4
2ej.
03088

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Unidad Acádemica de los Ciclos Profesional y de Posgrado, Colegio de Ciencias y Humanidades

El Mecanismo de Activación Transcripcional a Distancia en Bacteria y la Región de Control Positivo de NifA

Tesis para obtener el grado de Doctor en Biotecnología
por Víctor Manuel González Zúñiga

TESIS CON
FALLA DE ORIGEN

258144

1998

Universidad Nacional Autónoma de México
UNAM
Biblioteca Central
Dirección General de Bibliotecas de la UNAM

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.

PAGINACION

DISCONTINUA

La investigación descrita en esta tesis fué realizada en el Departamento de Reconocimiento Molecular y Bioestructura del Instituto de Biotecnología de la UNAM. La concepción inicial del problema fué hecha por el Dr. Enrique Morett, mientras que el enfoque estructural fué debido al Dr. Xavier Soberón. Ambos fungieron como co-tutores durante estos estudios.

Indice

Resumen I

Prólogo 1

Capítulo 1. La RNA polimerasa básica. 4

1.1. La estructura de la RNA polimerasa 4

1.2. Las subunidades β y β' 9

1.3. La subunidad α 10

Capítulo 2. Estructura y función del promotor. 12

2.1.El reconocimiento RNA polimerasa-promotor. 13

2.2. La estructura primaria de los promotores 14

2.3. Los factores sigma 16

2.3.1. La familia σ^{70}. 17

2.3.1.1. Las regiones conservadas de σ^{70} 18

2.3.2. La familia σ^{54} 21

2.3.2.1. Las regiones I y II: formación del complejo abierto 23

2.3.2.2. La región III: unión con la RNA polimerasa y el promotor 25

Capítulo 3. El mecanismo de inicio de la transcripción. 27

3.1. La formacion del complejo cerrado 27

3.2. La formación del complejo abierto. 29

Capitulo 4. Regulación de la transcripción. 31

4.1. La unión con el DNA. 32

4.2. La activación cercana. 33

4.2.1. La activación mediante CAP 34

4.2.2. Interacciones entre CAP y la RNA polimerasa. 35

4.2.3. Interacciones entre λ cI y la RNA polimerasa. 37

4.3. La activacion a distancia. 39

4.3.1. Los dominios funcionales de los activadores a distancia. 40

4.3.2. La estructura de las regiones de activación. 41

Capítulo 5. La familia de activadores bacterianos a distancia 44

5.1. Estructura en dominios. 44

5.2. La fijación de nitrógeno. 48

5.3. Regulación por oxígeno de NifA 49

5.4. Patrones evolutivos. 52

Capítulo 6. El modelo actual de activación transcripcional a distancia en bacteria. 55

6. 1. El modelo del asa 55

6.2. La función de los potenciadores 58

6.3. Oligomerización del activador 59

6.4. La hidrólisis de ATP. 61

6.5. El dominio Central de las EBPs es una GTPasa 63

Capítulo 7. Contribucion original: la función de control positivo de NifA 66

"*In vivo* studies on the positive control function of NifA: a conserved hydrophobic aminoacid patch at the central domain involved in transcriptional activation"

Epílogo 92

Referencias. 98

Gratitudes. 123

a Lilia...

Resumen

La regulación positiva de muchos operones bacterianos se lleva a cabo mediante proteínas que se unen cerca o lejos del promotor. Ambos tipos de regulación positiva son revisados abundantemente en este trabajo. En el primer caso los activadores ayudan a la RNA polimerasa a unirse al promotor, estabilizándola mediante interacciones proteìna-proteína. Las regiones de interacción de estos activadores han sido caracterizadas por su efectos en la capacidad de la proteína para estimular la transcripción. Esta clase de mutantes son llamadas de control positivo.

Los activadores de acción a distancia bacterianos (EBPs) se unen a secuencias especificas en el DNA apartadas del promotor.Una RNA polimerasa formada por el factor sigma alternativo σ^{54}, unida establemente en el promotor es el blanco de estos activadores. Para iniciar la transcripción las EBPs deben contactar la RNA polimerasa-σ^{54} e hidrolizar nucleótidos trifosfatados. Esta actividad enzimática está acoplada a la formación del complejo abierto.

Como contribución original presentamos el estudio genético de una región del dominio Central de NifA, una EBP que regula los genes de fijación de nitrógeno, que probablemente está involucrada en el control positivo. Nuestros resultados indican que un amino àcido aromático y un grupo hidroxilo, son características fundamentales para la activación transcripcional. Discutimos sí el papel de esta región es contactar la RNA polimerasa o bien es necesaria para la traducción de la señal de hidrólisis de ATP con un probable cambio conformacional acoplado al inicio de la transcripción.

Summary

Positive regulation of many bacterial operons is carry out by proteins which bind to DNA sites located near or distant from the promoter. Both classes of positive control are extensively reviewed in the present thesis. In the firts class, the activator helps RNA polymerase to bind properly on the promoter, by protein-protein interacctions. The interaction regions of these molecules has been characterized by their effects on the ability of the protein to stimulate transcription. This class of mutans are called positive control (PC).

The bacterial enhancer-binding proteins (EBPs), bind to specific DNA sequences far from the promoter. A RNA polymerase holoenzyme composed by the alternative σ factor , σ^{54}, is the target for the action of the EBPs. To initate transcription, the EBPs must touch the RNA pol-σ^{54}-promoter complex and hydrolyse ATP. This activity is coupled to open comlex formation.

In our original contribution is presented a genetic study of a conserved region of the Central domain of NifA, an EBP that regulates the nitrogen fixation genes, probably involved in positive control. Our results suggest that an aromatic amino acid and a hydroxyl group are important characteristics for the activity of the protein. The role of this region as a contact zone for the RNA polymerase or as a transducing effector of the ATP hydrolys is discussed.

Prólogo

El control de la transcripción es escencial para la homoestasis de todas las células, para el desarrollo ontogenético de los organismos multicelulares y para la respuesta a las condiciones externas en donde se encuentre cualquier forma de vida. Brillantemente Francois Jacob & Jacques Monod (1961) vislumbraron la existencia de reguladores en las células bacterianas que actúan sobre sitios específicos en el DNA modulando la expresión de los genes. Este conjunto de conceptos conocidos como el modelo del operón, ha sido el paradigma principal para el estudio de la regulación de la expresión genética (Beckwith, 1996). Sabemos que aún cuando hay diferencias en la complejidad del aparato transcripcional de los linajes principales de la vida (Dominios archaea, bacteria y eucaria, Woese, 1990), el mecanismo básico es el mismo. En todos los casos se necesita de una RNA polimerasa y de factores asociados para generar una molécula de RNA mensajero que sirve de molde para la síntesis de una proteína (Capítulos 1 y 2).

Durante el inicio de la transcripción ocurren interacciones complejas entre el DNA y las proteínas que constituyen el aparato transcripcional. El resultado del reconocimiento preciso entre éstas moléculas es la producción de RNA mensajeros en forma específica y altamente regulada. El presente estudio fué concebido para entender como una clase de reguladores, los activadores, reconocen la RNA polimerasa y promueven la transcripción (Capítulos 3 y 4).

Específicamente estudiamos el mecanismo de activación transcripcional a distancia en bacterias. Esta forma de regulación positiva depende de proteínas que se unen al DNA en sitios alejados del promotor. La activación a distancia fué descubierta en virus animales (virus de simio, SV40) y es común en eucariotes (Serfling *et. al.*, 1985). En bacterias, la familia de proteinas activadoras a distancia (llamadas en inglés Enhancer Binding Proteins,

EBPs) regulan una gran variedad de funciones y se hallan interdispersas en diferentes géneros bacterianos (Morett & Segovia, 1993; Shingler, 1996). Una de estas funciones es la fijación de nitrógeno. La transcripción de los genes para la enzima que reduce el nitrógeno atmósferico a amonio, la nitrogenasa-reductasa, (*nifHDK*), así como otros genes *nif* y *fix*, son controlados positivamente por una EBP, la proteína NifA (Capítulo 4 y 5).

Al momento de concebir el trabajo experimental que originó esta tesis, existía evidencia que indicaba que todas las EBPs tienen un mecanismo de activación común (Kustu *et. al.*, 1989). Se conocía que estas proteínas tienen una actividad de ATPasa, así como aspectos particulares de proteínas como NtrC, cuya forma activa está fosforilada, y de NifA, que es regulado por oxígeno. También estaba demostrado que estas proteínas se unen al DNA en secuencias específicas localizadas lejos del promotor (Reitzer & Magasanik, 1986; Morett & Buck, 1988; Sasse-Dwight & Gralla, 1988). Experimentos precedentes habían demostrado que las EBPs utilizan una clase especial de RNA polimerasa holoenzima compuesta por el factor alternativo σ^{54} (Hunt & Magasanik, 1985; Hirshmann *et. al.*, 1985). Este mecanismo de acción de las EBPs implica el reconocimiento entre dos complejos de proteínas: el activador en forma oligomérica unido a un sitio lejos del promotor y la RNA pol-σ^{54} que forma el complejo cerrado con el promotor. Como resultado, la secuencia de DNA entre los sitios de unión de estos complejos forma un asa (Su *et. al.*, 1991; Capítulo 6). Desde nuestro punto de vista, identificar la región de NifA que contacta con la RNA polimerasa es uno de los problemas que plantea el mecanismo de activación transcripcional a distancia. Un estudio genético *in extenso* sobre una región conservada de NifA probablemente involucrada en el reconocimiento de la RNA pol-σ^{54}, fué propuesto para su publicación en Julio de1997 a la revista Molecular Microbiology. Después de su revisión y discusión con los árbitros fué aceptada en Diciembre de 1997 (Capítulo 7).

Esta tesis no contiene las secciones usuales, pero se trató de que los propósitos y argumentos que dieron origen al estudio experimental se manifestaran a todo lo largo del texto. Se hizó enfásis en proporcionar una interpretación estructural de los datos, a tráves de los ejemplos de proteínas que intervienen en la transcripción cuya estructura tridimensional se conoce. El Dr. Miguel Cevallos y el Dr. Mario Soberón, que participaron en el comité asesor de doctorado, contribuyeron mejorar la calidad del trabajo experimental. Fué muy valiosa la discusión y el trabajo en común con Humberto Flores y Katy Juárez. Aprecio la revisión y la crítica que sobre este trabajo realizaron el Dr. Fernando Bastarrachea., el Dr. Gabriel Guarneros, el Dr. Julio Collado, la Dra. Gloria Soberón y la Dra. Laura Camarena. El apoyo técnico y económico al proyecto es señalado en el artículo en prensa. Solamente quiero enfatizar que conté con el apoyo del CONACyT para la realización de estos estudios.

Víctor González.

Cuernavaca, Mor., 13 de Febrero de 1998.

"voy por tu cuerpo como por el mundo...

Capítulo 1

La RNA polimerasa básica

1.1. La estructura de la RNA polimerasa. Las enzimas que catalizan la formación de enlaces fosfodiéster 3´-5´ entre ribonucleótidos trifosfatados en presencia de un molde de DNA, se conocen genéricamente como RNA polimerasas. En bacteria, archaea y eucaria las RNA polimerasas están compuestas de múltiples subunidades, pero las de los bacteriófagos T7 y SP6 y la encontrada en las mitocondrias consisten de una sola unidad polipeptídica (Chamberlin *et. al.*, 1970). Se han encontrado algunas regiones de similitud en la estructura primaria entre las RNA polimerasas de una sola subunidad y aquellas de múltiples subuinidades, por lo que se especula que tienen un ancestro común (Allison *et. al.*, 1993; Iwabe *et. al.*, 1991). Las RNA polimerasas eucarióticas forman un sistema multicomponente de 10 a 15 subunidades, entre las que se encuentra la "TATA-binding protein " (TBP), homóloga a la encontrada en arqueobacterias y filogenéticamente distante de las bacterianas (Huet *et. al.*, 1983; Klenk & Doolittle, 1994).). La RNA polimerasa bacteriana está compuesta de 4 subunidades (Figura 1).

La RNA polimerasa de *Escherichia coli* esta codificada por los genes *rpoA* (α), *rpoB* (β') y *rpoC* (β), cuyos productos se asocian como $\alpha_2\beta\beta'$ en la forma conocida como RNA polimerasa "core" (RNA pol ^{C} , básica o central). La RNA pol ^{C} es capaz de iniciar transcripción de manera inespecífica a partir de los extremos terminales de un fragmento

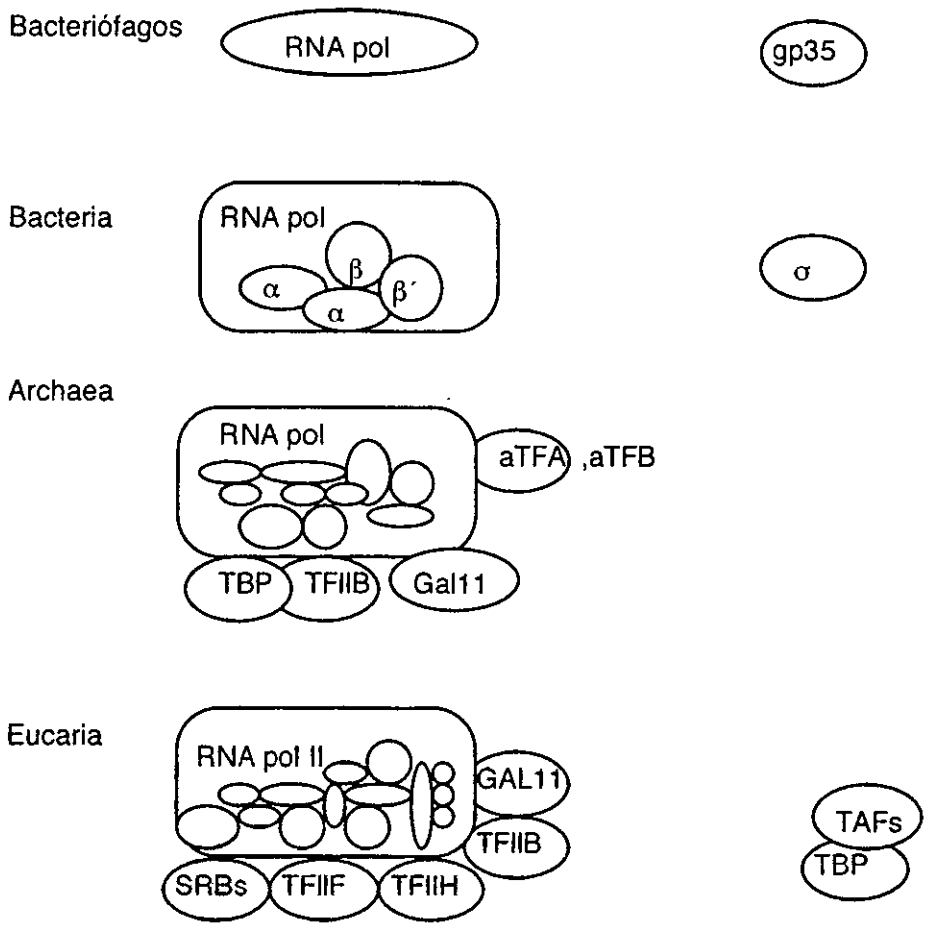

Figura 1. Organización en subunidades de los complejos transcripcionales básicos de los dominios de la vida.

de DNA, a partir de incisiones en una de las cadenas del DNA o a partir de aberturas en el DNA. La RNA pol ^{C} interviene en todos los pasos de la transcripción desde el inicio hasta la terminación, pero por sí misma es incapaz de iniciar la transcripción de manera específica. Richard Burguess *et. al.* (1969) descubrieron y caracterizaron una proteína adicional, que llamaron factor σ, que se asocia temporalmente con la RNA pol ^{C} para formar la RNA polimerasa holoenzima (RNA pol ^{H}). El factor σ confiere a la RNA pol ^{C} la capacidad de empezar la transcripción en sitios especificos del DNA, llamados promotores, que anteceden la región codificadora del gene. Travers & Burguess (1969) demostraron elegantemente que el factor σ se disocia de la RNA pol ^{C} inmediatamente

después de que el transcrito comienza a elongarse y se vuelve a asociar con otra RNApol ^{C} para comenzar un nuevo transcrito.

La RNA pol ^{C} es bastante estable a la disociación *in vivo* e *in vitro*. La ruta de ensamblado de la RNA pol ^{C} *in vitro* comienza con la dimerización de α, para formar α_2 enseguida se une la subunidad β y subsecuentemente β' (Zillig *et. al.*, 1976; Ishihama *et. al.*, 1981). No se conoce con detalle la estructura atómica de la RNA polimerasa bacteriana y solo se ha logrado la cristalización de un fragmento de digestión proteolítica del factor σ^{70} y la determinación por resonancia magnética nuclear (NMR) de la estructura del dominio carboxilo de la subunidad α (Jeon *et. al.*, 1995; Malhotra *et. al.*, 1996). Unicamente la RNA polimerasa del bacteriófago T7 ha sido cristalizada y su estructura ha sido determinada a una resolución de 3.3 Å (Sousa *et. al.*, 1993). La topología general de esta RNA polimerasa es parecida a un guante. Semejante forma también es parecida a la conformación del fragmento Klenow de la DNA polimerasa I y de la transcriptasa reversa del virus VIH (Ollis *et. al.*, 1985; Jacobo-Molina *et. al.*, 1993). Estas similitudes estructurales y la conservación de los residuos catalíticos son sugerentes de un origen común para todas las polimerasas RNA o DNA dependientes (Moras, 1993).

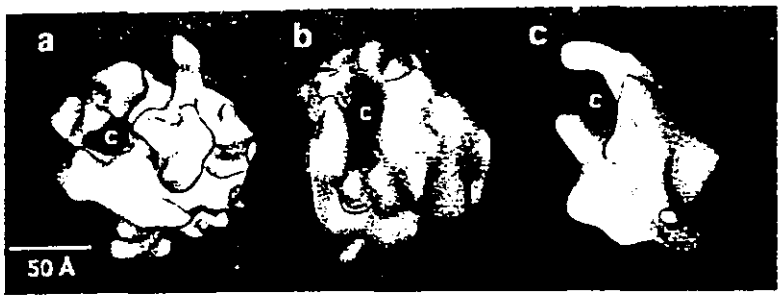

Figura 2. Comparación de la estructura de la RNA polimerasa II de levadura (a), RNA pol ^{C} (b) y RNA pol ^{H} (c) de E. coli (Polyakov et. al., 1995)

Información estructural de baja resolución (23-27 Å) de distintas RNA polimerasas ha sido obtenida mediante imágenes de microscopía electrónica (ME) sobre cristales bidimensionales (Kornberg & Darst, 1989). Entre estas se encuentran las RNA polimerasa I y II de *Sacharomyces cerevisae* (Darst *et. al.*, 1991; Shultz *et. al.*, 1993) y la RNA polimerasa de *E. coli* en sus formas básica y holoenzima (Darst *et. al.*, 1989; Polyakov *et. al.*, 1995). En todos estos casos la estructura resulta similar a la topología de "guante" determinada para la RNA polimerasa del bacteriófago T7. La RNA pol ^{H} de *E. coli* aparece como un objeto irregular de 90 x 95 x 165 Å (Darst *et. al.*, 1989), con un canal de aproximadamente 25 Å de diámetro por 55 Å de largo circundado por dos proyecciones alargadas (Figura 2). En este canal se especula que podría ser acomodada una molécula de DNA de doble cadena de 16 pares de bases. En la RNA pol ^{C} de *E coli* se encontró que el canal está completamente cerrado formando un túnel (Polyakov *et. al.*, 1995), de manera semejante a la RNA pol II de levadura (Darst *et. al.*, 1991). Estas observaciones

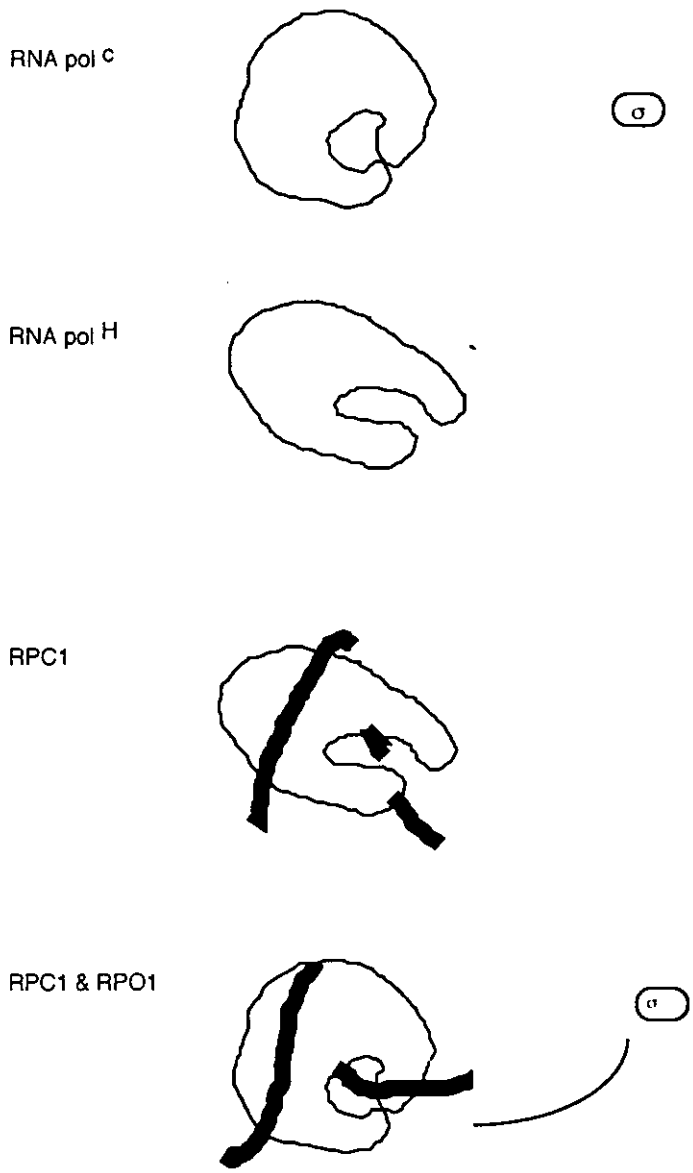

Figura 3. Cambios estructurales en la RNA polimerasa de E. coli durante el inicio de la transcripción.

condujeron a los autores a proponer que estas dos conformaciones de la RNA polimerasa corresponden a dos estados conformacionales y funcionales diferentes. La RNA polimerasa con el canal abierto representaría el estado de inicio de la transcripción, mientras que la forma cerrada, como túnel, representa el estado de apertura y elongación del transcrito (Figura 3; Polyakov *et. al.*, 1995). En ninguna de estas conformaciones es posible asignar una localización específica para cada subunidad de la RNA polimerasa. Sin embargo, tomando en cuenta el modelo atómico de mayor resolución de la RNA

polimerasa del bacteriófago T7, es posible asignar el sitio activo de la enzima en el lugar equivalente a la palma del guante, mientras que los dedos y el pulgar contendrían los determinantes estructurales para la unión del templado y los sustratos respectivamente (Moras, 1993).

1.2. Las subunidades β y β ´. Son conocidas como las subunidades mayores de la RNA polimerasa de *E. coli*, con 1342 (β) y 1407 (β´) aminoácidos, y se cree que son resultado de una duplicación ancestral (Onishi, 1985; Armaleo, 1987). Son muy parecidas a las subunidades mayores de la RNA polimerasa de archaea y eucarya, de tal manera que es posible que hayan estado presentes en el progenote (Allison *et. al.* 1993). Esta dos subunidades probablemente conformen gran parte de la estructura en forma de guante de las distintas polimerasas, obtenidas con ME. Recientes experimentos sugieren que el sitio activo de la RNA polimerasa está formado por ambas subunidades, en el lugar que se conoce como el canal (Polyakov *et. al.,* 1994). En ambas subunidades se han encontrado, utilizando puentes con agentes químicos fotoactivables, aminoácidos que unen la cadena de RNA en crecimiento (Borukovh *et. al.*, 1991; Khaslev *et. al.*, 1990; Mustaev *et. al.*, 1991). La subunidad β está involucrada en la unión con el DNA y contiene los determinantes funcionales para unir el primer ribonucleótido durante iniciación de la polimerización (Jin & Turnbough, 1994). Existen mutantes en la subunidad β que son incapaces de unir eficientemente los sustratos y abren el promotor de manera deficiente, abortando constantemente los transcritos (Weilbaecher *et. al.*, 1994). Este comportamiento es parecido al obtenido con el antibiótico rifampicina, letal para la mayoría de la bacterias. La rifampicina inhibe la iniciación de la transcripción interaccionando con la subunidad β. Esta inhibición es absolutamente dependiente de que la subunidad β este ensamblada con el resto de la polimerasa ya que la rifampicina no se une a la subunidad β no asociada (Jin & Gross, 1988). El antibiótico amanitina tiene un efecto similar a la rifampicina sobre la RNA polimerasas eucarióticas, lo cual sugiere una

similitud funcional entre ambas clases de polimerasas (Sentenac, 1985). Recientemente se localizó una región conservada de β′ en donde probablemente se une el ion magnesio (Zaychikov *et. al.* 1996). Dado que el polianion heparina compite *in vitro* por la unión de β′ con el templado parcialmente desnaturalizado, se ha especulado que la β′ tiene *in vivo*, el papel de unir el molde de DNA (Chamberlin, 1976).

1.3. La subunidad α. Esta subunidad de 332 aminoácidos tiene tres funciones conocidas: ensamblar la RNA polimerasa, participar en el reconocimiento del promotor y servir de blanco de regulación transcripcional (Ebright & Busby, 1994). Ensayos de proteólisis limitada muestran que α es una proteína con dos dominios, unidos por un conector flexible de composición variable (Blatter *et. al.*, 1994). El dominio amino terminal de la subunidad α (αNTD) contiene los determinantes para el ensamblado de la RNA polimerasa puesto que fragmentos de α que conservan este dominio son capaces de dimerizar y de interaccionar con las otras dos subunidades de la RNA polimerasa (Kimura *et. al.*, 1994; Kimura *et.al.*, 1995; Negishi *et. al.*, 1995). Estas RNA polimerasas son capaces de llevar a cabo transcripción basal pero no son activadas por proteínas reguladoras.

La subunidad α juega un papel importante en el reconocimiento del promotor interaccionando específicamente con secuencias que se localizan hacia arriba de la región -35 (Ross *et. al.*, 1993; Gaal *et. al.*, 1996). Esta región, llamada elemento UP (por "upstream"), se encuentra solo en promotores de genes cuyos productos se necesitan en gran cantidad, como por ejemplo los rRNA, codificados por *rrnb1*. Se ha demostrado mediante experimentos de protección a la DNAasa I, que el dominio carboxilo terminal de α (αCTD) ocupa el elemento UP (Ross *et. al.*, 1993).

La tercera función de la subunidad α es constituir el blanco de interacción para activadores transcripcionales. Distintos residuos localizados tanto en el dominio amino como en el carboxilo de esta subunidad han sido involucrados en la interaccion específica con los llamados activadores de clase I (Capítulo 3; Ebright, 1993; Ishihama, 1993). Recientemente se determinó con NMR la estructura en solución de αCTD (Jeon *et. al.*, 1995). αCTD forma una estructura compacta compuesta por 4 hélices y dos asas grandes, en la cual los determinantes para la unión al DNA y la activación parecen estar sobrelapados. Específicamente la hélice 1 ha sido involucrada en la activación y en el contacto con el DNA (residuos 260-270).

Es interesante señalar que la subunidad α también resulta ser el blanco de modificaciones postraduccionales, como la ADP-ribosilación durante la infección de *E. coli* con el bacteriófago T4 (Goff, 1984). Como consecuencia las RNA polimerasas ribosiladas no responden a la acción de la proteína activadora CAP y no promueven la transcripción de promotores con el elemento UP. Presuntamente esto incrementaría el número de RNA polimerasas disponibles para transcribir los genes virales (Ebright & Busby, 1994).

"voy por tus ojos como por el agua ...

Capítulo 2

Estructura y función del promotor.

Para iniciar transcripción la RNA pol ^{H} reconoce y se une con secuencias de DNA específicas, llamadas promotores, que anteceden una región de DNA codificante. Aunque la RNA pol ^{H} cubre de 40 a 50 pares de bases (pb) del promotor, es el factor σ la subunidad que contacta específicamente con el promotor. Se conoce la secuencia nucleotídica de 300 promotores de *E. coli*, *S. thyphimurium* y colifagos así como diversas clases de factores σ (Helmann & Chamberlin, 1986; Gross *et. al.*, 1992). En el genoma de *E. coli* secuenciado base por base en su totalidad, se han estimado 2405 regiones operadoras de las cuales 68% tendrian un promotor, 20% dos promotores y 12 % presentarían tres o más promotores (Blattner *et. al.*, 1997). Cada clase de factor σ funciona concertadamente con un tipo particular de promotor (Helmann & Chamberlin, 1988; Lonneto *et. al.* 1992).

En bacterias los factores σ constituyen un mecanismo mediante el cual diversos conjuntos de genes se expresan diferencialmente de acuerdo a las condiciones de crecimiento. En *Bacillus*, por ejemplo, hay mecanismos de regulación globales basados en la sustitución temporal de factores σ en una RNA polimerasa común, lo que resulta en un control de las fases vegetativa y de crecimiento sumamente eficiente (Losick & Pero, 1981). En *B. subtilis* hay 11 factores σ, mientras que en *E. coli* existen 6 factores σ con especificidades diferentes (Lonneto & Gross, 1996; Kuntz *et. al.*, 1997). De ellos σ^{70} es el responsable de la expresión de la mayoría de los genes del metabolismo durante la fase exponencial de crecimiento. Los otros factores son alternativos y se requieren para la expresión de genes de la fase estacionaria (σ^{38}), choque térmico (σ^{32} , σ^{24}) producción del flagelo (σ^{28}) y metabolismo nitrógenado (σ^{54}). En el resto de las especies bacterianas se han encontrado homólogos de cada uno de estos factores σ (Lonneto *et. al.*, 1992)

2.1. El reconocimiento RNA polimerasa-promotor. La RNA polimerasa localiza el promotor entre un exceso de secuencias no específicas. Un modelo simple es considerar la difusión directa de la RNA polimerasa desde el citoplasma hasta el promotor, pero las tasa de formación del complejo cerrado mediante la RNA polimerasa de bacteriófagos, es más rápida de lo que se esperaría por simple difusión (10^{-8} M^{-1} sec^{-1} *versus* 10^{-10} $M^{-1}sec^{-1}$). Un modelo de difusión facilitada se ajusta mejor a estas tasas, considerando que la RNA polimerasa inicialmente se une al DNA de manera inespecífica para luego desplazarse a lo largo de la doble hélice (difusión unidimensional), hasta encontrar el promotor (Von Hippel *et. al.*, 1984). La interacción inespecífica estaría dominada exclusivamente por fuerzas electrostáticas, y la específica por una matriz de puentes de hidrógeno orientada apropiadamente para resultar en una complementariedad entre el promotor y la RNA polimerasa. Alternativamente se puede considerar que la RNA polimerasa reconoce la topografía global del promotor, es decir, la amplitud de los surcos mayor y menor del DNA y/o la curvatura de la secuencia, para posteriormente interaccionar específicamente (Von Hippel *et. al.*, 1984).

Se ha pensado que la estructura de los promotores (dos regiones conservadas y un espaciador de secuencia variables), responde a los requerimentos para reconocer el promotor mediante la RNA polimerasa. Una hipótesis sencilla supone que la región distal de los promotores (-35 y -24, para σ^{70} y σ^{54} respectivamente) afecta la unión inicial de la RNA polimerasa y la región proximal (-10 y -12) funciona para la formación del complejo abierto (Gilbert *et. al.*, 1986). Sin embargo, los resultados de la alteración de cada una de estas regiones, no parecen ajustarse a éste modelo, ya que mutaciones en -35 del promotor *lacUV* también afectan la capacidad para producir transcritos y mutaciones en -10 parecen afectar la unión de la RNA pol ^{H} (Kobayashi *et. al.* 1990). Esto sugiere que las interacciones promotor-RNA pol ^{H}, es un proceso de varios pasos en donde participan ambas regiones del promotor.

2.2. La estructura primaria de los promotores. En el promotor de la mayoría de los genes de las bacterias se localizan dos grupos de hexanucleótidos altamente conservados situados a -10 (TATAAT, también llamada "Pribnow box") y -35 (TTGACA) del sitio de inicio de la transcripción (Figura 4 a). Generalmente estas dos regiones se encuentran separadas por una secuencia de alrededor de 17 pares de bases de composición variable. Diversos criterios genéticos y estadísticos han llevado a considerar que esta estructura representa el promotor ideal o consenso para la RNA pol-σ^{70} de *E. coli* (Hawley & McClure, 1983; Harley & Davidson, 1981; Lisser & Margalit, 1993). La hipótesis subyacente a este concepto está involucrada con la fuerza del promotor, es decir, que la capacidad para generar más transcritos está directamente relacionada con la similitud del promotor natural con respecto al consenso. De esta manera, aunque suelen ocurrir excepciones, mutaciones que incrementan la similitud del promotor con el consenso, también incrementan su fuerza, y viceversa (Reznikoff *et. al.*, 1985; Brunner & Bujard, 1987). La válidez del consenso ha sido apoyada por experimentos de diseño de promotores artificiales de acuerdo al consenso, los cuales resultan funcionales *in vivo* e *in vitro* (Rossi *et. al.*, 1983). Adicionalmente, experimentos de selección de promotores a partir de un conjunto de variantes producidas azar, han demostrado que los promotores obtenidos tienen una gran similitud a la estructura del promotor consenso (Oliphant & Struhl, 1988).

