₂* promoter completely overlaps the -10 region of the *fixRp₁* promoter.

Influence of an upstream promoter mutation on *fixRnifA* expression. We determined the effect of a chromosomally integrated mutation at position -68 {A(-68) to C [previously denominated A(-66) to C]} on the expression of the *fixRnifA* operon in the 8091 (wild-type), A8091 (*nifA*), and N8085 (*rpoN₁ rpoN₂*) strains. This mutation caused a strong reduction in aerobic expression but had only a moderate effect on anaerobic expression (Table 1). Figure 1 shows that in strain 8091, the A(-68)-to-C mutation specifically abolished expression of the P2 transcript but had no effect on P1 expression (lane 2). In the A8091 (lane 4) and N8091 (lane 6) strains, no transcripts were observed, because in these strains neither the *fixRp₁* nor the *fixRp₂* promoter can be activated. The use of an oligonucleotide complementary to the 16S rRNA in the same primer extension reactions as the control shows that in all lanes, the total amounts of RNA used were about the same (Fig. 1). The rRNA transcript (Pbj16S) also allowed us to determine the relative concentration of the *fixRnifA* transcripts in all mutant strains and growth conditions tested. The levels of the transcripts observed (data not shown) generally correlated with those of the *fixR-lacZ* expression, determined as β -galactosidase (Table 1).

The results presented above show that $E\sigma^{54}$ only directs the synthesis of P1 and that both the P2 transcript and the fraction of the longest transcript still present in the *nifA* and *rpoN₁ rpoN₂* backgrounds originate from *fixRp₂*, a promoter that is completely dependent on the integrity of the upstream -68 promoter region but independent of NifA and σ^{54} .

Effect of promoter mutations on *fixRnifA* expression. In an attempt to identify the nucleotides that constitute the *fixRp₂* promoter, we compared the promoter regions of two other *B. japonicum* genes whose expression is independent of σ^{54} and whose transcription start sites had been mapped. These genes are *glnA* and *fbcFH*, which code for a glutamine synthetase and for the cytochrome *bc₁* complex, respectively (7, 40a, 41). These genes code for unrelated functions, and at least the *fbcFH* operon is expressed under all growth conditions tested; therefore, it is likely that they are recognized by the RNA polymerase with the housekeeping sigma factor σ^{56} , the sigma factor that copurifies with the core RNA polymerase from *B. japonicum* (34). Figure 2 shows the sequences of these two promoters, with the +1 position used to generate the alignment. Two conserved regions starting 12 and 34 bp upstream from the transcription start were detected. Both regions have seven out of eight identical nucleotides in the two promoters. The sequences of these regions do not resemble any of those of

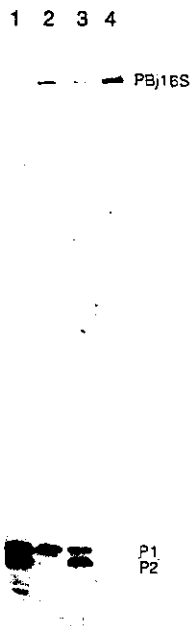


FIG. 3. Effect of promoter mutations on *fixRnifA* mRNA expression. The effect of mutations at the conserved σ^{54} -12 and -24 *fixRp*₁ promoter elements on transcription initiation was analyzed by primer extension with total RNA from anaerobic cultures from the following strains: lane 1, 8085 (wild type); lane 2, 8091 [A(-68)-to-C mutation in wild type]; lane 3, 7291R [TG(-26, -25) to CT]; lane 4, 7292R [GC(-13, -12) to AG]. Oligonucleotides lac4 and PBj16S were simultaneously used as primers for reverse transcription to detect the *fixRnifA* (P1 and P2) and 16S rRNA (PBj16S) transcripts respectively. The extension products were analyzed by electrophoresis on 6% polyacrylamide gels along with a G lane of a sequence ladder as described in Fig. 1.

the reported bacterial promoters, although the conserved stretches are positioned at the typical distance from the transcription start site. It is likely that these regions represent the consensus promoter sequences for the housekeeping *B. japonicum* RNA polymerase holoenzyme, $E\sigma^{96}$, and we tentatively named them the -35 and the -10 boxes.

We aligned the *fixR* promoter sequence with those of the *glnA* and *fbcFH* promoters by using the P2 start site as a reference (Fig. 2). Interestingly, in the -10 region, five positions are identical among these promoters, while in the -35 region, only four residues are conserved. The great majority of the positively regulated promoters have very poor or nonconserved -35 regions (10, 33). The *fixRp*₂ promoter is positively regulated; therefore, the lack of a strongly conserved -35 promoter element is consistent with this observation. The putative housekeeping -10 promoter element completely overlaps the σ^{54} -dependent -12 promoter element, while the -24 promoter element lies between the -10 and the -35 boxes in a region normally not critical for the function of -35/-10 promoters.

There is extensive evidence that mutations that disrupt any of the conserved -12 or -24 promoter elements of σ^{54} promoters have a strong promoter-down phenotype (3-5, 19, 31). We determined the effect of substitutions in these conserved regions on *fixR* promoter function. Replacement of the GC dinucleotide of the -12 promoter element by AG (strain 7292R) caused a severe reduction in *fixR-lacZ* expression (Table 1 [40]), and, consistently, no transcripts were observed (Fig. 3, lane 4), while changing the TG dinucleotide of the -24 promoter element by CT (strain 7291R) caused a slight reduction in *fixR-lacZ* expression (Table 1 [40]), and the transcript

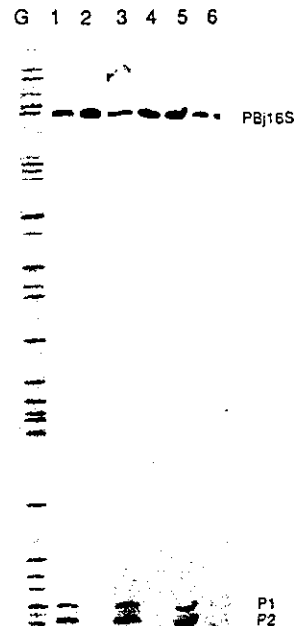


FIG. 4. Effect of oxygen on the *fixRnifA* transcripts. Total RNA was purified from the following strains grown aerobically in PSY medium (34): lane 1, 8085 (wild type); lane 2, 8091 [A(-68)-to-C mutation in wild type]; lane 3, A8085 (*nifA*); lane 4, A8091 [A(-68)-to-C mutation in *nifA* strain]; lane 5, N8085 (*rpoN*₁ *rpoN*₂); lane 6, N8091 [A(-68)-to-C mutation in *rpoN*₁ *rpoN*₂ strain]. Oligonucleotides lac4 and PBj16S were used as primers, and the cDNA was separated by electrophoresis in a 6% polyacrylamide gel.

pattern (Fig. 3, lane 3) was similar to that observed in the *rpoN*₁ *rpoN*₂ strain (Fig. 1, lane 5). We conclude from these observations that the GC dinucleotide, which forms part of the overlapping -10 and -12 promoter elements, is essential for the function of both promoters, whereas the -24 promoter element only contributes to the function of *fixRp*₁. This is in agreement with the fact that for all bacterial promoters except those of the -24/-12 type, the nucleotides between the -10 and the -35 regions are not critical for promoter function.

Influence of oxygen on *fixRnifA* expression. We analyzed the role of oxygen in *fixRnifA* expression by carrying out transcription mapping experiments with RNA isolated from aerobically grown cells. The NifA protein of *B. japonicum* is sensitive to oxygen in vivo, and the expression of the *rpoN*₁ gene is controlled by the FixLJ and FixK₂ proteins, such that it is only expressed in anaerobiosis (13, 17, 20). Therefore, we anticipated that in aerobically grown cells, the NifA-dependent promoter would not be active and that the only promoter used in these conditions is *fixRp*₂. Figure 4 shows that in the wild-type strain, the P1 transcript was greatly diminished compared with under anaerobic conditions (lane 1). The same pattern of transcripts was observed in the *nifA* (lane 3) and *rpoN*₁ *rpoN*₂ backgrounds (lane 5). Strains carrying the A(-68)-to-C promoter mutation did not show any transcript regardless of the background (lanes 2, 4, and 6), confirming that *fixRp*₁ is only expressed in anaerobic cultures and that no other protein of the σ^{54} family of positive regulators activates this promoter. β -Galactosidase expression from a *fixR-lacZ* gene fusion correlated with the levels of transcripts detected (Table 1).

In contrast to the anaerobic conditions, we observed in this experiment that the levels of *fixRp*₂ expression were about the

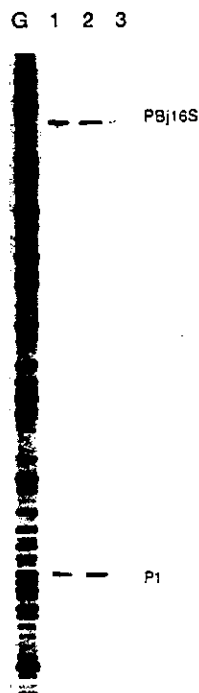


FIG. 5. Promoter usage in root nodule bacteroids. Total RNA was isolated from root nodule bacteroids 30 days after inoculation with the following strains: lane 1, 8085 (wild type); lane 2, 8091 [A(-68)-to-C mutation in wild type]; lane 3, N8085 (*rpoN₁ rpoN₂*). Transcripts P1 and PBj16S were detected as described in the legend to Fig. 1 alongside a G lane of a sequence ladder.

same in the *nifA* background as in the *rpoN₁ rpoN₂* background. Thus, it seems that in aerobically grown cells, $E\sigma^{54}$ does not compete with the other holoenzyme for the binding to the *fixR* promoter region. This probably reflects the lower level of $E\sigma^{54}$ in aerobic conditions (recall that *rpoN₁* is expressed only at reduced oxygen levels [13, 20]). Interestingly, this result also shows that oxygen modulates the expression of *fixRp₂*, because in the absence of σ^{54} , the anaerobic expression from this promoter is higher than that in aerobiosis (Table 1). This observation shows that the two regulatory systems that modulate *fixRnifA* expression are controlled at least by oxygen, although the regulation of the *fixRp₂* promoter varies only by a factor of 6. This result was previously obscured by the competition effect observed by the unproductively bound $E\sigma^{54}$ in the *nifA* background.

Expression of *fixRnifA* in bacteroids. To determine which *fixR* promoter is used during symbiosis, total RNA from bacteroids isolated from soybean root nodules 30 days after inoculation was purified and analyzed by primer extension. For this experiment, we analyzed the *fixR* transcripts of the wild-type and *rpoN₁ rpoN₂* strains. We also determined the effect of the A(-68)-to-C mutation in the *fixR* upstream promoter region. The necrotic phenotype of the nodule tissue elicited by a *nifA* mutant strain (14) did not allow us to use the A8085 and the A8091 strains in this experiment. Figure 5 shows the primer extension products. Surprisingly, the 8085 strain showed only the P1 transcript, showing that *fixRp₁* is the only promoter used in 30-day-old soybean bacteroids (Fig. 5, lane 1). It is consistent with this observation that the 8091 strain also produced only the P1 transcript, because the A(-68)-to-C mutation specifically affected the *fixRp₂* promoter (Fig. 5, lane 2). Although strain N8085 induces small Fix^- nodules (20), we were able to isolate enough RNA to carry out a primer extension experi-

ment. We did not detect any transcript (Fig. 5, lane 3), suggesting that the lack of expression of *fixRp₂* observed in the wild-type strain was not due to strong competition with $E\sigma^{54}$. Because the *fixRp₂* promoter is positively regulated by a protein that binds to the upstream -68 promoter region, lack of expression of this promoter could be the result of the absence of this protein in 30-day-old bacteroids. Alternatively, this regulatory protein could be inactive during this stage of the symbiosis.

Kullik et al. (20) showed that a *fixR-lacZ* fusion in an *rpoN₁ rpoN₂* background is expressed in 20-day-old bacteroids at about 70% of the level observed in a wild-type background. The apparent discrepancy with the results presented here could be due to the later time during symbiosis at which the RNA was isolated for the primer extension analysis (30-day-old bacteroids). It might be of interest to determine the temporal expression of this promoter in the course of development of the symbiosis.

Conclusions. The experiments presented here show that the *fixRnifA* operon is transcribed from two positively controlled overlapping promoters: one which is autoactivated by NifA and a second which is regulated at least by oxygen. These promoters are most likely recognized by different RNA polymerase holoenzymes. $E\sigma^{54}$ and, probably, $E\sigma^{96}$, that can initiate transcription at the same site, although the σ^{54} -independent polymerase initiates at the P1 start site less frequently than at P2. Such a pattern of transcription suggests that the two holoenzymes must bind to the same DNA region, which happens to contain two distinct and overlapping promoters. Overlapping promoters for different RNA polymerase holoenzymes have been documented in *Bacillus subtilis* and *Escherichia coli*. The *B. subtilis spoVC* and *spoVG* operons are transcribed by $E\sigma^{37}$ and $E\sigma^{32}$. In the *spoVG* operon, the two transcription start sites are 10 nucleotides apart, so that the promoters are partially overlapping, while in the *spoVC* operon the transcription start site is the same for both holoenzymes, suggesting that the two promoters must overlap considerably (18). The *E. coli gapA*, *rrnB*, and *dps* promoters are also transcribed from multiple promoters recognized by $E\sigma^{70}$ and $E\sigma^{32}$, or $E\sigma^s$ in the last example (1, 8, 30). However, naturally occurring, overlapping promoters for $E\sigma^{54}$ and another holoenzyme have not been described. The enterobacterial *glnAntrBC* and *glnHPQ* operons are transcribed from two promoters, one recognized by $E\sigma^{54}$ and the other recognized by $E\sigma^{70}$. These promoters, however, are far apart, such that there is no competition between the RNA polymerase holoenzymes for binding, although the binding sites for the activator protein NRI overlap with the σ^{70} promoter (9, 12, 37). We believe that in all overlapping promoters, including *fixRp₁* and *fixRp₂*, the RNA polymerase holoenzymes must contact the DNA in different fashions, perhaps by approaching the promoter DNA from different faces of the DNA helix and/or recognizing different nucleotides for binding. For the *E. coli rrnB* P1 promoter it has been shown by in vitro dimethyl sulfate and $KMnO_4$ footprinting experiments that the two RNA polymerase holoenzymes differ in their interaction with the promoter DNA (30). In vivo dimethyl sulfate genomic footprinting analysis of the *fixR* promoter region will now be carried out to determine the binding pattern of each holoenzyme.

ACKNOWLEDGMENTS

We are grateful to L. Oivera and P. Brower for technical assistance, P. Gaytán and E. López for oligonucleotide synthesis, and B. Thöny for providing *B. japonicum* 7291R and 7292R. We also thank J. Mora for fruitful discussions and X. Soberón for comments on the manuscript.

This work was supported by grants from Consejo Nacional de Ciencia y Tecnología, México (N9109-0670) and the Swiss National Foundation for Scientific Research. H.B. was supported by a CONACYT scholarship.

REFERENCES

- Altuvia, S., M. Almiron, G. Huisman, R. Kolter, and G. Storz. 1994. The *dps* promoter is activated by OxyR during growth and by IHF and σ^D in stationary phase. *Mol. Microbiol.* 13:265-272.
- Austin, S., M. Buck, W. Cannon, T. Eydman, and R. Dixon. 1994. Purification and *in vitro* activities of the native nitrogen fixation control proteins NifA and NifL. *J. Bacteriol.* 176:3460-3465.
- Brown, S. E., and F. M. Ausubel. 1984. Mutations affecting regulation of the *Klebsiella pneumoniae nifH* (nitrogenase reductase) promoter. *J. Bacteriol.* 157:143-147.
- Buck, M. 1986. Deletion analysis of the *Klebsiella pneumoniae* nitrogenase promoter: importance of spacing between conserved sequences around positions -12 and -24 for activation by the *nifA* and *nifC* (*glnG*) products. *J. Bacteriol.* 166:545-551.
- Buck, M., H. Khan, and R. Dixon. 1985. Site-directed mutagenesis of the *Klebsiella pneumoniae nifL* and *nifH* promoters and *in vitro* analysis of promoter activity. *Nucleic Acids Res.* 13:7621-7638.
- Cannon, W., and M. Buck. 1992. Central domain of the positive control protein NifA and its role in transcriptional activation. *J. Mol. Biol.* 225:271-286.
- Carlson, T. A., M. L. Guerinot, and B. K. Chelm. 1985. Characterization of the gene encoding glutamine synthetase I (*glnA*) from *Bradyrhizobium japonicum*. *J. Bacteriol.* 162:698-703.
- Charpentier, B., and C. Branlant. 1994. The *Escherichia coli gapA* gene is transcribed by vegetative RNA polymerase holoenzyme $E\sigma^{70}$ and by the heat shock RNA polymerase $E\sigma^{32}$. *J. Bacteriol.* 176:830-839.
- Claverie-Martin, F., and B. Magasanik. 1991. Role of integration host factor in the regulation of the *glnHp2* promoter of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 88:1631-1635.
- Collado-Vides, J., B. Magasanik, and J. D. Gralla. 1991. Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.* 55:371-394.
- Daniel, R. M., and C. A. Appleby. 1972. Anaerobic-nitrate, symbiotic and aerobic growth of *Rhizobium japonicum*: effects on cytochrome p450, other haemoproteins, nitrate and nitrite reductases. *Biochim. Biophys. Acta* 275:347-354.
- Dixon, R. 1984. Tandem promoters determine regulation of the *Klebsiella pneumoniae* glutamine synthetase (*glnA*) gene. *Nucleic Acids Res.* 12:7811-7830.
- Fischer, H.-M. 1994. Genetic regulation of nitrogen fixation in rhizobia. *Microbiol. Rev.* 58:352-386.
- Fischer, H. M., A. Alvarez-Morales, and H. Hennecke. 1986. The pleiotropic nature of symbiotic regulatory mutants: *Bradyrhizobium japonicum nifA* gene is involved in control of *nif* gene expression and formation of determinate symbiosis. *EMBO J.* 5:1165-1173.
- Fischer, H. M., D. Anthamatten, I. Kullik, E. Morett, G. Acuña, and H. Hennecke. 1991. Complex regulatory network for *nif* and *fix* gene expression in *Bradyrhizobium japonicum*, p. 203-210. In H. Hennecke and D. P. S. Verma (ed.), *Advances in molecular genetics of plant-microbe interactions*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Fischer, H. M., T. Bruderer, and H. Hennecke. 1988. Essential and non-essential domains in the *Bradyrhizobium japonicum* NifA protein: identification of indispensable cysteine residues potentially involved in redox activity and/or metal binding. *Nucleic Acids Res.* 16:2207-2224.
- Fischer, H. M., and H. Hennecke. 1987. Direct response of *Bradyrhizobium japonicum nifA*-mediated *nif* gene regulation to cellular oxygen status. *Mol. Gen. Genet.* 209:621-626.
- Johnson, W. C., C. P. Moran, and R. Losick. 1983. Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. *Nature (London)* 302:800-804.
- Khan, H., M. Buck, and R. Dixon. 1986. Deletion loop mutagenesis of the *nifL* promoter from *Klebsiella pneumoniae*: role of the -26 to -12 region in the promoter function. *Gene* 45:281-288.
- Kullik, I., S. Fritsche, H. Knobel, J. Sanjuan, H. Hennecke, and H.-M. Fischer. 1991. *Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the σ^{54} gene (*rpoN*). *J. Bacteriol.* 173:1125-1138.
- Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss. 1989. Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol. Rev.* 53:367-376.
- Lee, H.-S., D. K. Berger, and S. Kustu. 1993. Activity of purified NifA, a transcriptional activator of nitrogen fixation genes. *Proc. Natl. Acad. Sci. USA* 90:2266-2270.
- McClure, W. R. 1985. Mechanism and control of transcription in prokaryotes. *Annu. Rev. Biochem.* 54:171-204.
- Merrick, M. 1992. Regulation of nitrogen fixation genes in free-living and symbiotic bacteria, p. 835-847. In G. Stacey, R. Burris, and H. Evans (ed.), *Biological nitrogen fixation*. Chapman & Hall, New York.
- Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Morett, E., and M. Buck. 1988. NifA-dependent *in vivo* protection demonstrates that the upstream activator sequence of *nif* promoters is a protein binding site. *Proc. Natl. Acad. Sci. USA* 85:9401-9405.
- Morett, E., and M. Buck. 1989. *In vivo* studies on the interaction of RNA polymerase- σ^{54} with the *Klebsiella pneumoniae* and *Rhizobium meliloti nifH* promoters: the role of NifA in the formation of an open promoter complex. *J. Mol. Biol.* 210:65-77.
- Morett, E., H.-M. Fischer, and H. Hennecke. 1991. Influence of oxygen on DNA binding, positive control, and stability of the *Bradyrhizobium japonicum* NifA regulatory protein. *J. Bacteriol.* 173:3478-3487.
- Morett, E., and L. Segovia. 1993. The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* 175:6067-6074.
- Newlands, J. T., T. Gaal, J. Meccas, and R. L. Gourse. 1993. Transcription of the *Escherichia coli rnaB* promoter by the heat shock RNA polymerase ($E\sigma^{32}$) *in vitro*. *J. Bacteriol.* 175:661-668.
- Ow, D. W., Y. Xiong, Q. Gu, and S.-C. Shen. 1985. Mutational analysis of the *Klebsiella pneumoniae* nitrogenase promoter: sequences essential for positive control by *nifA* and *nifC* (*glnG*) products. *J. Bacteriol.* 161:868-874.
- Popham, D. L., D. Szeto, J. Keener, and S. Kustu. 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. *Science* 243:629-635.
- Raibaud, O., and M. Schwartz. 1984. Positive control of transcription in bacteria. *Annu. Rev. Genet.* 18:173-206.
- Regensburger, B., and H. Hennecke. 1983. RNA polymerase from *Rhizobium japonicum*. *Arch. Microbiol.* 135:103-109.
- Reibach, P. H., P. L. Mask, and J. G. Streeter. 1981. A rapid one-step method for the isolation of bacteroids from root nodules of soybean plants, utilizing self-generating percoll gradients. *Can. J. Microbiol.* 27:491-495.
- Reitzer, L. J., R. Bueno, W. D. Cheng, S. A. Abrams, D. M. Rothstein, T. P. Hunt, B. Tyler, and B. Magasanik. 1987. Mutations that create new promoters suppress the σ^{54} dependence of *glnA* transcription in *Escherichia coli*. *J. Bacteriol.* 169:4279-4284.
- Reitzer, L. J., and B. Magasanik. 1985. Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. *Proc. Natl. Acad. Sci. USA* 82:1979-1983.
- Sasse-Dwight, S., and J. D. Gralla. 1988. Probing the *Escherichia coli glnALG* upstream activation mechanism *in vivo*. *Proc. Natl. Acad. Sci. USA* 85:8934-8938.
- Thöny, B., D. Anthamatten, and H. Hennecke. 1989. Dual control of the *Bradyrhizobium japonicum* symbiotic nitrogen fixation regulatory operon *fixR nifA*: analysis of *cis*- and *trans*-acting elements. *J. Bacteriol.* 171:4162-4169.
- Thöny, B., H. M. Fischer, D. Anthamatten, T. Bruderer, and H. Hennecke. 1987. The symbiotic nitrogen fixation regulatory operon (*fixRnifA*) of *Bradyrhizobium japonicum* is expressed aerobically and is subject to a novel, *nifA*-independent type of activation. *Nucleic Acids Res.* 15:8479-8499.
- Thöny-Meyer, L., and H. Hennecke. Unpublished observations.
- Thöny-Meyer, L., D. Stax, and H. Hennecke. 1989. An unusual gene cluster for the cytochrome *bcl* complex in *Bradyrhizobium japonicum* and its requirement for effective root nodule symbiosis. *Cell* 57:683-697.
- Yanagi, M., and K. Yamasato. 1993. Phylogenetic analysis of the family of *Rhizobiaceae* and related bacteria by sequencing of 16S rRNA gene using PCR and DNA sequencer. *FEMS Microbiol. Lett.* 107:115-120.

