

00551



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

INSTITUTO DE BIOTECNOLOGIA

EVOLUCION DIRIGIDA DEL DOMINIO CENTRAL
DE LA PROTEINA TyrR DE *Escherichia coli* Y
ANALISIS *In silico* DEL RECLUTAMIENTO DE EBPs.

T E S I S

QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS BIOQUIMICAS
P R E S E N T A :
BIOL. ALFREDO MENDOZA VARGAS

TUTOR: DR. HUMBERTO FLORES SOTO



CUERNAVACA, MOR.

2004



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.

Esta tesis fue realizada en el laboratorio del Dr. Juan Enrique Morett Sánchez, en el Departamento de Ingeniería Celular y Biocatálisis del Instituto de Biotecnología de la Universidad Nacional Autónoma de México, bajo la asesoría del Dr. Humberto Flores Soto.

Esta tesis se la dedico a mis padres y hermanos por el apoyo que me han dado y en especial a Elizabeth por alentarme a seguir cumpliendo mis metas.

Gracias.

Agradecimientos.

Quiero agradecer primero a Humberto Flores por compartir sus conocimientos, por su paciencia y dedicación durante el desarrollo de este trabajo, y por otorgarme su confianza y su valiosa amistad.

De la misma manera agradezco al Dr. Enrique Morett por estar siempre apoyándome y por promover mi desarrollo académico compartiendo sus conocimientos y experiencias.

Agradezco a la Dra. Brenda Valderrama y a los Drs. Lorenzo Segovia, Víctor González, Ricardo Oropeza y Humberto Flores por sus valiosos comentarios para la mejora del trabajo y por el tiempo dedicado en la revisión de la tesis.

Agradezco a la Dra. Guadalupe Espín, y a los Drs. Miguel Ángel Cevallos, Humberto Flores y Enrique Morett por ser parte de mi comité tutorial y por darme los consejos siempre atinados durante el desarrollo de la tesis.

A Ricardo Grande por sus buenos comentarios sobre mi trabajo, por compartir sus conocimientos y por brindarme su amistad.

No podía faltar mi agradecimiento a Lety y Mary Olvera, quienes me han enseñado a ser eficiente en mi trabajo y claro por su gran amistad.

Agradezco a Ernesto (Dago), Yagul, Christian, Esmeralda, Gabriel, Joel, Dra. Saab, Fily, Luis (Ledezma), Olga, Etiene, Adrián, Mari Cruz, Juanita, Liliana, Carlos, Luis (Lozano), Santiago y no podía faltar Nelly, por compartir su espacio, sus conocimientos, y hacer más leve mi estancia en el laboratorio.

A los camaradas Luis, Pavel, Santiago, Yagul y Yimmy que siempre estuvieron en las buenas y en las malas, apoyándome y ayudándome durante el tiempo del posgrado, y por compartir junto con los compañeros de generación sus conocimientos.

A la Universidad Nacional Autónoma de México y al Instituto de Biotecnología por permitir superarme académicamente.

A la comunidad del IBT.

Finalmente quiero agradecer a CONACyT por financiar parte de la estancia en el posgrado y a PAPIIT-DGAPA (proyecto IN222803) por financiar mi estancia durante la escritura de la tesis.

INDICE GENERAL.

Indice de Figuras	ii
Indice de Tablas	iii
Abreviaturas	iii
Resumen	iv
Introducción	1
El sistema de regulación σ^{54}	2
Proteínas de unión a enhancer	4
TyrR de <i>Escherichia coli</i>	8
El regulón TyrR	9
TyrR y proteínas de unión a enhancer	12
Reclutamiento de EBPs	14
Evolución molecular <i>in vitro</i>	15
Planteamiento	17
Hipótesis y Objetivos	18
Materiales y Métodos	19
Resultados	28
I. Análisis <i>in silico</i> del sistema σ^{54}	28
Resultados y discusión	28
II. Evolución por mutagénesis del dominio central de TyrR de <i>E. coli</i>	44
Resultados	44
Discusión	60
Conclusiones	64
Apéndice	v
Bibliografía	viii

INDICE DE FIGURAS.

Figura 1	Dominios funcionales de las EBPs y regiones conservadas del dominio central.	5
Figura 2	Localización de cajas TyrR en las unidades transcripcionales del regulón TyrR.	11
Figura 3	Mapa de los genes estructurales del operón <i>phh</i> .	14
Figura 4	Mutagénesis dirigida al sitio de TyrR/PhhR y TyrR/PhhR+Arg.	21
Figura 5	Eliminación del gen <i>rpoN</i> con el uso del vector pK03.	28
Figura 6	Disposición de dominios de las EBPs únicas presentes en el par funcional.	31
Figura 7	Disposición de dominios de las EBPs presentes en organismos sin σ^{54} .	37
Figura 8	Alineamiento múltiple de secuencias del dominio central de las 14 EBPs encontradas en organismos sin σ^{54} .	42
Figura 9	Árbol filogenético de probables EBPs reclutadas.	43
Figura 10	Alineamiento de secuencias protéicas del dominio central de varias EBPs.	45
Figura 11	Actividad de β -galactosidasa <i>in vivo</i> de construcciones del dominio central de TyrR.	46
Figura 12	Alineamiento de la región N-terminal de los dominios de activación de distintas construcciones de TyrR y PspF.	49
Figura 13	Electroforesis en gel de agarosa al 1% para la caracterización genotípica por PCR de MC1061 $\Delta rpoN$.	51
Figura 14	Activación transcripcional de pUCTyrRZ y otras construcciones en presencia y ausencia de σ^{54} .	53
Figura 15	Actividad transcripcional de las mutantes M1-M5 obtenidas del banco mutagénico de TyrR/PhhR.	55
Figura 16	Alineamiento de las secuencias proteicas de las mutantes M1-M5.	56
Figura 17	Activación transcripcional de M4 y M5 en presencia y ausencia de σ^{54} .	57
Figura 18	Plásmidos reporteros para identificar la región en pMB210 que origina la activación transcripcional por las mut. M4 y M5.	58
Figura 19	Activación transcripcional de pMB210tt con las variantes M4 y M5 en presencia o ausencia de σ^{54} .	58
Figura 20	Activación transcripcional de M4 y M5 en promotor σ^{70} (pRO310).	59
Figura 21	Bloqueo de promotor σ^{54} en la activación transcripcional de M4 y M5.	60

INDICE DE TABLAS.

Tabla 1	Componentes del regulón TyrR.	9
Tabla 2	Oligonucleótidos.	22
Tabla 3	Plásmidos.	23
Tabla 4	Elementos del sistema σ^{54} en los genomas totalmente secuenciados	29
Tabla 5	Matriz de identidad de secuencia del grupo de Clamidias, proteína completa.	32
Tabla 6	Actividad de β -galactosidasa <i>in vivo</i> de pUCTyrRZ y pUCPspFZ.	49
Tabla 7	Caracterización fenotípica MC1061 Δ rpoN I	52
Tabla 8	Caracterización fenotípica MC1061 Δ rpoN II	52

ABREVIATURAS.

AAA	ATPasa asociada con actividades celulares
ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
ARNpol	ARN polimerasa
ATP	Adenosine triphosphate
CAP	Catabolite gene activating protein
cNMP	Cyclic nucleotide monophosphate
DMSO	Dimethyl sulfoxide
EBP	Enhancer binding protein
ELE	Enhancer like elements
E σ^{54}	ARN polimerasa-sigma54
GTP	Guanosine triphosphate
HVH	Hélice vuelta hélice
IHF	Integration host factor
LB	Luria-Bertani
ml	mililitro
NCBI	National center for biotechnology information
NJ	Neighbour joining
OD	Densidad óptica
ONPG	o-nitrofenil- β -D-galactopiranosido
PAS	Per, ARNT, and Sim
pb	Par de bases
PCR	Polymerase chain reaction
PRD	PTS regulation domain
UAS	Upstream activator sequence
X-gal	5-bromo-4cloro-3-indolil- β -D-galactosido

RESUMEN

Durante el inicio transcripcional en bacterias, la regulación dependiente de σ^{54} requiere de la hidrólisis de ATP que es catalizada por proteínas activadoras dependientes de enhancer (EBPs). Estos activadores transcripcionales se han identificado en una gran variedad de especies bacterianas y presentan una estructura característica con tres dominios funcionalmente independientes, donde el dominio central es el más conservado y contiene los requerimientos necesarios para la activación transcripcional. Este dominio alberga una región llamada C3 que ha sido considerada crítica para la interacción con σ^{54} . Se ha observado que TyrR de *E. coli* no presenta completa esta región, por lo tanto ha sido reclutada para funcionar de una manera independiente de σ^{54} . En este trabajo proponemos que en la familia de EBPs se han dado varios eventos de reclutamiento, similares al observado en TyrR. Para demostrar esto se realizó una búsqueda del par funcional σ^{54} -EBP en los genomas totalmente secuenciados. Se encontraron 14 eventos de reclutamiento en organismos que no presentaban σ^{54} , y otros casos interesantes que demuestran la plasticidad del sistema σ^{54} . También buscamos restablecer la activación dependiente de σ^{54} por TyrR utilizando técnicas de evolución dirigida. Sin embargo, no fue posible restaurar una actividad transcripcional dependiente de σ^{54} en el dominio central de TyrR, por lo que es necesario buscar nuevas modificaciones e implementar nuevos sistemas de tamizado (reporteros) para obtener las variantes deseadas.

INTRODUCCIÓN.

El proceso transcripcional es el mecanismo mediante el cual los genes son copiados en una molécula de ARN, que por lo general sirve como molde para la síntesis de una proteína. En bacterias, la enzima que lleva a cabo este proceso es la ARN polimerasa (ARNpol) dependiente de ADN. Esta enzima presenta la maquinaria necesaria para la síntesis de ARN, sin embargo no puede iniciar el proceso en lugares específicos del ADN, para ello requiere de factores transcripcionales presentes en la célula. El proceso transcripcional se puede dividir en varias etapas: *complejo cerrado*- se produce cuando la maquinaria transcripcional se une a la doble cadena del ADN en una secuencia específica llamada promotor, este complejo es transcripcionalmente inactivo; *complejo abierto*- se produce cuando se separan las cadenas mediante cambios conformacionales en el complejo cerrado dando lugar a la burbuja transcripcional, para permitir que la cadena templado esté disponible para la síntesis de la hebra de ARN; *iniciación*- es la polimerización de los primeros ribonucleótidos del ARN, sin embargo, este proceso es frecuentemente detenido por la aparición de eventos abortivos, que liberan transcritos cortos para iniciar nuevamente la síntesis de un ARN; *elongación*- ésta etapa involucra el desplazamiento de forma discontinua de la burbuja transcripcional y de la ARNpol sobre el ADN para el crecimiento de la molécula de ARN; *terminación*- existen señales en el ADN que reconoce la ARNpol para detener la polimerización, sin embargo para la liberación de la ARNpol es necesaria la formación de una estructura secundaria de tallo y asa rica en GC, ó la utilización de factores proteicos como Rho que facilitan la terminación (Platt, 1994; Richardson and Ruteshouser, 1986; Richardson, 2003).

En todos los organismos es necesario adecuar la expresión de sus genes a condiciones de crecimiento específicas. En bacterias como la traducción está acoplada a la transcripción, la regulación (adecuación) de la expresión genética se da principalmente a nivel transcripcional, aunque también hay mecanismos de regulación post-transcripcional.

En la transcripción bacteriana existe un mecanismo regulatorio que se produce durante la formación del complejo del inicio, mediado principalmente por factores transcripcionales que modulan el reclutamiento de las ARN polimerasas sobre elementos promotores en el ADN. El descubrimiento del factor transcripcional sigma (σ) que es una sub-unidad de la ARN polimerasa (ARNpol) permitió entender que este factor confería el

reconocimiento específico de los promotores durante el inicio transcripcional (Buck and Cannon, 1992; Goldfarb *et al.*, 1983; Sugiura *et al.*, 1970). Así, los patrones de expresión genética son derivados de la formación de complejos de la ARNpol con distintos factores σ . Un ejemplo clásico de este fenómeno, se observa en la cascada de factores σ que dirigen casi totalmente la expresión de los elementos necesarios para el proceso de esporulación en *Bacillus subtilis* (Kroos *et al.*, 1999).

La holoenzima ARNpol- σ ($E\sigma$) está formada de dos componentes: el núcleo, que es un heterodímero compuesto de cuatro sub-unidades ($\alpha[2]\beta\beta'$) que es el principal el responsable de la elongación de la cadena del ARN, y el factor σ , sub-unidad requerida para el reconocimiento del promotor en la etapa de inicio transcripcional.

Se han identificado diversos factores transcripcionales (σ^{70} , σ^{54} , σ^{32} , σ^{28} , σ^S , etc.), de los cuales, la mayoría muestran similitud funcional y estructural con el factor σ^{70} de *Escherichia coli*. Dentro de la clase σ^{70} , se incluye al factor responsable de la regulación transcripcional constitutiva y a la mayoría de factores que dirigen la transcripción de genes requeridos bajo ciertas condiciones especializadas (conocidos como factores σ alternativos) (Helmann, 1991; Studholme and Buck, 2000). Todos los factores de la familia σ^{70} , reconocen promotores compuestos por dos elementos localizados aproximadamente en las posiciones -10 y -35 con respecto al inicio de la transcripción. Las ARNpol's asociadas con cualquiera de estos factores son capaces en principio de isomerizar el complejo de inicio transcripcional de un estado cerrado a uno abierto en ausencia de otros elementos.

Se ha observado, que dentro de los factores σ existe uno que realiza una regulación más compleja y específica. Dicho factor σ , no muestra homología con σ^{70} ni con ninguna otra proteína y utiliza la energía de hidrólisis de ATP para promover activamente el inicio de la transcripción (Merrick, 1993; Studholme and Dixon, 2003; Weiss *et al.*, 1991), una característica no presente en la clase σ^{70} . Este elemento denominado factor sigma 54 (σ^{54}) es único en su tipo, y es conocido también como RpoN, NtrA o sigma N (Merrick, 1993; Ronson *et al.*, 1987).

El sistema de regulación σ^{54}

En un inicio, al sistema de regulación σ^{54} se le atribuyó principalmente el control de genes relacionados con la asimilación y fijación de nitrógeno (Kustu *et al.*, 1989;

Magasanik, 1989; Merrick, 1992). pero actualmente su función también está relacionada con actividades biológicas tales como la respuesta a la presencia de compuestos aromáticos, uso de alginatos, formación del pili o flagelo y transporte de ácidos dicarboxílicos entre muchas otras funciones (Kustu *et al.*, 1989 ; Studholme and Dixon, 2003; Thony and Hennecke, 1989). Es importante resaltar que en todos los genes regulados por este sistema, el mecanismo de acción es similar y la variabilidad regulatoria es proporcionada por los promotores y por un elemento adicional conocido como proteína de unión a enhancer (EBP).

El factor σ^{54} presenta una estructura primaria que consta de tres regiones: una región carboxilo terminal necesaria para la unión al ADN, una región central rica en residuos ácidos, que interactúa con la ARNpol para la desnaturalización del ADN en la región del promotor, y una región amino terminal rica en leucinas y glutaminas que proporciona los requerimientos necesarios para llevar a cabo la regulación. (Guo and Gralla, 1997; Hsieh and Gralla, 1994; Hsieh *et al.*, 1994; Syed and Gralla, 1998; Wong *et al.*, 1994). σ^{54} es el único factor transcripcional capaz de regular la expresión de promotores bacterianos que cuenten con elementos potenciadores conocidos como secuencias activadoras río arriba (UAS, upstream activator sequence o ELE, enhancer like elements) similares a los enhancers eucarióticos (Buck, 1986).

σ^{54} tiene la capacidad de unirse en ausencia de la ARNpol (Buck and Cannon, 1992) a promotores específicos que presenten las secuencias -12 (consenso TTGCA) y -24 (consenso TGGCAC) (Barrios *et al.*, 1999; Merrick, 1993; Morett and Buck, 1989). Sin embargo, el control transcripcional realizado por este factor sólo inicia cuando se asocia con el núcleo de la ARNpol (formación de la holoenzima ARNpol- σ^{54} [$E\sigma^{54}$]) dando como resultado la generación del complejo de iniciación cerrado. En este momento, la maquinaria transcripcional ($E\sigma^{54}$) se encuentra en un estado inactivo y solamente cuando recibe una señal de una proteína activadora es capaz de continuar el proceso transcripcional. Estas moléculas activadoras son conocidas como proteínas de unión a enhancers o potenciadores (enhancer binding proteins: EBPs) (Popham *et al.*, 1989). Los enhancers bacterianos (UAS), son secuencias en el ADN (palindrómicas o cercanamente palindrómicas) que se encuentran lejos del inicio de la transcripción. Durante la activación, la EBP unida a un enhancer interacciona transitoriamente con la $E\sigma^{54}$; esta interacción es frecuentemente

facilitada por el factor de integración del hospedero (IHF) (Hoover *et al.*, 1990). Dicho factor de integración permite que la curvatura del ADN sea la idónea para el contacto, sin embargo, en algunos casos el ADN puede presentar dicha curvatura de manera natural (Carmona and Magasanik, 1996). Posterior a la formación del complejo EBP- $E\sigma^{54}$ ocurre una apertura en el promotor, para la formación del complejo transcripcional abierto. Esta apertura es dependiente de la energía obtenida de la hidrólisis de ATP, que es realizada por la EBP (Weiss *et al.*, 1991). Por lo tanto, los activadores de $E\sigma^{54}$ acoplan la energía liberada de la hidrólisis de ATP o GTP a la isomerización de un complejo de inicio transcripcional cerrado a uno abierto (Lee *et al.*, 1994; Popham *et al.*, 1989; Sasse-Dwight and Gralla, 1988). Finalmente, una vez que el complejo transcripcional abierto es formado, la cadena de ADN templado puede ser leída, permitiendo la continuación del proceso transcripcional. En ese momento el factor σ^{54} es liberado de la ARNpol para ser reciclado en otro evento transcripcional.

La dependencia de enhancers o potenciadores, y de energía para abrir la doble cadena de ADN, son características de la ARNpol eucariótica del tipo II muy distinta al mecanismo transcripcional bacteriano utilizado por σ^{70} . Por lo tanto, el sistema de regulación σ^{54} es considerado como un híbrido en el sentido que tiene propiedades procarióticas y eucarióticas (Syed and Gralla, 1998). Sin embargo, esto no es un indicio que σ^{54} haya originado los sistemas de regulación de eucariontes.

Proteínas de unión a enhancer.

Las proteínas de unión a enhancer (EBPs) han sido identificadas en una gran variedad de especies bacterianas regulando la expresión de distintos genes. Estas proteínas se unen a sitios en el ADN, UAS, localizados comúnmente a 70-150 pb río arriba del inicio de la transcripción (Xu and Hoover, 2001). Estos elementos potenciadores contienen generalmente dos o más sitios para la unión de dímeros de EBPs. Los dímeros anclados en el ADN sirven como elementos nucleadores para la formación de los oligómeros funcionales (Porter *et al.*, 1993; Scholl and Nixon, 1996). Se ha observado que las formas oligoméricas de las EBPs tienen la capacidad de incrementar la tasa de hidrólisis de ATP que a su vez es acoplada a la formación de complejos transcripcionales abiertos de la $E\sigma^{54}$ (Xu and Hoover, 2001).

Estos activadores (EBPs) son proteínas modulares que generalmente presentan tres dominios funcionales y estructuralmente independientes: un dominio amino terminal, requerido para el reconocimiento de la señal regulatoria; un dominio central, que es responsable de la hidrólisis de ATP y de la activación transcripcional; y un dominio carboxilo terminal, que es el que se une al ADN (UAS) (Morett and Segovia, 1993). Además de los dominios, se ha observado en algunos casos que la unión (o linker) que conecta los dominios amino terminal y dominio central, también puede presentar funciones regulatorias (Garmendia and de Lorenzo, 2000) (fig. 1A).

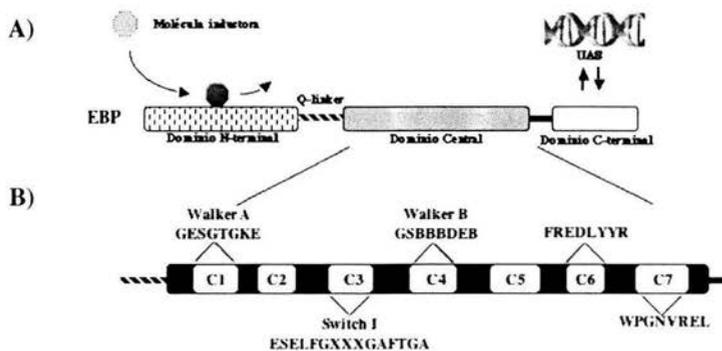


Figura 1. Dominios funcionales de las EBP y regiones conservadas del dominio central. A) Las EBP presentan 3 dominios característicos, el N-terminal que se encarga de responder a estímulos, el dominio central que es el catalítico para la activación dependiente de σ^4 , y el C-terminal que es el dominio de reconocimiento para las secuencias potenciadoras en el ADN. B) El dominio central presenta 7 regiones conservadas implicadas en su función y se muestran las secuencias consenso características para algunas de ellas. Las letras mayúsculas corresponden a la nomenclatura de aminoácidos, la letra B indica residuo hidrofóbico y la letra X indica cualquier residuo.

Dominio N-terminal. De los tres dominios presentes en las EBP, el dominio amino terminal es el menos conservado y permite clasificar a estos activadores transcripcionales de acuerdo al tipo de señal que recibe (Xu and Hoover, 2001). La mayoría de las EBP contienen dominios N-terminales del tipo de sistemas de dos componentes. Estas proteínas, que incluyen NtrC y DctD, son positivamente reguladas por la fosforilación en su amino terminal de un residuo aspartato. En NtrC la fosforilación dispara un cambio conformacional del dominio amino terminal, exponiendo una superficie hidrofóbica que probablemente transmite la señal al dominio central para la activación (Kern *et al.*, 1999). En contraste, en DctD el dominio receptor al recibir la señal por fosforilación, inhibe al

dominio central, reprimiendo su actividad constitutiva (Lee *et al.*, 1994). Un segundo grupo de activadores dependientes de σ^{54} , incluyen a las proteínas XylR y DmpR que interactúan directamente con pequeñas moléculas inductoras. Las interacciones con el inductor promueven la actividad de ATPasa, presumiblemente estimulando la oligomerización (Ng *et al.*, 1995) y de la misma forma que en DctD, los dominios de reconocimiento de señal de XylR y DmpR actúan como inhibidores intramoleculares de la capacidad de activación del dominio central (O'Neill *et al.*, 1998; Perez-Martin and De Lorenzo, 1995). El tercer grupo de activadores lo integran proteínas como NifA de *Klebsiella pneumoniae*, en la cual el dominio amino terminal regula la actividad a través de las interacciones con la proteína NifL que tiene la capacidad de sensar la concentración de O₂ (Berger *et al.*, 1994; Morett *et al.*, 1990; Shingler, 1996). Finalmente, existen EBPs que sólo tienen el dominio central y el dominio carboxilo terminal. La más caracterizada de éstas, es la proteína de respuesta a fago (PspF) (Elderkin *et al.*, 2002). La actividad de PspF es controlada por la formación de un complejo represivo con la proteína PspA (PspA-PspF), y se ha sugerido que PspA pudo en algún momento ser el dominio regulador amino terminal de una proteína PspF ancestral (Elderkin *et al.*, 2002). Studholme y Dixon (2003) describen una gran variedad de funciones asociadas al dominio N-terminal de las EBPs en un trabajo publicado recientemente.

Dominio C-terminal. El dominio C-terminal de diversas EBPs se encuentra medianamente conservado y se ha predicho que forma una estructura de hélice-vuelta-hélice (HVH) de unión a ADN (Drummond *et al.*, 1986). Los estudios de mutagénesis (Morett *et al.*, 1988), espectroscopía (Missailidis *et al.*, 1999) y resonancia magnética nuclear (P. Ray, K. J. Smith, R. A. Nixon, and E. I. Hyde, comunicación personal) en las proteínas NifA y NtrC apoyan ésta interpretación. También, se ha observado que éste dominio contiene los determinantes para la dimerización (Klose *et al.*, 1994) y probablemente también para la oligomerización en algunas EBPs (Porter *et al.*, 1993). Finalmente, se ha determinado por comparación de secuencias y por mutagénesis del arreglo HVH, la presencia de una glicina responsable para la formación de la segunda hélice (Aurora *et al.*, 1994), la cual confiere al activador un fuerte pegado a las secuencias activadoras (UAS) (Juarez *et al.*, 2000).

Dominio central. Una gran variedad de estudios genéticos, bioquímicos, de análisis de secuencia, y estructurales de los activadores transcripcionales dependientes de σ^{54} , han proporcionado una considerable cantidad de información sobre las funciones relacionadas con el dominio central como son, la activación transcripcional, la función de ATPasa y la oligomerización (fig. 1B) (Xu and Hoover, 2001). La estructura primaria del dominio central es la más conservada en las EBPs, y con base en ello, se han identificado siete regiones de importancia funcional (C1-C7 fig. 1B) (Morett and Segovia, 1993). El dominio central alberga los arreglos conocidos como Walker A y Walker B que participan en la unión e hidrólisis de ATP respectivamente (Walker *et al.*, 1982). En el arreglo Walker A (región C1), se ha observado una secuencia rica en glicinas que se encarga de interactuar con el fosfato β del ATP (Saraste *et al.*, 1990), mientras que, el arreglo Walker B (región C4) presenta un residuo muy conservado (ácido aspártico) que se encarga de coordinar el Mg^{2+} necesario para la hidrólisis del ATP (Hilgenfeld, 1995). Mutaciones en cualquiera de los residuos de estas regiones bloquean o disminuyen considerablemente la función de ATPasa, la cual es imprescindible durante la activación transcripcional. Además de estas regiones se ha sugerido que hay otros residuos en el dominio central que intervienen en la unión e hidrólisis de ATP, incluyendo ciertos residuos en la región C3, una arginina potencialmente catalítica localizada en la región C6, y algunos residuos de la región C7 que pueden estar participando en la unión de ATP y para la formación de oligómeros activos (Perez-Martin and de Lorenzo, 1996; Rombel *et al.*, 1999). Por estas características y por análisis de secuencia se ha determinado que el dominio central de las EBPs es miembro de la superfamilia AAA+ (ATPasa Asociada con una variedad de Actividades celulares), que utiliza mecanismos como la oligomerización y la conversión de la energía derivada de la hidrólisis de ATP en una fuerza mecánica para la remodelación de sus sustratos (Dixon *et al.*, 2002; Neuwald *et al.*, 1999; Vale, 2000).