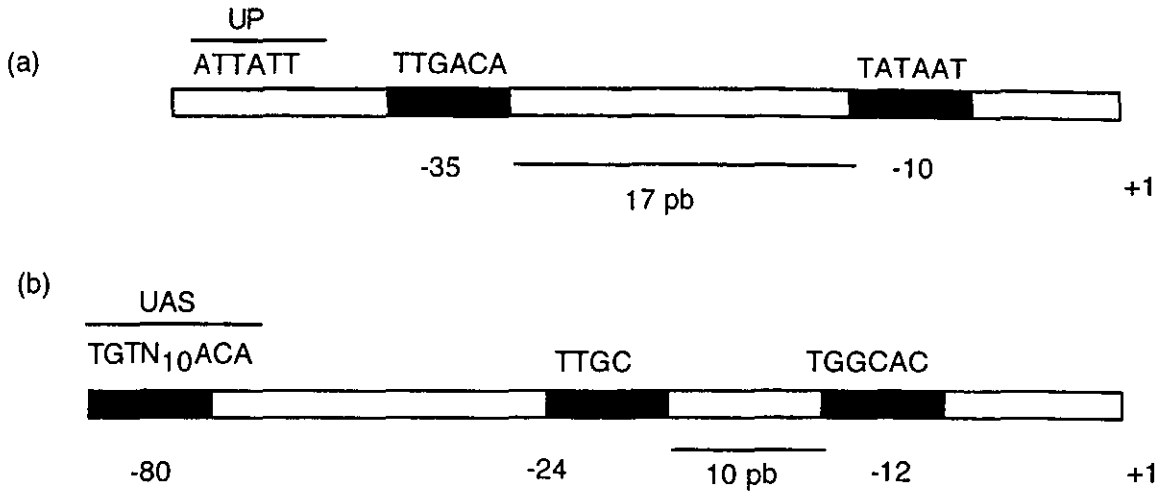

Figura 4. Estructura de los promotores σ^{70} (a) y σ^{54} (b)

Como se mencionó anteriormente, recientemente se ha encontrado que los promotores de algunos genes tienen una tercera región conservada justo hacia arriba de la región -35, el elemento UP, constituida por nucleótidos de adenina y timina (Ross *et. al.*, 1993). El reemplazamiento del elemento UP del gene para el RNA ribosomal (*rrnBP1*), con secuencias al azar reduce la capacidad de transcripción del promotor (Leirmo & Gourse, 1991). Por el contrario, la introducción del elemento UP en los promotores de algunos genes incrementa la tasa de transcripción (Newlands *et. al.*, 1991). El elemento UP participa en el reconocimiento de la subunidad α de la RNA pol ^{H} que contribuye a la eficiencia de iniciación de la transcripción (Ross *et. al.*, 1994). El elemento UP interacciona con la subunidad α de la RNA pol ^{H} tal vez estabilizando el complejo cerrado inicial (Rao *et. al.*, 1994). Es interesante señalar que ésta propuesta es análoga a la explicación ofrecida para el mecanismo de acción de una clase activadores, como CAP, que actúan sobre genes cuyos promotores tienen regiones -35 alejadas del consenso. Dado que estos activadores interaccionan a tráves de la subunidad α, probablemente contribuirían a estabilizar la RNA pol ^{H} en el promotor (Ishihama, 1993; Busby & Ebright, 1994).

En los promotores dependientes de σ^{54} la especificidad está determinada por las secuencias conservadas -12 (T**GG**CAC) y -24 (TT**GA**) separadas por 10 pares de bases (Figura 4 b). Los dinucleótidos GG y GA en estas dos posiciones son prácticamente invariantes en más de 120 promotores de esta clase (H. Barrios & E. Morett, sin publicar). Cualquier sustitución de estos dinucleótidos en el promotor *glnAp2* produce fenotipos deficientes en la transcripción (Merrick & Chambers, 1992). En estos promotores hay una estricta dependencia en el número de bases que tiene el espaciador. Cambios de una sola base en el espaciador inactivan totalmente el promotor (Buck *et. al.*, 1986; Buck *et. al.*, 1987). Esto indica que es necesario mantener ambas regiones conservadas en la misma cara del DNA. Esto no sucede con el espaciador entre las regiones -10 y -35 de los promotores reconocidos por la familia de factores σ^{70}, en donde alteraciones 2 o 3 pares de bases no afectan su funcionamiento (Warne & De Hasset, 1993).

2.3. Los factores σ. En las bacterias hay una amplia diversidad de factores σ agrupados en dos familias fisiológica y filogenéticamente no relacionadas (Lonneto *et. al.*, 1992; Merrick, 1993). La familia de factores emparentados con σ^{70} de *E. coli* es la más amplia y diversificada de las dos. Como ya citamos, en esta familia se encuentran todos aquellos factores que tienen el papel de transcribir genes indispensables para el crecimiento de la bacteria (genes básicos o "housekeeping genes"), así como de algunos de los factores σ alternativos que son requeridos para la expresión de genes solo en condiciones especiales de crecimiento como alta o baja temperatura, esporulación y motilidad (Helmann & Chamberlin, 1988; Losick & Pero, 1981).

La segunda familia está formada por un conjunto homógeneo de factores, estrechamente emparentados con el factor σ^{54} de *E. coli* (Rpo^N, σ^N), y otras proteobacterias y firmibacterias (Merrick, 1992; Shingler, 1996). Como se esperaría de la baja similitud en la

estructura primaria entre estas dos familias, también existen diferencias profundas en la función. Sin embargo, podemos anticipar al menos dos razgos funcionales comunes entre ambas familias: le brindan especificidad a la RNA polimerasa a tráves de reconocer una secuencia específica en el DNA, e inician la transcripción usando una RNA polimerasa común. Un estudio detallado de ejemplos representativos de factores sigma de las dos familias a llevado a entender mejor las bases moleculares de sus diferencias.

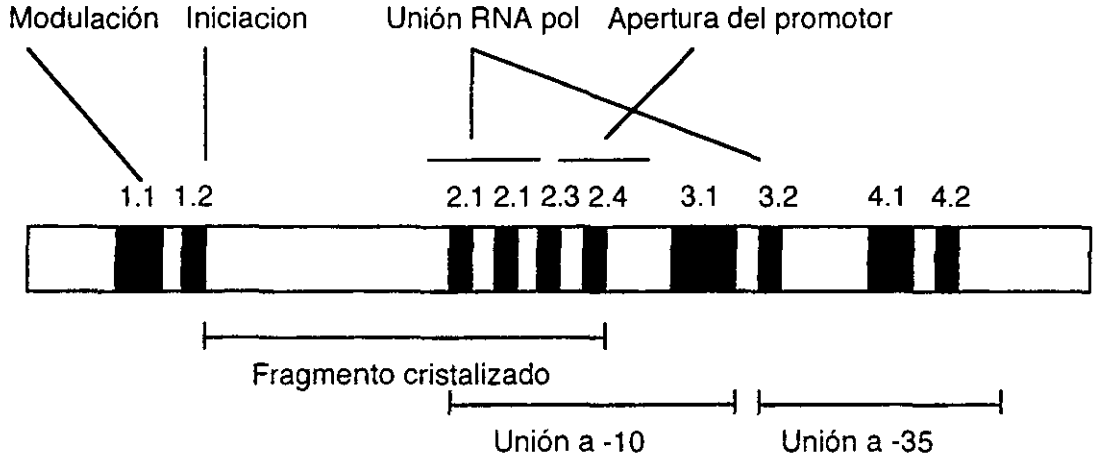

Figura 5. Regiones conservadas de σ^{70}.

2.3.1. La familia σ^{70}. Mediante alineamientos múltiples de las secuencias de aminoácidos de miembros de esta familia se han detectado cuatro regiones más conservadas que el resto de la proteína (Figura 5; Helmann & Chamberlin, 1988; Loneto *et. al.*, 1992). La determinación de la estructura cristalográfica de un fragmento de digestión tríptica de σ^{70} ha conducido a una revaloración del papel de cada una de estas regiones conservadas (Malhotra *et.al.*, 1996; Chan *et. al.*, 1996). El fragmento de σ^{70} que se logró cristalizar comprende parte de la región 1.2 hasta la región 2.3 de la proteína completa (amino ácidos114-446). La estructura de este dominio consiste primordialmente de α hélices conectadas con asas, semejando una V. En general se puede observar que varias de las

regiones conservadas se encuentran estrechamente asociadas en la estructura terciaria , formando un dominio mediante el cual σ se une con la RNA pol ^{C}, mientras que en el lado opuesto de la estructura se localizan los determinantes de unión con el DNA. En la estructura se puede observar un conector entre la region 1.1 y 2.1, de secuencia y longitud variable en los distintos factores σ. Este conector llega a tener 235 aminoácidos en el factor σ de *Pseudomonas* y es prácticamente inexistente en σ^{43} de *Bacillus subtilis*, donde las regiones 1.1 y 2.1 son casi contiguas. Estructuralmente el conector está formado por varias α hélices que sin embargo no tienen relación estructural con otras partes de σ que participan directamente en la función, justificando de alguna manera su ausencia en algunas de los factores σ conocidos (Lonneto *et. al.*, 1992).

2.3.1.1. Las regiones conservadas de σ^{70}

Región 1. Esta región localizada en el amino terminal de la proteína se encuentra sólo en los llamados factores σ primarios, es decir, en aquellos con la más alta similitud en la secuencia de aminoácidos con σ^{70} (Lonnneto *et. al.*, 1992; Hellmann & Chamberlain, 1988). Hay indicios de que el papel de esta región es modular la unión de la RNA pol ^{C} con el DNA (Dombroski *et. al.*, 1993). Fragmentos proteolitícos de esta región conservada inhiben *in vitro* la unión de la RNA pol ^{C} con el DNA, tal vez interaccionando con la región conservada 4 de σ^{70}, implicada en el reconocimiento de la región -35 del promotor (Dombroski *et. al.*, 1992). Experimentos recientes señalan que la región 1.1 podría estar involucrada en la iniciación de la transcripción, puesto que la remoción de este fragmento no permite que la isomerización del complejo cerrado al complejo abierto ocurra eficientemente (Wilson & Dombroski, 1997). Mutantes de σ^{70} que carecen de la región 1.2, forman holoenzimas incapaces de abrir el complejo cerrado.

Región 2. Esta región es la que presenta mayor similitud entre los distintos factores σ y consiste de cuatro subregiones. La remoción de la región 2.1 de σ^{70} y σ^{32} reduce la afinidad

de σ por la RNA pol ^{C} (Lesley et. al., 1991; Lesley & Burguess, 1989). Una mutación puntual en la región 2.1 de σ ^{E} de *Bacillus subtilis* (un cambio en un residuo expuesto arg 385 a tyr), también exhibe una baja afinidad por la RNApol ^{C} (Shuler *et. al.*, 1995). En la estructura de σ^{70} , la región 2.1 forma dos α-helices separadas por una inflexión. Esta hélice tiene relación estrecha con otras α hélices hacia al interior de la proteína y tiene una región de aminoácidos hidrofóbicos expuestos al solvente, donde se localizan algunos de los residuos implicados en la unión con la RNA pol ^{C} (Lesley & Burguess, 1989). La región 2.2 no ha sido implicada en ninguna de las funciones de σ a pesar de su extremada conservación. Estructuralmente la región 2.2 forma una hélice α atrapada entre otras dos. Un reporte reciente indica que una substitución conservativa en la región 2.2 de σ^{32} reduce la afinidad por la RNA pol ^{C} (Joo *et. al.* 1997).

Las regiones 2.3 y 2.4 están involucradas en el reconocimiento y la desnaturalización del promotor. Una prueba contundente de la relación de la región 2.4 con el reconocimiento de la región -10 del promotor es que mutaciones en el promotor son suprimidas con mutaciones en el factor σ (Kennedy *et. al.*, 1989; Siegele *et. al.*, 1989; Daniels *et. al.*, 1990; Waldburger *et. al.*, 1990; Tati *et. al.*, 1991). Esta región forma una hélice anfipática, cuyos residuos hidrofóbicos se encuentran hacia el interior de la proteína y los residuos que entran en contacto con la región -10, están expuestos al solvente (Malhotra *et. al.*, 1996). Helmann & Chamberlin (1988) definieron a los fatores σ como "proteinas desnaturalizadoras de DNA" y asignaron a la region 2.3 esta función debido a que contiene varios aminoácidos aromáticos que por analogía con lo que ocurre en las proteínas que unen RNA de cadena sencilla (Shamoo, 1995), pudieran servir para conferir estabilidad al DNA monocatenario, intercalándose entre las bases. En apoyo a ésta hipótesis, cuando cada uno de los residuos aromáticos de σ^{A} de *B. subtillis* fue sustituido por alanina, se observaron severas fallas en la apertura del promotor (Juang & Hellmann, 1994). De acuerdo con la estructura del fragmento σ^{70} estos aminoácidos aromáticos están

en una posición adecuada para interaccionar con la región del promotor alrededor de -10. Mutagénesis de los residuos aromáticos de σ^{70} apoyan esta hipótesis (Aiyar *et. al.*, 1994; Waldburger & Susskind, 1994). En este contexto, dos aminoácidos cargados positivamente cercanos a esta región auxiliarían a neutralizar la carga del DNA, interaccionando con los fosfatos. Experimentos de puenteo químico demostraron que σ se une a la cadena complementaria del templado durante la apertura del promotor (Buckle & Buc, 1989). Además, se ha observado que el fragmento de σ^{70} que fué cristalizado se une especificamente a un oligonucleótido que lleva la secuencia alrededor de -10 (Dombroski *et. al.*, 1992) Esto datos son congruentes con un modelo en el cual la región 2.3 y 2.4 actúan de manera concertada para abrir el promotor y estabilizar la hebra complementaria al templado interactuando directamente con ésta (Chan *et. al.*, 1996; Hellman & Chamberlin, 1988).

Región 3. Esta región consiste de dos subregiones. La 3.1 presuntamente funciona en el reconocimiento del promotor mediante una HTH, mientras que la 3.2, formada por residuos ácidos, probablemente esté involucrada en la unión con la RNA pol ^{C}. De hecho una mutante de σ^{32} que carece de esta región tiene una afinidad reducida para unirse a la RNApol ^{C}, dejando intactas las otras funciones de la proteína (Zhou *et. al.*, 1992). Algunos factores σ no aparentan tener esta región (como los factores σ de *B. subtilis*), por lo que se piensa que en esta clase de factores σ, esta superficie de interacción con la RNA pol ^{C} ha sido sustituida evolutivamenente.

Región 4. La región 4 participa en el reconocimiento de la región -35 del promotor y consiste de dos subregiones separadas por un conector corto. La región 4.1 se predice que forma una hélice anfipática y la región 4.2 tiene un alto parecido con secuencias que forman una hélice-vuelta-hélice. Dos residuos localizados en la hélice de reconocimiento han sido implicados en la unión con la región -35 el promotor (Kenney & Moran, 1991;

Siegele *et. al.*, 1989). Es interesante mencionar que cambios puntuales en la región 4.2 permiten la interconversión de las propiedades de reconocimiento del promotor de σ^B, σ^F, y σ^G (Margolis & Losick, 1991; Schmidt *et. al.* 1990).

2.3.2. La familia σ^{54}. El primer indicio de la existencia de este tipo de factores provino del estudio de la regulación de los genes del metabolismo nitrógenado (*ntr*) en *Salmonella thypimurium* (García *et. al.*, 1977). Posteriormente se logró la purificación y caracterización del producto del gene rpo^N (también llamado *glnF* o *NtrA*) tanto de *E. coli* como de *S. thypimurium* (Hunt & Magasanik, 1985; Hirshman *et. al.*, 1985). A la fecha, homólogos de *rpoN* se han encontrado en un buen número de especies bacterianas en una sola copia y, excepcionalmente en dos copias como en el genoma de *Bradyrhizobium japonicum* (Kullik *et. al.*, 1989).

Debido a que su presencia responde a funciones metabólicas específicas, no comprometidas con la viabilidad celular, σ^{54} es un factor de iniciación alternativo. Sin embargo, recientemente se reportó que la expresión de σ^{54} en *Myxococcus xanthus* resulta absolutamente indispensable para la viabilidad celular en varias condiciones de crecimiento. Esfuerzos por eliminar el gene ("knockout") de *M. xanthus*, mediante la remoción del gene o la inserción de elementos genéticos en *rpoN* resultaron inútiles, puesto que no se obtuvieron células viables, por lo tanto σ^{54} parece indispensable para el crecimiento vegetativo de esta bacteria (Keesleer & Kaiser, 1997).

Dos características fundamentales diferencian a los factores σ^{70} de los del tipo σ^{54}, una es que estos últimos en ausencia de la RNApol ^{C} se unen específicamente a ciertos promotores (Cannon & Buck, 1992), y la otra es que prácticamente no existe transcripción a partir de los promotores-σ^{54} en ausencia de una proteína activadora. Eσ^{54} forma complejos sumamente estables con el promotor incapaces de isomerizar hacia un

complejo abierto (Buck & Cannon, 1992), por lo que se ha especulado que los factores σ^{54} son deficientes en la desnaturalización del promotor, como contrapartida a la hipótesis de Helmann & Chamberlin que sugieren que los factores σ^{70} son proteínas que unen y desnaturalizan DNA (1988). Sin embargo, hallazgos recientes de mutantes de σ^{54} indican que más que una defiencia, este factor de iniciación está en una conformación que inhibe el inicio de la transcripción (Wang *et. al.* 1996).

Sasse-Dwigth & Gralla (1990), sugirieron que σ^{54} tiene una organización en dominios con firmas estructurales que recuerdan aquellas encontradas en los activadores transcripcionales eucarióticos. A partir de estudios genéticos y bioquímicos se han precisado tres regiones de σ^{54}. La región I ubicada en el amino terminal de la proteína es rica en leucina y glutamina y está probablemente involucrada en la activación de la transcripción. La región II, generalmente no conservada en secuencia y longitud pero que tiene comúnmente residuos ácidos. Esta región podría participar en la desnaturalización del promotor. Finalmente, la región III altamente conservada cuya parte amino sirve para la unión con la RNApol ^{C} y la parte carboxilo para la unión con el DNA (Figura 6).

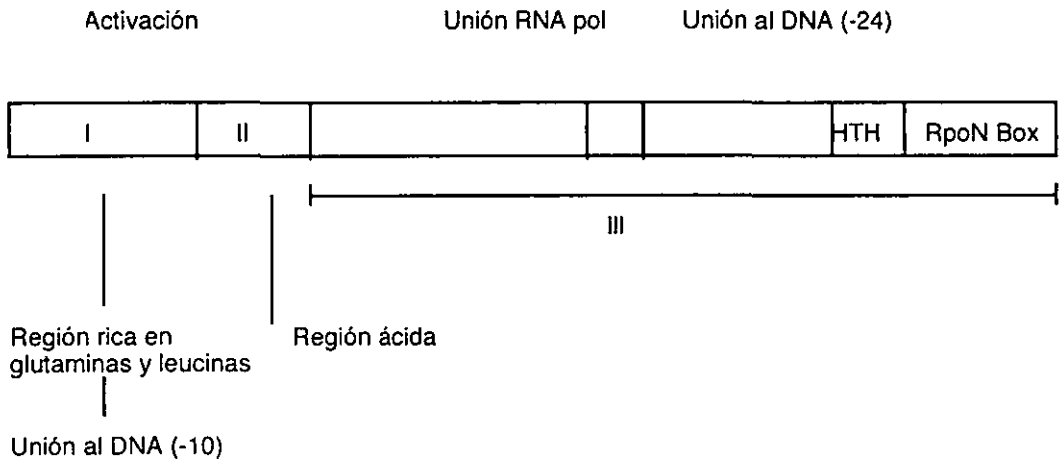

Figura 6. Estructura primaria de σ^{54}

2.3.2.1. Regiones I y II: la formación del complejo abierto. La llamada región de activación (región I) de σ^{54} está formada por heptadas de leucinas con glutaminas interdigitadas, involucradas en la modulación de la unión con la región -12 y en la apertura del promotor (Sasse-Dwight & Gralla, 1992; Hsieh *et. al.*, 1994). Se ha predicho que la estructura secundaria de estas heptadas pudiera corresponder a una hélice anfipática y se ha comprobado que situando por mutagénesis dirigida una prolina en las heptadas de leucina, la activación transcripcional es severamente afectada (Hsieh & Gralla, 1994). La remoción de 13 amino acidos del amino terminal de σ^{54} de *E. coli*, perturba tanto la actividad transcripcional como la unión con la región -12 del promotor, sin afectar la unión con la región -24 (Sasse-Dwight & Gralla, 1991). Debido a que la región de activación de algunos reguladores de eucariotes está compuesta por residuos de glutamina, se piensa que en σ^{54} una región similar podría tener la misma función (Sasse-Dwight & Gralla, 1991). La alteración de los residuos de glutamina ocasiona una pérdida progresiva, de acuerdo al número de residuos mutados, de la capacidad de iniciar transcripción (Hsieh *et. al.*, 1992). Una mutante con las cuatro glutaminas centrales sustituidas por serina o glutamato, está afectada en la formación del complejo abierto, pero su perfil de unión con el promotor permanece similar a la proteína silvestre. Mutantes múltiples en las glutaminas localizadas hacia el carboxilo ocasionan un aumento en la capacidad para abrir el promotor, pero esto no se traduce en una cantidad mayor de mRNA (Hsieh *et. al.*, 1993). Esto se ha interpretado como una falla en la estabilización del complejo abierto.

La sustitución progresiva de los residuos de leucina de esta region, produce una pérdida progresiva de la función de σ^{54}, pero la capacidad de unirse a la región -24 del promotor es mantenida (Hsieh & Gralla, 1994). Es sumamente llamativo que estos cambios, al igual que aquellos de las mutantes en la región rica en glutaminas no ocasionan la pérdida en la formación del complejo abierto. Esto significa que la contribución del activador se da en un paso posterior a la apertura del complejo cerrado, probablemente estabilizándolo. Muy

relacionada con esta idea, se encuentra el hallazgo de mutantes de σ^{54} que sobrepasan la necesidad de un activador alcanzando hasta un 15% de la transcripción obtenida usando σ^{54} y el activador NtrC (Wang *et. al.*, 1995). Estas variantes de σ^{54} llevan mutaciones sencillas o múltiples que afectan las leucinas entre las posiciones 25 y 30 (LLQLSTL) (Syed & Gralla. 1997). Para explicar estas observaciones, Jay Gralla, propone que esta región es un elemento para la inhibición de la apertura del promotor. Cuando esta región de σ^{54} es alterada por mutaciones, la RNA polimerasa holoenzima constituida es capaz de iniciar transcripción en ausencia de la proteína activadora (Wang & Gralla, 1996). La función asignada a la proteína activadora en este modelo es la de producir un cambio estructural de tal manera que la inhibición de la desnaturalización del promotor sea eliminada. En capítulos posteriores se considerará el papel de esta region como el blanco de las interacciones con el activador integrado a la concepción actual del mecanismo de activación a distancia.

La región II usualmente tiene entre 60 y 100 aminoácidos de longitud, entre los que se encuentran residuos de glutamato y aspartato cada tercer residuo, aparentemente en triadas repetidas. Esta región probablemente participa en los eventos que conducen a la desnaturalización del promotor (Sasse-Dwight & Gralla, et. al., 1990). La remoción de esta región de σ^{54} de *E. coli* reduce la tasa de formación del complejo abierto y una duplicación de una parte de esta región incrementa la tasa de formación del complejo abierto (Wong & Gralla, 1992). Se ha propuesto que esta región constituya un asa debido a que la remoción de parte de esta región o las inserciones, además de la variabilidad que tiene en distintos factores σ de la familia, no afectan severamente la apertura del promotor (Wong & Gralla, 1992; Merrick, 1993).

2.3.2.2. La región III:

La unión con la RNA polimerasa ^{c}. Debido a que σ^{70} y σ^{54} forman holoenzimas con una RNA polimerasa común se ha supuesto que los determinantes estructurales de esta asociación podrían estar conservados. Sin embargo no existe evidencia de alguna similitud en la estructura primaria de estas proteínas. De acuerdo con Cannon & Buck (1997), un fragmento de digestión proteolítica que abarca los aminoácidos 107-303 de σ^{54} de *K. pneumoniae* es suficiente para la unión con la RNA pol ^{C}. La remoción de una parte de esta región resulta en formas de σ^{54} incapaces de unirse con la RNA pol ^{C} pero aún capaces de unirse al DNA (Wang *et. al.* 1994). Mediante mutagénesis dirigida y mutagénesis con PCR se han aislado variantes de σ^{54} que pierden la capacidad de unirse con la RNA pol ^{C}. Estas mutaciones se localizan entre los aminoácidos 154 a 193. A la luz de la estructura tridimensional σ^{70}, se ha propuesto que esta región representa el equivalente estructural y funcional en σ^{54} para la unión con la RNA pol ^{C}. La region 2.1 de σ^{70} está constituida por residuos hidrofóbicos altamente conservados. En el modelo atómico de σ^{70} ésta región adopta una estructura secundaria conocida como "coiled-coil", por lo que se predice que en σ^{54} pudiera adoptar una estructura similar, dado que ambos factores de iniciación utilizan una RNA polimerasa común. A partir de un detallado mapeo de las actividades de unión a la RNA pol ^{C} y al DNA, Cannon y colegas (1997), lograron identificar un fragmento que va desde los aminoácidos 70 hasta el 184 todavía capaz de unirse a la RNA pol ^{C} pero que ya había perdido la capacidad de unirse al DNA.

La unión con el DNA. La unión de σ^{54} con el DNA se realiza a tráves de interacciones especificas entre la region III y el promotor. Estas interacciones han sido detectadas mediante puentes químicos y, probablemente son moduladas indirectamente por otras subregiones de la misma región, aparentemente organizadas como subdominios independientes (Sasse-Dwight & Gralla, 1990; Cannon *et. al.*, 1997). La remoción de fragmentos internos y mutantes sencillas en esta región de σ^{54} han demostrado producir

proteínas deficientes en la unión al DNA (Wang, *et. al.*, 1994; Merick & Chambers, 1992). En este lugar se localiza una hélice-vuelta-hélice y una región de ocho aminoácidos altamente conservados entre factores σ^{54} de distinto origen, llamada caja RpoN ("RpoN Box"; Coppard *et. al.*, 1991). Sustituciones en cualquiera de los ocho residuos que forman la caja RpoN, interrumpen la función de σ^{54} y disminuyen la afinidad de la holoenzima por el DNA. El dominio de unión al DNA (residuos 329-477) está constituido por dos subdominios (180-306 y 306-477) separables por digestión proteolitica (Cannon *et. al.*, 1997), y distinguibles por sus patrones de desnaturalización termal. Uno de estos subdominios lo constituye el fragmento 306-477 donde se localiza la hélice-vuelta-hélice y la caja RpoN que determinan la interacción directa con las secuencias consenso del promotor. Se ha encontrado que el subdominio 180-306 puede modular la unión con el DNA *in vitro* (Cannon *et. al.*, 1997). Cuando este subdominio se encuentra incluido en el polipéptido completo, la actividad de unión con el DNA se incrementa varias veces (Cannon *et. al.*, 1995). Por lo contrario, la remoción de un fragmento interno de este subdominio produce mutantes incapaces de unirse al DNA (Wong *et. al.*, 1994). Es interesante que estos subdominios estén separados por un conector corto sensible a digestión proteolítica, cuya función mínima es la de unir los dominios, pero también pudiera estar implicado en la interacción con el activador transcripcional. Hoover & Lee (1994) localizaron un residuo de cisteína en esta zona, cuya alteración impide la unión de σ^{54} con el activador, tal como es detectado en estudios de puentes químico.

.. ".voy por tu frente como por la luna...

Capítulo 3

El mecanismo inicio de la transcripción.

El inicio de la transcripción depende de la interacción especifica de los componentes mencionados anteriormente y ha sido descrita en los términos de la ecuación de Michaelis-Menten, usada para explicar la interacción de cualquier enzima (E) con un sustrato (S) para producir un producto (P) (Figura 7 a). De acuerdo con estos razonamientos, el inicio de transcripción es una reacción de la que es posible discernir los estados intermedios y de transición. Aunque experimentalmente se ha llegado a la conclusión de que el inicio de la transcripción es un proceso de múltiples pasos, se han descrito dos intermediarios principales: el complejo transcripcional cerrado (RPC) y el complejo transcripcional abierto (RPO).

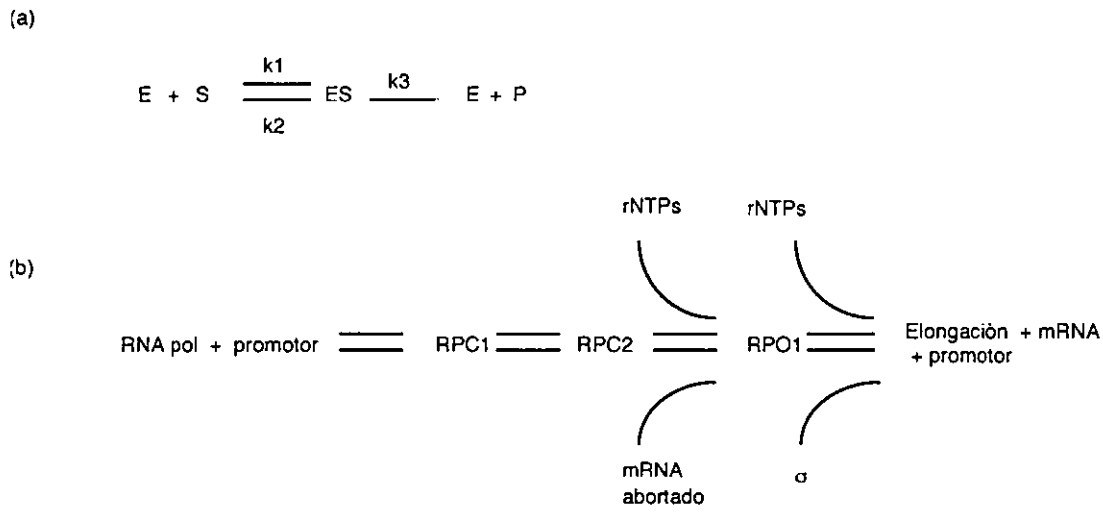

Figura 7. El mecanismo de inicio de la transcripción.

3.1. Formación del complejo cerrado. En este estado la doble hélice del DNA en el promotor permanece cerrada, es decir apareada, aún cuando la holoenzima se encuentre estrechamente unida. Dos clases de experimentos se han usado para medir las

interacciones RNA pol ^{H}-promotor en este estadío de la transcripción: la formación de complejos transcripcionales a baja temperatura y ensayos de interferencia o de protección del DNA a diferentes agentes químicos o a enzimas como la DNAasa I en presencia de la RNA pol ^{H} ("footprinting"). De esta clase de estudios se ha deducido que la primera interacción con la holoenzima ocurre en las regiones hacia arriba de la región -35 del promotor. Subsecuentemente diversos cambios conformacionales en la holoenzima extienden los sitios de contacto hacia el sitio de inicio de la transcripción. Estos pasos se han denominado como RPC1 y RPC2 respectivamente (Figura 7 b). RNA pol-σ^{70} y RNA pol-σ^{32} protegen el promotor en una sola cara del DNA durante RPC1 abarcando aproximadamente desde -55 hasta -5, sin llegar al sitio de inicio de la transcripción (Bokal *et. al.*, 1995; Hofer *et. al.*, 1985; Kovacic *et. al.*,1987; Schikor, 1990). Este patrón se ve alterado en RPC2 en donde la holoenzima cubre el sitio de inicio de la transcripción y los primeros 20 nucleótidos, sin alterar la protección de la región del promotor alrededor de -35. Se ha especulado que en RPC1 la holoenzima contacta el promotor a tráves del canal observado en la estructura obtenida con ME (Figura 4; Poliyakov *et. al.*, 1995), y dado que en RPC2 ambas caras del DNA en el promotor muestran algún tipo de protección, se cree que este paso pudiera ser aquel en el que se observa el canal completamente cerrado en la estructura obtenida por ME.

Estudios *in vivo* e *in vitro* han demostrado que la RNA pol-σ^{54} forma complejos cerrados estables con los promotores *glnAp2* de *K. pneumoniae*, *nifH* de *R. meliloti* y *nifH* y *nifU* de *K. pneumoniae*, a diferencia de aquellos formados por holoenzimas constituidas con factores de la familia σ^{70} (Sasse-Dwight & Gralla, 1991; Ninfa, *et. al.*, 1987; Popham *et. al.*, 1989; Buck & Cannon, 1992). Mediante ensayos de protección al dimetil sulfato se detectó que NtrC impide la metilación de las guaninas -13, -24 y -25 y -12, -14,-21 y -23 de la hebra no codificante de *glnAp2*. En el promotor *nifH* de *R meliloti* se protege la -11, -13, -22 y -24, por lo que ha concluido que las interacciones de la RNA pol-σ^{54} con estos

promotores son equivalentes (Morett & Buck, 1989). Las mutaciones que disminuyen la fuerza del promotor y que caen en -12, -13, -24 y 25 (Brown & Ausubel, 1984; Ow *et. al.*, 1985; Buck *et. al.*, 1985; Khan *et. al.*, 1986) pueden ser explicadas pensando en una unión deficiente del promotor con la RNA pol ^{H}. Otros estudios de protección y de interferencia con diversos agentes químicos han pemitido saber que, en promotores activos, ocurre una distorsión en la región adyacente a -12, dependiente de la unión con la RNA pol ^{C} y de la integridad del amino terminal del factor σ^{54} (Morris *et. al.*, 1994). Esto refleja un constreñimiento de la región del DNA que posteriormente será abierto. En el complejo cerrado los contactos están restringidos a una cara del DNA, en el surco mayor, mientras que el surco menor no se observa protegido. A diferencia de lo observado con la RNA pol-σ^{70} en donde la RNA pol ^{C} también contribuye a la unión con el DNA, en el complejo cerrado formado con la RNA pol-σ^{54} todos los contactos con el promotor han sido atribuidos al factor σ^{54}. De hecho se conoce que las RNA polimerasas constituidas con subunidades α carentes del dominio carboxilo, se asocian con el factor σ^{54} e inician transcripción *in vitro*, a partir de los promotores *glnAP2* o *nifH*, en presencia de las proteínas activadoras NtrC o NifA (Kustu *et. al.*, 1993). La subunidad α, como mencionamos en el Capítulo 1, interviene en la interacción con la región -35 del promotor.

La estabilidad del complejo cerrado formado por la RNA pol-σ^{54}, ha llevado a pensar que un complejo así constituido, representa un blanco fijo para la acción de la proteína activadora, dado que ésta se une lejos del promotor, y en ciertos casos puede activar sin estar unida al DNA. Según ésta idea una holoenzima que continuamente esté iniciando transcripción no podría ofrecer esta ventaja.

3.2. Formación del complejo abierto. Después de la estabilización del complejo cerrado en los promotores σ^{70}, ocurre la separación de las dos hebras del DNA. Hay un cambio en la reactividad de las bases pirimidinicas que se encuentran expuestas en cadenas sencillas

hacia el $KMNO_4$, desde las regiones proximales a -10 hasta una base más allá del sitio de inicio de la transcripción (Kierkegaard *et. al.*, 1983; Sasse-Dwight & Gralla, 1989; Sienbelist, 1979; Suh *et. al.*, 1993). Esta separación permite el acceso de radicales químicos de carga negativa, por lo que es posible detectarla utilizando este tipo de iones como sonda. Se supone que es en este paso donde se incorpora el primer rNTP (Craig *et. al.*, 1995; Suh *et. al.*, 1993). Sin embargo, el patrón de protección al DMS no cambia con respecto al estadío RPC2, indicando que la apertura inicial del promotor, cerca del inicio de transcripción, no requiere de cambios adicionales en los contactos de la holoenzima con el promotor.

El transcrito de RNA mensajero generalmente comienza en una base púrica situada entre los primeros 6 u 8 nucleótidos. Cuando ya se ha unido el primer rNTP, la holoenzima tiene una serie de cambios conformacionales que conducen a la completa desnaturalización del promotor. Después de la polimerización de 7 a 12 ribonucleótidos la holoenzima deja de contactar el promotor, libera el factor σ y empieza a elongar el transcrito, es decir a caminar sobre la hebra codificante del DNA, incorporando los ribonucleótidos respectivos. Este complejo es más estable a la disociación que cualquiera de los anteriores. El complejo abierto inicial (RPO1) y el de arranque (RPO2) se encuentran en equilibrio, de tal manera que se sintetizan transcritos cortos y se liberan continuamente hasta que por fin la RNA polimerasa procede a elongar la cadena de mRNA. La proporción de transcritos abortivos *versus* transcritos largos varía dependiendo del promotor, lo cual sugiere la existencia de elementos en la secuencia del promotor que influyen en la eficiencia de inicio de la transcripción. No se conoce la razón por la cual la holoenzima pocede de esta manera, pero se especula que esto representa un mecanismo adicional para asegurar la especificidad de la transcripción en los lugares correctos del genoma..

