**CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS**
**ESTACIÓN EXPERIMENTAL DEL ZAIDÍN**



**UNIVERSIDAD DE GRANADA**
**FACULTAD DE CIENCIAS**

# INTERACCIONES MOLECULARES DEL REGULADOR XylS Y LA ARN POLIMERASA CON EL PROMOTOR Pm DEL PLÁSMIDO TOL pWW0 DE *Pseudomonas putida*

Tesis Doctoral

**Patricia Domínguez Cuevas**
**2007**

Editor: Editorial de la Universidad de Granada
Autor: Patricia Domínguez Cuevas
D.L.: Gr. 1437 - 2007
ISBN: 978-84-338-4384-5

# INTERACCIONES MOLECULARES DEL REGULADOR XylS Y LA ARN POLIMERASA CON EL PROMOTOR Pm DEL PLÁSMIDO TOL pWW0 DE *Pseudomonas putida*

Memoria que presenta la licenciada
en Biología, Patricia Domínguez Cuevas,
para aspirar al Título de Doctor

Fdo.: Patricia Domínguez Cuevas

Vº Bº del Director

Fdo.: Juan Luis Ramos Martín
Doctor en Biología
Profesor de investigación del C.S.I.C.

Vº Bº de la Directora

Fdo.: Silvia Marqués Martín
Doctora en Biología
Investigadora Científica del C.S.I.C.

Universidad de Granada
2007

Esta Tesis Doctoral ha sido realizada en el
Departamento de Protección Ambiental de la
Estación Experimental del Zaidín (C.S.I.C.), Granada.

*A mis abuelas*
*A mis padres y hermana*
*A Chechu*

*"Si he conseguido ver más lejos,*
*es porque me he aupado en hombros de gigantes"*

*Isaac Newton*

*"El científico encuentra su recompensa en lo que*
*Henri Poincare llama el placer de la comprensión,*
*y no en las posibilidades de aplicación que*
*cualquier descubrimiento pueda conllevar"*

*Albert Einstein*

# *AGRADECIMIENTOS*

*Al Dr. Juan Luis Ramos, director del grupo de Degradación de Tóxicos Orgánicos de la Estación Experimental del Zaidín, y codirector de esta Tesis Doctoral; gracias por haberme permitido trabajar bajo tu dirección y por saber encontrar siempre el lado positivo de los resultados.*

*A la Dra. Silvia Marqués, codirectora de esta tesis doctoral, le agradezco haber confiado en mí desde el principio, formándome no solo científicamente sino como persona. Gracias Silvia por haber sido más que una "jefa", por respetar mis ideas y estar siempre dispuesta a discutirlas. Sin ti no lo habría conseguido.*

*A mis padres y hermana, de los que siempre he recibido un apoyo incondicional. Gracias papá y mamá por haberme inculcado la satisfacción que deja un trabajo bien hecho. A mi hermana por ser mi cómplice y la persona que mejor me conoce.*

*A mis niñas Ana (hola fea!!), Rebeca y Amada y mis chicos Jose Ángel y Antonio por demostrarme cada día lo que es la verdadera amistad, esa que perdura en el tiempo y la distancia.*

*A mis hermanitos del lab 2.6: Patri M, no tengo palabras, ni espacio suficiente para valorar tu infinita capacidad de trabajo y la inteligencia que posees, pero aún más importante para mí ha sido disfrutar de tu amistad, tu comprensión y complicidad a lo largo de estos años. A Alejo y Javi, mis dos hermanos mayores, que me hacen rabiar de vez en cuando, pero con los que siempre puedo contar, gracias .A la casi recién llegada, Águeda, que me ha hecho darme cuenta que ser "tan apretá" no puede ser bueno para la salud. Y a mi Paola!!, que no me olvido de ti, la niña más dulce de la EEZ, no cambies nunca tu forma de tratar con la gente.*

*A Euge, por su compresión y su ánimo que tanto me ha ayudado en los momentos difíciles, pero también y aún más importante por compartir alegrías y las ganas de hacer que la vida valga la pena. A mi Jesusito, mi niño desde que llegó al lab, por tu amistad sincera; me encanta esa manera tuya de tomarte la vida. A Wilson, por tu eterna sonrisa y por estar siempre dispuesto a discutir resultados y proponer soluciones.*

*A Teresa y Sandy a las que admiro por su transparencia y por su generosidad. A Antonio Jesús que siempre está dispuesto a resolver nuestros desastres informáticos sin perder su sonrisa maravillosa ni su risa "escandalosa".*

*To my Bham family, la resistencia extranjera, por vuestra cálida acogida y por hacerme sentir muy especial y muy querida. No hace falta mucho tiempo para conectar con gente tan especial como vosotros. I will never forget you!!*

*Al resto de mis compañeros dentro y fuera del grupo, los pasados y los presentes, que habida cuenta de los años que llevo en el "Zaidín", forman una lista tan larga de excelentes personas que me es imposible citarlos a todos. Todos vosotros habéis hecho de este tiempo una etapa para recordar.*

*A los Drs. Víctor de Lorenzo y Stephen Busby, por acogerme en sus respectivos laboratorios, lo que me ha permitido aprender y avanzar mucho más rápido en esta Tesis. Víctor, gracias además por tú amistad. Al Dr. José Eduardo González-Pastor, por hacer de mi estancia en Madrid una visita de placer más que de trabajo y por nuestras largas conversaciones telefónicas, de las que tanto disfruto.*

*A Chechu, al que nunca podré agradecerle lo suficiente todo su apoyo y las innumerables horas de trabajo. Gracias por creer en mí más de lo que yo lo hago, por ser tan generoso y por compartir conmigo tú pasión por el conocimiento. Te admiro mucho y te quiero aún más.*

íNDICE

# ÍNDICE GENERAL

| ABREVIATURAS | | ABBREVIATIONS | |
|---|---|---|---|
| 3MB | 3-metilbenzoato | 3MB | 3-methylbenzoate |
| - | - | Ap | ampicillin |
| ARNm | ARN mensajero | - | - |
| ARNP | ARN polimerasa | RNAP | RNA polymerase |
| - | - | BSA | bovine serum albumin |
| ChIP | inmunoprecipitación de cromatina | ChIP | chromatin immunoprecipitation |
| dNTPs | desoxinucleótidos | dNTPs | desoxynucleotides |
| E | núcleo de la ARNP | E | RNAP core |
| - | - | EDTA | ethylenediamine tetraacetic acid |
| - | - | EMSA | electrophoretic mobility shift assay |
| Fig | figura | Fig | figure |
| - | - | Gm | gentamycin |
| HTH | motivo hélice-giro-hélice | HTH | helix-turn-helix motif |
| - | - | IPTG | isopropyl-β-D-thiogalactopyranóside |
| - | - | Km | kanamycin |
| min | minuto | min | minute |
| NTPs | nucleótidos | NTPs | nucleotides |
| pb | par(es) de bases | bp | base pair(s) |
| PCR | reacción en cadena de la polimerasa | PCR | polymerase chain reaction |
| - | - | Rif | rifampicin |
| $RP_O$ | complejo promotor abierto | $RP_O$ | open complex |
| $RP_C$ | complejo promotor cerrado | $RP_C$ | closed complex |
| - | - | SDS | sodium dodecyl sulphate |
| - | - | Tc | tetracycline |
| - | - | X-gal | 5-bromo-4-chloro-3-indolyl-β-D-galactoside |

-

# ÍNDICE DE FIGURAS

**PÁG**

**INTRODUCCIÓN**

**PÁG**

**CAPÍTULO I**



**CAPÍTULO II**



**CAPÍTULO III**

**CAPÍTULO IV**



**CAPÍTULO V**

**SUPPLEMENTAL MATERIAL**

# ÍNDICE DE TABLAS

**SUPPLEMENTAL MATERIAL**

# INTRODUCCIÓN GENERAL

## REGULACIÓN DE LA TRANSCRIPCIÓN EN PROCARIOTAS.

Tras la resolución de la estructura tridimensional del ADN en 1953 (Watson y Crick, 1953), el estudio de la regulación del flujo de información genética desde el ADN hasta las proteínas, proceso conocido como expresión génica, ha constituido un tema central de investigación en biología. El control de la expresión génica resulta esencial en todos los procesos de adaptación y diferenciación que tienen lugar en los organismos vivos. En palabras de François Jacob y Jacques Monod:

*"According to the strictly structural concept, the genome is considered as a mosaic of independent molecular blueprints for the building of individual cellular constituents. In the execution of these plans, however, coordination is evidently of absolute survival value. The discovery of regulator and operator genes, and of repressive regulation of the activity of structural genes, reveals that the genome contains not only a series of blueprints, but a coordinated programme of protein synthesis and the means of controlling its execution." (Jacob and Monod, 1961).*

La transcripción constituye el primer paso en el proceso de expresión génica y uno de los puntos principales sujetos a regulación. El estudio de las respuestas adaptativas en bacterias ha llevado al descubrimiento de un abanico de reguladores que modulan la actividad de los promotores, ajustando la expresión de genes específicos a cambios en el ambiente.

La presente memoria de Tesis Doctoral versa acerca del estudio de la regulación de la transcripción del operón de la ruta *meta* del plásmido TOL pWW0 de *Pseudomonas putida*, que codifica los enzimas necesarios para el catabolismo de alquilbenzoatos. El regulador transcripcional XylS, que pertenece a la familia AraC, activa la transcripción desde el promotor del operón *meta* en respuesta a la presencia de benzoato y derivados. En dicha regulación se distinguen dos niveles principales, uno global en el que se integran las diferentes respuestas de *P. putida* a la presencia de compuestos aromáticos, y un nivel de regulación específico sobre el promotor del operón *meta* en el que intervienen varios factores de transcripción. Esta introducción pretende revisar el conocimiento actual sobre el proceso de transcripción en procariotas y los distintos niveles de regulación a los que está sometido, centrándonos en aquellos aspectos o mecanismos más relacionados con la regulación del operón *meta*.

# 1. LA TRANSCRIPCIÓN

La transcripción es el proceso mediante el cual la información contenida en el ADN es transcrita por la ARN polimerasa a una molécula de ARN de secuencia complementaria al ADN molde. La transcripción se divide en tres etapas principales: iniciación, elongación y terminación, que serán descritas en detalle a lo largo de esta introducción.