Por otro lado, la región C3 del dominio central de las EBPs es considerada crítica para la activación transcripcional de genes dependientes de σ^{54} . Mutantes en ésta región son incapaces de activar la transcripción, aunque presentan niveles significativos de actividad de ATPasa y buena unión al ADN (Gonzalez *et al.*, 1998; North *et al.*, 1996). También se ha observado que algunas mutantes en ésta región son defectuosas en la interacción con σ^{54} (Lee and Hoover, 1995; Morett and Segovia, 1993; Wang *et al.*, 1997), así como

desacopladores entre la hidrólisis del ATP y la activación (North *et al.*, 1996; Wang *et al.*, 1997). Finalmente, análisis comparativos de la secuencia han señalado que existe una clase de EBPs que presentan un dominio central con parte de la región C3 modificada (Foster-Hartnett and Kranz, 1992; Gonzalez *et al.*, 1998; Morett and Segovia, 1993; Pittard and Davidson, 1991). En dicha clase sólo se observan algunos casos como son NtrC de *Rhodobacter capsulatus* y TyrR de *Escherichia coli*. Esta última es parte fundamental de este trabajo, ya que despertó nuestra atención por llevar el control transcripcional con otra forma distinta a la holoenzima ARNpol- σ^{54} .

Todas las características antes mencionadas llevaron a Morett y Segovia (1993) a proponer que la región C3 de las EBPs constituye un elemento de reconocimiento y contacto con la E σ^{54} para activar la transcripción en los promotores dependientes de σ^{54} .

TyrR de *Escherichia coli*.

TyrR presente en *E. coli*, *Haemophilus influenzae*, *Pasteurella multocida*, *Vibrio cholerae* entre otras, es un regulador transcripcional de genes dependientes de E σ^{70} que reconoce los promotores -10 y -35 del inicio de la transcripción. TyrR en su forma monomérica tiene 513 aminoácidos y presenta una estructura modular conformada por tres dominios funcionalmente independientes (Cui and Somerville, 1993a). En la región carboxilo terminal se observa un arreglo de unión a ADN, hélice-vuelta-hélice (HvH), el cual es crítico para la regulación transcripcional (Hwang *et al.*, 1997). El dominio amino terminal alberga los residuos que son importantes para la activación transcripcional de promotores tales como *tyrP* y *mtr*, para la dimerización, y para la unión de aminoácidos aromáticos independiente de ATP (Cui and Somerville, 1993b; Kwok, 1998; Wilson *et al.*, 1995; Yang *et al.*, 1993). El dominio central es homólogo al dominio central de las EBPs (Morett and Segovia, 1993), y contiene un sitio de unión a ATP, un sitio de unión de aminoácidos aromáticos dependiente de ATP, además de estar implicado en la formación de hexámeros activos (Kwok, 1998; Wilson *et al.*, 1995; Yang *et al.*, 1993).

Como se ha descrito, la proteína TyrR presenta varias regiones que interactúan con moléculas tales como aminoácidos aromáticos (tirosina y fenilalanina) y/o cofactores (ATP). Estas múltiples interacciones le permiten a TyrR presentar al menos dos tipos de conformaciones que son importantes para su función. Se ha observado que en solución, y

en ausencia de cofactores o en la presencia de fenilalanina, la proteína TyrR existe como dímero. Sin embargo, con ATP, y tirosina, o ATP y concentraciones altas de fenilalanina la proteína TyrR se auto-agrega formando hexámeros (Wilson *et al.*, 1994).

El regulón TyrR

La proteína TyrR en *Escherichia coli* controla la expresión de un grupo de genes y operones cuyos productos están relacionados con la biosíntesis o transporte de aminoácidos aromáticos (Pittard, 1996) y puede actuar como represor o activador dependiendo del promotor y de la molécula co-efectora que tenga unida. Esta proteína reguladora junto con los genes que se encuentran bajo su control constituyen el regulón TyrR (tabla I).

Los reguladores transcripcionales negativos como TyrR se unen a regiones específicas llamadas operadores para reprimir o activar la transcripción de los genes bajo su control. Una característica general de los represores bacterianos es la de impedir la transcripción bloqueando el acceso de la ARNpol a los promotores (Choy *et al.*, 1997). Sin embargo, también se puede inhibir la transcripción en pasos posteriores al inicio transcripcional (Choy and Adhya, 1992; Monsalve *et al.*, 1996). TyrR reprime la transcripción de ambas maneras.

Tabla 1. Componentes del regulón TyrR (Tomado de Pittard y Davidson, 1991).

Unidad transcripcional	Función	Represión(R) ó Activación (A)	Efeotor
aroF-tyrA	DAHf sintetasa; prefenato deshidrogenasa	R	Tirosina
aroLM	Shikimato cinasa	R	Tirosina
tyrP	Sistema de transporte específico de tirosina	R/A	Tirosina/Fenilalanina
tyrB	Aminotransferasa de aminoácidos aromáticos	R	Tirosina/Fenilalanina
aroP	Sistema de transporte de aminoácidos aromáticos	R	Tirosina/fenilalanina/triptófano
tyrR	Proteína TyrR	R	Ninguno
mtr	Sistema transportador específico de triptófano	A	Fenilalanina/tirosina
aroG	DAHf sintetasa	R	fenilalanina

En general las secuencias operadoras del regulón TyrR, también llamadas cajas TyrR, son secuencias palindrómicas (similares a las UAS) y su secuencia consenso es

TGTAAAN6TTTACA. Se han descrito dos tipos de cajas TyrR, una caja fuerte y una débil, dependiendo de su afinidad con la proteína TyrR y de su similitud con la secuencia consenso (Andrews *et al.*, 1991; Kasian *et al.*, 1986; Yang *et al.*, 2002). La caja fuerte permite la interacción de un dímero o de un hexámero según sea el caso, pero una caja débil solamente permite la interacción del hexámero, siempre y cuando ésta se encuentre cercana a una caja fuerte y en la misma cara de la hélice del ADN (Hwang *et al.*, 1999).

En la represión por tirosina, los genes bajo el control de TyrR contienen dos cajas TyrR adyacentes (Kasian *et al.*, 1986). La localización de estas cajas TyrR depende de la unidad transcripcional bajo control y pueden estar sobrelapando al promotor -35, -10. Cuando la proteína TyrR interactúa con estas secuencias en el ADN por lo general se bloquea el posicionamiento de la maquinaria transcripcional, impidiendo la formación del complejo transcripcional cerrado (Gowrishankar and Pittard, 1998).

En la activación transcripcional llevada por TyrR con la fenilalanina unida, sólo se requiere la presencia de una caja fuerte posicionada apropiadamente río arriba del promotor, contrastando con la represión donde dos cajas son necesarias. Antes de la activación, TyrR se mantiene constitutivamente unida a la caja fuerte en su forma dimérica y al interactuar con fenilalanina adquiere la habilidad para activar la transcripción mediante el proceso de reclutamiento de la ARNpol (Hochschild and Dove, 1998; Ptashne and Gann, 1997). Se ha observado que durante la activación, TyrR incrementa la afinidad de la ARNpol por el promotor débil y estimula la formación del complejo transcripcional abierto, indicando que TyrR funciona como un activador transcripcional de la clase I (Lawley *et al.*, 1995; Yang *et al.*, 1996).

Como ya se ha descrito, la proteína TyrR reprime o activa la expresión de al menos ocho unidades transcripcionales que constituyen el regulón TyrR (Pittard and Davidson, 1991). Estas unidades transcripcionales se encuentran dispersas en el cromosoma de *E. coli* (Pittard *et al.*, 1987) y en la mayoría de los casos TyrR actúa como represor, y solamente en dos casos, *tyrP* y *mtr*, actúa como activador (tabla 1). Para una mejor comprensión de este regulón, a continuación se describirán brevemente algunos ejemplos de como TyrR realiza el control transcripcional.

Regulación de *tyrB*. El gene *tyrB*, que codifica para una aminotransferasa de aminoácidos aromáticos, presenta las cajas TyrR río abajo del promotor (fig. 2). Se ha observado que la

represión de este gene es dependiente de tirosina y fenilalanina, y que su efecto puede darse en distintos pasos de la transcripción (Yang *et al.*, 1999). En presencia de tirosina se impide la formación del complejo abierto, mientras que con fenilalanina se afectan los pasos posteriores a la formación del complejo abierto. Ambas cualidades son dadas por la unión de TyrR sobre sitios cercanos y río abajo del sitio de inicio de la transcripción (Yang *et al.*, 2002). Además, se ha observado que la presencia de las dos cajas en *tyrB* son importantes para la represión transcripcional por tirosina o fenilalanina (Yang *et al.*, 2002).

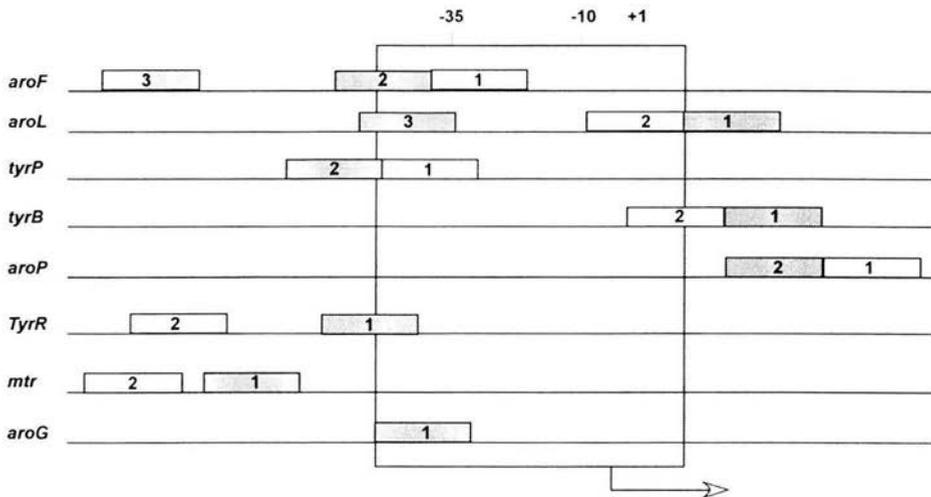


Figura 2. Localización de cajas TyrR en las unidades transcripcionales del regulón TyrR. Las diferentes regiones regulatorias son alineadas por sus inicios de la transcripción mostrado como +1, y se indican los elementos -10 y -35 del promotor. Las cajas claras indican una caja débil, mientras que las oscuras indican cajas fuertes (Tomado de Pittard y Davidson, 1991).

Regulación de *aroP*. El gen *aroP* codifica para una proteína de membrana implicada en el transporte de los tres aminoácidos aromáticos, presenta en su región de control 3 promotores (P1, P2 y P3) (Chye and Pittard, 1987; Wang *et al.*, 1997). TyrR reprime la expresión del promotor principal P1, desplazando a la ARN polimerasa hacia P3 que no es productivo y que se encuentra divergente (Wang *et al.*, 1998; Wang *et al.*, 1997). Por otro lado, el promotor P2 es parcialmente reprimido cuando la proteína TyrR está sola y sólo muestra un ligero incremento de la represión en presencia de tirosina o fenilalanina (Wang *et al.*, 1997). Los sitios de unión para TyrR en la región de control incluyen una caja fuerte

y una débil (fig. 2). TyrR interactúa con ambas cajas cuando están presentes ATP y tirosina, mientras que en presencia de fenilalanina TyrR sólo interactúa con la caja fuerte (Pittard and Davidson, 1991). La fenilalanina probablemente ayuda a la represión aumentando la afinidad de TyrR por el ADN y desplazando a la ARNpol (Yang *et al.*, 1999).

Regulación de *tyrP* y *mtr*. En *E. coli* la expresión de ambos genes *mtr* y *tyrP* que codifican para las permeasas de triptofano y tirosina respectivamente, es activada por TyrR. En el caso de *mtr*, la activación es mediada por tirosina o fenilalanina (Heatwole and Somerville, 1991), mientras que en el caso de *tyrP*, la fenilalanina promueve activación y la tirosina represión (Andrews *et al.*, 1991; Kasian *et al.*, 1986). Con estudios *in vivo* e *in vitro* se han identificado los sitios de unión de TyrR en regiones río arriba del inicio transcripcional de ambos genes (Yang *et al.*, 1996). En *mtr*, la caja TyrR fuerte que juega un papel importante durante la activación (fig. 2), se encuentra localizada aproximadamente a 80 pb río arriba del inicio transcripcional. En *tyrP*, la caja fuerte (fig. 2) se encuentra a 65 pb del inicio, sin embargo, ésta caja no está posicionada idealmente para la activación ya que se ha observado que la activación no es fuerte (Yang *et al.*, 1996).

TyrR y proteínas de unión a enhancer.

Como ya se ha mencionado, el dominio central de la proteína TyrR de *E. coli* presenta una estructura y función relacionadas con la familia de reguladores transcripcionales conocidos como proteínas de unión a enhancer (EBPs) (Morett and Segovia, 1993; Pittard and Davidson, 1991). TyrR, al igual que NtrC de *Rhodobacter capsulatus*, es claramente homóloga a las EBPs, presentando una identidad de secuencia del 35-40 % con los dominios central y carboxilo terminal de las proteínas NtrC, NifA, XylR, FhIA, HydG, DctD y PhhR, que son activadores *bonafide* de promotores σ^{54} (Schlensog and Bock, 1990; Stock *et al.*, 1989). Sin embargo, ambas proteínas son reguladores transcripcionales de promotores σ^{70} y se ha observado que no presentan algunos residuos en regiones del dominio central que son importantes para la activación transcripcional de promotores σ^{54} . El dominio central de la proteína TyrR no presenta parte de la región denominada C3 o switch I (Morett and Segovia, 1993; Osuna *et al.*, 1997). Le faltan los residuos GAFTGA, que han sido propuestos para la interacción con la ARNpol- σ^{54} , así su ausencia limitaría fuertemente su capacidad de interaccionar con σ^{54} y activar los

promotores dependientes de este factor (Foster-Hartnett and Kranz, 1992; Gonzalez *et al.*, 1998; Morett and Segovia, 1993; Pao and Saier, 1995; Pittard and Davidson, 1991). La importancia de estos seis residuos ha sido verificada en la estructura tridimensional de la EBP NtrC1 de *Aquifex aeolicus* recientemente resuelta, donde se observó que parte de la región C3 (residuos GAFTGA) es una asa (loop) expuesta en el heptámero activo que podría interactuar con σ^{54} (Lee *et al.*, 2003).

Además, el dominio central de TyrR también es parte de la superfamilia AAA+ (Dixon *et al.*, 2002; Neuwald *et al.*, 1999; Vale, 2000). Las marcas características de la superfamilia AAA+ para la unión e hidrólisis de ATP, se encuentran en las regiones Walker A y Walker B respectivamente, y ambas regiones están presentes tanto en TyrR como en las EBPs (Actividad de ATPasa) (Cui and Somerville, 1993a, b; Morett and Segovia, 1993). La actividad de ATPasa en las EBPs es requerida para la activación (Austin *et al.*, 1994; MacKintosh and MacKintosh, 1994), sin embargo en el control transcripcional llevado por TyrR no se ha establecido una relación directa con esta actividad (Zhao *et al.*, 2000).

Por otro lado, dentro de la familia de las EBPs, la proteína que presenta la mayor identidad con el dominio central de TyrR de *E. coli* es PhhR de *Pseudomonas aeruginosa*. PhhR es un activador de genes dependientes de σ^{54} y se ha observado que comparte algunas características a nivel funcional con la proteína TyrR de *E. coli*.

El gen denotado como *phhR* en *Pseudomonas aeruginosa* codifica para una proteína de 518 residuos que regula la transcripción del operón *phh* dependiente de σ^{54} , siendo la fenilalanina su inductor fisiológico. El operón está compuesto de tres genes estructurales (*phhA*, *phhB* y *phhC*). *phhA* codifica para una hidroxilasa de fenilalanina, *phhB* codifica para una deshidratasa de carbinolamina y finalmente *phhC* codifica una aminotrasferasa aromática; estas enzimas son las encargadas de las reacciones iniciales durante el catabolismo de fenilalanina (fig. 3) (Jensen and Gu, 1996).

La proteína PhhR que se transcribe de forma divergente al operón *phh* (fig. 3a) es una proteína modular que presenta una alta identidad en secuencia de aminoácidos con la proteína TyrR de *E. coli* (47%) (36% en el dominio amino terminal, 52% en el dominio central y 47% en dominio carboxilo terminal) indicando que son proteínas homólogas. También comparten similitud funcional, ya que existe evidencia experimental de que la

proteína PhhR de *P. aeruginosa* puede sustituir claramente a su homólogo TyrR de *E. coli* como represor transcripcional del operón *aroF-tyrA*, reconociendo e interaccionando con las cajas TyrR. Sin embargo, también se observó que PhhR es incapaz de reemplazar a TyrR como activador transcripcional del gen *mtr*. (Song and Jensen, 1996).

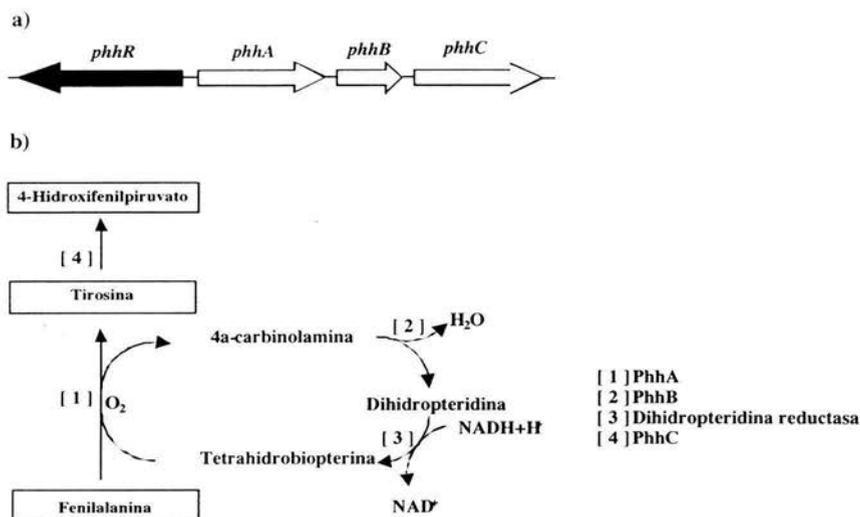


Figura 3. a) Mapa de los genes estructurales del operón *phh*. El gen regulador *phhR* es divergente (en negro) a los demás genes del operón *phh*. b) Reacciones iniciales del catabolismo de fenilalanina en mamíferos. Los tres genes estructurales del operón *phh* codifican enzimas que catalizan tres pasos de los cuatro de la vía (Tomado de Song and Jensen, 1996).

Reclutamiento de EBPs.

Se ha observado que una secuencia protéica con una función determinada puede diversificarse y realizar otras funciones, dando como resultado a los organismos portadores, alternativas para ser más eficientes en su medio o adaptarse a otras condiciones de crecimiento. La estrategia donde los elementos ya existentes se reutilizan en vez de ser creados de *novo* se le llama reclutamiento (Jensen, 1976), y puede ser una forma para facilitar el proceso evolutivo. En la naturaleza se han observado varios ejemplos que serán mencionados a continuación donde se ha demostrado el reclutamiento de proteínas.

Existe evidencia de que el ciclo de la urea fue formado a partir de cuatro enzimas involucradas en la biosíntesis de la arginina con la adición de la arginasa (Takiguchi *et al.*, 1989) y que el ciclo de Krebs se generó también por una combinación de enzimas pre-existentes en las rutas de biosíntesis de aspartato y glutamato (Melendez-Hevia *et al.*, 1996). Por otro lado, el caso de los barriles TIM nos indica claramente la re-utilización constante de un plegamiento para realizar distintas funciones (Copley and Bork, 2000). Finalmente existe evidencia filogenética de que la proteína Fis (factor for inversión stimulation) que une al ADN, fue generada del reclutamiento del dominio carboxilo terminal de una proteína de unión a enhancer (NtrC) durante un evento de transferencia horizontal (Morett and Bork, 1998).

Como ya se mencionó, las proteínas TyrR de *E. coli* y NtrC de *R. capsulatus* son homólogas a las EBPs pero no tienen la capacidad de activar la transcripción de promotores dependientes de σ^{54} y actúan sobre promotores σ^{70} . En organismos donde se presenta σ^{54} , la pérdida de parte de la secuencia de la región C3 (de importancia funcional) en el dominio central de la EBP, probablemente disparó el proceso de reclutamiento de estas proteínas hacia el sistema σ^{70} . El reclutamiento tal vez fue generado para impedir la pérdida del control transcripcional de sus genes blanco. Otra explicación de la existencia de estas proteínas, es que hayan sido originadas en un organismo que sufrió la pérdida de σ^{54} y obtenidas por *E. coli* y *R. capsulatus* por eventos de transferencia. Madan (2003), describe que cuando se pierden los factores transcripcionales, los genes bajo su regulación son perdidos también o reclutados para ser regulados por otro factor.

Evolución molecular *in vitro*.

Con la evolución molecular *in vitro* es posible modificar en el laboratorio, en poco tiempo, las proteínas con el fin de optimizar una función y/u obtener nuevas funciones, simulando los procesos naturales.

La idea de evolución de biomoléculas *in vitro* nació a partir de los trabajos pioneros de Spiegelman y colaboradores (Mills *et al.*, 1967) que fueron realizados sobre ácidos nucleicos. Posteriormente, el grupo de Arnold fue uno de los primeros en aplicar estos conocimientos en biocatálisis para el perfeccionamiento de enzimas implicadas en las rutas metabólicas (Chen and Arnold, 1993). Arnold y su grupo demostraron que con reiterados

ciclos de mutagénesis al azar, seguidos por el tamizado y/o selección de las librerías resultantes, es posible producir variantes de enzimas mucho más estables y activas que las parentales, todo esto en una escala de tiempo relativamente corto (Schmidt-Dannert, 2001). Por otro lado, Stemmer desarrolló un método conocido como DNA shuffling (Stemmer, 1994), el cual permite la recombinación de mutantes deletereas y benéficas (no caracterizadas) para obtener el fenotipo esperado. El método desarrollado por Stemmer ha sido extendido a la recombinación de familias de genes homólogos (family shuffling) (Cramer *et al.*, 1998) permitiendo de esta forma explorar un mayor espacio de secuencia.

Actualmente para la evolución de moléculas biológicas *in vitro* se han implementado una gran cantidad de estrategias con el fin de explorar más ampliamente el espacio de secuencia de cada proteína. La evolución molecular *in vitro* es una herramienta fuertemente utilizada para la modificación de proteínas, además de que puede aplicarse también para investigar las relaciones estructura-función (Schmidt-Dannert, 2001).

En esta sección se describirán las dos estrategias de evolución molecular utilizadas más frecuentemente. La primera, implica el diseño racional que es actualmente conocida como ingeniería de proteínas; la segunda, es un evento azaroso que imita el proceso evolutivo natural, comúnmente llamado mutagénesis al azar.

En el diseño racional (mutagénesis dirigida al sitio), cada uno de los aminoácidos que comprende una proteína puede ser modificado basándose en su importancia funcional previamente determinada. Para realizar esto, frecuentemente se utiliza la PCR (Polymerase Chain Reaction) donde se pueden introducir los cambios a través de los oligonucleótidos (primers) utilizados en la reacción. Con esta estrategia un residuo puede ser sustituido por cualquier aminoácido con el fin de verificar los cambios que puede tolerar una posición, y como estos alteran las propiedades de la proteína. La selección o el tamizado son mínimos debido a que se pueden ensayar independientemente todas las mutantes generadas. De acuerdo con lo anterior, para utilizar esta estrategia, es necesario manejar los conocimientos sobre las interacciones particulares de cada uno de los aminoácidos presentes en una proteína, ya que con esta información es posible predecir los efectos que puedan causar las sustituciones específicas (Skandalis *et al.*, 1997). Sin embargo, la aplicación de estos conocimientos se dificulta en la mayoría de las proteínas, principalmente por la carencia de datos estructurales en las bases de datos.

La otra manera de crear diversidad de secuencia a partir de una sola secuencia proteica es mediante mutagénesis al azar que utiliza técnicas como la PCR mutagénica (Cadwell and Joyce, 1992, 1994). En contraste con la mutagénesis dirigida al sitio, no se requiere un conocimiento previo de la estructura de la proteína (Skandalis *et al.*, 1997). La forma en que se introducen las mutaciones al azar dentro de un gene o un segmento del gene es mediante el uso de condiciones que incrementan la tasa de error de la ADN polimerasa durante la reacción. En la PCR mutagénica, la tasa de mutagénesis puede ser ajustada. Cuando se incrementa la tasa de mutagénesis, se incrementa considerablemente el número de posibles variantes generando un problema práctico en el momento de realizar el tamizado (librerías con miembros de 10^4 - 10^6 por tamizado). Esto además, aumenta la fracción de variantes no funcionales mientras que las mutaciones benéficas y la combinación de estas se ven disminuidas (Schmidt-Dannert, 2001).

La principal diferencia entre la evolución realizada *in vitro* y los procesos evolutivos naturales, es que *in vitro*, es un proceso guiado a una meta con la acumulación de mutaciones adaptativas; mientras que en la evolución natural la base es la presión de selección sometida por el ambiente (Schmidt-Dannert, 2001).

PLANTEAMIENTO.

En este trabajo proponemos restablecer en TyrR una probable actividad ancestral de activador dependiente de σ^{54} , basándonos en el actual conocimiento de las proteínas TyrR de *Escherichia coli* y de PhhR de *Pseudomonas aeruginosa*. Para ello generamos modificaciones en el dominio central, especialmente en la región C3, mediante mutagénesis dirigida al sitio y mutagénesis al azar. Esto, con el fin de poder entender qué modificaciones se produjeron en la proteína TyrR para el cambio de actividad. También, proponemos que la ausencia de alguno de los dos elementos del par funcional σ^{54} -EBP requeridos para el control transcripcional σ^{54} es un indicio de que el elemento conservado ha sido reclutado. Para ello analizamos la presencia de dicho par funcional en los genomas depositados, con el fin de determinar probables eventos de reclutamiento, tomando como base el evento sufrido por TyrR. Finalmente, con la información generada se podrán definir algunos ejemplos que muestren la plasticidad del sistema σ^{54} .

HIPÓTESIS.

La proteína TyrR de *E. coli* fué un activador de genes dependientes del factor σ^{54} , y al perder parte de su secuencia fue reclutada para la regulación de promotores dependientes de σ^{70} . El restablecimiento de su función original puede ser posible modificando parte de la secuencia del dominio central, principalmente en la región C3. De la misma manera que con TyrR de *E. coli*, muy probablemente se han dado más eventos de reclutamiento con otras EBPs, que indicaría una plasticidad estructural evolutiva en estas proteínas.

OBJETIVOS.

Generales.

Determinar los posibles eventos de reclutamiento de EBPs en los genomas bacterianos totalmente secuenciados y generar una variante del dominio central de la proteína TyrR que pueda activar promotores dependientes de σ^{54} .

Particulares.

- Realizar un análisis *in silico* de los elementos que participan en la regulación transcripcional dependiente de σ^{54} (par funcional σ^{54} -EBP) en los genomas bacterianos totalmente secuenciados.
- Búsqueda y asignación de la estructura modular de probables EBPs en organismos que no presenten σ^{54} .
- Asignar probables eventos de reclutamiento de EBPs en los genomas totalmente secuenciados.
- Obtener y clonar el dominio central de PhhR y determinar su activación transcripcional σ^{54} *en trans*.
- Generar por mutagénesis dirigida al sitio variantes del dominio central de TyrR con la región C3 similar a la secuencia de PhhR y determinar la activación transcripcional σ^{54} *en trans*.
- Crear nuevas mutantes del dominio central de la variante TyrR/PhhR con el uso de PCR mutagénica.
- Identificar, aislar y caracterizar las clonas del banco que presenten activación transcripcional dependiente del factor σ^{54} .