"voy por tu vientre como por tus sueños . . .

Capítulo 4

Regulación del inicio de la transcripción.

Aún cuando es posible regular la expresión genética prácticamente desde el inicio de la transcripción hasta la traducción del mRNA y postraduccionalmente, hacer o no un RNA mensajero es un mecanismo común y eficiente. Generalmente, los genes escenciales para la célula son expresados constitutivamente, y los genes cuyos productos son necesarios solamente en ciertas condiciones de crecimiento o de desarrollo vegetativo son regulados por medio de proteínas represoras (regulación negativa) o activadoras (regulación positiva). En ambos casos se necesita que la proteína reguladora se una a sitios operadores en el DNA que anteceden a los promotores. El ejemplo clásico de regulación negativa sin duda lo constituye el operón *lac*. El represor LacI se une al sitio operador en el DNA mantiendo prácticamente bloqueado el acceso de la RNA polimerasa al promotor del operón (*lacZYA*) cuando *E. coli* crece en ausencia de lactosa. Cuando hay lactosa presente en el medio de crecimiento, ésta actúa como un efector alostérico sobre LacI impidiéndole unirse con el DNA. Sin dejar de considerar que la regulación negativa tiene un papel importante en la homeoestasis celular y que algunas proteínas reguladoras tienen un papel dual como represores y activadores (Gralla & Collado-Vides, 1994), en esta revisión nos referiremos exclusivamente a los mecanismos de activación de la transcripción.

Existen numerosos genes bacterianos regulados positivamente por proteínas cuyos sitios de unión en el DNA están en la región proximal a los promotores, o por proteínas que se unen lejos del promotor (activación a distancia). Este último caso podría considerarse como un modo de inicio de la transcripción, más que de activación, porque la RNA pol-σ^{54} escencialmente no produce ningún transcrito en ausencia de la proteína activadora. En ambos casos de activación es necesario que el regulador interaccione con el DNA y con algún componente de la RNA polimerasa para modificar la tasa de iniciación

de la transcripción (Figura 8), pero existen diferencias profundas entre ambos mecanismos que detallaremos enseguida.

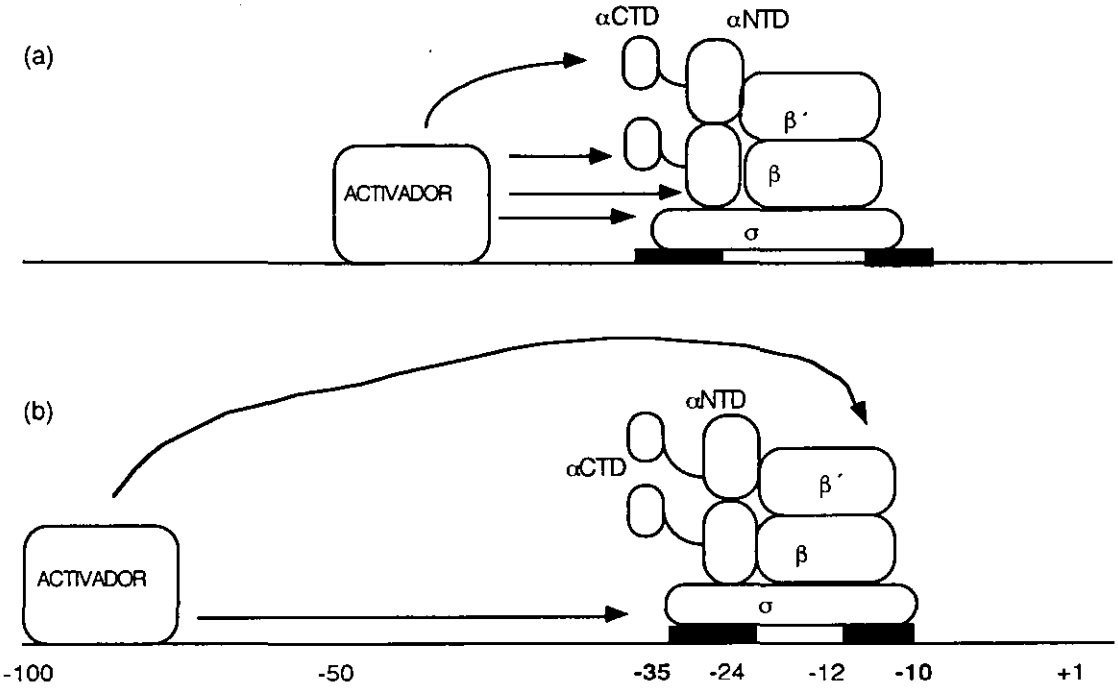

Figura 8. Interacciones proteína-proteína entre los activadores de acciòn cercana (a) y de acción a distancia (b).

4.1 La unión con el DNA. Los sitios de unión al DNA generalmente son simétricos, es decir están compuestos por una región conservada de DNA con mitades opuestas e iguales. Esto sugiere que los activadores se unen al DNA como moléculas diméricas. Algunos genes analizados presentan no solamente un sitio sino varios sitios de unión para el activador, no necesariamente iguales entre sí. En estos casos todos los sitios de unión son ocupados por la proteína activadora de manera cooperativa. La cooperatividad de la unión de las proteinas activadoras con el DNA depende de sitios operadores de alta afinidad para el pegado inicial de una molécula dimérica, y de las interacciones proteína-proteína entre los dímeros para favorecer la unión a un segundo o a un tercer sitio de pegado. Una demostración elegante de cooperatividad entre proteínas reguladoras se

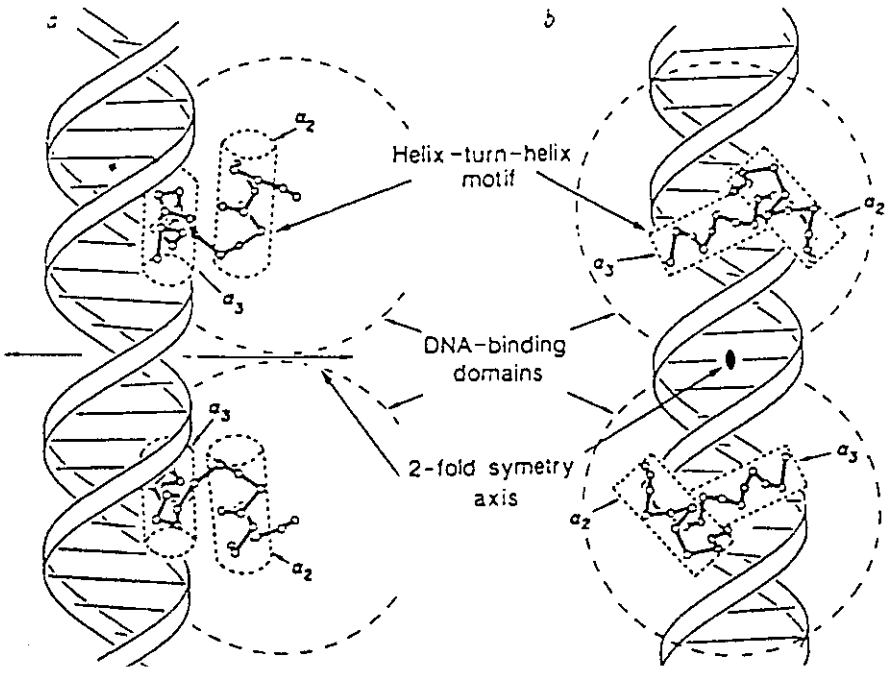

Figura 9. La hélice-vuelta-hélice, un motivo estructural común para la unión con el DNA.

encontró en el estudio de la decisión lisis-lisogenia del bacteriófago λ, donde es necesaria la interacción cooperativa de λ cI en tres sitios operadores para una total represión de un conjunto de genes y la activación de otros (Ptashne, 1992).

Las interacciones proteína-DNA de muchos activadores de bacterias y de algunos reguladores eucariotes se realizan a tráves de un módulo estructural conocido como hélice-vuelta-hélice (HTH; Figura 12). La HTH está compuesta generalmente por un segmento de 20 residuos que constituyen dos α hélices separadas en un ángulo de 120 Å (Paabo *et. al.*, 1984). Este módulo se encuentra embebido entre dominios estructuralmente diferentes (Agaarwal & Harrison, 1990) y localizada comúnmente en el extremo amino terminal de la proteína (Pérez-Rueda & Collado-Vides, 1997). La segunda hélice es denominada "de reconocimiento", debido a que cae sobre el surco mayor de la doble hélice y porque la substitución de los residuos expuestos resultan ser críticos para la especificidad (Wharton *et. al.*, 1984). La estructura cristalográfica de la HTH de varias proteínas es conocida y para otras se han hecho predicciones basándose en algoritmos de estructura secundaria (Brennan & Mathews, 1988; Dodd & Egan, 1990)

4.2. Activación cercana. Esta clase de proteínas activadoras es sin duda la más diversa en cuanto al tipo y número de genes que controla y ciertamente existe una diferenciación en la posición de los sitios que ocupan los activadores en la región de arriba del promotor, pero generalmente estos sitios se localizan muy cerca de la región -35 (Ishihama, 1994). Un análisis detallado de las posiciones de los sitios de unión de estas proteínas reguladoras con respecto al sitio de inicio de la transcripción en una colección de 132 promotores de la RNA pol-σ^{70} de *E. coli* fué realizado por Gralla & Collado-Vides (1994). Sus resultados muestran que 65 de estos promotores son activados y que los demás se dividen entre aquellos que son regulados por represión o pueden ser regulados de manera dual, tanto por represión como por activación. De acuerdo a su análisis, los activadores se unen preferentemente en sitios entre -80 y -30. Es notable que la mayoría de los activadores

que tienen sitios de unión distales, como CysB y TyrR, también tienen sitios de unión adicionales proximales al promotor. La localización de los sitios de unión en el DNA y además del espaciamiento de estos cuando existe más de un sitio, sugiere que la proteína activadora es capaz de tener contactos estéricos con la RNA polimerasa y que los activadores deben localizarse en una cara del DNA para quedar orientados en forma apropiada hacia la RNA polimerasa. Dos activadores de esta clase han sido bastante estudiados, la proteína catábolica activadora (CAP) y el represor λ cI.

4.2.1 Activación mediante CAP. La proteína catabólica activadora (CAP o CRP, proteína receptora de adenosín monofosfato cíclico, cAMP) es uno de los casos mejor caracterizados de activación cercana al promotor. En *E. coli* CAP activa la transcripción de más de 100 diferentes promotores uniéndose a los operadores en presencia de cAMP. Preferencialmente los operadores se encuentran a -40 y -60 con respecto al sitio de inicio de la transcripción (Pérez-Rueda, 1998). Los promotores dependientes de CAP han sido clasificados en dos clases de acuerdo a la posición de los sitios de unión a CAP respecto a las regiones -10, -35. En los de clase I, cuyo promotor típico es *lac*, el operador se encuentra centrado alrededor de -61.5 hasta -92.5 (Mandecki & Craruthers, 1984; Straney *et. al.*, 1989; Gaston *et. al.*, 1990; Ushida *et. al.*, 1990), pero siempre en la misma cara de la hélice en la cual se une la RNA polimerasa. En los promotores de Clase II, el operador sobrelapa la región -35, aparentemente reemplazando los determinantes de unión de la RNA polimerasa en esta región. El promotor prototipo de esta clase es *gal*, cuyo sitio de unión con CAP esta centrado en -41.5 . La posición del operador en estas clases de promotores ha sugerido a muchos, que los mecanismos de activación de la transcripción son diferente, aún cuando ambos dependan de CAP (Bell *et. al.*, 1990; Williams *et. al.*, 1991; Igarashi *et. al.*, 1991; Merkel *et. al.*, 1993; Pérez-Rueda, 1998)

CAP es un homodímero que ya ha sido cristalizado solo y en complejo con el DNA (Schultz *et. al.*, 1991) en donde cada monómero interactúa con una de las mitades del sitio

de unión. En estado cristalino CAP produce un doblamiento del DNA de aproximadamente 90°, de acuerdo con lo estimado en solución entre 80-100°. Presuntamente el doblamiento del DNA en la region cercana al promotor es una actividad importante de CAP, necesaria para una completa activación transcripcional (Collis et. al., 1989; Gartemberg & Crothers, 1988). Busby & Ebright (1997) han propuesto que CAP activa la transcripción mediante dos mecanismos diferentes de acuerdo a la arquitectura del promotor. Uno es mediante una interacción directa con la holoenzima en promotores de clase I (p*lac*) y otro ejerciendo un efecto indirecto eliminando una probable acción inhibitoria de la subunidad α en promotores de clase II (p*gal*).

4.2.2. Interacciones entre CAP y la RNA polimerasa. El mecanismo de activación transcripcional mediado por CAP requiere de su unión como dímero en la misma cara del DNA en que se encuentra la RNA pol ^{H}. Se conoce que CAP y la RNA pol ^{C} pueden interactuar en solución y que se unen de manera cooperativa en la región promotora (Straney *et. al.*, 1989; Ren *et. al.* 1988). Esto indica que en ambas moléculas deben existir una o varias regiones de interacción proteína-proteína. Alteraciones en regiones específicas de CAP únicamente afectan las propiedades transcripcionales de CAP, sin modificar ninguna otra como la unión al DNA, el doblamiento del DNA o la estructura cuaternaria de la proteína, por lo que se ha deducido que estas mutaciones alteran la superficie de reconocimiento con la RNA polimerasa (Bell *et. al.*, 1990; Eschenlauer & Reznikoff, 1991). Esta clase de mutantes descritas inicialmente en la proteína λ cI, son llamadas de control positivo (PC o "positive control"; Guarente *et. al.* , 1982). Mediante diversos procedimientos de mutagénesis, tres regiones de CAP han resultado involucradas en el control positivo. Una de ellas corresponde a los residuos 156-164, que estructuralmente forman un asa expuesta en la superficie de la proteina. En esta región es probable que los aminoácidos treo-158, pro-168 y met-163 participen contactando directamente la RNA polimerasa mediante su cadena lateral (Zhou *et. al.*, 1993; Eschenlauer & Reznikoff. 1991). Sustituciones en estas posiciones por aminoácidos que acortan o alargan la cadena

lateral producen proteínas completamente afectadas en su capacidad para activar la transcripción (Niu *et. al.*, 1994; Zhou *et. al.*, 1994). Hay diversas clases que señalan que el dominio carboxilo terminal de la subunidad α es el blanco de estas interacciones en la holoenzima. Genéticamente se ha demostrado que mutaciones puntuales o la remoción de éste dominio producen holoenzimas insensibles a la respuesta hacia CAP en los promotores de clase I (Igarashi & Ishihama, 1991; Zou *et. al.*, 1992). En el carboxilo terminal de α se localizan residuos especificos que pueden ser ligados quimicamente a CAP, mediante un agente quimico fotoactivable (Chen *et. al.*, 199; Kolb *et. al.*, 1993). Mediante espectrofotometría de fluorescencia se ha demostrado la asociación entre CAP y la subunidad α en solución en ausencia del promotor, la cual no ocurre con una mutante de CAP afectada en una región de control positivo (Heyduk *et. al.*, 1993). Adicionalmente, anticuerpos monoclonales dirigidos contra la subunidad α bloquean especificamente la acción de CAP (Sharif *et. al.*, 1994). Mediante heterodimeros orientados ha sido posible demostrar que el contacto de un solo monómero con la subunidad α, es suficiente para activar la transcripcion, siempre y cuando el monómero activo se encuentre unido al DNA en la proximidad con la RNA polimerasa (Zhou *et. al.* 1994; Zhou *et. al.*, 1993).

Una segunda región de control positivo definida por mutagenesis no involucra residuos cercanos en la secuencia primaria pero sí cercanos estructuralmente (his-19, his-21, glu- 96 y lys-101). Esta región funciona exlusivamente para la activación de los promotores de Clase II, pero no con los de clase I. Es llamativo que sustituciones en glu-96 resulten en formas de CAP superactivadoras de los promoteres de Clase II (Williams *et. al.*, 1996) Estos aminoácidos contactan a tráves de los residuos glu-162, glu-63, asp164 y glu-165 complementarios en carga con los de CAP, y se localizan en el amino terminal de la subunidad α (Niu *et. al.*, 1996), por lo que se ha supuesto que interacciones electrostáticas dominan este tipo de interacción. En este caso la activación se debe a un

incremento en la tasa de isomerización del complejo cerrrado a complejo abierto y no a un incremento en la afinidad de la RNA polimerasa con el promotor (Niu *et. al.*, 1996).

Algunos autores han sugerido que existe una tercera región para el control positivo en CAP, formada por los residuos 52-58 (Williams *et. al.*, 1991) debido que mutaciones en lys-52 revierten los efectos de mutaciones en el control positivo en la posición 159. Esto ha sido interpretado como resultado de la creación de otra zona de contacto con la RNA pol ^{H}, o debido a un efecto indirecto por la formación de un ambiente electrostático favorable para la interacción con la RNA polimerasa. Se conoce que estructuralmente la lisina 52 forma un puente salino con el glutamato 58, de tal modo que alteraciones en la lisina 52 podrian generar este ambiente al dejar libre el glutamato. Jin *et. al.*, (1995), han encontrado evidencia mediante puentes químicos de que en el promotor de galactosa (p*gal*) el asa 52-58 de CAP ésta en proximidad de σ en el complejo de preiniciación y que una mutación lys-52-asn, aumenta la eficiencia de la formación de complejos ligados.

4.2.3. Interaccciones entre λ cI y la RNA polimerasa. Las proteínas represoras cI de los fagos λ y 434 y c2 de p22 funcionan también como activadores transcripcionales. Se ha demostrado que la activación de la transcripción mediante estas proteínas involucra una hélice anfipática de la proteína que tiene uno o más residuos cargados negativamente accesibles al solvente (Guarente *et. al.*, 1982; Hostchild *et. al.*, 1983; Bushman &.Ptahsne, 1986). Substituciones de estos aminoácidos produce proteínas deficientes en activación pero competentes en el doblamiento y en la unión con el DNA. Algunos activadores eucarióticos como GAL4 y GCN4 de levadura, también tienen aminoácidos negativamente cargados en la región de activación (Ma & Ptashne, 1986; Giniger & Ptashne, 1986; Triezenberg *et. al.*, 1988). Ptashne y colaboradores han propuesto que el mecanismo de activación transcripcional se lleva a cabo mediante estas regiones ácidas localizadas en la superficie. Este punto de vista, ha resultado controvertido dado que, como hemos descrito para CAP y se ha observado en otras proteínas activadoras, no parece existir un

requerimento estricto de secuencias negativamente cargadas para activar la transcripción (Zhang *et. al.*, 1992). Por lo contrario, el análisis de una variedad de interacciones entre diferentes activadores y la RNA polimerasa sugiere que no hay ningún requerimento especial en secuencia en esta clase de interacciones proteína-proteína. Varios tipos de experimentos sugieren incluso que dominios de proteínas que no tienen relación con la activación transcripcional son capaces de funcionar en un sistema híbrido que contenga las propiedades de unión con el DNA en la región cercana al promotor. En el experimento de Dove *et. al.* (1997) por ejemplo, el dominio carboxilo de la subunidad α fué reemplazado por el dominio carboxilo del represor de λ cI, un dominio que tiene los determinantes para la oligomerización de la proteína. Cuando λ cI se une cerca de esta RNA polimerasa quimérica en una posición desde donde no puede usar su región de activación localizada en el amino terminal, interacciona mediante el dominio carboxilo con la subunidad α- cI, y dispara la transcripción (Figura 10).

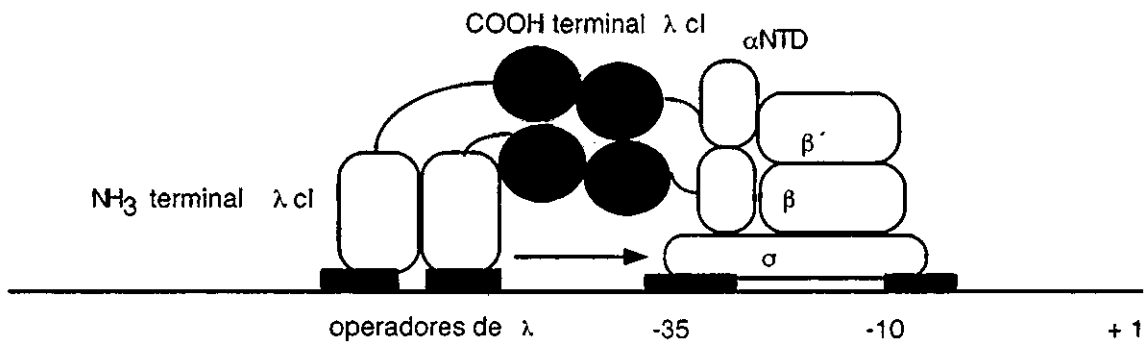

Figura 10 Activación de la RNA polimerasa mediante interacciones artificiales (Dove et. al ., 1997).

Diversos enfoques experimentales han sido utilizados para definir las interacciones entre los activadores y la RNA polimerasa. Como mencionamos arriba, CAP interacciona con la subunidad α y con σ^{70} (Figura 8). Li *et. al.* (1994) han encontrado que mutaciones en σ^{70} son capaces de suprimir mutaciones PC del represor cI de λ. La presencia de

múltiples sitios de interacción de los activadores con la holoenzima, se ha hecho evidente mediante la uitilización de sistemas de expresión artificiales. En una de las variaciones del experimento de Dove *et. al.*, (1997), si el sitio de unión con λ cI es colocado muy cerca de la RNA polimerasa, λ cI puede hacer dos contactos, el ya descrito que involucra la oligomerización con α-cI y otro con σ^{70}. Estas dos interacciones juntas tienen un efecto sinergístico sobre la transcripción, elevándola a niveles no alcanzados con una sola interacción por separado. Ejemplos de sinergismo también han sido descritos cuando dos proteínas activadoras (CAP y λ cI) comparten sitios en la vecindad del promotor, interaccionando especificamente con la holoenzima (λ cI con σ^{70} y CAP con α; Joung *et. al.*, 1994; Joung *et. al.*, 1993).

4.3. La activacion a distancia. Este tipo de regulación de la transcripción involucra interacciones entre la RNA polimerasa y proteínas activadoras unidas al DNA en sitios alejados del promotor (más allá de 100 pb hasta 2 o 3 kb). Esta clase de activación pertenece a un grupo más amplio de fenómenos de reconocimiento molecular que ocurren entre proteínas separadas por cierta distancia en el DNA (Ptashne, 1986; Wang & Giaever, 1988). La represión mediante AraC, λ cI, LacI y GalR y la acción de las recombinasas como HinF son ejemplos de interacciones de este tipo (Lobell & Schleif, 1990; Hanna-Rose & Hansen, 1996; Hoschild & Ptashne, 1988). Varias hipótesis han sido aventuradas para explicar el mecanismo molecular de la activación a distancia, pero sin duda la evidencia experimental apunta hacia el doblamiento y la formación de un asa con la secuencia de DNA entre el sitio de unión de la proteína reguladora y su blanco de interacción, la RNA polimerasa, de tal manera que el activador entra en contacto con la RNA polimerasa unida al promotor (Ptashne, 1986; Wang & Giaever, 1988; Hoschild & Ptashne, 1988). El "looping" es también el mejor modelo para explicar la represión a distancia. Por ejemplo, AraC se une al DNA en dos sitios separados por 200 pares de bases (Soisson *et. al.*, 1997). Cuando AraC está unido en el sitio proximal al promotor

activa la transcripción del operón *araBAD*. La represión ocurre cuando AraC se une simultáneamente a los sitios distal y proximal, y ambos dímeros se ponen en contacto. En cualquiera de estos dos casos, activación o represión a distancia, es necesario que las moléculas que interactúan estén en la misma cara del DNA. Alteraciones en la orientación de los sitios de unión con respecto a la periodicidad de la doble hélice conducen a graves fallas en la regulación de la transcripción.

4.3.1 Los dominios funcionales de los activadores a distancia. Las proteínas activadoras a distancia están formadas por módulos o dominios estructurales independientes que es posible separar (Frankel & Kim, 1991). Por ejemplo, un fragmento de las proteínas de *Sacharomyces*, Gal4 y GCN4, es capaz de unirse al DNA sin llevar a cabo la activación de la transcripción (Hope & Struhl, 1986). Estas proteínas truncadas pueden funcionar como represores de la transcripción cuando se unen

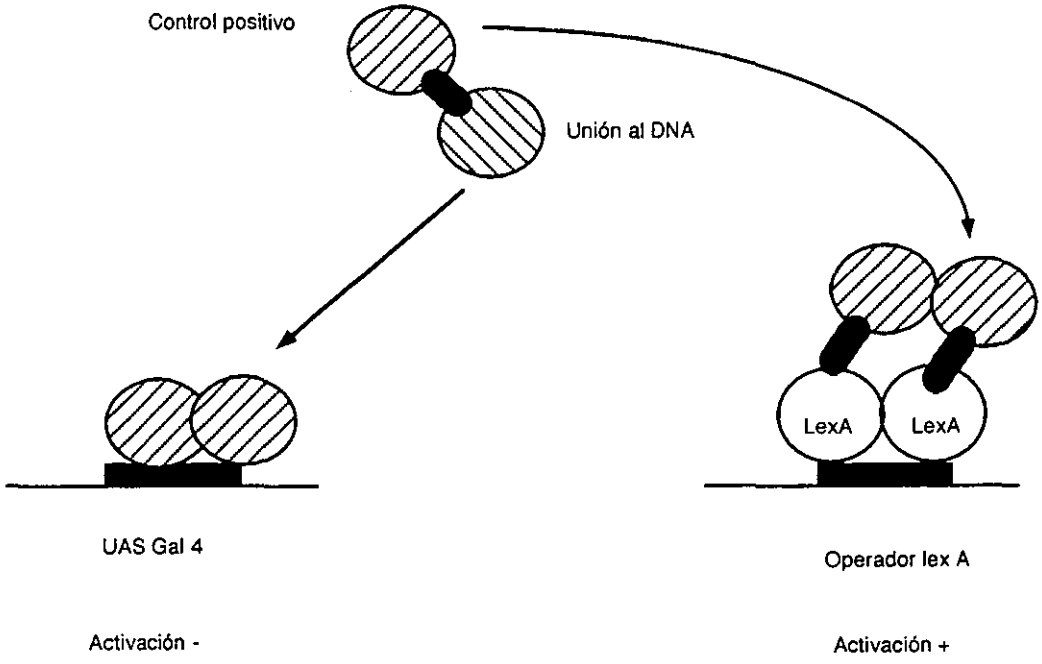

Figura 11 . Separación de las funciones de unión al DNA y de control positivo en un regulador eucariótico (Ptashne, 1986).

a sitios operadores intercalados entre el promotor. Brent & Ptashne (1985) demostraron que los dominios de unión al DNA y de activación son entidades independientes, mediante la construcción de proteínas híbridas entre Gal4 (74 aminoácidos terminales) y el represor LexA de *E. coli* (Figura 11). Esta proteína híbrida es capaz de activar la transcripción en levadura a partir de un promotor sintético que lleva el sitio de reconocimiento para LexA. La proteína híbrida se une a tráves del sitio operador para LexA y estimula la transcripción a partir de un dominio de Gal4. De la misma manera el híbrido LexA-Gal4-LexA se comporta como una proteína bifuncional que puede activar la transcripción a partir de la unión con cualquiera de los sitios de reconocimiento para LexA o Gal4 (Hope & Sthrull, 1986). Estos conceptos han sido aprovechados exitosamente en el diseño de sistemas para la detección de interacciones proteína-proteína (Fields & Song, 1989).

4.3.2. La estructura de las regiones de activación. Las regiones de activación de Gal4, GCN4 y PHO4 tienen en común un exceso de aminoácidos cargados negativamente aparentemente sin estructura definida (Hahn, 1993; Ma & Ptashne, 1987; Owaga & Oshima, 1990). Algunos investigadores han encontrado en Gal4 una relación entre la longitud de la región activadora y el número de residuos ácidos con la potencia del activador (Gill & Ptashne, 1987). Al igual que en los activadores bacterianos de acción cercana este punto de vista ha resultado sumamente controvertido. Existe evidencia en contra de que las regiones ácidas de las proteínas son indispensables para la activación transcripcional (Leuther *et. al.*, 1993; Van Hoy *et. al.*, 1993). La eliminación de los residuos ácidos por mutación o remoción completa de la región de activación de Gal4, o aún la adición de cargas positivas, solamente tiene efectos modestos en la activación transcripcional (Leuther *et. al.*, 1993). Estas contradicciones han sido consideradas como resultado de las diferencias en los métodos genéticos usados para obtener las variantes de Gal4, pues mientras Ptashne y colaboradores usaron proteínas híbridas, en este último caso se usó la proteína completa (Hahn *et. al.*, 1993). De cualquier manera no parece haber

un requerimento estricto en la composición de la secuencia de aminoácidos que constituyen la región de activación, e igualmente controvertido ha resultado asignar una estructura secundaria específica a estas regiones. De hecho, de acuerdo a los experimentos de Struhl, solamente es necesario que la proteína activadora se una al DNA y proporcione un elemento de contacto con la RNA polimerasa para que esta proceda a iniciar la transcripción, sin importar que este elemento este contenido en una proteína no activadora *per se* como la TATA binding protein (TBP; Chartejee & Struhl, 1995). Siegler *et. al.*, (1992), aventura la idea de que las regiones de activación carecen de estructura definida y pemanecen en un estado parecido al "molten globule", y sólo son estructuradas cuando interaccionan con el activador. Sin embargo, Leuther *et. al.* (1993), utilizando dicroismo circular sugiere que la región de activación de Gal4 forma una hoja β, y Ptashne ha defendido la idea de que los residuos expuestos al solvente en una hélice anfipática son los responsables de la interacción con la RNA polimerasa, como sucede en λ cI y otros activadores de los bacteriófago (Bushman & Ptashne, 1988; Bushman & Ptashne, 1986; Hoschild *et. al.*, 1983). La ausencia de estructuras definidas en las regiones de activación de los activadores bacterianos de acción cercana y en los activadores a distancia eucarióticos han conducido a la conclusión de que ambos utilizan un mecanismo de acción común (Ptashne & Gann, 1997). Este mecanismo, llamado reclutamiento consiste en ayudar a la RNA polimerasa a unirse con mayor eficiencia al promotor, dado que esta no forma complejos estables en el promotor (Choy & Green, 1993; Koleske & Young, 1994).

En bacterias la activación a distancia ocurre mediante proteínas de la familia de proteínas activadoras (EBPs), que funcionan concertadamente con la RNA polimerasa compuesta por el factor σ^{54} (Kustu *et. al.*, 1989; Morett & Segovia, 1993). Desde el punto de vista mecanístico existen varias características en la manera de actuar de las EBP que son escencialmente similares a los descritos para las proteínas eucarióticas Gal4 y GCN4. Una de estas similitudes es la estructura en dominios de las EBPs. Sin embargo una diferencia fundamental es la necesidad de una actividad de ATPasa asociada a las EBPs y

que se encuentra acoplada a la formación del complejo abierto (Weiss *et. al.*, 1991; Wedel *et. al.*, 1995). Además se ha demostrado que el dominio Central de algunas de estas proteínas es capaz de activar la transcripción desde el citoplasma bacteriano, sin necesidad de unirse al DNA (Huala & Ausubel, 1989; Huala *et. al.*, 1990; Flasner *et. al.*, 1996), lo cual no occurre en los activadores de eucariotes (Ptashne & Gann, 1997).

"voy por tu talle como por un rio...

Capitulo 5

La familia bacteriana de proteínas activadoras a distancia

En bacteria la activación de la transcripción a distancia no es tan común como en Eucaria, y hasta ahora no conocemos ningún caso en Archaea. La familia de las EBPs de origen bacteriano controlan la expresión de genes que sirven para diversas funciones que no son absolutamente indispensables para la vida de la bacteria en el laboratorio. Entre estas tareas celulares se encuentran desde sistemas de transporte específicos, de utilización de fuentes alternativas de carbono y de energía, producción de estructuras extracelulares y determinantes de virulencia (Kustu *et. al.*, 1989; Shingler, 1996). Debido a que estas son respuestas eventuales de la bacteria a las condiciones ambientales en las que se encuentre, es necesario que la respuesta transcripcional sea regulada eficientemente. En algunos casos la señal medio ambiental necesita ser traducida por proteínas auxiliares para generar una respuesta molecular adecuada que active la transcripción. Por ejemplo, en limitación de amonio NtrB fosforila NtrC, una EBP que promueve la transcripción de varios operones requeridos para el metabolismo nitrógenado (Nixon *et. al.*, 1986). Este tipo de regulación que consiste de una molécula receptora y otra efectora es denominado "sistema de dos componentes" (Albright *et. al.*, 1989). Esta es una estrategia de regulación ampliamente utilizada para el control de la expresión genética mediante otros activadores que no pertenecen a las EBPs tales como CheY (Stock *et. al.*, 1988). En otras EBPs la señal ambiental actúa directamente sobre la proteína reguladora (DmpR, XylR y FhlA) modificando su actividad. El trabajo experimental que llevamos a cabo trató sobre la proteína activadora NifA de *B. japonicum*, que controla los genes de fijación de nitrógeno y cuya actividad es regulada principalmente por las concentraciones de oxígeno.

5.1. Estructura en dominios. Abundantes evidencias genéticas, bioquímicas y evolutivas señalan que NifA y en general todas la EBPs, están formadas por tres dominios capaces de plegarse y funcionar por sí mismos (Drummnond *et. al.*, 1986; Drummond *et. al.*, 1990;

Morett & Segovia, 1993). El dominio Amino terminal sirve para la recepción de señales moleculares, el dominio Central tiene actividad de hidrólisis de ATP y la función de control positivo y el dominio Carboxilo sirve para la unión del activador con el DNA (Figura 12).