## 1.1. LA ARN POLIMERASA

La ARN polimerasa dependiente de ADN es la principal enzima implicada en el proceso de transcripción, y por lo tanto, la diana final sobre la que actúan la mayoría de los mecanismos de regulación de la expresión génica. La ARN polimerasa (ARNP) u holoenzima bacteriana está compuesta por un núcleo enzimático y un factor sigma. El núcleo enzimático, con un peso molecular relativo de aproximadamente 400.000 Da, consta de cinco subunidades ($\alpha_2\beta\beta'\omega$) conservadas evolutivamente en secuencia, estructura y función. El núcleo enzimático presenta actividad catalítica de elongación de la molécula de ARN, pero es incapaz de iniciar la transcripción de forma eficiente y específica, de manera que necesita asociarse a un factor sigma (σ) para formar la ARNP u holoenzima y adquirir de este modo la capacidad de reconocer secuencias específicas en el ADN, denominadas promotores.

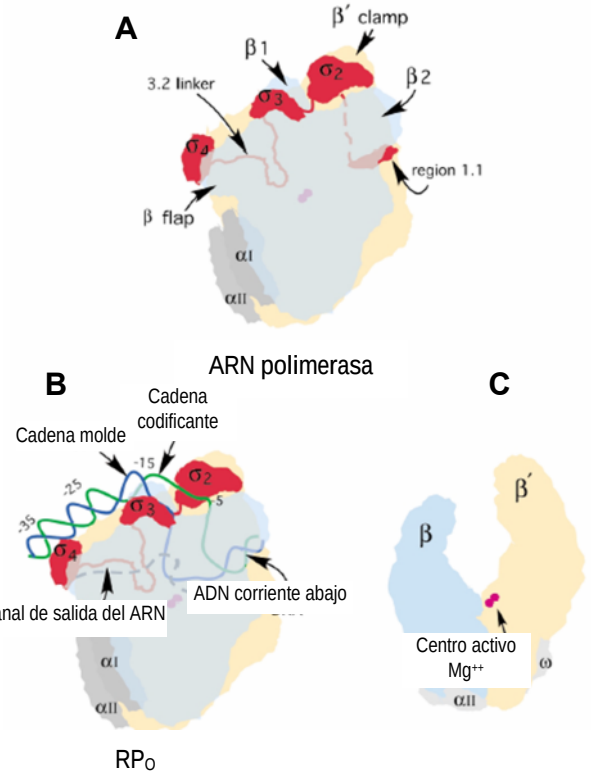



**Figura 1. Esquema de la estructura de la ARN polimerasa asociada a la subunidad sigma. A.** Estructura tridimensional de la ARNP. En la imagen se observa en rojo la estructura modular del factor sigma. **B.** Estructura de la ARNP formando parte de un complejo promotor abierto ($RP_O$) en el que la doble cadena de ADN forma una burbuja de transcripción entre las posiciones -12 y +2. **C.** Vista lateral del núcleo de la ARNP en la que se observan los dos iónes $Mg^{++}$ del centro activo asociados a la subunidad β'. Las distintas regiones de la subunidad sigma y del resto de subunidades de la ARNP se indican con sus respectivos nombres. El código de color es: rojo, σ; celeste, β; beige, β'; gris, αI, αII y ω. Obtenida de Hsu *et al*. 2002.

La ARN polimerasa bacteriana de *Thermus thermophilus* ha sido cristalizada y su estructura tridimensional resuelta con una resolución de 2,6 Å (Vassylyev *et al*., 2002). El núcleo adopta una estructura denominada "pinza de cangrejo" formada por las subunidades β y β' (Figura 1.C.) similar a la de la ARN polimerasa II de levaduras (Fu *et al*., 1999). El centro activo de la ARNP se localiza entre las subunidades β y β' e incluye los determinantes necesarios para la unión del ADN molde y del ARN generado durante la transcripción. El centro activo está constituido por tres residuos de aspártico de la subunidad β', que forman parte de la secuencia conservada (NADFDGD). Dichos residuos interaccionan con dos iones magnesio esenciales para el mecanismo de funcionamiento propuesto para todas las polimerasas de ácidos nucleicos descritas hasta la fecha (Sosunov *et al*., 2005).

Cada una de las dos subunidades α idénticas consta de dos dominios plegados de forma independiente que permanecen unidos por una conexión flexible de unos 20 aminoácidos de longitud. El dominio amino terminal de la subunidad α de la ARN polimerasa (aminoácidos 1-235) presenta los determinantes para la dimerización y participa en el ensamblaje de las subunidades β y β'. El dominio carboxilo terminal de la subunidad α es más pequeño (aminoácidos 250-329) y juega un papel importante en el reconocimiento de secuencias de ADN en ciertos promotores (ver apartado 1.4.). En cuanto a la subunidad ω, su pertenencia al complejo multiproteico de la ARNP es aún motivo de controversia. En un principio, se la consideró como parte de la ARNP debido a su co-purificación en cantidades estequiométricas con el núcleo, aunque no parecía jugar un papel directo en el proceso de transcripción, al ser la única subunidad prescindible de la ARNP (Burgess, 1971; Heil y Zillig, 1970). Sin embargo, resultados recientes apuntan a que la subunidad funciona como una chaperona, ayudando al correcto plegamiento y ensamblaje de la subunidad β' (Ghosh *et al*., 2001; Mathew y Chatterji, 2006; Mathew *et al*., 2005) e incluso podría cumplir alguna función en la respuesta estricta (del inglés "stringent") (Vrentas *et al*., 2005).

El ensamblaje de las subunidades de la ARNP sigue una secuencia específica, en la que en primer lugar se produce la asociación del dímero de subunidades α con la subunidad β. La subunidad ω mantiene a la subunidad β' soluble, ayudando al ensamblaje de la misma al complejo $\alpha_2\beta$. Por último el núcleo de la ARNP se asocia con un factor sigma para formar la holoenzima, capaz de reconocer secuencias promotoras y de iniciar la transcripción. El ensamblaje de las diferentes subunidades conlleva la formación de tres canales proteicos principales en la ARNP, cuya función es permitir la entrada del ADN molde de cadena sencilla al centro activo, la salida del ARN recién sintetizado y la entrada de los substratos (NTPs) al centro activo.

## 1.2. ETAPAS DEL PROCESO DE TRANSCRIPCIÓN

### 1.2.1. INICIACIÓN

Durante la iniciación, la holoenzima reconoce secuencias específicas de ADN denominadas promotores formando un complejo promotor cerrado ($RP_C$). A continuación, la ARN polimerasa desenrolla la doble cadena de ADN en la región -10 generando un complejo promotor abierto ($RP_O$) que, en presencia de nucleótidos trifosfato, inicia la transcripción. Una vez que la cadena de ARN ha alcanzado una longitud de 9-12 nucleótidos, de los cuales 8-9 están aún apareados con el molde de ADN formando un híbrido ADN-ARN, el complejo de transcripción pasa a la fase de elongación. A continuación, se describen en detalle cada una de las fases en que se divide la iniciación de la transcripción.

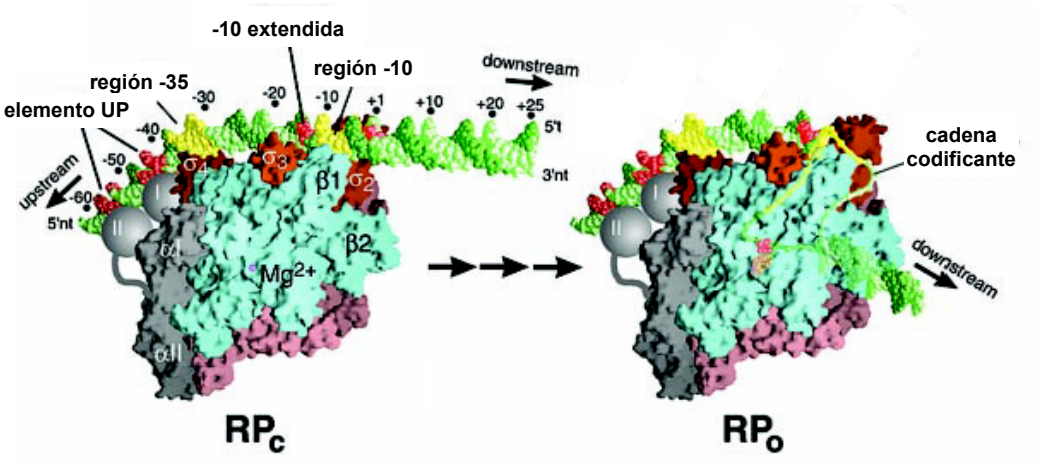



**Figura 2. Modelo de los complejos promotores cerrado ($RP_C$) y abierto ($RP_O$).** $RP_C$: la ARNP reconoce los elementos -10, -10 extendido, -35, y "UP" del promotor, estableciendo interacciones a través de las regiones 2.4, 3, 4.2 de la subunidad sigma y el dominio carboxilo terminal de la subunidad α, respectivamente. $RP_O$: la ARNP desenrolla la doble cadena de ADN, generando una burbuja de transcripción entre las posiciones -12 y +2 con respecto al punto de inicio de la transcripción, de manera que la cadena molde es dirigida hacia el centro activo de la ARNP. Código colores: verde, ADN; rojo, σ; celeste, β; magenta, β'; morado, ω; gris, αI y αII. Obtenido de Murakami *et al*. 2002.

**a.** Reconocimiento del promotor y formación del complejo promotor cerrado

La interacción de la ARNP con una región promotora en el ADN es el primer paso en la secuencia de eventos que llevan a la formación del complejo abierto y que culminan con la formación de un complejo de elongación. Se han identificado cuatro elementos en los promotores, que determinan el reconocimiento por la ARNP (Figura 2). Los dos elementos principales son dos hexámeros situados en posiciones -10 y -35 con respecto al punto de inicio de la transcripción: la región -10

del promotor es reconocida por el dominio 2.4 de la subunidad sigma de la ARNP, mientras que el dominio 4.2 de la misma subunidad reconoce la región -35 (Murakami *et al*., 2002a). Las secuencias consenso para ambos hexámeros dependen del factor sigma específico que reconoce cada promotor (Lisser y Margalit, 1993; McLean *et al*., 1997). La región -10 extendida y el elemento "UP" (del inglés upstream) constituyen los otros dos elementos del promotor importantes para el reconocimiento de ciertos promotores por la ARNP. La región -10 extendida incluye un motivo de 3-4 pb de longitud situado corriente arriba, adyacente al hexámero -10, y reconocido por la región 3 de la subunidad σ de la ARNP (Sanderson *et al*., 2003). El elemento "UP" es una secuencia rica en nucleótidos A/T de aproximadamente 20 pb de longitud, localizada corriente arriba de la región -35 del promotor y reconocida por los dominios carboxilo-terminales de las subunidades α de la ARNP (Ross *et al*., 2001) (Figura 2).