- Analizar las secuencias de las mutantes de TyrR con actividad transcripcional σ^{54} para identificar las regiones implicadas para realizar esta función.

MATERIALES Y MÉTODOS.

Análisis in silico.

Se realizó un análisis *in silico* de los elementos que participan en la regulación de la transcripción de genes dependientes del factor σ^{54} presentes en los 145 genomas totalmente secuenciados de la base de datos de NCBI (hasta diciembre del 2003). La búsqueda de σ^{54} y de probables EBPs se hizo utilizando el programa tblastn (PROTEIN query-TRANSLATED database) de NCBI (NCBI, <http://www.ncbi.nlm.nih.gov/>) con una matriz de sustitución BLOSUM62 y con un valor de corte de expectancia de 0.01. Los datos obtenidos de las probables EBPs fueron analizados usando bases de datos [PRODOM (<http://protein.toulouse.inra.fr/prodom/current/html/home.php>), SMART (<http://smart.embl-heidelberg.de/>) y PROSITE (<http://cn.expasy.org/prosite/>)], para determinar los dominios funcionales de estas proteínas. Además se realizó un alineamiento múltiple de secuencia del dominio central de estas proteínas y se generó por el método de Neighbour Joining (NJ) un árbol con el alineamiento usando la paquetería de CLUSTAL_X (MAC OSX). El árbol obtenido fue desplegado con el programa TreeView PPC (versión 1.6.5).

Análisis de secuencias de TyrR y PhhR.

Las secuencias de las proteínas TyrR de *E. coli* y PhhR de *P. aeruginosa* fueron obtenidas en la base de datos del National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>) y se analizaron a nivel de aminoácidos para identificar la identidad presente entre ambas proteínas y las regiones de conservación, usando los programas de alineamiento múltiple de CLUSTAL_X (MAC OSX) y de alineamiento de dos secuencias (bl2seq) de la página del NCBI.

Cepas de E. coli usadas en éste estudio.

Todas las construcciones de plásmidos fueron electroporadas y propagadas en *E. coli* JM101 [*F'* *traD36 proA⁺B⁺ lacI^f Δ(lacZ)M15/Δ(lac-proAB) glnV thi*] y MC1061 [*F* *araD139 Δ(ara-leu)7696 galE15 galK16 Δ(lac)X74 rpsL (Str^R) hsdR2 (r_k⁻ m_k⁺) mcrA mcrB1*]. Los ensayos de actividad de β-galactosidasa fueron realizados en *E. coli* JM101,

MC1061, MC1061 Δ *rpoN* (con las mismas características que MC1061 pero con la eliminación precisa del gen *rpoN*) y TH1 [*thi endA hsr Δ lacU169 hutC_{klebs} rpoN*]

Construcciones TyrRcentralWT y TyrRcentralGAFTGA.

En un estudio previo se generaron por PCR las siguientes construcciones: TyrRcentralWT que presenta el dominio central silvestre (Q191-R467) de TyrR de *E. coli* y TyrRcentralGAFTGA que contiene el mismo dominio central pero con una inserción en la región C3 que incorpora los residuos consenso para la familia de las EBPs (GAFTGA). Ambas construcciones fueron clonadas como péptidos de fusión en los sitios de restricción *EcoRI-HindIII* del vector pUC18 Kan^R. La construcción TyrRcentralGAFTGA posteriormente fue sub-clonada en *EcoRI-HindIII* del vector pUC18 Amp^R.

Amplificación de gen phhR.

El fragmento de gen de PhhR que corresponde al dominio central fue amplificado por PCR del genoma de *P. aeruginosa* obtenido de un cultivo lisado por calor. Las condiciones de la reacción fueron: un ciclo 92° C 5 min; 25 ciclos: 92° C 30 seg, 55° C 30 seg, 72° C 1 min; 1 ciclo a 72° C 5 min, utilizando los oligos PhhRKpnI 5' y PhhRXbaI 3' (tabla 2), la mezcla para la reacción de PCR fue la designada por el proveedor (apéndice), adicionando solamente DMSO al 14 % para disminuir la formación de estructura secundaria por el alto contenido de GC. La reacción de PCR se realizó en un Robocycler 40-Stratagene. El producto obtenido fue purificado con el kit de purificación ROCHE (apéndice). Cada oligonucleótido presentó un sitio de restricción (*KpnI* ó *XbaI*) para la clonación como péptido de fusión del fragmento obtenido en la región múltiple de clonación del vector pUC18 digerido previamente *KpnI-XbaI*.

Construcciones TyrR/PhhR y TyrR/PhhR+arg.

Con el uso de la técnica de PCR y con oligonucleótidos mutagénicos (mutagénesis dirigida al sitio) se generaron las construcciones TyrR/PhhR y TyrR/PhhR+Arg a partir de modificaciones puntuales en TyrRGAFTGA y se clonaron como péptidos de fusión en el pUC18. Para la generación de TyrR/PhhR, primero se amplificaron dos fragmentos con el uso de los oligonucleótidos reverso-tyrRTyG para el fragmento A y forward-tyrRS para el fragmento B (fig. 4). Los oligonucleótidos tyrRTyG y tyrRS (tabla 2) presentan los cambios a nivel de nucleótidos para modificar la región C3 consenso (ESELFGHEKGAFTGATAL) a la secuencia de PhhR (ETELFGYGPGAFEGATAL) y una

región donde pueden aparearse entre ellos. Los residuos en cursiva son las posiciones consenso para las EBPs que fueron introducidos en TyrRcentralGAFTGA. Con tyrRTyG-forward se amplificó un fragmento de 300 pb, mientras que con tyrRS se amplificó un fragmento de aproximadamente 650 pb. Las condiciones de la reacción de PCR para los fragmentos A y B fueron: un ciclo 92° C 5 min; 25 ciclos: 92° C 30 seg, 55° C 30 seg, 72° C 1 min; un ciclo a 72° C 5 min, en la mezcla de reacción para PCR. Ambos productos de PCR fueron purificados usando el Kit de ROCHE (apéndice). La secuencia se reconstruyó en dos pasos. Ambos fragmentos fueron alineados en la región complementaria y extendidos con 2u de *Taq* polimerasa (Paso 1: extensión: un ciclo 92° C 5 min; 5 ciclos: 92° C 30 seg, 55° C 1 min, 72° C 1 min; 1 ciclo a 72° C 5 min, en la mezcla típica para PCR sin oligonucleótidos) para generar un fragmento único de aproximadamente 950 pb. Este fragmento inmediatamente fue amplificado por PCR usando solamente los oligonucleótidos externos forward y reverse para pUC18 (Paso 2: amplificación: un ciclo 92° C 5 min; 25 ciclos: 92° C 30 seg, 55° C 30 seg, 72° C 1 min; un ciclo a 72° C 5 min, adición de oligonucleótidos externos). Este producto de PCR fue digerido *EcoRI-HindIII* y ligado en pUC18 para posteriormente secuenciarlo y hacer los ensayos de actividad de β -galactosidasa *in vivo* de esta nueva construcción.

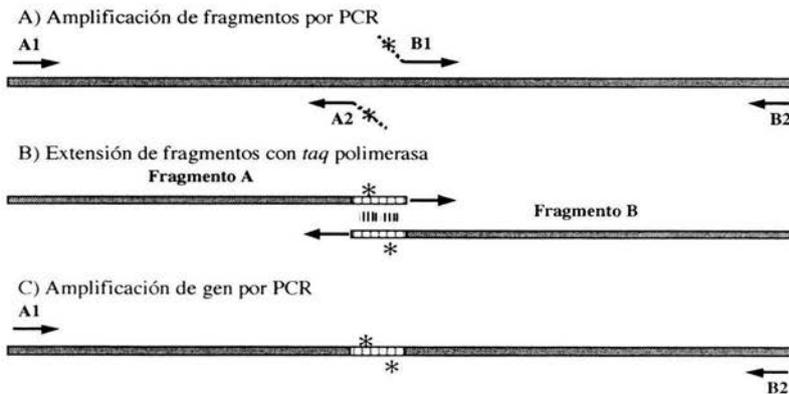


Figura 4. Mutagénesis dirigida al sitio de TyrR/PhhR y TyrR/PhhR+Arg. A) Por PCR y con el uso de oligonucleótidos mutagénicos (A2 y B1) se generan dos fragmentos con una región de complementación entre ellos. B) Las regiones complementarias son alineadas y extendidas con *Taq* polimerasa para reconstruir el gen con los cambios deseados. C) El gen reconstruido es sometido a otra ronda de amplificación por PCR utilizando los oligonucleótidos externos solamente.

De la misma forma en que se hizo esta construcción fue adicionada la arginina presente en la región C3 de PhhR para generar la construcción TyrR/PhhR+Arg. Las condiciones de amplificación por PCR y de reconstrucción del fragmento completo fueron similares, y sólo se utilizaron los oligonucleótidos TyrRc3c4A*-forward para el fragmento A1 y TyrRc3c4A1-reverso para el fragmento B1 (tabla 2, fig. 4). Las construcciones obtenidas también fueron secuenciadas y se les determinó la actividad transcripcional por ensayos de actividad de β -galactosidasa *in vivo*.

Tabla 2. Oligonucleótidos*.

Clave oligo	Secuencia 5' - 3'	Observaciones
PhhRkpnI5'	AGCTCGGTACCCCGCATCTACCATGTGCGCA	Amplificación de dominio central de PhhR <u>Sitio KpnI</u>
PhhRXbaI3'	CGACTCTAGATTACGCCGACGATGCGTCGAG	Amplificación de dominio central de PhhR <u>Sitio XbaI</u>
Foward	GTTTTCCCAGTCACGAC	Amplificación de región de polynker de pUC18
Reverso	AGCGGATAACAATTTACACAGGA	Amplificación de región de polynker de pUC18
TyrRTyG	CGGACCATAACCAACAGTTCCGTCTCGACCGCATC K/PE/GH/Y S/T	Amplificación de frag. A modificación de región C3 (S/T)(H/Y)(E/G)(K/P)
TyrRS	GGTTATGGTCCGGGCGCGTTTGAAGGCGCGACCGCG T/S	Amplificación de frag. B modificación de región C3 (T/S)
TyrRc3c4A1	CGCCCGCGGGCAAGCTCGGATTCTTTGAGCAGGCC	Amplificación de frag. A1 modificación de región C3+R (FPEQAN/LLELTA)
TyrRc3c4A*	GAGCTTGCCCTCCGGGCGCGCCCTTCAAACGCGCC	Amplificación de frag. B1 modificación de región C3+R (-TALKK/RPEGKL)
TyrRocr	GCCAAGCTTAAACGCGCCCGGACC	Amplificación de frag TyrRZ Codón de término (ocre) <u>Sitio HindIII</u>
PspFeco	ACGAATTCGATGGCAGAATACAAAGAT	Amplificación de frag PspFZ <u>Sitio EcoRI</u>
PspPhind	GCCAAGCTTAAACGCCCCGCTTC	Amplificación de frag TyrRZ Codón de término (ocre) <u>Sitio HindIII</u>
rpoNA1	AGGATCCCCGGGCGCCTCAGCGTTTACGATAACC	Amplificación de frag. A2 Eliminación de <i>rpoN</i> <u>Sitio SmaI</u>
rpoNA2	TCCTTATCGGTTGGGGTTCAGAATCGTACTCTCCT	Amplificación de frag. A2 Eliminación de <i>rpoN</i>
rpoNB1	ACGATTCTGAACCCCAACCGATAAGGAAGACACTA	Amplificación de frag. B2 Eliminación de <i>rpoN</i>
rpoNB2	TCTAGAGGATCCAGTGACGGGCTTCGCGTACA	Amplificación de frag. B2 Eliminación de <i>rpoN</i> <u>Sitio BamHI</u>
B7	GAAGAATTCAGACTCTGATAGG	Amplificación de 5' (N-terminal) de <i>rpoN</i> de <i>K. pneumoniae</i>
W302	CACTCGCCAGTTTGATCGACT	Amplificación de 5' (N-terminal) de <i>rpoN</i> de <i>K. pneumoniae</i>

*Todos los oligonucleótidos fueron sintetizados en la unidad de síntesis del Instituto de Biotecnología de la UNAM.

Determinación de la Actividad de β -galactosidasa in vivo.

Todas las actividades de β -galactosidasa *in vivo* reportadas en este trabajo fueron realizadas con el uso de los reporteros que se muestran en la tabla 3, de la siguiente manera: a 0.1 ml de cultivo (LB ó M9 a una OD₆₀₀ entre 0.4 y 0.6) se le adicionaron 0.9 ml de buffer Z (apéndice). Se agregaron 50 μ l de cloroformo y se agitó fuertemente (vortex) para lisar las células. Posteriormente, se añadieron 100 μ l del sustrato o-nitrofenil- β -D-galactopiranosido (ONPG) 8 mg/ml para la β -galactosidasa y esta reacción se incubó a temperatura ambiente. Finalmente, se detuvo la reacción con la adición de 0.5 ml de Na₂CO₃ 1M. El tubo de reacción fue centrifugado a 14 000 rpm durante un minuto y se determinó la OD₄₂₀ para estimar la actividad de β -galactosidasa. Para cada construcción se tomaron tres cultivos independientes. La actividad específica de la β -galactosidasa es reportada como nmoles de ONP producido por minuto en una cantidad determinada de células, y las unidades Miller (Miller, 1972) fueron calculadas como se describe a continuación.

Cálculo de la actividad de β -gal:

$$\beta\text{-gal Miller units} = \text{OD}_{420} / (\text{tiempo} \times \text{vol} \times \text{OD}_{600} \text{ células})$$

Donde:

tiempo: es el tiempo de la reacción dado en minutos.

vol: es el volumen de cultivo celular utilizado en ml (0.1 ml).

OD₄₂₀ es la densidad óptica de la reacción (coloración amarilla).

OD₆₀₀ es la densidad del cultivo celular.

Tabla 3. Plásmidos.

Plásmidos (reporteros)	Características	Cita
psU003	Región promotora <i>nifH</i> de <i>Rhizobium meliloti</i> (sin UAS) fusionada a <i>lacZ</i> , Cm ^r , originado de pACYC184	(Gonzalez et al., 1998)
pcU101	Región promotora <i>nifH</i> de <i>Rhizobium meliloti</i> (con UAS) fusionada a <i>lacZ</i> , Cm ^r , originado de pACYC184	(Gonzalez et al., 1998)
pMB210	Región promotora <i>nifH</i> de <i>Rhizobium meliloti</i> completa fusionada a <i>lacZ</i> , Tc ^r , originado de pRK290	(Ditta et al., 1985)
pMB210-UAS	Región promotora <i>nifH</i> de <i>Rhizobium meliloti</i> (sin UAS) fusionada a <i>lacZ</i> , Tc ^r , originado de pMB210	En este trabajo
pRT22	Región promotora <i>nifH</i> de <i>Klebsiella pneumoniae</i> fusionada a <i>lacZ</i> , Cm ^r , originado de pACYC177	(Tuli and Merrick, 1988)
pro310	Región promotora σ 70 <i>ompS1</i> de <i>Salmonella typhi</i> fusionada a <i>lacZ</i> , Tc ^r , originado de pro521	(Oropeza et al., 1999)

Tabla 3. Plásmidos continuación.

Plásmidos (Construc.)	Características	Cita
pUC18	Vector de clonación Amp ^R	(Vieira and Messing, 1982)
TyrRcentralWT	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de TyrR silvestre.	(Flores H. Datos no publicados)
TyrRcentral GAFTGA	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de TyrR con región C3 consenso (GAFTGA).	(Flores H. Datos no publicados)
TyrR/PhhR	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de TyrR con región C3 de PhhR.	En este trabajo
TyrR/PhhR+Arg	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de TyrR con regiones C3 y C4 de PhhR.	En este trabajo
TyrR191-285	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de TyrR con codón de término en región C3.	En este trabajo
M1-M5	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de mutantes TyrR/PhhR M1 a M5.	En este trabajo
pUC18PhhR	Derivado de PUC18 con inserto <i>KpnI-XbaI</i> . Dominio central de PhhR silvestre.	En este trabajo
PUCTyrRZ	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de TyrR hasta la región C3.	En este trabajo
PUCPspFZ	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Dominio central de PspF hasta la región C3.	En este trabajo
pUC18-MTIM	Derivado de PUC18 con inserto <i>EcoRI-HindIII</i> . Proteína mono TIM de <i>Trypanosoma brucei</i> .	(datos no publicados)
PM1-k	pWKS130 con sitio <i>KpnI</i> removido y <i>nifA</i> de pRJ7511 clonado con <i>BamHI-PstI</i> , Km ^R .	(Grande et al., 1999)
pmJ15	Derivado de pQE-32 con promotor T5. Inserto <i>NlaIII-HindIII</i> con dominio central de PspF (6XHisPspFAHTH)	(Jovanovic et al., 1999)
pKK223-3	Vector de clonación Km ^R	Pharmacia
pMM70	Derivado de PTZ19R con <i>rpoN</i> de <i>Klebsiella pneumoniae</i> expresado con su propio promotor	(Merrick and Gibbins, 1985)
pK03	Origen de replicación M13. Genes <i>sacB</i> (levan sacarasa) y <i>cat</i> (cm ^R). Región múltiple de clonación.	(Link et al., 1997)
pK03-RPON	Derivado de pK03 con inserto <i>BamHI-SmaI</i> en región múltiple de clonación. Regiones adyacentes de <i>rpoN</i> de <i>E. coli</i> .	En este trabajo

Construcción de los fragmentos N-terminal del dominio central de TyrR y PspF.

Con el uso de PCR y con oligonucleótidos mutagénicos se crearon las construcciones pUCTyrRZ y pUCPspFZ. En pUCTyrRZ, se generó un codón de termino en la región C3 (de GAFEGA a GAFZGA) utilizando el templado TyrR/PhhR y los oligonucleótidos reverse-TyrRocr. El oligonucleótido TyrRocr presenta el codón UAA (terminador ocre, tabla 2) seguido de un sitio *HindIII*. La reacción de amplificación del fragmento se desarrolló bajo las siguientes condiciones: un ciclo 92° C 5 min; 25 ciclos 92° C 30 seg, 55° C 30 seg, y 72° C 1 min; un ciclo a 72° C por 5 min. La reacción fue con la mezcla para PCR más la adición de DMSO al 10 %. La amplificación generó un fragmento que codifica parte de la región N-terminal de TyrR/PhhR y del dominio central de aproximadamente 300 pb que fue digerido *EcoRI-HindIII* para clonarlo como péptido de fusión en el vector pUC18 digerido previamente. De la misma forma, se construyó pUCPspFZ generando un fragmento de la región amino-terminal de la proteína PspF similar al de pUCTyrRZ. El templado utilizado para la amplificación de este fragmento fue el vector pMJ15 que presenta el dominio central de PspF de *E. coli* y los oligonucleótidos usados fueron PspFeco-PspFhind (tabla 2). Las condiciones de amplificación fueron: un ciclo 92° C 5 min; 25 ciclos 92° C 30 seg, 55° C 30 seg, 72° C 1 min; un ciclo a 72° C por 5 min, usando la mezcla común para la reacción de PCR (apéndice). El producto obtenido de aproximadamente 300 pb fue digerido *EcoRI-HindIII*, para ser clonado también como péptido de fusión en pUC18. Todos los fragmentos y vectores construidos fueron purificados con el kit de purificación ROCHE. A las nuevas construcciones también se les evaluó la actividad de β -galactosidasa *in vivo*.

Generación de banco mutagénico TyrR/PhhR.

Con el uso de la técnica de PCR mutagénica (Cadwell and Joyce, 1992) se mutagenizó la construcción TyrR/PhhR que contiene la región C3 del dominio central de PhhR. La reacción de PCR se realizó con una mezcla especial para PCR mutagénica y bajo las siguientes condiciones: un ciclo 92° C 5 min; 25 ciclos: 92° C 30 seg, 55° C 30 seg, 72° C 1 min; 1 ciclo a 72° C 5 min en un Robocycler 40-Stratagene. La reacción de 50 μ l contenía 7mM de MgCl₂, 0.5mM de MnCl₂, dNTPs desbalanceados (0.2mM dGTP, 0.2mM dATP, 1mM dCTP y 1mM dTTP), buffer Gibco 1X para PCR^{Mg} (apéndice), 10 pmol de cada uno de los oligonucleótidos (forward y reverse especiales para pUC18 por alinear con

las regiones externas a la zona de clonación múltiple), H₂O, *Taq* polimerasa 2 u y ADN templado (50 ng). El producto de PCR fue digerido *EcoRI-HindIII* y purificado con el Kit ROCHE para ser ligado (apéndice) en el vector pUC18. Células de la cepa MC1061 fueron electrotransformadas con el producto de la ligación. El banco fue recuperado y amplificado en cultivos con LB/Amp. Finalmente se purificó plásmido y se digirió con *EcoRI-HindIII* para ver la presencia del inserto. Además, se mandaron a secuenciar 10 clonas para verificar la tasa de mutagénesis de la reacción de PCR mutagénica.

Método de Tamizado (reportero pRT22).

El banco obtenido de la PCR mutagénica fue electroporado en la cepa MC1061 que contenía el plásmido pRT22. El vector pRT22 presenta la región de control de *nifH* de *Klebsiela pneumoniae* fusionada al gen *lacZ* y la activación de este promotor permite expresar la β-galactosidasa para la degradación del sustrato cromogénico X-Gal (5-bromo-4-cloro-3-indolil-β-D-galactosido). Las células recuperadas de la electroporación que contenían el banco mutagénico fueron tamizadas en cajas con medio M9/antibióticos/X-gal (apéndice). Las colonias que presentaron una mayor coloración azul fueron seleccionadas. Posteriormente, se generaron mezclas de las mutantes previamente seleccionadas para volver a tamizar de la misma manera y verificar el fenotipo observado. Las candidatas seleccionadas fueron secuenciadas y sometidas a ensayos de actividad de β-galactosidasa *in vivo* para verificar lo observado en el tamizado.

Construcción de la cepa MC1061ΔrpoN.

La eliminación del gen *rpoN* de *Escherichia coli* MC1061 se obtuvo por una doble recombinación homóloga, con el uso del vector suicida pK03 (fig. 5) (Link *et al.*, 1997). Para esto, se diseñaron dos juegos de oligonucleótidos para generar por PCR dos fragmentos que corresponden a las secuencias adyacentes 5' (fragmento A2: oligos rpoNA1 y rpoNA2, tabla 2) y 3' (fragmento B2: oligos rpoNB1 y rpoNB2, tabla 2) del gen *rpoN* de *E. coli* (condiciones de reacción de PCR para fragmentos A2 y B2: un ciclo 92° C 5 min; 25 ciclos: 92° C 30 seg, 55° C 30 seg, 72° C 1 min; un ciclo a 72° C 5 min, en la mezcla común para PCR). Cada fragmento fue de aproximadamente 500 pb y presentaba en los extremos un sitio de restricción (según sea el caso *BamHI* ó *SmaI*) y una región que complementaba aproximadamente con 15 pb del otro fragmento generado. Ambos fragmentos fueron apareados en dicha región complementaria al disminuir la temperatura de 92° C a 55° C y

extendidos con *Taq* polimerasa (Paso 1: extensión: un ciclo 92° C 5 min; 5 ciclos: 92° C 30 seg, 55° C 1 min, 72° C 1 min; un ciclo a 72° C 5 min, en la mezcla típica para PCR sin oligonucleótidos) para generar un fragmento único de 1000 pb con los sitios *Bam*HI y *Sma*I en los extremos. Este fragmento fue amplificado inmediatamente por PCR usando solamente los oligonucleótidos externos rpoNA1 y rpoNB2 (Paso 2: amplificación: un ciclo 92° C 5 min; 25 ciclos: 92° C 30 seg, 55° C 30 seg, 72° C 1 min; un ciclo a 72° C 5 min, adicionando de oligonucleótidos externos). El producto obtenido fue digerido con *Bam*HI-*Sma*I para posteriormente clonarlo en el vector pK03 previamente digerido con las mismas enzimas. El vector pK03 contiene una región de clonación múltiple y un origen de replicación sensible a temperatura (42° C), además de los genes *sacB* y *cat* para la levadura de sacarosa y para la cloranfenicol acetil transferasa respectivamente (Link *et al.*, 1997). El nuevo vector obtenido (pK03-RPON) con las secuencias adyacentes del gen *rpoN* de *E. coli* se introdujo por electroporación en células de *E. coli* cepa MC1061 para inducir el evento de doble recombinación. Posteriormente, una parte de las células electrotransformadas fueron sembradas en cajas de LB/Cm e incubadas a 42° C, mientras que otra parte, fue también crecida en cajas de LB/Cm, pero incubadas a 30° C para medir la frecuencia de la integración por recombinación del vector en el cromosoma durante el crecimiento a 42° C. De las cajas crecidas a 42° C, se seleccionaron de 1-5 colonias y se inocularon en 1 ml de medio LB líquido permitiendo su crecimiento durante 2 horas a 30° C. Este cultivo fue diluido y sembrado en cajas de LB/sacarosa al 5% para seleccionar aquellas clonas que no porten el plásmido. Finalmente, las colonias obtenidas en las placas fueron picadas en cajas de LB/sacarosa 5%, LB/Cm y en medio W-Arg para identificar las clonas que hubieran perdido el gen del cromosoma y no presentaran el vector pK03. La eliminación fue confirmada por PCR y por secuencia de la región cromosomal utilizando varios juegos de oligonucleótidos (oligonucleótidos para las regiones adyacentes del gen *rpoN* y oligonucleótidos para el reconocimiento de secuencia del gene *rpoN* de *Klebsiella pneumoniae*).

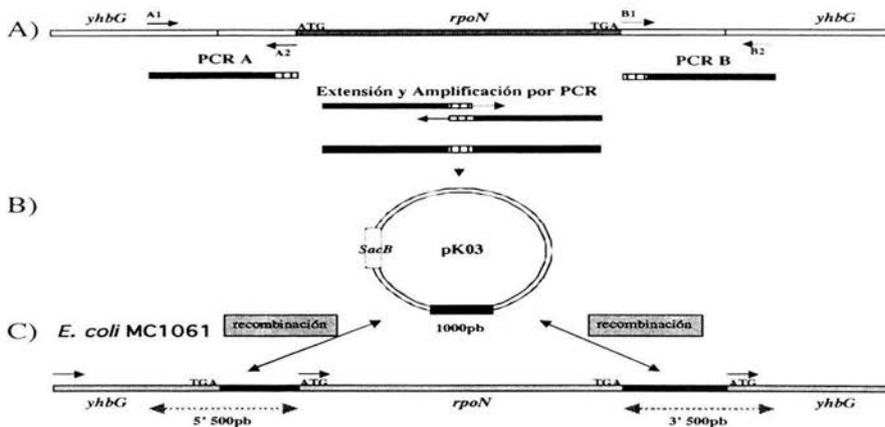


Figura 5. Eliminación del gen *rpoN* con el uso del vector pK03. A) Amplificación por PCR de las regiones 5' y 3' del gen *rpoN* para su alineamiento y extensión (fig. 4), para crear un fragmento mayor. B) Clonación del fragmento en el sitio de clonación múltiple de pK03. C) Transformación de la nueva construcción en *E. coli* MC1061 y facilitar la doble recombinación para la eliminación del gen.