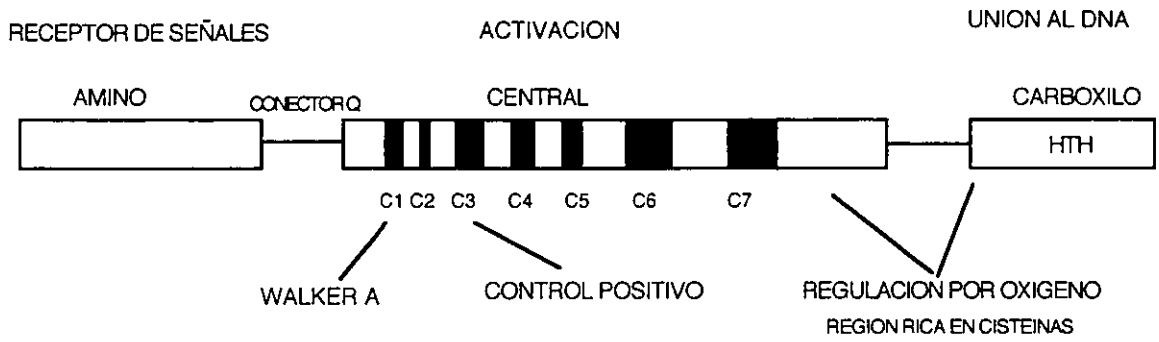

Figura 12. Estructura en dominios de las EBPs

El dominio Amino. Algunas EBPs necesitan ser modificadas postraduccionalmente para activar la transcripción (Ninfa *et. al.*, 1995). La fosforilación de aminoácidos específicos localizados en el dominio Amino de algunas EBPs, como el aspartato-54 de NtrC, estimula la unión cooperativa al DNA y la hidrólisis de ATP (Keener & Kustu, 1988). El mismo efecto se obtiene mediante la unión directa de un cofactor como fenol (DmpR), tolueno (XylR) o formato (FhlA) en el dominio Amino (Shingler & Moore, 1994; Delgado & Ramos, 1994; Shingler & Pavel, 1995; Hopper & Böck, 1995). La remoción del dominio Amino de estas proteínas y otras EBPs produce activadores constitutivos que ya no dependen de efectores aromáticos o de fosforilación para activar transcripción (Grimm *et. al.*, 1995; Gu *et. al.*, 1994; Huala *et. al*, 1992; Lee *et. al*, 1994; Pérez-Martín & de Lorenzo, 1996; Ronson *et. al.*, 1988; Xiao *et. al.*, 1994). Esto ha llevado a suponer que el dominio Amino ejerce un efecto inhibitorio sobre el dominio Central de la proteína y que esta inhibición es suprimida por la interacción con el efector o por fosforilación. Apoyando esta idea de interacción intradominios se ha encontrado que mutaciones en el

dominio Central de NtrC producen formas constitutivas independientes de fosforilación ($NtrC^{S160F}$; Weglenski *et. al.*, 1989) y que supresoras intragénicas en el dominio Central de DmpR restauran parcialmente la actividad de proteínas alteradas por mutaciones en el dominio Amino terminal (Ching *et. al.*, 1996).

El dominio Amino no tiene una función conocida en las NifA de los Rhizobia y probablemente sea dispensable para la activación transcripcional. NifA de *R. trifoli* carece del dominio Amino y una mutante de NifA de *R. meliloti* que carece de este dominio no presenta un gran defecto en la actividad transcripcional (Beynon *et. al.*, 1988; Iismaa *et. al.*, 1989; Huala & Ausubel, 1989). Recientemente se caracterizó otra proteína activadora dependiente de promotores σ^{54} llamada PspF que controla un operón que participa en la sobrevivencia de *E. coli* en la fase estacionaria a pH alcalino, pero que también se induce bajo otras condiciones como son la infección por fagos filamentosos (Model *et. al.*, 1997). PspF es activa constitutivamente *in vivo* e *in vitro*, lo cual es consistente con su carencia del dominio Amino terminal (Jovanovic *et. al.*, 1996).

El dominio Amino de NifA de *K. pneumoniae*, *A. vinelandii* y *Azospirillum* ha sido relacionado con la inactivación por oxígeno mediante la interacción con otra proteína sensora de oxígeno y amonio, llamada NifL (Hill *et. al.*, 1981; Narberhaus *et. al.*, 1995; Blanco *et.al.*, 1993; Elmerich *et. al.*, 1995). En el dominio amino de VnfA y AnfA de *A. vinelandii* se localiza un motivo estructural (Cys-X-Cys-X4-Cys y Ser-X-Cys-X3- Cys respectivamente), que probablemente forme un sitio sensitivo a la oxidoreducción, mediante la unión de un metal como Fe, Vn o Mo (Henderson *et al.*, 1989; Joergen *et. al.*, 1989). La remoción del dominio Amino de VnfA y AnfA produce una proteína constitutiva independientemente de la presencia de sus cofactores metálicos (Austin & Lambert, 1994).

El dominio Central. Una secuencia de aminoácidos rica en glutamina generalmente separa el dominio Amino del dominio Central de las EBPs, formando un conector de aproximadamente 20 residuos de longitud ("Q-linker"; Wooton & Drummond, 1989). La hidrofilicidad de este conector y su relativa tolerancia a las substituciones, hacen pensar que su función es contribuir a la flexibilidad entre ambos dominios. El dominio Central contiene los determinantes de control positivo, incluida una firma estructural característica de las proteínas que hidrolizan ATP (Saraste *et. al.*, 1991). Una gran cantidad de mutaciones que afectan la actividad transcripcional de las EBPs han sido localizadas en este dominio El dominio Central de NifA y DctD de *Rhizobium meliloti* es suficiente para activar la transcripción por sí mismo cuando es sobreexpresado en *E. coli* (Huala & Ausubel, 1989; Huala *et. al.*, 1992). Debido a estas razones, se piensa que al menos otras dos propiedades de NifA deben estar localizadas en el dominio Central: los determinantes de oligomerización y los de reconocimiento con la RNA polimerasa.

El dominio Carboxilo. En el dominio Carboxilo se encuentra una estructura necesaria para la unión con el DNA, que probablemente constituya una HTH (Drummond *et. al.*, 1986). Mutaciones en la HTH de NtrC y NifA suprimen mutaciones en los sitios de unión en el DNA (Contreras & Drummond, 1988; Morett & Buck, 1988). Se ha demostrado que el dominio de unión con el DNA de NifA de *K. pneumoniae* puede ser separado de la función de activación y que en altas concentraciones es capaz de asociarse específicamente a él, y reprimir *in vivo* algunos promotores (Morett *et. al.*, 1988; Lee *et. al.*, 1994). No obastante, mutaciones en la hélice de reconocimiento de la HTH de NtrC alteran significativamente la función transcripcional y la unión al DNA (Contreras *et. al.*, 1987; Drummond *et. al.*, 1986; Shiau *et. al.*, 1993). Esto indica que, a diferencia de lo que ocurre en NifA, en NtrC parece existir una mayor dependencia de la unión con el DNA para efectuar transcripción. Es probable que el dominio Carboxilo de NtrC contenga determinantes secundarios de oligomerización por analogía con Fis (factor de estimulación

de la inversión), además de los principales localizados en el dominio Central (North *et. al.*, 1993).

En la secuencia de aminoácidos que forman la HTH se encuentran algunos residuos muy conservados entre las EBPs (Morett & Segovia, 1993). Se sabe que las glicinas conservadas son responsable de la vuelta para formar la segunda hélice (Aurora *et. al.*, 1994). Sin embargo, NifA de *R. meliloti* carece de esta glicina y no muestra una gran afinidad por las UAS del *nifH* de *K. pneumoniae*, cuando se analiza mediante protección al DMS *in vivo*. A pesar de esto NifA de *R. meliloti* activa la transcripción eficientemente. En el laboratorio del Dr. Enrique Morett se han obtenido variantes de NifA de *R. meliloti* con substituciones de la leucina nativa por glicina en esta posición, de tal forma que, como se esperaba, NifA obtuvo una HTH similar al consenso suficiente para obtener una mejor unión al DNA y una mayor actividad del promotor *nifH* de *K. pneumoniae* (Morett *et. al.*, 1997).

5.2 La fijación de nitrógeno. La capacidad de reducir nitrógeno atmósferico en amonio se encuentra interdispersa en distintos géneros bacterianos y en *Methanococcus voltae* de archaea (Souilard & Sibold, 1986). Entre las bacterias fijadoras de nitrógeno se encuentran tanto gram positivas como gramnegativas de vida libre como *Klebsiella pneumoniae*, bacterias que establecen asociaciones mutualistas con las plantas como *Azospirillum*, y bacterias simbióticas de la Familia Rhizobeacea que forman asociaciones intracelulares con las raíces de las leguminosas, como *B. japonicum* (Young, 1992). La familia Rhizobeacea, comprende una gran diversidad de bacterias del suelo que interactúan con leguminosas cultivadas (frijol, alfalfa, trébol) y con leguminosas tropicales (leucaena, mimosa) (Martínez-Romero, 1995; Caballero-Mellado & Martínez-Romero, 1996). *B. japonicum* es el simbionte asociado con la soya.

Gran parte del avance en el conocimiento de la genética y biología molecular de la fijación de nitrógeno ha sido logrado estudiando a *K. pneumoniae* (Merrick, 1992). La fijación de nitrógeno occurre únicamente bajo condiciones especiales en donde la bacteria se desarrolla: ausencia de oxígeno y limitación de amonio. Por esta razón no es sorprendente que las proteínas reguladoras de los genes de fijación de nitrógeno (*nif*) respondan a estas señales medio ambientales. En las bacterias simbióticas se han caracterizado mutantes incapaces de reducir nitrógeno. Se ha demostrado que estas mutantes están afectadas en genes que intervienen en la producción y síntesis de la nitrogenasa, de los que generalmente se encuentra un homólogo en *K. pneumoniae* o bien son genes particulares del procesos simbiótico mismo (llamados genes *fix*) (Fischer, 1994). Actualmente se conocen más de 17 genes *nif* organizados en 7 u 8 operones (Merrick *et. al.*, 1992). Los productos de los genes *nifHDK*, constituyen la nitrogenasa reductasa, la enzima que reduce el nitrógeno molecular a amonio, a expensas de un considerable gasto energético. En *K. pneumoniae* y en casi todas las demás bacterias diazotrópicas, estos genes se hayan juntos formando un solo operón y en ciertas especies está repetido. En *R. etli* el operón completo se encuentra duplicado y, además, tiene una copia idéntica y funcional de *nifH* (Quinto *et. al.*, 1982; Valderrama *et. al.*, 1996). Entre los genes *nif* restantes hay algunos que intervienen en la biosíntesis de la nitrogenasa, y otros cuya función no se conoce. El gene *nifA* que produce el regulador positivo de la transcripción de la mayoría de estos genes, es común a todas las bacterias fijadoras de nitrógeno. En *A. vinelandii*, tres isoformas de NifA que responden a la disponibilidad de Fe, Vn y Mo, controlan la expresión de los operones correspondientes a cada tipo de nitrogenasa (Bishop & Premakumar, 1992).

5.3. Regulación por oxígeno de NifA. El proceso de fijación de nitrógeno es energéticamente muy demandante y necesita realizarse en ausencia de oxígeno, debido a la fragilidad de la nitrogenasa a este elemento. El control por oxígeno de la fijación de nitrógeno ocurre a nivel del inicio de la transcripción y traduccionalmente afectando

principalmente a NifA. En las bacterias de vida libre como *K. pneumoniae* la expresión de *nifA* es regulada transcripcionalmente por amonio a tráves de NtrC, y postraduccionalmente por oxígeno mediante NifL, cuyo gene es cotranscrito con NifA (Hill *et. al.*, 1981). En las bacterias simbióticas el oxígeno es el principal regulador, dado que presuntamente el bacteroide (la forma diferenciada de la bacteria en los nódulos) tiene un suministro asegurado de amonio de parte de la planta NifA de *K. pneumoniae*, *Azotobacter* y *Azospirillium* es inactivada por una interacción con NifL, una flavoproteína encargada de sensar los niveles de oxígeno y nitrógeno. Se ha propuesto que en esta clase de NifA el dominio Amino forma parte del sensor molecular al oxígeno, el cual, a tráves de su interacción con NifL inhibe la función transcripcional del dominio. Evidencia reciente obtenida *in vitro* usando los distintos dominios de NifA y NifL, señala que el llamado dominio Carboxilo de NifL de *A. vinelandii* y *K. pneumoniae* interacciona directamente con el dominio Central de NifA inhibiendo la transcripción (Narberhaus et. al., 1993; Lee *et. al.*, 1994; Berger *et. al.*, 1995) pero se desconoce cual es la señal *in vivo*, que recibe NifL para efectuar esta acción. Se conoce que la presencia de fierro (Fe^{++}) es requerida para eliminar la inhibición que NifL ejerce sobre NifA de *K. pneumoniae,* puesto que cultivos aeróbicos en presencia del agente quelante o-nitrofenantrolina, inactivan NifL (Schmitz *et. al.*, 1996). Sin embargo NifA ni tampoco NifL parecen interaccionar con el metal, por lo que se ha hipotetizado que existe una tercera proteína desconocida que efectúa esta tarea.

En las Rhizobeaceas la regulación por oxígeno ocurre a dos niveles. El primero a nivel de la transcripción mediante un par de proteínas (FixL y FixJ) de la familia de dos componentes. FixL es un sensor transmembranal que en condiciones microaeróbicas fosforila FixJ, un activador transcripcional que promueve la expresión de los genes *nifA* y *fixK* (David *et. al.*, 1988; De Philip *et. al.*, 1990). FixL es una hemoproteína capaz de autofosforilación y fosfotransferencia a FixJ (Gilles-González *et. al.*, 1991; Agron *et. al.*, 1993). FixK es la vez un activador de su propio gene y de *nifA* (Herting *et. al.*, 1989). El

segundo nivel de regulación por oxígeno en NifA de las bacterias simbióticas occurre mediante la inactivación directa de la proteína. A diferencia de NifA de *K. pneumoniae* que permanece activo transcripcionalmente en condiciones aeróbicas en ausencia de NifL. NifA de *B. japonicum* es rápidamenente inactivado a nivel postraduccional sin necesitar de NifL o de una proteína parecida (Fischer & Hennecke, 1987). Una característica común en varias EBPs inactivables por oxígeno como NifA de *B. japonicum*, NifA de *Rhodobacter capsulatus* y HprR de *Herbaspirillum*, es la presencia de cuatro císteinas localizadas entre el domino Central y Carboxilo (Weber *et. al.*, 1985; Buikema *et. al.*, 1985; Grönger *et. al.*, 1987) Mutantes de NifA de *B. japonicum* con sustituciones en cualquiera de estas cisteínas por histidina o serinas, o alteraciones en el espaciamiento entre ellas no son activas transcripcionalmente (Fischer *et. al.*, 1988; Fischer *et. al.*., 1989). Por otra parte, cuando *B. japonicum* crece aneróbicamente en presencia de un agente químico secuestrador de cationes como o-nitrofenantrolina, NifA es incapaz de activar la transcripción (Fischer *et. al.*, 1988). Un cofactor metálico presuntamente coordinado por las císteinas es sensible a los niveles de oxígeno intracelular a tráves de su estado de oxidoreducción.

La primera consecuencia de la inactivación de NifA de *B. japonicum* por oxígeno es la pérdida de la unión con el DNA, aún cuando la proteína persiste durante una hora en cultivos aeróbicos (Morett *et. al.*, 1991). Esto hace suponer que el cambio de NifA de un estado activo al inactivo involucra alteraciones en la estructura oligomérica de la proteína, probablemente mediada a tráves del estado de oxidoreducción del metal coordinado por las cisteínas (Osuna *et. al.*, 1997). Una variante de NifA ($NifA^{M217I}$) de *R. meliloti*, con un cambio en un aminoácido cercano al llamado Walker A (el sitio de unión con los fosfatos del monucleótido), resulta ser tolerante al oxígeno (Krey *et. al.*, 1992). Basándose en predicciones de estructura secundaria, estos autores proponen que una alteración del sitio de unión al nucleótido puede ser responsable de este fenotipo. Recientemente en el laboratorio del Dr. Morett se han obtenido variantes de NifA de *B. japonicum* tolerantes al oxígeno a tráves de esquemas de mutagénesis al azar seguido de selección positiva (H.

Flores & E. Morett sin publicar). Estas mutantes han sido consideradas como superactivadoras por su capacidad de estimular la transcripción del promotor *nifH* de *K. pneumoniae* fusionado al gene de cloramfenicol (*cam*). Aún cuando estas mutantes no revelan ninguna diferencia significativa en cuanto a su activación transcripcional en condiciones anaeróbicas se ha demostrado que son activas en presencia de oxígeno. Las mutaciones de estas variantes de NifA se localizan en distintas regiones conservadas del dominio Central, pero no alteran el conector de cisteínas presuntamente involucrado en la inactivación por oxígeno (Fischer *et. al.*, 1988). Es probable que la tolerancia al oxígeno en estas mutantes de NifA esté relacionado con un cambio en la conformación oligomérica de la proteína y/o con un cambio en la afinidad por el ATP.

5.4. Patrones evolutivos de NifA. Cuando se ha comparado la secuencia de aminoácidos entre las EBP, se ha encontrado que los dominios Amino y Carboxilo no son tan conservados como el dominio Central. Este dominio tiene más del 40% de similitud entre las proteínas más alejadas. Esta observación aunada al hecho de que las EBPs exhiben una estructura modular, es sugerente de intercambio de dominios ("domain shufling") durante la evolución. Sin embargo, no se han obtenido pruebas de que esta clase de eventos hayan ocurrido en ésta familia de proteínas (Morett & Segovia, 1993; Pao & Saier, 1997). De hecho, los cladogramas que relacionan a cada uno de los miembros de las EBP son escencialmente iguales entre sí, independientemente de cual dominio se escoja para hacer el dendrograma, por lo que se ha inferido que estos dominios han venido evolucionando juntos (Morett & Segovia, 1993). Los cladogramas construidos a partir de las similitudes entre las EBPs reflejan con claridad la divergencia natural de las especie bacteriana a la que cada EBP pertenece. Las EBPs se agrupan en clases definidas y en varios casos, como los subgrupos de NifA y NtrC, estos son consistentes con la separación de las subdivisiones α y γ de las proteobacterias (Morett & Segovia, 1993).

De acuerdo con Morett & Segovia (1993), se pueden encontrar hasta siete regiones de aminoácidos altamente conservados en el dominio Central de las EBPs (Figura 12) Entre estas se encuentra la firma G-G-GK, también conocida como "p-loop", Kinasa I o Walker A "motif", que la identifica como una GTPasa (Saraste *et. al.*, 1993; Walker *et. al.*, 1990; Fry *et. al.*, 1992). Esta firma se encuentra localizada en la región C1, y ha sido determinado que mutaciones en cualquiera de estos aminoácidos conservados afectan severamente la transcripción y la hidrólisis de ATP (Austin *et. al.*, 1990; Weiss *et. al.*, 1992; Cannon & Buck, 1992) Poco se conoce de la función de las restantes regiones conservadas. aunque en estas, como ya apuntamos, se hayan mapeado mutaciones que afectan la activación de la transcripción .

Una observación que sirvió como punto de partida para el trabajo experiemental que se presenta aquí, fué expuesta por Morett & Segovia (1993). Ellos se percataron de que una parte de la región conservada C3, no está en algunas de las proteínas reguladoras comparadas, a pesar de su gran similitud en el dominio Central (más del 50%; ver Figura 1, artículo Capítulo 7). Tyr R de *E. coli* carece de la parte carboxilo de la región C3 y NtrC de *R. capsulatus* la ha perdido por completo (Foster-Harnett *et. al.*, 1994). TyrR no es un "verdadero" activador a distancia, ni depende de $E\sigma^{54}$. TyrR controla la expresión de los genes para la biosíntesis de aminoácidos aromáticos, reprimiendo la mayoría de los operones que controla y activando solo algunos, como el *mtr* que codifica una permeasa (Pittard & Davidson, 1991). De cualquier modo, en este último caso, TyrR se comporta como un activador de acción cercana (Lawley *et. al.*, 1995; Yang *et. al.*, 1996). Los sitios operadores en *mtr* se localizan exactamente arriba de la región -35 del gene, lo cual hace que TyrR probablemente establezca contactos con la subunidad α de la RNA pol-σ^{70}, como en los activadores de clase I. TyrR tiene la actividad de ATPasa característica de los miembros de la familia, acoplada con la formación de oligomeros de alto orden necesarios para una completa represión de ciertos operones como *aroF*, *aroL*, *tyrB* y *tyrP* (Argaet *et. al.*, 1994; Wilson *et. al.*, 1994; Cui *et. al.*, 1993).

Por otro lado se han aislado mutaciones en la región C3 de NtrC de *S. thypimurium*, que afectan severamente la activación transcripcional conservando intactas las propiedades de hidrólisis de ATP y de unión al DNA (Ver Figura 1, artículo, Capítulo 7; Weiss *et. al.*, 1992; North *et. al.*, 1996). Estas dos observaciones condujeron a la hipótesis de que la región C3 de NifA constituye el elemento de reconocimiento del complejo cerrado promotor -RNA pol-σ^{54}, probablemente a tráves de un contacto directo.

"Voy por tu cuerpo como por un bosque . . .

Capítulo 6

El mecanismo de activacion transcripcional a distancia en Bacteria

La activación de la transcripción por proteínas que actúan a distancia en bacterias es un ejemplo bien documentado del modelo del asa mencionado en el capítulo 4. Los potenciadores transcripcionales ("enhancers") bacterianos fueron descubiertos en 1986 (Reitzer & Magasanik, 1986; Buck *et. al.*, 1986) como regiones del DNA localizadas a más de 100 pares de bases del promotor esenciales para la activación de la transcripción. La remoción de secuencias entre -72 y -184 del promotor *nifH* de *K. pneumoniae* reduce significativamente la transcripción (Buck *et. al.*, 1986). Al igual que los potenciadores eucarióticos, los bacterianos pueden situarse hasta 2 o 3 kilobases hacia arriba del promotor en cualesquiera de las dos orientaciones, sin pérdida total de la función (10%; Ninfa *et. al.*, 1987; Reitzer & Magasanik, 1987; Reitzer *et. al.*, 1989). Por sí mismos los potenciadores son inertes transcripcionalmente y actúan concertadamente con la RNA polimerasa, a tráves de la proteína activadora. El análisis de la secuencia nucleotídica de más de 20 potenciadores que forman parte de la región regulatoria de varios genes dependientes de σ^{54}, ha mostrado la conservación de las bases TGT N_{10} ACA (Buck *et. al.*, 1986). Estas secuencias, denominadas UAS ("Upstream Activating Sequences"), son los blancos de interacción con las EBPs, y se encuentran en una sola copia o en conjuntos de dos más copias (Cannon *et. al.*, 1990). Pruebas físicas del contacto de la proteína activadora con estas secuencias se han obtenido para varias EBPs mediante experimentos de protección de las UAS a diferentes agentes químicos tanto *in vivo* como *in vitro* (Sasse-Dwight & Gralla, 1988; Morett & Buck , 1988).

6.1 El modelo del asa. Una característica fundamental de la activación a distancia es la necesidad de que las UAS y el promotor se encuentren en la misma cara del DNA. En los potenciadores bacterianos, variaciones integrales en la periodicidad de la doble hélice con respecto al promotor (medias vueltas o vueltas completas) no alteran significativamente la

función (Buck *et. al.*, 1987; Reitzer *et. al.*, 1989; Minchin *et. al.*, 1989). Por lo contrario, cuando la helicidad es modificada en fracciones de 1 a 4, la transcripción es afectada seriamente. Esto llevó a la suposición de que para activar la transcripción NifA interactúa con la RNA polimerasa-σ^{54} unida al promotor y como consecuencia se forma un asa de DNA entre el promotor y las UAS (Buck *et. al.*, 1987; Ptashne, 1988). Un modelo alternativo considera que la proteína reguladora se desliza desde los sitios de unión en el DNA hasta la tocar la RNA pol -σ^{54} unida establemente al promotor. Dos experimentos arguyen en contra de este modelo. Uno de ellos es que la presencia de un sitio de unión para el represor *lacI* entre el potenciador y el promotor *nifH* no tiene un gran efecto sobre la activación de la transcripción mediante NifA, como podría esperarse de un interferencia estérica en el deslizamiento (Buck *et. al.*, 1986). En segundo lugar cuando el potenciador y el promotor de *glnA* se colocan en dos diferentes plásmidos concatenados, NtrC es capaz de activar la transcripción, no así cuando los plásmidos están decatenados (Wedel *et. al.*, 1990). Por lo consiguiente, la idea de la formación de un asa ha resultado más congruente con la evidencia experimental.

Un experimento elegante a favor del modelo del asa proviene de la detección de intermediarios con esta topología mediante micrografía electrónica. Su *et. al.* (1991) visualizaron complejos entre NtrC y la RNA polimerasa-σ^{54} con un asa de DNA de tamaño correspondiente al del número de bases que separan el potenciador del promotor. Recientemente, mediante el uso de una novedosa técnica de microscopía de alta resolución ("Scanning Force Microscopy", SFM), ha sido posible reunir evidencia adicional que refuerza la idea de la formación del asa. Esta técnica ha sido usada exitosamente para obtener imágenes de muestras biológicas en aire o en líquido, sin usar ningún medio de contraste. Esto permite un mínimo de manipulaciones de la muestra, de tal manera que se han podido observar diversos estados de oligomerización de NtrC, y su interacción con la RNA pol-σ^{54} (Wymann *et. al.*, 1997; Rippe *et. al.* 1997).

En la figura 13 se representa el mecanismo de activación de la transcripción de acuerdo a las observaciones experimentales comentadas. La proteína activadora, en forma de dímero se encuentra en solución, pero puede acceder a los sitios de unión en el DNA , mientras que la RNA pol-σ^{54} se une al promotor para formar un complejo cerrado estable, en donde el molde de DNA se encuentra en uno de los lados de la holoenzima. De acuerdo con las observaciones de Wyman *et. al.* (1997), en las UAS se forma un complejo oligomérico que consiste probablemente de ocho subunidades. El siguiente intermediario detectado es un complejo entre el activador y la RNA pol-σ^{54}, y la formación de un asa entre las UAS y el promotor. Esto requiere del doblamiento de la secuencia de DNA intermedia. En algunos promotores dependientes de σ^{54} la presencia de la proteína promotora de recombinación IHF ("Integration Host Factor") es necesaria para una activación eficiente (Cannon *et. al.*, 1990; Hoover *et. al.*, 1990). En los promotores donde IHF no parece requerirse, se ha propuesto que tienen una curvatura intrínseca, que ayudaría a suplir la actividad de doblamiento de esta proteína (Carmona *et. al.*, 1997). Finalmente, la isomerización del complejo cerrado al complejo abierto ocurre cuando el activador contacta la RNA pol-σ^{54} e hidroliza las moléculas de ATP.

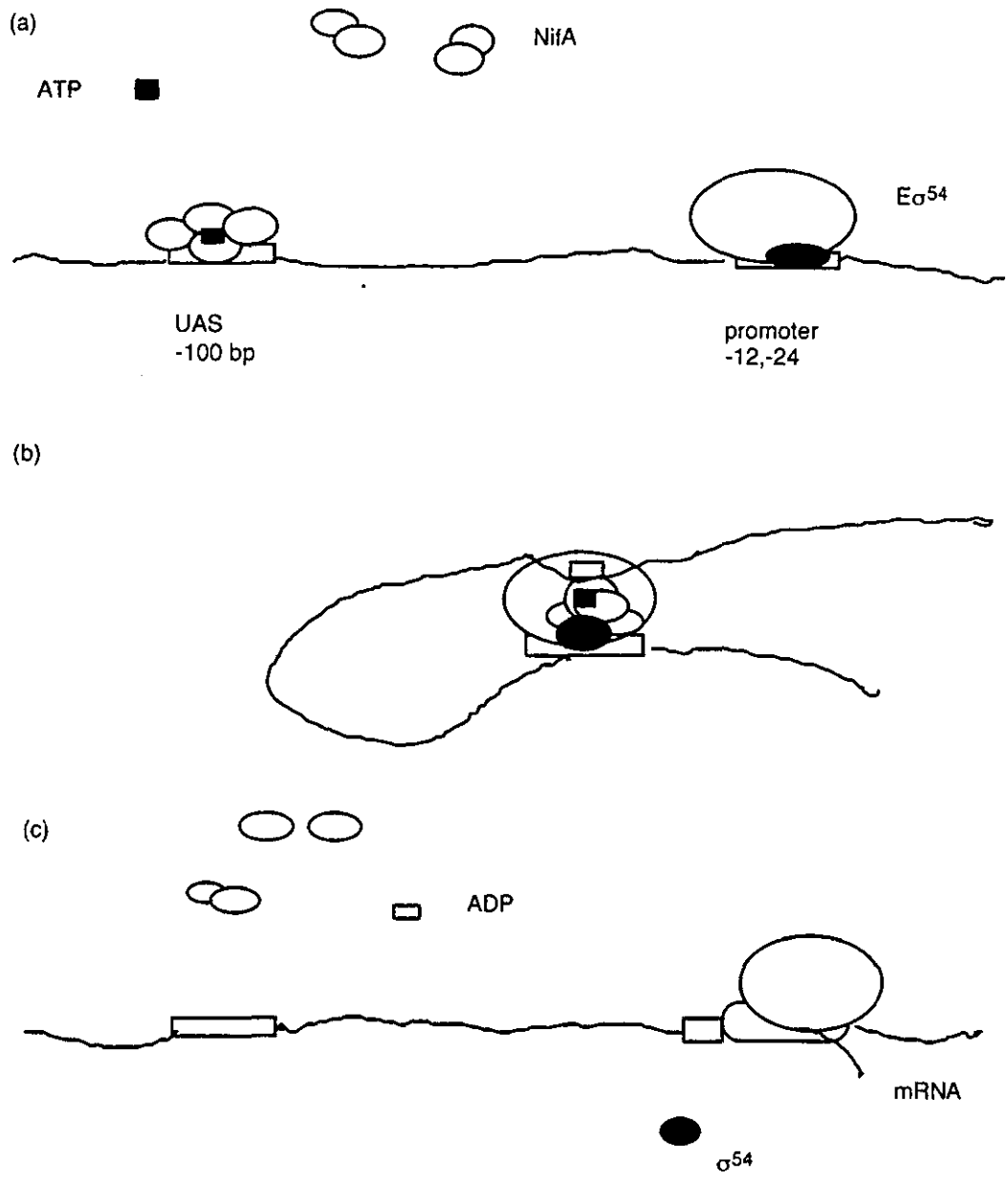

Figura 13. El mecanismo de activación mediante las EBPs.

6.2 Las función de los potenciadores. Una de las funciones probables de los potenciadores es brindar especificidad a la transcripción de los genes dependientes de σ^{54}, puesto que la RNA pol-σ^{54} no es específica por sí misma. Asegurar esto debe ser importante debido a la coexistencia de varios activadores dependientes de σ^{54} en una misma célula. Por ejemplo, en *E. coli* podrían coexistir 7 activadores de esta clase. De esta manera, los genes *nif* son activados únicamente por NifA y no por NtrC, en parte debido a

la especificidad de la unión con las UAS. Sin embargo, cuando uno de estos activadores es sobreexpresado puede activar promotores no-homólogos (Buck *et. al.*, 1986). Esto presuntamente se debe a efectos *in trans* del activador, es decir activa la transcripción desde el citoplasma, sin estar unido a las UAS.

Se ha pensado también que el potenciador sirve para incrementar la concentración local de la proteína activadora en la vecindad del promotor facilitando el encuentro con la RNA pol-σ^{54}. De acuerdo con esta propuesta, cuando el potenciador se aleja artificialmente del promotor, la actividad transcripcional disminuye (Buck *et.al.* 1986). Recientemente se ha propuesto que la función del potenciador es la de organizar un oligomero de alto orden, crucial para activar la transcripción, el cual aparentemente cuelga desde el potenciador hasta el promotor contactando la RNA pol-σ^{54} (Wyman *et. al.*, 1997). La activación promiscua podría ser evitada, incrementando la complejidad del sistema. Requerir un oligomero de alto orden para activar la transcripción puede ser uno de estos factores, haciendo improbable la formación de un oligomero a partir de sitios no especificados en orientación y composición.

6.3 Oligomerización del activador. Se ha demostrado que las EBPs interaccionan con diversas moléculas y cada interacción repercute en la organización cuaternaria de la proteína. Mediante el uso de análogos de ATP no hidrolizables se ha demostrado que la unión del mononucleótido provoca un cambio conformacional que capacita a estos activadores para ensamblarse en un oligomero y unirse a las UAS (Fárez-Vidal *et. al.*, 1996; Pérez-Martin & De Lorenzo, 1996). Estos cambios conformacionales se reflejan en alteraciones en la sensitividad de NtrC y XylR a las proteasas que afectan principalmente el dominio Central. Se ha propuesto que el ATP en XylR dispara un ciclo de multimerización y desmultimerización, que comienza con la unión del ATP a XylR preactivado por la unión con el efector *m*-xileno. Esto fomenta las interacciones proteína-proteína entre los dímeros de XylR unidos a las UAS. Una vez que el complejo

multiproteínico es ensamblado ocurre la hidrólisis de ATP y consecuentemente el inicio de la transcripción (Pérez-Martín & De Lorenzo, 1997)

La forma activa de las EBPs es probablemente un oligomero de alto orden, tal vez constituido por ocho subunidades (Wyman *et. al.*, 1997). Un solo dímero de NtrC fosforilado resulta insuficiente para activar la transcripción o de hidrolizar ATP (Porter *et. al.*, 1993; Austin & Dixon, 1992). Ambas funciones son estimuladas por la unión de NtrC a las UAS. NtrC no fosforilado permanece en un equilibrio monómero-dímero en solución (Klose *et. al.*, 1994) y se une débilmente a las dos UAS del promotor *glnAp2* (Weiss *et. al.*, 1992). Cuando NtrC es fosforilado por NtrB, ocurre un desplazamiento de este equilibrio hacia la formación de oligomeros y un incremento en la unión cooperativa de los dímeros con las UAS (Mettke *et. al.*, 1995; Chen & Reitzer, 1995). Para activar la transcripción a partir de una sola UAS es necesaria una alta concentración de NtrC. Supuestamente, altas concentraciones de NtrC estimulan las interacciones proteína-proteína. En experimentos de complementación de funciones, utilizando una mutante de NtrC incapaz de unirse a las UAS y una mutante de NtrC afectada en las propiedades de transcripción pero efieciente para unirse a las UAS, se alcanzan niveles de activación transcripcional similares a los obtenidos con la proteína silvestre (Wyman *et. al.*, 1977). A partir de estas observaciones se ha concluido que la función de las UAS es mediar la asocioación entre dímeros fosforilados y organizar una estructura oligomérica para contactar el complejo $E\sigma^{54}$-promotor. A diferencia de los activadores eucarióticos, en las que cada elemento está unido al DNA, algunos de los dímeros de NtrC son mantenidos unidos solamente por interacciones proteína-proteína. .

TyrR, es necesario recalcar que no es un activador σ^{54}, responde a varios efectores alostéricos como fenilalanina y tirosina. De la unión de estos efectores depende cuales operones van a ser reprimidos. Se ha reportado que el cambio en la asociación dímero-hexámero de TyrR require de la presencia de ATP (pero no la hidrólisis) y de amino

ácidos aromáticos como efectores (fenilalanina y tirosina). Dado que esta asociación ocurre a concentraciones de TyrR que probablemente sean las que se alcanzan en la célula, se cree que estas transiciones son importantes para la unión cooperativa de TyrR a sus sitios de unión en el DNA. Mutaciones en los sitios probables de unión de TyrR con sus efectores alostéricos conducen a fallas para reprimir algunos operones (Kwok *et. al.*, 1995).

6.4. La hidrólisis de ATP. Se ha demostrado ampliamente la necesidad de la hidrólisis de mononucleótidos trifosfatados para promover la isomerización del complejo cerrado formado por la RNA pol-σ^{54} hacia el complejo abierto, en el cual el DNA en la vecindad del sitio de inicio de la transcripción está parcialmente desnaturalizado. Como ya hemos mencionado, ésta actividad enzimática se localiza en el dominio Central de las EBPs. Tempranamente se notó que en este dominio había un sitio probable de unión de nucleótidos, debido a la presencia de un grupo de aminoácidos conservados, característicos del llamado "fosfato-binding loop", rico en glicinas, encontrado en las proteínas que unen e hidrolizan ATP y GTP (Drummond *et. al.* 1986; Saraste *et. al.*, 1990). Consistente con ésta observación, mutaciones varios de los residuos que conforman el asa afectan severamente tanto la hidrólisis de ATP como la transcripción (Austin *et. al.* 1991; Weiss *et. al.*, 1991; Cannon & Buck, 1992).