Discusión y conclusiones del apéndice I

El hecho de que el operón *fixRnifA* se expresara en condiciones aeróbicas y se hubiese identificado el pegado de una proteína sobre la región -68; Sugiera la existencia de un doble sistema de regulación transcripcional para dicho operón. Esto nos motivo ha estudiar el inicio de transcripción, ya que probablemente estos inicios podrían ser diferentes para el sistema NifA dependiente en comparación con aquel que une una proteína específicamente sobre la región -68. La información generada a partir de este interés la describo a continuación.

Hemos comprobado la existencia de un doble sistema de regulación transcripcional que controla la expresión de este operón *fixRnifA*. Los resultados muestran que la expresión de este operón es a partir de dos promotores diferentes que se encuentran sobrelapados sobre la región 5' de este operón. Uno de los promotores se denominó *fixRp₁* y se utiliza para la expresión de *fixRnifA* en condiciones microaeróbicas y simbiosis. La utilización de mutantes *nifA* y *rpoN_{1/2}*, así como la comparación con datos ya reportados, corroboran que la expresión de *fixRp₁* es dependiente tanto de σ^{54} como de la autoactivación de NifA para su expresión (56,95). El otro promotor, denominado *fixRp₂* se expresa en condiciones aeróbicas y microaeróbicas pero no a los 30 días de simbiosis; su expresión es regulada por oxígeno y depende de la integridad de la región -68.

Con el mapeo de los inicios de transcripción, se determinó que el inicio para *fixRp₁* se encuentra a la distancia óptima de una secuencia promotora -24/-12; por otro lado, el inicio de transcripción preferente para *fixRp₂* es solamente dos pares de bases más corto que el anterior aunque tiene la capacidad de iniciar en el sitio de *fixRp₁* (Apéndice I, Fig.1). Estos datos, aunados a la utilización de mutantes sobre la región -12 que afectan totalmente la expresión de *fixRnifA* en todas las condiciones, nos sugieren fuertemente que ambos promotores se encuentran sobrelapados. Con la comparación entre las secuencias promotoras de los genes *fbcHF*, *glnA* y *fixR* hemos propuesto que el promotor *fixRp₂* presenta las características de un promotor

reconocido por un miembro de la familia de σ^{70} (Figura 16).

De esta manera tenemos que el promotor *fixRp₁* lo reconoce una RNA polimerasa- σ^{54} , además de necesitar de la proteína NifA para su expresión en microaerobiosis y simbiosis; a este sistema de regulación lo denominamos NifA dependiente. En cambio, el promotor *fixRp₂* lo reconoce una RNA polimerasa holoenzima de la familia de σ^{70} y necesita de la integridad de la región -68 para su expresión en condiciones aeróbicas y microaeróbicas; a este sistema de regulación lo denominamos RegR dependiente.

PROMOTOR *fixR* p1

-24	-12	P1
TTTACGTACAT-CGCAGT <u>TGGCGCAAATCCTGCT</u> ACGCGTGCGCGG@ <i>fixR</i>		
	<u>TGGCACRNNNNTTGCW</u>	σ^{54} consenso

PROMOTOR *fixR* p2

"-35"	"-10"	P2
AATCC <u>GGGCTGG</u> ACGCTATCTGAGCC <u>GGTGCT</u> ATCCCGCATTTT@ <i>glnA</i>		
TTAGCGGGCT <u>TGCGCTCGCCCGGCAAGGGTGCCA</u> ATAGAACGATT@ <i>fbcFH</i>		
TTTACGTACAT-CGCAGTGGCGCAAATCCT <u>TGCT</u> ACGCGTGCGCGG@ <i>fixR</i>		

Fig. 16. A) Comparación del promotor *fixRp₁* contra el consenso del promotor σ^{54} a partir de su inicio de transcripción. B) Comparación entre las secuencias promotoras de los genes *fbcHF*, *glnA* y *fixR*, ejemplificando al promotor *fixRp₂*.

El patrón de expresión de *fixRnifA* observado en condiciones microaeróbicas en un fondo silvestre, comparado con la expresión en un fondo *rpoN_{1/2}*, muestra claramente una competencia de ambas RNA-polimerasas por el reconocimiento de la región promotora (Apéndice I, Tabla.1). Está bien documentado que mientras más cercano al consenso se encuentren los promotores -24/-12, la RNA-pol- σ^{54} tiene la capacidad de formar complejos más estables (12). Sugerimos que el promotor *fixRp₂* no tiene secuencias lo suficientemente cercanas al consenso como para ser utilizado preferentemente para la expresión de *fixRnifA*; de ahí que la holoenzima que reconoce el promotor *fixRp₂* sea desplazada por la RNA-pol- σ^{54} . La necesidad de utilizar el activador transcripcional RegR que interactúa en la región -68, no implica que este compensando la carencia de un promotor fuerte. Al igual que el promotor *fixRp₁* el cual, aun teniendo un promotor consenso -24/-12, necesita de la intervención de NifA como sucede en la familia de las EBPs para activar la transcripción.

Hemos reportado que la expresión del operón *fixRnifA*, a los 30 días de entrar en simbiosis, se controla exclusivamente por el promotor *fixRp₁*. Esto sugiere que *fixRp₂* no es funcional al menos en esta etapa de la simbiosis; sin embargo, detectamos la expresión de ambos promotores en condiciones microaeróbicas (Apéndice I, Figura 1 y 3). Probablemente la expresión de *fixRp₂* en aerobiosis es basal, ya que lo reconoce una RNA polimerasa de la familia de σ^{70} que tiene la capacidad de transcribir constitutivamente a bajos niveles. Al entrar en microaerobiosis, la expresión de *fixRp₂* aumenta por lo menos 6 veces, influenciada por RegR que se regula por oxígeno y contacta a la región -68. Así podríamos sugerir que el promotor *fixRp₂* se necesita para la expresión de la proteína FixR o algún otro blanco, el cual puede estar jugando un papel importante durante la microaerobiosis y lo mismo puede estar pasando durante las etapas tempranas de la simbiosis donde *fixRp₂* pudiera ser esencial. Esto se ha sugerido después de la identificación de la proteína denominada RegR que interacciona en la región -68. Las mutaciones en

regR afectan drásticamente la simbiosis y provocan una pérdida casi total en la fijación de nitrógeno. Esto apoya la propuesta de su importancia en las etapas tempranas de la simbiosis y su posible intervención en la respiración microaeróbica (8).

Otra propuesta sobre el uso de un doble sistema de regulación para la expresión de *fixRnifA*, es que el promotor *fixRp₂* activa basalmente la expresión del operón en condiciones aeróbicas, para sintetizar constantemente a NifA. La síntesis de NifA puede ser necesaria cuando la bacteria cambia de un estilo de vida libre a simbiosis. Sin embargo, el gasto energético de la bacteria por la síntesis de una proteína que es lábil en aerobiosis no tendría mucho sentido (73). Si se considera que FixR tiene similitud con una deshidrogenasa dependiente de NAD y se sintetiza en las mismas condiciones que NifA, podríamos sugerir que la función de FixR es llevar a un estado reducido a NifA cuando ésta se encuentra oxidada en aerobiosis dejándola en su forma activa. Se sabe que NifA de *B.japonicum* después de entrar en contacto con el oxígeno se oxida e inactiva, además de que es propensa a degradarse por proteólisis, siendo incapaz de reactivarse de nuevo. Sin embargo, estos experimentos se han realizado en *E. coli* y en ausencia de FixR: Hasta ahora no hay datos que soporten la hipótesis de reactivación de NifA por FixR aunque es poco probable que suceda (57,73).

Hasta este momento se conocen algunos ejemplos en bacterias, donde se observa el uso de dos polimerasas distintas con sitios de reconocimiento sobrelapados. Entre ellos tenemos los genes *yvyD* de *B.subtilis*, que utiliza σ^H y σ^B , *groESL* de *N.gonorrhoeae*, que utiliza σ^{32} y σ^{70} ; *gal* de *E.coli* que utiliza σ^{38} y σ^{70} , *groESL* de *P.aeruginosa* que utiliza σ^{32} y σ^{70} , entre otros (27,31,55,92). El uso de promotores sobrelapados es importante en la expresión diferencial de un gen o de un operón, dependiendo de las condiciones fisiológicas en las que se encuentra la célula. Esto puede implicar que mientras más complejos son los sistemas de regulación, más frágiles tienden a ser. A pesar de ello, la utilización de promotores

múltiples puede servir como una ventaja evolutiva, la cual permite a la célula utilizar más de un sistema para la expresión de genes comprometidos en funciones específicas para su sobrevivencia.

APENDICE II

Barrios H, Grande R, Olvera L and Morett E. (1998). In vivo genomic footprinting analysis reveal that the complex *Bradyrhizobium japonicum* *fixRnifA* promoter region is differently occupied by two distinct RNA polymerase holoenzymes. *Proc Natl Acad Sci USA* . 95:1014-1019.

In vivo genomic footprinting analysis reveals that the complex *Bradyrhizobium japonicum fixRnifA* promoter region is differently occupied by two distinct RNA polymerase holoenzymes

(σ^{54} /open complex/transcription/protein-DNA interactions)

HUMBERTO BARRIOS*, RICARDO GRANDE*, LETICIA OLVERA, AND ENRIQUE MORETT†

Departamento de Reconocimiento Molecular y Bioestructura, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apartado Postal 510-3, Cuernavaca, Morelos, Mexico

Communicated by Sydney Kustu, University of California, Berkeley, CA, November 20, 1997 (received for review August 11, 1997)

ABSTRACT The *Bradyrhizobium japonicum fixRnifA* operon is transcribed from two promoters: *fixRp1*, a -24/-12 promoter recognized by the σ^{54} -holoenzyme form of the RNA polymerase, and *fixRp2*, a -35/-10 promoter that is transcribed by a second, unidentified, form of RNA polymerase holoenzyme. The *fixRp1* promoter is autoregulated during microaerobiosis by NifA, whereas *fixRp2* is also activated, but by a different regulatory protein. The main transcription start sites for these promoters are just two nucleotides apart, such that the conserved -12 and -10 regions of *fixRp1* and *fixRp2*, respectively, must overlap each other, whereas the -24 and -35 regions lie one DNA helical turn apart. Using *in vivo* genomic dimethyl sulfate and KMnO₄ footprinting, we showed that the promoter region is differentially protected, depending upon which holoenzyme is bound. Mutagenesis analyses indicated that positions from -12 to -14 are critical for the activity of both promoters, whereas mutations at -10 and -11 affected mainly *fixRp2* expression. When the sequence of the putative -35 region of *fixRp2* was modified to match the putative consensus, expression from this promoter was increased 3-fold and the reactivity toward KMnO₄, but not the transcriptional start site, moved two nucleotides further upstream, indicating that the altered promoter forms a different open complex. Additionally, we detected NifA-dependent methylation protection of two atypical NifA binding sites and protection of guanine -75. The latter residue is located in a region critical for *fixRp2* promoter activation. The results present direct physical evidence of the complexity of the organization, regulation, and function of the *fixRnifA* promoter region.

In *Bradyrhizobium japonicum*, the root nodule endosymbiont of soybean plants, as in most of the nitrogen fixing proteobacterial species, genetic control of free-living (*nif*) and symbiotic (*fix*) nitrogen fixation gene expression is exerted by NifA. This protein belongs to the enhancer-binding protein (EBP) family of regulators that activate transcription from promoters recognized by RNA polymerase holoenzyme with the alternative sigma factor σ^{54} ($E\sigma^{54}$) (reviewed in refs. 1-3). Regulation of *nif* and *fix* gene expression is finely controlled by oxygen, which regulates both the expression and the activity of NifA (4). Promoters recognized by $E\sigma^{54}$ are unique in having two DNA boxes unusually close together, centered at -12 and -24 nucleotides upstream from the transcription start site, instead of the more common -10 and -35 boxes found in most of the bacterial promoters (5). $E\sigma^{54}$ binds to these sequences and forms a stable closed complex that is isomerized to the

transcriptionally active open promoter complex only by the appropriate EBP, in a process that requires ATP hydrolysis (6-9). To activate transcription the EBP binds to enhancer-like elements located far from the promoter, typically more than 100 bp upstream (10-12). Genetic and *in vivo* footprinting analyses have shown that NifA binds to DNA sequences of the type TGT-N₁₀-ACA, called upstream activator sequences (UASs) (11, 13). The *nif* and *fix* genes have up to three UASs, and the number of these elements in a given promoter contributes to the level of NifA-mediated activation (14, 15).

The *B. japonicum* NifA protein is encoded in the *fixRnifA* operon. This operon is positively autoregulated under anaerobiosis and, unexpectedly, it is also expressed under aerobic conditions, albeit at a lower level (16, 17). Deletion, primer extension, and mutagenesis analyses have shown that a region around position -68 is required for the aerobic expression (16, 17). Gel retardation experiments, using cell extracts, showed that this region is the binding site for a protein (17). We have previously shown that the pattern of expression of *fixRnifA* is due to the presence of two overlapping promoters: *fixRp1*, which belongs to the -24/-12 class, and *fixRp2*, which shares sequence identity with the -35 and -10 regions found in a set of other *B. japonicum* promoters (16). The nucleotide sequence of the latter promoters does not show similarity to any other known bacterial promoter and probably represents a new type of promoter.

Primer extension analyses showed that *fixRp1* directs the synthesis of a transcript, called p1, that starts 12 nucleotides downstream of the -12 region. In addition to σ^{54} , p1 is dependent on NifA, and it is observed only at low oxygen tensions. Transcripts originating from *fixRp2* start at two sites: one coincides with the start site for p1, whereas the more abundant transcript, designated p2, initiates just 2 nucleotides further downstream (16). Expression from *fixRp2* requires the integrity of the upstream -68 promoter region but it is independent of σ^{54} and NifA. This promoter is expressed under aerobic and anaerobic conditions but it is not expressed in root nodule bacteroids (16). Mutations in the -24 region of the σ^{54} -dependent *fixRp1* promoter impair only the p1 transcript, showing that this region is not required for the expression of *fixRp2*. Conversely, mutants in the conserved -12 region of the *fixRp1* promoter do not show any transcript, suggesting that these changes simultaneously disrupt the overlapping -10 region of the *fixRp2* promoter. In agreement with this observation, in a strain lacking σ^{54} the expression from the *fixRp2* promoter is enhanced severalfold, suggesting that, in the wild-type strain, the two RNA polymerase holoenzymes compete for binding to the same promoter region (16).

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.

© 1998 by The National Academy of Sciences 0027-8424/98/951014-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

Abbreviations: DMS, dimethyl sulfate; UAS, upstream activator sequence.

*H.B. and R.G. contributed equally to this work.

†To whom reprint request should be addressed. e-mail: emorett@ibt.unam.mx.

In this report we analyze the *in vivo* chromosomal occupancy and isomerization of *fixRnifA* wild-type and mutant promoters, using dimethyl sulfate (DMS) and KMnO_4 , respectively. Results establish that the two RNA polymerase holoenzymes bind differently to a common promoter region and that *fixRp1* is completely within the *fixRp2* promoter. We also show NifA-dependent protection of two imperfect UASs at -100 and -121, and protection by a second regulatory protein of a guanine residue in the -68 region.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. All the strains used in this work were constructed by chromosomal integration by means of homologous recombination of the appropriate *fixR-lacZ* fusion (16). Chromosomal integration results in the duplication of the *fixRnifA* promoter region up to the site of the *lacZ* gene fusion (codon 8 of *fixR*). Strain 8085 contains the wild-type *fixR-lacZ* fusion in the wild-type *B. japonicum* 110*spc4* strain, whereas strain 8091 has an A-to-C transversion at position -68 (18). In strains A8085 and N8085 the *fixR-lacZ* fusion was integrated into the *nifA* mutant strain A9 (19), or into the *rpoN1/rpoN2* double mutant strain N50-97 (20), respectively.

Growth Conditions. *B. japonicum* cultures were normally grown aerobically in PSY medium (18) at 30°C with the following antibiotics: kanamycin at 100 µg/ml, tetracycline at 50 µg/ml, streptomycin at 100 µg/ml, and spectinomycin at 100 µg/ml. Anaerobic cultures were grown in yeast extract/mannitol medium with 10 mM KNO_3 as the electron acceptor.

DMS Genomic Footprinting. Anaerobic cultures (160 ml) were grown until an OD_{540} of 0.5-0.6, and DMS was added to a final concentration of 0.1%. After 1 min the cells were collected by centrifugation and washed twice with saline phosphate solution (11). When required, the cultures were preincubated 10 min with rifampicin (200 µg/ml) as previously reported (5). The methylated DNA was purified according to standard procedures and was cleaved by incubation with piperidine as described (21). The methylation pattern was obtained by primer extension using 9 µg of DNA, 0.5 pmol of the appropriate ^{32}P 5'-end-labeled oligonucleotide [lac4 (16) or *fixR* upstream (5'-GTGAAAGCGAACGCGGGC-3')]. These oligonucleotides prime 119 nucleotides downstream or 204 nucleotides upstream from the *fixR-lacZ* +1, respectively, and were extended by linear amplification using *Taq* DNA polymerase (GIBCO/BRL) by 20 cycles (1 min at 92°C, 1 min at 55°C, and 90 sec at 72°C).

Potassium Permanganate Reactivity Analysis. Anaerobically grown 160-ml cultures were incubated with rifampicin for 10 min before exposure to KMnO_4 (10 mM), as previously reported (5). The DNA was isolated as described above and the oxidized pyrimidine residues were detected by using the same conditions as for the methylated DNA. Primer extension products derived from the DMS or KMnO_4 modifications were separated by electrophoresis on 6% polyacrylamide gels and exposed on hyperfilm β -Max films (Amersham) or scanned in a Molecular Dynamics PhosphorImager.

Site-Directed Mutagenesis. The -10 and -35 promoter regions were mutagenized by a direct oligonucleotide method based on PCR (22). A degenerate oligonucleotide with a mixture of the four nucleotides at positions -10, -11, and -14 (CGCAAATCCNGCNGCGTGCG) was used to generate the double and triple mutants described here. For the construction of the -35 mutant promoter an oligonucleotide with four substitutions was synthesized (GCATTACGGGCTTG-GCAGTGCC). The oligonucleotides were extended by using as template plasmid pRJ7211, which contains a *fixR-lacZ* fusion, or total DNA of strain 8091, which carries a *fixR* A⁻⁶⁸-to-C promoter mutation also fused to *lacZ*, and an oligonucleotide that primes in the coding region of the *lacZ*

gene. These PCR products were used as megaprimers and extended further downstream, as reported (22). The digested products were used to replace the wild-type 331-bp *SmaI*-*Bam*HI promoter fragment, and the mutations were confirmed by sequencing the entire cloned fragment.

Chromosomal Integration of the Mutant *fixR-lacZ* Fusions. To determine the effect of the different *fixR* mutations on promoter performance, the various promoter mutations were cloned into the suicide plasmid pSUP202. This plasmid carries a tetracycline-resistance gene and cannot replicate in *B. japonicum*. Therefore when it is introduced by conjugation into *B. japonicum* and selection for tetracycline resistance is applied, integration by homologous recombination is achieved. The four mutant *fixR-lacZ* fusion derivatives were cloned into *EcoRI*/*NcoI*-digested pSUP202, as *EcoRI*-*DraI* fragments, and conjugated into the *B. japonicum* wild-type strain, and tetracycline-resistant derivatives were isolated. To confirm the site of recombination (the plasmids could recombine at low frequencies between the promoter and the *lacZ* gene, yielding a wild-type *fixR-lacZ* fusion and a mutant promoter in front of the *fixRnifA* operon) total DNA from all the transconjugant strains was purified, the *fixR-lacZ* promoter region was selectively amplified, and the resulting PCR products were sequenced. In all the strains probed the recombination event took place at the desired location, rendering the mutant promoters in front of the transcriptional *fixR-lacZ* fusions. The β -galactosidase activity of each fusion was determined as described (16).

Primer Extension Analysis. To determine the transcription start sites for the mutant *fixR-lacZ* promoters, total RNA from the *B. japonicum* transconjugant strains was purified and analyzed by primer extension, as previously described (16).

RESULTS AND DISCUSSION

NifA-Dependent DMS Protection of Two Imperfect UASs. *In vivo fixR-lacZ* expression analyses have shown that NifA autoregulates its expression during microaerobic growth (16, 17). This activation is partially dependent on sequences upstream from the promoter (17), suggesting that NifA must interact with the DNA. Fig. 1A shows the nucleotide sequence of the *fixRnifA* promoter region. Two elements resembling NifA binding sites (UAS), located at -100 (TGT-N₁₀-CCA)

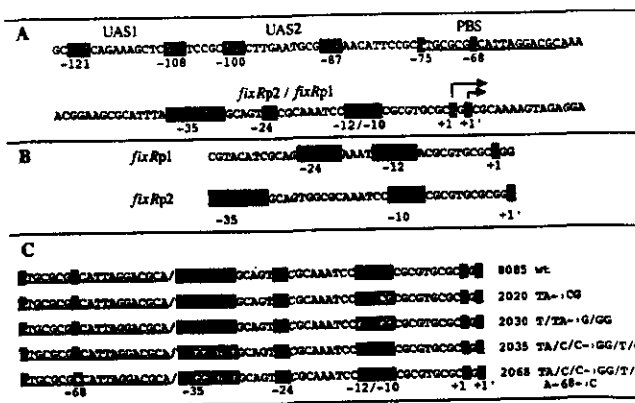


FIG. 1. Structure of the *B. japonicum fixRnifA* promoter region. (A) Nucleotide sequence of the *fixRnifA* upstream promoter region, showing the putative NifA binding sites (UAS1, centered at -121, and UAS2, centered at -100), and the protein binding site (PBS) in the -68 region. The overlapping promoter region is also shown (*fixRp2/fixRp1*). +1 indicates the start of the σ^{54} -dependent p1 transcript; +1' indicates the major transcription start site, p2, for the *fixRp2* transcript. (B) Positions presumed to serve as conserved sequences for each promoter. (C) Mutations constructed at the *fixRnifA* promoter region. Relevant positions for each element are shaded, whereas positions in boxes indicate nucleotide substitutions in the mutant promoters.

and -121 (TGT-N₁₀-ACC) nucleotides upstream from the σ^{54} -dependent transcription start site, are indicated. These sequences have five of the six positions that conform to the canonical UAS elements (10, 11). To find out whether these putative UAS elements are contacted by NifA, we carried out *in vivo* genomic DMS protection experiments in wild-type and *nifA*⁻ backgrounds carrying chromosomally integrated *fixR-lacZ* fusions.