La contribución relativa de cada uno de los cuatro elementos descritos a la especificad de unión de la ARNP difiere de un promotor a otro. No se ha encontrado ningún promotor natural que posea secuencias idénticas a las consenso en los cuatro elementos, de modo que variaciones en alguno de los elementos se ven compensadas por la presencia de otros. Una vez que la ARNP reconoce los elementos presentes en un promotor estableciendo interacciones con los mismos, forma un complejo promotor cerrado ($RP_C$), en el que las regiones 3.2 y 1.1 de la subunidad sigma ocupan, respectivamente, el canal de salida del ARN y el canal de entrada del ADN molde al centro activo (Figura 3). La localización de las citadas regiones de sigma cumple funciones determinadas en el proceso de inicio de la transcripción que serán discutidas posteriormente.

**b.** Isomerización a complejo abierto

Tras el reconocimiento del promotor por la ARNP, la región 2.3 de la subunidad σ de la ARNP sufre a una isomerización estableciendo interacciones con la hebra no codificante en la región -10 del promotor (Tomsic *et al*., 2001; Tsujikawa *et al*., 2002). Esto desenrolla la doble hélice de ADN entre las posiciones -12 y +2, formando la burbuja de transcripción y generando el complejo promotor abierto ($RP_O$). A continuación, la cadena sencilla de ADN no codificante se desplaza hacia el centro activo a través de un canal proteico cargado positivamente, formado por las regiones 2 y 3 de sigma y por las subunidades β y β' (Figura 3) (Murakami *et al*., 2002a). Durante el transcurso de éste proceso, la región 1.1 de la subunidad σ desaloja dicho canal mediante un mecanismo que aún está por determinar (Mekler *et al*., 2002). Por último, el ADN de doble cadena localizado corriente abajo es "pinzado" entre las posiciones +5 y +12 en el interior de otro canal proteico que se forma entre las subunidades β y β' (Craig *et al*., 1995; Mecsas *et al*., 1991).

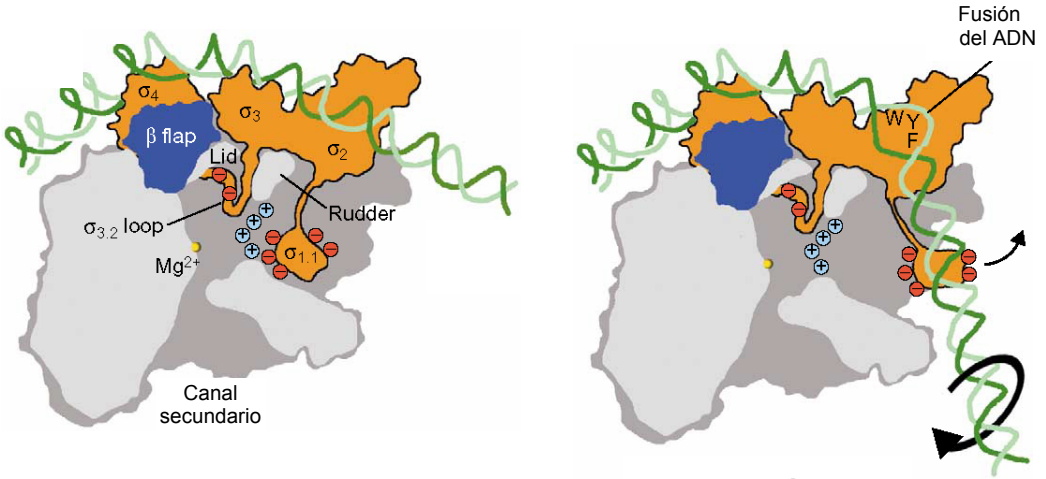


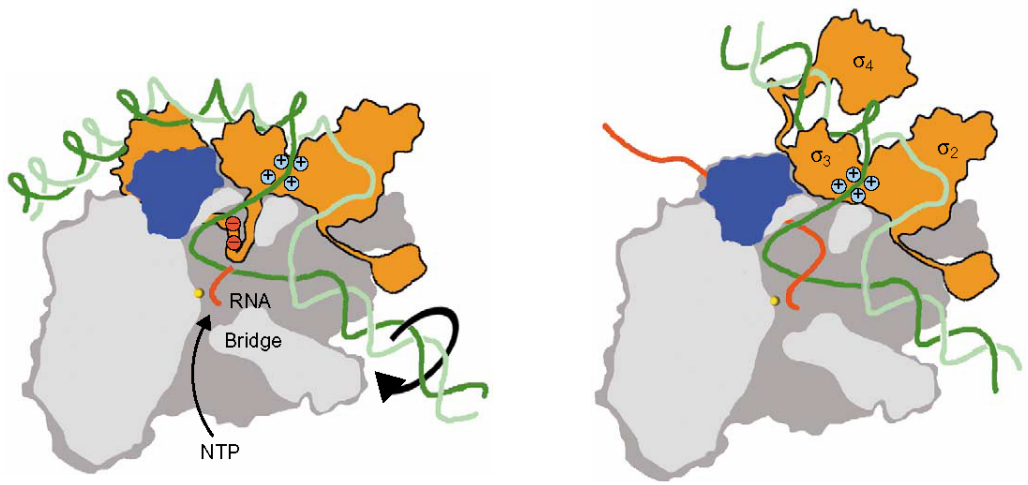

**Figura 3. Cambios estructurales en la ARNP durante el inicio de la transcripción I.** La mayor parte de la subunidad β se ha omitido, de manera que quede al descubierto el centro activo de la ARNP donde se encuentran los dos iones $Mg^{++}$. Código de color: gris, β'; azul, β; naranja, σ; verde claro, cadena codificante; verde oscuro, cadena molde o no codificante; gris oscuro, señala los dos canales proteicos. Obtenida de Murakami *et al*., 2003.

**c.** Iniciación abortiva

Una vez formado el $RP_O$, la ARNP comienza a sintetizar la molécula de ARN. Los substratos necesarios para la reacción, los nucleótidos trifosfato (NTPs), acceden al centro activo a través de un canal proteico secundario formado por las subunidades β y β' de la ARNP. Cuando el transcrito alcanza una longitud de unos pocos nucleótidos contacta con un lazo de la región 3.2 de la subunidad sigma situado en el canal de salida del ARN, dando lugar al proceso de iniciación abortiva (Murakami *et al*., 2002b). En cada nuevo paso de adición de un nucleótido antes de formar un complejo de elongación estable, la cadena de ARN puede, bien desplazar el lazo de la región 3.2 de σ o bien disociarse del complejo (Figura 4). De este modo, la ARNP inicia la transcripción varias veces, generando y liberando transcritos cortos (2-12 nucleótidos de longitud) sin disociarse del promotor. Finalmente la cadena de ARN alcanza una longitud de 12 nucleótidos, suficiente para completar el híbrido ARN-ADN dentro de la burbuja de transcripción, y desplazar a la región 3.2 de la subunidad σ fuera del canal de salida del ARN. Cuando la molécula de ARN recién sintetizada alcanza el canal de salida termina el proceso de iniciación abortiva.

**d.** Escape del promotor

El escape o abandono del promotor consiste en el paso de complejo abierto ($RP_O$) a un complejo de elongación estable con la consecuente disociación de la subunidad sigma de la holoenzima. El movimiento del lazo de la región 3.2 de sigma acopla la síntesis de la cadena de ARN con el proceso de escape del promotor (Figura 4). Dicho desplazamiento desestabiliza las interacciones entre la región "β-flap" de la subunidad β y la región 4 de σ y, en consecuencia, las interacciones entre la región 4 de σ y la región -35 en el promotor, lo que permite a la ARNP abandonar el mismo y desplazarse corriente abajo, alongando el transcrito de ARN (Bar-Nahum y Nudler, 2001; Mukhopadhyay *et al*., 2001).




**$RP_O$ e iniciación abortiva** **Escape del promotor**

**Figura 4. Cambios estructurales en la ARNP durante el inicio de la transcripción II**. La mayor parte de la subunidad β se ha omitido, de manera que quede al descubierto el centro activo de la ARNP. Código de color: gris, β'; azul, β; naranja, σ; verde claro, cadena codificante; verde oscuro, cadena molde o no codificante; rojo, cadena de ARN sintetizada. Obtenida de Murakami *et al*., 2003.

**e.** Pausas próximas al promotor

Este mecanismo regulatorio, conservado tanto en procariotas como en eucariotas, consiste en la detención de la ARN polimerasa por un espacio de tiempo corto en secuencias de ADN próximas al promotor y situadas corriente abajo del punto de inicio de la transcripción. Esta parada está causada por la asociación de la subunidad σ de la ARNP a secuencias de ADN con homología a la secuencia consenso de la región -10 del promotor. La respuesta estricta, mediada por el nucleótido ppGpp (tetrafosfato de guanosina), parece estar relacionada con las pausas próximas al promotor (ver apartado 2.3) (Ring *et al*., 1996).

### 1.2.2. ELONGACIÓN

Al igual que la iniciación de la transcripción, las etapas de elongación y terminación constituyen puntos importantes en la regulación de la transcripción, en los que participan varias proteínas reguladoras. La transición desde la fase de iniciación hasta la de elongación se caracteriza por el escape de la ARN polimerasa del promotor, la disociación del factor sigma del núcleo, y la formación de un complejo ternario de elongación altamente procesivo. La disociación del factor sigma es motivo de controversia en la actualidad y de hecho, se ha propuesto un modelo alternativo en el cual el factor sigma permanece unido al núcleo dentro del complejo ternario durante la fase de elongación (Bar-Nahum y Nudler, 2001; Mukhopadhyay *et al*., 2001) (Ver apartado 1.3.3). Durante la fase de elongación se producen pausas múltiples y específicas de la ARNP, cuya duración depende de la secuencia que está siendo transcrita, del establecimiento de interacciones específicas con proteínas reguladoras, de la concentración de los substratos (NTPs) y de errores de transcripción en el extremo 3' de la cadena de ARN sintetizada. Las paradas en la transcripción pueden dar lugar a una desestabilización del complejo y finalmente a la disociación de la ARNP del ADN, liberando el transcrito (von Hippel, 1998). Se han descrito una serie de proteínas que modifican las propiedades de la ARNP durante las fases de elongación y terminación, alterando la procesividad de la misma al provocar pausas, paradas, terminación o anti-terminación de la transcripción. En relación con esto, hay que destacar que la mayoría de los mutantes afectados en las fases de elongación y terminación presentan mutaciones localizadas en el interior de los tres canales proteicos principales de la ARNP: el canal primario de entrada del ADN, el canal secundario o de entrada de los NTPs, y el canal salida del ARN (Vassylyev *et al*., 2002; Zhang *et al*., 1999), que son a su vez los puntos de regulación donde actúan las proteínas reguladoras.