Se ha propuesto que NTRC y en general todas las EBPs son máquinas moleculares simples en el sentido usado por Alberts & Miaky-Lye (1992), es decir utilizan energía para realizar un trabajo molecular específico (Porter *et. al.*, 1993; Wedel & Kustu, 1995; Pérez-Martín & De Lorenzo, 1996). Sin embargo, en contraste con las proteínas G (Bourne *et. al.*, 1991; Webb 1992) en NtrC la hidrólisis de ATP no tiene un papel regulatorio. La unión del nucleótido trifosfatado en las proteínas G les confiere una conformación activa, mientras que la unión con el nucleótido difosfato produce conformaciones inactivas. En el caso de NtrC la unión de ADP o de análogos de ATP no

hidrolizables no es suficiente para activar la transcripción, la cual solamente ocurre cuando el oligomero hidroliza concertadamente más de dos moléculas de ATP. A pesar de que se ha observado que cada monómero o dímero son capaces de hidrolizar el ATP por sí mismo, solamente occurre cuando existen grandes concentraciones de proteína.. El rápido decaimiento de la activación transcripcional de NtrC, debido a la presencia de un competidor del ATP no hidrolizable como ATPγ S, indica un desacoplamiento energético en el oligomero que impide iniciar la transcripción.

De acuerdo con Wedel & Kustu (1995), la activación de la transcripción en los promotores σ^{54} tiene que superar una barrera cinética y otra termodinámica. La reacción de inicio de transcripción en estos promotores procede tan lentamente que prácticamente no existe síntesis de nuevos transcritos, en ausencia de hidrólisis de ATP. La apertura inicial del promotor ocurre concomitantemente con la hidrólisis de ATP. Sin embargo, los complejos abiertos recién formados, inmediatamente decaen en ausencia del activador y ATP, indicando que el equilibrio termodinámico se desplaza a expensas de la energía libre. Por analogía con lo que occure con la RNA pol II en donde el concurso de la hidrólisis de ATP no es es necesario para iniciar transcripción sobre promotores sintéticos heteroduplex, se pensó que NtrC pudiera funcionar de manera similar. Un experimento análogo demostró que aún usando promotores en plásmidos superenrollados o bien promotores heteroduplex, NtrC y ATP son aún absolutamente necesarios. Estas observaciones sugieren que NtrC no actúa directamente desnaturalizando el promotor, como lo hacen las helicasas. Por lo consiguiente, es factible que el mecanismo de activación mediante las EBPs involucre un cambio en la configuración de la RNA pol-σ^{54}. Estos posibles cambios conformacionales suceden concomitantes con la apertura del complejo abierto, como se infiere del incremento en el patrón de protección de las bases en la región del promotor hacia el sitio de inicio de la transcripción. El inicio de la transcripción con otras holoenzimas (RNA pol-σ^{32} y RNA pol-σ^{70}), a pesar de no ser dependiente de una actividad enzimática como la hidrólisis de ATP, también parece estar limitado por un

cambio cambio conformacional más que por la desnaturalización local del promotor (Buc, 1989; Leirmo & Record, 1989). Es probable que el inicio de la transcripción mediante la RNA pol-σ^{54} involucre cambios conformacionales comunes con estas holoenzimas. Usando proteólisis limitada, recientemente se han inferido cambios conformacionales en regiones de σ^{54} expuestas en la superficie, durante la formación de los complejos transcripcionales de iniciación (Casaz & Buck, 1997). Estos cambios son el resultado de la interacción con la RNA pol ^{C} y se encuentran asociados a la transición del complejo cerrado al abierto.

6.5. El dominio Central de las EBPs es una GTPasa. Recientemente investigadores de nuestro laboratorio, propusieron un modelo estructural del dominio Central de NifA basándose en la predición de la estructura secundaria y en algoritmos que permiten el reconocimiento de plegamientos comunes (Osuna *et. al.*, 1997; Figura 14). Ellos sugieren que el dominio Central es una GTPasa. La familia de las GTPasas incluye proteínas muy diversas que dependen estrechamente de la hidrólisis de mononucleótidos para su función (Bourne *et. al.*, 19). Entre ellas se encuentran EFTu, el oncogene *ras* y las proteínas G. El modelo propuesto, basándose con mucho, en la estructura ya conocida de EFTu determinada por cristalografía, consiste de un polipéptido que adopta una topología α/β. Dado que no se conoce la función precisa de la mayoría de las regiones conservadas del dominio Central propuestas inicialmente por Morett & Segovia, en el modelo estructural del dominio Central se hizó una asignación tentativa, basándose en analogías con miembros de las GTPasas y con ejemplos de mutaciones caracterizadas en miembros de la familia de las EBP (Morett & Segovia, 1993; Osuna *et. al.*, 1997). El motivo estructural conocido como "p-loop" o Walker A, se localiza en un asa entre la β1 y la αA en la región C1 (Morett & Segovia, 1997). Esta región es la que sostiene el trifosfato, tal vez dejándolo en la posición apropiada para un ataque nucleofílico de parte de los residuos catalíticos. La presencia de residuos de glutamato en la región C1 y C4 sugiere que uno de estos puede ser un residuo catalítico. Los autores proponen que el llamado Walker B este

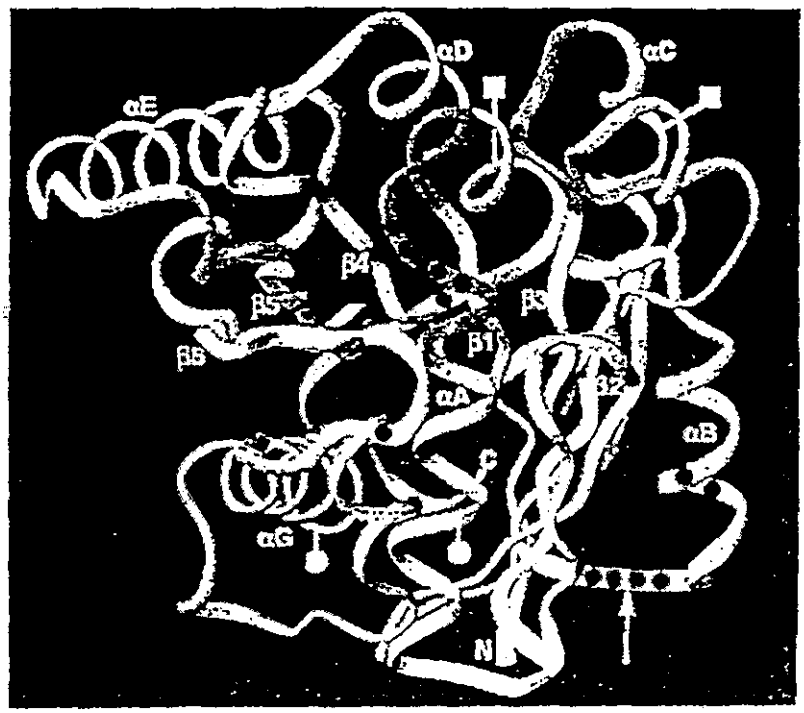

Figura 14. El modelo estructural del dominio Central de NifA (Osuna, *et. al.*, 1997).

contenido precisamente en la region C4 (β3) en el cual se localiza un aspartato altamente conservado. Este residuo puede estar involucrado en la coordinación del ion magnesio, observado en las GTPasas. La base nucleotídica por otra parte es sostenida por aminoácidos de las regiones C6 y C7 (αD, αF, β5 y β7), en donde una asa entre αD y β5 es rica en residuos de arginina. Una serie de residuos invariantes se encuentra en el asa que conecta la αF con la αG. De acuerdo con su conservación, las mutaciones localizadas en ésta asa resultan en proteínas incapaces de activar la transcripción y de hidrolizar ATP.

Dado que la oligomerización de NifA es un paso crucial para la activación de la transcripción, es importante conocer en donde se localizan los determinantes de oloigomerización de la proteína. Del análisis de mutantes de deleción de TyrR (Cui & Sommerville, 1993), se ha sugerido que las regiones que comprenden las αE y αG, del dominio Central son esenciales para la oligomerización de la proteína. Como mencionamos anteriormente NifA es una molécula que se inactiva en la presencia del oxígeno atmosférico. Esto se lleva a cabo a tráves de un ion metálico coordinado por residuos de císteina entre el dominio Central y el dominio Carboxilo (Fisher *et. al.*, 198). Es posible que la inactivación de la proteína involucre un cambio en el estado oligomérico, dado que mutaciones en la región probable de oligomerización es homóloga de NtrC perjudican la actividad de la proteína

Las similitudes entre las GTPasas y las EBPs, indican también que la región de control positivo de las EBPs pudiera ser equivalente a la región conocida como efectora/transducta de señales encontrada en las GTPasas. La región efectora en EFTu sirve como blanco de interacciones proteína-proteína con el ribosoma, y supuestamente inducen la hidrólisis de ATP y la consecuente liberación del aminoacil-tRNA acarreado por EfTu. En las proteínas G y en *ras* estas regiones son conocidas como Switch I y Switch II, las cuales sirven para interaccionar con las GAPs ("Guanosin Activating Proteins") u otras proteínas efectoras. De varias de estas proteínas se conoce la estructura cristalográfica en

asociación con análogos de GTP y en un caso asociada con la molécula efectora (RBD en complejo con Rap1A). En *ras* la molécula efectora interacciona con el Switch I provocando una alteracion conformacional que consiste en el cambio de orientación de la cadena lateral de varios aminoácidos (Wittinghofer & Nassar, 1996). Esto tiene como consecuencia un incremento en la tasa de hidrólisis de ATP y la liberación del complejo proteiníco. Se ha pensado que la región conservada C3 de las EBPs es la más probable candidata a tener las propiedades de efector/transductor. En el modelo estructural del dominio Central se predice que la región C3 está constituida por la αb (ESELFGHEK) hacia el amino terminal, mientras que para la region carboxilo (GAFTGA) no se predice con exactitud que tipo de estructura secundaria adopta. Es conveniente recalcar que la región carboxilo no se encuentra en la proteína TyrR, ni en otras proteínas con un dominio Central homólogo a NifA, pero que no son activadores dependientes de σ^{54}. También es oportuno señalar que en la región C3 se han mapeado mutaciones que alteran severamente la actividad de distintas EBPs, pero que no afectan la hidrólisis de ATP. Por estas razones se ha propuesto que la región C3 representa la región de control positivo de NifA y de otros activadores de la familia. Un estudio genético de la subregion GAFTGA se presenta en el siguiente capítulo.

"voy por tu frente como por la luna . . .

Capítulo 7

Contribución original: La región de control positivo de NifA.

Como hemos venido sugiriendo desde capítulos precedentes, uno de los requisitos para activar la transcripción mediante cualquier proteína activadora es establecer contactos con la RNA polimerasa. Se ha propuesto que los activadores transcripcionales de levadura y los activadores bacterianos de acción cercana funcionan mediante un mecanismo de acción similar, a tráves de reclutar a la RNA polimerasa (Ptashne & Gann, 1997). Esto quiere decir que ejercen su acción auxiliando a la RNA polimerasa a situarse apropiadamente sobre el promotor. La unión del activador en la región proximal del promotor es absolutamente necesaria en este caso, para poner en cercanía al activador con la RNA polimerasa. Diversas clases de interacciones se originan de este encuentro. Como discutimos en el capítulo referente a la activación cercana, las regiones de activación en CAP y CI están constituidas por segmentos pequeños de aminoácidos localizados en la superficie de la proteína (Ishihama, 1992). Cualquier cambio en estas secuencias produce alteraciones en la capacidad para activar la transcripción, pero no tiene ningún efecto sobre otras propiedades de la proteína como la unión con el DNA o el doblamiento de este. En esta clase de activación parecen ser suficientes las interacciones proteína-proteína para estimular la transcripción, aún cuando provengan de proteínas sin propiedades de activación *per se* (Dove *et. al.*, 1997)

En contraste, los activadores bacterianos de acción a distancia, actúan sobre una RNA polimerasa unida establemente al promotor, y requieren de la hidrólisis de ATP para incrementar la tasa de isomerización del complejo cerrado al complejo abierto. En altas concentraciones de la proteína activadora no es necesario que ésta se una al DNA, ejerciendo su acción directamente desde el citoplasma. Necesariamente el activador transcripcional debe contactar la RNA polimerasa, ¿cuál es entonces la región del activador que tiene esta propiedad?. Debido a que una parte de la región C3 está ausente

en TyrR, una proteína no dependiente de σ^{54}, Morett y Segovia (1993) propusieron que la región C3 es el sitio de reconocimiento para la RNA polimerasa. Alteraciones en varios de los aminoácidos que forman región C3 de NtrC y DcTD se comportan como mutantes en el control positivo (North *et. al.*, 1996; Wang *et. al.*, 1997). En esta contribución examinamos exhaustivamente los requisitos funcionales y estructurales de la región C3 de NifA para activar transcripción. Para ello utilizamos un enfoque genético. Concluimos que esta es una zona de control positivo en donde la presencia de un residuo aromático y un grupo hidroxilo son estrictamente requeridos para activar la transcripción.

molecular microbiology

Bob Simons, Editor
University of California, Los Angeles
Department of Microbiology and
Molecular Genetics
1602 Molecular Sciences Building
Los Angeles, CA 90095-1489 USA

Phone 1 310 206-9736
FAX 1 310 206-9737
Email molmicro@lifesci.ucla.edu

December 30, 1997

Dr. Enrique Morett
Departamento de Reconocimiento
Molecular y Bioestructura
Instituto de Biotecnologia
Universidad Nacional Autonoma de Mexico
P.O. Box 510-3
Cuernavaca, Morelos 62250
MÉXICO

Dear Dr. Morett:

I am pleased to inform you that your manuscript entitled "In vivo studies on the positive control function of NifA: A conserved hydrophobic.. " (MS No UCLA-97-1195R2) has been accepted for publication in *Molecular Microbiology*, and forwarded to the publishers. The date of acceptance is the date of receipt at the production office.

You should receive proofs about 4 weeks after receipt of your paper by the production office (provided that there are no production problems, e.g. missing figures or tables, in which case you will be contacted). It is particularly important that you read the proofing instructions, which will be sent with your proofs. You should return your proofs within 3 days of receipt (if necessary, arrange for someone to correct them in your absence). You must return your proofs even if you have no corrections (or inform the production office of this by email).

The manuscript that I have accepted is the final version: only errors introduced by the copy-editor or typesetter can be corrected on the proofs. Any requests for changes to the data or the text will cause the paper to be referred back to me, and possibly to the referees, and will delay publication of your article. Publication will normally be within 5 weeks of receipt of your corrected proofs by the production office. For questions or concerns about publication, please contact Helen MacDonald at:

Helen MacDonald
Molecular Microbiology
Prepress Projects
4 Dower Place (off Hay Street)
Perth PH1 5HU
SCOTLAND

FAX: +44-1738-622467
mole@prepress-projects.co.uk

Thank you for submitting your paper to *Molecular Microbiology*. I hope you will submit future papers to the journal.

Sincerely yours,

Bob Simons
BS/lc

***In vivo* Studies on the Positive Control Function of NifA: A Conserved Hydrophobic Amino Acid Patch at the Central Domain Involved in Transcriptional Activation.**

Víctor González, Leticia Olvera, Xavier Soberón and Enrique Morett*

Departamento de Reconocimiento Molecular y Bioestructura, Instituto de Biotecnología, Universidad Nacional Autonóma de México. A.P 510-3. Cuernavaca, Morelos 62250. México.

Running title: an aminoacid patch of NifA involved in transcriptional activation.

* To whom correspondence should be addressed: Telephone number (53)(73) 291665; Fax number: (53)(73) 172388; E-mail address: emorett@ibt.unam.mx

Summary

The Eubacterial Enhancer-Binding Proteins activate transcription by binding to distant sites and, simultaneously, contacting the RNA-polymerase σ^{54}-promoter complex ($E\sigma^{54}$). The positive control function is located at the Central domain of these proteins, but it is not know which specific region has the determinants for the interaction with $E\sigma^{54}$. Here we present genetic evidence that a small region of hydrophopic amino acids, previously denominated C3, at the Central domain of *Bradyrhizobium japonicum* NifA, is involved in positive control. We obtained twenty-six missense mutants along this conserved region. Among these, only strains expressing the $NifA^{F307Y}$ and $NifA^{A310S}$ mutant proteins retained some of the transcriptional activity (<20%), while those carrying $NifA^{E298D}$ and $NifA^{T308S}$ had very low but detectable activity (<1.0 %). The rest of the NifA mutants did not induce any measurable transcriptional activity. When expressed in the presence of wild type NifA, the great majority of the mutants displayed a dominant phenotype, suggesting that their oligomerization determinants were not altered. *In vivo* dimethyl-sulfate footprinting experiments for a subset of the NifA mutants showed that they were still able to bind specifically to DNA. Analysis of intragenic supressors highlight the functional relevance of a hydroxyl group at position 308 to activate transcription.

Key Words: σ^{54}, NifA, positive control, intragenic suppression, GTPase.

Introduction.

The positive control NifA protein activates the expression of several *nif* and *fix* genes in response to the cellular oxygen and /or nitrogen status (Fischer, 1994). NifA belongs to a large family of Enhancer-Binding Proteins (EBPs) which activate transcription from distant sites and use the sigma 54 (σ^{54})-containing holoenzyme form of the RNA polymerase ($E\sigma^{54}$) (Kustu *et. al.*, 1989; Morett and Segovia, 1993; Shingler, 1996). σ^{54} is an evolutionarily unique class of initiation factor that binds to the -12/-24 promoters ($TGGCACN_5TTGCA$) in the absence of core RNA polymerase (Buck and Cannon, 1992; Lonneto *et. al.*, 1992; Cannon *et. al.*, 1997). $E\sigma^{54}$ is unable to carry out transcription unless an EBP, such as NifA, is present (Popham *et. al.*, 1989; Merrick, 1993; Lee *et. al.*, 1994). In the current model for transcriptional activation, an oligomeric form of the EBP, bound far from the promoter at the Upstream Activating Sequences (UAS), contacts the $E\sigma^{54}$-closed promoter complex, so it is converted to an open, transcriptionally active form, in a process shown to require ATP hydrolysis (Weiss *et. al.*, 1991; Berger *et. al.*, 1995) . A DNA loop corresponding to the intervening segment between the promoter and the UAS has been observed in NtrC of *Salmonella thypimurium*, the prototype member of the EBP family (Su *et. al.*, 1990). Recently, it has been shown that DctD, an EBP of *Rhizobium meliloti*, can be photocrosslinked to the β subunit of RNA polymerase and to σ^{54}, so it is likely that they interact (Lee and Hoover, 1995; Wang and Hoover, 1997). Nevertheless, the details of such molecular recognition events are not known.

NifA, as well as the other EBPs, is a three-domain protein (Drummond *et. al.*, 1986; Morett and Segovia, 1993; Berger *et. al.*, 1995), whose Central domain of about 220 amino acids is sufficient, by itself, to activate transcription (Huala and Ausubel,1989). The Amino terminal domain function is unknown in NifA, while the Carboxy terminal domain contains a helix-turn-helix motif that is required for binding to the UAS (Morett

and Buck, 1988; Morett *et. al.*, 1988). An interdomain linker with four conserved cysteine residues between the Central domain and the Carboxy terminal domain of *Bradyrhyzobium japonicum* NifA, has been implicated in sensing oxygen, presumably through the coordination of a metal cofactor (Fischer *et. al.*, 1988; Morett *et. al.*, 1991).

Seven highly conserved regions (C1 to C7) have been distinguished at the Central domain of the EBPs (Morett and Segovia, 1993). C1 has the typical sequence known as Walker A motif, present in a large number of proteins that bind and hydrolyze mononucleotides (Walker *et. al.*, 1982). Accordingly, mutations at this region affect ATP hydrolysis (Austin *et. al.*, 1991; Weiss *et. al.*, 1991). On the other hand, the precise function of the rest of the conserved regions at the Central domain remains to be established. It is noteworthy that two regulators, TyrR of *Escherichia coli* and NtrC of *Rhodobacter capsulatus* contain a domain with high sequence similarity to the Central domain of the EBPs, but they are not σ^{54}-dependent activators (Pittard and Davidson, 1991; Morett and Segovia, 1993; Pao and Saier, 1995; Foster-Harnett *et. al.*, 1994). Notably, the TyrR protein also hydrolyzes ATP in association with its physiological activity, despite having a small deletion that removed part of the C3 region, the highly conserved GAFTGA amino acid patch (Cui *et. al.*, 1993; Morett and Segovia, 1993; Fig. 1 a). Based on this observation and on the fact that mutations at the C3 region of NtrC affect the positive control function, but they do not interfere with ATP hydrolysis or DNA-binding ($NtrC^{repressor}$, Weiss *et. al.*, 1991; North *et. al.*, 1996), we proposed that this region may be involved in recognition of the $E\sigma^{54}$-promoter-complex (Morett and Segovia, 1993). Furthermore, we recently modeled the structure of the Central domain of NifA and showed that it might adopt a topology related to the nucleotide-binding fold found in members of the GTPase superfamily, such as the bacterial elongation factor EFTu and the Ras oncogene (Osuna *et. al.*, 1997; Bourne *et. al.*, 1990). All members of the GTPase superfamily have sites for recognition their effectors (Switches I and II; Hilgenfield, 1995). In the proposed model, the C3 region lies at the equivalent position as

switch I of EFTu, such that it is a likely candidate for protein-protein interactions and / or signal transduction (Osuna *et. al.*, 1997).

To examine the role of the GAFTGA amino acid patch of the C3 region of *B. japonicum* NifA in positive control we made an extensive mutagenesis analysis of it and conducted *in vivo* experiments to characterize the mutants. Our results showed that the majority of the amino acid substitutions severely affected the transcriptional activity of NifA, but had little effect on DNA-binding and on the ability to interact with wild type NifA *in trans*. These results support the view that region C3 of NifA is critical for positive control.

Results

Construction of NifA missense mutants. The positive control determinants of the EBPs reside in the Central domain. We have previously proposed that the highly conserved C3 region may function in protein-protein contacts with $E\sigma^{54}$, required for isomerization of the closed promoter complex to the active, open form (Morett & Segovia, 1993; Osuna *et. al.*, 1997). Here we focused on the study of the carboxy terminal moeity (GAFTGA) of that region, absent in TyrR but highly conserved in the rest of the EBPs (Fig. 1 a). We followed a saturation, oligonucleotide directed, mutagenesis scheme (Merino *et. al.*, 1992) to substitute the phe 307 and thr 308 residues. Ala 306 and ala 310 residues were reeplaced only by amino acids of similar size. Since glycines are normally critical for the structure of proteins they were not substituted. In six independent experiments, a 218 base pairs (bp) DNA fragment with the desired single, double or triple nucleotide substitutions was synthesized by PCR and used to replace the corresponding DNA fragment of the wild type *nifA* gene, cloned in plasmid pRJ7511 (Fischer and Hennecke, 1989). After transformation into *E. coli* JM101 several clones were identified by DNA sequencing. The entire 218 bp synthetic DNA fragments inserted in these clones were sequenced and y the *nifA* genes that had single amino acid substitutions were selected for further analysis. Figure 1 b and Table 1 show the resulting twenty six single amino acid substitutions and their corresponding codon changes. All but two of these mutants were at the desired positions. In $NifA^{E298D}$ and $NifA^{K304N}$, unexpected third base substitutions (A—>T) appeared, while the target for mutagenesis remained without base alterations.

Even thought we did not get all codon changes expected from the mutagenesis strategy, we obtained a well represented sample of biochemically and structurally different amino acids, allowing us to examine the sequence requirements and informational content of the function for the C3 region.

***In vivo* activities of the NifA mutants.** The transcriptional activation properties of each NifA mutant was studied by measuring the expression of the *Klebsiella pnumoniae nifH (KpnifH*p) promoter fused to the *lacZ* gene. β- galactosidase activities of cell extracts grown under anaerobic conditions are shown in Table 2. A complete loss of the activation property of NifA was the most common effect, regardless of the biochemical nature of the amino acid introduced at the indicated positions. Only the $NifA^{A310S}$ and $NifA^{F307Y}$ mutants retained significant transcriptional activation, while the conservative amino acid substitutions of mutants $NifA^{T308S}$ and $NifA^{E298D}$ resulted in residual, but detectable, β-galactosidase activity (Table 2). It is remarkable that these conservative amino acid substitutions displayed such a drastic effect. It seems likely that the hydroxyl group at position 308 could play a critical role in transcriptional activation. Alternatively, the position of the methyl group of threonine could be the relevant feature for activity. To discriminate between such possibilities we substituted thr 308 for the isosteric amino acid valine by directed mutagenesis, with a procedure that also provided substitutions by cysteine and glycine. Neither of these mutants ($NifA^{T308V}$, $NifA^{T308C}$ and $NifA^{T308G}$) showed detectable transcriptional activity (Table 2), further strengthening the notion for a requirement of the hydroxyl group for transcriptional activation. In a similar fashion we also examined the informational content of the phe 307 residue, by substituting it with the conservative tyrosine residue. This mutant, $NifA^{F307Y}$, activated transcription at about 30% of the wild type level (Table 2). Thus, among the 9 substitutions examined at this position only tyrosine was accepted, indicating that the aromatic ring is a stringent requirement for activation.

The severe loss of transcriptional activation found among the NifA mutants described here demonstrates that the amino acid sequence at the C3 region is very tightly constrained, as almost every alteration lead to a great defect in transcriptional activation

***In vivo* expression of the altered NifA proteins.** To rule out that the NifA mutants failed to activate transcription due to a defect in their expression or stability, we performed immunodetection experiments using antibodies raised against a synthetic oligopeptide corresponding to a part of the Central domain of *B. japonicum* NifA (Morett *et. al.*, 1991). As shown in Figure 2, a polypeptide of 63 Kd was immunodetected in crude cell extracts of anaerobically grown cultures carrying the *nifA* wild type gene as well as the *nifA* alleles tested. The majority of the NifA mutant proteins described here were present in very similar amounts to that of wild type NifA; only the $NifA^{F307A}$, $NifA^{F307R}$, $NifA^{T308R}$, $NifA^{T308P}$, $NifA^{T308L}$ and $NifA^{T310S}$ proteins seemed to be present at slightly lower levels (Fig. 2, lanes 5, 10, 15-18 and 21). We infer from these results that the strong defect in activation displayed by these NifA mutants is not due to a failure in gene expression or to instability of the NifA mutant proteins, but to an alteration in the determinants for transcriptional activation, DNA-binding or oligomerization.

DNA-binding properties of the NifA mutants. To investigate if the mutations at the C3 region of NifA had an effect on DNA-binding, we performed *in vivo* dimethyl sulfate (DMS) footprinting experiments for a subset of the NifA mutants here described. We probed the occupancy of the *KpnifH* UAS and the divergent *nifJ* UAS at the upstream region of the *nifH* promoter by either wild type or mutant NifA proteins. The DMS patterns obtained from cells in the presence of the wild-type NifA showed protection of guanines -123, -126 and -127 and hypermethylation of guanines -125 and -133 at the bottom strand of the *nifH* UAS, as previously reported (Morett and Buck, 1988; compare lane 1 and 2 of Fig. 3 a). A DMS protection profile equivalent to that of the wild type NifA was observed upon the presence of the NifA mutants altered at the thr 308 residue ($NifA^{T308A}$, $NifA^{T308R}$, $NifA^{T308L}$, $NifA^{T308P}$, $NifA^{T308S}$), although it was not as strong as the protection observed with the wild type NifA (Figure 3 a, lanes 3-7). To further asess the occupancy of the *nifH* UAS by the thr 308 NifA mutants, we also analyzed the DMS protection pattern of the top strand. When wild type NifA was present, protection of

guanine -136 was detected (compare lanes 8 and 9 of Fig. 3 b). In addition, the guanine -182 of the divergent *nifJ* UAS, which is contained in the same DNA fragment, was protected by the wild type protein. A similar methylation protection was also detected with all the thr 308 mutants, although it was not as strong as with the wild type protein (Fig. 3 b, lanes 10-14). In particular, the NifA^{T308R} produced a weak DMS protection pattern at the top strand, although the *nifH* UAS bottom strand appeared well protected by this mutant protein.

The DNA-binding activity of one mutant for each amino acid position altered, except for ala 306 was also analyzed (NifA^{E298D}, NifA^{K304N}, NifA^{F307E} and NifA^{A310D}). These mutants produced a clear protection of guanine -136 and guanine -182 at the top strand in a similar fashion to that of the thr 308 mutants (Fig. 3 c, 15-18).

These experiments indicated that the DNA-binding activity of NifA was not strongly perturbed by mutations affecting the C3 region. Thus, it cannot account for the major defects on the transcriptional activity displayed by these mutants.

Functional inactivation of wild type NifA by altered NifA proteins. The active form of NifA is likely to be an oligomer, as it has been shown for its homologues NtrC, TyrR and XylR proteins (Porter *et. al.*, 1993; Wilson *et. al.*, 1994; Pérez Martín and de Lorenzo, 1996). If the inactive NifA mutant proteins were not altered in their oligomerization properties we could reasonably expect that in merodiploid experiments they would exert an inhibitory effect over the wild type NifA (Herskowitz, 1987). We addressed this proposal performing a dominance test for every *nifA* allele, based on a compatible multiple plasmid system, designed such that the amount of the NifA mutant polypeptides was greater than that of the wild type products (see experimental procedures). We tested the dominance effect on the weak *KpnifH*p fused to *lacZ* (Table 2, column 4). In this experiment the most common effect was dominance, with a complete

inhibition of the NifA wild type activity by the presence of the NifA mutant polypeptides tested. In contrast the NifA mutants carrying the most conservative amino acid substitution showed either partial dominance (up to 1 000 Miller units; $NifA^{E296D}$ and $NifA^{T308S}$) or behaved as recessive alleles ($NifA^{A310S}$), which had no effect on transcription.

These results suggest that mutantions at the C3 region did not alter the oligomerization properties of NifA. Alternatively the dominance effect could be due to exclusion of the NifA wild type protein from the UAS by a higher concentration of the mutant protein, resulting in the inability of the wild type protein to activate the weak *KpnifH*p *in trans.* To exclude this possibility, dominance tests with the *R. meliloti nifH* promoter (*RmnifH*p) devoid of the UAS were carried out. This is an strong promoter that can be partially activated in the absence of UAS (Morett and Buck, 1989; Buck and Cannon, 1989). Five selected NifA mutants displaying recessive, partially dominant or dominant phenotypes were tested for their effect over wild type NifA. Figure 4 shows that these mutant proteins had the same transdominance phenotype in both the *RmnifH*p with (Fig. 4 a) or without the UAS (Fig. 4 b). As expected, the activation levels in the absence of UAS were lower than in the promoter with an intact regulatory region. $NifA^{T308A}$ and $NifA^{T308V}$, two of the dominant mutants on the *KpnifH*p also lead to a marked reduction in *RmnifH*p'-*lacZ* expression by the wild type NifA. In agreement with the dominance test performed with the *KpnifH*p'-*lacZ*, the NifA mutants with the conservative amino acid substitutions $NifA^{T308S}$ and $NifA^{E298D}$ exerted a moderate inhibitory effect, while the $NifA^{A310S}$ displayed a recessive phenotype. Thus, we conclude that the inhibitory effect of the C3 region mutants is not due to a competition for binding to the UAS, but to a functional inactivation of wild type NifA by its association with different mutant polypeptides.

Intragenic supressors of NifA^{E298D} and NifA^{T308S}. To assess structural and functional relationships between specific amino acids of the C3 region with other regions of the Central domain of NifA, we searched for intragenic supressors of *nifA*^{E298D} and *nifA*^{T308S} alleles. These mutants are suitable candidates to be suppressed, because of the conservative nature of their substitutions and their residual transcriptional activity. A PCR-mutagenic procedure was implemented to obtain a random collection of Central domain mutants on DNA templates containing either mutant allele (see experimental procedures). Libraries of ~ 15 000 variants of *nifA*^{T308S} and ~ 60 000 of *nifA*^{E298D} were obtained. After a library amplification step (see experimental procedures) and transformation into *E. coli* JM101 cells containing the reporter *KpnifHp'-lacZ* fusion, several putative suppressor clones were chosen based on their blue colony appearance on X-Gal plates opposite to the slightly blue appearance displayed by the parental *nifA*^{E298D} and *nifA*^{T308S} alleles. The entire synthetic fragment, coming from PCR amplification, of each suppressor clone was sequenced and β-galactosidase assays were performed to measure their transcriptional activity. Table 3 describes one second-site suppressor mutation for NifA^{T308S} (NifA$^{T308S/V364I}$), and two for NifA^{E298D} (NifA$^{E298D/S356G}$ and NifA$^{E298D/N357S}$). The NifA$^{T308S/V364I}$ had 16-fold more β-galactosidase activity than the parental single mutant NifA^{T308S}, but it was still lower than wild type. The double mutants, NifA$^{E298D/N357S}$ and NifA$^{E298D/S356G}$, raised the transcriptional activity to about half of the wild type level.

To determine the allele specificity of these suppressor mutants we exchanged the secondary mutations between the two mutant *nifA* alleles and the wild type (Table 3). The secondary substitutions did not have detrimental effects on transcription by themselves in a wild type context (NifA^{N357S}, NifA^{S356G} and NifA^{V364I}). Interestingly, NifA^{V364I} was unable to suppress the transcriptional defect of NifA^{E298D} as it did in the context of the NifA^{T308S} mutation (compare NifA$^{T308S/V364I}$ with NifA$^{E298D/V364I}$). Thus, it displayed allele-specific suppression. NifA^{S356G} and NifA^{N357G} were able to suppress the

missense NifA^{T308S} partially (NifA$^{T308S/S356}$) or totally (NifA$^{T308S/N357S}$), thus acting as general suppressors.

As described before, the low levels of transcriptional activity of the NifA^{T308S} underscore the relevance of the hydroxyl group, considering that NifA^{T308V} was unable to activate transcription (Table 2). The importance of the hydroxyl group is supported by the fact that combining NifA^{T308V}, with each of the three suppressor mutations was never able to alleviate the transcriptional defect, as they did for the NifA^{T308S} (Table 3).

Discussion

The mechanism of transcriptional activation by the EBPs involves interactions between the regulatory protein and $E\sigma^{54}$ bound at the promoter (Su *et. al.*, 1990; Lee and Hoover, 1995). Ample genetic and biochemical evidence indicate that the Central domain of the EBPs contains the structural and functional determinants for positive control (Wei *et. al.*, 1981; McNeil *et. al.*, 1982; Austin *et. al.*, 1991; Huala and Ausubel, 1992; North *et. al.*, 1996; Pérez-Martín and de Lorenzo, 1996). Several observations point to the conserved C3 region of the Central domain as an element likely involved in the recognition of the $E\sigma^{54}$-promoter complex (Morett and Segovia, 1993; North *et. al.*, 1996; Wang and Hoover, 1997). Here, we present genetic evidence to support that the C3 region of NifA is a positive control region which does not participate in either oligomerization or DNA-binding.