RESULTADOS.

I. Análisis *in silico* del sistema σ^{54} .

Resultados y Discusión.

El rápido avance en la secuenciación genómica de los organismos, ha proporcionado una gran cantidad de datos que pueden ser explotados en una infinidad de áreas biológicas. Utilizando esta información buscamos determinar probables eventos de reclutamiento de EBPs, similar al ocurrido en TyrR de *Escherichia coli*. Para ello, realizamos un análisis *in silico* de los elementos que participan en la regulación transcripcional de genes dependientes del factor σ^{54} (σ^{54} y EBPs) en los 145 genomas bacterianos completamente secuenciados y depositados hasta diciembre del 2003 en la base de datos del National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>).

De la base de datos se eliminaron las secuencias que pertenecían a serotipos de la misma especie evitando una posible duplicación de información, reduciéndola finalmente a 17 arqueas y 100 eubacterias (tabla 4). Sobre estos 117 genomas, se hizo una búsqueda de probables σ^{54} y EBPs, utilizando como elementos de búsqueda RpoN de *E. coli* y el dominio central de NtrC de *Rhizobium meliloti* (L139-A367) respectivamente. La búsqueda fue realizada con el programa tblastn (PROTEIN query-TRANSLATED database) del

NCBI, con una matriz de sustitución BLOSUM62 y con un valor mínimo de expectativa de 0.0001. Con los resultados obtenidos se hicieron nuevas búsquedas con otras secuencias de EBPs y de RpoNs con el fin de verificar lo observado inicialmente.

Los resultados de la búsqueda se observan en la tabla 4. En Arqueas no se encuentra presente el sistema σ^{54} , mientras que en eubacterias, 58 organismos presentan los dos elementos requeridos en la búsqueda con un número diverso de probables EBPs. Dentro de este grupo de organismos, 7 contienen más de un gene *rpoN*. También se puede observar que en 10 organismos se encuentra presente el par funcional (σ^{54} -EBPs) con un sólo miembro de cada elemento. En *Neisseria meningitidis*, que se abordará independientemente, sólo se observaron restos de elementos que muy probablemente pertenecían a un sistema σ^{54} y la evidencia de un evento de reclutamiento. En los restantes 41 genomas de eubacterias donde no está presente σ^{54} , se observaron 12 genomas donde la búsqueda localizó 14 probables EBPs.

Tabla 4. Elementos del sistema σ^{54} en los genomas totalmente secuenciados.

	EBPs	<i>rpoNs</i>		EBPs	<i>rpoNs</i>
ARCHEA			Proteobacteria		
17 genomes	---	---	alpha subdivision		
BACTERIAL			Others		
Actinobacterial			<i>Bradyrhizobium japonicum</i> USDA 110 *	11	2
<i>Bifidobacterium longum</i> NCC2705	---	---	<i>Brucella melitensis</i> 16M	3	1
<i>Corynebacterium diphtheriae</i>	---	---	<i>Brucella suis</i> 1330	3	1
<i>Corynebacterium efficiens</i> YS-314	---	---	<i>Caulobacter crescentus</i> CB15	4	1
<i>Corynebacterium glutamicum</i> ATCC 13032	---	---	<i>Mesorhizobium loti</i> *	7	2
<i>Mycobacterium bovis</i> subsp. <i>bovis</i> AF2122/97	---	---	Rhizobiaceae		
<i>Mycobacterium leprae</i>	---	---	<i>Azrobacterium tumefaciens</i> str. C58 (Cereon)	6	1
<i>Mycobacterium tuberculosis</i> CDC1551	---	---	<i>Sinorhizobium meliloti</i>	7	1
<i>Streptomyces avermitilis</i> MA-4680 ***	1	---	Rickettsiales		
<i>Streptomyces coelicolor</i> A3(2) ***	1	---	<i>Rickettsia conorii</i> ***	1	---
<i>Tropheryma whipplei</i> TW08/27	---	---	<i>Rickettsia prowazekii</i> ***	1	---
Chlamydiae			beta subdivision		
<i>Chlamydia muridarum</i> **	1	1	Bordetella		
<i>Chlamydia trachomatis</i> **	1	1	<i>Bordetella bronchiseptica</i> *	6	2
<i>Chlamydophila caviae</i> GPIC **	1	1	<i>Bordetella parapertussis</i> *	6	2
<i>Chlamydophila pneumoniae</i> AR39 **	1	1	<i>Bordetella pertussis</i>	3	1
Cyanobacteria			Burkholderiaceae		
<i>Gloeobacter violaceus</i> ***	1	---	<i>Ralstonia solanacearum</i>	13	1
<i>Nostoc</i> sp. PCC 7120 ***	1	---	Neisseriaceae		
<i>Prochlorococcus marinus</i> str. MIT 9313 ***	1	---	<i>Chromobacterium violaceum</i> ATCC 12472	8	1
<i>Synechococcus</i> sp. WH 8102	---	---	<i>Neisseria meningitidis</i> MCS8 ***	1	1
<i>Synechocystis</i> sp. PCC 6803 ***	1	---	Others		
<i>Thermosynechococcus elongatus</i> BP-1	---	---	<i>Nitrosomonas europaea</i> ATCC 19718	3	1
Firmicutes			epsilon subdivision		
Bacillales			<i>Campylobacter jejuni</i> subsp. <i>jejuni</i> **	1	1
<i>Bacillus anthracis</i> str. Ames	6	1	<i>Helicobacter hepaticus</i> ATCC 51449 **	1	1
<i>Bacillus cereus</i> ATCC 14579	7	1	<i>Helicobacter pylori</i> 26695 **	1	1
<i>Bacillus halodurans</i>	9	1	<i>Wolfinella succinogenes</i> *	2	2
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	4	1	gamma subdivision		
<i>Listeria innocua</i>	3	1	Enterobacteriales		
<i>Listeria monocytogenes</i> EGD-e	3	1	<i>Buchnera aphidicola</i> str. APS (Acyrthosiphon pisum)	---	---
<i>Oceanobacillus iheyensis</i> HTE831	5	1	<i>Candidatus Blochmannia floridanus</i>	---	---
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2	---	---	<i>Wigglesworthia glossinidia</i> endosymbiont of Glossi	---	---

Tabla 4. Elementos del sistema σ^{54} en los genomas totalmente secuenciados (continuación).

<i>Staphylococcus epidermidis</i>	---	---	<i>Escherichia coli</i> K12	12	1
Clostridia			<i>Photobacterium luminescens</i> subsp. laumondii T101	6	1
<i>Clostridium acetobutylicum</i>	3	1	<i>Salmonella enterica</i> subsp. enterica serovar Typhi	12	1
<i>Clostridium perfringens</i> str. 13 **	1	1	<i>Salmonella typhimurium</i> LT2	15	1
<i>Clostridium tetani</i> F88	6	1	<i>Shigella flexneri</i> 2a str. 2457T	9	1
<i>Thermoanaerobacter tengcongensis</i>	10	1	<i>Yersinia pestis</i> KIM	5	1
Lactobacillales			Others		
<i>Enterococcus faecalis</i> V583	5	1	<i>Coxiella burnetii</i> RSA 493	---	---
<i>Lactobacillus plantarum</i> WCFS1 **	1	1	<i>Shewanella oneidensis</i> MR-1	11	1
<i>Lactococcus lactis</i> subsp. lactis	---	---	Pasteurellaceae		
<i>Streptococcus agalactiae</i> 2603V/R	---	---	<i>Haemophilus ducreyi</i> 35000HP ***	1	---
<i>Streptococcus mutans</i> UA159	---	---	<i>Haemophilus influenzae</i> Rd KW20 ***	1	---
<i>Streptococcus pneumoniae</i> R6	---	---	<i>Pasteurella multocida</i> ***	2	---
<i>Streptococcus pyogenes</i> M1 GAS	---	---	Pseudomonadaceae		
Mollicutes			<i>Pseudomonas aeruginosa</i> PAO1	22	1
<i>Mycoplasma gallisepticum</i> R	---	---	<i>Pseudomonas putida</i> KT2440	22	1
<i>Mycoplasma genitalium</i>	---	---	<i>Pseudomonas syringae</i> pv. tomato str. DC3000	23	1
<i>Mycoplasma penetrans</i>	---	---	Vibrionaceae		
<i>Mycoplasma pneumoniae</i>	---	---	<i>Vibrio cholerae</i> O1 biovar eltor str. N16961	14	1
<i>Mycoplasma pulmonis</i>	---	---	<i>Vibrio parahaemolyticus</i> RIMD 2210633	17	1
<i>Ureaplasma urealyticum</i>	---	---	<i>Vibrio vulnificus</i> YJ016	14	1
Others			Xanthomonadaceae		
<i>Aquifex aeolicus</i> VF5	6	1	<i>Xanthomonas axonopodis</i> pv. citri str. 306 *	9	2
<i>Bacteroides thetaiotaomicron</i> VPI-5482	9	1	<i>Xanthomonas campestris</i> pv. campestris str. ATCC 33913 *	8	2
<i>Chlorobium tentidum</i> TL5	3	1	<i>Xylella fastidiosa</i> Temecula1	2	1
<i>Deinococcus radiodurans</i> R1	---	---	Spirochaetales		
<i>Fusobacterium nucleatum</i> subsp. nucleatum ***	2	---	<i>Borrelia burgdorferi</i> B31 **	1	1
<i>Pirellula</i> sp. 1	16	1	<i>Leptospira interrogans</i> serovar lai str. 56601	2	1
<i>Porphyromonas gingivalis</i> W83	3	1	<i>Treponema pallidum</i> subsp. pallidum str. Nichols	2	1
<i>Thermotoga maritima</i>	---	---			

Búsqueda de EBP's con el dominio central NtrC de *R. meliloti*, y de σ^{54} con RpoN de *E. coli*. *genomas con dos genes *rpoNs*. **genomas con el par funcional σ^{54} -EBP. ***genomas con probable evento de reclutamiento de EBP's.

El par funcional σ^{54} -EBP.

A partir de los datos obtenidos en la búsqueda, se seleccionaron y analizaron los casos donde se observó solamente un miembro de cada elemento del par funcional (1-EBP/1- σ^{54}), esto con el fin de identificar los dominios presentes en la EBP y describir las probables características funcionales. De los diez casos descritos en la tabla 4 (no tomando en cuenta el caso de *Neisseria meningitidis*), cuatro tienen los tres dominios característicos para las EBP's (N-terminal, central y C-terminal; fig. 6). Las proteínas HH1635 (FlgR) de *Helicobacter hepaticus*, CJ1024C de *Campylobacter jejuni* y BB0763 de *Borrelia burgdorferi* presentan un arreglo similar a NtrC que es miembro del sistema de dos componentes. Un sistema de dos componentes permite a la bacteria detectar responder adecuadamente tanto a señales de estrés como a ambientales, y un buen representante de estos casos es el sistema NtrB/NtrC. La proteína NtrC que regula la asimilación de nitrógeno, es activada por una histidina cinasa (NtrB), mediante la fosforilación de un residuo de aspartato presente en el dominio sensor. Esta fosforilación permite la oligomerización necesaria para la activación transcripcional (Austin and Dixon, 1992; Mettke *et al.*, 1995). La proteína CPE2358 de *Clostridium perfringens* es la única de este

grupo con dos dominios en su N-terminal, uno conocido como PAS, similar al presente en la proteína TyrR de *E. coli* y que es utilizado para detectar metabolitos del medio, mientras que el otro asignado como PTS-HP_r, que es un dominio de fosforilación (Studholme and Dixon, 2003).

Por otro lado, existen seis casos donde no se observa el dominio de unión a ADN (C-terminal, HVH), cuatro se encuentran distribuidos en Clamidiae (CP0162 de *Chlamydomphila pneumoniae*, CCA00155 de *Chlamydomphila caviae*, TC0753 de *Chlamydia muridarum* y CT468 de *Chlamydia trachomatis*), mientras que los otros dos son HP0703 de *Helicobacter pylori* y lp_0585 de *Lactobacillus plantarum* (fig. 6).

En el grupo de proteínas de Clamidiae (4 miembros) se observó una identidad alta entre ellas (66 % mínimo en proteína completa, tabla 5), indicando que muy probablemente son genes ortólogos. Los genes ortólogos son genes homólogos que se encuentran realizando la misma función en distintos organismos (Gogarten, 1994). Además, igual que las EBPs identificadas en *H. hepaticus* (HH1635), *C. jejuni* (CJ1024C) y *B. burgdorferi* (BB0763), presentan un arreglo de sus dominios similar a NtrC. En cuanto a la proteína lp_0585 de *L. plantarum*, se observó que presenta el dominio central menos conservado, además de no tener un dominio regulador en el N-terminal. Sin embargo, esta proteína probablemente pueda seguir detectando a través de un dominio de fosforilación (PRD) ubicado en la región C-terminal, disposición no común para las EBPs (Studholme and Dixon, 2003).

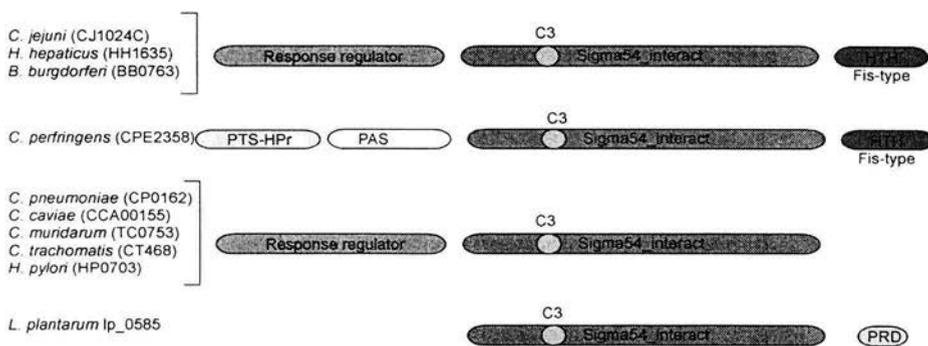


Figura 6. Disposición de dominios de las EBPs únicas presentes en el par funcional. Con el análisis de dominios funcionales (bases de datos SMART, Prodom, Prosite), se observó que las mayoría de las EBPs descritas aquí presentan un dominio N-terminal de tipo del sistema de dos componentes (response regulator). Además, se observa que en el grupo de clamidias, las EBPs no presentan el dominio de unión a ADN, indicando una probable activación *in trans* llevada por estas proteínas. Finalmente en *L. plantarum* la EBP presenta un arreglo poco común para esta familia de activadores transcripcionales.

Tabla 5. Matriz de identidad de secuencia del grupo de clamidias, proteína completa (programa BioEdit Sequence Alignment Editor).

Seq->	CT468	TC0753	CCA00155	CP0162
CT468	1.000	0.896	0.720	0.659
TC0753	---	1.000	0.723	0.668
CCA00155	---	---	1.000	0.769
CP0162	---	---	---	1.000

Entre los casos que más despertaron nuestra atención, se encuentran el de *H. pylori* y los del grupo de clamidias, ya que aún con la pérdida del dominio C-terminal (fig. 6), situación curiosa en las EBPs, tienen la posibilidad de activar en *trans* a promotores σ^{54} (Mathews and Timms, 2000; Spohn and Scarlato, 1999). Existe evidencia de que sólo el dominio central de NifA, una EBP que activa genes para la fijación de nitrógeno en *Rhizobium meliloti*, puede activar la expresión en *trans* de genes dependientes del sistema σ^{54} (Huala and Ausubel, 1989). El dominio C-terminal, le podría conferir especificidad a las EBPs cuando se encuentran coexistiendo varias EBPs dentro de un organismo, como es el caso de *E. coli* que tiene 12 activadores de esta familia. Nosotros sugerimos que probablemente la presencia de este dominio C-terminal en las EBPs únicas, hasta cierto punto sería innecesaria, por la capacidad de activación en *trans* del dominio central y por la ausencia de una probable activación cruzada con otros activadores dependientes de σ^{54} inexistentes. Sin embargo, para que se logre una activación eficiente bajo estas características, sugerimos también la presencia de al menos dos elementos: un dominio central que genere una actividad en *trans* muy robusta y promotores fuertes -12 y -24, con el fin de que con la mínima interacción se promueva la transcripción. Juárez *et al* (2002), identificaron por mutagénesis al azar cuatro residuos del dominio central de NifA de *Bradyrhizobium japonicum* (P291S, S356G, N271D y S381G), que al ser modificados incrementaban una activación originalmente disminuida por una deficiencia en el pegado a las UAS. Ellos también observaron que los cambios sencillos y multiples incrementaban de forma distinta la activación transcripcional. En las EBPs del grupo de Clamidias y la de *H. pylori* se observa al menos una posición (S356G) de las cuatro antes mencionadas, además de presentar las regiones del domino central de importancia funcional muy conservadas. También, en *H. pylori* se han localizado cinco operones con promotores -12, -24, en los

cuales, se observan secuencias consenso que indican una fuerte interacción con σ^{54} (Spohn and Scarlato, 1999).

Todas las EBPs descritas en esta sección presentan las regiones altamente conservadas del dominio central, a excepción de lp_0585 de *L. plantarum*, así como los aminoácidos conservados que estudios genéticos han demostrado que son importantes para su función (principalmente regiones C1, C3, C4 y C7) (Gonzalez *et al.*, 1998; Hilgenfeld, 1995; Morett and Segovia, 1993; Perez-Martin and de Lorenzo, 1996; Rombel *et al.*, 1999; Saraste *et al.*, 1990). Esto nos permite sugerir que estos activadores son funcionales y que se mantienen junto con σ^{54} regulando la expresión de genes. Existen datos experimentales en *H. pylori* y *C. jejuni* que confirman dicha funcionalidad. En ambos casos, se observó que la expresión de los operones con promotores σ^{54} se relaciona con la estructuración del flagelo y que mutantes en la EBP, restringen la formación de dicha estructura motriz (Jagannathan *et al.*, 2001; Spohn and Scarlato, 1999).

Estos datos en conjunto sugieren que los factores σ^{54} en estos casos se encuentren probablemente coevolucionado junto con su EBP, incrementando su especificidad por ella y perdiendo la capacidad de interactuar con otras EBPs. En *Rhodobacter sphaeroides* que se vera más adelante existen factores σ^{54} específicos para ciertas EBPs (Poggio *et al.*, 2002).

Más de un elemento del par funcional.

La mayor parte de los genomas analizados, presentan desde 2 hasta 23 EBPs (tabla 4), y sus funciones han sido descritas con mayor detalle por Studholme y Nixon (2003). Además de esta información, los resultados de la tabla 4, muestran a los organismos que presentan varias copias del gen *rpoN* (σ^{54}). De los organismos que presentan el sistema σ^{54} , *Bradyrhizobium japonicum*, *Mesorhizobium loti*, *Bordetella bronchiseptica*, *Bordetella parapertussis*, *Wolinella succinogenes*, *Xanthomonas axonopodis* y *Xanthomonas campestris*, contienen 2 genes *rpoN* sugiriendo la posibilidad, de que puedan estar funcionando independientemente con determinadas EBPs, como es el caso de *Rhodobacter sphaeroides*. En *R. sphaeroides*, un genoma no analizado en este trabajo, se ha observado que al menos 2 de sus 4 copias de *rpoN* son requeridas para expresar de forma independiente grupos de promotores. El gen *rpoN2* regula la expresión de genes del flagelo,

mientras que el responsable de la expresión de genes relacionados con la asimilación de nitrógeno es *rpoN1* (Poggio *et al.*, 2002).

El caso Neisseria.

En el curso de este trabajo, encontramos un caso que nos dio pauta para reconstruir un probable evento evolutivo reciente, que finalizó en la eliminación gradual del sistema σ^{54} y en el reclutamiento de algunos de sus elementos. En el genoma de *Neisseria meningitidis* (tabla 9) se observa la presencia de ambos elementos del par funcional (σ^{54} -EBP), sin embargo, la secuencia de RpoN carece de regiones de importancia funcional indicándonos la inactividad del sistema σ^{54} . Además, en la especie cercana *Neisseria gonorrhoeae* resalta la existencia de promotores σ^{54} funcionales que han sido verificados experimentalmente (Carrick *et al.*, 1997). Dadas estas observaciones, nos propusimos a analizar por separado y más a fondo este caso.

Promotores sigma 54: En *Neisseria meningitidis* hasta ahora no se conoce la existencia de promotores dependientes de σ^{54} . Sin embargo, en *Neisseria gonorrhoeae* la expresión de la proteína Pile, que es una subunidad estructural del pili y que también está presente en *N. meningitidis*, es controlada por un promotor dependiente del factor σ^{70} (P1), sobrelapado con un promotor σ^{54} (P3). El promotor P3 no es activo en *N. gonorrhoeae*, pero en *E. coli* y en *P. aeruginosa* sí es dependiente de σ^{54} (Carrick *et al.*, 1997). Esto nos indica que la expresión del gene *pilE*, necesario para la adherencia específica sobre células de mucosa en humanos, cambió su modo transcripcional a σ^{70} al perderse el sistema σ^{54} . Madan Babu (2003), describe que la pérdida de factores σ dispara la acumulación de pseudogenes y que si algún gene dependiente del factor perdido es esencial, se modifica el promotor para ser reconocido por cualquier otro factor σ , muy probablemente fue lo que sucedió en *N. gonorrhoeae*.

Sigma 54: Existe evidencia experimental de que en las Neisserias se expresa un factor σ^{54} por detección inmunológica usando anticuerpos anti-caja RpoN, que es una región en el C-terminal de la proteína que reconoce la región -24 del promotor, sin embargo no es funcional (Laskos *et al.*, 1998). En la secuencia de la probable RpoN de *N. meningitidis* obtenida en la búsqueda, observamos la conservación de las regiones N-terminal (rica en glutaminas) y caja RpoN, sin embargo, también se observa la eliminación de la región

donde se encuentra el arreglo HVH (hélice-vuelta-hélice), importante para reconocimiento de la región -12 del promotor. En cambio, en la RpoN de *N. gonorrhoeae*, Laskos *et al* (1998) observaron por análisis de secuencia un cambio en el marco de lectura de todo el N-terminal (incluyendo la HVH), además de la eliminación presente en *N. meningitidis*. Estas características imposibilitaron a esta proteína (RpoN *N. gonorrhoeae*) para complementar el sistema σ^{54} de *E. coli* TH1 (*rpoN* interferido por un transposon) (Laskos *et al.*, 1998). Esta información sugiere que la pérdida del sistema σ^{54} producida por la eliminación de parte de la secuencia de σ^{54} en ambas especies, se dio antes de la especiación entre *N. gonorrhoeae* y *N. meningitidis*, y que en *N. gonorrhoeae* posteriormente, se modificó el marco de lectura inicial. Además, es de destacar que ambas proteínas conservan la caja RpoN y probablemente puedan seguir funcionando como represor transcripcional ya que se mantiene su afinidad de unión a ciertas regiones en el ADN.

Proteínas de unión a enhancer: Como resultado de este análisis, encontramos una probable EBP en el genoma de *N. meningitidis* (NMB0115, fig 7; tabla 4) que se expresa bajo ciertas condiciones (Grifantini *et al.*, 2003). La proteína NMB0115 presenta un dominio central poco conservado con regiones funcionales tales como Walker A y la región C7, pero carece por completo de la región C3, indicando la pérdida de su función como activador σ^{54} . Por otro lado, en la especie *N. gonorrhoeae*, se ha descrito que la fusión de una probable EBP con una proteína histidina cinasa (Pils), generó una proteína quimérica conocida como Rsp (regulator-sensor protein) (Carrick *et al.*, 2000). Rsp presenta el dominio transmembranal de Pils fusionado a parte del dominio central (C4 en adelante) y el dominio C-terminal (unión a ADN, HVH) de la EBP, dicha fusión, ejemplifica claramente un evento de reclutamiento de EBPs en esta especie. Las secuencias del dominio central de NMB0115 (*N. meningitidis*) y de la parte correspondiente de Rsp (*N. gonorrhoeae*), resultaron distintas cuando se analizaron por alineamiento múltiple utilizando Clustal_X. Esta información nos indicó que cada organismo reclutó distintas EBPs posterior a la pérdida del sistema σ^{54} .

En este caso en particular, se pudo observar que el sistema σ^{54} sí existió en Neisserias dada la presencia del par funcional y de un tercer elemento que corresponde al promotor -12 y -24, desechando completamente que la EBP localizada fuera un evento aislado de transferencia horizontal. En trabajos posteriores se necesitaría realizar una

búsqueda de promotores σ^{54} en los genomas donde solamente se observaron EBPs, con el fin de determinar cuales organismos llegaron a presentar un sistema σ^{54} y poder asignar cuales genes pudieron estar bajo su control, como lo observado con PilE de *N. gonorrhoeae*.

Reclutamiento de EBPs: evidencia de genomas sin σ^{54} .

A continuación describiremos las características estructurales y funcionales de 14 probables EBPs obtenidas en la búsqueda, en los genomas secuenciados sin σ^{54} (tabla 4). Nosotros sugerimos que estas proteínas sufrieron un evento de reclutamiento que fue disparado por la pérdida del factor σ^{54} , y que su estado funcional probablemente se ha mantenido en base al análisis que a continuación presentamos.

En el genoma del patógeno oportunista *Fusobacterium nucleatum*, que presenta 4 factores σ y 43 reguladores transcripcionales (Kapatral *et al.*, 2002), se localizaron las proteínas FN1321 y FN1831 descritas por Studholme y Nixon (2003). Ambas proteínas presentan una disposición de dominios similar a NtrC con un dominio central de activación, un dominio C-terminal para unión a ADN y un dominio sensor N-terminal perteneciente al sistema de dos componentes (fig. 7). Por lo tanto, en *F. nucleatum* las proteínas FN1321 y FN1831, muy probablemente puedan seguir detectando señales y actuando sobre nuevos blancos, tales como promotores dependientes de otros factores σ .

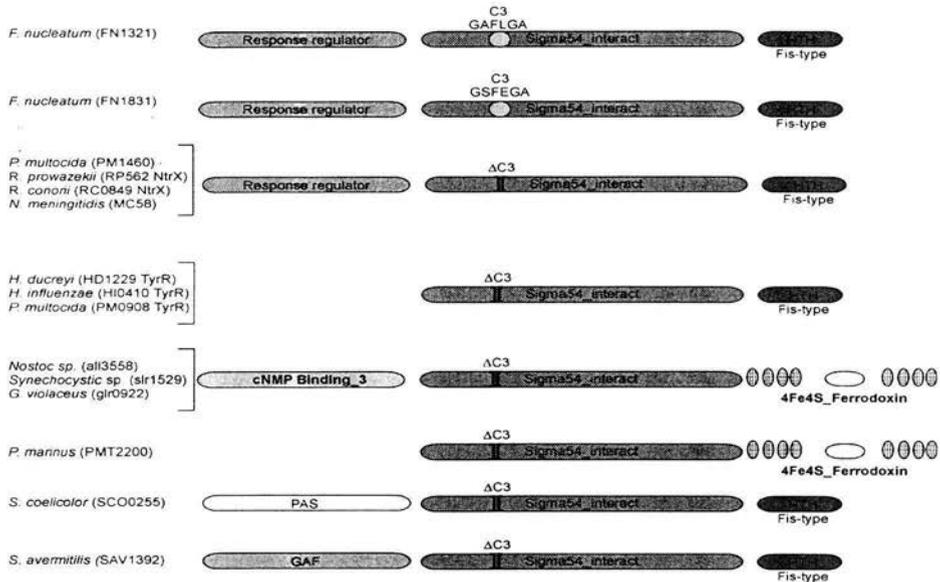


Figura 7. Disposición de dominios de las EBP presentes en organismos sin σ^{54} . En la mayoría de las proteínas se identificaron dominios preestablecidos para las EBP (Studholme y Nixon, 2003) como son el arreglo similar a NtrC y los arreglos de las TyrR's y del género *Streptomyces*. Sin embargo, en cianobacterias todos los dominios presentes, a excepción del dominio central, son atípicos para la familia. La pérdida de la región C3 en estas proteínas corresponde a Δ C3. Dominios asignados por las bases de datos SMART, Prodom y Prosite.