Entre los reguladores que intervienen en esta etapa de la transcripción se encuentran las proteínas GreA y GreB, NusA y Mfd. Los reguladores GreA y GreB actúan durante las pausas o paradas de la transcripción, escindiendo el extremo 3' del ARN transcrito, lo que permite a la ARNP transcribir de nuevo la zona del ADN que ha originado la parada (Borukhov *et al*., 1993; Fish y Kane, 2002). El factor NusA estimula ciertos tipos de pausas tales como las causadas por el atenuador del operón de biosíntesis de triptofano, o la terminación dependiente de Rho (ver 1.2.3.b) y, por otro lado, puede inducir anti-terminación tanto en terminadores dependientes como independientes de Rho (Nudler y Gottesman, 2002; Richardson, 1996). NusA compite con el factor sigma por la unión al núcleo de la ARNP y una vez que forma parte del complejo ternario de elongación establece interacciones con la cadena no codificante del ADN provocando las pausas en la transcripción. Por último, el factor Mfd actúa reactivando la transcripción en aquellas ARNP que se

encuentran paradas o detenidas durante la elongación. Su mecanismo de acción consiste en desplazar a la ARNP hasta que ésta reengancha el extremo 3' del ARN a su centro activo. Mfd es capaz incluso de reclutar a la maquinaria de reparación del ADN a puntos del ADN dañados y que serán reparados durante la transcripción (Roberts y Park, 2004).

### 1.2.3. TERMINACIÓN

En la última fase del proceso de transcripción, la ARN polimerasa se disocia del ADN molde debido a la presencia de unas estructuras secundarias en el ARN transcrito denominadas terminadores, liberando el transcrito de ARN y finalizando el proceso de transcripción. Existen dos clases de terminadores en el genoma de *Escherichia coli* que difieren tanto en sus características estructurales como en su mecanismo de acción:

**a.** Terminadores intrínsecos o independientes de Rho: consisten en repeticiones invertidas, ricas en pares GC presentes en la secuencia de ADN, seguidas de un tramo de timinas. Las repeticiones invertidas generan una estructura en forma de horquilla en el transcrito de ARN mensajero, estructura que causa la parada de la ARN polimerasa. La estabilidad del complejo ternario ARNP-ADN-ARN disminuye en el híbrido ADN-ARN formado por el emparejamiento de timinas y uracilos, induciendo la disociación del ARN, el ADN molde y la ARN polimerasa sin la intervención de proteínas auxiliares (Kingsford *et al*., 2007; Reynolds *et al*., 1992).

**b.** Terminadores dependientes de Rho: requieren para su funcionamiento de una proteína denominada factor Rho. El factor Rho es una proteína universal bastante conservada en bacterias con actividad helicasa dependiente de ATP. Rho es un homohexámero de protómeros de 46,8 kDa que se une a una secuencia no traducida de ARN localizada corriente arriba del punto de parada de la transcripción, denominada secuencia rut (del inglés Rho utilization) (Figura 5). La formación del complejo entre Rho y la hebra de ARN sucede en tres pasos, en los cuales se generan tres conformaciones intermedias previas a la formación de una asociación estable, momento en el que se activa la hidrólisis de ATP. La hidrólisis de ATP por el complejo Rho-ARN induce su actividad helicasa, que desenrolla el duplex ARN-ADN en dirección 5'→ 3' (Richardson, 2002), permitiendo a Rho translocarse a lo largo de la hebra de ARN (Richardson, 2003). Una vez que Rho alcanza el complejo ternario de elongación (ARN-ADN-ARNP) actúa sobre el mismo, causando la terminación de la transcripción y la disociación del ARN en el punto de parada de la transcripción (Richardson, 1990; Richardson y Richardson, 1996).

En algunas especies bacterianas, como es el caso de *Escherichia coli*, la terminación de la transcripción está mediada por Rho en la mayoría de los casos, mientras que en otras bacterias, como por ejemplo en *Bacillus subtilis*, las proteínas homólogas a Rho juegan un papel menor, siendo la norma la terminación intrínseca.

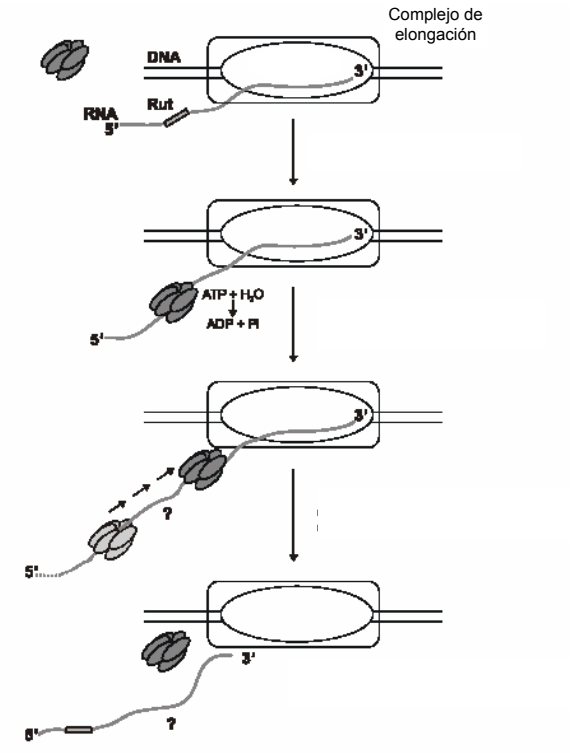



**Figura 5. Etapas del proceso de terminación dependiente de Rho.** El factor Rho reconoce la secuencia rut en el ARN transcrito y gracias a la hidrólisis de ATP se transloca a lo largo de la hebra de ARN hasta alcanzar a la ARNP, provocando el disociación del complejo ternario de elongación y la terminación de la transcripción. Obtenido de Banerjee, *et al*. 2006.

## 1.3. LA SUBUNIDAD SIGMA

La subunidad sigma de la ARN polimerasa se describió originalmente como un factor que confería al núcleo de la enzima la capacidad de reconocer secuencias promotoras específicas (Burgess *et al*., 1969), aunque esta hipótesis no se confirmó experimentalmente hasta años después (Gross *et al*., 1998). Además de su función en el reconocimiento de secuencias específicas en los promotores, la subunidad sigma provoca un cambio conformacional en el núcleo de la ARNP, necesario para que ésta adquiera su forma activa, a la vez que actúa como diana de una gran variedad de reguladores de la transcripción. La mayoría de las especies bacterianas poseen varios factores sigma diferentes, entre los cuales se incluye un factor sigma principal encargado de la mayor parte de la expresión génica durante la fase exponencial de crecimiento. El resto de factores sigma alternativos de una bacteria permiten a la célula responder a cambios específicos en las condiciones

ambientales dirigiendo a la ARNP hacia determinados grupos de promotores (Ishihama, 2000).

La unión de la subunidad sigma al núcleo de la ARNP es reversible, de modo que la expresión desde las diferentes clases de promotores depende en cada momento del nivel intracelular del factor sigma que reconoce cada promotor y de la capacidad de dicho factor para asociarse al núcleo de la ARNP dirigiéndola hacia el promotor (ver apartado 1.3.5).

### 1.3.1. Familias de factores sigma

Los factores sigma se agrupan en dos grandes familias que difieren en secuencia, estructura y mecanismo de acción: la familia $\sigma^{70}$ y la familia $\sigma^{54}$.

**a.** La familia de factores $\sigma^{70}$

La mayoría de los factores sigma alternativos pertenece a la familia de factores $\sigma^{70}$. Los factores sigma de esta familia se agrupan a su vez según su relación filogenética (Lonetto *et al*., 1992).

- Grupo 1: Incluye al factor $\sigma^{70}$ y sus ortólogos. Los factores sigma pertenecientes a este grupo son proteínas esenciales, responsables de la transcripción de los denominados genes “de mantenimiento” (housekeeping).
- Grupo 2: Estrechamente relacionados con los factores sigma del grupo 1, pero dispensables para la supervivencia celular. El factor $\sigma^{38}$ ó $\sigma^{S}$ de *E. coli*, regulador general en respuesta a estrés, es el miembro más estudiado del grupo.
- Grupo 3: Más distantes en homología de secuencia con respecto al factor $\sigma^{70}$. Puede dividirse a su vez en varios subgrupos de proteínas relacionadas evolutivamente entre sí y que comparten funciones similares. Integran este grupo los factores sigma de choque térmico ($\sigma^{32}$ ó $\sigma^{H}$), de síntesis del flagelo ($\sigma^{28}$ ó $\sigma^{F}$) o de control del proceso de esporulación ($\sigma^{B}$).
- Grupo 4: Constituyen un grupo bastante alejado filogenéticamente del resto de los factores pertenecientes a la familia $\sigma^{70}$. Se denominaron originalmente factores extracitoplasmáticos por su relación con la regulación de distintos procesos asociados a la envuelta celular y procesos de transporte (Helmann, 2002). Frecuentemente se encuentran formando parte de un operón que también codifica una proteína transmembrana que actúa como factor anti-sigma, por ejemplo ($\sigma^{E}$-RseA de *E. coli*).
- Grupo 5: En 2002, Helmann y colaboradores propusieron la creación un nuevo grupo de factores sigma, con miembros muy distantes filogenéticamente, que incluye reguladores positivos de la síntesis de toxinas y bacteriocinas.

**b.** La familia de factores $\sigma^{54}$

Como se ha mencionado anteriormente, los factores de la familia $\sigma^{54}$ difieren de los factores $\sigma^{70}$ además de en secuencia y estructura, en el mecanismo que emplean para la formación del complejo abierto. Mientras que la ARNP/$\sigma^{70}$ es capaz de llevar a cabo este proceso por sí misma, la ARNP/$\sigma^{54}$ requiere tanto de un activador específico como de la hidrólisis de ATP por parte de dicho activador para generar un complejo promotor abierto (Buck *et al*., 2000; Studholme y Buck, 2000). Los activadores específicos de la ARNP/$\sigma^{54}$ pertenecen a la familia AAA (del inglés ATPases associated with various cellular activities) y se agrupan en la denominada, familia EBP (del inglés enhancer binding proteins) por su analogía con los "enhancers" eucarióticos. Aunque la mayoría de las eubacterias suelen poseer varios representates de la familia $\sigma^{70}$, generalmente no disponen de más de un factor de la familia $\sigma^{54}$. De hecho, algunas especies bacterianas Gram positivas con alto contenido en G+C en sus genomas y las cianobacterias carecen de este factor sigma (Studholme y Buck, 2000). Los promotores dependientes de $\sigma^{54}$ se caracterizan por poseer dos motivos conservados localizados en posiciones -12 y -24 con respecto al punto de inicio de la transcripción, a diferencia de los motivos -10 y -35 propios de los promotores reconocidos por los factores de la familia $\sigma^{70}$ (Barrios *et al*., 1999).