The rationale for doing the footprinting experiments directly over genomic DNA is derived from previous observations which established that the complex pattern of *fixRnifA* gene expression is detected only in chromosomally integrated *fixR-lacZ* fusions (17). If the fusions are plasmid-born, autoactivation by NifA is obscured (23), probably by a mechanism resembling the multicopy inhibition of nitrogen fixation by plasmid-born *nif* promoter sequences observed in *Klebsiella pneumoniae* (24). Genomic *in vivo* modified DNA was isolated and used as template for cycling primer extension with *Taq* DNA polymerase using a 5'-³²P-end-labeled synthetic oligonucleotide. With this method we were able to directly analyze the *in vivo* reactivity toward DMS of a discrete zone of the chromosomal DNA, and the influence that bound proteins have in such reactivity. Fig. 2 shows the pattern of cleaved products of the *fixRnifA* upstream regulatory region after exposure of anaerobically grown 8085 (wild type, lane 2), A8085 (*nifA*⁻, lane 1), 8091 (A⁻⁶⁸ to C in wild-type background, lane 3), and N8085 (*rpoN1*⁻/*rpoN2*⁻, lane 4) strains to DMS. In the wild-type strain diminished products corresponding to positions -100 and -121 were observed compared with the *nifA*⁻ strain, indicating a NifA-dependent protection from methylation of guanine residues of the TGT motifs of both putative UASs. Strain 8091, which does not express the *fixRp2* promoter, showed a protection pattern similar to that of the wild-type strain, whereas strain N8085 showed a clear protection of UAS1 and a partial protection of UAS2. Thus, the

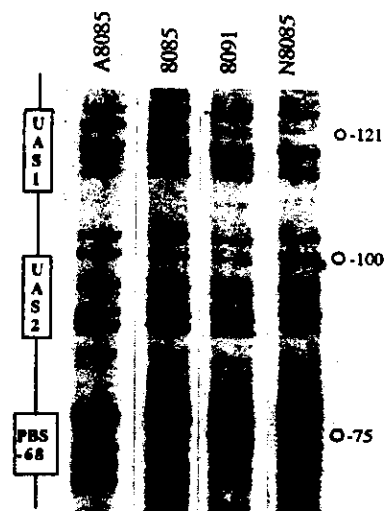


FIG. 2. Genomic DMS footprinting analysis of the *B. japonicum fixRnifA* upstream promoter region. The *B. japonicum fixRnifA* upstream promoter DNA region, integrated on the chromosome as a *fixR-lacZ* fusion, was analyzed by DMS footprinting in anaerobic cultures of strains A8085 (*nifA*⁻ background), 8085 (wild-type background), 8091 (mutant A⁻⁶⁸ to C promoter in wild-type background), and N8085 (*rpoN1*⁻/*rpoN2*⁻ background). Primer extension products, obtained by linear amplification of the genomic DNA with *Taq* DNA polymerase, were separated by gel electrophoresis and detected by autoradiography. Guanine residue -121 of the top DNA strand was protected from methylation in all the strains except in the *nifA*⁻ background, whereas guanine -100 was clearly protected in the 8085 and 8091 strains, and less protected in strain N8085. Guanine residue -75 of the protein binding site (PBS) is also shown (see Fig. 3).

protection of the UASs by NifA is independent of the activity of the promoters.

The above results show that the requirement for upstream regions for activation by NifA is due to the presence of *bona fide* UASs, and that the binding of this protein is independent of the expression of the promoters. Protection of a UAS that does not fully match the consensus has previously been reported for the *nifHc* gene of *Rhizobium etli* (25) and the *nifJ* gene of *K. pneumoniae* (26). However, these observations have been made in heterologous systems having *nifA* expressed from a strong promoter on a multicopy plasmid, whereas in our experiments the only source of NifA was the chromosomal *nifA* gene. To our knowledge this is the first report in which the chromosomal occupancy of a bacterial enhancer has been demonstrated.

Methylation Protection of the -68 Region. The *fixRp2* promoter is expressed regardless of the nitrogen and carbon sources and it is only partially regulated by oxygen; however it is not expressed in 30-day soybean root nodule bacteroids (16). The regulatory mechanism controlling this expression pattern has not been fully determined. However, genetic analysis showed that the integrity of the upstream -68 region is essential for expression from the *fixRp2* promoter. A single A-to-C transversion at position -68 abolishes *fixRp2* expression (16, 17). These observations suggest that the *fixRp2* promoter is positively regulated by a protein that binds to the -68 region. To find out whether this region is protected from methylation *in vivo* we performed genomic DMS footprinting experiments over the wild-type promoter in different genetic backgrounds and compared them with the promoter bearing the A⁻⁶⁸-to-C transversion. Fig. 3 shows the methylation pattern of this region obtained from anaerobically grown 8085, 8091, and N8085 strains (see also strain A8085 in Fig. 2, lane 1). Compared with the strain having the A⁻⁶⁸-to-C mutation (8091), a remarkable protection of G -75 and also hypermethylation of G -62 were observed in the strains with the wild-type promoter region, in all the genetic backgrounds tested. This result shows that the -68 region is a strong protein-binding site that extends at least from -68 to -75, and that the A⁻⁶⁸-to-C mutation drastically impairs the *in vivo* binding of the activator protein, as suggested by the gel retardation experiments.

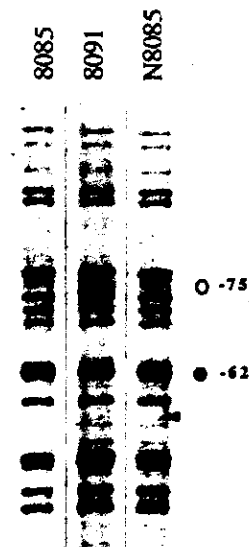


FIG. 3. Genomic DMS footprinting analysis of the *B. japonicum fixRnifA* -68 promoter region. The *in vivo* methylation pattern of the -68 region obtained from anaerobically grown 8085, 8091, and N8085 strains is shown. Protection of G -75 (○) and hypermethylation of G -62 (●) are indicated.

Interestingly, the nucleotide sequence of the -68 region showed a palindromic motif (TGCG-N₁₁-CGCA), from -74 to -56 , starting just beside the protected -75 guanine and encompassing the -68 position. Additional mutations in this region are necessary to define unambiguously the extent of the protein-binding site.

Mutagenesis Analysis of the *fixRnifA* Promoter Region. On the basis of the comparison of the nucleotide sequence upstream of the transcription start site of the *glnA*, *fbfH*, and *fixRp2* genes of *B. japonicum*, we proposed a new type of promoter with no sequence similarity to any other known bacterial promoter (Fig. 1B) (16). Because of the diverse functions accomplished by these gene products we proposed that this could be the "housekeeping" *B. japonicum* promoter, recognized by a σ^{70} homologue. Hennecke and collaborators (27, 28) have recently determined the transcription start for the 16S ribosomal RNA. They found a motif highly resembling the consensus $-35/-10$ boxes for σ^{70} holoenzymes at the right distance from the transcription start site (TTGACA-N₁₇-TATAAc-N₇+1), suggesting that the typical housekeeping promoter does exist in this organism. Moreover, they proposed that the *glnA* and *fbfH* promoters have weak sequence similarity to the canonical $-35/-10$ promoters (27). Thus, although the *fixRp2* promoter does not show sequence similarity to the canonical $-35/-10$ promoters for σ^{70} holoenzymes it could be transcribed by the most abundant (housekeeping) RNA polymerase.

To determine which nucleotides are critical for *fixRp2* we constructed a set of mutations around the conserved GC dinucleotide of the *fixRp1* -12 region (Fig. 1C), and integrated them by homologous recombination into the chromosome of the wild-type *B. japonicum* strain (see *Materials and Methods*). Table 1 shows the expression of each of the *fixR-lacZ* fusions in aerobic and anaerobic cultures. A triple mutant in which the -14 , -11 , and -10 positions were changed to Gs showed no expression in any growth condition, indicating the nucleotides replaced are required for the activity of both promoters (strain 2030). In contrast, strain 2020, in which the TA dinucleotide at $-11/-10$ was changed to CG, showed no expression under aerobic conditions, but the expression under anaerobic conditions was only slightly reduced (Table 1). This finding indicates that the *fixRp2* promoter was drastically impaired by the double mutation, whereas it had only a moderate, if any, effect on the *fixRp1* promoter. The TA dinucleotide is the only element that might conform to a canonical -10 promoter region for σ^{70} holoenzyme.

The above results are consistent with the reported extent of the $-12/-24$ promoters, which shows that there is no conservation downstream of position -11 (5). Thus, the *fixRp2* promoter extends at least to the -10 nucleotide and therefore the *fixRp1* promoter is completely within the *fixRp2* sequence.

It has been documented extensively that positively regulated promoters have a weakly conserved to nonconserved -35 region, compared with promoters expressed constitutively or subject to negative control [reviewed in ref. 29]. Our proposal of the relevant nucleotides for the *fixRp2* promoter suggests that *fixRp2* has a poorly conserved -35 region (16), in

agreement with being subject to positive regulation. To increase the similarity of the -35 region to the sequence of the *glnA* and *fbfH* promoters, we constructed a multiple replacement at this region (Fig. 1C) and integrated it into the chromosome of the wild-type *B. japonicum*. This mutant form also increases the similarity of the *fixRp2* promoter to the canonical -35 promoter region for σ^{70} holoenzyme. The resulting strain (2035) showed a marked increase in the expression of the *fixR-lacZ* fusion under both aerobic and anaerobic conditions (Table 1). Our interpretation of this result is that the multiple replacement generated a stronger *fixRp2* promoter that is expressed at higher levels, under both aerobic and anaerobic conditions (see below). It is likely that this mutation did not affect *fixRp1* because the increase in the aerobic vs. anaerobic levels of expression was similar to that in the strain with the wild-type promoter (Table 1, and see below). To evaluate whether the mutation rendered the *fixRp2* promoter less dependent on positive control, the same multiple replacement was also constructed in a background carrying the A⁻⁶⁸-to-C transversion. If the multiple nucleotide substitution resulted in a constitutive promoter the level of expression would be similar regardless of the integrity of the -68 region. Table 1 shows that the resulting strain (2068) did not express the *fixR-lacZ* fusion under aerobic conditions, whereas the anaerobic expression was only marginally reduced, as expected if only *fixRp1* were active. Thus, we conclude that this mutant promoter, despite being highly expressed, is still dependent on the binding of the regulatory protein to the -68 region. It is likely that the multiple substitution increased the affinity of the RNA polymerase for the *fixRp2* promoter. However, *in vivo* genomic footprinting analysis of this mutant promoter did not show any additional protection (data not shown).

Mapping of the Mutant *fixRp2* Transcription Start. As mentioned above, the expression from each of the *fixR* promoters can be distinguished by the transcription start site. To corroborate that the enhanced expression of strain 2035 comes from *fixRp2*, total RNA from cell cultures grown anaerobically was purified and subjected to primer extension analysis. Fig. 4 shows the extension products corresponding to RNA isolated from strains 2068, 2035, and 8085. It is remarkable that the short extension product in strain 2035 is at least 3 times more intense as compared with the long product. The former product is the main transcript of *fixRp2*, because it is missing in the strain (2068) that has in addition the A⁻⁶⁸-to-C mutation that renders *fixRp2* inactive (16). Thus, the primer extension

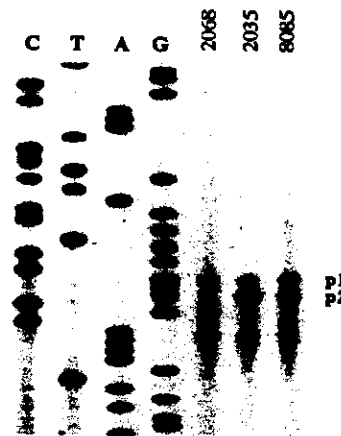


FIG. 4. Transcription start site mapping of the -35 mutant *fixRp2* mRNA. Total RNA was purified from the following strains grown anaerobically: 2068 (mutant at the -35 region in the A⁻⁶⁸-to-C *fixRnifA* promoter region), 2035 (mutant at the -35 region in wild-type *fixRnifA* promoter region), and 8085 (wild-type *fixRnifA* promoter). Transcripts p1 (mainly from the *fixRp1* promoter), and p2 (from the *fixRp2* promoter) are indicated.

Table 1. Expression of the *fixR-lacZ* promoter mutants

Strain	Relevant characteristic of <i>fixR</i> promoter*	β -Galactosidase activity, Miller units	
		Aerobic	Anaerobic
8085	Wild type	501 \pm 63	887 \pm 8
2020	TA \rightarrow CG	22 \pm 6	684 \pm 54
2030	T/TA \rightarrow G/GG	3 \pm 1	4 \pm 1
2035	TA/C/C \rightarrow GG/T/G	1,599 \pm 319	2,389 \pm 17
2068	TA/C/C \rightarrow GG/T/G	23 \pm 3	610 \pm 54

*See Fig. 1.

experiments demonstrate that the -35 mutation resulted in a more active *fixRp2* promoter and did not have any effect on *fixRp1*. Taken together, these results strongly support our interpretation that the *glnA*, *fbcFH*, and *fixRp2* promoters share functional sequence similarity.

RNA Polymerase-DNA Contacts in Open Promoter Complexes. Results of primer extension (16) and of mutagenesis experiments shown above and previously published (23) suggest that the -12 and the -10 regions of the *fixRp1* and *fixRp2* promoters, respectively, overlap each other, whereas the -24 and the -35 regions are located one helical turn apart, as depicted in Fig. 1B (16). This implies that the two holoenzymes could interact with the same residues in the $-12/-10$ region. To physically localize the contact points for each RNA polymerase holoenzyme on the promoter DNA, we carried out genomic DMS footprinting experiments in *B. japonicum* cells previously treated with rifampicin (this antibiotic inhibits transcription elongation, resulting in accumulation of stable open complexes). In the wild-type strain residues -24 , -23 , and -13 of the top strand (Fig. 5A, lane 2) and -12 of the bottom strand (Fig. 5B, lane 2), the most conserved positions of the $-24/-12$ promoters, were protected from methylation, compared with the pattern obtained with DNA methylated *in vitro* in the absence of proteins (Fig. 5, lanes 1). Hypermethylation of G -8 of the top strand was also observed (Fig. 5A, lane 2). A similar protection and hypermethylation pattern has been observed at several other σ^{54} -dependent promoters (5, 30-35). Analogous protection was observed in strain 8091 (Fig. 5, lanes 3), where the *fixRp2* promoter is not active. Conversely, the protection of the top strand was not observed in strain N8085, which lacks σ^{54} (Fig. 5A, lane 4), indicating that the protection resulted from the interaction of $E\sigma^{54}$ with the promoter DNA. Interestingly, protection of G -12 of the bottom strand was detected in this strain, when compared with the DNA methylated *in vitro* (Fig. 5B, lane 4). This result shows that the -12 position is contacted by a holoenzyme other than $E\sigma^{54}$. Moreover, because the protection was also detected in the 8091 strain, which does not express the *fixRp2* promoter

(Fig. 5B, lane 3), and this position is protected by $E\sigma^{54}$ in all the $-24/-12$ promoters analyzed to date, it is likely that G -12 is also contacted by $E\sigma^{54}$.

In a previous study we unexpectedly found that the expression of *fixRp2* was severalfold higher in a strain lacking σ^{54} than in a strain devoid of NifA. This paradoxical observation has been interpreted as the result of the formation of a stable $E\sigma^{54}$ -*fixRp1* complex that in the absence of NifA cannot proceed to clear the promoter, resulting then in a partial occlusion of *fixRp2* (16). Our observation that the -12 position is contacted by both holoenzymes is consistent with a model in which both RNA polymerases compete for the binding to the same DNA, and also with the genetic data that show that disruption at the -12 position strongly affects the expression of both promoters (16, 17).

It is remarkable that guanine -12 was the only nucleotide protected by the second polymerase. This could be due to a weak interaction of this holoenzyme with other guanine residues. Alternatively, the open complex formed by this holoenzyme may not be as stable as the ones formed by $E\sigma^{54}$.

DNA Melting at *fixRp1* and *fixRp2* Promoters. Besides the specific recognition of promoter DNA, the σ factor confers on the core RNA polymerase the ability to initiate transcription on duplex templates. This observation led to the early proposal that σ might melt the DNA by tightly binding to one DNA strand (36). Protein-DNA crosslinking studies have shown that both the β and β' subunits of the RNA polymerase and σ^{70} bind to the nontemplate strand (37-39). Also, Buck and collaborators (40) have shown crosslinking of the σ^{54} protein with promoter DNA. deHaseth and Helmann (41) have proposed that the stabilization of single-stranded DNA in the open complex could involve interactions between σ and certain groups of the nucleotide bases not accessible on duplex DNA, such as hydrogen bonding to the nucleosides, ionic interactions with the phosphate backbone, and sequence-independent base stacking interactions with specific amino acid residues. Thus it is likely that the open complexes formed by each holoenzyme over the *fixRnifA* promoter region could be different.

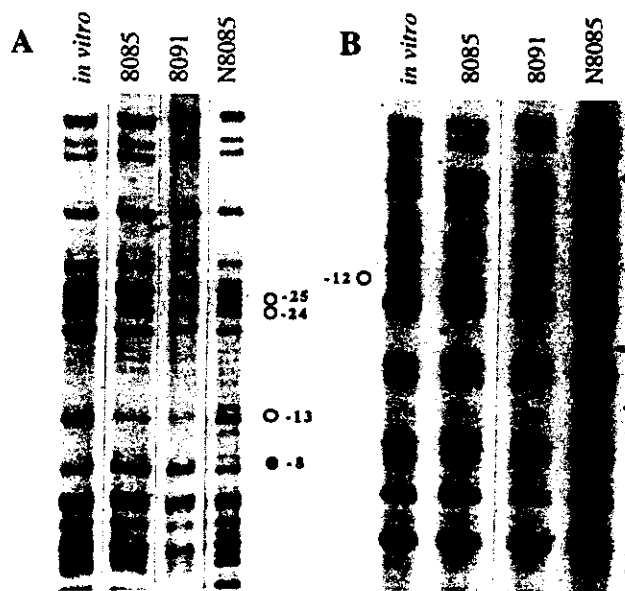


FIG. 5. *In vivo* genomic DMS footprinting analysis of the *B. japonicum fixRnifA* promoter region. Anaerobic cultures of strains 8085, 8091, and N8085 were grown for 3 days and incubated with rifampicin before the exposure to DMS. Primer extension pattern of the *fixRnifA* top (A) and bottom (B) strands is shown. Control DNA methylated *in vitro* in the absence of proteins is shown on the left. Protected positions are indicated with \circ , whereas the hypermethylated position is indicated with \bullet .

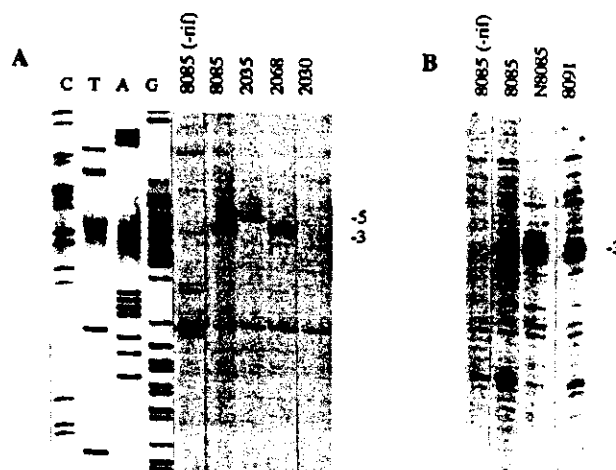


FIG. 6. *In vivo* genomic reactivity of the *fixRnifA* promoter region to potassium permanganate. Anaerobic cultures were incubated with rifampicin for 10 min and exposed to $KMnO_4$ for another 10 min, and the total DNA was purified. The first lanes of each panel show the primer extension products of DNA samples from strain 8085 that were not treated with rifampicin. (A) Effect of promoter mutations in open complex formation. Strains 8085 and 2068 showed hyperreactivity to $KMnO_4$ of position -3 , whereas in strain 2035 the hyperreactivity moved to position -5 . The triple mutant strain 2030, for which no *fixR-lacZ* expression was detected (Table 1), did not show any hyperreactivity to $KMnO_4$. (B) Comparison of $KMnO_4$ reactivity of strains 8085, N8085, and 8091. Lanes C, T, A, and G are a sequence ladder obtained with the same oligonucleotide used for the primer extension.

To investigate the nature of the open complexes, anaerobic *B. japonicum* cultures, treated with rifampicin, were exposed to KMnO_4 . This single-stranded DNA selective reagent has been very useful for detecting the extent of strand separation in several prokaryotic and eukaryotic promoters (34, 42). Genomic DNA oxidized *in vivo* was isolated and used as template for cycling primer extension with *Taq* DNA polymerase, in a manner similar to that for the DMS-treated DNA. The extent of the reactive DNA in the *fixRnifA* promoter region is shown in Fig. 6. In the strain with the wild-type *fixR* promoter region, position -3 showed a marked reactivity compared with the DNA modified in the absence of rifampicin (compare lanes 1 and 2 in Fig. 6A and B). Similar reactivity was observed when either *fixRp2* or *fixRp1* was the only active promoter (strains N8085 and 8091; Fig. 6B, lanes 3 and 4, respectively). This indicates that even though the two holoenzymes interact differently with the promoter DNA, and the transcription start sites are displaced by two nucleotides, the two holoenzymes form similar open complexes. Unexpectedly, when mutant 2035, which showed enhanced expression of *fixRp2*, was analyzed the reactivity at -3 was not observed whereas position -5 became hyperreactive (Fig. 6A, lane 3). Thus, changing the -35 promoter region resulted in the formation of a different open complex. Mutant 2068, which has the same -35 mutations in the A⁻⁶⁸-to-C background, showed hyperreactivity at -3 (Fig. 6A, lane 4). As expected, mutant 2030, which did not show expression under any conditions, did not show sensitivity to KMnO_4 (Fig. 6A, lane 5). Our interpretation of these results is that each holoenzyme generates a similar open complex; however, when *fixRp2* was modified to match the consensus, and therefore expressed at higher levels, the reactivity moved upstream, indicating that a different DNA conformation was present.

The experiments presented here provide physical and genetic evidence that the *fixRnifA* operon is transcribed from two overlapping promoters, recognized by two different RNA polymerase holoenzymes. We showed that the mechanism of activation, promoter DNA recognition, and even melting in strain 2035 is different for each holoenzyme.

This article is dedicated to the memory of Dr. Jorge Calderón. We are grateful to F. Romero for technical assistance and to P. Gaytán and E. López for oligonucleotide synthesis. We also thank J. Aguirre, M. A. Cevallos, D. Romero, and S. López for fruitful discussions and H.-M. Fischer, X. Soberón, and B. Valderrama for their comments on the manuscript. This work was supported in part by grants from the Consejo Nacional de Ciencia y Tecnología, México (N9109-0670 and 400344-5-0132P-N). H.B. and R.G. are recipients of Consejo Nacional de Ciencia y Tecnología scholarships.