Con las mismas características que FN1321 y FN1831, podemos observar a las proteínas encontradas en los genomas del género *Rickettsia* (*Rickettsia prowazekii*: RP562; y *R. conorii*: RC0849; fig 7). Las rickettsias son parásitos intracelulares obligados que contienen sólo los factores transcripcionales σ^{70} y σ^{32} , así como un número reducido de genes regulatorios. También presentan una tendencia a la reducción genómica (Andersson *et al.*, 1998; Ogata *et al.*, 2001). Las proteínas RP562 y RC0849 han sido anotadas como NtrX, que es parte del sistema NtrY/NtrX, similar a NtrB/NtrC, y que fue determinado por primera vez en *Azorhizobium caulinodans* (Pawlowski *et al.*, 1991). También con el mismo arreglo podemos observar a PM1460, una de las dos proteínas localizadas en el genoma de *Pasteurella multocida*. Los miembros de este género son considerados como oportunistas ó invasores secundarios en vertebrados, y *P. multocida* considerado como un patógeno importante (May *et al.*, 2001). A diferencia de las proteínas de *F. nucleatum* RP562, RC0849 y PM1460 no presentan parte de la región C3 que es de gran importancia para la activación dependiente de σ^{54} (ver siguiente sección).

Nosotros también localizamos otro grupo de proteínas conocidas como TyrR's en Pasteurellaceas (patógenos oportunistas), agrupadas por su alta identidad con el dominio central de la proteína TyrR de *E. coli*. Como ya se mencionó, TyrR en *E. coli* regula al menos ocho genes implicados en la biosíntesis y transporte de aminoácidos aromáticos. Las proteínas de *Haemophilus influenzae* (HI0410), *Haemophilus ducreyi* (HD1229) y *Pasteurella multocida* (PM0908) a diferencia de su homólogo en *E. coli*, carecen del dominio N-terminal (fig. 7). Sin embargo, su funcionalidad no está en duda, ya que existe evidencia experimental de que al menos una de ellas (TyrR de *H. influenzae*; HI0410) actúa como represor transcripcional de promotores dependientes de σ^{70} igual que TyrR de *E. coli* (Pittard and Davidson, 1991; Zhu *et al.*, 1997).

De todas las EBPs presentes en el la fig. 7, las proteínas de las cianobacterias (all3558, slr1529 y glr0922) fueron las que presentaron los arreglos modulares más atípicos entre la familia de las EBPs. Estos organismos son fotosintéticos y presentan una gran variedad de sistemas de transducción de señales y de factores σ (Kaneko *et al.*, 1996; Kaneko *et al.*, 2001; Nakamura *et al.*, 2003). Su disposición de dominios no caracterizado hasta ahora en las EBPs, contiene un dominio N-terminal de unión a cNMP (nucleótido cíclico monofosfato), un dominio central típico de las EBPs, y una región C-terminal con un dominio del tipo ferredoxina (4Fe4S), rodeado de varios elementos transmembranales. Entre las proteínas que contienen el dominio de unión a cNMP se encuentran la CAP (catabolite gene activator protein), proteínas cinasas y canales iónicos. Mientras que el arreglo del C-terminal se encuentra distribuido principalmente en ferredoxinas del tipo bacteriano y en proteínas hierro/azufre para el transporte de electrones. En la proteína de *Prochlorococcus marinus* (PMT2200), también se observa un arreglo 4Fe4S. *P. marinus* es una pequeña cianobacteria marina con pocos genes regulatorios (factores σ , reguladores transcripcionales y sistemas de dos componentes cinasa/sensor) (Rocap *et al.*, 2003). Por lo tanto, estas deficiencias implicarían el uso de vías alternativas de señalización donde pudieran estar integrándose proteínas como PMT2200.

Nosotros proponemos que el tipo de arreglos modulares presentes en las cianobacterias, solamente pudieron ser creados por el intercambio de dominios (Barnett *et al.*, 2000; Gilbert, 1978), donde un dominio central de una EBP con capacidad de detectar

ATP del medio y de formar oligómeros funcionales, se fusionó con otros dominios para realizar una nueva función que tal vez ya no intervenga en la activación transcripcional.

Finalmente, la mayoría de estructuras descritas en esta sección, presentan módulos de sistemas de dos componentes (FN1321, FN1831, PM146, RP562 y RC0849), ó son similares a TyrR de *E. coli* (HD1229, HI0410 y PM0908), ó presentan los arreglos descritos para la cianobacterias (all3558, slr1529, glr0922 y PMT2200). Existen otros dominios en el N-terminal presentes en las proteínas del género *Streptomyces* (fig. 7). En SCO0255 de *S. coelicolor* y SAV1392 de *S. avermitilis* observamos a los dominios PAS y GAF respectivamente, ya descritos anteriormente para otras EBPs, junto con un dominio central y un dominio C-terminal. Estas proteínas serían los primeros elementos de un sistema σ^{54} descritos para las actinobacterias. En el género *Streptomyces* abundan los reguladores transcripcionales del tipo de dos componentes y los factores σ . Sin embargo, nunca se habían identificado ni el factor σ^{54} ni sus activadores (EBPs) (Bentley *et al.*, 2002; Ikeda *et al.*, 2003).

Conservación de secuencia del dominio central de probables EBPs.

Como se puede observar, en el alineamiento de secuencia del dominio central de las 14 EBPs (fig. 8), existe conservación en ciertas regiones importantes ya descritas anteriormente, y que fueron designadas como C1-C7. Sin embargo, se observa que la región C3 de importancia funcional para la activación transcripcional dependiente de σ^{54} , en la mayoría de nuestras probables EBPs, se encuentra incompleta. Esta ausencia específica de los residuos GAFTGA, indica que estas proteínas no esta actividad regulatoria. Existe un número considerable de estudios mutagénicos, de análisis de secuencia y estructurales como el de NtrC1 de *Aquifex aeolicus* (PDB 1NY6) que demuestran la importancia de estos residuos (Gonzalez *et al.*, 1998; Lee *et al.*, 2003; Morett and Segovia, 1993; Osuna *et al.*, 1997). Hay dos casos que sí presentaron esta región completa y son localizados en el genoma de *F. nucleatum*, FN1321 y FN1831. Studholme y Dixon (2003) sugieren que *F. nucleatum*, perdió σ^{54} o ganó estas EBPs por transferencia horizontal, y que su mecanismo de regulación transcripcional era dependiente de σ^{70} debido a la ausencia de σ^{54} . Observamos que algunas regiones de éstas proteínas presentan aminoácidos poco comunes en la familia, por ejemplo en la región C3 de FN1321 la

treonina cambia a leucina (GAFT/LGA), mientras que en FN1831 cambia a ácido glutámico (GSFT/EGA)(figs. 7 y 8). Hasta ahora el cambio T/E de la región C3 de FN1831 sólo se ha detectado en la proteína PhhR *Pseudomonas aeruginosa*, que depende de σ^{54} . Sin embargo, en ésta misma proteína se observa un cambio adicional en la región (GA/SFEGA), representado en activadores como PilR de *P. aeruginosa* (Ishimoto and Lory, 1992), AnfA y VnfA de *Azotobacter vinelandii* (Walmsley *et al.*, 1994), y LevR de *Bacillus subtilis* (Debarbouille *et al.*, 1991). Estas modificaciones se salen del consenso de la familia (GAFTGA), pero al estar presentes en otros activadores funcionales nos sugieren que FN1831 podría ser funcional en un sistema σ^{54} .

Por otro lado, las proteínas que presentan la región C1 ó Walker A con alta conservación son: FN1321 y FN1831 de *F. nucleatum*, las TyrR's (HD1229, HI0410 y PM0908) y las proteínas del género *Streptomyces* (SCO0255 y SAV1392), indicándonos que dichas proteínas tienen una alta probabilidad de unir ATP. En contraste, las proteínas de cianobacterias son más divergentes en ésta región (fig. 8). También, a pesar de que algunas proteínas presentan la región C1 consenso, la región implicada en la hidrólisis de ATP (C4 ó Walker B) tiene modificaciones importantes. Por lo general, en la región C4 están presentes un ácido aspártico y un ácido glutámico (DE) después de una serie de residuos hidrofóbicos. Se ha determinado que el ácido aspártico 238 en *A. aeolicus* es requerido para la interacción con los cationes divalentes (Mg^{+2}) que interactúan con el ATP (Lee *et al.*, 2003). Un ejemplo donde estos residuos no están conservados, se observa en ambas proteínas de *F. nucleatum* (FN1321 y FN1831), donde estas posiciones cambian a ED (FN1321) y AN (FN1831) sugiriendo posibles defectos en la hidrólisis de ATP y que tal vez estas proteínas no requieran de la función de ATPasa para realizar su función. Esta región también se encuentra modificada en las proteínas de cianobacterias y las del género *Streptomyces*. Por el contrario, solamente las candidatas observadas en ambas rickettsias (RP562 y RC0849) presentan esta región más conservada.

Otras regiones importantes que han sido caracterizadas para las EBPs se pueden observar en el alineamiento de la figura 8. La secuencia FREDLY (región C6) no es muy conservada, y en los casos de cianobacterias existe una eliminación de secuencia. La arginina 293 (*Aquifex aeolicus*) de esta secuencia sólo se conserva en algunas EBPs descritas en este estudio. R293 probablemente puede estar contactando el fosfato γ del ATP

(Lee *et al.*, 2003). Esto nos confirma que este residuo junto con la región C1 de estas proteínas sí pueden interactuar con el ATP. Finalmente, tenemos a la región C7, que junto con la región C3 son características para definir diferencias entre las EBPs y la familia AAA+. Esta región C7 es muy conservada, sin embargo, se puede observar que algunas candidatas presentan modificaciones. La región C7 contiene un residuo conocido como sensor II (R357 *A. aeolicus*) que le permite responder a la EBP a los cambios en el balance de nucleótido fosfatado (relación NDP/NTP).

Inferencias filogenéticas. Para tener una idea de cuantas veces ha sucedido este evento de reclutamiento dentro de la familia de las EBPs, nosotros realizamos un análisis de distancias filogenético de las 14 EBPs designadas como reclutadas. El árbol generado por NJ (Clustal_X), es mostrado en la figura 9. La mayoría de las EBPs aquí descritas se encuentran en un grupo, y sólo RP562 (*R. prowazekii*) y RC0849 (*R. conorii*) se encuentran junto con las EBPs funcionales en otro grupo (XylR, AtoC, NifA etc., fig. 9). Estos datos sugieren que el reclutamiento probablemente ha sucedido sólo dos veces en la familia de las EBPs.

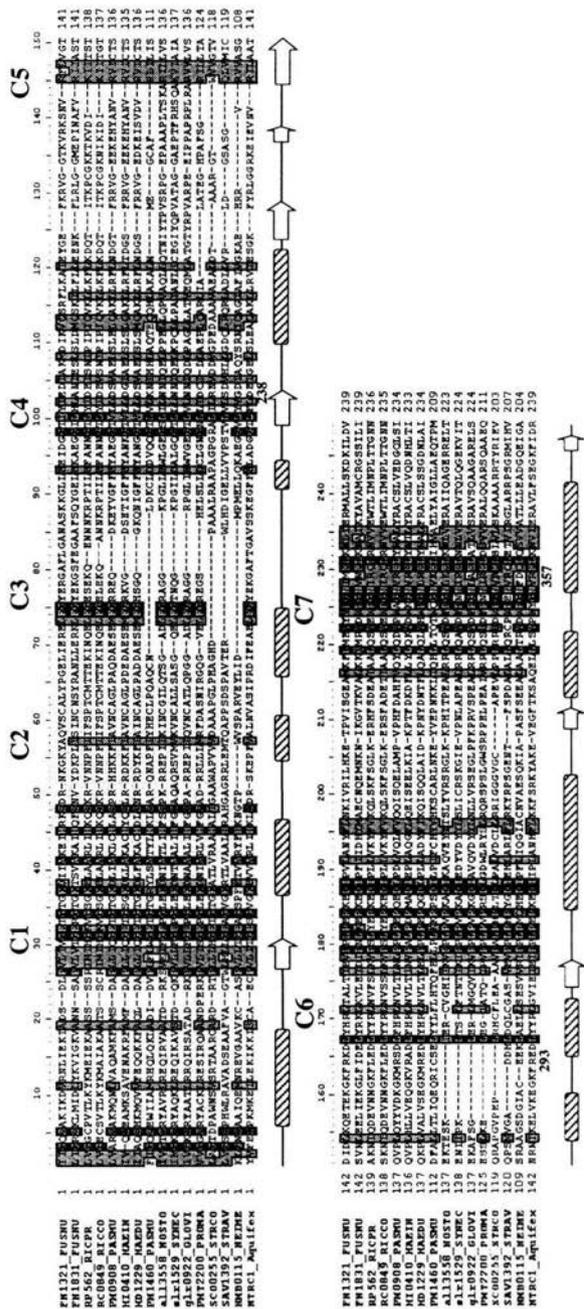


Figura 8. Alineamiento múltiple de secuencias del dominio central de las 14 EBFs encontradas en organismos sin σ^54 . En la figura se pueden observar las regiones de importancia funcional para la familia de las EBFs conocidas como Walker A (región C1), región C3, Walker B (región C4), y regiones C6 y C7. A excepción de las proteínas FN1321 y 1831 de *F. nucleatum*, las demás candidatas presentan la región C3 incompleta. La estructura secundaria esquematizada bajo el alineamiento esta basada en la estructura de NtrC de *A. aeolicus* y los números representan algunas posiciones importantes, ver texto.

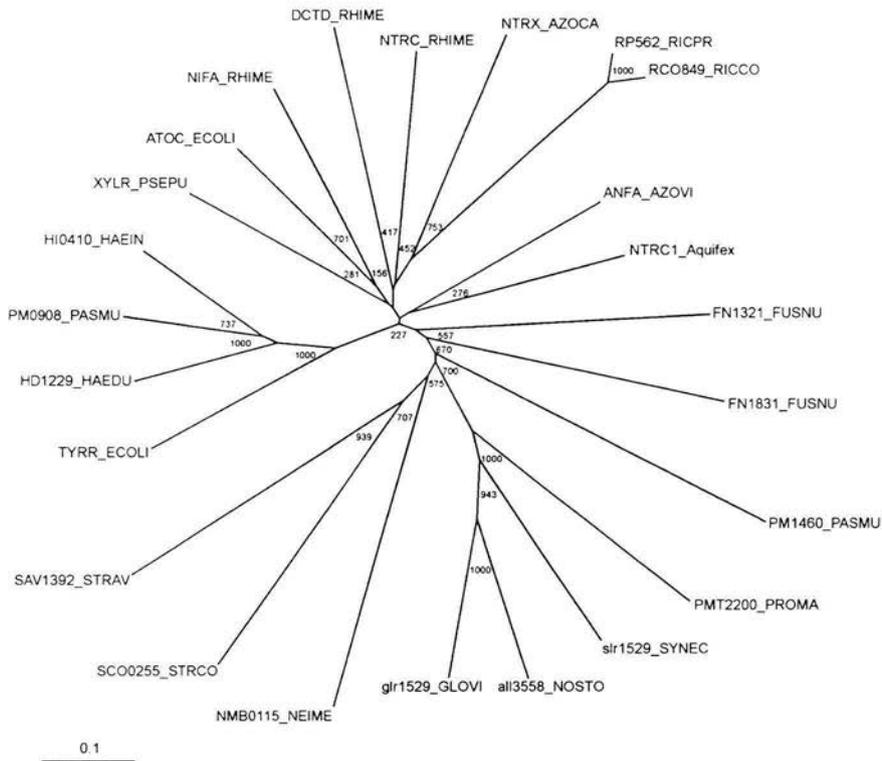


Figura 9. Arbol de agrupamiento del dominio central de probables EBPs reclutadas. En este árbol basado en distancias, se observa que en la familia de EBPs han sucedido 2 eventos globales de reclutamiento. Uno donde se encuentran la mayoría de las EBPs obtenidas en la búsqueda, mientras que en el otro se encuentran las EBPs del género *Rickettsia* RP562_RICPR y RC0849_RICCO junto con EBPs funcionales (previamente descritas en la literatura). El árbol fue generado por Neighbour Joining (NJ) de la paquetería de Clustal_X. Los valores internos indican el valor de bootstrap.

II. Evolución por Mutagénesis del Dominio Central de TyrR de *E. coli*.

Resultados.

Alineamiento de los dominios centrales de TyrR y PhhR.

Dentro de las EBPs, la proteína PhhR de *Pseudomonas aeruginosa* es la que comparte mayor identidad de secuencia con TyrR de *E. coli*. Utilizando las secuencias de aminoácidos de los dominios centrales, obtenidas en la base de datos del NCBI (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>), se realizó un alineamiento múltiple usando la paquetería del programa Clustal_X para identificar residuos atípicos en las regiones de conservación que comparten estas proteínas con las EBPs. En los resultados de este análisis (fig. 10), se puede observar que el dominio central de PhhR presenta los determinantes importantes para la activación dependiente de σ^{54} , tales como las regiones conservadas C1, C3, C4 y C7 descritas por Morett y Segovia (1993). Sin embargo, el dominio central de PhhR presenta en las regiones C3 y C4 residuos poco comunes para la familia de las EBPs, tales como la treonina 274, el ácido glutámico 285, la adición de la arginina 288 y la glicina 308 (fig.10). También se observa en el alineamiento que la región C3, una de las más conservadas y relevante para la activación σ^{54} , no está presente en el dominio central de TyrR como ya se había mencionado anteriormente. Para lograr nuestro objetivo, nos basamos en la relación de identidad que existe entre el dominio central de PhhR con el dominio central de TyrR (49 % a nivel de aminoácidos, bl2seq NCBI), y en el hecho de que PhhR puede sustituir a TyrR en algunas de sus funciones (Song and Jensen, 1996).

Construcciones TyrRcentralWT y TyrRcentralGAFTGA.

En un estudio previo (Flores H, datos no publicados) se observó por ensayos de β -galactosidasa *in vivo* en *E. coli* JM101, que el dominio central silvestre de TyrR de *E. coli* (TyrRcentralWT) no activa la transcripción de promotores dependientes de σ^{54} (pCU101 y pSU003, tabla 3). Tampoco se detectó la capacidad de activar transcripcional con la construcción TyrRcentralGAFTGA que contiene los residuos de la región C3 consenso para las EBPs, ausentes en TyrR silvestre. Estas observaciones fueron corroboradas en el presente trabajo (tablas 6 y 7) y dieron la pauta para probar nuevas posiciones en dicha región C3 con el fin de obtener una variante que activara genes dependientes de σ^{54} .

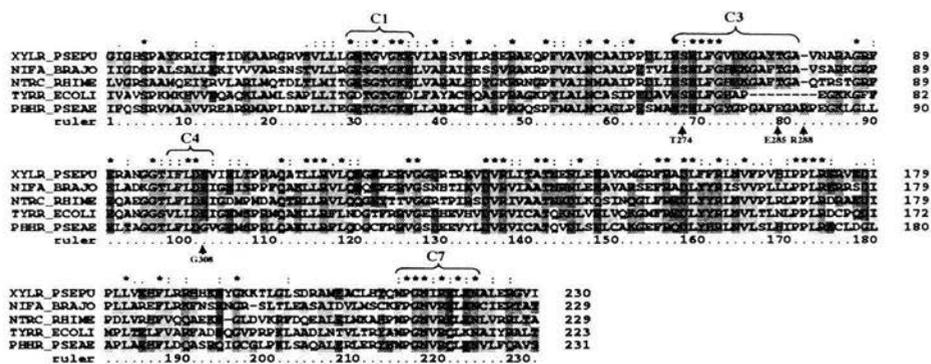


Figura 10. Alineamiento de las secuencias proteicas del dominio central de varias EBPs. Las regiones conservadas y de importancia funcional son señaladas como C1, C3, C4 y C7 (Morett and Segovia, 1993). PhhR de *Pseudomonas aeruginosa* es la más cercana a TyrR de *E. coli* y presenta una región C3 distinta a los demás miembros de las EBPs (residuos indicados con flechas). También se observa que la región C3 de la proteína TyrR de *E. coli* no esta completa. XYLR_PSEPU, proteína XylR de *Pseudomonas putida*; NIFA_BRAJO, proteína NifA de *Bradyrhizobium japonicum*; NTRC_RHIME, proteína NtrC de *Rhizobium meliloti*; TYRR_ECOLI, proteína TyrR de *Escherichia coli*; PHHR_PSEAE, proteína PhhR de *Pseudomonas aeruginosa*. Alineamiento por Clustal_X, Mac OSX.

Activación en trans de PhhR.

Lo primero que hicimos fue verificar si el dominio central de PhhR presentaba una capacidad robusta de activar en trans promotores dependientes de σ^{54} . La construcción pUC18PhhR que expresa el dominio central de la proteína PhhR de *Pseudomonas aeruginosa* fue obtenida utilizando distintas condiciones en la reacción de PCR debido al gran contenido de GC presente en el genoma de *P. aeruginosa*. La capacidad de activación en trans por éste dominio fue evaluada mediante ensayos de actividad de β -galactosidasa *in vivo*. Para estos ensayos se utilizó la cepa JM101 de *E. coli* y el plásmido reportero pCU101 (materiales y métodos, tabla 3) que contiene la región promotora -12, -24 y las UAS de *nifH* de *Rhizobium meliloti* fusionada a *lacZ*.

Los resultados del ensayo son mostrados en la figura 11. La activación en trans por el dominio central de PhhR (pUC18PhhR: 1543 unidades) es 15 veces menor que la activación transcripcional generada por la proteína NifA completa de *Bradyrhizobium japonicum* (pM1-k: 25080 unidades). La activación por pUC18PhhR es ligeramente mayor que la del dominio central silvestre de TyrR (TyrRcentralWT: 1254 unidades) y que la del dominio central de TyrR con parte de la región C3 al consenso (TyrRcentralGAFTGA:

1511 unidades). Probablemente la baja activación transcripcional del pUC18PhhR observada, sea el resultado de la región C3 poco común que presenta PhhR (fig. 10). Por ejemplo, PhhR no presenta en su región C3 una treonina implicada en la función de las EBPs. En NifA de *B. japonicum*, se ha demostrado que esta treonina (308) en la región C3 (GAFTGA) juega un papel crucial para la activación transcripcional del sistema σ^{54} (González *et al.*, 1998). Sin embargo, la capacidad de activación de PhhR ha sido demostrada por la expresión y detección inmunológica de PhhA, una proteína bajo control de PhhR (Song and Jensen, 1996). También la baja activación se puede deber a que utilizamos solamente el dominio central de PhhR y no la proteína completa. Por otro lado, quizá σ^{54} de *E. coli* sea distinto e incapaz de complementar a σ^{54} de *P. aeruginosa*, o que PhhR presente una baja función de activación de manera natural. Por lo tanto, los datos observados en la figura 11 indican que PhhR presenta una baja activación en *trans* y cualquiera de estas razones puede ser la respuesta.

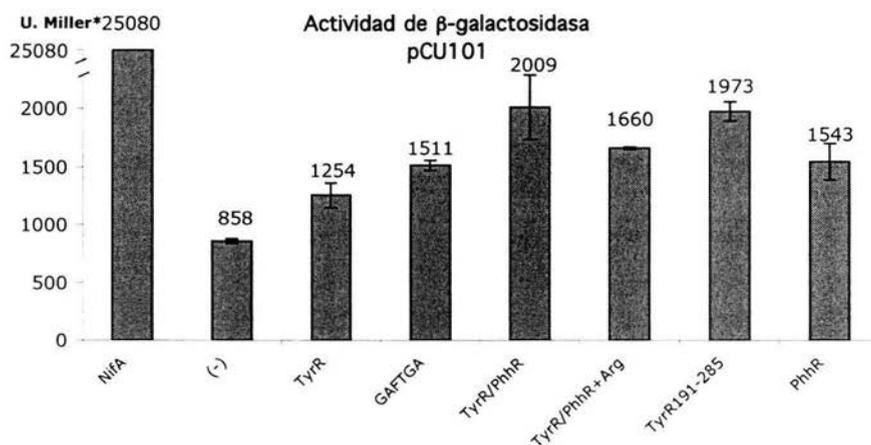


Figura 11. Actividad de β -galactosidasa *in vivo* de construcciones del dominio central de TyrR. Se observa que el dominio central de PhhR presenta una baja capacidad de activación en *trans*, mientras que las mutantes TyrR/PhhR y TyrR191-285 son las que generan mayor activación. NifA: pM1-k; (-): pKK223-3; TyrR: TyrRcentralWT; GAFTGA: TyrRcentralGAFTGA; PhhR: pUC18PhhR. Los ensayos se realizaron en la cepa *E. coli* JM101 con crecimiento anaeróbico a 30° C y en medio LB. Se utilizó el reportero pCU101 (tabla 3). *Unidad Miller: Cantidad de enzima que produce 1nmol de o-nitrofenol/min/OD₆₀₀ de cultivo celular (Miller, 1972).

Activación transcripcional de TyrR/PhhR.

Para lograr el objetivo del trabajo y de acuerdo a lo observado en el alineamiento de la figura 10 se construyó una variante del dominio central de TyrR que presentara una región C3 semejante a la de PhhR de *Pseudomonas aeruginosa* (TyrR/PhhR). Para ello se utilizó mutagénesis dirigida al sitio (materiales y métodos). La alta identidad de secuencia entre TyrR y PhhR nos hizo suponer que la restitución de la región C3 en TyrR con residuos de PhhR podría generar una variante con la capacidad de activar genes dependientes de σ^{54} . Esta región en la proteína PhhR completa si es dependiente de σ^{54} . Por otro lado, con el uso de una secuencia C3 consenso en TyrR no se observaron buenos resultados al buscar restablecer ésta función.