### 1.3.2. Estructura de los factores de la familia $\sigma^{70}$

Los miembros de la familia $\sigma^{70}$ son proteínas modulares que constan de cuatro regiones conservadas subdivididas a su vez en subregiones, a las que se les han asignado diferentes funciones, como el reconocimiento de secuencias promotoras, la unión al núcleo de la ARNP o la apertura de la doble cadena de ADN durante la transcripción (Figura 6, ver apartado 1.2.1) (Gruber y Bryant, 1998). Como se describió en 1.2.1, las subregiones 2.4, 3.0 y 4.2 contienen los determinantes de unión al ADN y reconocen respectivamente los elementos -10, -10 extendido y -35 en los promotores. La región 1.1 sólo está presente en el grupo 1 de factores sigma, donde actúa como dominio auto-inhibidor enmascarando a los determinantes de unión al ADN (ver apartado 1.3.4). La región 2.3 participa en la fusión de la doble cadena de ADN (Murakami *et al*., 2002a).

La estructura cristalina de la subunidad σ principal de *Thermus aquaticus* revela la existencia de tres dominios que se pliegan de forma independiente (Campbell *et al*., 2002). El dominio 2 ($\sigma_2$) comprende las regiones 1.2-2.4, incluyendo una región no conservada presente sólo en algunos factores sigma. El dominio 3 ($\sigma_3$) engloba las regiones 3.0 y 3.1; y el dominio 4 ($\sigma_4$), que ha sido cristalizado en complejo con el ADN, comprende las regiones 4.1 y 4.2. Estos dominios están conectados entre sí por uniones flexibles y cada uno de ellos incluye tanto un determinante para la unión

a la ARNP, como un motivo de unión al ADN. La estructura de la región 1.1 no ha podido ser definida en estructuras cristalinas de alta resolución.

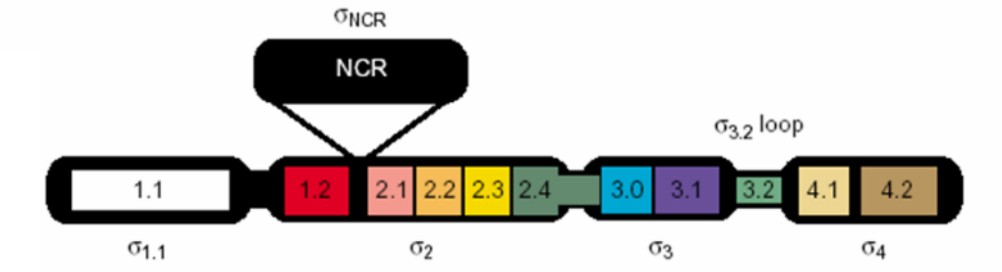



**Figura 6. Regiones conservadas de la familia de factores σ$^{70}$.** Las distintas subregiones presentan un código de color que se mantiene en la estructura tridimesional de la figura 7. Los cuatro dominios estructurales que se pliegan de forma independiente están indicados debajo de la figura. Obtenido de Murakami *et al*., 2003.

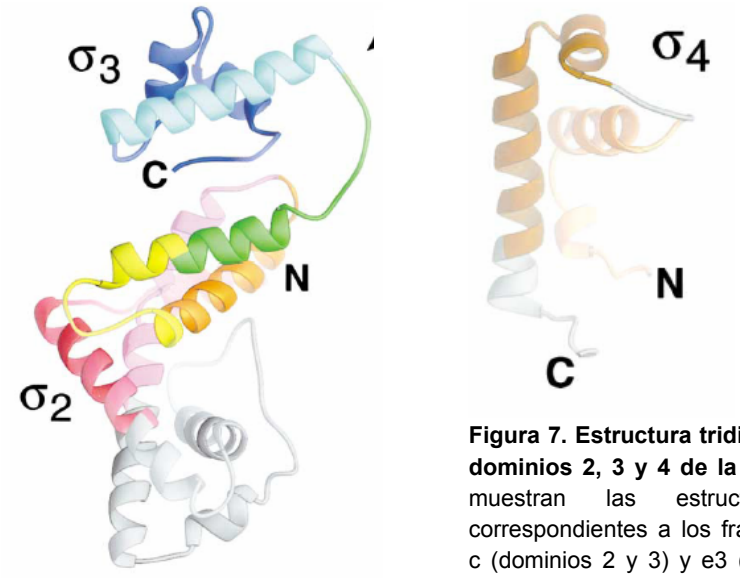



**Figura 7. Estructura tridimensional de los dominios 2, 3 y 4 de la subunidad σ.** Se muestran las estructuras cristalinas correspondientes a los fragmentos estables c (dominios 2 y 3) y e3 (dominio 4) de los cinco resultantes de la proteolisis del factor sigma. Obtenido de Campbell *et al*., 2002.

### 1.3.3. El ciclo del factor sigma

Es el proceso en el cual el factor sigma se asocia al núcleo de la ARNP para dirigir el inicio de la transcripción, disociándose de la misma cuando se completa el paso a un complejo de elongación estable. Durante la terminación de la transcripción, la ARNP se disocia del ADN molde y del ARN transcrito, quedando libre para volver a unirse a una subunidad sigma y comenzar un nuevo ciclo de transcripción. La unión del factor sigma para formar la holoenzima produce un

cambio conformacional tanto en el factor sigma como en el núcleo de la ARNP (Murakami *et al*., 2002a). El cambio conformacional que sufre el factor sigma desenmascara sus dominios de unión al ADN, y por tanto sólo es capaz de unirse a las secuencias promotoras cuando forma parte de la holoenzima. Este ciclo permite a la célula ajustar sus patrones de transcripción rápidamente, en respuesta a cambios en las condiciones externas o a señales intracelulares, activando la expresión de determinados grupos de genes a través de distintos factores sigma, que muestran afinidad por diferentes secuencias promotoras.

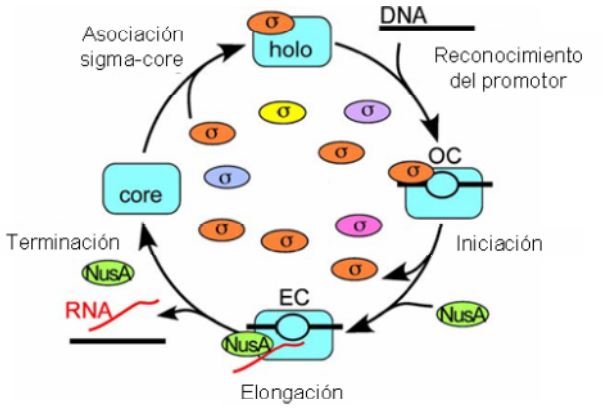



**Figura 8. Modelo del ciclo del factor sigma.** Los distintos factores σ presentes en la célula compiten por la unión al núcleo de la ARNP. Una vez completada la fase de inicio de la transcripción, el factor sigma se disocia de la ARNP. La unión competitiva del factor de elongación NusA (ver apartado 1.2.2) puede desplazar a la subunidad σ de la ARNP. Obtenido de Money, *et al*., 2005

Este concepto tradicional del ciclo del factor sigma ha cambiado en los últimos años en los que han surgido nuevas hipótesis que sugieren que el factor sigma podría jugar un papel regulatorio durante la etapa de elongación a través tanto de una asociación transitoria con el complejo de elongación (Kapanidis *et al*., 2005; Mooney y Landick, 2003), como de una asociación permanente (Bar-Nahum y Nudler, 2001; Mukhopadhyay *et al*., 2001) o de un cambio conformacional en el complejo de elongación que perdura tras la disociación del factor sigma (Berghofer-Hochheimer *et al*., 2005).

Estudios recientes acerca de la dinámica del factor sigma parecen confirmar que, principalmente, se disocia del complejo de elongación de modo estocástico tras el escape del promotor, permitiendo reprogramar la maquinaria transcripcional de acuerdo con el ciclo tradicional (Raffaelle *et al*., 2005). Estos resultados no contradicen el hecho de que el factor sigma pueda permanecer unido al complejo de elongación en algún caso o que incluso pueda volver a asociarse al complejo tras su disociación inicial (Mooney y Landick, 2003), regulando procesos tales como las pausas próximas al promotor (Kapanidis *et al*., 2005; Mukhopadhyay *et al*., 2001; Nickels *et al*., 2004), las pausas transcripcionales durante la elongación o incluso la terminación de la transcripción (Berghofer-Hochheimer *et al*., 2005).

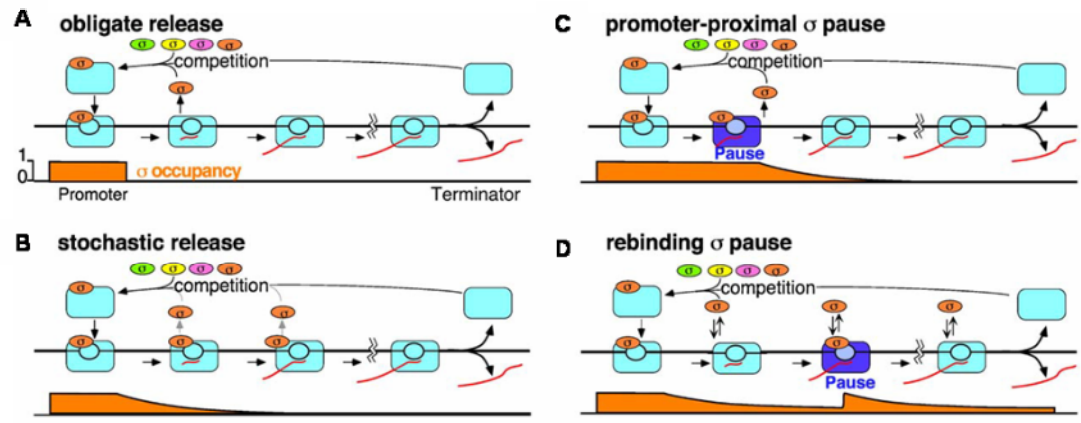



**Figura 9. Disociación del factor σ de la ARNP.** La figura describe las distintas hipótesis vigentes que explican la dinámica del factor σ. **A.** Disociación obligatoria: El factor sigma se disocia del núcleo de la ARNP con el paso a la etapa de elongación. **B.** Disociación estocástica: la afinidad del factor sigma por la ARNP desciende con el paso a un complejo de elongación, de manera que el factor sigma se disocia de la ARNP a lo largo del proceso de transcripción, con mayor o menos rapidez dependiendo de la concentración de sigma libre y de otras proteínas competidoras. **C.** Pausas próximas al promotor: en algunos promotores la subunidad sigma establece interacciones con el ADN corriente abajo del punto de inicio de la transcripción antes de disociarse de la ARNP, provocando una parada en la transcripción. **D.** Reasociación del factor sigma: la subunidad sigma puede volver a asociarse al complejo ternario de elongación durante dicha fase, estableciendo interacciones con secuencias en el ADN, lo que provoca una pausa en la transcripción. La barra de color naranja representa la fracción de moléculas de ARNP asociadas al factor sigma en cada momento durante proceso de transcripción.Obtenido de (Mooney *et al*., 2005).