Functional and structural constraints of the GAFTGA amino acid patch: a positive control determinant of NifA. By means of genetic analysis we showed that the carboxy motif of the C3 region (GAFTGA) has strict structural requirements for positive control and that it is essentially refractory to amino acid changes. Our initial mutagenesis approach relied on obtaining an extensive number of amino acid substitutions per position, independently of selective schemes, so that the sequence requirements of the GAFTGA region for transcriptional activation could be analyzed. Despite this motif is highly conserved, it was intriguing to find that almost all amino acid substitutions led to a great defect in transcriptional activity. So, it is likely that the GAFTGA patch conform a highly structured region where each amino acid has an strict commitment for the positive control function. Only four NifA mutants ($NifA^{F307Y}$, $NifA^{A310S}$, $NifA^{T308S}$, $NifA^{E298D}$) retaining some transcriptional activity were obtained in this work. These amino acids substitutions are structurally very similar to the wild type amino acids in volume and/or biochemical properties. Notably, there are some proteins along the evolutionary

conserved EBP family that precisely contain these amino acids (Morett and Segovia, 1993; Pao and Saier, 1995; Osuna *et. al.*, 1997).

At position 308, only NifA^{T308S} had detectable transcriptional activity, while all of the other eight mutants at the same position did not show activity. It is noteworthy that NifA^{T308V}, which carries the isosteric valine, was not tolerated, perhaps due to the lack of the hydroxyl group. Thus, we propose that the thr 308 residue plays a crucial role in transcriptional activation, probably by making a highly specific hydrogen bond with other residue of NifA or with another protein of the Eσ^{54} promoter-complex.

An aromatic residue at position 307 seems also to be relevant for transcription, as the phenylalanine to tyrosine substitution at this position was the only active mutant obtained. This is not unexpected since some EBPs have tyrosine in a context in which other residues of the C3 region remain conserved (Morett and Segovia, 1993).

From the amino acid substituitions at position 310 only NifA^{A310S} showed partial transcriptional activation. Interestingly, when this position was substituted with other small amino acid such as glycine (NifA^{A310G}), or amino acids with a side chain with potential for hydrogen-bonding or negative charge (NifA^{A310N} and NifA^{A310D}), the resultant proteins were unable to activate transcription. NtrC and DctD mutations at the equivalent position (NtrC^{A220T} and DctD^{A225T}) are also deleterious (Weiss *et. al.*, 1991; North *et. al.*, 1996; Wang and Hoover, 1997).

Missense mutations at the C3 region have been mapped on NtrC mutants isolated by *in vivo* screening of *Salmonella* glutamine auxotrophs (Wei *et. al.*, 1981; Weiss *et. al.*, 1991; North *et. al.*, 1996; Figure 1 b). These NtrC mutants, denominated NtrC^{repressor}, absolutely lack transcriptional activity, as it was shown by their further *in vitro* characterization (North *et. al.*, 1996). Recently, Wang and Hoover (1997) selected DctD

mutants unable of activate transcription, some of them also mapped at the C3 region (Figure 1 b). It seems likely that the NifA mutants described here, and these NtrC and DctD mutants, are affected in a similar way.

Suppressor genetic analysis was used here as a mean to identify specific structural and functional interactions between the C3 region and other regions of the Central domain of NifA. Restoration of the activity may result from the reestablishment of the protein-protein contacts, or from an alteration of the structural organization of the mutant protein leading to some activity (Travormina *et. al.*, 1996). Here, we describe two neighboring intragenic mutations able to raise the activity of the two conservative substitutions made at the C3 region ($NifA^{E298D}$ and $NifA^{T308S}$). Of the suppressor mutations found, $NifA^{V364I}$ behaved as allele-specific since it was unable to relieve the defect of the $NifA^{E298D}$. However, this mutant did not impair transcriptional activation in a wild type context. This effect is suggestive of a direct contact between the residue 364 and the C3 region (Table 3). It is noteworthy, that neither of the second site mutations were able to correct the defects of $NifA^{T308V}$, so this fact provides further indication of the requirement of an hydroxyl group at position 308 to activate transcription.

DNA-binding and oligomerization properties were maintained in the NifA mutants. There is broad evidence that the DNA-binding determinants of NifA reside at the Carboxy terminal domain (Morett and Buck, 1988; Morett *et. al.*, 1989), thus, we anticipated that the mutants at the C3 region of the Central domain would not have strong DNA-binding defects. Figure 3 shows that the NifA mutants produced a DMS protection pattern similar to the wild type, indicating that they maintained the ability to bind to the UAS. Unexpectedly, the DMS protection patterns produced by the mutants were not as strong as the one observed with the wild type protein. Since high levels of protein had been proven to be necessary for a total UAS occupation (Morett and Buck, 1988), we examined the expression levels of the mutant proteins in total cell extracts. None of the

mutants was expressed significantly below the level of the wild type protein. Thus, mutations at the C3 region could have an indirect effect on the DNA-binding but not as strong to eliminate them. Interestingly, North *et. al.*, (1996) noted that among the $NtrC^{repressor}$ mutants, $NtrC^{G219K}$ has an increased affinity for the DNA. These observations suggest that the C3 region could be located at the surface of NifA, close to the helix-turn-helix, or facing to the DNA.

There is some experimental evidence suggesting that the oligomerization determinants of the EBPs lie at the Central domain (Cui and Somerville, 1993; North *et. al.*, 1995; Osuna *et. al.*, 1997; Wyman *et. al.*, 1997). The transdominance experiments in merodiplod cells, aimed at analyzing the oligomerization properties of the NifA mutants, showed that the majority of them displayed a dominant phenotype over the wild type. The mutants that retained residual transcriptional activity were partially dominant. Conversely, a mutant with appreciable activity ($NifA^{T308S}$) was recessive. We interpret the distinct dominant phenotypes as a result of the formation of heteromeric complexes composed by inactive and active subunits. The defects, in the case of the NifA mutants that retained residual activity ($NifA^{E298D}$ and $NifA^{T308S}$), could be partially suppressed by the fully active wild type protein. That is, the defects of these mutants are so moderate that they can be corrected by wild type polypeptides in the oligomer. On the contrary, the heteromers formed by totally inactive NifA proteins were inactive, despite carrying wild type NifA subunits, indicating that these mutations severely impaired the function of the protein and therefore precluded any kind of functional complementation. Our transdominance test carried out with the strong *RmnifHp* showed a similar dominance phenotype in the presence and in the absence of UAS. This indicates that the reduction of the activation by the wild type NifA was not due to its exclusion from the UAS by the transcriptionally inactive but DNA-binding proficient, NifA mutants.

The above results suggest that the initial events in the enhancer-mediated trancriptional activation, such as DNA-binding and multimerization steps, were not strongly altered in the NifA C3-mutants. Hence, their defect might be in subsequent events towards the formation of an active $E\sigma^{54}$-promoter complex.

The C3 region in the structural model of the Central domain. The conserved amino acid regions showed by Morett and Segovia (1993), may have a structural correlation with the predicted fold for the Central domain (Osuna *et.al.*, 1997). Based on this model the C3 region, presumably located at the surface of the protein, might play a role equivalent to the switch I region of some GTPases, like EFTu (Osuna, *et. al.*, 1997). The switch I region is both a contact for guanosine activating proteins (GAPs) and a region subject to allosteric transitions upon the binding and hydrolysis of GTP (Wittinghofer and Nassar, 1996; Stouten *et. al.*, 1993). The C3 region was modeled as an α helix (ESELFGHEK) followed by segment (GAFTGA) whose secondary structure was not well predicted (Osuna *et. al.*, 1997). In the fold model the C3 region neither interact directly with the mononucleotide nor contain residues for catalysis or binding of ATP.

Several groups have shown that NifA proteins from different eubacterial species are very difficult to purify and assay in *in vitro* systems (Lee *et. al.*, 1994; Austin *et. al.*, 1994). So, this difficulty precluded us to directly address the ATPase properties of the NifA derivatives. However, $NtrC^{repressor}$ and DctD mutants with alterations at the C3 region, retain the ATPase activity (Weiss *et. al.*, 1991; North *et. al.*, 1996; Wang and Hoover, 1997). Thus, it is likely that the NifA-C3 mutants here described might conserved ATPase activity.

Our genetic suppressor analysis suggest that the secondary mutations restored the defects of $NifA^{T308S}$ and $NifA^{E298D}$ in different ways (Table 3). For example, $NifA^{V364I}$ displayed allele-specific suppression, thus, it is tempting to speculate that region C3

might encode two different functions. Taking the GTPases as a working model for the study of the function of NifA, such functions could be related to protein-protein interactions with the $E\sigma^{54}$-closed complex (GAFTGA) and to the commitment of ATP hydrolysis to an allosteric change (ESELFGHEK), required for open complex formation. The $TyrR^{E274Q}$ mutant, that fails to repress transcription of operons which depend on higher order oligomers in response to tyrosine, is located at the second glutamate of the ESELFGHEK subregion (Kwok *et. al.*, 1995). It has been proposed that this mutant cannot couple the binding and/or hydrolysis of ATP to an allosteric transition needed to bind its co-effector tyrosine and multimerize in the appropriate way (Kwok *et. al.*, 1995). Although some DctD mutants at the GAFTGA subregion could be photocrosslinked to β and σ^{54} subunit (Lee & Hoover, 1995; Wang and Hoover, 1997), we favor the view that the ESELFGHEK subregion present in all the EBPs is a common allosteric element of a putative switch, while the GAFTGA subregion, that is not present in TyrR, may be a recognition element for the $E\sigma^{54}$ -promoter complex.

Experimental procedures

Bacterial strains and plasmids. *Escherichia coli* JM101 [D (*lac-pro* AB) *sup* E *thi* /F'*lac* Iq DM15 *tra* D36 *pro* AB+] (Messing *et. al.*, 1981) was used for all *in vivo* activation and footprinting experiments. Plasmid pRJ7511 (Tc^r) which carries the *B. japonicum nifA* gene constitutively expressed from the *cat'* promoter (Fischer *et. al.*, 1989), or mutant derivatives of it with an altered *nifA* gene, were used through the work. Plasmid pRT22 (Cm^r) that carries the *nifH* promoter fused to the *lacZ* gene assays (Tuli and Merrick, 1988) was used as reporter for the *in vivo* transcriptional activation tests and for the DMS footprinting. A low copy number plasmid called pM1-K was constructed (H. Flores and E. Morett, unpublished) by subcloning the entire *cat'nifA* transcriptional fusion from pRJ7511 into the pWSK130 plasmid (Wang and Kushner, 1991). Plasmid pCU101 containing the *RmnifHp'-lacZ* fusion was constructed by subcloning a corresponding PCR DNA fragment from pMB210 (Better *et. al.* , 1985). pSU003 is a derivative of pCU101 that lacks of the UAS (S. Dávila and E. Morett, unpublished results).

Growth conditions. *E. coli* cultures were normally grown at 37°C in Luria broth (LB) (Miller, 1972). For the activation assays and the DMS footprinting experiments, cells were cultured anaerobically at 30°C in a modified NFDM medium prepared as described (Morett *et. al.* 1991). Antibiotics used for plasmid selection were: chloramphenicol (25 μg/ml), tetracyline (10 μg/ml), and kanamycin (25 μg/ml). X-Gal (40 μg/ml).

Directed mutagenesis of the *nifA* gene. A direct oligonucleotide mutagenesis method based on PCR was employed (Merino, *et. al.* 1992). It consists on the synthesis, by PCR, of a mutagenic megaprimer using an oligonucleotide altered at the chosen codon (see below),and one Initiator oligonucleotide (5'-GTATGCCGGATCCTGCTCGCGAGCGC-3') that primes far from the mutagenesis target, towards the 5' end of *nifA*.. After purification, the mutagenic megaprimer was extended towards the 3' of the *nifA* gene using a

Termination oligonucleotide (5'-CGCAGTCTGCAGAAGCTTCATGCAAAGTCACT-3'). Specific rescue of the synthetic mutant DNA fragments was accomplished adding to the extension PCR-product a short oligonucleotide (5'-GTATCGCGGATCCT-3') complementary to a tail of non-homologous sequence to *nifA* incorporated by the initiator oligonucleotide. The synthetic DNA fragment obtained (718 bp) carries part of the Central domain of *nifA* with alterations at predetermined codons. Six mutagenic oligonucleotides complementary to the bottom strand were used in this work: For the ala 306 codon we synthesized a mix of oligonucleotides (RRC mutagenesis; see Table 1) as follows: 5'-GACAGCGGTAAAAGYYACCTTTCTCATGG-3'). For the phe 307 codon, we utilized a SNN mix to obtain the other nineteen amino acid replacements at this position plus a stop codon: 5'GACAGCACCGGTSNNGGCACCACCTTTCTC-3'. Similarly, for the thr 308 codon we employed a SNN mix: 5'-GACAGCACCSNNAAAGGCACC-3'. Additionally, a GMM mix (KKG mutagenesis in Table 1) was synthesized to obtain the $NifA^{T308C}$, $NifA^{T308G}$ and $NifA^{T308V}$ variants: 5'-GACAGCACCGMMAAAGGCACC-3'. For the alanine 310 codon a AYY mix (RRT mutagenesis; see Table 1): 5'GCGGGCGCTGACAYYACCGGTAAAGGC-3' was used. Finally, for the directed phe 307 to tyr substitution a single oligonucleotide with the appropriated codon (TAC) was synthesized: 5'-GACAGCACCGGTGTAGGCTTCCTC-3'.

PCR-mutagenesis. We followed the Leung's PCR-random mutagenesis procedure to generate a collection of Central domain variants on the $nifA^{E298D}$ and the $nifA^{T308S}$ alleles as templates (Leung *et. al.* 1989). We used 10 pmol of Initiator and Termination oligonucleotides in a reaction mix containing 0.050 µg of DNA template, 2.5 u of Taq polymerase (Gibco-BRL), 1mM each dGTP, dCTP and dTTP and 0.2 mM dATP, 10 nM β-mercaptoethanol , 10 µl dimethyl sulfoxide, 6.1 mM $MgCl_2$ and 0.5 mM $MnCl_2$. The reactions were performed in an Stratagene Robocycler for 25 cycles.

Cloning of *nifA* mutant genes. In six independent directed-mutagenesis experiments, one for each mutagenic oligonucleotide mix, and two PCR-mutagenesis experiments we synthesized a DNA fragment that contained part of the Central domain of *nifA* (718 bp) and carried the desired codon changes, as described above. The DNA fragment from directed-mutagenesis was restricted with *KpnI* and *BglII* endonucleases to produce a 218 bp internal fragment used to substitute for the equivalent DNA fragment at the complete *nifA* wild type gene on plasmid pRJ7511. The mutants were identified by DNA sequencing after electrotransformation of the ligation products into *E. coli* JM101 cells. The nucleotide sequence for each individual *nifA* mutant covered the entire 218 bp synthetic fragment and only single codon changes at the predetermined positions were selected for their further analysis. In some cases (Table 1), the entire PCR fragment of 718 bp was cloned into the pMOSt vector (Amersham Life Sciences, Cleveland, OH). These *nifA* mutants were also identified by sequencing and the internal 218 bp fragment subcloned into pRJ7511.

A similar replacement procedure was followed to clone the Central domain DNA fragments from the PCR-mutagenesis experiments but using the unique *KpnI* and *EcoRI* restriction sites. The libraries obtained after electrophoration of JM101 cells, were amplified extracting plasmid DNA of the whole collection of recombinant clones and transforming it into JM101 cells. The transformant clones were mixed and subject to plasmid purification. This enriched *nifA* mutant DNA was used to transform JM101 cells containing the reporter plasmid pRT22, to select for colony appearance the reversion phenotypes on X-Gal plates.

DNA sequencing. All the DNA sequencing was done with Taq polymerase and cycling reactions according to the manufacturer's instructions (Gibco-BRL). The oligonucleotide used for sequencing (5'-CAGAACTCGCAGCAACTTCG-3'), complementary to a DNA region located proximal to the *Bgl II* endonuclease site, was ^{32}P end-labeled at the

5'terminus with T4 polynucleotide kinase (Amersham Life Science, Cleveland, OH). The sequence products were separated in 7% polyacrilamide, 8M urea gels. Autoradiograms were obtained exposing the dry gel 12 to 24 hours with Kodak X-Omat films or in a Phosphorimager apparatus.

***In vivo* activation assays.** We measured the β-galactosidase activity from anaerobic cultures of JM101 cells carrying the reporter plasmid pRT22 (*KpnifHp'-lacZ*) in the presence of wild type NifA or the mutant protein variants, according to the Miller's method (Miller, 1972), with the following modifications: 1 ml of cells from an anaerobic culture ($D.O._{600}$ = O.6- 0.7) was collected by centrifugation (10,000 g) and resuspended in 0.9 ml of Z buffer (0.06 M $Na_2HPO_4.7H_20$; O.4 M $NaH_2P0_4.7H_20$; .001 M KCl; 0.01 M $MgSO_4.7H_20$; 0.05 M β-mercaptoethanol). The cells were permeabilized by adding 0.1% SDS and chloroform. ONPG (16 mg/ml) was used to develop the reaction at 30°C, and stopped by addition of 0.5 ml of 0.2 M Na_2CO_3.

DMS footprinting. Footprinting of the *K. pneumoniae nifH* and *nifJ* UAS top and bottom strand were performed as described by Morett *et.al.* (1991) with the following modifications: anaerobic cultures of *E. coli* JM101 cells which carried the compatible plasmids pRT22 and the pRJ7511 (or the derivatives of it with the different *nifA* mutant genes), were collected at an $D.0._{600}$ of 0.6-0.8. A freshly prepared 2% (v/v) DMS solution was added to each culture at a final concentration of 0.1%. After 1 min, the cultures were centrifuged at 7,000 rpm for 5 min at 4°C. The cell pellet was washed twice with 50 ml of ice-cold saline phosphate solution (150 mM NaCl, 40 mM K_2HPO4, 22 mM KH_2PO_4, pH 7.2). The methylated plasmid was obtained by an alkaline lysis standard procedure and purified by phenol-chloroform extraction and ethanol precipitation. Plasmid DNA was cleaved at the methylated positions by incubation at 90°C with 10% piperidine. The piperidine excess was eliminated by repeated lyophilization of the samples in a DNA Speed-Vac 110 (Savant Co.). Finally, the plasmid DNA was resuspended in 50 ml of distillate water.

Primer extension reactions were performed using a synthetic oligonucleotide 5′ labeled with [γ^{32} ATP] complementary to the bottom (5′-ACCGTAAATAGCCGATTG-3′) or top (5′-AAGAGATTGCGTGGAATAAGA-3′) strands of the *K. pneumoniae* upstream promoter region in pRT22. The primer extension reactions (2.5 mM $MgCl_2$, 2.5 mM each dNTP, 0.5 μg methylated DNA and 5 units of Taq polymerase in a final volume of 50 μl), were performed with Taq Polymerase (Gibco-BRL) by 10 rounds of denaturation-annealing-extension cycles (92°, 55° and 72°C) in a Stratagene Robocycler (Stratagene Inc). The extension products were concentrated adding 1 ml of n-butanol and washed with 70% ethanol. 5 of 10 μl of concentrated extension products were separated by electrophoresis in a 7% polyacrilamide gel containing 8 M urea. Autoradiograms were obtained using a Phosphorimager and analyzed by the ImageO2 program, or by exposing the gel on Kodak X-Omat films at -70°C during three to six days.

Western blot techniques. For the detection of the NifA wild type or mutant proteins we used a polyclonal antibody kindly donated by Dr. H. Hennecke, and described in Morett *et. al.*, (1991). This antibody recognize an epitope at the Central domain of NifA. Crude cell extracts were prepared by concentrating a 3 ml of anaerobic culture in 200 ml of sample buffer (10% glycerol; 5% 2-mercaptoethanol; 2.3% SDS; tris-HCl 0.5 M pH 6.8 and 0.05% bromophenol blue) and boiling for 10 minutes at 90°C. 15 μl samples were subject to standard electrophoresis and electroblotting to Hybond-C extra nitrocellulose membranes (Amersham Life Sciences, Cleveland, OH). Alkaline phospatase-conjugated goat anti-rabbit antibodies (Bio-Rad Laboratories, Richmond, CA) served as secondary antibodies, and they were used according to the manufacturers instructions.

Oligonucleotide Synthesis. All the oligonucleotides used in this work were synthethized by the phosphoramidite method on an Applied Biosystem 381A DNA synthesizer at the Unidad de Síntesis de Macromoléculas, IBT-UNAM.

Acknowledgments. We wish to thank to M.A. Cevallos, M. Soberón, H. Flores and E. Horjales for helpful discussions. Thanks are also extended to M. Olvera, K. Nevárez, F. Romero, E. Flores, E. López, P. Gaytán and F. Sánchez for technical assistance. This work was supported by grants from PADEP-UNAM 030301 to V.G., DGAPA-UNAM IN208894 to E.M. and X.S. and European Union CI1*-CT94-0060 to EM. V.G. was recipient of a CONACyT scholarship.

Figure legends

Figure 1. (a) Amino acid comparison of the well characterized region C1, or Walker A motif and region C3 of four well-known EBPs. The underlined amino acids at region C1 shows the sequence characteristic of proteins that bind and hydrolyze ATP. Detailed aligments with more than 40 activator have been published elsewhere (Morett and Segovia, 1993; Pao and Saier *et. al.*., 1995; Osuna *et. al.*, 1997). Despite the conservation of region C1 and a part of region C3, *E. coli* TyrR lacks of the GAFTGA region. Numbers on the top indicate the amino acid position of the conserved regions C1 and C3 according to *B. japonicum* NifA sequence.

(b) Amino acid substitutions at region C3 of NifA obtained in this work (top) and for comparison *S. thyphimurium* NtrC, *E. coli* TyrR and *R. meliloti* DctD mutants characterized previously at the same region (bottom; Weiss *et. al.*, 1992; Kwok *et. al.*, 1995; North *et. al.*, 1996; Wang and Hoover, 1997).

Figure 2. Immunodetection of wild type and mutant NifA proteins. NifA proteins in soluble cell extracts of anaerobic cultures carrying the wild type or mutant *nifA* alelles were detected by western immunoblotting as described in experimental procedures. Lane 0, molecular weight markers; lane1, NifA wild type; lane 2, No NifA; lanes 3 to 22, NifA mutants (below the sticky bar) as indicated for each position (over the sticky bar) in a single letter amino acid code.

Figure 3. *In vivo* occupancy of the *K. pneumoniae* *nifH* and *nifJ* UAS bottom and top strand by wild type NifA and for a subset of NifA mutants. Footprinting assays were conducted as described in experimental procedures. (a) and (b) DMS protection patterns of the *nifH* UAS bottom and top strand respectively for NifA mutants at position 308 (T308 under the bar). Lanes 1 and 8, NifA wild type; lanes 2 and 9, no NifA; lanes 3 to 14, indicate the different NifA mutants tested showing the amino acid substitution which

they carried, as shown by a single letter code below the bar. (c) DMS protecction patterns of the top strand of *nifH* and *nif J* UAS for a subset of NifA mutants. The amino acid altered is indicated under the bar and the substitution analyzed is indicated bellow each bar (lane 15 to 18). Open circles shows the protected guanines. Hypermethylated guanines are showed by black circles.

Figure 4. Transdominance effect of NifA mutants on the NifA wild type protein in the *R. meliloti nifH* promotor with UAS (a) and without UAS (b). Transcriptional activity is indicated as Miller Units plus standard desviation of three independent assays. The last columns in (a) and (b) indicated the background β-galactosidase activity attained by the plasmids pCU101 and pSU003 respectively in absence of NifA. In (b) this background value was substracted for each column.

Figure 1

(a)

	region C1	region C3
	252 268	294 313
NifA *B. japonicum*	.. VLLRGESGTGKELVAKA	ESELFGHEKGAFTGAVSARK
NtrC *S. typhimurium*	. VLINGESGTGKEVVATL	ESELFGHEKGAFTGANTIRQ
TyrR *E. coli*	LLITGDTGTGKDLFAYA	ESELFGH .....APEGKK.
DctD *R. meliloti*	VLVAGETGSGKEVVATL	ESELFGHERGAFTGAVKKIR

(b)

											Y	G			
											N	C			
											E	V			
											H	S			
											R	R			
											P	P		S	
											I	M		G	
										N	L	N		N	
NifA mutants			D					N		D	A	A		D	
NifA wild type	E	S	E	L	F	G	H	E	K	G	A	F	T	G	A
NtrC repressor		F	Q							V				K	T
TyrR			Q												
DctD			V							D	V	L	I		T
											D				

Figure 2

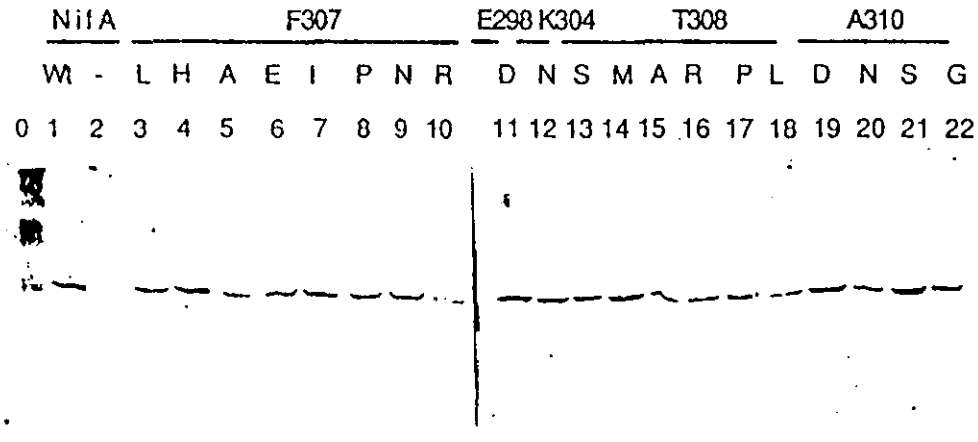
NifA
F307
E298 K304
T308
A310
W t - L H A E I P N R D N S M A R P L D N S G
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

figure 3

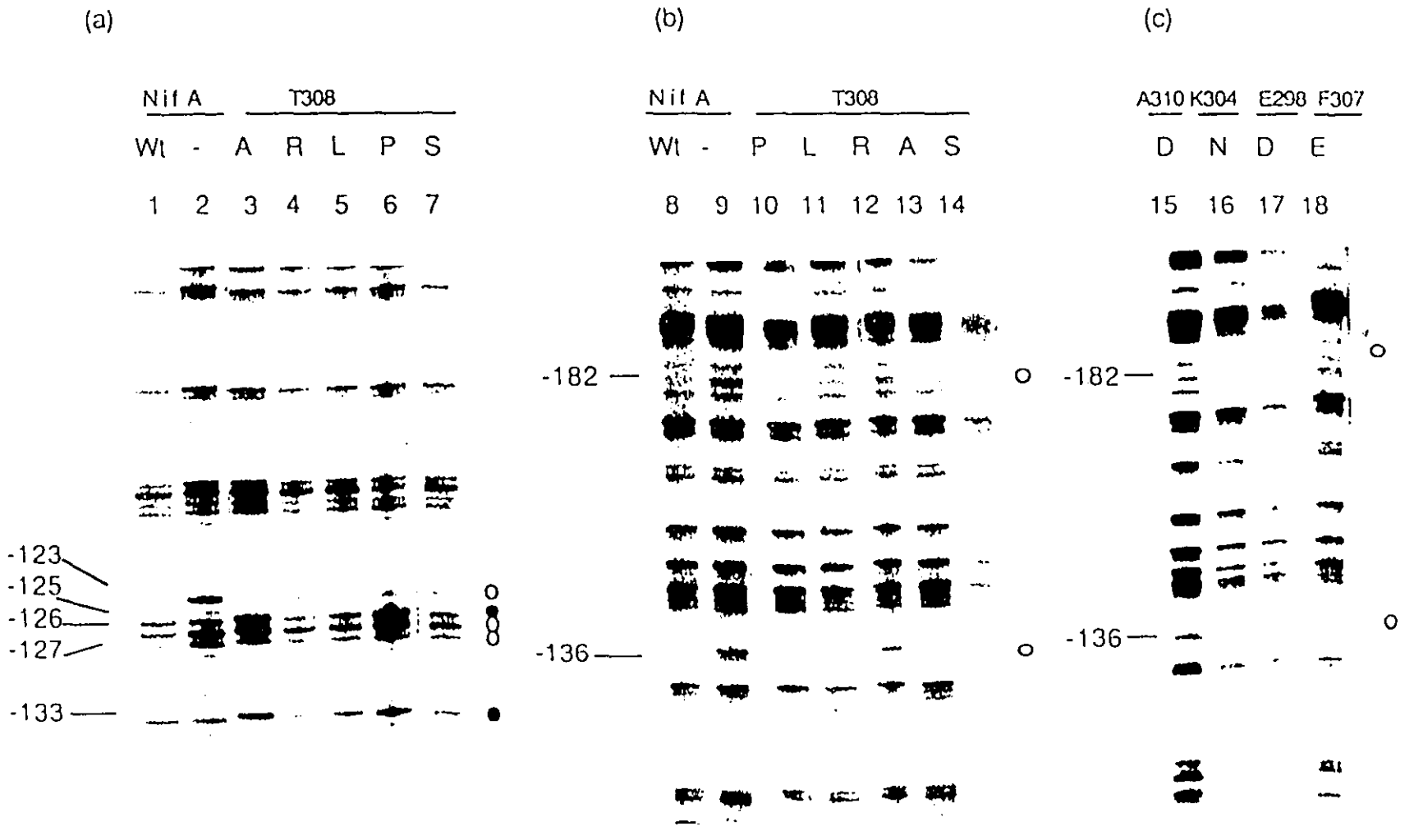
(a)
NifA
T308
Wt - A R L P S
1 2 3 4 5 6 7
-123
-125
-126
-127
-133
(b)
NifA
T308
Wt - P L R A S
8 9 10 11 12 13 14
-182
-136
(c)
A310 K304 E298 F307
D N D E
15 16 17 18
-182
-136

Figure 4

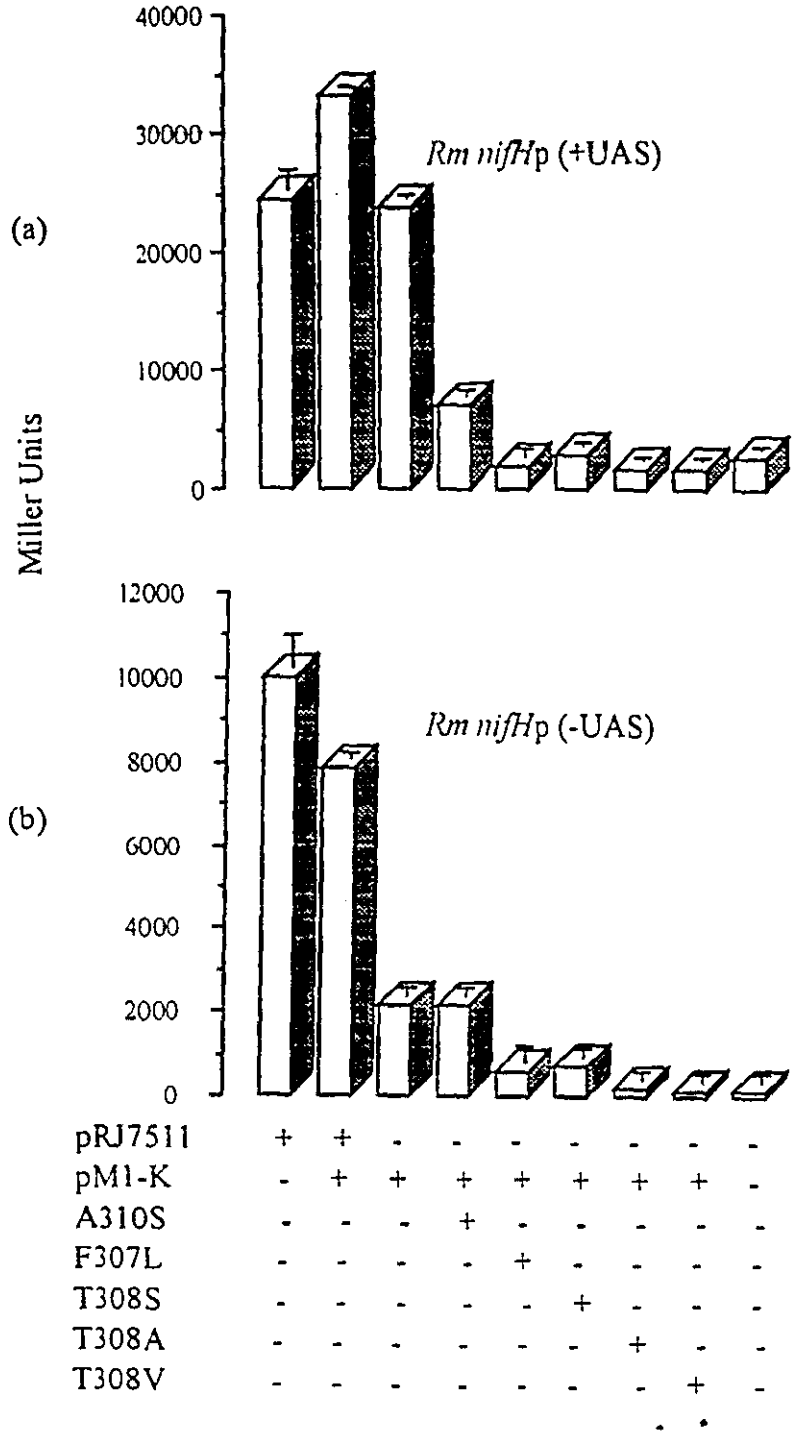
(a)
Miller Units
40000
30000
20000
10000
0
Rm nifHp (+UAS)
(b)
12000
10000
8000
6000
4000
2000
0
Rm nifHp (-UAS)
pRJ7511 + + - - - - - - -
pM1-K - + + + + + + + -
A310S - - - + - - - - -
F307L - - - - + - - - -
T308S - - - - - + - - -
T308A - - - - - - + - -
T308V - - - - - - - + -

Table 1. NifA missense mutants at region C3.

Mutant	Replacement		Codon Change#		Mutagenesis±
A306D¶	Ala 306 —>	Asp	GCC —>	GAC	RRC
A306N¶	"	Asn	"	AAC	"
F307A	Phe 307 —>	Ala	TTT —>	GCC	NN G/C
F307L	"	Leu	"	CTC	"
F307I	"	Ile	"	ATG	"
F307P	"	Prol	"	CCC	"
F307E	"	Glu	"	GAG	"
F307H	"	His	"	CAC	"
F307R	"	Arg	"	AGG	"
F307N	"	Asn	"	GAG	"
F307Y¶	"	Tyr	"	TAC	TAC (directed)
T308A	Thr 308—>	Ala	ACC —>	GCC	NN G/C
T308L	"	Leu	"	CTC	"
T308M	"	Met	"	ATG	"
T308P	"	Prol	"	CCC	"
T308R	"	Arg	"	CGC	"
T308S	"	Ser	"	AGC	"
T308V¶	"	Val	"	GTC	KKG
T308G¶	"	Gly	"	GGC	"
T308C¶	"	Cys	"	TGC	"
A310S	Ala 310 —>	Ser	GCT —>	AGT	RRT
A310N	"	Asn	"	AAT	"
A310D	"	Asp	"	GAT	"
A310G	"	Gly	"	GGT	"
K304N	Lys 304 —>	Asn	AAA —>	AAT	Taq pol error[1]
E298D	Glu 296 —>	Asp	GAA —>	GAT	Taq pol error[2]

underlined are the position changed.
¶ These mutants were first cloned at the pMOSt vector (Amersham, Co.)
± The resulting codon changes are indicated: N = G, A, T or C; R = A or G ; K = T or G (see experimental procedures).
[1, 2] Mutants obtained from the thr 308 NNG/C and ala 310 RRT mutagenesis experiment respectively.