La nueva construcción TyrR/PhhR fue evaluada como activador de genes dependientes de σ^{54} . Para ello se cuantificó la actividad de β -galactosidasa *in vivo* en *E. coli* JM101 transformada con el reportero pCU101 que contiene la fusión *nifH-lacZ* de *R. meliloti* (pCU101, tabla 3). La actividad de la variante es mostrada en la figura 11. La construcción TyrR/PhhR (2009 unidades) mostró ser más activa que las construcciones TyrRcentralWT (1254 unidades), TyrRcentralGAFTGA (1511 unidades) y pUC18PhhR (1543 unidades). Estos resultados indican que los cambios introducidos incrementan la capacidad de activar la expresión del gen reportero. Sin embargo, la variante TyrR/PhhR mostró aún baja actividad, por lo que se buscó incrementarla. Para ello, generamos por mutagenénesis dirigida al sitio nuevas modificaciones en la región C3, creando la mutante TyrR/PhhR+Arg (materiales y métodos, tabla 2). Esta variante presenta en la región C3 la adición de una arginina (288 en PhhR, fig. 10) y los siguientes cambios (mostrados en negrillas):

PhhR WT	ESMAE <u>T</u> ELFGY <u>G</u> PGAF <u>E</u> GAR <u>P</u> EGKL <u>G</u> LLELTAGGTLFLDG
TyrR/PhhR	EDAV <u>E</u> TELFGY <u>G</u> PGAF <u>E</u> GA- TALKKGF FEQANGGSVLLDE
TyrR/PhhR+Arg	EDAV <u>E</u> TELFGY <u>G</u> PGAF <u>E</u> GAR <u>P</u> EGKL <u>G</u> LLELTAGGSVLLDE

La actividad transcripcional de TyrR/PhhR+Arg (1660 unidades) resultante de los ensayos de β -galactosidasa *in vivo*, se reporta en la figura 11, donde se puede observar una disminución en la activación con relación a TyrR/PhhR (2009 unidades). Este efecto negativo producido por la introducción de la arginina y los otros residuos, se puede deber a

que este grupo de residuos en PhhR pueden estar contactando con otros residuos que no están presentes en la proteína TyrR. Por último, la arginina 288 presente en PhhR y no en otras EBPs pudo ser la causa de que el dominio central de PhhR presentara una baja activación en *trans* de manera natural.

Durante la obtención de la construcción TyrR/PhhR+Arg, se generó una mutante espontánea, que presenta un codón de término ocre (UAA) en la treonina de la región C3 (GAFT/ZGA)(fig. 12). Esto origina que sólo se exprese la región N-terminal del dominio central de TyrR/PhhR, ya que la cepa utilizada (JM101) no es supresora para este codón de término. Esta mutante denominada TyrR191-285 mostró una actividad transcripcional (1973 unidades) similar a la construcción TyrR/PhhR (2009 unidades) (fig. 11). Este resultado inesperado nos hizo analizar con más detalle esta variante.

Activación transcripcional de pUCTyrRZ y pUCPspFZ.

Para verificar que la actividad transcripcional generada por el fragmento proteico resultante de la expresión de TyrR191-285 (fig. 11) no fuera un artificio, nos dimos a la tarea de contruir dos nuevas mutantes, pUCTyrRZ y pUCPspFZ. Ambas fueron clonadas como péptidos de fusión en pUC18 (*EcoRI-HindIII*). La construcción pUCPspFZ es originada del activador transcripcional PspF que controla genes de respuesta a fago dependientes de σ^{54} en *E. coli*. Tanto pUCTyrRZ como pUCPspFZ consisten únicamente de un dominio de activación (central), interrumpido por un codón de término en la treonina de la región C3 y la eliminación del resto de la secuencia (fig. 12). Esta interrupción fue generada por PCR usando oligonucleótidos mutagénicos (materiales y métodos, tabla 2).

TyrR/PhhR	MITNSQLQNV	AAQDVSAFSQ	IVAVSPKMKH	VVEQAQKLAM	LSAPLLITGD
TyrR191-285	MITNSQLQNV	AAQDVSAFSQ	IVAVSPKMKH	VVEQAQKLAM	LSAPLLITGD
TyrRZ	MITNSQLQNV	AAQDVSAFSQ	IVAVSPKMKH	VVEQAQKLAM	LSAPLLITGD
PspF^{WT}	MRGSHHHHHH	GIRMAEYKDN	LLGEANSFLE	VLEQVSHLAP	LDKPVLIIGE
PspFZ	MITN.....	..SMAEYKDN	LLGEANSFLE	VLEQVSHLAP	LDKPVLIIGE
				Región C3	
					↑
TyrR	TGTGKDLFAY	ACHQASPRAG	KPYLALNCAS	TPEDAVETEL	FG YGP GAFEG
TyrR191-285	TGTGKDLFAY	ACHQASPRAG	KPYLALNCAS	TPEDAVETEL	FG YGP GAFEG
TyrRZ	TGTGKDLFAY	ACHQASPRAG	KPYLALNCAS	TPEDAVETEL	FG YGP GAFEG
PspF^{WT}	RGTGKELIAS	RLHYLSSRWQ	GPFISLNCAA	LNENLLDSEL	FGHEAGAF TG
PspFZ	RGTGKELIAS	RLHYLSSRWQ	GPFISLNCAA	LNENLLDSEL	FGHEAGAF TG
				Región C3	
				↑	
TyrR/PhhR	AT.ALKKGFF	EQANGGSVLL			
TyrR191-285	ARPEGKFGFF	EQANGGSVLL			
TyrRZ			
PspF^{WT}	AQ.KRHPGRF	ERADGGTLFL			
PspFZ			

Figura 12. Alineamiento de la región N-terminal de los dominios de activación de las distintas construcciones TyrR y PspF. En la región C3 se generaron mutaciones puntuales para crear un codón de término igual al presente en TyrR191-285. Las construcciones resultantes fueron pUCTyrRZ (TyrRZ) y pUCPspFZ (PspFZ). En negritas y cursivas se señalan las modificaciones iniciales tomadas de la secuencia de PhhR e introducidas en el dominio central de TyrR. En sombreado se indica la señal de terminación (ocre) adicionada en la secuencia en pUCTyrRZ y pUCPspFZ, los aminoácidos siguientes son parte de la secuencia del vector pUC18, mientras que, en TyrR191-285 son la continuación del dominio central. Alineamiento hecho en Clustal_X (Mac OSX).

La capacidad de activación transcripcional de ambas construcciones fue determinada por ensayos de β -galactosidasa *in vivo* y los resultados son mostrados en la tabla 6. Se puede observar que la activación de pUCTyrRZ (797 unidades) y TyrR191-285 (792) son muy similares y que la construcción pUCPspFZ (674 unidades) se encuentra por debajo de estas. Esto nos indica que no existe una supresión del codón de término en TyrR191-285, que podría estar generando un dominio central completo. Sin embargo, el efecto de este codón de término en pUCPspFZ, sí influyó en su capacidad de activar la transcripción en nuestro sistema de activación pCU101, como se observa claramente en la tabla 6.

Tabla 6. Actividad β -galactosidasa *in vivo* de pUCTyrRZ y pUCPspFZ. Crecimiento de *E. coli* MC1061 anaeróbico a 30° C en medio mínimo M9. Uso de reportero pCU101 (tabla 3).

Clona	Unidades Miller	Desviación estándar (+/-)
pM1-k NifA <i>B. japonicum</i>	17636	834
pKK223-3	405	23
TyrRcentralWT	617	24
pMJ15 Dom central PspF ^{WT}	3784	211
TyrR/PhhR	697	14
TyrR191-285	792	24
pUCTyrRZ	797	43
pUCPspFZ	674	20

*Unidad Miller: Cantidad de enzima que produce 1nmol de o-nitrofenol/min/OD₆₀₀ de cultivo celular (Miller, 1972).

Los resultados obtenidos con las construcciones pUCTyrRZ y TyrR191-285 también indican que el fragmento obtenido de la expresión de ambas construcciones, puede contener los requerimientos necesarios para generar una activación transcripcional del sistema σ^{54} . Por lo tanto, es necesario demostrar si esta activación transcripcional observada con el fragmento pUCTyrRZ, es dependiente de σ^{54} . Para esto, decidimos construir una cepa de *E. coli* sin σ^{54} .

Construcción de cepa MC1061 Δ rpoN.

La cepa *E. coli* MC1061 sin σ^{54} (gene *rpoN*), se generó por un evento de doble recombinación utilizando el vector pK03, como se indica en materiales y métodos. La cepa MC1061 Δ rpoN, fue caracterizada por su crecimiento en medio W-arginina, por activación transcripcional de genes dependientes de σ^{54} y por PCR de regiones dentro y fuera del gen *rpoN*.

Primero se verificó por PCR la eliminación del gen *rpoN* en la cepa MC1061 Δ rpoN, usando oligonucleótidos que amplifican las secuencias adyacentes 5' y 3' del gen *rpoN* (*rpoNA1* y *rpoNB2*, materiales y métodos, tabla 2) y oligonucleótidos que amplifican la región N-terminal del gene *rpoN* de *K. pneumoniae* (B7 y W302, materiales y métodos, tabla 2). La PCR usando ADN cromosomal de MC1061 y los oligonucleótidos *rpoNA1* y *rpoNB2* generó un fragmento de 2432 pb, que corresponde a las regiones adyacentes 5' y 3' (500 pb de cada una), más las 1432 pb de la región codificante del gene *rpoN* (fig. 13 banda A). En MC1061 Δ rpoN, con los mismos oligonucleótidos, la PCR generó un producto de 1000 pb que corresponde sólo a las regiones adyacentes 5' y 3' (fig. 13 banda B), indicándonos la eliminación de la secuencia del gen *rpoN* en este locus. También por PCR, se verificó si la secuencia eliminada no se había desplazado a otra región por algún evento de recombinación durante el proceso de eliminación del gene *rpoN*. Para determinar esto, se utilizaron los oligonucleótidos B7 y W302 que amplifican aproximadamente 800 pb de la región codificante 5' del gene *rpoN* de *K. pneumoniae*. En la figura 13 se observa claramente que la PCR con ADN de MC1061 Δ rpoN, no produce la banda esperada, mientras que con MC1061 esta banda si se observa (fig. 13 banda C). Finalmente, el producto de la amplificación de MC1061 Δ rpoN generado con los oligonucleótidos *rpoNA1* y *rpoNB2* (fig. 13 banda B) fue secuenciado por la unidad de secuencia del Instituto de

Biocología de la UNAM, y al analizar su secuencia se observó que durante la eliminación del gen *rpoN* no se afectaron las regiones vecinas.

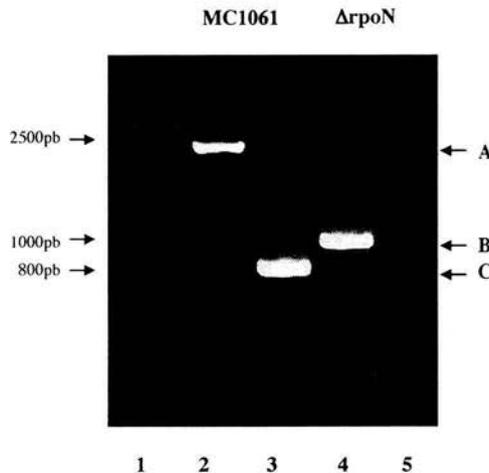


Figura 13. Electroforesis en gel de agarosa al 1% para la caracterización genotípica por PCR de MC1061 Δ *rpoN*. La banda A es el producto de amplificación del gen *rpoN* más las regiones adyacentes 5' y 3'; la banda B corresponde a las regiones adyacentes únicamente; y la banda C es la región interna del gene *rpoN*. Carril 1: Marcador GenRuler Fermentas; carriles 2 y 4: amplificación con oligos *rpoNA1* y *rpoNB2* (tabla 2) de MC1061(2) y MC1061 Δ *rpoN*(4); carriles 3 y 5: amplificación con oligos B7 y W302 (tabla 2) de MC1061(3) y MC1061 Δ *rpoN*(5).

Los resultados de la caracterización fenotípica se reportan en la tabla 7 y destaca que la cepa MC1061 Δ *rpoN* sólo creció en W-arginina cuando fue complementada con un gen *rpoN* silvestre de *Klebsiella pneumoniae*. El medio W-arginina induce la expresión del promotor *glnAp2* dependiente de σ^{54} , que controla al operón *glnALG* (*glnA*:glutamina sintetasa; *glnL*:NtrB; y *glnG*:NtrC) necesario para la asimilación de nitrógeno en *E. coli* (Ninfa *et al.*, 1987; Syed and Gralla, 1998). La ausencia de σ^{54} en *E. coli* MC1061 Δ *rpoN* impide la expresión de *glnAp2*, reflejándose en una disminución considerable en el crecimiento.

Para la caracterización funcional de la cepa MC1061 Δ *rpoN*, primero se observó el fenotipo generado por la expresión de un gen reportero bajo el control de σ^{54} en cajas de LB/X-Gal/Cm y posteriormente se cuantificó la activación por ensayos de β -galactosidasa *in vivo*. En MC1061 Δ *rpoN*/pCU101 sólo se observó la activación transcripcional basal del

reportero, mientras que, las colonias de una cepa MC1061 Δ rpoN/pCU101 complementada con σ^{54} de *K. pneumoniae* produjeron un azul más intenso por la degradación del X-Gal, indicando la activación transcripcional realizada por los activadores presentes en las células (tabla 7). Finalmente, también se verificó por ensayos de β -galactosidasa *in vivo* (materiales y métodos), la activación transcripcional σ^{54} , utilizando la proteína NifA de *Bradyrhizobium japonicum* como activador del sistema σ^{54} (pM1-k). En los resultados de la tabla 8, se observa que en presencia del activador NifA y un σ^{54} funcional (σ^{54} *E. coli*: 3467 unidades ó σ^{54} *K. pneumoniae*: 3315 unidades), se incrementa considerablemente la activación transcripcional del reportero, mientras que en la cepa MC1061 Δ rpoN, con el activador NifA, no se observa activación transcripcional considerable (403 unidades).

Tabla 7. Caracterización fenotípica de MC1061 Δ rpoN I. Crecimiento de MC1061 y MC1061 Δ rpoN en medios W-arginina y W-arginina/X-Gal. Actividad de la fusión *nifH-lacZ* (pCU101, tabla 3).

Cepa	Crecimiento		Actividad fusión <i>nifH-lacZ</i> (pCU101). LB/x-Gal
	LB	W-Arg	
MC1061wt	+	+	Observable
MC1061 Δ rpoN	+	-	Mínima
MC1061 Δ rpoN/ <i>rpoN K. pneumoniae</i>	+	+	Observable

Tabla 8. Caracterización fenotípica de MC1061 Δ rpoN II. Actividad de β -galactosidasa *in vivo* de *E. coli* MC1061 Δ rpoN+pWKS NifA y el reportero pCU101. Crecimiento aeróbico a 37° C en LB.

Cepa/reportero/construcción	Unidades Miller*	Desviación Estándar
MC1061wt/pCU101	391	6.3
MC1061wt/pCU101/pM1-k	3467	224.1
MC1061 Δ rpoN/pCU101	390	28.7
MC1061 Δ rpoN/pCU101/pM1-k	403	62.3
MC1061 Δ rpoN/pCU101/pM1-k/ <i>rpoN K. pneumoniae</i>	3315	229.7

*Unidad Miller: Cantidad de enzima que produce 1nmol de o-nitrofenol/min/OD₆₀₀ de cultivo celular (Miller, 1972).

Todos los resultados anteriores demuestran que la cepa de *E. coli* MC1061 Δ rpoN no presenta el gen *rpoN*.

Actividad transcripcional de pUCTyrRZ en la cepa MC1061 Δ rpoN.

Utilizando la cepa MC1061 Δ rpoN, verificamos si la activación transcripcional observada con la construcción pUCTyrRZ era dependiente de σ^{54} , ó si fue producto de otros factores celulares.

En la figura 14 se observa que la variante pUCTyrRZ con el codón de término muestra una activación mayor que los controles en presencia de σ^{54} . Sin embargo, en los resultados de las actividades transcripcionales realizadas en MC1061 Δ rpoN no se observa claramente una disminución en la activación para la construcción pUCTyrRZ. Estos resultados indican que probablemente la activación observada no es dependiente de σ^{54} . Quizá en un trabajo posterior sería muy interesante identificar esta activación por ensayos transcripcionales *in vitro* con el fin tener un mejor panorama de lo que se esta observando en este trabajo.

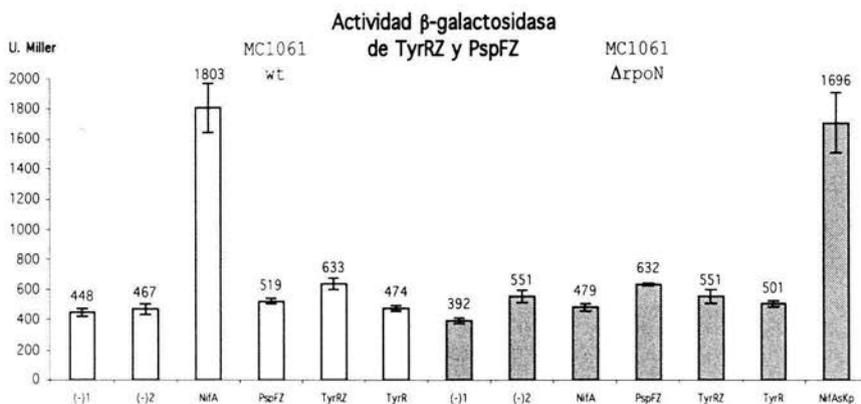


Figura 14. Activación transcripcional de pUCTyrRZ y otras construcciones en presencia y ausencia de σ^{54} . En las actividades realizadas en MC1061 (σ^{54} wt) se observa que sólo el control positivo (NifA) y pUCTyrRZ activan la transcripción por arriba de las demás construcciones evaluadas. Sin embargo en la ausencia de σ^{54} (MC1061 Δ rpoN), la actividad de pUCTyrRZ no disminuye de manera considerable. Esta disminución en la actividad no se mantiene constante en otros ensayos realizados, indicando que no es posible determinar la dependencia de esta activación por σ^{54} . [-]1: cepa/pCU101; [-]2: pUC18; NifA: pM1-k; TyrR: TyrRcentralWT. La última barra representa la activación con NifA en la cepa MC1061 Δ rpoN complementada con σ^{54} de *Klebsiella pneumoniae*. Uso de reportero pCU101 (tabla 3). Crecimiento aeróbico en LB y a 37° C.

Construcción de banco mutagénico de TyrR/PhhR.

A partir de los datos mostrados en la figura 11 para la construcción TyrR/PhhR que contiene parte de la región C3 del dominio central de PhhR, decidimos generar variantes al azar de esta para incrementar su actividad. Esta construcción fue sometida a mutagénesis por PCR y se generó un banco de 2.3×10^5 clonas. La tasa de mutagénesis del gen fue del 2.12 %, esta tasa de mutagénesis indica que por cada inserto de 850 pb (inserto de TyrR/PhhR en pUC18) existían aproximadamente 18 cambios. El banco obtenido y clonado en pUC18 (*EcoRI-HindIII*) fue electroporado en la cepa *E. coli* MC1061 que contenía el reportero pRT22. El vector pRT22 (materiales y métodos, tabla 3) contiene la región de control *nifH* de *K. pneumoniae* fusionada a *lacZ* y decidimos utilizarlo en vez de pCU101 por presentar una baja actividad basal. Esto con el fin de identificar diferencias durante el tamizado del banco.

En el tamizado del banco bajo el sistema pRT22, se exploraron 50 000 clonas en 50 cajas con medio mínimo M9-X-Gal. Para una buena observación del fenotipo (colonias azules), es necesario tener las colonias aisladas en las cajas de petri, de lo contrario se tendrían que aislar nuevamente si se encuentran mezcladas. De las clonas exploradas se seleccionaron por color 65 candidatas que fueron nuevamente tamizadas bajo las mismas características. Finalmente 5 clonas con una mayor coloración fueron sometidas a ensayos de β -galactosidasa *in vivo*.

Actividades de las mutantes seleccionadas (reportero pCU101).

Para verificar si la activación de las 5 mutantes obtenidas durante el tamizado con el promotor *nifH* de *K. pneumoniae* (pRT22) se mantenía por arriba de TyrR/PhhR utilizando otro promotor, estas fueron sometidas a ensayos de actividad transcripcional de β -galactosidasa *in vivo* con el reportero pCU101 (promotor *nifH* de *R. meliloti*). Como se observa en la figura 15, las 5 mutantes presentan actividades superiores a la construcción parental TyrR/PhhR (M1, 909 unidades; M2, 985 unidades; M3, 949 unidades, M4, 849 unidades; M5 1186 unidades; y TyrR/PhhR, 665 unidades), donde M5 sobresale con casi el doble de actividad.

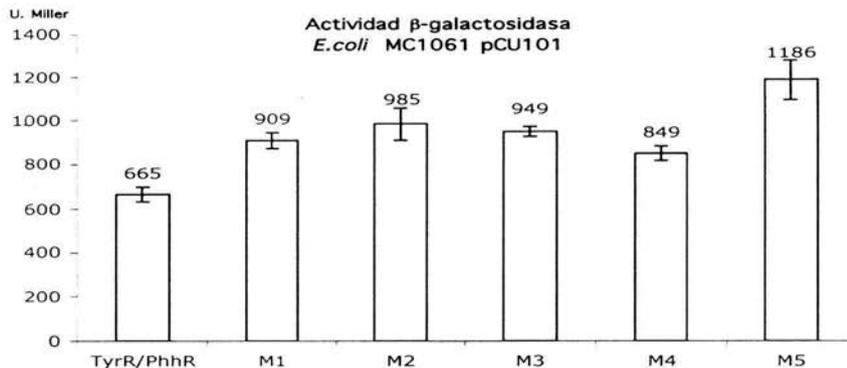


Figura 15. Actividad transcripcional de las mutantes M1-M5 obtenidas del banco mutagénico de TyrR/PhhR. Se observa claramente que las 5 mutantes seleccionadas por el tamizado con el vector pRT22 conservan el incremento en la activación (con relación a la construcción parental) bajo el sistema pCU101. La mutante 5 es la más activa, con casi el doble de actividad de TyrR/PhhR. Los ensayos *in vivo* fueron realizados en *E. coli* MC1061. Crecimiento aeróbico en LB y a 37° C.

Cuando se analizaron las secuencias a nivel de proteína de las 5 candidatas (fig. 16), se pudo observar que sólo existían dos secuencias diferentes. Las candidatas M1, M2 y M5, presentan un codón de término en el inicio de la secuencia, que impediría la formación del dominio central completo. Este codón de término (ambar) generado en la PCR mutagénica no puede ser suprimido por la cepa utilizada en el ensayo (MC1061). Una posibilidad para explicar este fenómeno es que la traducción del ARN de estas mutantes pueda estar iniciando en un segundo codón ATG (metionina) ó GTG (valina) presente en el gen y que se encuentre en fase con el resto de la proteína. Esto acortaría la proteína en 20 aminoácidos. Sin embargo, para esto sería necesaria la cercanía de una región similar a la Shine-Dalgarno para el reconocimiento por el ribosoma. Las mutantes M3 y M4, no presentan el codón de término y la mayoría de los cambios observados son distintos a las primeras mutantes. Los cambios de todas las mutantes son descritos en la figura 15. Los cambios no afectan las regiones de importancia funcional que han sido descritas para las EBPs (C1, C3, C4, y C7). Sin embargo, cuando se identificaron estos cambios en la estructura de NtrC1 de *Aquifex aeolicus* (Lee *et al.*, 2003), se observó que la mayoría se localizaban en las regiones de interfase entre los monómeros del heptámero funcional.

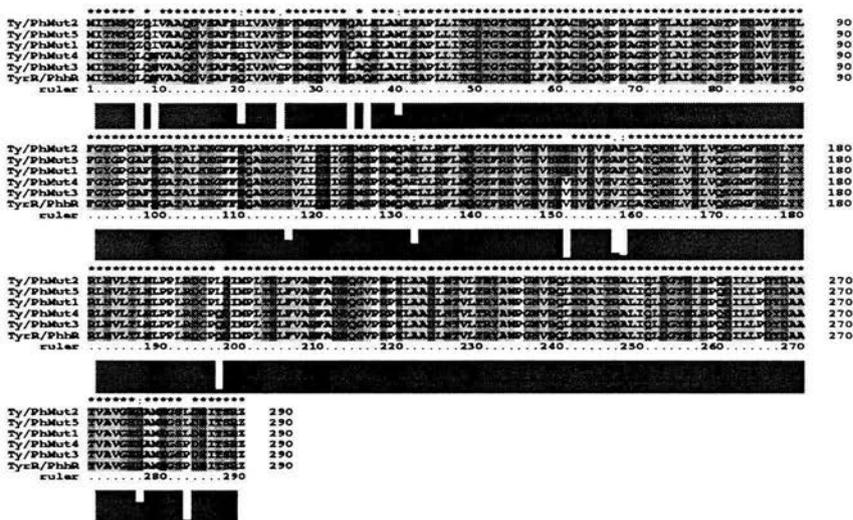


Figura 16. Alineamiento de las secuencias proteicas de las mutantes M1-M5. Se observa la disposición de aminoácidos de los dos grupos de mutantes hermanas (grupo 1: Mut 1, Mut 2 y Mut 5; grupo 2: Mut 3 y Mut 4). Cada grupo mantiene sus cambios característicos no localizados en regiones de importancia funcional. Alineamiento realizado con Clustal_X (Mac OSX).

Al observar los resultados del análisis de secuencia, se seleccionó una candidata de cada grupo (M4 y M5) y ambas fueron sometidas nuevamente a ensayos de activación transcripcional *in vivo*, pero en la cepa MC1061 Δ rhoN y con el reportero pMB210 (materiales y métodos, tabla 3). Esto se realizó con el fin de determinar si la activación producida por estas construcciones es dependiente de σ^{54} . En la figura 17, se observa claramente que las mutantes activan mejor que la construcción parental en presencia de σ^{54} . Sin embargo, en ausencia de σ^{54} también se observa este comportamiento. Estos datos, sugieren que las mutantes probablemente puedan estar activando la fusión por otro mecanismo transcripcional independiente de σ^{54} . Quizá la existencia de un promotor σ^{70} sobrelapado en la región de control de *nifH* de *R. meliloti* (pMB210), o una activación proveniente de la región 5' del promotor -12, -24 sea la respuesta. Existe evidencia experimental de la presencia de promotores dependientes de σ^{70} sobrelapados en promotores -12, -24, como es el caso de *fixR* de *Bradyrhizobium japonicum* (Barrios *et al.*, 1995). Por lo tanto, diseñamos varias modificaciones en el vector pMB210 para evaluar estas posibilidades.

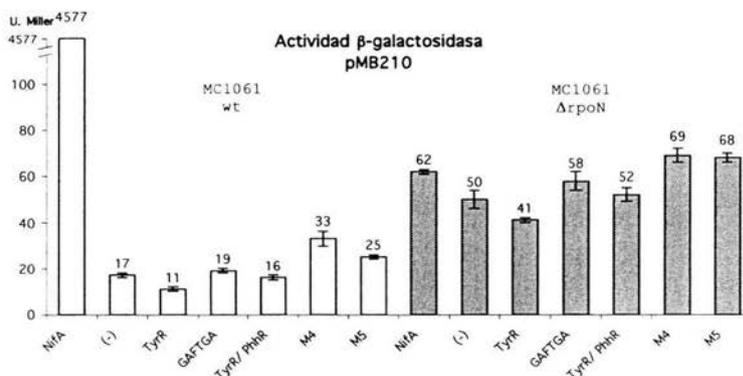


Figura 17. Activación transcripcional de M4 y M5 en presencia y ausencia de σ^{54} . A primera vista se puede observar una mayor activación transcripcional en la cepa MC1061 Δ rhoN que no presenta σ^{54} . Las mutantes M4 y M5 en MC1061 mantienen los niveles de activación por arriba de los controles observados en ensayos anteriores, sin embargo, estas diferencias se conservan en MC1061 Δ rhoN, indicando probablemente una independencia de σ^{54} . NifA:pM1-k; [-]:pUC18; TyrR:TyrRcentralWT; GAFTGA: TyrRcentralGAFTGA. Se usó el reportero pMB210 (tabla 3). Crecimiento aeróbico en LB y a 30° C.