- Kustu, S., Santero, E., Popham, D., Keener, J. & Weiss, D. (1989) *Microbiol. Rev.* **53**, 367-376.
- Merrick, M. J. (1993) *Mol. Microbiol.* **10**, 903-909.
- Morett, E. & Segovia, L. (1993) *J. Bacteriol.* **175**, 6067-6074.
- Fischer, H. M. (1994) *Microbiol. Rev.* **58**, 352-386.
- Morett, E. & Buck, M. (1989) *J. Mol. Biol.* **210**, 65-77.
- Austin, S., Buck, M., Cannon, W., Eydmann, T. & Dixon, R. (1994) *J. Bacteriol.* **176**, 3460-3465.
- Austin, S. & Dixon, R. (1992) *EMBO J.* **11**, 2219-2228.
- Lee, H.-s., Berger, D. K. & Kustu, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2266-2270.
- Weiss, D. S., Batut, J., Klose, K. E., Keener, J. & Kustu, S. (1991) *Cell* **67**, 155-167.
- Buck, M., Miller, S., Drummond, M. & Dixon, R. (1986) *Nature (London)* **320**, 374-378.
- Morett, E. & Buck, M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9401-9405.
- Reitzer, L. J. & Magasanik, B. (1986) *Cell* **45**, 785-792.
- Morett, E., Cannon, W. & Buck, M. (1988) *Nucleic Acids Res.* **16**, 11469-11488.
- Alvarez-Morales, A., Betancourt-Alvarez, M., Kaluza, K. & Hennecke, H. (1986) *Nucleic Acids Res.* **14**, 4207-4227.
- Gubler, M. (1989) *Mol. Microbiol.* **3**, 149-159.
- Barrios, H., Fischer, H. M., Hennecke, H. & Morett, E. (1995) *J. Bacteriol.* **177**, 1760-1765.
- Thöny, B., Anthamatten, D. & Hennecke, H. (1989) *J. Bacteriol.* **171**, 4162-4169.
- Regensburger, B. & Hennecke, H. (1983) *Arch. Microbiol.* **135**, 103-109.
- Fischer, H.-M., Alvarez-Morales, A. & Hennecke, H. (1986) *EMBO J.* **5**, 1165-1173.
- Kullik, I., Fritsche, S., Knobel, H., Sanjuan, J., Hennecke, H. & Fischer, H.-M. (1991) *J. Bacteriol.* **173**, 1125-1138.
- Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.
- Merino, E., Osuna, J., Bolivar, F. & Soberón, X. (1992) *BioTechniques* **12**, 508-510.
- Thöny, B., Fischer, H.-M., Anthamatten, D., Bruderer, T. & Hennecke, H. (1987) *Nucleic Acids Res.* **15**, 8479-8499.
- Reidel, G. E., Brown, S. E. & Ausubel, F. (1983) *J. Bacteriol.* **153**, 45-56.
- Valderrama, B., Davalos, A., Girard, L., Morett, E. & Mora, J. (1996) *J. Bacteriol.* **178**.
- Charlton, W., Cannon, W. & Buck, M. (1993) *Mol. Microbiol.* **7**, 1007-1021.
- Beck, C., Marty, R., Klausli, S., Hennecke, H. & Gottfert, M. (1997) *J. Bacteriol.* **179**, 364-369.
- Kundig, C., Beck, C., Hennecke, H. & Gottfert, M. (1995) *J. Bacteriol.* **177**, 5151-5154.
- McClure, W. R. (1985) *Annu. Rev. Biochem.* **54**, 171-204.
- Cannon, W. V., Kreuzer, R., Kent, H. M., Morett, E. & Buck, M. (1990) *Nucleic Acids Res.* **18**, 1693-1701.
- Minchin, S. D., Austin, S. & Dixon, R. A. (1989) *EMBO J.* **8**, 3491-3499.
- Morett, E., Fischer, H.-M. & Hennecke, H. (1991) *J. Bacteriol.* **173**, 3478-3487.
- Popham, D. L., Szeto, D., Keener, J. & Kustu, S. (1989) *Science* **243**, 629-635.
- Sasse-Dwight, S. & Gralla, J. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8934-8938.
- Whitehall, S., Austin, S. & Dixon, R. (1992) *J. Mol. Biol.* **225**, 591-607.
- Hinkle, D. C. & Chamberlin, M. (1970) *Cold Spring Harbor Symp. Quant. Biol.* **35**, 65-72.
- Brodolin, K. L., Studitsky, V. M. & Mirzabekov, A. D. (1993) *Nucleic Acids Res.* **21**, 5748-5753.
- Harrison, C. A., Turner, D. H. & Hinkle, D. C. (1982) *Nucleic Acids Res.* **10**, 2399-2414.
- Hillel, Z. & Wu, C. (1978) *Biochemistry* **17**, 2954-2960.
- Buck, M. & Cannon, W. (1992) *Nature (London)* **358**, 422-424.
- deHaseth, P. L. & Helmann, J. D. (1995) *Mol. Microbiol.* **16**, 817-824.
- Mirkovitch, J. & Darnell, J. E., Jr. (1992) *Mol. Biol. Cell* **3**, 1085-1094.

Discusión y conclusiones del apéndice II

Con los resultados del trabajo anterior teníamos evidencias genéticas de la posible existencia de dos polimerasa y dos activadores distintos que controlaban la expresión del operón *fixRnifA*. Nuestro objetivo en esta parte del proyecto era encontrar evidencias directas de la interacción física de las diferentes proteínas involucradas en la regulación de este operón.

En este trabajo hicimos una caracterización de los diferentes elementos involucrados en el control de la expresión del operón *fixRnifA*. Esto se realizó mediante experimentos de "footprinting" genómico *in vivo*, con dimetil sulfato (DMS) y permanganato de potasio (KMnO₄), además de la introducción de algunas mutaciones sobre la región promotora. Con los resultados de este trabajo demostramos la interacción *in vivo* de la holoenzima E- σ^{54} con la región promotora ya que las guaninas de las posiciones -25, -24 y -13 de la cadena codificante se protegen de la metilación en una cepa silvestre, pero no en una doble mutante *rpoN*_{1/2} que codifica para σ^{54} . Al realizar el análisis sobre la cadena no codificante se encontró que la posición -12 se protege de metilación en las cepas silvestre, en la doble mutante *rpoN*_{1/2} y en la mutante -68. Nuestra interpretación es que la posición -12 del promotor *fixRnifA*, es contactada por una holoenzima diferente a E- σ^{54} y que probablemente pertenece a la familia de E- σ^{70} . Por lo tanto, tenemos evidencias físicas del contacto de dos factores sigma distintos sobre la misma región promotora, donde se comparte al menos la posición -12 para la interacción de ambos sigmas (Apéndice II, Figura 5).

Con el análisis de las regiones por arriba del inicio de la transcripción, se determinó un patrón de protección sobre las guaninas -100 y -121, las cuales forman motivos TGT-N10-ACC y TGT-N10-CCA, por lo que los proponemos como UAS reconocidos por NifA (71). En la mutante NifA no se observó esta protección, con lo que concluimos que estos motivos aunque imperfectos, son reconocidos por NifA

(Apéndice II, Figura 2). En este mismo análisis se identificó que la posición -75 se protege de metilación específicamente en la cepa silvestre, pero no en la cepa con el cambio de A por C en la posición -68. Esto sugiere que la posición -75 es un sitio de contacto para el activador transcripcional RegR del sistema aeróbico y la mutación en -68 afecta el reconocimiento de este activador al DNA (Apéndice II, Figura 3).

Se ha propuesto que el factor sigma para funciones básicas o "housekeeping" de *B.japonicum* (σ^{80}) reconoce al promotor *fixRp2*; la secuencia promotora reconocida por este factor sigma es idéntica a la de σ^{70} de *E.coli* (9).

Con las evidencias de la interacción de dos holoenzimas distintas sobre la misma región promotora y con la comparación con otras secuencias promotoras de *B.japonicum* (Apéndice I, Figura 16), se construyeron mutaciones. Estas con el objetivo de ver cuales son los nucleótidos importantes para la funcionalidad del promotor *fixRp2* y poderlo separar del promotor *fixRp1*. Este trabajo específico, fue realizado por Ricardo Grande, que se encontraba incorporado al proyecto.

Mutaciones en las dos posiciones por abajo del dinucleótido conservado -12 (GC) del promotor *fixRp1*, que representan las posiciones más conservadas para el promotor reconocido por σ^{80} (TA por CG), reducen hasta en un 5% la actividad aeróbica de *fixRp2*. Por otro lado, la expresión aeróbica de un promotor con un tercer cambio en la posición contigua por arriba del (GC) (cepa 2030), reduce 7 veces su actividad en comparación con la doble mutante (cepa 2020). Si el promotor *fixRp2* fuera reconocido por σ^{80} , cualquier cambio conjuntamente en las posiciones TA, que representan los sitios más conservados de los promotores σ^{70} de *E.coli*, daría como resultado una expresión nula de este promotor (66,84). Al contrario de lo esperado, este promotor todavía puede reducir su expresión con un tercer cambio, que se encuentra por afuera de la posible región -10 propuesta para ser reconocida por σ^{80} . Con estos datos, nosotros sugerimos que el promotor *fixRp2* no es reconocido por el factor σ^{80} "housekeeping" de *B.japonicum* y sí por otro factor alternativo de la

familia de σ^{70} (Apéndice II, Figura 1 y Tabla 1) (Figura 17).

σ^{70} Tata box

<u>TTGACA</u>	<u>TATAAT</u>	E.c consenso σ^{70}
ATTGGCGTTGACAGCCCGGAAGGTGGGGCCTATAACCCAACC		B.j consenso σ^{80}
TACGTACATCGCAGTGGCGCAAA-TCCTGCTACGCGTGCGCGGG		B.j <i>fixR</i>
		p2

Figura 17. Comparación entre la secuencia consenso del los promotores σ^{70} de *E.coli*, σ^{80} de *B.japonicum* y del promotor *fixR* de *B.japonicum*.

En base a la comparación de las regiones promotoras que se muestran en el apéndice I (Figura 16), se realizaron cambios sobre la posible región -35 para acercarla al consenso que proponemos. Este experimento se propuso con la finalidad de probar si hacercamos al promotor *fixRp₂* a su consenso, este puede hacerse independiente de regulación positiva por RegR. Los resultados indican que al llevar esta región al consenso aumenta su expresión en un 100% en condiciones aeróbicas (cepa 2035). Al utilizar los mismos cambios de la región -35, pero sobre el promotor mutante -68, la expresión del promotor *fixRp₂* disminuyó casi totalmente (cepa 2068). Está bien documentado que conforme los promotores del tipo σ^{70} se alejan del consenso, necesitan de un regulador positivo que se une en las regiones por arriba del promotor para favorecer la transcripción (67). En el caso del promotor *fixRp₂* aunque se le llevó al consenso propuesto y aumentó considerablemente su expresión, éste no dejó de ser dependiente de la integridad de la región -68. Consideramos que los cambios realizados en la región -35 aumentan la afinidad de la polimerasa por el promotor pero no sin dejar de ser dependiente de la regulación positiva (Apéndice II, Tabla 1). Con estos datos y la identificación de la proteína

RegR, sabemos que el sistema que utiliza el promotor *fixRp₂* es controlado positivamente por el activador transcripcional RegR y muy probablemente regulado por oxígeno, aunque no existen evidencias concluyentes al respecto.

Tenemos documentado que la expresión del operón *fixRnifA* se encuentra bajo el control de dos RNA polimerasas distintas. Tomando en cuenta que ambas reconocen diferentes secuencias promotoras e inician la transcripción en dos sitios distintos, pero muy cercanos, analizamos los patrones de formación de complejos abiertos de la transcripción. El objetivo de estos experimentos fue ver si el promotor de *fixRnifA*, que es contactado por dos polimerasas distintas, formarían complejos abiertos de la transcripción distintos. En los casos de los promotores silvestres *fixRp₁* y *fixRp₂* la formación del complejo abierto es idéntica para ambos; la excepción fue el promotor que llevamos al consenso en -35, cuyo complejo abarcó dos pares de bases más (Apéndice II, Figura 6). Sugerimos que la modificación de la secuencia promotora -35, afecto directamente la estructura del DNA y en consecuencia la formación del complejo abierto de la transcripción. Esta propuesta se respalda al tomar en cuenta que el complejo abierto que forma el promotor que tiene los cambios en -35 y -68 (cepa 2068), es idéntico al del promotor silvestres.

Con las evidencias físicas y genéticas de la existencia de dos sistemas de regulación para la expresión de *fixRnifA*, sabemos que el promotor *fixRp₁* utiliza a E- σ^{54} y NifA para su autoactivación. NifA se une a las UAS imperfectas de este promotor en microaerobiosis y simbiosis. Nosotros sugerimos que una holoenzima con un factor sigma alternativo, perteneciente a la familia de σ^{70} y que no es σ^{80} , reconoce al promotor *fixRp₂*. Este promotor necesita de la integridad de la región -68 para su expresión y hemos identificado la unión de la proteína RegR sobre esta región, el cual es el activador positivo del sistema RegR.



APENDICE III

(Datos Adicionales)

Apéndice III (Datos adicionales)

En este punto del trabajo, quedaban algunas incógnitas por resolver sobre el doble sistema de regulación transcripcional que controla la expresión del operón *fixRnifA*. Dentro de estas se encuentra la identidad del gen de la proteína que codifica para el activador transcripcional del sistema aeróbico, así como cuál es el factor sigma que reconoce el promotor *fixRp₂*.

Por lo tanto, el interés del proyecto se enfocó en identificar al activador transcripcional del sistema -68 dependiente, para lo cual seguimos la estrategia de mutagénesis al azar con el transposón Tn5. Estos experimentos se realizaron sobre una cepa de *B.japonicum*, con una fusión transcripcional *fixR-lacZ* mutante en la región -24 integrada a genoma. Esta mutación impide el uso del sistema NifA dependiente, por lo tanto, la única expresión monitoreable proviene del sistema RegR dependiente. La generación de mutantes se realizó mediante un método en el que una cruce triparental se incubaba durante 2 ó 3 días, después de lo cual se recuperan las células y se crecen por 12 horas en medio líquido con antibióticos para su selección.

Las células recuperadas se siembran en cajas con antibióticos y se incuban por 10 días aproximadamente. Ya que existen colonias lo suficientemente desarrolladas, éstas se estrían sobre cajas con antibióticos y X-gal para detectar la expresión de la fusión transcripcional, la que tarda aproximadamente 15 días. Durante este experimento, se rayaron 80,000 colonias y solamente se obtuvieron 2 candidatas las cuales resultaron ser hermanas. El análisis de estas candidatas mostró que ocurrió una integración del plásmido que contenía el transposón Tn5 sobre la fusión transcripcional en genoma, lo que resultaba en un fenotipo de expresión espurio. A la par con esta estrategia se siguieron otras más en sistemas heterólogos para encontrar al activador que se une a la región -68, pero tampoco dieron resultados adecuados.

Al mismo tiempo que realizábamos nuestros experimentos, en el lab del Prof. Hauke Hennecke en Suiza, se trabajaba simultaneamente en la identificación del activador transcripcional del sistema RegR dependiente, mediante la estrategia de microsecuenciación. Finalmente el activador fue identificado por el grupo suizo. Este activador pertenece a la familia de las proteínas de dos componentes, donde la proteína receptora de la señal es una histidina cinasa típica con capacidad de unirse a membrana, a la que denominaron RegS. Por otro lado, la proteína que se encarga de regular la respuesta tiene la capacidad de unirse establemente a la región -68, fue denominada RegR. Mutaciones en *regR* son incapaces de activar la transcripción a partir del promotor *fixRp₂* en condiciones aeróbicas y microaeróbicas, sabiendo ahora que esta proteína es la encargada de regular la expresión del sistema RegR dependiente (8). Se ha visto que está mutante es incapaz de formar una simbiosis óptima, además de su incapacidad de fijar nitrógeno. El estímulo que controla la expresión de RegR, así como de sus genes blanco se desconocen, aunque se propone que está relacionado con los cambios intracelulares de oxido-reducción (8). En la figura 18 se muestra el modelo de los sistemas de regulación que controlan la expresión del operón *fixRnifA*.

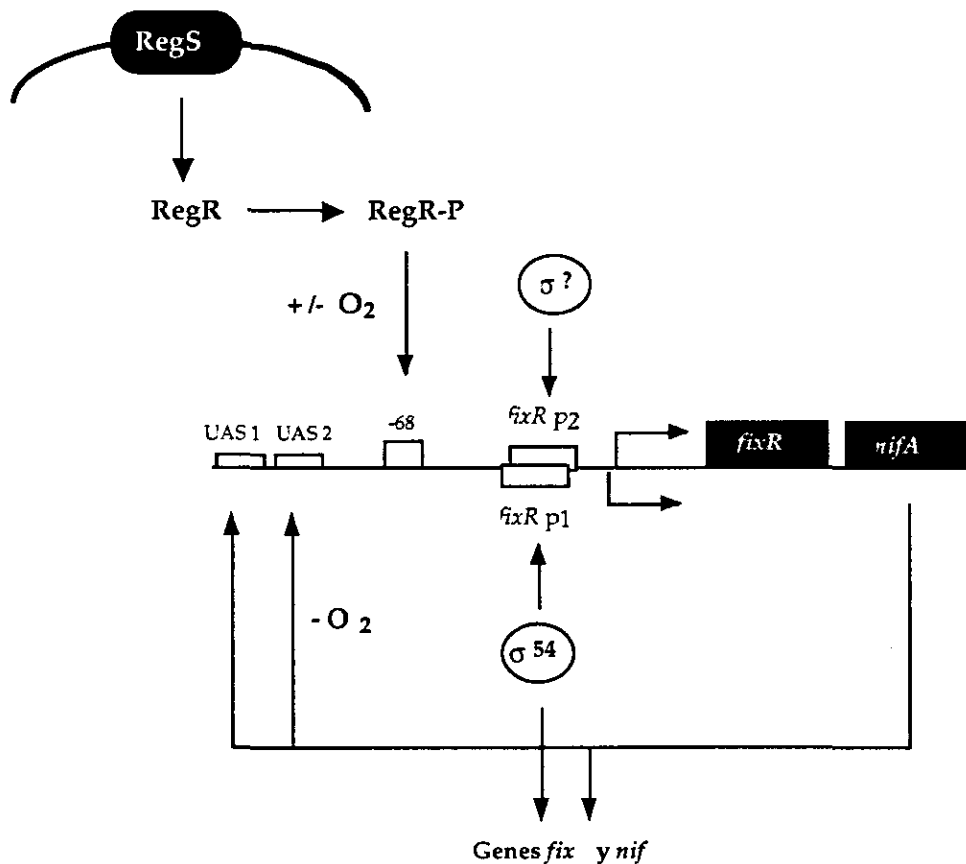


Figura 18. Esquema de los sistemas de regulación que controlan la expresión del operón *fixRnifA* en *Bradyrhizobium japonicum*.

Con los resultados anteriores nuestros objetivos se enfocaron entonces en dos cuestiones muy particulares. La primera es comprobar si los palíndromes existentes en la región -68 son los sitios primarios de reconocimiento de la proteína RegR; la segunda era saber si los dos motivos UAS que tienen secuencias imperfectas funcionan de manera independiente o cooperativa para unir multiméricamente a la proteína NifA. Para abordar ambos problemas se llevó a cabo una mutagénesis sitio dirigida, siguiendo el protocolo de Sarkar *et al* 1990 (87). Las mutaciones obtenidas se integraron al genoma de una cepa silvestre (USDA 110) de *B.japonicum* por doble recombinación homóloga, como se muestra en la figura 19.

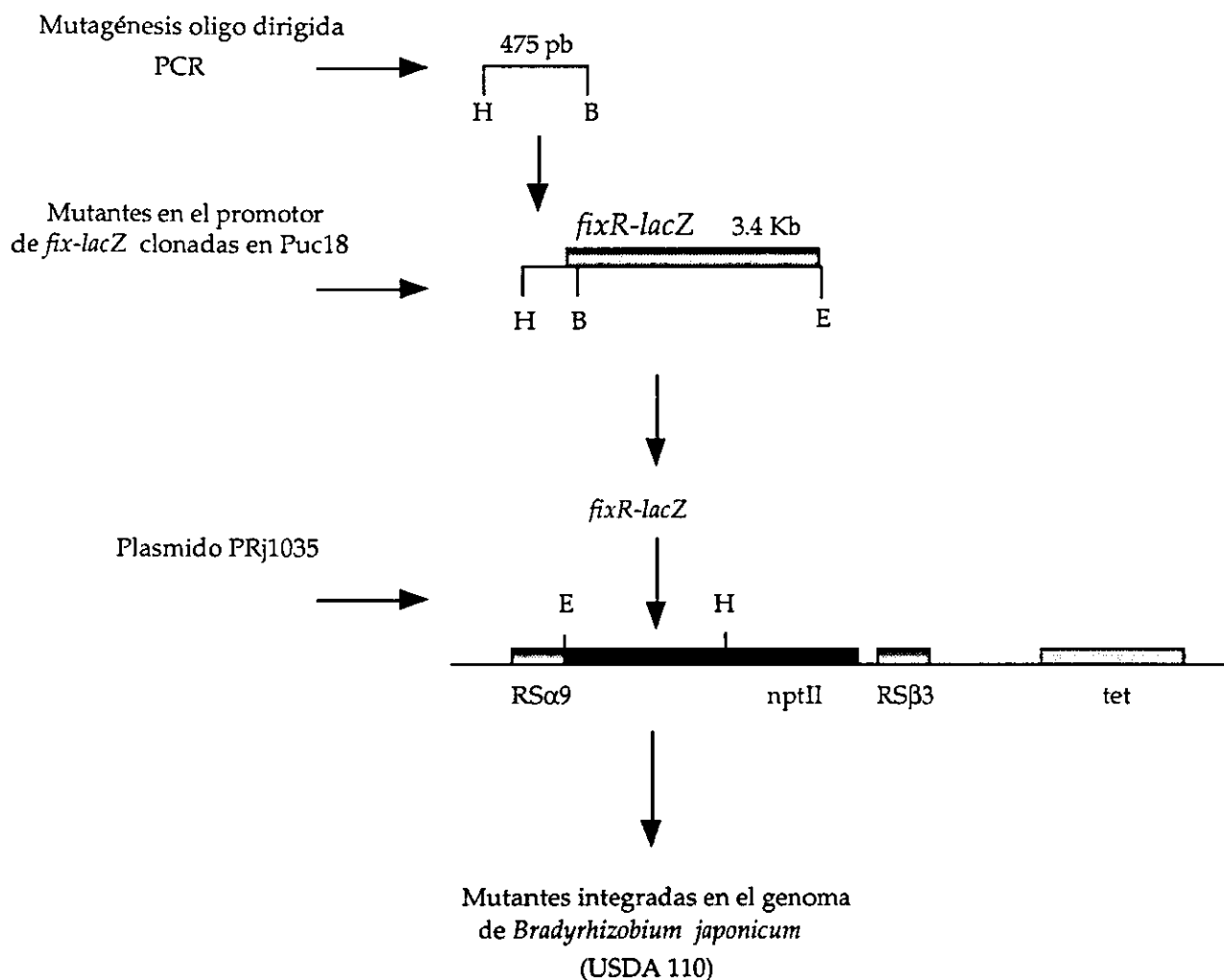


Figura 19. Para la mutagénesis en la región -68 y en las UAS se utilizaron los siguientes oligonucleótidos respectivamente: -75 (I) 5' [AAACATTCCGMGKGMGMGAMATTAGGACGC] 3', -75 (II) 5' [GCGCGACATTMGGMMGM MAAACGGAAGCGC] 3' y MUT-UAS 5' [CACCATCGCTKTCAGAAAGCTCACCTCCGCTKTCTTGAATGC] 3'. Donde M = A o C y K = G o T. Los productos de PCR fueron clonados en una fusión silvestre que se encontraba en el vector PUC18. En éste se intercambié el fragmento de 475pb H/B proveniente de la mutagénesis. Todas las mutantes se secuenciaron y reclonaron como fusión completa de 3.4Kb H/E sobre el plásmido PRj1035. Este plásmido tiene las secuencias repetidas R α 9 y R β 3, necesarias para la doble recombinación homóloga en el genoma de *B.japonicum*. Todas las mutantes se secuenciaron de genoma para verificar los cambios.

En la tabla I se muestran los resultados de las mutaciones sobre la región -68. Estas mutaciones se realizaron sobre un molde de *fixR-lacZ* con cambios en las posiciones (GG) -25/-24, esta fusión es incapaz de utilizar el sistema NifA dependiente, por lo tanto toda la expresión detectable vendría de la expresión del sistema RegR dependiente.

El análisis de expresión de todas las clonas con dobles cambios seleccionados carecen de actividad y se encuentran entre las posiciones -76 y -56. Es importante señalar que por lo menos una posición de cada doble mutante se encuentra sobre alguna de las secuencias palindrómicas. Las mutantes sencillas Bj74G y Bj56 que se realizaron sobre cada uno de los palíndromes suprimen claramente la actividad de la fusión, por otro lado, el cambio de la mutante sencilla Bj70 que se encuentra un nucleótido por afuera del palíndrome más distal de (+1), carece de actividad al igual que las anteriores.