### 1.3.4. Contactos entre el factor sigma y el núcleo de la ARNP

Cada dominio de la subunidad sigma establece numerosas interacciones de forma simultánea con el núcleo de la ARNP contribuyendo a la gran afinidad de unión entre sigma y el núcleo (Murakami *et al*., 2002b). Durante el proceso de iniciación de la transcripción se produce la disociación gradual de dominios individuales de la subunidad sigma unidos al núcleo, dando lugar a la liberación eventual de la misma (ver apartado 1.2.1.d) (Murakami y Darst, 2003). El mayor número de interacciones se establece con el segundo dominio estructural de la subunidad sigma (regiones 1.2-2.4) que adopta forma de "U", plegándose alrededor de la región coiled-coil de la subunidad β' (Figura 10c). La región 1.1 del factor sigma, únicamente presente en los factores sigma del grupo 1, posee una gran cantidad de amino ácidos cargados negativamente y se localiza en el canal proteico de entrada del ADN, extendiéndose hasta el centro activo de la holoenzima. Durante la formación del complejo abierto, la región 1.1 se desplaza fuera de dicho canal, de manera que se ha propuesto que la presencia de la región 1.1 en el interior del canal tiene como función aumentar la anchura del mismo, facilitando la entrada del ADN

hacia el sitio activo y acelerando la formación del complejo abierto (Murakami *et al*., 2002b).

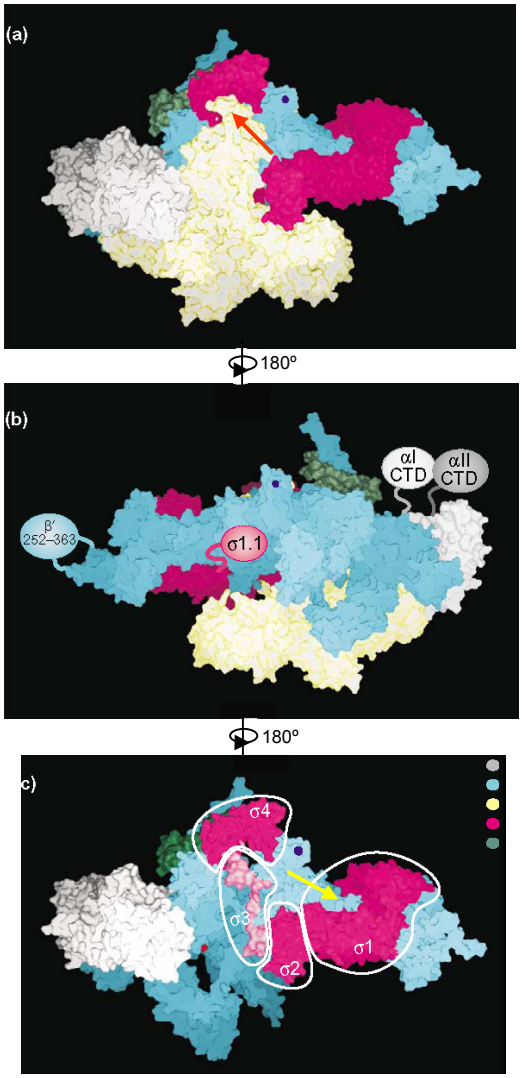



**Figura 10. a) Vista frontal de la ARN polimerasa.** El código de color para las distintas subunidades es: σ, magenta; β', azul; β, amarillo; α, gris; ω, verde oscuro. La flecha de color rojo indica la posición de la región "flap" de la subunidad β **b) Vistal posterior de la ARNP**, obtenida por rotación de 180º sobre el eje vertical. Los elementos de los que no se ha podido determinar su estructura están representados como elipses y conectados con el resto de la ARNP por lazos flexibles. **c) Vista de la ARNP sin la subunidad β.** Se observan los distintos dominios de la subunidad sigma y como el canal de salida del ARN está ocupado por la región 3 de sigma. La flecha amarilla señala la posición de la región "coiled-coil" de β'. Los dos iones magnesio del centro activo están representados como dos esferas rojas. Obtenido de (Borukhov y Nudler, 2003).

La región 3.2 de la subunidad sigma ocupa el canal de salida del transcrito de ARN en el complejo de elongación. Como se ha descrito anteriormente (ver apartado 1.2.1.c), la elongación de la cadena de ARN provoca un impedimento estérico, de modo que durante la iniciación abortiva, se disocian del complejo ARN pequeños recién sintetizados. El proceso se repite hasta que un ARN de 13-14 nucleótidos de longitud alcanza el canal de salida, desplazando a la región 3.2. La región 4 de la subunidad sigma contiene un motivo hélice-giro-hélice (HTH), y presenta gran similitud con activadores transcripcionales de la familia FixJ y con el dominio de unión a ADN de NarL (Campbell *et al*., 2002). La estructura de dicha región adopta forma de "C", con un bolsillo hidrofóbico orientado hacia la ARNP, que

establece interacciones con la hélice situada en el extremo de la región "β-flap" de la subunidad β (Campbell *et al*., 2002) (Figura 10a). La curvatura provocada en el ADN por la unión de la región 4 de sigma al elemento -35 del promotor, altera la trayectoria del ADN corriente arriba, aproximándolo a la ARNP y facilitando tanto las interacciones entre el dominio carboxilo-terminal de la subunidad α y el ADN, como las interacciones entre activadores de tipo II y la subunidad sigma (Murakami y Darst, 2003). Durante el proceso de escape de la ARNP del promotor, el movimiento de la región 3.2 fuera del canal de salida del ARN desestabiliza las interacciones entre la región 4 de sigma y la región "β-flap".

### 1.3.5. Competencia entre factores sigma

La idea central del modelo de competencia entre factores sigma en la célula se basa en la existencia de una cantidad limitante de núcleo de la ARNP libre en relación a la cantidad total de los distintos factores sigma (Ishihama, 2000). Durante la fase estacionaria de crecimiento, la competencia por la unión al núcleo de la ARNP es aún mayor debido al descenso de los niveles de núcleo disponibles (Kawakami *et al*., 1979). Por lo tanto, la proporción relativa de las diferentes especies de holoenzima va a depender fundamentalmente de dos parámetros estrechamente regulados: la concentración intracelular de cada uno de los factores sigma y la afinidad relativa de unión de cada uno de ellos al núcleo de la ARNP.

**a.** Regulación de los niveles intracelulares de cada factor sigma

La concentración intracelular de cada factor sigma varía dependiendo de las condiciones a las que se encuentra sometida la célula. Los mecanismos de regulación de los niveles dependen tanto del factor sigma, como de la especie bacteriana en cuestión. Como se ha mencionado anteriormente, el factor sigma principal $\sigma^{70}$ es el responsable de la mayor parte de la expresión génica en la célula, mientras que los factores sigma alternativos regulan grupos específicos de genes en respuesta a cambios en las condiciones ambientales. De acuerdo con las funciones descritas la concentración intracelular del factor principal $\sigma^{70}$ es la más elevada de todos los factores sigma, tanto en fase exponencial como en fase estacionaria de crecimiento, e incluso bajo condiciones de estrés (Jishage y Ishihama, 1995; Jishage *et al*., 1996). Los niveles del factor $\sigma^{38}$ aumentan con la entrada en la fase estacionaria de crecimiento, así como en respuesta a factores de estrés. Su regulación tiene lugar a varios niveles e involucra numerosos mecanismos y factores de regulación (Hengge-Aronis, 2002). A nivel transcripcional, intervienen en la regulación factores como la proteína CRP (catabolic repression protein) (Lange *et al*., 1995; Lange y Hengge-Aronis, 1994), algunos factores de "quorum sensing" (Whiteley *et al*., 2000) y el regulador transcripcional PsrA (Kojic y Venturi, 2001). El control de la traducción de $\sigma^{38}$ constituye el proceso en el que intervienen un mayor

número de mecanismos y factores de regulación. Entre ellos se encuentran la proteína Hfq de unión a ARN, ARNs pequeños no codificantes como DsrA y OxyS, proteínas asociadas al nucleoide como HU y H-NS (Zhou y Gottesman, 2006), proteínas del choque térmico o de choque frío como DnaK y CspC, respectivamente, y moléculas pequeñas como UDP-glucosa o la molécula ppGpp.

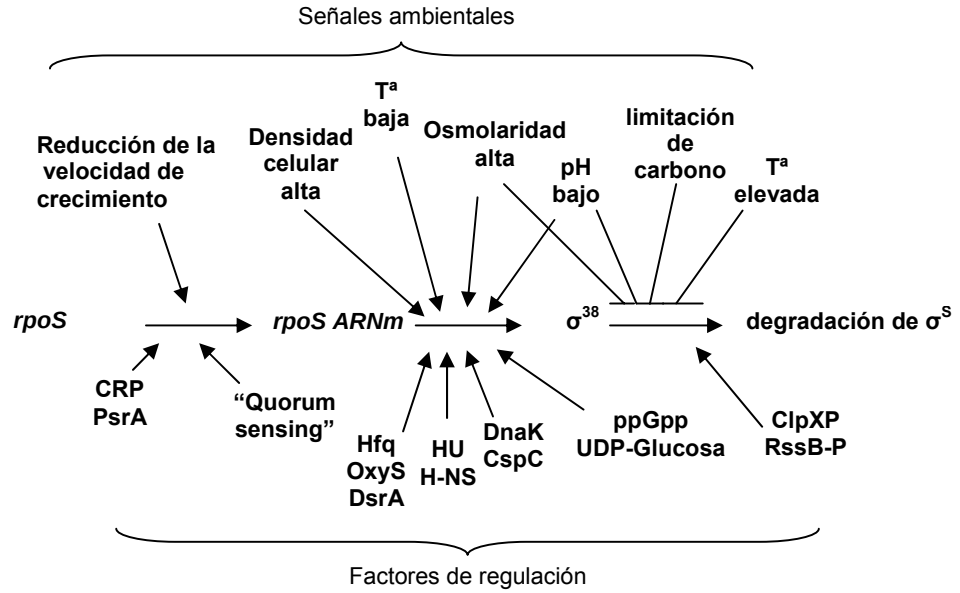



**Figura 11. Regulación de los niveles del factor $\sigma^{38}$.** Los niveles del factor $\sigma^{38}$ se regulan a tres niveles: transcripcional, traduccional y post-traduccional, en los que interviene un amplio abanico de reguladores y mecanismos diferentes que responden a señales ambientales diversas. Adaptado de Hengge-Aronis *et al*., 2002.