Nosotros construimos 3 reporteros (fig. 18) a partir del plásmido pMB210 (Ditta *et al.*, 1985) para determinar una posible actividad transcripcional generada en la región 5' del promotor *nifH*. El plásmido pMB210 contiene aproximadamente 750 pb de la región de control *nifH* de *Rhizobium meliloti*. Esta región comprende el promotor -12, -24 y las secuencias activadoras upstream (UAS). La primera construcción (pMB210tt) presenta en el extremo 5' de las 750 pb de *R. meliloti* los terminadores transcripcionales de los genes ribosomales de *E. coli*, *rrnBT1* y *rrnBT2*. Esto con el fin de evitar una posible transcripción originada en la secuencia del pMB210. La segunda construcción (pMB210-A1), contiene la secuencia de *R. meliloti* hasta las secuencias activadoras (UAS). Finalmente, la tercera (pMB210-A2), presenta solamente el promotor -12, -24 de la región de control de *nifH*, eliminando el resto de la secuencia de *R. meliloti* incluyendo las UAS.

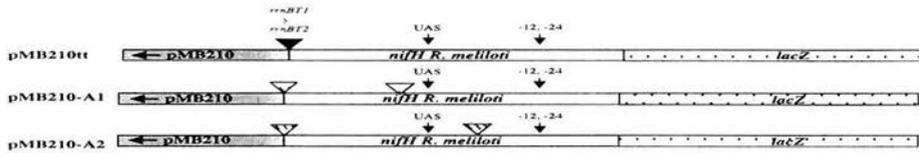


Figura 18. Plásmidos reporteros para identificar la región en pMB210 que origina la activación transcripcional por las mutantes M4 y M5. A pMB210tt se le adicionaron terminadores transcripcionales en el extremo 5' de las 750 pb de *nifH* de *R. meliloti*. pMB210-A1, contiene la secuencia del promotor *nifH* hasta las secuencias activadoras (UAS). pMB210-A2 contiene sólo el promotor -12, -24 de *nifH*. Todas las construcciones se basan en el plásmido pMB210. El triángulo con color sólido indica inserción de secuencia. Las regiones entre los triángulos rayados son secuencias eliminadas.

Con cada uno de ellos se realizaron ensayos de activación transcripcional *in vivo* con las mutantes M4 y M5, en presencia y en ausencia de σ^{54} . Los resultados de los ensayos fueron muy similares entre sí, por lo que sólo se presenta en la figura 19 los obtenidos con pMB210tt. En este ensayo (fig. 19), se observa claramente que ambas mutantes (M4 y M5) en la presencia de σ^{54} (MC1061) mantienen la misma tendencia que en ensayos transcripcionales anteriores. Cuando σ^{54} no está presente (MC1061 Δ *rpoN*), se observa el mismo incremento en las mutantes M4 y M5. Estos resultados indican que no existe una transcripción originada en la secuencia 5' de la región de control *nifH*, que sea independiente de σ^{54} .

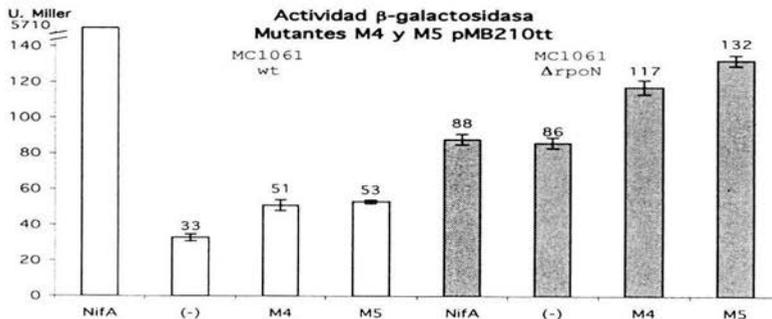


Figura 19. Activación transcripcional de pMB210tt con las variantes M4 y M5 en presencia o ausencia de σ^{54} . Los datos de esta gráfica reflejan los resultados realizados también con las construcciones pMB210-A1 y pMB210-A2. Las mutantes M4 y M5 en ausencia de σ^{54} mantienen la activación transcripcional de nuestros reporteros por arriba de los controles, indicando finalmente una independencia de la activación por el factor σ^{54} . NifA:pM1-k; [-]:pUC18. Uso de reporteros pMB210tt, pMB210-A1 y pMB210-A2. Crecimiento aeróbico en LB y a 30° C.

Decidimos probar nuestras mutantes M4 y M5 en un sistema bajo un promotor dependiente de σ^{70} , con el fin de determinar si estas activaban dicho promotor. Utilizando el vector pRO310 (Oropeza *et al.*, 1999), que presenta la región promotora -10, -35 del gene *ompS1* de *Salmonella typhi* fusionada a *lacZ*, realizamos ensayos de activación transcripcional *in vivo* con las mutantes M4 y M5 en la cepa MC1061 Δ *rpoN*. Los resultados descritos en la figura 20, indican claramente que ambas mutantes son incapaces de activar el promotor σ^{70} presente en el vector pRO310 y que la activación observada es solamente la basal.

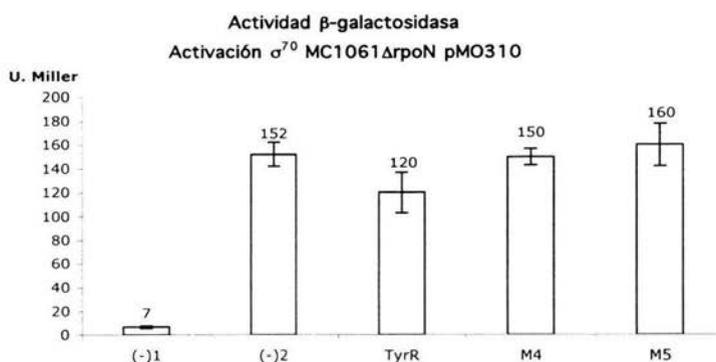


Figura 20. Activación transcripcional de M4 y M5 en un promotor σ^{70} (pRO310). En el ensayo transcripcional se utilizó un reportero que presenta una región de control σ^{70} fusionado a *lacZ*. Se observó que las mutantes M4 y M5 no generan una activación mayor que la obtenida por los controles evaluados. El ensayo se realizó en la cepa MC1061 Δ *rpoN*. [-]1: cepa sola; [-]2: cepa/pRO310. Crecimiento aeróbico en LB y a 37° C.

Finalmente, se realizó un ensayo con las mutantes M4 y M5, donde se utilizó una mutante σ^{54} de *Klebsiella pneumoniae* incapaz de activar la transcripción (construida en el Laboratorio, datos no publicados). Al utilizar esta mutante, que se une al ADN pero no activa la transcripción, bloqueamos el promotor -12, -24, del plásmido pMB210. Los resultados del ensayo son mostrados en la figura 21. En este ensayo se observa una disminución general de las actividades, que puede deberse a la utilización del gen σ^{54} de *K. pneumoniae*. Sin embargo, en M4 y M5 la baja activación observada en presencia de la mutante σ^{54} , nos sugiere que muy probablemente dicha activación, no proviene de la región 3' del promotor, sino que está relacionada con una región muy cercana al promotor. Estos datos por lo tanto, fortalecen nuestra hipótesis de un promotor sobrelapado en el promotor

-12, -24, de *nifH*. La actividad transcripcional alta observada en TyrRcentralWT en este ensayo, quizá sea un efecto del estado metabólico de las células, ya que en otros ensayos se mantuvo por debajo de las mutantes M4 y M5.

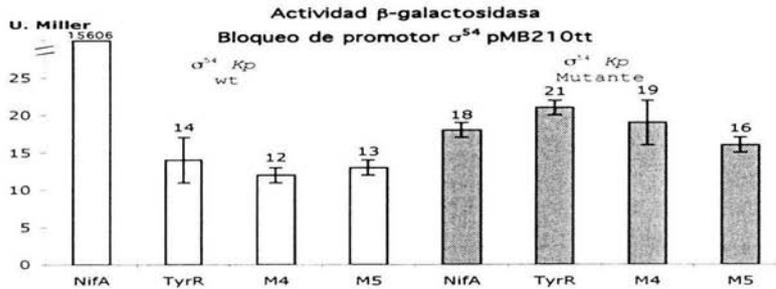


Figura 21. Bloqueo de promotor σ^{54} en la activación transcripcional de M4 y M5. Se observa claramente la disminución en la activación transcripcional generada por las mutantes M4 y M5 cuando la cepa MC1061 Δ *rpoN* es complementada con un σ^{54} incapaz de activar la transcripción dependiente de σ^{54} . Estos datos dan fuerza a nuestra hipótesis de un promotor sobrelapado ó muy cercano al promotor σ^{54} presente en el pMB210tt (pMB210tt, presenta terminadores transcripcionales en el 5' del promotor). Crecimiento aeróbico en LB y a 30° C.

Discusión.

En general, el amplio conocimiento sobre evolución molecular *in vitro* (llámese ingeniería de proteínas o evolución dirigida), se ha enfocado a la búsqueda de nuevas propiedades funcionales, o a la optimización de las ya existentes. En una menor proporción, este conocimiento se ha utilizado para el esclarecimiento de los múltiples procesos evolutivos, hasta ahora poco estudiados. Nosotros buscamos restablecer una función ancestral del dominio central de TyrR de *Escherichia coli*, modificando parte de su secuencia con el fin de comprender cuales fueron los determinantes para la activación dependiente de σ^{54} que ha perdido esta proteína al adquirir su nueva función.

Como ya se ha detallado anteriormente, TyrR es un regulador (activador/represor) de genes dependientes de σ^{70} (Pittard and Davidson, 1991), mientras que las EBPs con un mecanismo totalmente distinto, activan los genes bajo la dependencia del factor σ^{54} (Morett and Segovia, 1993). Nosotros esperábamos una mayor activación en *trans* con el dominio central de PhhR de *Pseudomonas aeruginosa* que la que se observó (fig. 11). Aunque las diferencias observadas en los resultados no fueron muy contundentes, pUC18PhhR siempre se mantuvo por arriba de los controles evaluados (fig. 11). La activación en *trans* por el

dominio central ha sido determinada experimentalmente en varias EBPs siendo un claro ejemplo la proteína NifA de *Rhizobium meliloti* (Huala and Ausubel, 1989). Entre las probables causas de la baja activación en *trans* observada con pUC18PhhR se encuentra la presencia en el dominio central de regiones C3 y C4 distintas al consenso de las EBPs ya descritas por Song y Jensen (1996). Sin embargo, estas modificaciones pueden ser una respuesta natural a los requerimientos celulares. También, la ausencia en la construcción pUC18PhhR de los dominios adicionales (N-terminal: sensor de la señal; y C-terminal: pegado a las UAS) podría estar relacionada con la baja activación observada. Esto es respaldado por la observación en NifA de *Bradyrhizobium japonicum*, de una unión fuerte por el C-terminal a las secuencias activadoras upstream, compensando una baja capacidad de activación del dominio central (Juarez *et al.*, 2000). Finalmente, queda también la posibilidad de que σ^{54} de *Escherichia coli* presente en los ensayos transcripcionales *in vivo*, no funcione adecuadamente con las EBPs de *Pseudomonas aeruginosa*, como lo propusieron Song y Jensen (1996) cuando realizaron los experimentos del activador PhhR en *E. coli*.

Al observar que el dominio central presentaba una actividad baja, decidimos introducir sólo parte de la región C3 de PhhR a TyrR, tomando en cuenta que esta región sí es funcional en PhhR (Song and Jensen, 1996) y que TyrR presenta las regiones de unión e hidrólisis de ATP, así como los requerimientos para la formación de oligómeros funcionales más conservadas que PhhR (Dixon *et al.*, 2002; Pittard and Davidson, 1991). Todas estas regiones son necesarias para la activación transcripcional de genes dependientes de σ^{54} . En los resultados descritos en la figura 11, se observó que TyrR/PhhR incrementó la activación transcripcional del reportero con relación a las construcciones TyrRcentralWT y TyrRcentralGAFTGA. TyrRcentralGAFTGA presenta la secuencia consenso de las EBPs para la región C3 y en un estudio anterior se había observado que no activaba la transcripción dependiente de σ^{54} (Flores H., datos no publicados). Por lo tanto, el incremento en la activación es un indicio de que los cambios introducidos lograron nuestro propósito.

Al observar estos resultados, también generamos una nueva mutante adicional (TyrR/PhhR+arg) con otras modificaciones en la región C3 buscando incrementar más la activación transcripcional. Sin embargo, con ésta construcción disminuyó

considerablemente la activación (fig. 11). Estos datos sugieren que los nuevos residuos de PhhR introducidos en TyrR/PhhR, causaron un efecto negativo en la activación transcripcional. Sería necesario cambiar un residuo a la vez para identificar cual ó cuales son los cambios nocivos en la construcción. Probablemente estos residuos se encuentren interactuando con otras regiones presentes en la secuencia de PhhR, pero no en la de TyrR/PhhR.

Durante la construcción de la mutante TyrR/PhhR+arg, se generó la mutante TyrR191-285 que expresa la región amino del dominio central por la presencia de un codón de término en la región C3. Esta mutante, junto con pUCTyrZ, presentaron un incremento en la activación con relación a los controles TyrRcentralWT y TyrRcentralGAFTGA (tabla 6). Al observar estos datos, pensamos que la región expresada era importante para la activación transcripcional de genes dependientes de σ^{54} y que probablemente presentara los requerimientos mínimos para la activación. Sin embargo, cuando la mutante pUCTyrRZ fue evaluada en ausencia de σ^{54} (MC1061 Δ rp0N) (fig. 14) no se pudo observar claramente una activación transcripcional dependiente de σ^{54} . Nosotros proponemos para un trabajo posterior la caracterización de esta mutante con ensayos de transcripción *in vitro*.

Por otro lado, al aplicar evolución dirigida a la mutante TyrR/PhhR para incrementar su capacidad de activación transcripcional, se obtuvieron 5 mutantes. De ellas, M4 y M5 fueron sometidas a ensayos transcripcionales (figs. 15 y 17). Ambas en presencia de σ^{54} (MC1061) activaron más la transcripción de nuestro reportero que la mutante parental TyrR/PhhR. Existen estudios en EBPs donde se ha aplicado evolución dirigida, con el fin de incrementar la activación transcripcional y en donde han logrado buenos resultados (Juarez *et al.*, 2000). Cuando se analizó la secuencia de la mutantes M4 y M5 (fig. 16), se observó que los cambios presentes en ambas, no se localizaban en las regiones de importancia funcional descritas por Morett y Segovia (1993). Sin embargo, cuando fueron sobrepuestas en la estructura de NtrCI de *Aquifex aeolicus* (Lee *et al.*, 2003) se observó que la mayoría se encontraban en la región de interfase monómero-monómero del oligómero funcional. Esto nos indicó que las modificaciones generadas probablemente promuevan la formación de los oligómeros activos en M4 y M5.

Cuando se verificó la actividad transcripcional dependiente de σ^{54} en ensayos transcripcionales con la cepa MC1061 Δ rp0N (fig. 17), se observó que las mutantes M4 y

M5 continuaban activando la transcripción del reportero aún en ausencia de σ^{54} . Estos resultados indicaban que la activación observada probablemente era originada por un promotor que estuviera presente en el reportero pMB210 y que fuera dependiente de otro factor σ . Esto nos llevo a la búsqueda del origen de esta activación.

Se construyeron 3 vectores a partir del pMB210 para buscar una posible activación en la región 5' de la región de control *nifH* de *R. meliloti*, también, se realizaron ensayos transcripcionales con un reportero σ^{70} dependiente, y se buscó bloquear con un σ^{54} mutante el promotor -12, -24, del pMB210. Los resultados de estos experimentos (figs. 19, 20 y 21) indicaron un probable solapamiento de promotores σ^{54} y σ^{70} , dando como resultado la activación observada en los ensayos transcripcionales anteriores por las mutantes M4 y M5. Se ha determinado por datos experimentales la presencia de promotores solapados en la región de control *nifA* de *Bradyrhizobium japonicum* inicialmente σ^{54} dependiente (Barrios *et al.*, 1995). También se observó que en ausencia de σ^{54} la activación transcripcional basal de los reporteros de la serie pMB210 fue mayor que cuando se encontraba σ^{54} presente, indicando un fuerte bloqueo de la transcripción generada en la región de control *nifH* de *R. meliloti*.

Los resultados obtenidos demuestran que existen formas distintas de activar los promotores σ^{54} dependientes ensayados en este trabajo. Nosotros al buscar variantes de TyrR que activaran más estos promotores, logramos obtener candidatas, aunque no las esperadas. Entre las causas probables de esta selección de variantes estaría la existencia de un promotor solapado en la región de control *nifH* de *R. meliloti*, por lo que se debe evitar en estudios posteriores este tipo de activación. Otra explicación sería la presencia en la región de control -12 y -24 de una estructura característica que permita una mayor afinidad de la ARNpol por este tipo de regiones en el ADN y que esta haya sido incrementada por proteínas como las mutantes M4 y M5.

También, creemos que es necesario explorar más el espacio de secuencia presente en el dominio central de la proteína TyrR para poder lograr nuestro objetivo principal, y proponemos la creación de nuevos bancos mutagénicos de la construcción TyrR/PhhR para ver si es posible obtener el fenotipo esperado. También con el uso de nuevas estrategias, como son la recombinación entre genes homólogos y entre genes no homólogos (por la baja

identidad de secuencia que presenta el dominio central de TyrR con relación a otros activadores transcripcionales de la familia de las EBPs), se pueda lograr la meta planteada.

Los limitados conocimientos sobre la relaciones estructura función de la proteínas que integran las EBPs nos dificulta el uso de la ingeniería de proteínas, que junto con la mutagénesis al azar y la recombinación *in vitro*, son las únicas herramientas desarrolladas para explorar los mecanismos que se generan durante la evolución natural de las proteínas.

CONCLUSIONES.

Con la búsqueda de EBPs en los genomas bacterianos totalmente secuenciados, identificamos 14 eventos de reclutamiento en 12 especies que no presentan σ^{54} . Estas proteínas presentaron diversas estructuras modulares y en la mayoría se observó pérdida de secuencia en la región C3 con excepción de las proteínas de *Fusobacterium nucleatum*.

También se identificaron y describieron con detalle casos de interés como el evento de reclutamiento reciente en Neisserias, y de la presencia de EBPs únicas del par funcional EBP- σ^{54} en clamidias y *Helicobacter pylori*, las cuales no presentaron el dominio de unión a las secuencias activadoras (C-terminal).

Por otra parte, con los resultados obtenidos en la evolución dirigida de TyrR, no fue posible restaurar la actividad dependiente de σ^{54} al dominio central de TyrR de *E. coli*. Las modificaciones dirigidas a la región C3 y las generadas por mutagénesis al azar no fueron suficientes para restaurar dicha actividad, indicando que existen otras posiciones no exploradas que son necesarias para lograr esta meta.

Finalmente, los métodos de tamizado y selección utilizados en esta parte permitieron seleccionar algunas variantes, aunque no con las características esperadas. Por lo tanto, es necesario modificar los reporteros usados para obtener las variantes deseadas.

APÉNDICE.

Medios de cultivo.

LB

10 g triptona.
 5 g extracto de levadura.
 5 g NaCl.
 1 ml NaOH 1N.
 Antibióticos.
 H₂O cuanto baste para 1 litro.
 * Para LB sólido, añadir 15 g de agar.

LB sacarosa 5 %

10 g triptona.
 5 g extracto de levadura.
 5 g NaCl.
 1 ml NaOH 1N.
 Antibióticos
 H₂O cuanto baste para 1 litro.
 5 % sacarosa
 * Para LB sólido, añadir 15 g de agar.

M9

6 g Na₂HPO₄
 3 g KH₂PO₄
 1 g NH₄Cl
 0.5 g NaCl
 1 ml MgSO₄•7H₂O
 10 ml glucosa 20 %
 0.1 ml vitamina B1 (tiamina) 0.5 %
 5 ml casaminoácidos 20 %
 Antibióticos.
 * Para M9 sólido, añadir 15 g de agar.

W-arg

10.4 g K₂HPO₂
 5.2 g KH₂PO₄
 20 ml glucosa 20%
 1 ml tiamina 5 %
 0.860 ml MgSO₄ 1M
 1 g L-arginina
 5 ml de L-leucina 200x
 Antibióticos.
 * W-arg sólido, añadir 15 g de agar.

Técnicas de Biología Molecular.

PCR (reacción).

Buffer de amplificación 10X (Gibco)	5 µl
500 mM KCl	
100 mM Tris pH 8.3	
15 mM MgCl ₂	
0.1 % gelatina	
Mezcla de 4 dNTPs Biolabs (c/u a 2.5 mM)	5 µl
Oligonucleótido 1	10 pmoles
Oligonucleótido 2	10 pmoles
ADN templado	5 µg
<i>Taq</i> polimerasa (invitrogen).	2.5 unidades
H ₂ O	para volúmen final 50 µl

* Uso Dimetil sulfóxido (DMSO) dependiendo de la reacción.

PCR Mutagénica (reacción).

Buffer de amplificación 10X (Gibco)	5 µl
500 mM KCl	
100 mM Tris pH 8.3	
70 mM MgCl ₂	
0.1 % gelatina	
Mezcla 10X de 4 dNTPs Biolabs	5 µl
2 mM dGTP	
2 mM dATP	
10 mM dCTP	
10 mM dTTP	
Oligonucleótido 1	10 pmoles
Oligonucleótido 2	10 pmoles
ADN templado	5 µg
MnCl ₂ 5 mM	5 µl
Taq polimerasa (invitrogen).	2.5 unidades
H ₂ O	para volúmen final 50 µl
* Uso Dimetil sulfoxido (DMSO) dependiendo de la reacción.	

Ligaciones en Vectores (reacción).

Buffer para ADN ligasa T4 10X (Roche).	2 µl
660 mM Tris-Hcl	
50 mM MgCl ₂	
50 mM DTT	
10 mM ATP	
pH 7.5	
ADN ligasa T4 (Roche)	1 u
Vector digerido	20 fentomoles
Inserto digerido	60 fentomoles
H ₂ O	para volúmen final 20 µl
* Temperatura de incubación, 16° C toda la noche.	

Digestiones con Enzimas de Restricción (reacción).

Buffers 10X (Amersham's) dependiendo enzima	1 µl
ADN (vector/producto de PCR)	5 µg
Enzima 1 (Amersham)	3 u
Enzima 2 (Amersham)	3 u
H ₂ O	para volúmen final 10 µl
* Temperatura de reacción dependiendo la enzima.	
** Uso de BSA (Amersham) dependiendo la enzima.	

Purificación de Producto de PCR (Roche).

1. Identificar y cortar la banda de ADN del gel de agarosa para introducirla a un tubo de 1.5 ml de microcentrifuga.
2. Adicionar 30 μ l de buffer de unión por cada 100 mg de agarosa.
3. Disolver la agarosa agitando el tubo e incubar a 56° C por 10 min.
4. Adicionar 150 μ l de isopropanol por cada 100 mg de agarosa disuelta y agitar fuertemente.
5. Pasar por columna el contenido del tubo y centrifugar a 14 000 rpm 1 min.
6. Eliminar el líquido y adicionar a la columna 500 μ l de buffer de lavado para centrifugar a 14 000 rpm 1 min.
7. Eliminar el líquido y adicionar 200 μ l de buffer de lavado para centrifugar a 14 000 rpm 1 min.
8. Eliminar el líquido y adicionar 60 μ l de buffer de elución a la columna para centrifugar a 14 000 rpm 1 min.
9. Almacenar la solución eluida que contiene el ADN purificado.

Purificación de Plásmido (Roche).

1. Centrifugar el cultivo de toda la noche a 9 000 rpm 30 seg y resuspender el pellet en 250 μ l de buffer de suspensión+RNasa.
2. Mezclar suavemente con 250 ml de buffer de lisis e incubar a temperatura ambiente por 5 min.
3. Mezclar suavemente con 350 ml de buffer de unión e incubar en hielo por 5 min.
4. Centrifugar a 14 000 rpm por 10 min.
5. Pasar por columna el sobrenadante y centrifugar a 14 000 rpm 1 min.
6. Eliminar el líquido y adicionar a la columna 700 μ l de buffer de lavado para centrifugar a 14 000 rpm 1 min.
7. Eliminar los restos del líquido por centrifugación a 14 000 rpm 1 min.
8. Adicionar 60 μ l de buffer de elución a la columna y centrifugar a 14 000 rpm 1 min.
9. Almacenar la solución eluida que contiene el ADN purificado.

Buffers.

Buffer Z

- 16.1 g $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$
- 5.5 g $\text{Na}_2\text{PO}_4 \cdot \text{H}_2\text{O}$
- 0.75 g KCl
- 0.246 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$
- 2.7 ml β -mercaptoetanol
- H_2O para volumen final de 1000 ml
- Ajustar pH a 7.0

BIBLIOGRAFÍA.

- Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H., and Kurland, C.G. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**: 133-140.
- Andrews, A.E., Dickson, B., Lawley, B., Cobbett, C., and Pittard, A.J. (1991) Importance of the position of TYR R boxes for repression and activation of the tyrP and aroF genes in *Escherichia coli*. *J Bacteriol* **173**: 5079-5085.
- Aurora, R., Srinivasan, R., and Rose, G.D. (1994) Rules for alpha-helix termination by glycine. *Science* **264**: 1126-1130.
- Austin, S., and Dixon, R. (1992) The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent. *Embo J* **11**: 2219-2228.
- Austin, S., Buck, M., Cannon, W., Eydmann, T., and Dixon, R. (1994) Purification and in vitro activities of the native nitrogen fixation control proteins NifA and NifL. *J Bacteriol* **176**: 3460-3465.
- Barnett, P., Tabak, H.F., and Hetteema, E.H. (2000) Nuclear receptors arose from pre-existing protein modules during evolution. *Trends Biochem Sci* **25**: 227-228.
- 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.
- Barrios, H., Valderrama, B., and Morett, E. (1999) Compilation and analysis of sigma(54)-dependent promoter sequences. *Nucleic Acids Res* **27**: 4305-4313.
- Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J., and Hopwood, D.A. (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**: 141-147.
- Berger, D.K., Narberhaus, F., and Kustu, S. (1994) The isolated catalytic domain of NIFA, a bacterial enhancer-binding protein, activates transcription in vitro: activation is inhibited by NIFL. *Proc Natl Acad Sci U S A* **91**: 103-107.
- Buck, M., and Cannon, W. (1992) Specific binding of the transcription factor sigma-54 to promoter DNA. *Nature* **358**: 422-424.
- Buck, M., S. Miller, M. Drummond, and R. Nixon (1986) *Upstream activator sequence are present in the promoters of nitrogen fixation genes*. London: Nature.
- Cadwell, R.C., and Joyce, G.F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Appl* **2**: 28-33.
- Cadwell, R.C., and Joyce, G.F. (1994) Mutagenic PCR. *PCR Methods Appl* **3**: S136-140.
- Carmona, M., and Magasanik, B. (1996) Activation of transcription at sigma 54-dependent promoters on linear templates requires intrinsic or induced bending of the DNA. *J Mol Biol* **261**: 348-356.