La transversión A por C de la mutante sencilla Bj63, que se encuentra entre la secuencia que separa a los palíndromes propuestos, conserva hasta un 29 por ciento de actividad en comparación con la silvestre. Por otro lado, la mutante sencilla Bj74C, que se encuentra dentro de la secuencia palindrómica distal, tiene una transición de T por C en la posición -74, ésta conserva la misma actividad de una cepa silvestre. Estos datos respaldan la propuesta de la existencia e importancia de los palíndromes en el reconocimiento de la región -68 por RegR, ya que la transversión de T por G en la misma posición -74 anula completamente la expresión (cepa Bj74G). Por otro lado, mutaciones en las secuencias intermedias de los palíndromes tienen efectos moderados en comparación con aquéllas que caen dentro de ellos. Estos datos sugieren que la integridad de los palíndromes y posiciones contiguas a estos son esenciales para la interacción de RegR a la región -68. Dando lugar a la expresión del sistema regulador que controla la actividad transcripcional del operón *fixRnifA* por medio del promotor *fixRp₂* (Tabla I).

ACTIVIDAD DE B-GALACTOSIDASA (UNIDADES MILLER)
DE LAS MUTANTES EN LA REGION -68

CEPA	-75 ↓ T G C G * * * *	-68 ↓ G A C A T T A G G A * * * *	-56 ↓ C G C A * * *	AERO	ANAER
Bj 74-70	G	A		146	140
Bj 74-67	G	A		9	5
Bj 76-74	A G			3	4
Bj 72-70		A A		30	52
Bj 59-56			A C	22	31
Bj 59-57			A A	17	17
Bj 57-56			A C	16	22
Bj 70		A		47	32
Bj 56			C	4	6
Bj 63			C	514	670
Bj 74G	G			4	13
Bj 74C	C			1851	1112
Bj Wt 30		Wt-68	MUT -24	1789	1170
Bj Wt 25		Wt-68	Wt -24	1676	2208

Tabla I. Actividades de β -galactosidasa medida de 3 experimentos independientes. Los asteriscos representan la posiciones mutadas.

Hasta ahora se han establecido ejemplos donde un mayor número de motivos funcionales UAS eleva los niveles de transcripción a partir de un promotor -24/-12 (3). Lo mismo ocurre en relación al consenso de dichos motivos, donde la cercanía de la secuencia al consenso determina un nivel más alto de transcripción (referencias en la cita 33).

En la tabla II se muestran los efectos de las mutaciones en los motivos UAS del promotor *fixRnifA*. Estas mutaciones se realizaron sobre un molde mutante en la posición -68, lo que permite monitorear exclusivamente la actividad del sistema NifA dependiente. Estos motivos UAS son imperfectos y presentan las secuencias TGT N₁₀ CCA y TGT N₁₀ ACC a 100 y 121 pares de bases respectivamente por arriba del inicio de transcripción, en lugar de la secuencia consenso TGT N₁₀ ACA. Las mutantes creadas de las UAS₁, UAS₂ y de ambas UAS_{1/2} contienen transversiones de G por T. En los tres casos la expresión de la fusión *fixR-lacZ* en condiciones aeróbicas y microaeróbicas fue nula. Pensamos que la carencia de secuencias consenso en estos motivos, impide el reconocimiento de NifA individualmente por alguno de estos motivos. Por lo tanto es necesaria la unión de NifA en ambos motivos al mismo tiempo para activar la transcripción del operón *fixRnifA* a partir del promotor *fixRp₁*, trabajando de manera cooperativa. Casos como este se han reportado para la proteína NtrC (102).

Los datos obtenidos a partir de la mutagénesis realizada tanto en la región -68 como en los motivos UASs, dan evidencias adicionales de la complejidad existente del mecanismo de regulación transcripcional bajo el cual se encuentra el operón *fixRnifA* de *B.japonicum*. La descripción de la posible funcionalidad de los palíndromes descritos, puede llegar a utilizarse para detectar más genes blanco y correlacionarlos con la posible señal a la cual responde este sistema de dos componentes RegS-RegR.

ACTIVIDAD DE B-GALACTOSIDASA (UNIDADES MILLER)
DE LAS MUTANTES EN LA UASs

CEPA	T G T *	N10	A C C	N5	T G T *	N10	C C A	AERO	ANAER
Bj mut. UAS1	T							4	12
Bj mut. UAS2					T			5	9
Bj mut. UAS1/2	T				T			6	17
Bj Wt 25		Wt -68				Wt -24		1676	2208
Bj Wt 30		Wt -68				MUT -24		1789	1170
Bj 8091		MUT -68				Wt -24		6	1078

Tabla II. Actividades de 3 experimentos independientes, los asteriscos representan los sitios mutados.

APENDICE IV

Barrios H, Valderrama B and Morett E. (1999). Compilation and analysis of σ_{54} dependent promoter sequences. (Sometido en Nucleic Acids Research)

Compilation and analysis of σ^{54} -dependent promoter sequences

HUMBERTO BARRIOS, BRENDA VALDERRAMA AND ENRIQUE MORETT*

Departamento de Reconocimiento Molecular y Bioestructura, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62271, México.

* Corresponding author. Mailing address : Instituto de Biotecnología, Universidad Nacional Autónoma de México, A. Postal 510-3, Cuernavaca Morelos 62271, México. Phone: 52 (7) 3291605. Fax: 52 (7) 3172388. Email: emorett@ibt.unam.mx.

ABSTRACT

Promoters recognized by the RNA-polymerase with the alternative sigma factor σ^{54} ($E\sigma^{54}$) are unique in having conserved positions around -24 and -12 nucleotides upstream from the transcriptional start site, instead of the typical -35 and -10 boxes. Here we compile 186 -24/-12 promoter sequences reported in the literature and generate an updated and extended consensus sequence. The use of the extended consensus increases the probability of identifying genuine -24/-12 promoters. The effect of several reported mutations at the -24/-12 elements on RNA-polymerase binding and promoter strength is discussed in the light of the updated consensus.

INTRODUCTION

The upstream regulatory region of all bacterial genes or operons contains the promoter, that is the DNA sequence which determines specific recognition by the RNA polymerase to initiate transcription. The eubacterial RNA polymerase is a heterotetrameric enzyme constituted by one β , one β' and two α subunits (1). This "core" enzyme interacts with the initiation factor s to form the transcriptionally active holoenzyme ($E\sigma$). The addition of the σ factor confers upon core the specificity for the recognition of promoter sequences.

All known s factors belong to two different families: those evolutionarily related to the *E. coli* housekeeping factor σ^{70} , and those related to the alternative factor σ^{54} (2). Each family of sigma factors have different promoter sequence recognition, isomerisation and regulation properties. $E\sigma^{70}$ does not form stable closed-promoter complexes, therefore transcription can be initiated spontaneously in the absence of activator proteins (13). In contrast, $E\sigma^{54}$ forms stable closed-promoter complexes but it is unable to initiate transcription spontaneously. This polymerase is absolutely dependent on additional transcriptional factors, denominated Enhancer Binding Proteins (EBP) to initiate RNA synthesis. $E\sigma^{54}$ controls several ancillary processes including the degradation of xylene and toluene, dicarboxylic acids transport, pillin synthesis, nitrogen fixation, hydrogen biosynthesis (reviewed in 3-5), flagellar assembly (6), arginine catabolism (7), alginate production (8), rhamnolipid production (9), acetoin catabolism (10), mannose uptake (11) and proline iminopeptidase activity (12).

The basic promoters recognized by the $E\sigma^{70}$ family, although diverse in sequence, are normally conformed by two hexamers centered between -10 and -35 nucleotides upstream from the transcriptional start site (14). In contrast, $E\sigma^{54}$ recognizes promoter sequences with conserved GG and GC elements located -24 and -12 nucleotides upstream from the transcriptional start site +1, that is one DNA helical turn closer than the -35/-10 promoters (15). The aim of the present work was the compilation and analysis of the reported -24/-12 promoter sequences, either experimentally confirmed or identified by sequences similarity, and the discussion of possible regulatory implications. With the information obtained we refined the consensus sequence of σ^{54} -dependent promoters and carried out a comparative analysis of the effect of reported mutants of some of these promoters.

THE -24/-12 PROMOTER CONSENSUS SEQUENCE

We followed three criteria to identify putative -24/-12 promoters: a) mapped transcriptional start site with or without additional genetic evidence for the promoter; b) genetic evidences (mutation or heterologous gene expression); and c) putative promoters reported in the literature on the basis of sequence similarity to the -24/-12 elements. A collection of 186 sequences satisfying any of the three selection criteria is shown in Table 1. The nucleotide sequences were obtained from GenBank through the National Institute of Health server (<http://www3.ncbi.nlm.nih.gov/entrez/>). The sequence of each promoter including 10 nucleotides upstream and 10 nucleotides downstream from the highly conserved -24/-12 elements is shown in Table 1. For 85 of these promoters their transcriptional start site has been experimentally determined. A list of identified EBPs activating each promoter or the presence of putative EBPs binding sites are included in Table 2 for the cases reported. The relative nucleotide frequency at each position and the consensus sequence for each subgroup and for the whole collection is shown in Figure 1.

The consensus sequences shown in Figure 1 are depicted following a previously reported definition (14), where any nucleotide occurring with a frequency of more than 6 standard deviations from the expected random occurrence of each nucleotide (0.25) is denominated highly conserved (upper case), between 3 and 6 standard deviations is denominated weakly conserved (lower case), and below 3 standard deviations is not significant (N). As previously reported, the highest conservation was found around the -24 and the -12 elements (15). Around the -24 element, from position -31 to -20, there are eight highly conserved and three weakly conserved nucleotides with the sequence mrNrYTGGCACG (see Fig.1). Around the -12 element (positions -15 to -8), there are 5 highly conserved and one weakly conserved nucleotides with the sequence TTGCWNNw. Thus, the mrNrYTGGCACG-N4-TTGCWNNw sequence conform the updated consensus for the -24/-12 promoters.

The subgroup with mapped transcriptional start site presented the extended consensus sequence but not the subgroup without mapped transcriptional start site (compare Fig. 1B and 1C). This discrepancy probably reflects the inclusion of incorrectly assigned -24/-12 promoters sequences in the compilation of the latter

group. When the consensus sequence is displayed as the most frequent dinucleotide (Fig. 1D), it becomes apparent that at some positions one type of dinucleotide is strongly favoured. For example, at position -14 T is by far the most abundant nucleotide (85%) but the presence of a pyrimidine is almost 100% conserved. Similarly, at positions -20 and -21 R and Y are more than 90% conserved.

A comparison of the consensus sequence derived by group and the relative occurrence of the most common nucleotide at each position are shown in Table 3. In the subgroup with mapped transcriptional start site, the G at the -24 position is 100% conserved; the G at the -25 position is 99% conserved, with the exception of *Myxococcus xanthus* gene 4521 promoter (16,17); the G at the -13 position is 96% conserved with the exception of the *Pseudomonas aeruginosa oprE* (18), the *Escherichia coli glnH* and the *Neisseria gonorrhoeae pip* (12,19) promoters; while the C at the -12 position is 96% conserved with the exceptions of the *Alcaligenes eutrophus hoxK* (20), the *E. coli pspA* (21), the *Rhizobium leguminosarum* biovar *viciae nifH* (22), and the *Rhodospirillum rubrum glnB* (23) promoters.

CONTACTS BETWEEN σ^{54} AND THE PROMOTER IN CLOSED COMPLEXES AND RELEVANCE OF EACH POSITION.

Comparison of the deduced protein sequences of the σ^{54} factors showed that they are highly conserved (24): the amino terminus (region I) is comprised by 50 amino acids rich in glutamine and leucine residues that have been shown to be involved in activation (4). This region is followed by a non-conserved stretch of 27 to 110 amino acids (region II) with a high content of acidic residues. The rest of the protein is a long conserved region of approximately 400 residues (region III) involved in closed complex formation, having determinants for core interaction and promoter recognition, specifically a well characterized helix-turn-helix and a highly conserved sequence of 10 amino acids known as RpoN-box. These motifs are involved in the recognition of the -24 and -12 promoter elements, respectively (25-27).

The promoter residues contacted by $E\sigma^{54}$ have been extensively described (28-31). From different chemical and enzymatic protection assays it has been shown that σ^{54} binds specifically around the -24 and -12 promoter elements in the absence of core polymerase (31). Different reactivities have been observed along the

promoter sequence, both on the DNA major groove and the backbone, but always facing the same side of the double helix. These experiments demonstrate that σ^{54} contacts actually extend from positions -31 to -5, although only for the GG and GC dinucleotides at positions -24 and -12 specific protein-DNA interaction have been detected (30). Further addition of the core polymerase induced a distortion downstream from the -12 element but the reactivity towards methylating reagents did not change, indicating that the core subunits do not contact the promoter directly but through σ^{54} (29,30,32). Changes of the highly conserved -24/-12 dinucleotides prevent the binding of σ^{54} to the promoter (31,33), while increasing the similarity to the consensus upstream from the -12 position enhances the stability of the closed complex and therefore the binding affinity for σ^{54} (28,30,31).

Mutations reported elsewhere helped to identify positions critical for the promoter function (16,27,34-38). Table 4 shows several reported mutation in -24/-12 promoters and a quantitative analysis of their effect. All the changes in the highly conserved positions -25/-24 and -13/-12 are down-mutations, reducing drastically the expression, except for the *Myxococcus xanthus* 4521 gene promoter which has an A at -25 (Table 4, lane E) (39). Qualitatively, similar changes reported in the *Azorhizobium caulinodans nifA* (40), the *Caulobacter crescentus flbG* and *flaN* (41-43) and the *Klebsiella pneumoniae nifH* (44) promoters, also presented a down-phenotype. Mutations in the *K. pneumoniae nifH* promoter at the -26 and -16 positions decreased the expression by 25% and at the -15 position by 75% (35,37), while simultaneous mutations at the -17, -16 and -15 positions decreased the expression by 55% (Table 4, lane C) (35). The substitution of the same residues for three Ts enhanced the σ^{54} and E σ^{54} binding affinities both *in vivo* and *in vitro* (31,40,45). Nevertheless it has been shown that this T-rich stretch could be slightly deleterious for the activity of other promoters (35).

Mutations around the -24 region that increased the similarity to the consensus, for example, a change of G for A at the -22 position in the *K. pneumoniae nifL* promoter, enhanced the expression level more than two-fold (47). Moreover, mutations at positions -13 or -12 that reduced promoter strength could be compensated by other recognition enhancing mutations that brought the promoter sequence closer to the consensus (46). However, It has been suggested that a very strong contact between the promoter and σ^{54} could prevent the formation of new transcriptional complexes (48).

TRANSCRIPTIONAL START SITES IN -24/-12 PROMOTERS

σ^{70} -dependent promoters generally initiate transcription at a purine, being adenine more frequently utilised than guanine (14). The selection of this nucleotide is influenced by the sequence around -35 and by the compositions of the -2 to -5 positions (49,50). In our collection of 85 -24/-12 promoters with mapped transcriptional start site a significant proportion of them (64%) also initiate transcription at a purine (see Fig. 2). In σ^{70} -dependent promoters the spacing between the first nucleotide of the -10 element and the transcriptional start site is usually 6 or 7 nucleotides, although functional examples between 4 and 10 nucleotides have been reported (14). In contrast, -24/-12 promoters start transcription generally between 11 to 13 nucleotides downstream from the -12 position (GC) (Fig. 2). The shortest distance reported is 8 nucleotides for the *Parasponia rhizobium nifH* (51), the *Rhodospirillum centenum cheAY* (52) and the *Rhodospirillum rubrum nifj* (53) promoters. The longest distance reported is 21 nucleotides for the *Pseudomonas aeruginosa algD* promoter (54). At present, there are no reported examples of transcriptional starts between 17 and 20 nucleotides. Based on these results it is not possible to accurately predict a transcriptional start site for -24/-12 promoters based solely on the promoter sequence.

SPACING BETWEEN THE -24 AND -12 CONSERVED ELEMENTS

The optimal spacing between the -35 and -10 elements in σ^{70} -dependent promoters is 17 +/- 1 nucleotides, but functional promoters with spacing between 15-20 nonconserved nucleotides have been reported (14,55). In contrast, deletions of one or more nucleotides in the stretch between the -24/-12 elements abolished promoter function (16,36,41-43). Thus, a stringent requirement for these motifs to be positioned on the same face of the DNA helix seems to be a necessary condition for the binding of Es⁵⁴. Although there are no published examples of nucleotide insertions between the -24/-12 elements, it is reasonable to predict that any insertion might severely disrupt promoter function.

CONCLUSION

Here we present an updated compilation of σ^{54} -dependent promoters and the derivation of an extended consensus sequence. The new consensus extends from positions -8 to -31 relative to the transcriptional start site (Fig. 1 and Table 3). Interestingly, the observed contacts of σ^{54} with the promoter DNA span exactly this length (30). Promoter sequences from 47 organisms are represented in our study. However, only for 33 of them the *rpoN* gene has been identified and sequenced.

Experimental data show that the promoter binding affinity for E σ^{54} is independent of the actual activity observed *in vivo* (46). It has been proposed that the -24 element functions as an attachment determinant for E σ^{54} while the -12 element is involved in the fine modulation of an already established closed-complex towards its isomerization (46). Individual changes around the -12 element that brought the sequence closer to the consensus enhanced the binding affinity of the promoter but did not necessarily increase the transcriptional activity (31,35). Furthermore, Gralla and coworkers showed for the *glnAp2* promoter that single changes that deviate the sequences from the consensus increased the transcriptional activity, as long as the rest of the element sequence remains consensus (46). Thus, it seems that there should be a balance between the E σ^{54} binding affinity and the promoter strength.

The first group of -24/-12 promoters being characterized initiate transcription precisely 12 nucleotides downstream from the conserved GC element, this coincidence lead to the operational designation of these promoters as -24/-12 (15). Our updated compilation of 85 promoters with mapped transcriptional start site from 28 different species provides evidences that the initial nucleotide selection is more flexible than initially considered. In Figure 2 we present these data as a Gaussian distribution similar to that obtained from the analysis of σ^{70} -dependent promoters from *B. subtilis* and *E. coli* (56,57), with the significant difference that instead of the preferential use of a single position, the use of nucleotides 11, 12 and 13 downstream from the GC is roughly equally frequent (24%, 20% and 27% respectively). A few examples of transcripts starting three nucleotides before and after positions 11 to 13 have been also described and were included in the analysis. Nevertheless we consider that the only case reported of a transcript starting 21 nucleotides downstream from the GC seems to be wrongly designated as dependent

of σ^{54} , although it was included in this compilation on the basis of satisfying our selection criteria (54,58),

The consensus sequence derived from our collection of 186 -24/-12 promoters from 47 different bacterial species shows a remarkable conservation both in sequence and in structure in contrast to the flexible consensus derived from s^{70} -dependent promoters. This conservation symbolizes the strict requirements for promoter recognition and function required for a highly controlled regulation.

ACKNOWLEDGEMENTS

We wish to thank Ricardo Grande and Mariana Peimbert for their comments on the manuscript. HB is recipient of a Consejo Nacional de Ciencia y Tecnología scholarship.

Figure legends

Figure 1. Relative frequency of each nucleotide per position and consensus sequences of the -24/-12 promoters. A) Entire collection of the 186 -24/-12 promoters; B) the 85 promoters with mapped transcriptional start site; C) the 101 promoters without mapped transcriptional start site; D) dinucleotide frequency composition of the 85 promoters with mapped transcriptional start site. R stands for purines; Y for pyrimidines; W is A or T. Poisson statistics were used to express the square standard deviation for each nucleotide as $np\sqrt{1-px}$, where n is the number of bases at that position and px is the occurrence frequency for each nucleotide (56).

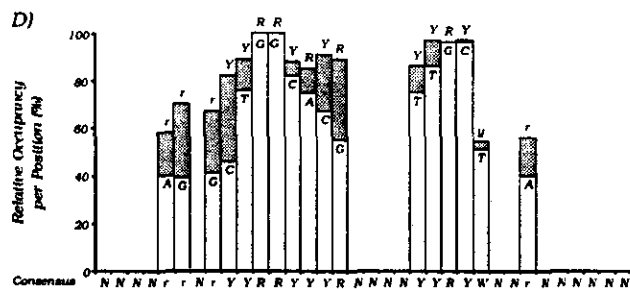
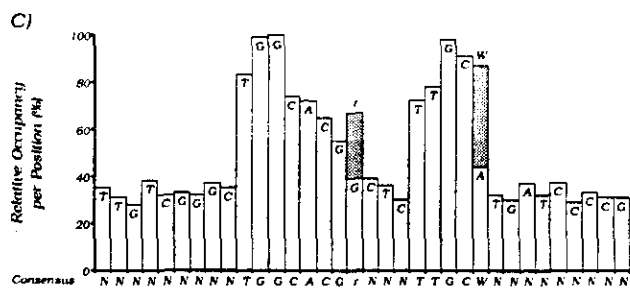
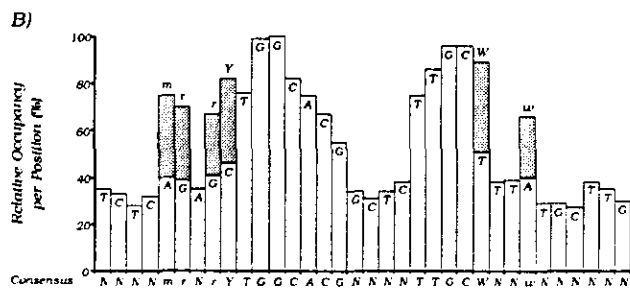
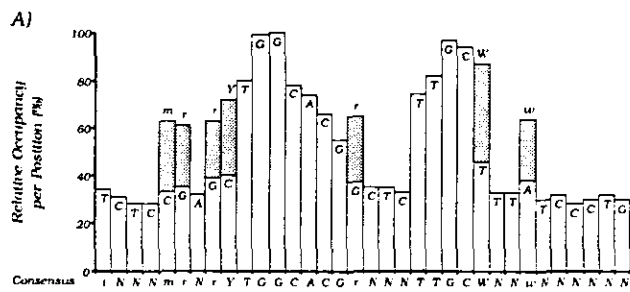
Figure 2. Distribution of transcriptional start sites for the 85 -24/-12 promoters. The reference position is the conserved C of the -12 element.