A nivel postraduccional, el complejo proteasa ClpXP es el encargado de degradar al factor $\sigma^{38}$, previa interacción con la forma fosforilada de la proteína RssB (Studemann *et al*., 2003), proceso en cuya regulación intervienen la proteína IraP (Bougdour *et al*., 2006) y otros mecanismos descubiertos recientemente (Fredriksson *et al*., 2007).

El factor $\sigma^{32}$ está sometido a una complicada regulación que incluye el control de su actividad (Blaszczak *et al*., 1999), de su estabilidad (Tatsuta *et al*., 2000) y de su traducción, a través de la desestabilización de una estructura secundaria en su ARNm que solapa con el punto de inicio de la traducción (El-Samad *et al*., 2005; Morita *et al*., 2000). La proteasa FtsH controla la estabilidad del factor $\sigma^{32}$ (Herman *et al*., 1995; Tatsuta *et al*., 2000) y las chaperonas DnaK, DnaJ y GroELS regulan su actividad y su degradación (Guisbert *et al*., 2004; Tatsuta *et al*., 2000).

**b.** Afinidad por la ARNP

La afinidad de unión al núcleo de la ARNP es diferente para cada factor sigma. Las afinidades de unión relativas calculadas *in vitro* para los siete factores sigma presentes en *E. coli* muestran que el factor principal $\sigma^{70}$ presenta la mayor afinidad por el núcleo, seguido de $\sigma^{54}$, $\sigma^{F}$, $\sigma^{32}$, $\sigma^{FecI}$, $\sigma^{E}$ y $\sigma^{38}$ en orden decreciente de afinidad (Maeda *et al*., 2000). Por otro lado, las condiciones intracelulares y otros factores adicionales afectan a las afinidades de unión relativas, regulando el reemplazamiento de la subunidad sigma asociada al núcleo. Por ejemplo, cambios en la composición química del citosol (ej: concentración de polifosfatos inorgánicos, concentración de glutamato) afectan a la utilización preferencial de $\sigma^{38}$ sobre $\sigma^{70}$ (Ding *et al*., 1995; Kusano y Ishihama, 1997a; Kusano y Ishihama, 1997b).

La molécula efectora de la respuesta estricta (ppGpp) afecta la transcripción desde determinados tipos de promotores, al unirse a las subunidades β y β' de la ARNP (Chatterji *et al*., 1998; Toulokhonov *et al*., 2001) alterando la afinidad relativa por los factores sigma (Jishage *et al*., 2002; Laurie *et al*., 2003). Asi, en presencia de niveles altos de ppGpp aumenta la fracción de holoenzima asociada a factores sigma alternativos, reduciéndose la fracción de holoenzima unida al factor principal $\sigma^{70}$, situación que ocurre en respuesta a condiciones nutricionales o ambientales desfavorables para el crecimiento (Nystrom, 2004) (Ver apartado 2.3).

### 1.3.6. Factores anti-sigma

Los factores anti-sigma son proteínas que interaccionan con un factor sigma específico, impidiendo su asociación con el núcleo de la ARNP e inhibiendo la transcripción desde los promotores dependientes de dicho factor (Helmann, 1999). La región 4 de la subunidad sigma es, frecuentemente, la diana de los factores anti-sigma, aunque también pueden establecer interacciones a través de otras regiones de la misma (Chadsey y Hughes, 2001; Decatur y Losick, 1996). Algunos factores anti-sigma no impiden la unión de σ al núcleo, sino que actúan impidiendo el inicio de la transcripción desde determinados tipos de promotores, por ejemplo AsiA-$\sigma^{70}$ (Hinton y Vuthoori, 2000). Como ejemplo de parejas sigma-antisigma cabe citar, $\sigma^{38}$-RssB (Mika y Hengge, 2005), $\sigma^{70}$-Rsd (Westblade *et al*., 2004), $\sigma^{F}$-FlgM (Schmitt *et al*., 1996), $\sigma^{E}$-RseA (Missiakas *et al*., 1997), $\sigma^{FecI}$-FecR (Enz *et al*., 2000).

### 1.3.7. Solapamiento funcional entre factores sigma.

Por lo general, los distintos factores sigma presentes en la célula reconocen secuencias promotoras diferentes, de manera que la gran mayoría de genes son transcritos por una única ARNP. A pesar de esto, análisis recientes del solapamiento de secuencias promotoras reconocidas por distintos factores sigma en un mismo

gen (Wade *et al*., 2006) apuntan a que algunos factores sigma alternativos podrían haber evolucionado para reconocer promotores dependientes de $\sigma^{70}$, aumentando su nivel de transcripción. Este solapamiento funcional entre factores sigma permitiría la expresión de determinados genes bajo múltiples condiciones ambientales, lo que provocaría un aumento en la complejidad de los patrones de regulación. Sin embargo, no parece probable que todos los factores sigma alternativos compartan este solapamiento funcional con el factor sigma principal $\sigma^{70}$, puesto que algunos de ellos transcriben grupos de genes relacionados con funciones muy especializadas, que sólo se transcriben bajo circunstancias específicas, como por ejemplo los factores sigma de esporulación en *Bacillus subtilis*.

El promotor Pm del operón meta del plásmido TOL, objeto de estudio de esta tesis doctoral, representa un ejemplo de solapamiento funcional en el que dos factores sigma alternativos relacionados con la respuesta a estrés permiten la transcripción constante desde un único promotor en respuesta a la presencia de un compuesto aromático, el 3-metilbenzoato (3MB).

## 1.4. EL PROMOTOR

Los promotores son secuencias de ADN reconocidas por la subunidad sigma de la ARNP que permiten la unión de la misma para dar comienzo al proceso de transcripción. Atendiendo al tipo de factor sigma que reconoce un promotor, éste presenta una serie de características de secuencia y estructurales. Las secuencias consenso definidas para los distintos factores sigma que se describen a continuación, se han deducido del análisis de promotores de *Escherichia coli*. La alta homología de secuencia que comparten los factores sigma ortólogos de las distintas especies, en especial en la regiones implicadas en el reconocimiento del promotor, se corresponde con la homología que presentan las secuencias consenso de los promotores que reconocen (Domínguez-Cuevas y Marqués, 2004; Helmann, 1995). Los factores de la familia $\sigma^{70}$ reconocen secuencias consenso en las regiones -10 y -35, mientras que el factor $\sigma^{54}$ establece interacciones con secuencias consenso localizadas en posiciones -12 y -24 con respecto al punto de inicio de la transcripción.

### 1.4.1. Estructura de los promotores dependientes de $\sigma^{70}$

Los promotores dependientes del factor $\sigma^{70}$ se caracterizan por presentar dos hexámeros de secuencia conservada, TTGACA y TATAAT centrados en posiciones -35 y -10, respectivamente, con respecto al punto de inicio de la transcripción (Hawley y McClure, 1983). La distancia que separa ambos hexámeros es variable, y oscila entre las 15-20 pb con un valor óptimo de 17±1 pb (Dombroski *et al*., 1996). En promotores que carecen de un buen consenso en la región -35, es frecuente que

aparezca una nueva región conservada corriente arriba de la región -10, denominada región -10 extendida (TG en posiciones -15 y -14), que es reconocida por la región 3.0 de la subunidad sigma (Barne *et al*., 1997). Además corriente arriba de la región -35 los promotores dependientes de $\sigma^{70}$ pueden presentar secuencias de ADN de unos 20 pb ricas en nucleótidos A/T denominadas "elemento UP", que ayudan a la estabilización de la unión de la ARNP mediante contactos de las subunidades α de la ARNP con dichas secuencias o con proteínas reguladoras (Benoff *et al*., 2002). Concretamente, el elemento UP puede dividirse en dos subregiones de 10 pb conocidas como subelementos UP distal y proximal atendiendo a su localización respecto al punto de inicio de la transcripción, que son reconocidas por los dominios carboxilo terminales de cada una de las subunidades α de la ARNP. Un promotor dado puede presentar tanto un elemento UP completo como los subelementos UP distal y proximal por separado.

### 1.4.2. Estructura de los promotores dependientes de $\sigma^{38}$

El factor $\sigma^{38}$ posee gran similitud de secuencia con el factor $\sigma^{70}$ en las regiones 2.4 y 4.2 de reconocimiento de los motivos -10 y -35 del promotor. Este hecho se ve reflejado en la capacidad que tienen ambas polimerasas de reconocer los mismos promotores *in vitro* (Gaal *et al*., 2001). Sin embargo la situación *in vivo* es diferente y ambas polimerasas controlan regulones diferentes y juegan papeles complementarios. El solapamiento existente en cuanto al reconocimiento de promotores por ARNP asociadas a los factores $\sigma^{38}$ ó $\sigma^{70}$ tiene su explicación en la variedad de factores que afectan de forma diferente a la asociación de ambas polimerasas. En primer lugar, la ARNP/$\sigma^{38}$ es menos sensible a variaciones de secuencia que se apartan del consenso en las regiones -10 y -35, así como a cambios en la distancia de separación entre ambos motivos (Lacour *et al*., 2003; Typas y Hengge, 2006). La ARNP/$\sigma^{38}$ reconoce una secuencia -10 extendida distinta a la reconocida por ARNP/$\sigma^{70}$, concretamente en las posiciones -13 (C) y -14 (T) con respecto al punto de inicio de la transcripción y muestra preferencia por secuencias ricas en A y T entre la región -10 y el punto de inicio de la transcripción (Gaal *et al*., 2001; Lee y Gralla, 2001; Weber *et al*., 2005). Las diferencias entre estos dos tipos de promotores se resumen en la figura 11. Cada una de las características descritas anteriormente afecta en mayor o menor grado a la selectividad por la ARNP/$\sigma^{38}$. Los promotores dependientes de ARNP/$\sigma^{38}$ son modulares, en el sentido de que diferentes combinaciones de estas características pueden permitir el reconocimiento por $\sigma^{38}$ (Typas *et al*., 2007). A nivel global, niveles bajos de superenrollamiento negativo ejercen un efecto menor sobre la transcripción dependiente de la ARNP/$\sigma^{38}$, comparado con el efecto sobre la transcripción dependiente de la ARNP/$\sigma^{70}$. Esto se corresponde con la bajada en el grado de superenrollamiento negativo que tiene lugar durante la fase crecimiento de estacionaria, momento en el

que se produce la activación de genes dependientes de la ARNP/$\sigma^{38}$ (Bordes *et al*., 2003; Kusano *et al*., 1996).