- Carrick, C.S., Fyfe, J.A., and Davies, J.K. (1997) The normally silent sigma54 promoters upstream of the pilE genes of both *Neisseria gonorrhoeae* and *Neisseria meningitidis* are functional when transferred to *Pseudomonas aeruginosa*. *Gene* **198**: 89-97.
- Carrick, C.S., Fyfe, J.A., and Davies, J.K. (2000) The genome of *Neisseria gonorrhoeae* retains the remnants of a two-component regulatory system that once controlled piliation. *FEMS Microbiol Lett* **186**: 197-201.
- Chen, K., and Arnold, F.H. (1993) Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc Natl Acad Sci U S A* **90**: 5618-5622.
- Choy, H.E., and Adhya, S. (1992) Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. *Proc Natl Acad Sci U S A* **89**: 11264-11268.
- Choy, H.E., Hanger, R.R., Aki, T., Mahoney, M., Murakami, K., Ishihama, A., and Adhya, S. (1997) Repression and activation of promoter-bound RNA polymerase activity by Gal repressor. *J Mol Biol* **272**: 293-300.
- Chye, M.L., and Pittard, J. (1987) Transcription control of the aroP gene in *Escherichia coli* K-12: analysis of operator mutants. *J Bacteriol* **169**: 386-393.
- Copley, R.R., and Bork, P. (2000) Homology among (betaalpha)(8) barrels: implications for the evolution of metabolic pathways. *J Mol Biol* **303**: 627-641.
- Cramer, A., Raillard, S.A., Bermudez, E., and Stemmer, W.P. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**: 288-291.
- Cui, J., and Somerville, R.L. (1993a) The TyrR protein of *Escherichia coli*, analysis by limited proteolysis of domain structure and ligand-mediated conformational changes. *J Biol Chem* **268**: 5040-5047.
- Cui, J., and Somerville, R.L. (1993b) A mutational analysis of the structural basis for transcriptional activation and monomer-monomer interaction in the TyrR system of *Escherichia coli* K-12. *J Bacteriol* **175**: 1777-1784.
- Debarbouille, M., Martin-Verstraete, I., Klier, A., and Rapoport, G. (1991) The transcriptional regulator LevR of *Bacillus subtilis* has domains homologous to both sigma 54- and phosphotransferase system-dependent regulators. *Proc Natl Acad Sci U S A* **88**: 2212-2216.
- Ditta, G., Schmidhauser, T., Yakobson, E., Lu, P., Liang, X.W., Finlay, D.R., Guiney, D., and Helinski, D.R. (1985) Plasmids related to the broad host range vector, pRK290, useful for gene cloning and for monitoring gene expression. *Plasmid* **13**: 149-153.
- Dixon, M.P., Pau, R.N., Howlett, G.J., Dunstan, D.E., Sawyer, W.H., and Davidson, B.E. (2002) The central domain of *Escherichia coli* TyrR is responsible for hexamerization associated with tyrosine-mediated repression of gene expression. *J Biol Chem* **277**: 23186-23192.
- Drummond, M., Whitty, P., and Wootton, J. (1986) Sequence and domain relationships of ntrC and nifA from *Klebsiella pneumoniae*: homologies to other regulatory proteins. *Embo J* **5**: 441-447.
- Elderkin, S., Jones, S., Schumacher, J., Studholme, D., and Buck, M. (2002) Mechanism of action of the *Escherichia coli* phage shock protein PspA in repression of the AAA family transcription factor PspF. *J Mol Biol* **320**: 23-37.

- Foster-Hartnett, D., and Kranz, R.G. (1992) Analysis of the promoters and upstream sequences of *nifA1* and *nifA2* in *Rhodobacter capsulatus*; activation requires *ntnC* but not *rpoN*. *Mol Microbiol* **6**: 1049-1060.
- Garmendia, J., and de Lorenzo, V. (2000) The role of the interdomain B linker in the activation of the XylR protein of *Pseudomonas putida*. *Mol Microbiol* **38**: 401-410.
- Gilbert, W. (1978) Why genes in pieces? *Nature* **271**: 501.
- Gogarten, J.P. (1994) Which is the most conserved group of proteins? Homology-orthology, paralogy, xenology, and the fusion of independent lineages. *J Mol Evol* **39**: 541-543.
- Goldfarb, D.S., Wong, S.L., Kudo, T., and Doi, R.H. (1983) A temporally regulated promoter from *Bacillus subtilis* is transcribed only by an RNA polymerase with a 37,000 dalton sigma factor. *Mol Gen Genet* **191**: 319-325.
- González, V., Olvera, L., Soberon, X., and Morett, E. (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.
- Gowrishankar, J., and Pittard, A.J. (1998) Superimposition of *tyrR* protein-mediated regulation on osmoresponsive transcription of *Escherichia coli* *proU* in vivo. *J Bacteriol* **180**: 6743-6748.
- Grande, R.A., Valderrama, B., and Morett, E. (1999) Suppression analysis of positive control mutants of NifA reveals two overlapping promoters for *Klebsiella pneumoniae* *rpoN*. *J Mol Biol* **294**: 291-298.
- Grifantini, R., Sebastian, S., Frigimelica, E., Draghi, M., Bartolini, E., Muzzi, A., Rappuoli, R., Grandi, G., and Genco, C.A. (2003) Identification of iron-activated and -repressed Fur-dependent genes by transcriptome analysis of *Neisseria meningitidis* group B. *Proc Natl Acad Sci U S A* **100**: 9542-9547.
- Guo, Y., and Gralla, J.D. (1997) DNA-binding determinants of sigma 54 as deduced from libraries of mutations. *J Bacteriol* **179**: 1239-1245.
- Heatwole, V.M., and Somerville, R.L. (1991) The tryptophan-specific permease gene, *mtr*, is differentially regulated by the tryptophan and tyrosine repressors in *Escherichia coli* K-12. *J Bacteriol* **173**: 3601-3604.
- Helmann, J.D. (1991) Alternative sigma factors and the regulation of flagellar gene expression. *Mol Microbiol* **5**: 2875-2882.
- Hilgenfeld, R. (1995) Regulatory GTPases. *Curr Opin Struct Biol* **5**: 810-817.
- Hochschild, A., and Dove, S.L. (1998) Protein-protein contacts that activate and repress prokaryotic transcription. *Cell* **92**: 597-600.
- Hoover, T.R., Santero, E., Porter, S., and Kustu, S. (1990) The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. *Cell* **63**: 11-22.
- Hsieh, M., and Gralla, J.D. (1994) Analysis of the N-terminal leucine heptad and hexad repeats of sigma 54. *J Mol Biol* **239**: 15-24.
- Hsieh, M., Tintut, Y., and Gralla, J.D. (1994) Functional roles for the glutamines within the glutamine-rich region of the transcription factor sigma 54. *J Biol Chem* **269**: 373-378.
- Huala, E., and Ausubel, F.M. (1989) The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti* *nifH* promoter. *J Bacteriol* **171**: 3354-3365.

- Hwang, J.S., Yang, J., and Pittard, A.J. (1997) Critical base pairs and amino acid residues for protein-DNA interaction between the TyrR protein and tyrP operator of *Escherichia coli*. *J Bacteriol* **179**: 1051-1058.
- Hwang, J.S., Yang, J., and Pittard, A.J. (1999) Specific contacts between residues in the DNA-binding domain of the TyrR protein and bases in the operator of the tyrP gene of *Escherichia coli*. *J Bacteriol* **181**: 2338-2345.
- Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S. (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **21**: 526-531.
- Ishimoto, K.S., and Lory, S. (1992) Identification of pilR, which encodes a transcriptional activator of the *Pseudomonas aeruginosa* pilin gene. *J Bacteriol* **174**: 3514-3521.
- Jagannathan, A., Constantinidou, C., and Penn, C.W. (2001) Roles of rpoN, fliA, and flgR in expression of flagella in *Campylobacter jejuni*. *J Bacteriol* **183**: 2937-2942.
- Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. *Annu Rev Microbiol* **30**: 409-425.
- Jensen, R.A., and Gu, W. (1996) Evolutionary recruitment of biochemically specialized subdivisions of Family I within the protein superfamily of aminotransferases. *J Bacteriol* **178**: 2161-2171.
- Jovanovic, G., Rakonjac, J., and Model, P. (1999) In vivo and in vitro activities of the *Escherichia coli* sigma54 transcription activator, PspF, and its DNA-binding mutant, PspFDeltaHTH. *J Mol Biol* **285**: 469-483.
- Juarez, K., Flores, H., Davila, S., Olvera, L., Gonzalez, V., and Morett, E. (2000) Reciprocal domain evolution within a transactivator in a restricted sequence space. *Proc Natl Acad Sci U S A* **97**: 3314-3318.
- Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions (supplement). *DNA Res* **3**: 185-209.
- Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S., Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A., Nakazaki, N., Shimpo, S., Sugimoto, M., Takazawa, M., Yamada, M., Yasuda, M., and Tabata, S. (2001) Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res* **8**: 205-213; 227-253.
- Kapatral, V., Anderson, I., Ivanova, N., Reznik, G., Los, T., Lykidis, A., Bhattacharyya, A., Bartman, A., Gardner, W., Grechkin, G., Zhu, L., Vasieva, O., Chu, L., Kogan, Y., Chaga, O., Goltzman, E., Bernal, A., Larsen, N., D'Souza, M., Walunas, T., Pusch, G., Haselkorn, R., Fonstein, M., Kyrpides, N., and Overbeek, R. (2002) Genome sequence and analysis of the oral bacterium *Fusobacterium nucleatum* strain ATCC 25586. *J Bacteriol* **184**: 2005-2018.
- Kasian, P.A., Davidson, B.E., and Pittard, J. (1986) Molecular analysis of the promoter operator region of the *Escherichia coli* K-12 tyrP gene. *J Bacteriol* **167**: 556-561.

- Kern, D., Volkman, B.F., Luginbuhl, P., Nohaile, M.J., Kustu, S., and Wemmer, D.E. (1999) Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* **402**: 894-898.
- Klose, K.E., North, A.K., Stedman, K.M., and Kustu, S. (1994) The major dimerization determinants of the nitrogen regulatory protein NTRC from enteric bacteria lie in its carboxy-terminal domain. *J Mol Biol* **241**: 233-245.
- Kroos, L., Zhang, B., Ichikawa, H., and Yu, Y.T. (1999) Control of sigma factor activity during *Bacillus subtilis* sporulation. *Mol Microbiol* **31**: 1285-1294.
- Kustu, S., Santero, E., Keener, J., Popham, D., and Weiss, D. (1989) Expression of sigma 54 (ntrA)-dependent genes is probably united by a common mechanism. *Microbiol Rev* **53**: 367-376.
- Kwok, T. (1998) The domain structure of the regulatory protein TyrR from *E. coli* K-12.: University of Melbourne.
- Laskos, L., Dillard, J.P., Seifert, H.S., Fyfe, J.A., and Davies, J.K. (1998) The pathogenic neisseriae contain an inactive rpoN gene and do not utilize the pilE sigma54 promoter. *Gene* **208**: 95-102.
- Lawley, B., Fujita, N., Ishihama, A., and Pittard, A.J. (1995) The TyrR protein of *Escherichia coli* is a class I transcription activator. *J Bacteriol* **177**: 238-241.
- Lee, J.H., Scholl, D., Nixon, B.T., and Hoover, T.R. (1994) Constitutive ATP hydrolysis and transcription activation by a stable, truncated form of *Rhizobium meliloti* DCTD, a sigma 54-dependent transcriptional activator. *J Biol Chem* **269**: 20401-20409.
- Lee, J.H., and Hoover, T.R. (1995) Protein crosslinking studies suggest that *Rhizobium meliloti* C4-dicarboxylic acid transport protein D, a sigma 54-dependent transcriptional activator, interacts with sigma 54 and the beta subunit of RNA polymerase. *Proc Natl Acad Sci U S A* **92**: 9702-9706.
- Lee, S.Y., De La Torre, A., Yan, D., Kustu, S., Nixon, B.T., and Wemmer, D.E. (2003) Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. *Genes Dev* **17**: 2552-2563.
- Link, A.J., Phillips, D., and Church, G.M. (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* **179**: 6228-6237.
- MacKintosh, C., and MacKintosh, R.W. (1994) Inhibitors of protein kinases and phosphatases. *Trends Biochem Sci* **19**: 444-448.
- Magasanik, B. (1989) Regulation of transcription of the *glnALG* operon of *Escherichia coli* by protein phosphorylation. *Biochimie* **71**: 1005-1012.
- Mathews, S.A., and Timms, P. (2000) Identification and mapping of sigma-54 promoters in *Chlamydia trachomatis*. *J Bacteriol* **182**: 6239-6242.
- May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whittam, T.S., and Kapur, V. (2001) Complete genomic sequence of *Pasteurella multocida*, Pm70. *Proc Natl Acad Sci U S A* **98**: 3460-3465.
- Melendez-Hevia, E., Waddell, T.G., and Cascante, M. (1996) The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J Mol Evol* **43**: 293-303.

- Merrick, M.J., and Gibbins, J.R. (1985) The nucleotide sequence of the nitrogen-regulation gene *ntrA* of *Klebsiella pneumoniae* and comparison with conserved features in bacterial RNA polymerase sigma factors. *Nucleic Acids Res* **13**: 7607-7620.
- Merrick, M.J. (1992) *Regulation of nitrogen fixation genes in free-living and symbiotic bacteria*. New York: Chapman Hall.
- Merrick, M.J. (1993) In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). *Mol Microbiol* **10**: 903-909.
- Mettke, I., Fiedler, U., and Weiss, V. (1995) Mechanism of activation of a response regulator: interaction of NtrC-P dimers induces ATPase activity. *J Bacteriol* **177**: 5056-5061.
- Miller, H. (1972) *Experiments in Molecular Genetics*: Cold Spring Harbor Laboratory.
- Mills, D.R., Peterson, R.L., and Spiegelman, S. (1967) An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc Natl Acad Sci U S A* **58**: 217-224.
- Missailidis, S., Jaseja, M., Ray, P., Chittock, R., Wharton, C.W., Drake, A.F., Buck, M., and Hyde, E.I. (1999) Secondary structure of the C-terminal DNA-binding domain of the transcriptional activator NifA from *Klebsiella pneumoniae*: spectroscopic analyses. *Arch Biochem Biophys* **361**: 173-182.
- Monsalve, M., Mencia, M., Rojo, F., and Salas, M. (1996) Activation and repression of transcription at two different phage phi29 promoters are mediated by interaction of the same residues of regulatory protein p4 with RNA polymerase. *Embo J* **15**: 383-391.
- Morett, E., Cannon, W., and Buck, M. (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.
- Morett, E., and Buck, M. (1989) In vivo studies on the interaction of RNA polymerase-sigma 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., Kreutzer, R., Cannon, W., and Buck, M. (1990) The influence of the *Klebsiella pneumoniae* regulatory gene *nifL* upon the transcriptional activator protein NifA. *Mol Microbiol* **4**: 1253-1258.
- Morett, E., and Segovia, L. (1993) The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J Bacteriol* **175**: 6067-6074.
- Morett, E., and Bork, P. (1998) Evolution of new protein function: recombinational enhancer Fis originated by horizontal gene transfer from the transcriptional regulator NtrC. *FEBS Lett* **433**: 108-112.
- Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N., Shimpo, S., Takeuchi, C., Yamada, M., and Tabata, S. (2003) Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. *DNA Res* **10**: 137-145.

- Neuwald, A.F., Aravind, L., Spouge, J.L., and Koonin, E.V. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* **9**: 27-43.
- Ng, L.C., Poh, C.L., and Shingler, V. (1995) Aromatic effector activation of the NtrC-like transcriptional regulator PhhR limits the catabolic potential of the (methyl)phenol degradative pathway it controls. *J Bacteriol* **177**: 1485-1490.
- Ninfa, A.J., Reitzer, L.J., and Magasanik, B. (1987) Initiation of transcription at the bacterial *glnAp2* promoter by purified *E. coli* components is facilitated by enhancers. *Cell* **50**: 1039-1046.
- North, A.K., Weiss, D.S., Suzuki, H., Flashner, Y., and Kustu, S. (1996) Repressor forms of the enhancer-binding protein NtrC: some fail in coupling ATP hydrolysis to open complex formation by sigma 54-holoenzyme. *J Mol Biol* **260**: 317-331.
- O'Neill, E., Ng, L.C., Sze, C.C., and Shingler, V. (1998) Aromatic ligand binding and intramolecular signalling of the phenol-responsive sigma54-dependent regulator DmpR. *Mol Microbiol* **28**: 131-141.
- Ogata, H., Audic, S., Renesto-Audiffren, P., Fournier, P.E., Barbe, V., Samson, D., Roux, V., Cossart, P., Weissenbach, J., Claverie, J.M., and Raoult, D. (2001) Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* **293**: 2093-2098.
- Oropeza, R., Sampieri, C.L., Puente, J.L., and Calva, E. (1999) Negative and positive regulation of the non-osmoregulated *ompS1* porin gene in *Salmonella typhi*: a novel regulatory mechanism that involves *OmpR*. *Mol Microbiol* **32**: 243-252.
- Osuna, J., Soberon, X., and Morett, E. (1997) A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. *Protein Sci* **6**: 543-555.
- Pao, G.M., and Saier, M.H., Jr. (1995) Response regulators of bacterial signal transduction systems: selective domain shuffling during evolution. *J Mol Evol* **40**: 136-154.
- Pawlowski, K., Klosse, U., and de Bruijn, F.J. (1991) Characterization of a novel *Azorhizobium caulinodans* ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation and metabolism. *Mol Gen Genet* **231**: 124-138.
- Perez-Martin, J., and De Lorenzo, V. (1995) The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor. *Proc Natl Acad Sci U S A* **92**: 9392-9396.
- Perez-Martin, J., and de Lorenzo, V. (1996) ATP binding to the sigma 54-dependent activator XylR triggers a protein multimerization cycle catalyzed by UAS DNA. *Cell* **86**: 331-339.
- Pittard, A.J., and Davidson, B.E. (1991) TyrR protein of *Escherichia coli* and its role as repressor and activator. *Mol Microbiol* **5**: 1585-1592.
- Pittard, J. (1996) The various strategies within the TyrR regulation of *Escherichia coli* to modulate gene expression. *Genes Cells* **1**: 717-725.
- Pittard, W.B., 3rd, Sorensen, R.U., and Schnatz, P.T. (1987) Pregnancy outcome in mothers with cystic fibrosis: normal neonatal immune responses. *South Med J* **80**: 344-346.
- Platt, T. (1994) Rho and RNA: models for recognition and response. *Mol Microbiol* **11**: 983-990.
- Poggio, S., Osorio, A., Dreyfus, G., and Camarena, L. (2002) The four different sigma(54) factors of *Rhodobacter sphaeroides* are not functionally interchangeable. *Mol Microbiol* **46**: 75-85.

- Popham, D.L., Szeto, D., Keener, J., and Kustu, S. (1989) Function of a bacterial activator protein that binds to transcriptional enhancers. *Science* **243**: 629-635.
- Porter, S.C., North, A.K., Wedel, A.B., and Kustu, S. (1993) Oligomerization of NTRC at the *glnA* enhancer is required for transcriptional activation. *Genes Dev* **7**: 2258-2273.
- Ptashne, M., and Gann, A. (1997) Transcriptional activation by recruitment. *Nature* **386**: 569-577.
- Richardson, J.P., and Ruteshouser, E.C. (1986) rho factor-dependent transcription termination. Interference by a mutant rho. *J Mol Biol* **189**: 413-419.
- Richardson, J.P. (2003) Loading Rho to terminate transcription. *Cell* **114**: 157-159.
- Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., Arellano, A., Coleman, M., Hauser, L., Hess, W.R., Johnson, Z.I., Land, M., Lindell, D., Post, A.F., Regala, W., Shah, M., Shaw, S.L., Steglich, C., Sullivan, M.B., Ting, C.S., Tolonen, A., Webb, E.A., Zinser, E.R., and Chisholm, S.W. (2003) Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042-1047.
- Rombel, I., Peters-Wendisch, P., Mesecar, A., Thorgeirsson, T., Shin, Y.K., and Kustu, S. (1999) MgATP binding and hydrolysis determinants of NtrC, a bacterial enhancer-binding protein. *J Bacteriol* **181**: 4628-4638.
- Ronson, C.W., Nixon, B.T., Albright, L.M., and Ausubel, F.M. (1987) *Rhizobium meliloti* *ntrA* (*rpoN*) gene is required for diverse metabolic functions. *J Bacteriol* **169**: 2424-2431.
- Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990) The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**: 430-434.
- Sasse-Dwight, S., and Gralla, J.D. (1988) Probing the *Escherichia coli* *glnALG* upstream activation mechanism in vivo. *Proc Natl Acad Sci U S A* **85**: 8934-8938.
- Schlenso, V., and Bock, A. (1990) Identification and sequence analysis of the gene encoding the transcriptional activator of the formate hydrogenlyase system of *Escherichia coli*. *Mol Microbiol* **4**: 1319-1327.
- Schmidt-Dannert, C. (2001) Directed evolution of single proteins, metabolic pathways, and viruses. *Biochemistry* **40**: 13125-13136.
- Scholl, D., and Nixon, B.T. (1996) Cooperative binding of DctD to the *dctA* upstream activation sequence of *Rhizobium meliloti* is enhanced in a constitutively active truncated mutant. *J Biol Chem* **271**: 26435-26442.
- Shingler, V. (1996) Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism. *Mol Microbiol* **19**: 409-416.
- Skandalis, A., Encell, L.P., and Loeb, L.A. (1997) Creating novel enzymes by applied molecular evolution. *Chem Biol* **4**: 889-898.
- Song, J., and Jensen, R.A. (1996) PhhR, a divergently transcribed activator of the phenylalanine hydroxylase gene cluster of *Pseudomonas aeruginosa*. *Mol Microbiol* **22**: 497-507.
- Spohn, G., and Scarlato, V. (1999) Motility of *Helicobacter pylori* is coordinately regulated by the transcriptional activator FlgR, an NtrC homolog. *J Bacteriol* **181**: 593-599.
- Stemmer, W.P. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**: 389-391.

- Stock, J.B., Ninfa, A.J., and Stock, A.M. (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol Rev* **53**: 450-490.
- Studholme, D.J., and Buck, M. (2000) Novel roles of sigmaN in small genomes. *Microbiology* **146** (Pt 1): 4-5.
- Studholme, D.J., and Dixon, R. (2003) Domain architectures of sigma54-dependent transcriptional activators. *J Bacteriol* **185**: 1757-1767.
- Sugiura, M., Okamoto, T., and Takanami, M. (1970) RNA polymerase sigma-factor and the selection of initiation site. *Nature* **225**: 598-600.
- Syed, A., and Gralla, J.D. (1998) Identification of an N-terminal region of sigma 54 required for enhancer responsiveness. *J Bacteriol* **180**: 5619-5625.
- Takiguchi, M., Matsubasa, T., Amaya, Y., and Mori, M. (1989) Evolutionary aspects of urea cycle enzyme genes. *Bioessays* **10**: 163-166.
- Thony, B., and Hennecke, H. (1989) The -24/-12 promoter comes of age. *FEMS Microbiol Rev* **5**: 341-357.
- Tuli, R., and Merrick, M.J. (1988) Over-production and characterization of the nifA gene product of *Klebsiella pneumoniae*--the transcriptional activator of nif gene expression. *J Gen Microbiol* **134** (Pt 2): 425-432.
- Vale, R.D. (2000) AAA proteins. Lords of the ring. *J Cell Biol* **150**: F13-19.
- Vieira, J., and Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**: 259-268.
- Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo J* **1**: 945-951.
- Walmsley, J., Toukdarian, A., and Kennedy, C. (1994) The role of regulatory genes nifA, vnfA, anfA, nfrX, ntrC, and rpoN in expression of genes encoding the three nitrogenases of *Azotobacter vinelandii*. *Arch Microbiol* **162**: 422-429.
- Wang, P., Yang, J., Ishihama, A., and Pittard, A.J. (1998) Demonstration that the TyrR protein and RNA polymerase complex formed at the divergent P3 promoter inhibits binding of RNA polymerase to the major promoter, P1, of the *aroP* gene of *Escherichia coli*. *J Bacteriol* **180**: 5466-5472.
- Wang, Y.K., Lee, J.H., Brewer, J.M., and Hoover, T.R. (1997) A conserved region in the sigma54-dependent activator DctD is involved in both binding to RNA polymerase and coupling ATP hydrolysis to activation. *Mol Microbiol* **26**: 373-386.
- Weiss, D.S., Batut, J., Klose, K.E., Keener, J., and Kustu, S. (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.
- Wilson, T.J., Maroudas, P., Howlett, G.J., and Davidson, B.E. (1994) Ligand-induced self-association of the *Escherichia coli* regulatory protein TyrR. *J Mol Biol* **238**: 309-318.
- Wilson, T.J., Argæet, V.P., Howlett, G.J., and Davidson, B.E. (1995) Evidence for two aromatic amino acid-binding sites, one ATP-dependent and the other ATP-independent, in the *Escherichia coli* regulatory protein TyrR. *Mol Microbiol* **17**: 483-492.
- Wong, C., Tintut, Y., and Gralla, J.D. (1994) The domain structure of sigma 54 as determined by analysis of a set of deletion mutants. *J Mol Biol* **236**: 81-90.

- Xu, H., and Hoover, T.R. (2001) Transcriptional regulation at a distance in bacteria. *Curr Opin Microbiol* **4**: 138-144.
- Yang, J., Camakaris, H., and Pittard, A.J. (1993) Mutations in the *tyrR* gene of *Escherichia coli* which affect TyrR-mediated activation but not TyrR-mediated repression. *J Bacteriol* **175**: 6372-6375.
- Yang, J., Camakaris, H., and Pittard, A.J. (1996) In vitro transcriptional analysis of TyrR-mediated activation of the *mtr* and *tyrP+3* promoters of *Escherichia coli*. *J Bacteriol* **178**: 6389-6393.
- Yang, J., Wang, P., and Pittard, A.J. (1999) Mechanism of repression of the *aroP P2* promoter by the TyrR protein of *Escherichia coli*. *J Bacteriol* **181**: 6411-6418.
- Yang, J., Camakaris, H., and Pittard, J. (2002) Molecular analysis of tyrosine- and phenylalanine-mediated repression of the *tyrB* promoter by the TyrR protein of *Escherichia coli*. *Mol Microbiol* **45**: 1407-1419.
- Zhao, S., Zhu, Q., and Somerville, R.L. (2000) The $\sigma(70)$ transcription factor TyrR has zinc-stimulated phosphatase activity that is inhibited by ATP and tyrosine. *J Bacteriol* **182**: 1053-1061.
- Zhu, Q., Zhao, S., and Somerville, R.L. (1997) Expression, purification, and functional analysis of the TyrR protein of *Haemophilus influenzae*. *Protein Expr Purif* **10**: 237-246.