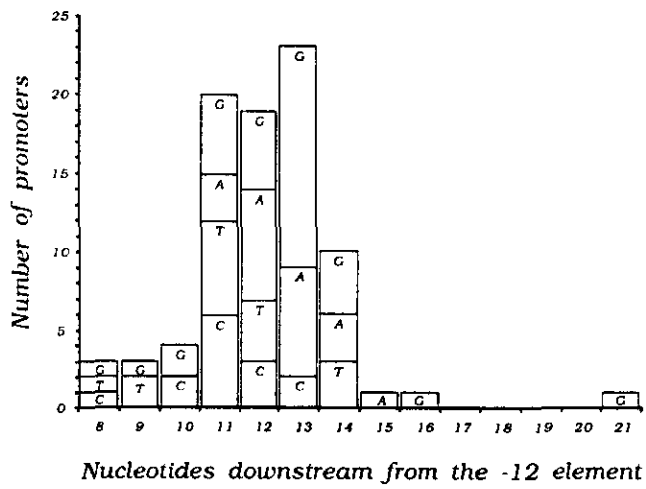
Table 1. Nucleotide sequence of the 186 -24/-12 promoters used in this study. The first column is the gene name and the second column the organism name abbreviated as follows: Al.den = *Alcaligenes denitrificans*, Al.eut, *Alcaligenes eutrophus*; Al.fae, *Alcaligenes faecalis*; Az.bra, *Azospirillum brasilense*; Az.cau, *Azorhizobium caulinodans*; Az.vin, *Azotobacter vinelandii*; Az.chr, *Azotobacter chroococcum*; Bc.sub, *Bacillus subtilis*; Br.jap, *Bradyrhizobium japonicum*; Cw.rhi, *Cowpea Rhizobium*; Cm.col, *Campylobacter coli*; Cl.cre, *Caulobacter crescentus*; Er.chr, *Erwinia chysanthemi*; Es.col, *Escherichia coli*; Hr.ser, *Herbaspirillum seropedicae*; En.clo, *Enterobacter cloacae*; En.agg, *Enterobacter agglomerans*; Kl.aer, *Klebsiella aerogenes*; Kl.pne, *Klebsiella pneumoniae*; Kl.oxy, *Klebsiella oxytoca*; Mb.the, *Methanobacterium thermoautotrophicum*; Mc.cap, *Methylococcus capsulatus*; Mx.xan, *Myxococcus xanthus*; Ns.gon, *Neisseria gonorrhoeae*; Pr.rhi, *Parasponia Rhizobium*; Pt.vul, *Proteus vulgaris*; Ps.aer, *Pseudomonas aeruginosa*, Ps.put, *Pseudomonas putida*, Ps.syr, *Pseudomonas syringae*; Rh.leg, *Rhizobium leguminosarum*; b.pha, *Rhizobium leguminosarum* biovar *phaseoli*; b.tri, *Rhizobium leguminosarum* biovar *trifolii*; b.vic, *Rhizobium leguminosarum* biovar *viciae*; Rh.pha, *Rhizobium phaseoli*; Rh.mel, *Sinorhizobium meliloti*; Rh.sp, *Rhizobium* sp; Rh.BT, *Rhizobium* BTAi1; Rh.tri, *Rhizobium trifolii*; Rh.pNG, *Rhizobium* plasmid NGR2; Rb.cap, *Rhodobacter capsulatus*; Rs.rub, *Rhodospirillum rubrum*; Rs.cen, *Rhodospirillum centenum*; Th.fer, *Thiobacillus ferrooxidans*; Sl.typ, *Salmonella typhimurium*; Vb.alg, *Vibrio alginolyticus*; and Vb.cho, *Vibrio cholerae*. TSS means transcriptional start site.

Table 2. Regulatory features derived from the analysis of the 186 -24/-12 promoters used in this study. Organism names are as in Table 1. TSSD is distance from the transcription start site to the C of the conserved -12 element; EBP means experimentally identified activator of the EBP family; Bin.site means possible EBPs binding site identified by sequence similarity; Distance is the distance from the second G of the -24 element to the EBP binding site.

Table 3. Comparison of -24/-12 promoter consensus sequences; A) Consensus sequence from the 186 compiled promoters. B) Consensus sequence from 85 promoters with mapped transcriptional start site. C) Consensus sequence from 101 promoters without mapped transcriptional start site. D) Consensus sequence obtained from the comparison of groups A, B and C. Numbers under the sequences represent the relative occurrence of the nucleotide at that position (percent)

Table 4. Effect of some mutations in -24/-12 promoters. Relative values are shown in brackets: A) Consensus sequence from the -24/-12 promoters with mapped transcriptional start site derived from this analysis. B) Mutants of the *K. pneumoniae nifL* promoter (34,37). C) Mutants of the *K. pneumoniae nifH* promoter (35-37), D) Mutants of the *B. subtilis levD* promoter (38). E) Mutants of the *M. xanthus* 4521 gene promoter (16). F) Mutants of the *K. pneumoniae glnA* promoter (27).





Organ.	Gene	TSSD	KBP	Bin.site	Distance	Reference	Organ.	Gene	TSSD	KBP	Bin.site	Distance	Reference
Al.den	azu	-	-	INF	49	(59)	Kl.pne	nifB	11pb	-	INF, NifA	91	(132, 145)
Al.out	hoxF	14pb	-	INF, HoxA	40, 143	(61, 62)		nifK	11pb	-	INF, NifA	3, 33	(133, 200)
	hoxK	13pb	-	-	-	(20)		nifF	11pb	-	NifA	-	(60)
	uraDI	-	-	WtrC	163	(XKY1373)		nifH	13pb	-	INF, NifA	9, 96	(63, 64)
	acoX	13pb	-	-	-	(10)		nifU	10pb	-	INF, NifA	14, 107, 153	(64)
	acoX	14pb	-	-	-	(68)		nifL	11pb	-	WtrC	137	(65, 66)
Al.fae	nifH	-	-	NifA	100	(69)		nifM	11pb	-	-	-	(67)
As.bra	glnB	11pb	-	-	96	(71)		nifE	-	-	-	-	(67)
	glsE	-	-	INF, NifA	31, 174	(73)		nifU	11pb	-	INF, NifA	2, 76, 85	(70)
	ORF2	-	-	NifA	76	(75)	Kl.oxy	nifL	13pb	-	-	-	(72)
	nifW	-	-	-	-	(76)		nifB	-	-	NifA	91	(72, 74)
	nifB	-	-	NifA	64, 105	(78)		nifT	-	-	-	-	(72)
	nifH	11pb	-	NifA	70, 79	(80)	Mo.the	nifH	-	-	-	-	(77)
As.cau	fixA	-	-	NifA	-	(82)	Mc.cap	glnA	-	-	WtrC	41	(79)
	glnB	13pb	-	-	-	(83)	Mc.man	abhA	14pb	-	-	-	(17, 81)
	nifO	-	-	-	-	(82)		4521	14pb	-	-	-	(16, 17)
	nifA	13pb	-	-	-	(84)		pilA	16pb	-	-	-	(17)
As.vin	fixA	-	-	-	-	(X65523)	Ms.gou	pip	13pb	-	-	-	(12)
	nifB	11pb	-	NifA	95	(85)	Pr.rhi	nifD	12pb	-	NifA	69	(51)
	nifE	-	-	NifA	63	(87, 88)		nifH	8pb	-	-	-	(51)
	nifF	-	-	NifA, NifA	113	(90)	Pr.vul	glnA	-	-	-	-	(86)
	nifH	-	-	INF, NifA	87	(88)	Ps.aer	algC	11pb	AlgR	AlgR	61, 146	(8, 89)
	nifH2	-	-	-	-	(92)		oprK	10pb	-	-	-	(18)
	nifM	-	-	-	-	(88)		PhhA	-	PhhR	PhhR	51	(91)
	nifU	-	-	-	-	(88)		cpG2	14pb	-	-	-	(93, 96)
	anfH	11pb	-	INF, AnfA	63, 212	(94, 95)		asu	-	-	-	-	(59)
	ORF11	-	-	NifA	45	(88)		nirA	-	-	-	-	M97294
	ORF10	-	-	NifA	81	(88)		rhIAB	14pb	RhIR	RhIR	13, 151	(9, 96)
	ORF12	-	-	NifA	106	(88)		algD	21pb	AlgR	INF	33, 391, 468	(54, 58)
	ORF5	-	-	NifA	96	(88)		flaSR	13pb	FlaQ	INF, NifA	5, 42	(6)
	ORF6	-	-	NifA	94	(88)	Ps.put	ompK	-	-	INF	28	(97)
	ORF8	-	-	NifA	87	(88)		xy1A	13pb	-	INF, XylR	33, 100, 131	(98-100)
	vnfEXX	-	-	-	-	(103)		xy1S	11pb	-	INF, XylR	6, 113, 143	(100-102)
	anfA	-	-	-	-	(88)	Ps.syr	hrpL	-	-	-	-	136536
	vnfA	-	-	-	-	(88)		avrD	-	-	-	-	(104)
As.cbr	nifH	11pb	-	INF, NifA	19, 100	(107)		pelE	-	-	-	-	(105)
	vnfD	-	-	-	-	(109)	Rh.log	fixA	-	-	-	-	(106)
	vnfH	-	-	-	-	(110)		orf240	-	-	-	-	(108)
	nifH	-	-	-	-	(114)		dctA	-	DctD	DctD	75, 108	(106)
	nifM	-	-	-	-	(116)		fixZ	-	-	-	-	(111, 113)
	nifU	-	-	-	-	(116)		glnA	-	-	-	-	(115)
Bc.sub	levD	14pb	LevR	INF, LevR	41, 80	(38, 119)		glnXI	12pb	-	-	-	(117)
	rocD	14pb	RocR	RocR	69, 110	(7)		fixA	-	-	NifA	67	(120)
	rocA	14pb	RocR	INF, RocR	18, 88, 132	(7, 121)	b.pha	fixA	-	-	-	-	(122)
Er.fad	glnXI	11pb	-	WtrC	78	(123)	b.tri	nifB	-	-	-	-	X16311
	fixB	11pb	-	-	-	(124)	b.vic	glnB	13pb	-	-	-	(122)
	fixR	12pb	NifA	RegR, NifA	43, 60, 81	(125, 126)		nifB2	-	-	-	-	70
	glnB	12pb	-	WtrC	-	(127)		nifH1	-	-	NifA	91	X11084
	nifB	13pb	-	-	-	(123)		nifH	13pb	-	NifA	85	(22)
	nifD	13pb	-	NifA	69	(128, 129)	Rh.pha	nifHa	-	-	NifA	90	(130)
	nifH	10pb	-	INF, NifA	71, 100	(131, 132)		nifHc	-	-	-	-	(130)
	nifS	11pb	-	NifA	52	(133)	Rh.mel	D-ala	-	-	-	-	(134)
	ndp	13pb	-	INF, NifA	46	(135)		glnXI	12pb	-	WtrC	73	(136)
	groKSL3	13pb	-	-	-	(137)		dctB	-	-	-	-	(112)
	fixA	12pb	-	-	-	(138)		dctA	-	-	-	-	(111, 139)
	hupS	14pb	-	-	-	(140)		fixA	13pb	-	NifA	77	(141)
Cv.rhi	nifD	15pb	-	NifA	71	(142)		fixD	-	-	-	-	(63)
	nifH	11pb	-	-	-	(142)		glnB	-	-	-	-	050385
Cm.col	flgH	9pb	-	-	-	(143)		nosB	-	-	-	-	023753
	flaB	13pb	-	-	-	(144)		nifH	13pb	-	-	-	(63)
Cl.cre	flaG	9pb	-	-	-	(41)		nifB	-	-	NifA	41, 61	(145)
	flgL	-	-	FlhD	82	(146, 147)		nifJ	12pb	-	NifA	34	(148, 149)
	flhP	13pb	-	-	-	(150)		nifJ2	12pb	-	NifA	88	(148, 149)
	flhW	13pb	-	-	-	(151)		nifM	-	-	NifA	88	(152)
	flhG	-	-	FlhD	75	(41, 43, 153)		PI	12pb	-	NifA	81	(154)
	flgP	12pb	FlhD	INF	21, 51	(153, 155, 156)	Rh.sp	nifH	-	-	NifA	67	(158)
	flaM	10pb	FlhD	INF, ftr	34	(41, 43)	Rh.SP	nifH	11pb	-	-	-	(160)
	flgZ	-	-	FlhD	16, 37, 71	(153, 156, 157)	Rh.tri	nifH	-	-	NifA	88	(161)
	pillin	11pb	-	-	-	(159)	Rh.pSU	nifQ	-	-	-	-	(163)
	flgW	13pb	FlhD	ftr	88	(153, 156)		fixA	-	-	NifA	74	(163)
Er.cbr	celY	-	-	-	-	(162)		fixB	-	-	NifA	74	(163)
	hrpW	-	-	-	-	(164)	Mo.cap	nifB	-	-	-	-	(168)
Es.col	argT	-	-	-	-	(168)		nifE	-	-	-	-	(170)
	fdhF	12pb	-	PhlA	84	(166, 167)		nifH	13pb	-	INF, NifA	49	(132, 173)
	fixA	-	-	-	-	(169)		ORF6	-	-	-	-	(175)
	hycA	12pb	-	INF, PhlA	6, 36	(171, 172)		nifU	-	-	NifA	41	(178)
	hycB	12pb	-	PhlA	-	(172, 176)		nifP	-	-	NifA	-	(180)
	glnA	12pb	-	WtrC	76, 100	(3, 176, 177)		fdhD	13pb	-	-	-	(173)
	nac	12pb	-	WtrC, WtrC	42, 118	(179)		anfH	-	-	-	-	(183)
	glnM	13pb	-	INF, WtrC	74, 86	(19, 181)	Ms.rab	glnB	13pb	-	WtrC	79	(23)
	popA	12pb	PopP	INF, PopP	2, 61, 84	(21, 182)		nifH	-	-	INF, NifA	90, 121	(132, 186)
Er.sar	nifA	-	-	NifA	96	(184)		nifU	9pb	-	-	-	(53)
	nifH	-	-	NifA	92, 112	(188)	Ms.com	chaAY	9pb	-	-	-	(52)
	nifB	-	-	NifA	35, 97	(184)	Sh.fer	nifH	-	-	INF, NifA	74, 119	(189-191)
Er.clo	nifLa	-	-	-	-	(187)		acrB	-	-	WtrC	115	(189)
Er.gsf	nifL	-	-	WtrC	163	(188)	Sl.typ	argT	-	-	-	-	(193)
	nifB	-	-	NifA	93	(188)		dhaA	-	-	WtrC	103	(192, 194)
	nifE	-	-	NifA	94	(192)		glnA	9pb	-	WtrC	74	(195)
	nifH	-	-	-	-	(X99694)		prpB	-	-	-	-	(196)
	nifZ	-	-	-	-	(X99694)	Vb.alg	glnA	-	-	-	-	(149)
	nifY	-	-	NifA	107	(192)	Vb.cho	fixB	-	-	-	-	(198)
Kl.aer	nac	13pb	-	WtrC, WtrC	42, 119	(179, 197)		fixA	-	-	-	-	(199)
Kl.pne	glnA	12pb	-	WtrC	74	(166)		rfaD	-	-	-	-	(201)

COMPARISON OF σ^{54} -DEPENDENT PROMOTER SEQUENCES

-25 -24

-13 -12

A)	N	N	N	N	N	N	N	N	N	N	T	G	G	C	A	C	G	r	N	N	N	T	T	G	C	W	N	N	N	N	N	N	N	N	N
											83	99	100	74	72	65	55	67							73	79	99	92	87						
B)	N	N	N	N	m	r	N	r	Y	T	G	G	C	A	C	G	N	N	N	N	N	T	T	G	C	W	N	N	w	N	N	N	N	N	
					75	70		67	82	76	99	100	82	75	67	55									75	86	96	96	89		66				
C)	i	N	N	N	m	r	N	r	Y	T	G	G	C	A	C	G	r	N	N	N	N	T	T	G	C	W	N	N	w	N	N	N	N	N	y
	34				63	61		63	72	80	99	100	78	74	66	55	64								74	82	97	94	87		64				62
D)	N	N	N	N	m	r	N	r	Y	T	G	G	C	A	C	G	N	N	N	N	N	T	T	G	C	W	N	N	w	N	N	N	N	N	

	-25	-24	-23	-12																	
A)	T	G	G	C	A	C	G	N	N	N	T	T	G	C	T/A	N	N	A	N	N	N
B) <i>Kp-ni/L</i>	G	G	C	C	G	C	A	C	G	G	T	T	G	C	A	T	G	G	T	F	A
	A(3)	A(4)	T(118)	A(261)	T(116)	T(112)	A(83)	A(93)					A(3)	T(13)							
C) <i>Kp-ni/H</i>	T	G	G	T	A	T	G	T	T	C	C	C	T	G	C	A	C	T	T	C	T
	C(75)							A(102)	T(109)	T(78)	C(8)	A(1)	T(7)								
D) <i>Bs-levD</i>	T	G	G	C	A	C	G	A	T	C	C	T	T	G	C	A	T	T	A	T	A
	C(1)	A(1)	T(2)			G(17)						A(2)	A(1)								C(18)
E) <i>Mx-4521</i>	G	A	G	C	A	C	G	C	G	T	C	T	T	G	C	T	T	T	G	G	C
	A(99)	T(76)	T(4)	T(9)	A(9)	T(13)	T(128)	C(76)	T(122)					T(2)	C(29)	A(147)	T(135)	A(116)	T(81)	C(62)	
F) <i>Kp-glnA</i>	T	G	G	C	A	C	A	G	A	T	T	T	C	G	C	T	T	T	A	T	A
	A(13)	A(10)												A(6)	A(12)						
	C(13)	C(14)												C(14)	G(12)						
	T(11)	T(9)												T(26)	T(14)						

REFERENCES

1. Helmann, J. D. & Chamberlin, M. (1988) *Annu. Rev. Biochem.* **57**, 839-872.
2. Gross, C.A., Loneto, M. and Losick, R. (1992) Bacterial sigma factors. In: *Transcriptional regulation*, edited by McKnight, S.L. and Yamamoto, K.R. pp 129-176. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
3. Kustu, S., Santero, E., Keener, J., Popham, D. & Weiss, D. (1989) *Microbiol. Rev.* **53**, 367-376.
4. Merrick, M. (1993) *Mol. Microbiol.* **10**, 903-909.
5. Fischer, H. M. (1994) *Microbiol. Rev.* **58**, 352-386.
6. Arora, S. K., Ritchings, B. W., Almira, E. C, Lory, S. & Ramphal, R. (1997) *J. Bacteriol.* **179**, 5574-5581.
7. Gardan, R., Rapoport, G. & Débarbouillé, M. (1995) *J. Mol. Biol.* **249**, 843-856.
8. Zielinski, N. A., Maharaj, R., Roychoudhury, S., Danganan, C. E., Hendrickson, W. & Chakrabarty, A. M. (1992) *J. Bacteriol.* **174**, 7680-7688.
9. Pearson, J. P., Pesci, E. C. & Iglewski, B. H. (1997) *J. Bacteriol.* **179**, 5756-5767.
10. Priefert, H., Krüger, N., Jendrossek, D., Schmidt, B. . & Steinbüchel, A. (1992) *J. Bacteriol.* **174**, 899-907.
11. Martin-Verstraete, I., Stulke, J., Klier, A. & Rapoport, G. (1995) *J. Bacteriol.* **177**, 6919-6927.
12. Albertson, N. H. & Koomey, M. (1993) *Mol. Microbiol.* **9**, 1203-1211.
13. Gralla, J. D. (1990) *Meth. Enzymol.* **185**, 37-54.
14. Hawley, D. K. & McClure, W. R. (1983) *Nucleic Acids Res.* **11**, 2237-2255.
15. Thöny, B. & Hennecke, H. (1989) *FEMS Microbiol. Rev.* **63**, 341-358.
16. Keseler, I. M. & Kaiser, D. (1995) *J. Bacteriol.* **177**, 4638-4644.
17. Wu, S. S. & Kaiser, D. (1997) *J. Bacteriol.* **179**, 7748-7758.
18. Yamano, Y., Nishikawa, T. & Komatsu, Y. (1998) *FEMS Microbiol. Lett.* **162**, 31-37.
19. Nohno, T. & Saito, T. (1987) *Nucleic Acids Res.* **15**, 2777.
20. Schwartz, E., Gerischer, U. & Friedrich, B. (1998) *J. Bacteriol.* **180**, 3197-3204.
21. Jovanovic, G. & Model, P. (1997) *Mol. Microbiol.* **25**, 473-481.
22. Roelvink, P. W., Harmsen, M., van Kammen, A. & van den Bos, R. C. (1990) *Gene* **87**, 31-36.
23. Johansson, M. & Nordlund, S. (1996) *Microbiology* **142**, 1265-1272.
24. Morett, E. & Segovia, L. (1993) *J. Bacteriol.* **175**, 6067-6074.

25. Coppard, J. R. & Merrick, M. J. (1991) *Mol. Microbiol.* **5**, 1309-1317.
26. Taylor, M., Butler, R., Chambers, S., Casimiro, M., Badii, F. & Merrick, M. (1996) *Mol. Microbiol.* **22**, 1045-1054.
27. Merrick, M. & Chambers, S. (1992) *J. Bacteriol.* **174**, 7221-7226.
28. Buck, M. & Cannon, W. (1994) *Nucleic Acids Res.* **22**, 1119-1120.
29. Morris, L., Cannon, W., Claverie-Martin, F., Austin, S. & Buck, M. (1994) *J. Biol. Chem.* **269**, 11563-11571.
30. Cannon, W., Austin, S., Moore, M. & Buck, M. (1995) *Nucleic Acids Res.* **23**, 351-356.
31. Buck, M. & Cannon, W. (1992) *Nature* **358**, 422-424.
32. Wang, X. Y., Kolb, A., Cannon, W. & Buck, M. (1997) *Nucleic Acids Res.* **25**, 3478-3485.
33. Brown, S. E. & Ausbel, F. M. (1984) *J. Bacteriol.* **157**, 143-147.
34. Khan, H., Buck, M. & Dixon, R. (1986) *Gene* **45**, 281-288.
35. Buck, M. & Cannon, W. (1989) *Nucleic Acids Res.* **17**, 2597-2612.
36. Buck, M. (1986) *J. Bacteriol.* **166**, 545-551.
37. Buck, M., Khan, H. & Dixon, R. (1985) *Nucleic Acids Res.* **13**, 7621-7638.
38. Martin-Verstraete, I., Debarbouille, M., Klier, A. & Rapoport, G. (1992) *J. Mol. Biol.* **226**, 85-99.
39. Keseler, I. M. & Kaiser, D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 1979-1984.
40. Stiger, J., Schneider, M. & de Bruijn, F. J. (1993) *Mol. Plant Microbe Interac.* **6**, 238-252.
41. Mullin, D. A. & Newton, A. (1989) *J. Bacteriol.* **171**, 3218-3227.
42. Mullin, D. A. & Newton, A. (1993) *J. Bacteriol.* **175**, 2067-2076.
43. Ninfa, A. J., Mullin, D. A., Ramakrishnan, G. & Newton, A. (1989) *J. Bacteriol.* **171**, 383-391.
44. Ow, D. W., Xiong, Y., Qu, Q. & Shen, S. -C. (1985) *J. Bacteriol.* **161**, 868-874.
45. Morett, E. & Buck, M. (1989) *J. Mol. Biol.* **210**, 65-77.
46. Wang, L. & Gralla, J. D. (1998) *J. Bacteriol.* **180**, 5626-5631.
47. Garcia, E., Bancroft, S., Rhee, S. G. & Kustu, S. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1662-1666.
48. Tintut, Y., Wang, J. T. & Gralla, J. D. (1995) *Genes & Devel.* **9**, 2305-2313.
49. Fredrick, K. & Helmann, J. D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 4982-4987.
50. Jeong, W. & Kang, C. (1994) *Nucleic Acids Res.* **22**, 4667-4672.