Table 2. Activation and transdominance properties of the NifA mutant proteins

NifA protein	Colony appeareance	Activation[1]	Transdominance[1]
None	white	12 (±10)	n.t.[2]
NifA w.t.(pM1-K)	dark blue	10 487 (±827)	n.t.
NifAw.t.(pRJ7511)	dark blue	17 184 (±1680)	11 935 (±638)
A306D	white	4 (±4)	n.t.
A306N	white	8 (±3)	n.t.
F307A	white	27 (±6)	569 (±114)
F307H	white	20 (±5)	105 (±5)
F307E	white	15 (±3)	74 (±13)
F307L	white	18 (±3)	3 489 (±276)
F307M	white	15 (±3)	108 (±12)
F307P	white	13 (±3)	32 (±8)
F307R	white	11 (±4)	n.t.
F307N	white	13 (±3)	n.t.
F307Y	blue	6 220 (±556)	n.t.
T308A	white	8 (±2)	62 (±4)
T308L	white	7 (±1)	67 (±14)
T308M	white	10 (±1)	53 (±3)
T308S	slight blue	82 (±33	1 477 (±81)
T308P	white	7 (±3)	69 (±16)
T308R	white	17 (±7)	95 (±24)
T308V	white	7 (±3)	n.t.
T308C	white	8 (±2)	n.t.
T308G	white	7 (±1)	n.t.
A310S	blue	3 645 (±181)	11 012 (±1193)
A310N	white	13 (±5)	418 (±106)
A310D	white	15 (±4)	200 (±93)
A310G	white	17 (±2)	1 089 (±273)
K304N	white	17 (±4)	242 (±20)
E298D	slight blue	96 (±19)	1 786 (±210)

[1] Activation and transdominance are expressed in Miller units. The values represent the mean and the standard deviation of at least three independent experiments.

[2] Not tested.

Table 3. Intragenic suppressor analysis

NifA protein	Activity[1]
NifA w.t.	23 466 (±1250)
E298D	75 (±12)
E298D/S356G	6154 (±1067)
S356G	31 460 (±1403)
E298D/N357S	9 454 (±771)
N357S	26 100 (±141)
T308S	109 (±6)
T308S/V364I	1 633(±52)
V364I	29 067 (±2759)
E298D/V364I	346 (±55)
T308S/S356G	1 840 (±339)
T308S/N357S	26 636 (±2 759)
T308V	39 (± 11)
T308V/S356G	159 (±72)
T308V/N357S	370 (±98)
T308V/V364I	160 (±16)

[1] Expressed as Miller units. Data indicate the mean value and the standard deviation of three independent experiments

Epílogo.

A lo largo de este trabajo se describió el problema de comunicación entre moléculas durante el inicio de la transcripción. En particular nos concentramos en como los activadores bacterianos a distancia reconocen y estimulan la RNA pol-σ^{54}. Proponemos que en NifA, y por extensión en todos los demás activadores de la familia de las EBPs, se encuentra una región que contiene los determinantes estructurales y funcionales para el control positivo. Mutaciones en la denominada región C3 de NifA, NtrC y DctD, severamente afectan la actividad transcripcional de esta proteína sin alterar de manera importante otras funciones de la proteína como son la hidrólisis de ATP, la oligomerización y la unión con el DNA (North *et. al.*, 1997; Wang & Hoover, 1997; González *et. al.*, 1998 en prensa). Por lo consiguiente estas variantes se comportan como mutantes en el control positivo (Hoschild *et. al.*, 1983; Guarente *et. al.*, 1981; Eschenlauer & Reznikoff, 1991; Zhou *et. al.*, 1993).

No conocemos como la región C3 contribuye a la actividad transcripcional de la proteína. Es probable que este constituida por dos subregiones con atributos funcionales distintos. El hecho de que una parte de la región C3 (GAFTGA) no se encuentra en proteínas reguladoras como TyrR, que no dependen de la RNA pol-σ^{54} sino de la RNA polimerasa-σ^{70}, hizó pensar que esta zona es importante para reconocer el complejo cerrado RNA pol-σ^{54} (Morett & Segovia, 1993). Como discutimos en el capítulo 7 y fué previamente propuesto por Osuna *et. al.*, (1995), la región C3 puede representar una zona de interacción con la RNA pol -σ^{54} y/o una zona de transducción de la señal de hidrólisis de ATP, por analogía con EfTu, una proteína de la familia de las GTPasas (Bourne, 1991).

El trabajo desarrollado recientemente por el grupo de T. Hoover en Athens, GA, sugiere que la región C3 efectivamente partipa en interacciones proteína-proteína con la RNA

polimerasa, pero no a tráves de la subregión carboxilo (GAFTGA) sino de la subregión amino (ESELFGHEK) (Wang *et. al.*, 1997). Mediante puentes químicos fotoestimulables ellos encontraron que DctD puede unirse a las subunidades β´ y σ^{54} (Wang & Hoover, 1995). Mutantes sencillas de DctD alteradas en una parte de la región C3 (ESELFGHEK) y que son incapaces de realizar transcripción no pueden unirse a las subunidades mencionadas, mientras que mutantes alteradas en la región GAFTGA, que tampoco son capaces de iniciar transcripción, retienen esta capacidad. Esto ha sido interpretado como que el contacto inicial, o al menos el detectable por puentes químicos, se lleva a cabo mediante la subregión ESELFGHEK, mientras que la GAFTGA funcionaría para la transducción de la señal de hidrólisis de ATP. En el mismo sentido, North *et. al.*, (1997), han argumentado que la GAFTGA está encargada de transducir la señal de hidrólisis de ATP, basándose en que los aminoácidos que la conforman no tienen grupos funcionales reactivos. A pesar de no contar con evidencias directas, en nuestro análisis genético, demostramos la importancia de mantener un residuo aromático y un grupo hidroxilo en las posiciones 307 y 308 de NifA, para activar la transcripción *in vivo*, aunque ciertamente a muy bajos niveles. Es llamativo que las regiones de interacción proteína-proteína contengan comúnmente residuos aromáticos y en aquellas que son transitorias contengan residuos con un grupo hidroxilo (Jones & Thornton, 1996.). Residuos aromáticos e hidrofóbicos también han sido encontrados frecuentemente en activadores transcripcionales virales como VP16 y en GCN4 de levadura (Regier *et al.*,., 1993; Jackson *et. al.*, 1996).

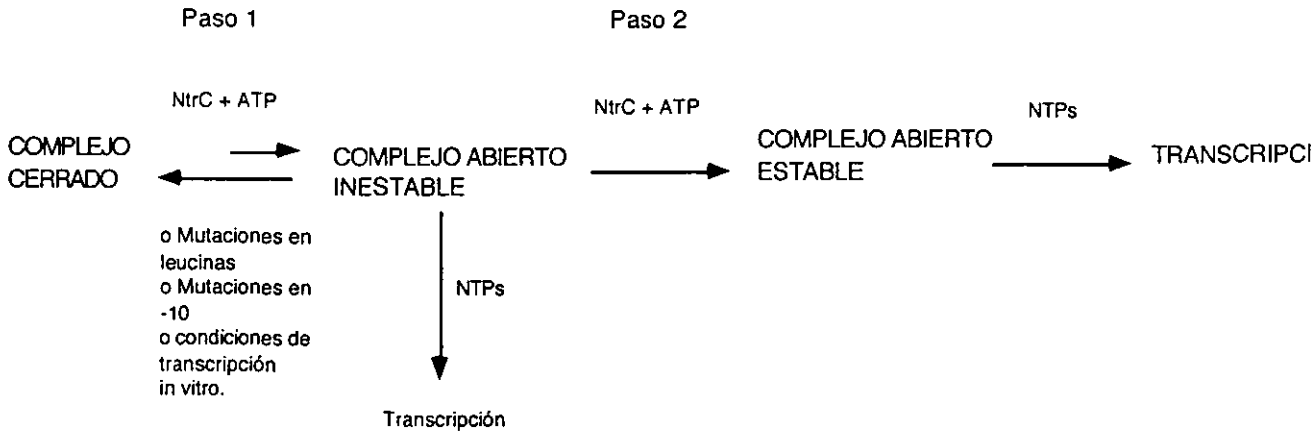

Figura 15. El modelo extendido de activación de la transcripción por NtrC (Wang & Gralla, 1996).

Podemos tratar de interpretar estos resultados a la luz del mecanismo de inicio de la transcripción propuesto por Wang & Gralla (1996; Figura 15). En este modelo se considera que la RNA pol-σ^{54} en el complejo cerrado tiene bloqueadas las funciones de apertura del promotor. Estas funciones han sido atribuidas al dominio amino terminal de σ^{54}, el cual es requerido para la unión con la región -12 del promotor (Sasse-Dwight & Gralla, 1990). El concurso de la proteína activadora y del ATP es absolutamente necesario para superar esta inhibición de la transcripción. Sin embargo, mutantes de σ^{54} afectadas en varias de las leucinas localizadas en el amino terminal, mutantes en la región -12 del promotor y ciertas condiciones de transcripción *in vitro* como la fuerza iónica o la temperatura, confieren al sistema independencia del activador (Wang *et. al.*, 1997). A pesar de esto, la transcripción no alcanza los niveles obtenidos en presencia de la proteína activadora y la holoenzima constituida con una mutante de σ^{54} aún responde a la acción de NtrC y del ATP. En contraste con la resistencia al polianion heparina que exiben los complejos abiertos formados normalmente, los complejos abiertos formados por estas mutantes de σ^{54} o simulados por las condiciones de transcripción, son muy sensibles a la heparina (Wang & Gralla, 1996). Se ha calculado que aproximadamente un décimo de los

complejos formados en estas condiciones están abiertos. Dadas estas observaciones se ha concluido que el activador y el ATP participan en un segundo paso estabilizando el complejo abierto de vida media más larga (Wedel & Kustu, 1995; Wang & Gralla, 1996). Recientemente se ha observado que una mutante de σ^{54} que carece del dominio amino terminal, es capaz de iniciar transcripción a partir de los promotores *glnAp2* y *glnH* bajo ciertas condiciones *in vitro*, pero no *in vivo* (Sasse-Dwight & Gralla, 1990; Gralla *et. al.*, 1997). Sin embargo esta mutante es incapaz de responder al activador NtrC. Este resultado ha sido tomado como evidencia de que el activador interacciona directamente con el dominio amino de σ^{54}. Estudiando la formación de complejos de iniciación en los promotores dependientes de σ^{54} mediante proteólisis limitada, Casaz & Buck (1997), han encontrado cambios conformacionales en varias regiones de σ^{54}. Particularmente interesante es la sensibilidad a las proteasas mostrada por el glu 36 localizado en el dominio amino de σ^{54}, indicando que esta proteína puede ajustar su conformación después de la interacción alguno de sus ligandos.

Un mecanismo de dos pasos para la formación del complejo abierto también ha sido sugerido basándose en las propiedades termodinámicas de los complejos transcripcionales formados por NtrC (Wedel & Kustu,1995). Estos investigadores demostraron que los complejos abiertos rápidamente decaen cuando NtrC es retirado del los complejos transcripcionales por medio de un tamizado a tráves de columnas de sefarosa. Por lo consiguiente NtrC es necesario para la estabilización del complejo abierto. (Wedel & Kustu, 1995).

En el modelo estructural del dominio Central la subregión ESELFGHEK forma una α hélice. La subsititución de cada uno de estos residuos por alanina revela que solo alteraciones en glutamato 211, serina 212 y fenilalanina 215 afectan severamente la transcripción , mientras que los cambios en los otros residuos solamente la reducen (Wang

et. al., 1997). esto hace pensar que los residuos importantes para el contacto con la RNA polimerasa se encuentran en la cara de la hélice expuesta al solvente. Estas mutantes de DctD son incapaces de interaccionar con σ^{54}, pero conservan afinidad por la subunidad β'. Esto sugiere que es necesario el contacto simultáneo del activador con ambas subunidades de la RNA pol-σ^{54} para isomerizar el complejo cerrado, tal vez mediante distintos monómeros (Wang *et. al.*, 1997).

El segundo glutamato de la subregión ESELFHEK ha sido alterado en NtrC, DctD, TyrR y NifA produciendo proteínas severamente afectadas transcripcionalmente pero que al menos en las tres primeras proteínas, con suficiente actividad de hidrólisis de ATP (North *et. al.*, 1996; Wang *et. al.*, 1997; Kwok *et. al.*, 1995; González *et al.*, en prensa). Tres diferentes substituciones en esta posición en DctD (alanina, glicina y valina) tampoco están alteradas en el reconocimiento con las subunidades β 'y σ^{54} (Wang *et. al.*, 1997). Esto tiene sentido sí consideramos que este residuo se encuentra en la cara opuesta de la hélice, en relación a aquellos residuos implicados en el contacto con la RNA pol-σ^{54}. En TyrR, la substitución de este glutamato por glutamina, produce una proteína deficiente en la unión tirosina dependiente de ATP (Kwok *et. al.*, 1995). Esta mutante resulta incapaz de reprimir operones que requieren de la formación de un oligomero de alto orden. Es posible que este residuo conservado se encuentre en la ruta de los cambios conformacionales que ocurren en la proteína después de la hidrólisis de ATP, es decir puede ser parte de la función de transducción de energía para acoplarla a la apertura del complejo abierto.

Dado que las mutantes en la región GAFTGA de DctD aún son capaces de unirse eficientemente a β' y σ^{54}, y en general conservan la actividad de la ATPasa es posible que la función que se este alterando en estas proteínas sea la de transducción de la señal de hidrólisis de ATP. No es claro como sucedería esto pero es importante señalar que tal vez

involucre contactos directos con el complejo RNA pol-σ^{54}-promotor tan débiles que no son detectados mediante un ligamiento químico *in vitro*. Otra alternativa es considerar que la GAFTGA participa en la estabilización posterior del complejo recién abierto. El requisito de un amino ácido aromático y un residuo con un grupo hidroxilo en esta región sugieren que pudieran participar en interacciones proteína-proteína o proteína DNA, sucesos que deben occurrir durante la isomerización del complejo cerrado. Dos experimentos en progreso en el laboratorio pueden ayudar a clarificar este punto. Uno de ellos es examinar la formación de los complejos abiertos en presencia de las mutantes de NifA afectadas en la GAFTGA que llevan substituciones conservativas. Este experimento asume que, dado que la subregión ESELFGHEK esta intacta, las interacciones con la RNA pol-σ^{54} están preservados y que esto no es suficiente para activar la transcripción. No sabemos sí los amino ácidos de la región GAFTGA contribuyen a abrir el promotor o a estabilizarlo. Sí resultara ser ésta última opción, se esperarían cambios en los patrones de apertura del complejo abierto por exposición de las cadenas sencillas al $KMnO_4$ Un segundo experimento consiste en la búsqueda de supresoras extragénicas de mutaciones en la región C3 de NifA. En este caso, los productos de los genes *rpoC* y *rpoN* son los candidatos obvios a contener aminoácidos que sirvan de interacción con el activador.

"voy por tus pensamientos afilados .. .

PIEDRA DE SOL (Octavio Paz)

Referencias [1]

Agron, P. G., Ditta, G. S., & Helinski, D. R. (1993) Oxygen regulation of *nifA* transcription *in vitro*. *Proc. Natl. Acad. Sci. USA* **90**: 3506-3510.

Agaarwal, A. K., & Harrison S. C. (1990) *Ann Rev Biochem* **59**: 933-969.

Ayers, D., Sunshine, M., Six, W., Christie, G. (1994) Mutations affecting two adjacent amino acids in the α subunit of RNA polymerse block transcriptional activation by the bacteriphage P2 Ogr protein. *J Bacteriol* **176**: 7511-7518.

Alberts, B., & Miake-Lye, R. (1992). Unscrambling the puzzle of biological machines: The importance of the details. *Cell* **68**: 13179-13184.

Albright, L. M., Huala, E., & Ausubel, F. M. (1989) Prokaryotic signal transduction mediated by sensor and regulator protein pairs. *Annu Rev Genet* **23**: 311-336.

Allison., L. A., Moyle, M., Shales, M., & Ingles, C. P. (1985). Extensive homology between the largest subunits of eukaryotic and prokaryotic RNA polimerases. Cell 42: 599-610.

Argaet, V., Wilson, T. J., & Davidson, B. E. (1994) Purification and analysis of the *Escherichia coli* regulatory protein TyrR and analysis of its interactions with ATP, tyrosine, phenylalanine, and trytophan. *J Biol Chem* **269**: 5171-5178.

Armaleo. D. (1987) Structure and evolution of prokaryotic and eukaryotic RNA polymerases: a model. *J Theor Biol* **127**: 301-314.

Aurora, R., Srinivasan, R., & Rose, G. (1994) Rules of a helix termination by glycine. **Science** *264*: 1126:1130.

Austin, S., Kundrot, C. , & Dixon, R. (1991) Influence of mutation in the putative binding site of the nitrogen regulatory protein NtrC on its positive control function. *Nuc Acids Res* 19: 2281-2287.

Austin, S., & Dixon, R. (1992). The prokaryotic enhancer binding protein NtrC has and ATPase activity wich is phosphorylation and DNA dependent. *EMBO J.* **11**: 2219-2228.

Austin, S., Buck, M., Cannon, W., Eydmann, T., & Dixon, R. (1994) Purification and *in vitro* activities of the native nitrogen fixation control proteins NifA and NifL. *J Bacteriol* 176: 3460-3465.

[1] Las referencias en negritas aparecen citadas en el artículo original.

Beckwith, J. (1996) The operon: an historical account. In *Escherichia coli & Salmonella*, cellular and molecular biology,2nd ed. (Neidhart, F. C., ed) p. 1227-1231. ASM press. Washintong, D. C.

Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., Buc, H., Minchin, S., Williams, J., & Busby, S. (1990) Mutations that alter the ability of the *Escherichia coli* cyclic cAMP protein to activate transcription. *Nucleic Acids Res* **18**: 7243-7250.

Berger, D., Narberhaus, F., & Kustu, S. (1995). The isolated catalytic domain of NIFA, a bacterial enhancer-binding protein , activates transcription *in vitro*: activation is inhibited by NifL. *Proc. Natl. Acad. Sci. USA*. **91**: 103-107.

Berger, D., Narberhaus, F., Lee, H.S. & Kustu, S. (1995) *In vitro* studies of the domains of the nitrogen fixation regulatory protein NifA. *J Bacteriol* 177: 191-199.

Better, M., Ditta, G., & Helinski, D. (1985) Deletion analysis of *Rhizobium meliloti* symbiotic promoters. *EMBO J* 10: 2419-2424.

Beynon, J., Williams, M. K., & Cannon, F. C. (1988) Expression and functional analysis of the *Rhizobium meliloti nifA* gene. *EMBO J* 7: 7-14.

Bishop, P. E. & Premakumar, R. (1992). Alternative nitrogen fixation systems. In Biological nitrogen fixation (Stacey, G., Evans, H. J., & Burris, R. H., eds). p.736- 762. Chapman & Hall, New York

Blanco, G., Drummond, M., Woodley, P., & Kennedy, C. (1993) Sequence and molecular analysis of the **nifL** gene of Azotobacter vinelandii. *Mol Microbiol* **9**: 869-879.

Blatter, E. E., Ross, H., Tang, R. L. Gourse, R., & Ebright R. (1994) Domain organization of RNA polymerasa α subunit : C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**: 889-896.

Blattner, F. R. *et. al.* (1997) The complete geenome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-1462.

Bokal. A. J., Ross, W., and Gourse, R. L. (1995). The transcriptional activator protein FIS: DNA interaction and cooperative interactions with RNA polymerase at the *Escherichia coli rrnB* P1 promoter. *J Mol Biol* **245**: 197-207.

Borukhov, S., Sagitov, S., Josaitis, R., Gourse, R. , & Goldarf, A. (1993) Two modes of transcription initiation in vitro at the *rrnB* P1 promoter of *Escherichia coli. J Biol Chem* **268**: 23477-23482.

Bourne, H. R., Sanders, D. A. , & McCormick, F. (1990) The GTPasa superfamily: a conserved switch for diverse cell functions. ***Nature*** **348: 643-653.**

Brent. R., & Ptashne, M. (185) A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**: 729-736.

Brenann, R. G., & Mattews, B. W. (1989) *J. Biol Chem* **264**: 1903-1906.

Brown, S. E., & Ausubel, F. M., (1984) Mutations affecting regulation of the *Klebsiella pneumoniae nifH* (nitrogenase reductase promoter*). J Bacteriol* **157**: 143-147.

Brunner, M., & Bujard, H. (1987) Promotor recognition and promotor strength in the *Escherichia coli* system. *EMBO J.* **16**: 3139-3144.

Buck, M. (1986) Deletion analysis of the *Klebsiella pneumoniae* nitrogenase promoter: importance of spacing between conserved sequences around positions -12 and -24 for activation by the *nifA* and *ntrC* (*glnG*) products. *J Bacteriol* **166**: 545-551.

Buck, M., Miller, S., Drummond, M., & Dixon, R. (1986) Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature* **320**: 374-378.

Buck, M.,., Woodcock, J. & Cannon, W (1987) Transcriptional activation of the *Klebsiella pneumonia* nitrogenase promoter may involve DNA-loop formation. *Mol Microbiol* 1: 243-249.

Buck, M. & Cannon, W. (1989) Mutations in the RNA polymerase recognition sequence of the *Klebsiella pneumoniae nifH* promoter permitting transcriptional activation in the absence of NifA binding to upstream activator sequences. ***Nuc Acids Res*** **17: 2597-2612.**

Buck, M. & Cannon, W. (1992) Specific binding of the transcription factor sigma-54 to promoter DNA. ***Nature*** **358: 422-423.**

Buckle, M., & Buc, H. (1989) Fine mapping of DNA single stranded regions using base-specific chemical probes: study of an open complex formed between RNA polymerase and the *lac* UV promoter. *Biochemistry* **28**: 4328-4396.

Buikema, W. J., Szeto, W. W., Lemley, P. V., Orme-Johnson, W. H.,& Ausubel, F. M.(1985) Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen fixation gene *ntrC* of *K. pneumoniae. Nucleic Acids Res* **13**: 4339-4555.

Burguess, R. R., Travers, A. A., Dunn, J. J., & Bautz, E. K.F. (1969) Factor stimulating transcription by RNA polymerase. *Nature* **221**: 43-46.

Busby, S., & Ebright, R. H. (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell* **79**: 743-746.

Busby, S., & Ebright, R. H. (1997) Transcription activation at class II CAP-dependent promoters. *Mol Microbiol* **23**: 853-859.

Bushman, F. D., Shang, C., & Ptashne, M. (1989) A single glutamic acid residue plays a key role in the transcriptional activation function of lamda repressor. *Cell* **58**: 1163-1171.

Bushman, F., & Ptashne, M. (1986) Activation of transcription by the bacteriophage 434 repressor. *Proc. Natl. Acad. Sci. USA* 83: 9353-9357

Bushman, F., & Ptashne, M. (1988) Turnig l Cro into a transcriptional activator. *Cell* **54**: 191-197.

Cannon, W. V., Kreutzer, R., Kent, H. M., Morett, E., & Buck, M. (1990). Activation of the *Klebsiella pneumoniae nifU* promoter: identification of multiple and overlapping upstream NifA binding sites. *Nucl Acids Res* **18**: 1693-1701.

Cannon, W. V., & Buck, M. (1992). The central domain of the positive control protein NifA and its role in transcriptional activation . *J.Mol Biol* **225**: 271-286.

Cannon, W. V., Chaney, M. K., Wang, X. Y., & Buck, M. (1997) Two domains within σ^N (σ^{54}) cooperate for DNA binding. *Proc Natl Acad Sci USA* 94: 5006-5011.

Carmona, M., Claverie-Martin, F., & Magasanik, B. (1997). DNA bending and the initiation of transcription at s54 dependent bacterial promoters. *Proc. Natl. Acad. Sci. USA* **94**: 9568-9572.

Cannon, W. V., Claverie-Martin , F.., Austin, S., & Buck, M. (1994). Identification of a DNA-contacting surface in the transcription factor sigma-54. *Mol. Microbiol* **11**: 227-236.

Chamberlin, M.J. 1976. In RNA polymerase (Losick, R., & Chamberlin, M., eds) p. 17-67 and 159-191. Cold Spring Harbor Laboratory. Cold Spring Harbor. New York.

Chartejee, S., & Struhl, K. (1995) Connecting a promoter bound protein to TBP bypasses the need for a transcriptional domain. *Nature 374*: 820-822.

Choy, B., & Green, M. (1993). Eukaryotic activators function during multiple steps of preinitiation complex assembly. *Nature* **366**: 531-536.

Cazas, P., & Buck, M., (1997). Probing the assembly of transcription initiation complexes trough changes in σ^N protease sensitivity. *Proc. Natl. Acad. Sci. USA* **94**: 12145-12150.

Chan, C. L., Lonetto, M., & Gross, C. (1996) Sigma domain structure: one down, one to go. *Structure* **4**: 1235-1238.

Chen, P., & Reitzer, L., (1995) Active contribution of two domains to cooperative DNA binding of the enhancer-binding protein nitrogen regulator I (NtrC) of *Escherichia coli*: stimulation by phosphorylation and the binding of ATP. *J. Bacteriol.* **177**: 2490-2496.

Chen, Y., Ebright, Y. W., & Ebright, R. H. (1994) Identification of the target of a transcriptional activator protein by protein-protein phohocrosslinking. *Science* **265**: 90-92.

Ching, L., O'Neill, E., & Shingler, V. (1996) Genetic evidence for interaction regulation of the phenol-responsive σ^{54}-dependent activator DmpR. *J Biol Chem* **271**: 17281-17286.

Contreras, A., & Drummond, M. (1987) The effect on the function of the transcriptional activator NtrC from *Klebsiella pneumoniae* of mutations in the DNA recognition helix. *Nucl. Acids Res.* **16**: 4025-4039.

Craig, M. L., Suh, W. C., & Record, T. H. (1995) HO and DNAase I probing of Eσ^{70}RNA polymersase-λ Pc promoter open complexes: Mg^{2+} binding and its structural consequences at the transcription start site. *Biochemistry* **34**: 15624- 15632.

Coppard, J. R., & Merrick, M. J. (1991) Cassette mutagenesis implicates a helix-turn-helix motif in promoter recognition by the novel RNA polymerase sigma factor σ^{54}. *Mol Microbiol* **5**: 1309-1327.

Cui, J., & Sommerville, R. (1993) A mutational analysis of the structural basis for trasncriptional activation and monomer-monomer interction in the TyrR system of *Escherichia coli. J Bacteriol* 175: 1777-1784.

Cui, J, & Sommerville, R. (1993) The TyrR protein of *Escherichia col:* analysis by limited proteolysis and ligand mediated conformational changes. *J. Biol. Chem.* **268:** 5040-5047.

Cui, J., Ni, L., & Somerville, R. (1993) ATPasa activity of TyrR, an transcriptional regulatory protein for σ^{70} RNA polymerase. *J Biol Chem* 268: 13023-13025.
Daniels, D., Zuber., P., & Losick, R. (1990) Two amino acids in an RNA polymerase s factor involved in the recognition of adjacent base pairs in the -10 region of a cognate promoter. *Proc Natl Acad Sci USA* **87**: 8075-8079.

Darst, S. A., Edwards, S. M., Kubalek, E. W., & Kornberg, R. D. (1991) Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. Cell **66**: 121-128.

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

David, M., Daveran, M.L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P., & Khan, D. (1988). Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. *Cell* **54**: 671: 683.

Delgado, A., & Ramos, J. L. (1994) Genetic evidence for activation of the positive transcriptional regulator XylR, a member of the NtrC family of regulators, by effector binding. *J Biol Chem* **269**: 8059-8062.

de Philip, P., Batut, J., Boistard, P. (1990) *Rhizobium meliloti* FixL is an oxygen sensor and regulates *R. meliloti* nifA and FixK genes differently in *Escherichia coli*. *J Bacteriol* **172**: 4255-4262.

deHaseth, P. L., & Helmann, J. D. (1995) Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerse-induced strand separation of double helical DNA. *Mol. Microbiol.* **16**: 817-824.

Dodd, I., & Egarn, J. B. (1990) Improved detection of helix-urn-helix DNA binding motifs in protein sequences. Nucleic Acids Res 18: 5019-5026.

Dombroski, A., Walter, W., & Gross, C. (1993). Amino-terminal amino acids modulate sigma-factor DNA-binding. *Genes Dev.* 7: 2446-2455.

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

Dove, S. L., Joung, J. K., & Hoschild, A. (1997) Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386**: 627.

Drummond, M., Whitty, P & Wooton, J. (1986) Sequence and domain relationships of NtrC and NifA from *Klebsiella pneumoniae*: Homologies to other regulatory proteins. *EMBO J* 5: 441-447.

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

Ebrigth, R. H. (1993) Transcription at class I CAP dependent promoters. *Mol. Microbiol.* **8**: 797-802.

Ebrigth, R. H., & Busby, S. (1994) The *Escherichia coli* RNA polymerase α subunit: structure and function. *Current Op Gen Genet* **5**: 197-203.

Eschenlauer, A. C. & Reznikoff, W. S. (1991) *Escherichia coli* catabolite gene activator protein mutants defective in positive control of *lac* operon transcription. *J. Bacteriol.* **173**: 5024-5029.

Fields, S., & Song, O. K. (1989). *Nature* **340**: 245-246.

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

Fischer, H. M., & Hennecke, H. (1987) Direct response of *Bradyrhizobium japonicum* NifA-mediated nif gene regulation to cellular oxygen status. *Mol Gen Genet* 209: 621-626.

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

Flashner, Y., Weiss, D., Keener, J., & Kustu, S. (1995). Constitutive forms of the enhacer-binding protein NtrC: evidence that essential oligomerization determinants lie in the central activation domain. *J Mol Biol* **249**: 700-713.

Foster-Hartnett, D., Cullen, P. J., Monika, E. M., & Kranz, R. G. (1994) A new type of NtrC transcriptional activator. *J Bacteriol* 176: 6157-6187.

Frankel, A. D., & Kim, P. S. (1991) Modular structure of transcription factors: implications for gene regulation. *Cell* **65**: 717-719.

Fry, D. C., Kuby, S. A., Mildvan, A. S. (1986) ATP-binding site of adenylate kinase: mechanistic implications of its homology with ras-encoded p21, F1-ATPase and other nucleotide-binding proteins. *Proc Natl Acad Sci USA* **83**: 907-911.

Gaal, T., Ross, W., Blatter, E. E., Tang, H., Jia, X., Krishnan, V. V. , Assa-Munt, N., Ebright, R. H., & Gourse, R.L., (1996) DNA-binding determinants of the α subunit of RNApolymerase: novel DNA-binding domain architecture. *Genes Dev* **10**:16-26.

Gilles-González, M. A., Ditta, G. S., & helinski, D. R. (1991) *Nature* **350**: 170-172.

Guarente, L., Nye., J. S., Hoschild, A., & Ptashne, M. (1982) A mutant lamda repressor with specific defect in its positive control function . *Proc Natl Acad Sci USA* **79**: 2236-2240.

García, E., Bancroft, S., Rhee, G., & Kusty, S. (1977) The product of the newly identified gene, *glnF*, is required for the synthesis for the synthesis of glutamine synthethase. *Proc Natl Acad Sci USA* **74**: 1662- 1666.

Gaston, K., Bell, A., Kolb, A., Buc, H., & Busby, H., (1990). Stringent spacing requirements for transcription activation by CRP. *Cell* **62**: 733-743.

Gilbert, W. (1976) Starting and stopping sequences for the RNA polymerase. In RNA polymerase (Losick, R., & Chamberlin, M., eds). p. 193-206. Cold Spring Harbor Laboratory. . Cold Spring Harbor, New York.

Gill. G., & Ptashne, M. (1987). Mutants of Gal4 protein altered in an activation function. *Cell* **51**: 121-216.

Ginniger, E., & Ptashne, M. (1987) Transcription in yeast activated by a putative amphiphatic α helix linked to a DNA. *Nature* **330**: 670-673.

Goff, C., (1984) Coliphage-induced ADP-ribosylation of *Escherichia coli* RNA polymerse. *Methods Enzymol* **106**: 418-429.

Gross, C., Lonetto, M., & Losick, R., (1992). Bacterial sigma factors. In Transcriptional Regulation (McKnight S. & Yamamoto, K., eds) p. 129-176. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Gu, B., Lee, J. H., Hoover, T. R., Chool, D., & Nixon, B. T. (1994) *Rhizobium meliloti* DctD, a σ^{54} transcriptional activator , may be negatively controled by a subdomain in the C-terminal en of itsTwo-component receiver module. *Mol Microbiol* **13**: 51-66.

Hahn, S. (1993) Structure (?) and function of acidic transcription activators. *Cell* **72**: 4891-483.

Harley, D., & Reynolds, R. P. (1987) Analysis of E. coli promoter sequences. *Nucleic Acids Res* 15:2343-2361.

Hawley, D., & McClure, W. R. (1983) Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res* **11**: 2237-2255.

Heyduk, T., Lee, J. C., Ebright, Y. W., Blatter, E. E., Zhou, Y., & Ebright R. H. (1993) CAP interacts with RNA polymerase in solution in absence of promoter DNA. *Nature* **364**: 548-549.

Helmann, J. D., & Chamberlin, R. (1988) Structure and function of bacterial sigma factors. *Annu Rev. Biochem.* **57**: 839-872.

Henderson, N., Austin, S., & Dixon, R. (1989). Role of metal ions in negative regulation of nitrogen fixation by the nifL gene product from *Klebsiella pneumoniae. Mol Gen Genet* **216**: 484-491.

Herskowitz, I (1987) Functional inactivation of genes by dominant negative mutations. ***Nature*** **329: 219-222.**

Hill, S., Kennedy, C., Kavangh, E., Goldberg, R. B., & Dixon, R. (1981) Nitrogen fiaxation gene (*nifL*) involved in oxygen regulation of nitrogenase synthesis in *Klebsiella pneumoniae*. *Nature* **290**: 424-426.

Hirshman, J., Wong, P. K., Sei, K., Keener, J., & Kustu, S. (1985) Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transciption in vitro: evidence that the *ntrA* product is a sigma factor. *Proc Natl Acad Sci USA* **82**: 7525-7529.

Hoover, T. R., Santero, E., Porter, S., & Kustu, S. (199) The integration host factor stimulates interaction of RNA polymerase with NifA, the transcriptional activator of nitrogen fixation operons. *Cell* **63**: 11-22.