51. Weinman, J. J., Fellows, F. F., Gresshoff, P. M., Shine, J. & Scott, K. F. (1984) *Nucleic Acids Res.* **12**, 8329-8344.
52. Jiang, Z-Y. & Bauer, C. E. (1997) *J. Bacteriol.* **179**, 5712-5719.
53. Lindblad, A., Jansson, J., Brostedt, E., Johansson, M., Hellman, U. & Nordlund, S. (1996) *Mol. Microbiol.* **20**, 559-568.
54. Wozniak, D. J. (1994) *J. Bacteriol.* **176**, 5068-5076.
55. de Haseth, P. L. & Helmann, J. D. (1995) *Mol. Microbiol.* **16**, 817-824.
56. Harley, C. B. & Reynolds, R. P. (1987) *Nucleic Acids Res.* **15**, 2343-2361.
57. Helmann, J. D. (1995) *Nucleic Acids Res.* **13**, 2351-2360.
58. Kato, J. & Chakrabarty, A. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1760-1764.
59. Hoitink, C. W., Woudt, L. P., Turenhout, J. C., van de Kamp, M. & Canters, G. M. (1990) *Gene* **90**, 15-20.
60. Arnold, W., Rump, A., Klipp, W., Prierer, U.B. and Pühler, A. (1988) In: *Nitrogen fixation: Hundred years after*, eds Bothe, H., de Bruijn, F.J. and Newton. p. 303. Gustav Fischer, Stuttgart.
61. Tran-Betcke, A., Warnecke, U., Böcker, C., Zaborosch, C. & Friedrich, B. (1990) *J. Bacteriol.* **172**, 2920-2929.
62. Zimmer, D., Schwartz, E., Tran-Betcke, A., Gewinner, P. & Friedrich, B. (1995) *J. Bacteriol.* **177**, 2373-2380.
63. Weber, G., Reiländer, H. & Pühler, A. (1985) *EMBO J.* **4**, 2751-2756.
64. Charlton, W., Cannon, W. & Buck, M. (1993) *Mol. Microbiol.* **7**, 1007-1021.
65. Ow, D. W., Sundaresan, V., Rothstein, D. M., Brown, S. E. & Ausubel, F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2524-2528.
66. Minchin, S. D., Austin, S. & Dixon, R. A. (1988) *Mol. Microbiol.* **2**, 433-442.
67. Paul, W. & Merrick, M. (1987) *Eur. J. Biochem.* **170**, 259-265.
68. Priefert, H., Hein, S., Kruger, N., Zeh, K., Schmidt, B. & Steinbuchel, A. (1991) *J. Bacteriol.* **173**, 4056-4071.
69. Hai, W., Zheng, H., You, C. & Wang, B. (1992) *I. Chuan. Hsueh. Pao* **19**, 369-377.
70. Cannon, W. V., Kreutzer, R., Kent, H. M., Morett, E. & Buck, M. (1990) *Nucleic Acids Res.* **18**, 1693-1701.
71. de Zamaroczy, M., Delorme, F. & Elmerich, C. (1990) *Mol. Gen. Genet.* **224**, 421-430.
72. Kim, Y. M., Ahn, K. J., Beppu, T. & Uozumi, T. (1986) *Mol. Gen. Genet.* **205**, 253-259.

73. Mandal, A.K. & Ghosh, S. (1993) *J. Bacteriol.* **175**, 8024-8029.
74. Kim, Y. M., Hidaka, M., Masaki, H., Beppu, T. & Uozumi, T. (1991) *Agric. Biol. Chem.* **55**, 3121-3123.
75. Frazzon, J. & Schrank, I. S. (1998) *FEMS Microbiol. Lett.* **15**, 151-158.
76. Milcamps, A., Keyers, V. & Vanderleyden, J. (1993) *Biochim. Biophys. Acta* **28**, 237-238.
77. Hochheimer, A., Schmitz, R. A., Thauer, R. K. & Hedderich, R. (1995) *Eur. J. Biochem.* **235**, 910-920.
78. Liang, Y. Y., de Zamaroczy, M., Arsene, F., Paquelin, A. & Elmerich, C. (1992) *FEMS Microbiol. Lett.* **79**, 113-119.
79. Cardy, D. L. N. & Murrell, J. C. (1990) *Journal of General Microbiology.* **136**, 343-352.
80. de Zamaroczy, M., Delorme, F. & Elmerich, C. (1989) *Mol. Gen. Genet.* **220**, 88-94.
81. Romero, J. M. & Zusman, D. R. (1991) *J. Bacteriol.* **173**, 2969-2976.
82. Kaminski, P. A., Norel, F., Desnoues, N., Kush, A., Salzano, G. & Elmerich, C. (1988) *Mol. Gen. Genet.* **214**, 496-502.
83. Michel-Reydellet, N., Desnoues, N., Elmerich, C. & Kaminski, P. A. (1997) *J. Bacteriol.* **179**, 3580-3587.
84. Loroch, A. I., Nguyen, B. G. & Ludwig, R. A. (1995) *J. Bacteriol.* **177**, 7210-7221.
85. Joerger, R. D. & Bishop, P. E. (1988) *J. Bacteriol.* **170**, 1475-1487.
86. Steglitz-Morsdorf, U., Morsdorf, G. & Kaltwasser, H. (1993) *FEMS Microbiol. Lett.* **106**, 157-164.
87. Brigle, K. E. & Dean, D. R. (1988) *Nucleic Acids Res.* **16**, 5214.
88. Joerger, R. D., Jacobson, M. R. & Bishop, P. E. (1989) *J. Bacteriol.* **171**, 3258-3267.
89. Fujiwara, S., Zielinski, N. A. & Chakrabarty, M. (1993) *J. Bacteriol.* **175**, 5452-5459.
90. Bennett, L. T., Jacobson, M. R. & Dean, D. R. (1988) *J. Biol. Chem.* **263**, 1364-1369.
91. Song, J. & Jensen, R. A. (1996) *Mol. Microbiol.* **22**, 497-507.
92. Raina, R., Reddy, M. A., Ghosal, D. & Das, H. K. (1988) *Mol. Gen. Genet.* **214**, 121-127.
93. Minton, N. P., Atkinson, T., Bruton, C. J. & Sherwood, R. F. (1984) *Gene* **31**, 31-38.
94. Austin, S. & Lambert, J. (1994) *J. Biol. Chem.* **269**, 18141-18148.
95. Joerger, R. D., Jacobson, M. R., Premakumar, R., Wolfinger, E. D. & Bishop, P. E. (1989) *J. Bacteriol.* **171**, 1075-1086.
96. Ochsner, U. A., Fiechter, Armin. & Reiser, Jakob. (1994) *J. Biol. Chem.* **269**,

19787-19795.

97. Nordlund, I., Powlowski, J. & Shingler, V. (1990) *J. Bacteriol.* **172**, 6826-6833.
98. Inouye, S., Nakazawa, A. & Nakazawa, T. (1985) *J. Bacteriol.* **163**, 863-869.
99. Pérez-Martín, J. & de Lorenzo, V. (1996) *Cell* **86**, 331-339.
100. Pérez-Martín, J. & de Lorenzo, V. (1996) *J. Mol. Biol.* **258**, 562-574.
101. Bertoni, G., Pérez-Martín, J. & de Lorenzo, V. (1997) *Mol. Microbiol.* **23**, 1221-1227.
102. Gallegos, M. T., Marqués, S. & Ramos, J. L. (1996) *J. Bacteriol.* **178**, 2356-2361.
103. Wolfinger, E. D. & Bishop, P. E. (1991) *J. Bacteriol.* **173**, 7565-7572.
104. Yucel, I., Boyd, C., Debnam, Q. & Keen, N. T. (1994) *Mol. Plant Microbe Interac.* **7**, 131-139.
105. Bauer, D. W. & Collmer, A. (1997) *Mol. Plant Microbe Interac.* **10**, 369-379.
106. Hontelez, J. G. J., Lankhorst, R. K., Katinakis, P., van den Bos, R. C. & van Kammen, A. (1989) *Mol. Gen. Genet.* **218**, 536-544.
107. Jone, R., Woodley, P., Birkmann-Zinoni, A. & Robson, R. L. (1993) *Gene* **123**, 145-146.
108. Colonna-Romano, S., Arnold, W., Schlüter, A., Boistard, P., Pühler, A. & Priefer, U. B. (1990) *Mol. Gen. Genet.* **223**, 138-147.
109. Robson, R. L., Woodley, P. R., Pau, R. N. & Eady, R. R. (1989) *EMBO J.* **8**, 1217-1224.
110. Fallik, E. & Robson, R. L. (1990) *Nucleic Acids Res.* **18**, 4616.
111. Engelke, T., Jording, D., Kapp, D. & Pühler, A. (1989) *J. Bacteriol.* **171**, 5551-5560.
112. Jiang, J., Gu, B. H., Albright, L. M. & Nixon, B. T. (1989) *J. Bacteriol.* **171**, 5244-5253.
113. Lederbur, H., Gu, B., Sojda, J. III & Nixon, B. T. (1990) *J. Bacteriol.* **172**, 3888-3889.
114. Robson, R. L., Woodley, P. R. & Jones, R. (1986) *EMBO J.* **5**, 1159-1163.
115. Rossen, L., Ma, Q. -S., Mudd, E. A., Johnston, A. W. B. & Downie, J. A. (1984) *Nucleic Acids Res.* **12**, 7123-7134.
116. Evans, D. J., Jones, R., Woodley, P. R., Wilborn, J. R. & Robson, R. L. (1991) *J. Bacteriol.* **173**, 5457-5469.
117. Colonna-Romano, S., Riccio, A., Guida, M., et al. (1987) *Nucleic Acids Res.* **15**, 1951-1964.
118. Patriarca, E. J., Chiurazzi, M., Manco, G., et al. (1992) *Mol. Gen. Genet.* **234**, 337-345.

119. Martin-Verstraete, I., Débarbouille, M., Klier, A. & Rapoport, G. (1994) *J. Mol. Biol.* **241**, 178-192.
120. Michiels, J. & Vanderleyden, J. (1993) *Biochem. Biophys. Acta.* **13**, 232-233.
121. Calogero, S., Gardan, R., Glaser, P., Schweizer, J., Rapoport, G. & Debarbouille, M. (1994) *J. Bacteriol.* **176**, 1234-1241.
122. Chiurazzi, M. & Iaccarino, M. (1990) *Mol. Microbiol.* **4**, 1727-1735.
123. Carlson, T. A., Martin, G. B. & Chelm, B. K. (1987) *J. Bacteriol.* **169**, 5861-5866.
124. Gubler, M. & Hennecke, H. (1988) *J. Bacteriol.* **170**, 1205-1214.
125. Barrios, H., Fischer, H. M., Hennecke, H. & Morett, E. (1995) *J. Bacteriol.* **177**, 1760-1765.
126. Barrios, H., Grande, R., Olvera, L. & Morett, E. (1998) *Proc. Natl. Acad. Sci. USA* **95**,
127. Martin, G. B., Thomashow, M. F. & Chelm, B. K. (1989) *J. Bacteriol.* **171**, 5638-5645.
128. Kaluza, K. & Hennecke, H. (1984) *Mol. Gen. Genet.* **196**, 35-42.(Abstract)
129. Alvarez-Morales, A., Betancourt-alvarez, M., Kaluza, K. & Hennecke, H. (1986) *Nucleic Acids Res.* **14**, 4207-4227.
130. Quinto, C., de la Vega, H., Flores, M., et al. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1170-1174.
131. Fuhrmann, M. & Hennecke, H. (1984) *J. Bacteriol.* **158**, 1005-1011.
132. Hoover, T. R., Santero, E., Porter, S. & Kustu, S. (1990) *Cell* **63**, 11-22.
133. Ebeling, S., Noti, J. D. & Hennecke, H. (1987) *Nucleic Acids Res.* **15**, 9598.
134. Leong, S. A., Williams, P. H. & Ditta, G. S. (1985) *Nucleic Acids Res.* **13**, 5965-5976.
135. Weidenhaupt, M., Fischer, H. M., Acuña, G., Sanjuan, J. & Hennecke, H. (1993) *Gene* **129**, 33-40.
136. Shatters, R. G., Somerville, J. E. & Kahn, M. L. (1989) *J. Bacteriol.* **171**, 5087-5094.
137. Fischer, H. M., Babst, M., Kaspar, T., Acuña, G., Arigoni, F. & Hennecke, H. (1993) *EMBO J.* **12**, 2901-2912.
138. Fuhrmann, M., Fischer, H. M. & Hennecke, H. (1985) *Mol. Gen. Genet.* **199**, 315-322.
139. Watson, R. J. (1990) *Mol. Plant Microbe Interac.* **3**, 174-181.
140. Black, L. K. & Maier, R. J. (1995) *Mol. Microbiol.* **16**, 405-413.
141. Earl, C. D., Ronson, C. W. & Ausubel, F. M. (1987) *J. Bacteriol.* **169**, 1127-1136.

142. Yun, A. C. & Szalay, A. A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7358-7362.
143. Kinsella, N., Guerry, P., Cooney, J. & Trust, T. J. (1997) *J. Bacteriol.* **179**, 4647-4653.
144. Alm, R. A., Guerry, P. & Trust, T. J. (1993) *J. Bacteriol.* **175**, 4448-4455.
145. Buikema, W. J., Klingensmith, J. A., Gibbons, S. L. & Ausubel, F. M. (1987) *J. Bacteriol.* **169**, 1120-1126.
146. Schoenlein, P. V., Gallman, L. S., Winkler, M. E. & Ely, B. (1990) *Gene* **93**, 17-25.
147. Gober, J. W. & Shapiro, L. (1990) *Genes & Devel.* **4**, 1494-1504.
148. Soto, M. J., Zorzano, A., Mercado-Blanco, J., Lepek, V., Olivares, J. & Toro, N. (1993) *J. Mol. Biol.* **229**, 570-576.
149. Sanjuan, J. & Olivares, J. (1991) *Arch. Microbiol.* **155**, 543-548.
150. Ramakrishnan, G., Zhao, J. L. & Newton, A. (1991) *J. Bacteriol.* **173**, 7283-7292.
151. Dingwall, A., Gober, J. W. & Shapiro, L. (1990) *J. Bacteriol.* **172**, 6066-6076.
152. Aguilar, O. M., Reiländer, H., Arnold, W. & Pühler, A. (1987) *J. Bacteriol.* **169**, 5393-5400.
153. Wu, J., Benson, A. K. & Newton, A. (1995) *J. Bacteriol.* **177**, 3241-3250.
154. Better, M., Lewis, B., Corbin, D., Ditta, G. & Helinski, D. R. (1983) *Cell* **35**, 479-485.
155. Marques, M. V. & Gober, J. W. (1995) *Mol. Microbiol.* **16**, 279-289.
156. Dingwall, A., Garman, J. D. & Shapiro, L. (1992) *J. Mol. Biol.* **228**, 1147-1162.
157. Khambaty, F. M. & Ely, B. (1992) *J. Bacteriol.* **174**, 4101-4109.
158. Badenoch-Jones, J., Holton, T. A., Morrison, C. M., Scott, K. F. & Shine, J. (1989) *Gene* **77**, 141-153.
159. Minnich, S. A. & Newton, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1142-1146.
160. Legocki, R. P., Yun, A. C. & Szalay, A. A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5806-5810.
161. Scott, K. F., Rolfe, B. G. & Shine, J. (1983) *DNA* **2**, 149-155.
162. Guiseppi, A., Aymeric, J. L., Cami, B., Barras, F. & Creuzet, N. (1991) *Gene* **106**, 109-114.
163. Freiberg, C., Perret, X., Broughton, W. & Rosenthal, A. (1996) *Genome. Res.* **6**, 590-600.
164. Mukherjee, A., Cui, Y., Liu, Y. & Chatterjee, A. K. (1997) *Mol. Plant Microbe Interac.* **10**, 462-471.
165. Nonet, M. L., Marvel, C. C. & Tola, D. R. (1998) *J. Biol. Chem.* **5**, 12209-12217.

166. Dixon, R. (1984) *Nucleic Acids Res.* 12, 7811-7830.
167. Birkmann, A. & Bock, A. (1989) *Mol. Microbiol.* 3, 187-195.
168. Masepohl, B., Klipp, W. & Puhler, A. (1988) *Mol. Gen. Genet.* 212, 27-37.
169. Eichler, K., Buchet, A., Bourgis, F., Kleber, H. P. & Mandrand-Berthelot, M. A. (1995) *J. Basic. Microbiol.* 35, 217-227.
170. Moreno-Vivian, C., Schmehl, M., Masepohl, B., Arnold, W. & Klipp, W. (1989) *Mol. Gen. Genet.* 216, 353-363.
171. Lutz, S., Bohm, R., Beier, A. & Bock, A. (1990) *Mol. Microbiol.* 4, 13-20.
172. Hopper, S., Babst, M., Schlensong, V., Fischer, H-M., Hennecke, H. & Böck, A. (1994) *J. Biol. Chem.* 269, 19597-19604.
173. Willison, J. C., Pierrard, J. & Hübner, P. (1993) *Gene* 133, 39-46.
174. Lutz, S., Jacobi, A., Schlensong, V., Böhm, R., Sawers, G. & Böck, A. (1991) *Mol. Microbiol.* 5, 123-135.
175. Masepohl, B., Angermüller, S., Hennecke, S., Hübner, P., Moreno-Vivian, C. & Klipp, W. (1993) *Mol. Gen. Genet.* 238, 369-382.
176. Reitzer, L. J. & Magasanik, B. (1986) *Cell* 45, 785-792.
177. Ninfa, A. J., Reitzer, L. J. & Magasanik, B. (1987) *Cell* 50, 1039-1046.
178. Cullen, P. J., Foster-Hartnett, D., Gabbert, K. K. & Kranz, R. G. (1994) *Mol. Microbiol.* 11, 51-65.
179. Muse, W. B. & Bender, R. A. (1998) *J. Bacteriol.* 180, 1166-1173.
180. Gennaro, G., Hubner, P., Sandmeier, U., Yakunin, A. F. & Hallenbeck, P. C. (1996) *J. Bacteriol.* 178, 3949-3952.
181. Claverie-Martin, F. & Magasanik, B. (1991) *Proc. Natl. Acad. Sci. USA* 88, 1631-1635.
182. Weiner, L., Brissette, J. L., Ramani, N. & Model, P. (1995) *Nucleic Acids Res.* 23, 2030-2036.
183. Schuddekopf, K., Hennecke, S., Liese, U., Kutsche, M. & Klipp, W. (1993) *Mol. Microbiol.* 8, 673-684.
184. Souza, E. M., Funayama, S., Rigo, L. U., Yates, M. G. & Pedrosa, F. O. (1991) *Journal of General Microbiology.* 137, 1511-1522.
185. Machado, I. M., Yates, M. G., Machado, H. B., Souza, E. M. & Pedrosa, F. O. (1996) *Braz. J. Med. Biol. Res.* 29, 1599-1602.
186. Lehman, L. J., Fitzmaurice, W. P. & Roberts, G. P. (1990) *Gene* 95, 143-147.
187. Deng, X. & Shen, S. (1995) *SCI. CHINA. B.* 38, 60-66.

188. Siddavattam, D., Steibl, H. D., Kreutzer, R. & Klingmüller, W. (1995) *Mol. Gen. Genet.* **249**, 629-636.
189. Kilkenny, C. A., Berger, D. K. & Rawlings, D. E. (1994) *Microbiology* **140**, 2543-2553.
190. Berger, D. K., Woods, D. R. & Rawlings, D. E. (1990) *J. Bacteriol.* **172**, 4399-4406.
191. Pretorius, I. M., Rawlings, D. E., O'Neill, E. G., Jones, W. A., Kirby, R. & Woods, D. R. (1987) *J. Bacteriol.* **169**, 367-370.
192. Kreutzer, R., Singh, M. & Klingmüller, W. (1989) *Gene* **78**, 101-109.
193. Higgins, C. F. & Ames, G. F. L. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1083-1087.
194. Ferro-Luzzi, G. & Nikaido, K. (1998) *EMBO J.* **4**, 539-547.
195. Klose, K. E. & Mekalanos, J. J. (1997) *Infect. Immun* **65**, 587-596.
196. Horswill, A. R. & Escalante-Semerena, J. C. (1997) *J. Bacteriol.* **179**, 928-940.
197. Schwacha, A. & Bender, R. A. (1993) *J. Bacteriol.* **175**, 2107-2115.
198. Klose, K. E. & Mekalanos, J. J. (1998) *Mol. Microbiol.* **28**, 501-520.
199. Klose, K. E. & Mekalanos, J. J. (1998) *J. Bacteriol.* **180**, 303-316.
200. Cannon, W., Charlton, W. & Buck, M. (1991) *J. Mol. Biol.* **220**, 915-931.
201. Stroehler, U. H., Karageorgos, L. E., Morona, R. & Manning, P. A. (1995) *Gene* **155**, 67-72.

Discusión y conclusiones del apéndice IV

Los promotores reconocidos por la RNA-pol- σ^{54} presentan un alto grado de conservación, tanto en su estructura nucleotídica como en su mecanismo transcripcional ampliamente distribuido en varios generos bacterianos. Esta bien documentado que los genes que se encuentran regulados por este tipo de sistemas presentan funciones metabólicas alternativas para la célula y que llegan a ser muy variables. El objetivo de este trabajo fue actualizar el conocimiento que se tiene de los promotores dependientes de σ^{54} , mediante la recopilación y analisis de la secuencias nucleotídicas de este tipo de promotores reportados.

Los 186 promotores recopilados dependientes σ^{54} se agruparon para su analisis en aquellos con inicio de transcripción mapeado experimentalmente y aquellos que carecian de esta evidencia. Tomando en cuenta que una de sus características mas comunes es la conservación de los dinucleotidos GG y GC en las regiones -24/-12, realizamos un alineamiento de las secuencias promotoras en base a estos dos elementos. La comparación de las secuencias consenso obtenidas de ambos grupos y de la totalidad de las secuencias, nos ha llevado a la descripción de una secuencia consenso extendida para los promotores dependientes de σ^{54} (mrNrYTGGCACG-N₄ -TTGCAWNNw) (7,96).

Cualquiera que sea la contribución real de estos nuevos nucleotidos en este consenso extendido, esta tendra que ser elucidada experimentalmente. Sin embargo, consideramos que estadisticamente la aparición de las nuevas posiciones en este consenso extendido son significativas. La relevancia de estas posiciones se puede encontrar en el fino proceso del reconocimiento y en la capacidad de iniciar de la transcripción por la RNA-pol- σ^{54} , jugando un papel adicional en el mecanismo de regulación transcripcional.

Otra de las características muy peculiar de este tipo de promotores es la distancia que hay entre los dos elementos conservados -24 y -12, la cual consta de 10 nucleótidos. El aumento o disminución de un nucleótido en esta distancia da como resultado promotores no funcionales. En contraste los promotores reconocidos por σ^{70} pueden llegar a tener una distancia variable de 15 a 20 nucleótidos no conservados entre los elementos -35 y -10. La posible rigidez en la distancia de los promotores dependientes σ^{54} , puede estar dada por la necesidad de tener a los elementos -24/-12 posicionados en la misma cara de la doble hélice de DNA como una condición básica para el reconocimiento del promotor por la E- σ^{54} .

La distancia que existe entre el elemento -12 al inicio de la transcripción es variable y más del 70% de los promotores con inicio de transcripción mapeado presentan distancias entre 10 y 12 nucleótidos. Actualmente se sabe que los determinantes en el reconocimiento y apertura del DNA se encuentran en el factor sigma, se ha sugerido que la variabilidad en la distancia para iniciar la transcripción depende evidentemente de la secuencia promotora sin embargo hasta ahora no hay evidencia directa sobre esta idea.

El análisis de los datos existentes sobre la interacción de E- σ^{54} y σ^{54} sobre la región promotora, muestran que los nucleótidos más conservados en este tipo de promotores son las posiciones utilizadas para ser contactadas. Mutaciones sobre esta región muestran que mientras más cercano al consenso se encuentran las secuencias promotoras, tienen la capacidad de formar complejos más estables con E- σ^{54} . Sin embargo mientras más estables son estos complejos su actividad transcripcional disminuye. Se propone que esto es debido a que la E- σ^{54} se une tan establemente al promotor que es incapaz de iniciar un nuevo ciclo de transcripción.

Dentro de las características distintivas de los promotores σ^{54} dependientes, se encuentra la secuencia consenso mrNrYTGGCACG-N₄-TTGCAWNNw, la distancia única de 10 pares de bases entre los motivos -12/-24 y su mecanismo de transcripción altamente conservado en este tipo de promotores.

FALTA PAGINA

No.

55

REFERENCIAS

1. Agron, P.G. and D.R. Helinski. 1995. Symbiotic expression of *Rhizobium meliloti* nitrogen fixation genes is regulated by oxygen p. 275-287. In J.A. Hoch and T.J. Silhavy (ed.), Two-component signal transduction. ASM, Washington, D.C..
2. Albright, L. M., E. Huala, and F. M. Ausubel. 1989. Prokaryotic signal transduction mediated by sensor and regulator protein pairs. *Ann. Rev. Genet.* 23:311-336.
3. Alvarez-Morales, A., M. Betancourt-Alvarez, K. Kaluza, and H. Hennecke. 1986. Activation of the *Bradyrhizobium japonicum* *nifH* and *nifDK* operons is dependent on promoter-upstream DNA sequences. *Nucleic Acids Res.* 14:4207-4227.
4. Anthamatten, D. and H. Hennecke. 1991. The regulatory status of the *fixL*-like and *fixJ*-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*. *Mol. Gen. Genet.* 225:38-48.
5. Austin, S. and R. Dixon. 1992. The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent. *EMBO J* 11:2219-2228.
6. Barrios, H., H. M. Fischer, H. Hennecke, and E. Morett. 1995. Overlapping promoters for two different RNA polymerase holoenzymes control *Bradyrhizobium japonicum* *nifA* expression. *J. Bacteriol.* 177:1760-1765.
7. Barrios, H., B. Valderrama, and E. Morett. 1999. Extended compilation and analysis of σ^{54} -dependent promoters. Sin publicar
8. Bauer, E., T. Kaspar, H. M. Fischer, and H. Hennecke. 1998. Expression of the *fixR-nifA* operon in *Bradyrhizobium japonicum* depends on a new response regulator, RegR. *J. Bacteriol.* 180:3853-3863.
9. Beck, C., R. Marty, S. Kläusli, H. Hennecke, and M. Göttfert. 1997. Dissection of the transcription machinery for housekeeping genes of *Bradyrhizobium japonicum*. *J. Bacteriol.* 179:364-369.
10. Bertoni, G., N. Fujita, A. Ishihama, and V. de Lorenzo. 1998. Active recruitment of σ^{54} -RNA polymerase to the *Pu* promoter of *Pseudomonas putida*: role of IHF and α CTD. *EMBO J.* 17(17):5120-5128.

11. Buck, M. 1986. Deletion analysis of the *Klebsiella pneumoniae* nitrogenase promoter: importance of spacing between conserved sequences around positions -12 and -24 for activation by the *nifA* and *ntrC* (*glnG*) products. *J. Bacteriol.* 166:545-551.
12. Buck, M. and W. Cannon. 1992. Specific binding of the transcription factor sigma-54 to promoter DNA. *Nature* 358:422-424.
13. Buck, M., W. Cannon, and J. Woodcock. 1987. Transcriptional activation of the *Klebsiella pneumoniae* nitrogenase promoter may involve DNA loop formation. *Mol. Microbiol.* 1:243-249.
14. Buck, M., S. Miller, M. Drummond, and R. Dixon. 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature* 320:374-378.
15. Burgess, R.R., B. Erickson, D. Gentry, M. Gribskov, D. Hager, S. Lesley, M. Strickland, and N. Thompson. 1987. Bacterial RNA polymerase subunits and genes p. 3-15. In W. Reznikoff, R.R. Burgess, J.E. Dahl, C.A. Gross, M.T. Record, and M.P. Wickens (ed.), *RNA polymerase and the regulation of transcription*. Elsevier, New York.
16. Burton, Z., R. R. Burgess, J. Lin, D. Moore, S. Holder, and C. A. Gross. 1981. The nucleotide sequence of the cloned *rpoD* gene for the RNA polymerase sigma subunit from *E.coli* K-12. *Nucleic Acids Res.* 9:2889-2903.
17. Cannon, W., F. Claverie-Martin, S. Austin, and M. Buck. 1994. Identification of a DNA-contacting surface in the transcription factor sigma-54. *Mol. Microbiol.* 11:227-236.
18. Cannon, W. V., M. K. Chaney, X. -Y. Wang, and M. Buck. 1997. Two domains within σ_N (σ_{54}) cooperate for DNA binding. *Proc. Natl. Acad. Sci. USA* 94:5006-5011.
19. Cannon, W. V., R. Kreutzer, H. M. Kent, E. Morett, and M. Buck. 1990. Activation of the *Klebsiella pneumoniae nifU* promoter: identification of multiple and overlapping upstream NifA binding sites. *Nucleic Acids Res.* 18:1693-1701.
20. Daniels, D., P. Zuber, and R. Losick. 1990. Two amino acids in an RNA polymerase sigma factor involved in the recognition of adjacent base pairs in the -10 region of a cognate promoter. *Proc. Natl. Acad. Sci. USA* 87:8075-8079.

21. Darst, S. A., E. W. Kubalek, and R. D. Kornberg. 1989. Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. *Nature* 340:730-732.

22. de Haseth, P. L. and J. D. Helmann. 1995. Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. *Mol. Microbiol.* 16:817-824.

23. de Philip, P., E. Soupène, J. Batut, and P. Boistard. 1992. Modular structure of the FixL protein of *Rhizobium meliloti*. *Mol. Gen. Genet.* 235:49-54.

24. Doi, R.H. 1991. Regulation of gene expression p. 15-34. In U.N. Streaps and R.E. Yasbin (ed.), *Modern Microbial Genetics*. Wiley-Liss, Inc., USA.

25. Dombroski, A. J., W. A. Walter, M. T. Record, D. A. Siegele, and C. A. Gross. 1992. Polypeptides containing highly conserved regions of transcription initiation factor σ^{70} exhibit specificity of binding to promoter DNA. *Cell* 70:501-512.

26. Drummond, M. H., A. Contreras, and L. A. Mitchenall. 1990. The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of *Klebsiella pneumoniae*. *Mol. Microbiol.* 4:29-37.

27. Drzewiecki, K., C. Eymann, G. Mittenhuber, and M. Hecker. 1998. The *yvyD* gene of *Bacillus subtilis* is under dual control of σ^B and σ^H . *J. Bacteriol.* 180:6674-6680.

28. Estrem, S. T., T. Gaal, W. Ross, and R. L. Gourse. 1998. Identification of an UP element consensus sequences for bacterial promoters. *Proc. Natl. Acad. Sci. USA* 95:9761-9766.

29. Fischer, H. M. 1994. Genetic regulation of nitrogen fixation in Rhizobia. *Microbiol. Rev.* 58:352-386.

30. Fischer, H. M., T. Bruderer, and H. Hennecke. 1988. Essential and non-essential domains in the *Bradyrhizobium japonicum* NifA protein: identification of indispensable cysteine residues potentially involved in redox activity and/or metal binding. *Nucleic Acids Res.* 16:2207-2224.

31. Fujita, M., A. Amemura, and H. Aramaki. 1998. Transcription of the *groESL* operon in *Pseudomonas aeruginosa* PAO1. *FEMS Microbiol. Lett.* 163:237-242.

32. Gaal, T., W. Ross, E. E. Blatter, H. Tang, X. Jia, V. V. Krishnan, N. Assa-Munt, R. H. Ebright, and R. L. Gourse. 1996. DNA-binding determinants of the β subunit of RNA polymerase: novel DNA-binding domain architecture. *Genes & Devel.* 10:16-26.
33. Glubler, M. 1989. Fine-tuning of *nif* and *fix* gene expression by upstream activator sequences in *Bradyrhizobium japonicum*. *Mol. Microbiol.* 3:149-159.
34. Gonzalez, V., L. Olvera, X. Soberón, and E. Morett. 1998. In vivo studies on the positive control function of NifA: a conserved hydrophobic amino acid patch at the central domain involved in transcriptional activation. *Mol. Microbiol.* 28:55-67.
35. Gross, C.A., M. Loneto, and R. Losick. 1992. Bacterial sigma factors p. 129-176. In S.L. McKnight and K.R. Yamamoto (ed.), *Transcriptional regulation*. Cold Spring Harbor Laboratory Press, New York.
36. Harley, C. B. and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* 15:2343-2361.
37. Hawley, D. K. and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* 11:2237-2255.
38. Helmann, J. D. and M. Chamberlin. 1988. Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* 57:839-872.
39. Hoover, T. R., E. Santero, S. Porter, and S. Kustu. 1990. The integration host factor stimulates interaction of RNA polymerase with NifA, the transcriptional activator for nitrogen fixation operons. *Cell* 63:11-22.
40. Hsieh, M. and J. D. Gralla. 1994. Analysis of the N-terminal leucine heptad and hexad repeats of Sigma 54. *J. Mol. Biol.* 239:15-24.
41. Igarashi, K., N. Fujita, and A. Ishihama. 1992. Identification of a subunit assembly domain in the alpha subunits of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* 218:1-6.
42. Iismaa, I. and M. Watson. 1989. The *nifA* gene product from *R. leguminosarum* *bv. trifolii* lacks the N-terminal domain found in other NifA proteins. *Mol. Microbiol.* 3:943-945.
43. Ishihama, A. 1993. Protein-protein communications within the transcription apparatus. *J. Bacteriol.* 175:2483-2489.

44. Ishihama, A. and K. Ito. 1972. Subunits of RNA polymerase in function and structure.II. Reconstitution of *Escherichia coli* RNA polymerase from isolated subunits. J. Mol. Biol. 72:111-123.

45. Ishihama, A. and T. Saitoh. 1979. Subunids of RNA polymerase in fuction and structure. IX. Regulation of RNA polymerase activity by stringent starvation protein (SSP). J. Mol. Biol. 129:517-530.

46. Ishihama, A., T. Saitoh and R. Fukuda. 1976. Control of formation of RNA polymerase in *Escherichia coli*. Cold Spring Harbor Lab, Cold Spring Harbor, NY.

47. Ishihama, R. 1981. Subunit assembly of *Escherichia coli* RNA polymerase. Advan. Biophys 14:1-35.

48. Jeong, W. and C. Kang. 1994. Start site selection at *lacUV5* promoter affected by the sequence context around the initiation sites. Nucleic Acids Res. 22:4667-4672.

49. Jovanovic, G., L. Weiner, and P. Model. 1996. Identification, nucleotide sequence, and characterization of PspF, the transcriptional activator of the *Escherichia coli* stress-induced *psp* operon. J. Bacteriol. 178:1936-1945.

50. Kahn, D. and G. Ditta. 1991. Modular structure of FixJ:homology of the transcriptional activator domain with the -35 binding domain of sigma factors. Mol. Microbiol. 5:987-997.

51. Keener, J. and S. Kustu. 1988. Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: Roles of the conserved amino-terminal domain of NTRC. Proc. Natl. Acad. Sci. USA 85:4976-4980.

52. Keesler, I. M. and D. Kaiser. 1997. σ^{54} , a vital protein for *Myxococcus xanthus*. Proc. Natl. Acad. Sci. USA 94:1979-1984.

53. Kimura, M., N. Fujita, and A. Ishihama. 1994. Functional map of the alpha subunit of *Escherichia coli* RNA polymerase. J. Mol. Biol. 242:107-115.

54. Klose, K. E., A. K. North, K. M. Stedman, and S. Kustu. 1994. The mayor dimerization determinants of the nitrogen regulatory protein NTRC from enteric bacteria lie in its carboxy-terminal domain. J. Mol. Biol. 241:233-245.

55. Kolb, A., D. Kotlarz, S. Kusano, and A. Ishihama. 1995. Selectivity of the *Escherichia coli* RNA polymerase $E\sigma^{38}$ for overlapping promoters and ability to support CRP activation. *Nucleic Acids Res.* 23:819-826.
56. Kullik, I., S. Fritsche, H. Knobel, J. Sanjuan, H. Hennecke, and H. M. Fischer. 1991. *Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the σ_{54} gene (*rpoN*). *J. Bacteriol.* 173:1125-1138.
57. Kullik, I., H. Hennecke, and H. M. Fischer. 1989. Inhibition of *Bradyrhizobium japonicum* *nifA*-dependent *nif* gene activation by oxygen occurs at the NifA protein level and is irreversible. *Arch. Biochem.* 151:191-197.
58. Kustu, S., A. K. North, and D. S. Weiss. 1985. Prokaryotic transcriptional enhancers and enhancer binding proteins. *Trends Biochem. Sci.* 16:397-402.
59. Kustu, S., A. K. North, and D. S. Weiss. 1991. Prokaryotic Transcriptional Enhancers and Enhancer-Binding Proteins. *Trends. Biochem. Sci.* 16:397-402.
60. Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss. 1989. Expression of σ_{54} (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol. Rev.* 53:367-376.
61. Lesley, S. A. and R. R. Burgess. 1989. Characterization of the *Escherichia coli* Transcription factor σ^{70} : Localization of a region involved in the interaction with core RNA polymerase. *Biochemistry* 28:7728-7734.
62. Liu, X. and P. Matsumura. 1996. Differential regulation of multiple overlapping promoters in flagellar class II operons in *Escherichia coli*. *Mol. Microbiol.* 21:613-620.
63. Lois, A. F., G. S. Ditta, and D. R. Helinski. 1993. The oxygen sensor FixL of *Rhizobium meliloti* is a membrane protein containing four possible transmembrane segments. *J. Bacteriol.* 175:1103-1109.
64. Lois, A. F., M. Weinstein, G. S. Ditta, and D. R. Helinski. 1993. Autophosphorylation and phosphatase activities of the oxygen -sensing protein FixL of *Rhizobium meliloti* are coordinately regulated by oxygen. *J. Biol. Chem.* 268:4370-4375.
65. Lonetto, M., M. Gribskov, and C. A. Gross. 1992. The σ^{70} family: Sequence conservation and evolutionary relationship. *J. Bacteriol.* 174:3843-3849.

66. Malhotra, A., E. Severinova, and S. A. Darst. 1996. Crystal structure of a σ^{70} subunit fragment from *E.coli* RNA polymerase. *Cell* 87:127-136.
67. McClure, W. R. 1985. Mechanism and control of transcription in prokaryotes. *Annu. Rev. Biochem.* 54:171-204.
68. Merrick, M. 1993. In a class of its own - the RNA polymerase sigma factor σ^{54} (σ^N). *Mol. Microbiol.* 10:903-909.
69. Merrick, M.J. 1992. Regulation of nitrogen fixation genes in free-living and symbiotic bacteria p. 835-876. In G. Stacey, R.H. Burris, and H.J. Evans (ed.), *Biological Nitrogen Fixation*. Chapman and Hall, New York.
70. Monson, E. K., M. Weinstein, G. S. Ditta, and D. R. Helinski. 1992. The FixL protein of *Rhizobium meliloti* can be separated into a heme-binding oxygen-sensing domain and a functional C-terminal kinase domain. *Proc. Natl. Acad. Sci. USA* 89:4280-4284.
71. Morett, E. and M. Buck. 1988. NifA-dependent in vivo protection demonstrates that the upstream activator sequence of *nif* promoters is a protein binding site. *Proc. Natl. Acad. Sci. USA* 85:9401-9405.
72. Morett, E., W. Cannon, and M. Buck. 1988. The DNA-binding domain of the transcriptional activator protein NifA resides in its carboxy terminus, recognises the upstream activator sequences of *nif* promoters and can be separated from the positive control function of NifA. *Nucleic Acids Res.* 16:11469-11488.
73. Morett, E., H. M. Fischer, and H. Hennecke. 1991. Influence of oxygen on DNA binding, positive control, and stability of the *Bradyrhizobium japonicum* NifA regulatory protein. *J. Bacteriol.* 173:3478-3487.
74. Morett, E. and L. Segovia. 1993. The σ^{54} bacterial enhancer-binding protein family: Mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* 175:6067-6074.
75. Negishi, T., N. Fujita, and A. Ishihama. 1995. Structural map of the alpha subunit of *Escherichia coli* RNA polymerase: Structural domains identified by proteolytic cleavage. *J. Mol. Biol.* 248:723-728.
76. Nellen-Anthamatten, D., P. Rossi, O. Preisig, I. Kullik, M. Babst, H. M. Fischer, and H. Hennecke. 1998. *Bradyrhizobium japonicum* FixK2, A crucial distributor in the FixLJ-dependent regulatory cascade for control of genes inducible by low oxygen levels. *J. Bacteriol.* 180:5251-5255.

77. Ninfa, A. J., L. J. Reitzer, and B. Magasanik. 1987. Initiation of transcription at the bacterial *glnAp2* promoter by purified *E. coli* components is facilitated by enhancers. *Cell* 50:1039-1046.

78. North, A. K., D. S. Weiss, H. Suzuki, Y. Flashner, and S. Kustu. 1996. Repressor forms of the enhancer-binding protein NtrC: Some fail in coupling ATP hydrolysis to open complex formation by σ^{54} -holoenzyme. *J. Mol. Biol.* 260:317-331.

79. Panny, S. R., A. Heil, B. Mazus, P. Palm, W. Zillig, S. Z. Mindlin, T. S. Ilyina, and R. B. Khesin. 1974. A temperature-sensitive mutation of the β' -subunit of DNA-dependent polymerase from *E. coli* T16. *FEBS Lett.* 48:241-245.

80. Polyakov, A., E. Severinova, and S. A. Darst. 1995. Three-dimensional structure of *E. coli* core RNA polymerase: Promoter binding and elongation conformations of the enzyme. *Cell* 83:365-373.

81. Popham, D., D. Szeto, J. Keener, and S. Kustu. 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. *Science* 243:629-635.

82. Record, M.T., W.S. Reznikoff, M.L. Craig, K.L. Mcquade and P.J. Schlx. 1996. *Escherichia coli* and *Salmonella*. ASM, Washington, D.C..

83. Rippe, K., M. Guthold, P. H. von Hippel, and C. Bustamante. 1997. Transcriptional activation *via* DNA-looping: visualization of intermediates in the activation pathway of *E. coli* RNA polymerase. σ^{54} holoenzyme by scanning force microscopy. *J. Mol. Biol.* 270:125-138.

84. Roberts, C. W. and J. W. Roberts. 1996. Base-specific recognition of the nontemplate strand of promoter DNA by *E. coli* RNA polymerase. *Cell* 86:495-501.

85. Roberts, G. P. and W. J. Brill. 1981. Genetics and regulation of nitrogen fixation. *Ann. Rev. Microbiol.* 35:207-235.

86. Ross, W., S. E. Aiyar, J. Salomon, and R. E. Gourse. 1998. *Escherichia coli* promoters with UP elements of different strengths: modular structure of bacterial promoters. *J. Bacteriol.* 180:5375-5383.

87. Sarkar, G. and S. S. Sommer. 1990. The "megaprimer" method of site-directed mutagenesis. *BioTechniques* 8:404-407.

88. Sasse-Dwight, S. and J. D. Gralla. 1990. Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ_{54} . *Cell* 62:945-954.
89. Schmidt, R., P. Margolis, L. Duncan, R. Coppolecchia, C. P. Moran, and R. Losick. 1991. Control of developmental transcription factor σ^F by sporulation regulatory proteins SpoIIAA and SpoIIAB in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 87:9221-9225.
90. Spohn, G., D. Beier, R. Rappuoli, and V. Scarlato. 1997. Transcriptional analysis of the divergent *cagAB* genes encoded by the pathogenicity island of *Helicobacter pylori*. *Mol. Microbiol.* 26:361-372.
91. Su, W., S. Porter, S. Kustu, and H. Echols. 1990. DNAP-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. *Proc. Natl. Acad. Sci. USA* 87:5504-5508.
92. Tauschek, M., C. W. Hamilton, L. A. Hall, C. Chomvarin, J. A. Fyfe, and J. K. Davies. 1997. Transcriptional analysis of the *groESL* operon of *Neisseria gonorrhoeae*. *Gene* 189:107-112.
93. Taylor, M., R. Butler, S. Chambers, M. Casimiro, F. Badii, and M. Merrick. 1996. The RpoN-box motif of the RNA polymerase sigma factor σ_N plays a role in promoter recognition. *Mol. Microbiol.* 22:1045-1054.
94. Thöny, B., D. Anthamatten, and H. Hennecke. 1989. Dual control of the *Bradyrhizobium japonicum* symbiotic nitrogen fixation regulatory operon *fixRnifA*: Analysis of *cis*- and *trans*-acting elements. *J. Bacteriol.* 171:4162-4169.
95. Thöny, B., H. M. Fischer, D. Anthamatten, T. Bruderer, and H. Hennecke. 1987. The symbiotic nitrogen fixation regulatory operon (*fixRnifA*) of *Bradyrhizobium japonicum* is expressed aerobically and is subject to a novel, *nifA*-independent type of activation. *Nucleic Acids Res.* 15:8479-8499.
96. Thöny, B. and H. Hennecke. 1989. The -24/-12 promoter comes of age. *FEMS Microbiol. Rev.* 63:341-358.
97. Tintut, Y., C. Wong, Y. Jiang, M. Hsieh, and J. D. Gralla. 1994. RNA polymerase binding using a strongly acidic hydrophobic-repeat region of σ_{54} . *Proc. Natl. Acad. Sci. USA* 91:2120-2124.
98. Travers, A. A. 1987. Structure and function of *E. coli* promoter DNA. *CRC Crit. Rev. Biochem.* 22:180-219.

99. Wang, L. F., S. S. Park, and R. H. Doi. 1999. A novel *Bacillus subtilis* gene, *antE*, temporally regulated and convergent to and overlapping *dnaE*. *J. Bacteriol.* 181:353-356.
100. Weiss, D. S., J. Batut, K. E. Klose, J. Keener, and S. Kustu. 1991. The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. *Cell* 67:155-167.
101. Weiss, D.S., K.E. Klose, T.R. Hoover, A.K. North, S.C. Porter, A.B. Wedel, and S. Kustu. 1992. Prokaryotic transcriptional enhancers p. 667-694. In S.L. McKnight and K.R. Yamamoto (ed.), *Transcriptional regulation*. Cold Spring Harbor Laboratory Press, New York.
102. Weiss, V., F. Claverie-Martin, and B. Magasanik. 1988. Phosphorylation of nitrogen regulator-I of *Escherichia coli* induces strong cooperative binding to DNA essential for activation of transcription. *Proc. Natl. Acad. Sci. USA* 89:5088-5092.
103. Weiss, V. and B. Magasanik. 1988. Phosphorylation of NRI of *E. coli*. *Proc. Natl. Acad. Sci. USA* 85:8919-8923.
104. Wilson, C. and A. J. Dombroski. 1997. Region 1 of σ^{70} is required for efficient isomerization and initiation of transcription by *Escherichia coli* RNA polymerase. *J. Mol. Biol.* 267:60-74.
105. Witty, J.F., F.R. Michin, L. Skot, and J.E. Sheehy. 1986. Nitrogen fixation and oxygen in legume root nodules p. 275-314. In B.J. Mifflin (ed.), *Oxford surveys of plant molecular and cell biology*. Oxford University, Oxford.
106. Wong, C. and J. D. Gralla. 1992. A role for the acidic trimer repeat region of transcription factor σ_{54} in setting the rate and temperature dependence of promoter melting in vivo. *J. Biol. Chem.* 267:24762-24768.
107. Wootton, J. C. and M. Drummond. 1989. The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Prot. Eng.* 2:535-543.
108. Wyman, C., I. Rombel, A. K. North, C. Bustamante, and S. Kustu. 1997. Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science* 275:1658-1661.
109. Yura, T. and A. Ishihama. 1979. Genetics of bacterial RNA polymerase. *Ann. Rev. Genet.* 13:59-97.