Hsieh, M., Tintut, Y., & Gralla, J. D. (1994), Functional roles for the glutamines within the glutamine-rich region of the transcription factor σ^{54}. *J. Biol. Chem.* **269**: 373-378.

Hilgenfield, R. (1995) Regulatory GTPasas. ***Curr Op Struct Biol*** **5, 810-817.**

Hofer, B., Muller, D., Koster, H. (1985) The pathway of *E. coli* RNA polymerase - promoter-complex formation as visulaized by footprinting. *Nucleic Acids Res* **16**: 5995-6013.

Hope, I. A., & Struhl, K. (1986) Functional dissection of a eucaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**: 885-894.
Hoschild, A., Irwin, N., & Ptashne, M. (1983). Repressor structure and the mechanism of positive control. *Cell* **32**: 319-325.

Hoschild, A., & Ptashne, M. (1988). Interaction at a distance between λ repressors disrupt gene activation. *Nature* **336**: 353-357.

Huala, E. & Ausubel, F. (1989) The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from *R. meliloti nifH* promoter.*J Bacteriol* 171: 3354-3365.

Huala, E., Stitger J., & Ausubel, F. (1992) The central domain of Rhizobium leguminosarum DctD functions independently to activate transcription. *J. Bacteriol.* **174**: 1428-1431.

Huet, J., Schnabel, R., Sentenac, A., & Zillig, W. (1983) Archaeobacteria and eucryotes posses DNA-dependent RNA polymerases of a common type. *EMBO J.* **2**: 1291- 1294.

Hunt, T. P., & Magasanik, B. (1985) Transcription of glnA by purified components: core RNA polymerase and products of gln F, glnG and glnL. *Proc Natl Acad Sci USA* **82**: 8453-8457.

Igarishi, K., Fujita, N., & Ishihama, A. (1991) Identification of the a subunit assembly domain in the alpha subunit of *Escherichia coli* RNA polymerse. *J Mol Biol* **218**: 1-6.

Igarishi, K., Fujita, N., & Ishihama, A. (1991) Bipartite functional map of the *Escherichia coli* RNA polymerse alpha subunit: involvement of the C-terminal region in trannscription activation by cAMP-CRP. *Cell* **65**: 2640-2649.

Ishihama, A. (1981). Subunit assembly of *Escherichia coli* RNA polymerase. *Adv Biophys* **14**: 1-35.

Ishihama,.A. (1993) Protein-protein communiction within the transcription apparatus. *J. Bacteriol* **175**: 2483-2489.

Iwabe, N., Kei-Ichi, K., Kishino, H., Hasegawa, M., & Miyata, T. (1991). Evolution of RNA polymerases and branching patterns of the three major groups of archaeobacteria. *J Mol Evol* **32**: 70-78.

Jackson, B., Drysdale, C. M., Natarajan, K., & Hinnebush, A. (1996). Identification of seven hydrophobic clusters in GCN4 making redundant contributions to transcriptional activation. *Mol Cell Biol* **16**: 5537-5571.
Jacob , F. & Monod,.J. (1961) Genetic regulatory mechanisms in the synthesis of proteins, *J Mol Biol* **3**: 318- 356.,

Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Lu, X., Tantillo, C., Willimas, R. L., Kramer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., & Arnold, E. (1993). Crystal structure of human inmunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. ***Proc Natl Acad Sci USA* 90:** 6320-6324.

Jeon, Y. H., Negishi, T., Shirakawa, M., Yamazaki, T., Fujita, N., Ishihama, A., Kyogoku, Y. (1995) Solución structure of the activator contac domain of the RNA polymerase α subunit. *Science* **270**: 1485-1497.

Jin, D. J., & Turnbough, C. L., (1994). An *Escherichia coli* RNA polymerase defective in transcription due to its overproduction of abortive initiation products. *J Mol Biol* **236**: 72-80.

Jin, R., Sharif., K. A. & Krakov, J. S. (1995) Evidence for contact between the cyclic AMP recepto protein and the subunit of *Escherichia coli* RNA polymerase. *J Biol Chem* **270**: 19213-19216.

Joo, D. M., Ng, N., & Calendar, R. (1997) A s32 mutant with a single amino acid change in the highly conserved region 2.2 exhibits reduced core RNA polymerase affinity. ***Proc Natl Acad Sci USA* 94**: 4907-4912.

Jones, S., & Thornoton, J. (1996) Principles of protein-protein interactions. ***Proc Natl Acad Sci USA* 93**: 13-20.

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

Joung, J., Lee, L., & Hoschild, A. (1993) Synergistic activation of transcription by Escherichia coli cAMP receptor protein. *Proc Natl Acad Sci USA* **90**: 3083-3087.

Joung, J., Koepp, D., & Hoschild, A. (1994) Synergistic activation of transcription by l cI protein and E. coli cAMP receptor protein . *Science* **265**: 1863-1866.

Joerger, R. D., Jacobson, M.R., & Bishop, P. E. (1989) *J Bacteriol* **171**: 3258-3267.

Juang, Y. L., & Helmann, J. D. (1994) A promoter melting region in the primary σ factor of *Bacillus subtilis* . Identification of functionally important aromatic amino acids. *J. Mol. Biol.* **235**: 1470-1488.

Kashlev, M., Lee, J., Zalenskaya, V., Nikiforov, V., & Goldfard, A. (1990) Blocking of the initiation to elongation transition by a transdominant RNA polymerase mutation. *Science* **248**: 1006-1009.

Keener, J. D., & Kustu, S. (1988). The protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NtrC of enteric bacteria: roles of the conserved N-terminal domain of NTRC. ***Proc Natl Acad Sci USA*** **85**: 4976-4980.

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

Kierkegard, K., Buc, H., Spasski, A., & Wang, J. C. (1983). Mapping of single-stranded regions in duplex DNA at the sequence level: Single-strand-specific citosine methylation in RNA polymerase-promoter complexes. *Biochemistry* **80**: 2544-2548.

Kimura, M., Fujita, N., Ishihama, A. (1994) Functional map of the alpha subunit of *Escherichia coli* RNA polymerase. *J Mol Biol.* **1994**: 242: 107-105.

Kimura, M., & Ishihama, A. (1995) Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: insertion analysis of the amino terminal assembly domain. *J Mol Biol* **216**: 756-767

Klenk, H. P., & Doolittle, W. F. (1994) Archaea and eucaryotes versus bacteria?. *Current Biol* **4**: 920-922.

Klose, K. E., North, A. K., Stedman, K. M., & 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.* **232**: 67-68.

Kobayashi, M. , Nagata, K., & ishihama, A. (1990) Promoter selectivity of *Escherichia coli* RNA polymerase: effect of base substitutions in the promoter -35 region on promoter strength. *Nucleic Acids Res* **18**: 7367-7372.-

Kolb, A., Igarashi, A., Ishihama. A., Lavigne, M., Buckle, M., & Buc, H. (1993) *E. coli* RNA polymerase, deleted in the C-terminal part of its α subunit, interacts differently with the cAMP-CRP complex at the *lac*P1 and at the *gal*P1 promoter. *Nucleic Acid Res* **21**: 319-326.

Koleske, A., & Young, R. A. 1994 An RNA polymerase II holoenzyme responsive to activators. *Nature* **368**: 466-469.

Kovacic, R. T (1987) The O° C closed complexes between *Escherichia coli* RNA polymerase and two promoters , T78-A3 and *lacUV5*. *J Biol Chem* **262**: 13654-13661

Krey, R., Pühler, A., & Klipp, W. (1992) A defined amino acid exchange close to the putative nucleotide binding site is responsible for an oxygen -tolerant variant of the *Rhizobium meliloti* NifA protein. *Mol Gen Genet* **234**: 433-441.

Kullik, I., Fritsche, S., Knobel, H., Sanjuan, J., Hennecke, H., & Fischer, H.M. (1991). *Bradyrhizobium japonicum* has two differntially regulated functional homologs of the σ^{54} (*rpoN*). *J Bacteriol* **173**: 1125-1138.

Kustu, S., Santero., E., Popham, D., & Keener, J. (1989) Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol Rev* 53: 367-376.

Kustu, S., North, A. K., & Weiss, D. S. (1991). Prokayotic transcriptional enhancers and enhancer-binging proteins. *Trends Biochem Sci* **16**: 397-402.

Kwok. , T., Yang, J., Pittard, A. J., Wilson, T. H., & Davidson, B. (1995) Analysis of an *Escherichia coli* mutant TyrR with impaired capacity for tyrosine-mediated repression, but still able to activate at σ^{70} promoters. *Mol Microbiol* 17: 471-481.

Lawley, B., Fujita, N., Ishihama, A., & Pittard, A. J. (1995) The TyrR protein of *Escherichia coli* is a class I transcription activator. *J Bacteriol* **177**: 238-241.

Lee, H. S. , Berger, D. K. & Kustu, S. (1994) Activity of purified NifA, a transcriptional activator of nitrogen fixation genes. *Proc Natl Acad Sci USA* 90: 2266-2270.

Lee, J. H., & Hoover, T. H. (1995) Protein crosslinking studies suggest that *Rhizobium melitoti* C4-dicarboxylic acid transport protein D, a σ^{54}-dependent transcriptional activator, interacts with σ^{54} and the β subunit of RNA polymerase. *Proc Natl Acad Sci USA*. 92: 9702-9706.

Liermo, S., & Gourse, R. L. (1991) Factor-independent activation of *Escherichia coli rrnA* transcription I. Kinetic analysis of the roles of the upstream activator region and supercoiling on transcription of the *rrnB* P1 promoter in vitro. *J Mol Biol* **220**: 555-568.

Leung, D. W., Chen, E., and Goeedel, D.V. (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* 1: 11-15.

Lesley, S, A. Brow, M.A.D., & Burguess, R. R. (1991) Use of *in vitro* protein synthesis from polymerase chain reaction-generated templates to study interaction of *Escherichia coli* transcription factors with core RNA polymerase and for epitope mapping of monoclonal antibodies. *J Biol Chem* **266**: 2632-2638.

Lesley, S, A. & Burguess, R. R. (1989) Characterization of the *Escherichia coli* transcription factor σ^{70} : localization of a region involved in the interaction with core RNA polymerase. *Biochemistry* **28**: 7728-7734.

Leuther, K. K., Salmeron, J., & Johnston, S. A. (1993) Genetic evidencve that an activation domain of GAL4 does not require acidicity and may form a b sgeet. *Cell* **72**: 575-585.

Li, M., Moyle, H., & Susskind, M. (1994) Target of the transcriptional activation function of phage λ cI protein. *Science* **263**: 75-77.

Lisser, S., & Margalit, H. (1993) Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res* **21**: 1507-1516.

Lobell, R., & Schleif, R. (1990) DNA looping and unlooping by the AraC protein*Science* **250**: 528.

Losick, R. & Pero, (1981) Cascades of sigma factors. *Cell* **25**: 1606-1614.

Lonneto, M., Gribskov, M., & Gross, C.A. (1992) The σ^{70} family: sequence conservation and evolutionary relationships. *J Bacteriol* 174: 3843-3849.

Martínez-Romero, E., & Caballero-Mellado, J. (1996) Rhyzobium phylogenies & bacterial genetic diversity. *Crit Rev Plant Sci* **15**: 113-140

McNeil, T., Roberts, G. P., McNeil, D., & Tyler, B. (1982) The products of *glnL* and *gl G* are bifunctional regulatory proteins. *Mol Gen Genet* 188: 325-333.

Mecsas, J. , Cowing, D. W., & Gross, C. A. (1991) Development of RNA polymerase-promoter contacts during open complex formation. *J.Mol. Biol.* **220**: 585-597.

Miller, J. H. (1972). *Experiments in molecular genetics.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Merino, E., Osuna, J., Bolívar, F., & Soberón, X. (1992). A general, PCR method for single or combinatorial oligonucleotide-directed mutagenesis on pUC/M13 vectors. *Biotechniques* 12: 508-509.

Merrick, M. J. (1992). Regulation of nitrogen fixation genes in free-living and symbiotic bacteria, p. 835-847. In G. Stacey, R. Burris and H. Evans (ed). Biological nitrogen Fixation. Chapman Hall, New York.

Merrick, M. J. (1993) In a class of its own -the RNA polymerase sigma factor σ54 (σN). *Mol Microbiol* 10: 903-909.

Merrick, M. J. & Chambers, S. (1992) The helix-turn helix motif of σ^{54} is involved in recognition of the -13 promoter region. *J Bacteriol* **174**: 7221-7226.

Mettke, I., Fielder, U., & Weiss, V. (1995) Mechanism of activation of a response regulator: interaction of NtrC-dimers induces ATPase activity. *J. Bacteriol.* **177**: 5056-5061.

Messing, J., (1981) A system for shotgun sequencing. *Nuc Acids Res* 9: 259.

Minchin, S. D., Austin, S. & Dixon, R. (1988). Transcriptional activation of the *Klebsiella pneumoniae nifLA* promoter by NTRC is face-of the helix dependent and the activator stabilises the interaction of sigma 54-RNA polymerase with the promoter. *EMBO J.* **8**: 3491-3499

Moras. D. (1993) Two sisters and their cousin. *Nature* **364**: 572-573.

Mustaev, A. M., Kashlev, M., Lee, Y., Polyakov, A., Lebedev, A., Zalenskaya, A., Grachev, M., Goldfard, A., & Nikiforov, V. (1991) Mapping of the priming substrate contacts in the active center of the *Escherichia coli* RNA polymerase. *J Biol Chem* **266**: 23927-23931.

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

Ninfa, A., Atkinson, M. R., Kamberov, E. S., Feng, J., & Ninfa, E. G. (1995) Control of nitrogen assimlation by the NRII-NRI two-component system of enterica bacteria. In Two-Component Signal Transuction (Hoch, J., A., & Silhavy, T. J., eds) pp 67-88. American Society for Microbiology, Washington, D. C.

Niu, W., Zhou, Y., Dong, Q., Y. W. Ebright, & Ebright, R. H. (1994) Characterization of the activating region of *Escherichia coli* catabolite activator protein (CAP). I. Saturation and alanine scanning mutagenesis. *J Mol Biol* **43**: 595-602.

Nixon, B. T., Ronson, C. W., & Ausubel, F. M. (1989) Two-component regulatory systems responsive to enviromental stimuli share strongly conserved domains with the nitrogen assimilation refulatory genes *ntrB* and *ntrC*. ***Proc Natl Acad Sci USA*** **83**: 7850-7854.

Ma, J., & Ptashne, M. (1987) A new class of transcriptional activators. *Cell* **51**: 113-119.

Malhotra, A., Severinova, E., & Darst, S. A. (1996) Crystal structure of σ^{70} subunit fragment from *E. coli* RNA polymerse. *Cell* **87**: 127-136.

Model, P.ç, Jovanovic, G., & Dworkin, J. (1997) The *Escherichia coli* phage-shock-protein (*psp*) operon. *Mol Microbiol* 24: 255-261.

Moore, J. B., Shiau, S. P., & Reitzer, L.J. (1993) Alterations of highly conserved residues in the regulatory domain of nitrogen regulator I (NtrC) of *Escherichia coli*. *J Bacteriol* **175**: 2692-2701.

Morett, E., & Buck, M. (1988) NifA-dependent *in vivo* protection demostrates that the upstream activator sequence of *nif* promoters is a protein binding site. *Proc Natl Acad Sci USA* 85: 9401-9405.

Morett, E. & Buck, M., (1989) *In vivo* studies on the interaction of RNA polymerase-σ^{54} with the *Klebsiella pneumoniae* and *Rhizobium meliloti nifH* promoters. The role of NifA in the formation of an open complex. *J Mol Biol* 210: 65-77.

Morett, E., Kreutzer, R., Cannon, W., & Buck, M. (1990). The influence of the *Klebsiella pneumoniae* gene *nifL* upon the transcriptional activator protein NifA. Mol. Microbiol. 4: 1253-1258.

Morett, E., Fisher, H. M., & Hennecke, H. (1991) Influence of oxygen on DNA binding, positive control and stability of the *Bradyrhizobium japonicum* NifA regulatory protein. *J Bacteriol* 173: 3478-3487.

Morett, E., & Segovia, L. (1993) The σ^{54} bacterial enhancer-binding protein family: Mechanism of action and phylogenetic relationship of their functional domains. *J Bacteriol* 175: 6067-6074.

Narberhaus, F., Lee, H.S., Scmitz, R. A., He, L., & Kustu, S. (1995) The C-terminal domain of NifL is sufficient to inhibit NifA activity *in vivo*. *J Bacteriol* **177**: 5078-5087.

Negishi, T., Fujita, N., & Ishihama, A. (1995) Structural map of the alpha subunit of *Escherichia coli* RNA polymersae: structural domains identified by proteolytic cleavaje. *J Mol Biol* **248**: 723-728.

Newlands, J. T., Ross, W., Gosink, K., & Gourse, R. L. (1991) Factor-independent activation of rRNA transcriptiom¿ II, Characterization of complexes of *rrnB* P1 promoters containing or lacking the upstream activator region with *E. coli* RNA polymerase. *J Mol Biol* **220**: 569-583.

North, A. K., Klose, K. E., Stedman, K. M., & Kustu, S. (1993) Prokaryotic enhancer-binding protein reflect eucaryotic like modularity: the puzzle of nitrogen regulatory protein C. *J Bacteriol.* 175: 4267-4273.

North, A. K., Weiss, D.S., Susuki, H., Flashner, Y., & Kustu, S. (1996) Repressor forms of the enhancer binding protein NtrC: Some fail in coupling ATP hydrolysis to open complex formation by σ^{54}-holoenzyme. *J Mol Biol* 26: 317-331.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, G., & Steitz, T. A. (1985) Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**: 762-766.

Oliphant, A., & Struhl, K. (1988). Defining the consensus sequences of *E. coli* promoter elements by random selection. *Nucleic Acids Res* **15**: 7673-7682.

Ohnishi, K. (1985) Amino acid and nucleotide sequence homologies among *E. coli* RNA polymerase core enzyme subunits, DNA primase, elongation factor Tu, F1-ATPase a, ribosomal protebn L3, DNA polymersae I, T7 DNA polymerse, and MS2 phage RNA replicase β subunit. *Nucleic Acid Sym. Ser.* **22**: 253-256.

Osuna, J., Soberón X., & E. Morett. (1997) A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. *Protein Science* 6: 543-555.

Ow, D. W., Xiong, Y., Gu, Q., & Shen, S. C. (1985) Mutational analysis of the *Klebsiella pneumoniae* nitrogenase promoter: sequences essential for positive control by *nifA* and *ntrC* (*gnl G*) products. *J Bacteriol* **161**: 868-874.

Pabo, C. O. (1990). *Science* **248**: 1210-1213.

Parkinson, J. S., & Kofoid, E. C. (1992) Communication modules in bacterial signaling proteins. *Ann Rev Genet* **26**: 71-112.

Ptashne, M. (1986). Gene regulation by proteins acting nerby and at a distance. Nature 322: 697-701.

Ptashne, M. (199). A genetic Switch.

Ptashne, M., & Gann, A. (1997) Transcriptional activation by recruitment. *Nature* **386**: 569-577.

Pao, G.M., & Saier, Jr., M.H. (1995) Response regulator of bacterial signal transducing systems: selective domain shuffling during evolution. ***J Mol Evol*** **40:136-154.**

Pérez-Martìn, J., & de Lorenzo, V. (1996) ATP binding to the σ54-dependent activator XylR triggers a protein multimerization cycle catalyzed by UAS. ***Cell*** **86: 331-339.**

Pittard, J., & Davidson, J. (1991) TyrR protein of *Escherichia coli* and its role as repressor and activator. ***Mol Microbiol*** **5: 1585-1592.**

Polyakov, A., Severinova, E., & Darst, S. A. (1993) Three dimensional structure of *E. coli* RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell* **83**: 365-373.

Porter, S. C., North, A. K., Wedel, A. B., & Kustu, S. (1993) Oligomerization of NTRC at the *gln* A enhancer is required for transcriptional activation. ***Genes Dev*** **7: 2258-2273.**

Popham, D. L, Szeto, D., Keener, J., & Kustu, S. (1989) Function of a bacterial activator protein that binds to transcriptional enhancers. ***Science*** **243: 629, 635.**

Quinto, C., de la Vega, H., Flores, M., Fernández, L. , Ballado, T., Soberón, G., & Palacios, R. (1983) Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli. Nature* **299**: 724-726.

Rao, L., Ross, W., Appleman, J. A., Gaal, T., Leirmo, S., Schlax, P. J., Record. T. M., & Gourse, R. L. (1994) Factor-independent activation of *rrnB* P1: an "extended" promoter with an upstream element that dramatically increases promoter strenght. *J Mol Biol* **235**: 1421-1435.

Reitzer, L., & Magasanik, B. (1986) Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. *Cell* **45**: 785-792.

Reitzer, L., Movsas, B., & Magasanik, B., (1989) Activation of *glnA* transcription by nitrogen regulator I (NR_I)-phosphate in *Escherichia coli*: evidence for long-range physical interaction between NR_I-phosphate and RNA polymerase. *J. Bacteriol.* **171**: 5512-5522.

Regier, J. L., Shen, F., & Treizenberg, S. J. (1993) Pattern of aromatic and hydrophobic amino acids critical for one of the VP16 trasncriptional activator. ***Proc Natl Acad Sci USA*** **90**: 883-887.

Ren, Y. L., Garges, S., Adhya, S., & Krakow, S. (1988) Cooperative binding of heterologous proteins: evidence for contact between the cyclic AMP receptor protein and RNA polymerase. ***Proc Natl Acad Sci USA*** **85**: 4138-4142.

Reznikoff, W. S. , Siegele, D. A., Cowing, D. W., & Gross, C. A. (1985) The regulation of trannscription initiation in bacteria. *Ann Rev Genet* **19**: 355-387.

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

Rippe, K., Von Hippel, P., Langowski, J. (1995) Action at a distance : DNA-looping and initiation of transcription. *Trends Biochem Sci* **20**: 500-506.

Rossi, J. J., Soberón, X., Marumoto, Y., McMahon, J., & Itakura, K. (1983). Biological espression of an Escherichia coli consensus promoter and some mutant derivatives. ***Proc Natl Acad Sci USA*** **80**: 3203-3207.

Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., & Gourse, R. (1993). A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science* 262: 1407-1413.

Saraste, M., Sibbald, P. R., & Wittinghofer, A. (1990) The P-loop a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* 15: 430-434.

Sasse-Dwigth, S., & Gralla, J. D. (1988) Probing the Escherichia coli glnALG upstream activation mechanism in vivo. ***Proc Natl Acad Sci USA*** **85**: 8934-8938.

Sasse-Dwigth, S., & Gralla, J. D. (1990). Role of eukaryotic -type functional domains found in the prokaryotic enhancer receptor sigma 54. *Cell* **62**: 945-954.

Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90°C. *Science* **253**: 1001-1007.

Schultz, P., Celia, H., Riva., M., Sentenac, A., & Oudet, P., (1993) Three-dimensional structure of yeast RNA polymerase I determined by electron microscopy of two dimensional crystals. *EMBO J.* **12**: 2601-2607.

Schuler, M. F., Tatti, K. M., Wade, K. H., & Moran, C, P. (1995) A single amino acid substitution in σ^E affects its ability to bind core RNA polymerase. *J Bacteriol* **177**: 3687-3694.

Shingler, V., & Pavel, H. (1995) Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds. *Mol Microbiol* **17**: 505-513.

Shingler, V., & Moore, T. (1994) Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas* sp strain CF600. ***J. Bacteriol*** **176**: 1555-1560

Siegele, D., Hu, J., Walter, W., & Gross, C. (1989) Altered promoter recognition by mutant forms of the s70 subunit of *Escherichia coli* RNA polymerase. *J Mol Biol* **206**: 5901-603.

Sienbelist, U. (1979) RNA polymerase unwinds an 11 base pair of a phage T7 promoter. *Nature* **279**: 269-281.

Straney, D. C.., Straney, S. B., & Crothers, D. M. (1989) Synergy between *Escherichia coli* CSP protein and RNA polymersae at the *Escherichia coli lac* UV5 promoter. *J Mol Biol* 206:41-47

Sousa, R., Chung. Y. T., Rose, J. P., & Wang, B. C. (1993) Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Åresolution. *Nature* **364**: 593-599.

Spassky. A., Kierkegaard, K., & Buc, H. (1985). Changes in the DNA structure of the *lac UV5* promoter during formation of an open complex with *Escherichia coli* RNA polymerase. *Biochemistry* **24**: 2723-2731.

Shariff, K., Fujita, N., Jin, R., Igarashi. K., & Ishihama, A. (1994) Epitope mapping and functional characterization of monoclonal antibodies specific for the alpha subunit of *Escherichia coli* RNA polymerase. *J Biol Chem* **1994**: 23655-23660.

Shingler, V. (1996) Signal sensing by σ^{54} dependent regulators: derepression as a control mechanism. ***Mol Microbiol*** **19: 409-416.**

Schmitz, R., He, L., & Kustu, S. (1996) Iron is requires to relieve inhibitory effects of NifL on transcriptional activation by NifA in *Klebsiella pneumoniae. J Bacterio.l* **178**: 4679-4687.

Soisson, S. M., MacDougall-Shackleton, B., Schleif, & Wolberger, C. (1997) Structural basis for ligand-induced oligomerization of AraC. *Science* **276**: 421-425.

Souillard, N., & Sibold, L., (1986) Primary structure and expression of a gene homologous to nifH (nitrogenase Fe protein) from archeobacterium *Methanococcus voltae. Mol Gen Genet* **203**: 21-28.

Stock, A. M., Chen. T., Welsh, D., & Stock. (1988). CheY protein, a central regulator of bacterial chemotaxis, belongs to a family of proteins that controls gene expression in response to changing enviromental conditions. *Proc Natl Acad Sci USA* **85:**1403-1407

Stouten, F. W., Sander, C., Wittinghofer, A., & Valencia, A. (1993). How does the switch II region of G-domains work. ***FEBS*** **1: 1-6.**

Su, W., Porter, S, Kustu, S, & Echols, H. (1990) DNA-looping and enhanecr activity. Asssotiation between DNA-bound NtrC activator and RNA polymerase at the bacterial *gln* A promoter. ***Proc Natl Acad Sci USA*** **87: 5504-5508.**

Suh, W. C., Ros, W. & Record, T. M. (1993) Two open complexes and a requirement of Mg^{2+} to open the λ pR transcription start site. *Science* **259**: 358-361.

Syed, A., & Gralla, J. D. (1997) Isolation and properties of enhancer bypass mutants of sigma 54. *Mol Microbiol* **23**: 987-995.

Travers, A. A., & Burguess, R. R. (1969) Cyclic reuse of sigma factor. *Nature* **222**: 527-540.

Thony, B, & Hennecke, H. (1989). The -24 /-12 promote comes of age. *FEMS Microbiol Rev.* 63: 341-358.

Tintut, Y., Wong, C. Jiang, Y., Hsieh, M., & Gralla, J. D. (1994) RNA polymerase binding using a strongly acidic hydrophobic -repeat region of s54. *Proc Natl Acad Sci USA* 91: 2120-2124.

Tintut, Y., & Gralla, J. D. (1995) PCR-mutagenesis identifies a polymerase-binding sequence of sigma 54 that includes a sigma 70 homology region. *J. Bacteriol* **177**: 58818-5825.

Travormina, P., Reznikoff, W., & Grosss, C.A. (1996). Indentifying interacting regions in the *b* subunit of *Escherichia coli* RNA polymerase. *J Mol Biol* 258: 213-223.

Tuli, R., & Merrick, M. J. (1988). Over-production and chracterization of the *nif A* gene product of *Klebsiella pneumoniae* -the transcriptional activator of nif gene expression. *J Gen Microbiol* 134: 425-432.

Valderrama, B., Dávalos, A., Girard, M. L., Morett, E., & Mora, J. (1996) Regulatory proteins and *cis*-acting elementes involved in the transcriptional control of *Rhizobium etli* reiterated *nifH* genes. *J Bacteriol* **178**: 3119-3126.

Van Hoy, M., Leuther, K. K., Kodadek, T., & Johnston, S. (1993). The acidic activation domains of the GCN4 and GAL4 proteins are not α helical but β sheets. *Cell* **72**: 584-597.

Von Hippel, P. H., Bear, D. G., Morgan, W. D., & McSwiggen, J. A. (1984) Protein-nucleic acid interactions in transcription: a molecular analysis. *Ann Rev Biochem* **53**: 389-446.

Waldburger, C., Gardella, T., Wong, R., & Susskind, M. (1990) Changes in the conserved region 2 of *Escherichia coli* σ^{70} affecting promoter recognition. *J Mol Biol* **215**: 267-276.

Waldburger, C., & Susskind, M. (1994) Probing the informational content of region *Escherichia coli* σ^{70} region 2.3 by combinatorial cassette mutagenesis. *J Mol Biol* **235**:1489-1500.

Wang, J. C., & Giaever, G. N. (1988) Action at a distance along DNA. *Science* **240**: 300-304.

Wang, J., Syed, A., Hsieh, M., & Gralla, J. D. (1995). Coverting Escherichia coli RNA polymerase into an enhancer-responsive enzyme: role of an NH_2.terminal acidic hydrophobic -repeat region of σ^{54}. *Science* **270**: 992-994.

Wang, R., & Kushner, S. R., (1991). Construction of versatiles low copy-number vector for cloning, sequencing and gene expression in *Escherichia coli. Gene* 100: 195-199.

Wang, Y., & Hoover, T. R. (1997). Alterations whitin the activation domain of the σ^{54} -dependent activator DctD that prevent transcriptional activation. *J. Bacteriol.* 179: 5818-5819.

Wang, Y., Lee, J., Brewer, J. M., & Hoover, J. M. (1997). A conserved region in the σ^{54}-dependent activator DctD is involved in both binding to RNA polymerase and coupling ATP hydrolysis to activation. *Mol Microbiol* **26**: 373-386.

Wang. X. Y., Kolb, A., Cannon,.W., & Buck, M. (1997). Nucleoprotein complex formation by the enhancer binding protein NifA. *Nuc Acids Res* **25**: 3478-3485.

Walker, J. H., Saraste, M., Runswick, M. J., & Gay, N. J. (1982). Distantly related sequences in the α and β subunits of ATP synthase, myosin, kinases and the ATP requering enzymes and a common nucleotide binding fold. *EMBO J* 1: 945-951.

Warne, S. E. & de Haseth, P. L. (1993) Promoter recognition by *Escherichia coli* RNA polymerase. Effects of single pair deletions and inserions in the spacer DNA separating -10 and -35 regions are dependent on spacer DNA sequence. *Biochemistry* **32**: 6134-6140

Wedel, A. B., Weiss, D., Popham, D., Droge, P., & Kustu, S., (1990) A bacterial enhancer functions to tehter a transcriptional activator near a pomoter. *Science* **248**: 486-490.

Wedel, A. B., & Kustu, S. (1995). The bacterial enhancer-binding protein NtrC is a molecular machine: ATP hydrolisis is coupled to transcriptional activation. *Genes Dev* **9**: 2042-2052.

Wei, R. G. & Kustu, S. (1981) Glutamine auxothrophs with mutations in a nitrogen regulatory gene, *ntrC*, that is near *glnA*. *Mol Gen Genet* 183: 392-399.

Weilbaecher, R., Hebron, C., Feng, C., & Landick, R. (1994) Termination-altering amino acid substitutions in the β′ subunit of *Escherichia coli* RNA polymerase identify regions involved in RNA chain elongation. *Genes Dev* **8**: 2913-2927.

Weiss, D., Batut, J., Klose, K. E., Keener, J. & 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.

Weiss, V., Claverie-Martín, F., & Magasanik, B. (1992). Phosphorylation of nitrogen regulator I of *Escherichia coli* induces strong cooperative binding to DNA essential for activation of transcription. *Proc. Natl. Acad. Sci. USA.* **89**: 5088-5092.

Williams, R., Bell, A., Sims, G., & Busby, S. (1991). The role of two surface exposed loops in transcription activation by the Escherichia coli CRP and FNR proteins. *Nucleic Acids Res* **19**: 6705-6712.

Wilson, T. J., Maroudas, P., Howlet, G. J., & Davidson, B. E. (1994) Ligand-induced self- association of the Escherichia coli regulatory protein TyR. *J Mol Biol* 238: 309-318.

Wilson. C., & Dombroski, A. J. (1997) Region 1 of σ^{70} is required for efficient isomerization and initiation of transcription by *Escherichia coli* RNA polymerase. *J Mol Biol* **267**: 60-74.

Wittinghofer, A., & Nassar, N. (1996) How ras-related proteins talk to their effectors. *Trend Bioch Sci* 21: 488-491.

Woese, C. R., Kandler, O., Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria and Eucarya. *Proc. Natl. Acad. Sci. USA* **87**: 4576-4579.

Wong, C., Tintut, Y., & 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.

Wootton, J, C., & Drummond, M. (1989). The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Protein Eng* **2**: 535-543.

Wyman, C., Rombel, I., North, A. K., Bustamante, C., & Kustu, S. (1997) Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science* 275: 1658- 1661.

Yang, J., Camarakis, H., & Pittard, A. J. (1996) Further analysis of the activation function of the TyrR regulatory protein of *Escherichia coli. J.Bacteriol* **178**: 1120-1125

Young, J. P. W. (1992). Phylogenetic clasification of nitrogen -fixing organisms. In Biological nitrogen fixation (Stacey, G., Evans, H. J., & Burris, R. H., eds). p. 43-86. Chapman & Hall, New York

Zhang, X., Zhou, Y., Ebright, Y. W., & Ebright, R. H. (1992) Catabolite gene activator protein (CAP) is not an "acidic activating region" transcription activator protein. *J Biol Chem* **167**: 8136-8139.

Zhou, Y. & Gross, C. (1992) How a mutation in the gene encoding σ^{70} supressed the defective heat shock response caused by a mutation in the gene encoding σ^{32}. *J Bacteriol* **174**: 7128-7137.

Zhou, Y., Zhang, X., & Ebright, R. H. (1993) Identification of the activation region of catabolite activator protein (CAP): isolation and characterization of mutants of CAP specifically defective in transcription activation. *Proc. Natl. Acad. Sci. USA*. 90: 6081-6085.

Zhou, Y., Busby, S., & Ebrigtht, R. H. (1993) Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. **Cell 73**: 375-379.

Zhou, Y., Shanon Pendergrast, P., Bell, A., Williams, A., Busby, S., & Ebrigtht, R. H. (1994) The functional subunit of a dimeric transcription activbator protein depends on promoter architecture. *EMBO J* **13**: 4549-4557.

Zhou, Y., Busby, S., & Ebrigtht, R. H. (1993)

Zhou, Y., Merkel, T. J., & Ebright, R. H, (1994) Characterization of the activating region of Escherichia coli gene activator protein (CAP) II. Role at Class I and Class II CAP-dependent promoters. J Mol Biol 243: 603-610.

Zillig, W., Palm, P., & Heil, A. (1976) Function and reassambly of subunits of DNA-dependent RNA polymerase. In RNA polymerase (Losick, R. & Chamberlin, R., eds). Cold Spring Harbor Laboratory, Cold Spring Harbor , New York.

Gratitudes

A mi papá Fidencio y a mi mamá Tolla. A mis hermanos y mis tíos Rodolfo, Sofía y Evaristo. A mis amigos, "*landscapes & adaptative surfaces*". . .