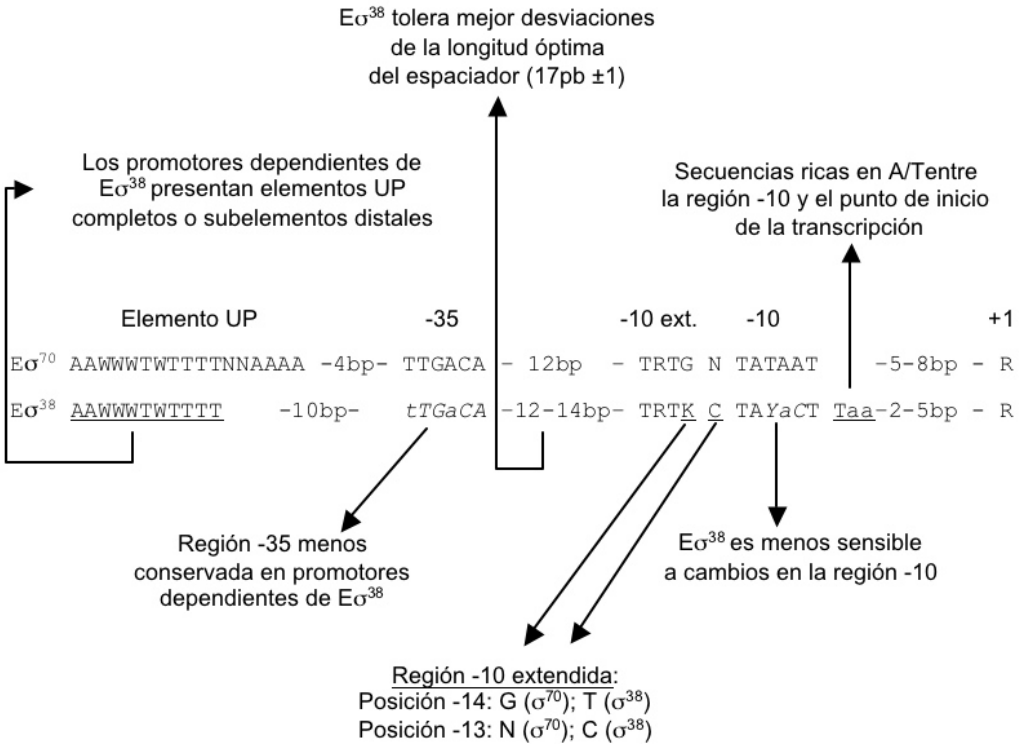



**Figura 12. Comparación de las secuencias consenso de promotores reconocidos por Eσ$^{70}$ y Eσ$^{38}$.** Los elementos de la secuencia resaltados aparecen en la mayor parte de los promotores dependientes de los factores sigma correspondientes. Las bases menos conservadas de los elementos -10 y -35 aparecen en cursiva. Los elementos subrayados contribuyen a la dependencia por $\sigma^{38}$, junto con la presencia de regiones -10 y -35 degeneradas. Adaptado de Typas *et al*., 2007.

### 1.4.3. Estructura de los promotores dependientes de $\sigma^{32}$

El factor $\sigma^{32}$ ó $\sigma^{H}$ es el factor sigma que regula la respuesta a choque térmico. El análisis de promotores dependientes de $\sigma^{32}$, así como el estudio de los promotores de los genes inducidos tras un choque térmico, ha permitido definir las características de los promotores dependientes de este factor sigma. Las secuencias consenso de las regiones -10 (5'-CCCCATWT-3') y -35 (5'-TTGAAA-3') difieren notoriamente de las de los promotores dependientes del factor $\sigma^{70}$ (Cowing *et al*., 1985; Nonaka *et al*., 2006). La distancia entre las regiones -10 y -35 oscila entre 13-14 nucleótidos. Al igual que ocurre con los promotores dependientes de $\sigma^{38}$ y de forma contraria a los promotores dependientes de $\sigma^{70}$, los promotores de genes del choque térmico suelen presentar elementos UP completos o sólo con el

subelemento distal, pero no suelen presentar un subelemento UP proximal aislado. Esto probablemente se deba a que los aminoácidos que en $\sigma^{70}$ contactan con la subunidad α de la ARNP, que se une al subelemento UP proximal, no están conservados en las otras sigma alternativas (Nonaka *et al*., 2006; Typas y Hengge, 2005).

## 2. REGULADORES DE LA TRANSCRIPCIÓN

Bajo esta denominación se agrupan todos aquellos factores, proteicos o no, y mecanismos que detectan señales intra- o extracelulares cuya función es transmitirla, regulando de forma coordinada el proceso de transcripción.

### 2.1. FACTORES DE TRANSCRIPCIÓN

Son factores de regulación proteicos. Algunos factores de transcripción controlan la expresión de un gran número de genes, mientras que otros modulan de forma específica la expresión de unos pocos. Estos factores de transcripción se conocen como reguladores globales y específicos, respectivamente.

#### 2.1.1. Reguladores globales

En *E. coli* se estima que siete factores de transcripción son los responsables del control del 50% de todos los genes regulables de la célula. Entre estos reguladores globales se encuentra CRP (del inglés cyclic AMP receptor protein), también conocida como CAP (del inglés catabolite gene activator protein), que regula el mecanismo de represión catabólica en dicha bacteria (Grainger *et al*., 2005; Zheng *et al*., 2004). Otro regulador importante es la proteína FNR (fumarate y nitrite reductase) que regula un gran número de genes en respuesta a condiciones de limitación de oxígeno (Constantinidou *et al*., 2006; Grainger *et al*., 2007). IHF (integration host factor) y Fis (del inglés factor for inversion stimulation) detectan el grado de superenrollamiento del ADN en la célula y regulan la transcripción de genes que codifican topoisomerasas (ver apartado 2.2) además de jugar un papel en la infección por ciertos bacteriófagos. Como reguladores específicos responden a señales diversas y regulan un gran numero de genes relacionados con diferentes funciones (Arfin *et al*., 2000; McLeod y Johnson, 2001; Schneider *et al*., 2000). Los tres reguladores globales restantes del grupo son ArcA (anaerobic respiratory control), NarL (nitrate regulation) y Lrp (leucine regulatory protein) (Hung *et al*., 2002; Liu y De Wulf, 2004; Overton *et al*., 2006).

#### 2.1.2. Factores de transcripción específicos.

Son proteínas que controlan la expresión de uno o pocos genes, regulando procesos celulares específicos uniéndose al promotor de los mismos. La regulación

puede ocurrir mediante activación o represión de la transcripción. Algunos factores de transcripción actúan únicamente como activadores o como represores de la transcripción, mientras que otros pueden cumplir ambas funciones dependiendo del promotor al que se asocien (Pérez-Rueda y Collado-Vides, 2000).

**a)** Represores

Los represores reducen la frecuencia de iniciación de la transcripción en sus promotores diana. En la mayoría de los casos el mecanismo de represión es simple e involucra a un solo represor. Existen cuatro mecanismos generales de represión:

1. Represión por impedimento estérico, en el que el sitio de unión del represor está localizado muy cerca o solapando con los elementos de reconocimiento de la ARNP (Müller-Hill, 1996).

2. Represión por interferencia con pasos siguientes al reclutamiento de la ARNP, como la isomerización a complejo abierto y el escape del promotor (Müller-Hill, 1998).

3. Represión por la formación de un lazo en el ADN mediante la unión de varios represores a sitios distales del promotor. La estructura en forma de lazo impide a la ARNP iniciar la transcripción (Choy, 1996).

4. Represor por anti-activación. El represor unido al ADN actúa impidiendo la activación al establecer interacciones con un activador (Shin *et al*., 2001).

**b)** Activadores

Los activadores favorecen la iniciación de la transcripción acelerando distintas etapas del proceso. Pueden actuar mejorando la afinidad de la ARNP por el promotor, efecto que ejercen una vez unidos a sus secuencias específicas en el promotor (reclutamiento) o interaccionando con la ARNP en solución antes de unirse al ADN (pre-reclutamiento) (Martin *et al*., 2002). También pueden estimular el paso del complejo promotor cerrado a abierto y/o favorecer el escape de la ARNP del promotor. Los mecanismos de activación simple (aquellos en los que interviene un solo activador) se dividen en tres grupos generales:

1. Activación de clase I, en la que el activador se une a secuencias de ADN localizadas corriente arriba de la región -35 del promotor y recluta a la ARNP mediante interacciones con la subunidad α de la misma (Ebright, 1993). La flexibilidad de la conexión entre los dominio carboxilo y amino terminales de la subunidad α de la ARNP permite que los sitios de unión del activador puedan estar a mayor o menor distancia corriente arriba del promotor.

2. Activación de clase II. El activador reconoce secuencias de ADN que solapan con la región -35 del promotor, y establece contactos con la subunidad

σ de la ARNP (Dove *et al*., 2003), contactos que aumentan la afinidad de ésta por el promotor y alteran los pasos subsiguientes en el proceso de inicio de la transcripción. Además de los contactos que establecen con la subunidad σ, en algunos casos los activadores de clase II también interaccionan con la subunidad α de la ARNP.

3. Activación por cambio conformacional. El activador altera la conformación del promotor de manera que posibilita la interacción de la ARNP con los elementos -10 y -35. En la familia de activadores transcripcionales MerR, la distancia entre las regiones -10 y -35 no es óptima para la unión de la ARNP, de modo que los activadores de esta familia se unen a la secuencia espaciadora entre -10 y -35, ocasionando un cambio en la estructura del ADN que permite re-orientar dichos elementos (Brown *et al*., 2003).

En algunos promotores la regulación responde a más de una señal ambiental y depende de más de un factor de transcripción. Esto implica que debe de haber una integración de señales. Cuando la regulación depende de una combinación de activadores y represores, éstos suelen actuar de forma independiente y alternativa. Sin embargo, en promotores regulados por una combinación de activadores los mecanismos son más complejos. Se han definido cuatro mecanismos generales: cambio en la posición del activador (Richet *et al*., 1991), contactos independientes de cada activador con la ARNP (McLeod y Johnson, 2001; Tebbutt *et al*., 2002), unión cooperativa de los activadores (Wade *et al*., 2001) y anti-represión (Browning *et al*., 2002).

### 2.1.3. Mecanismo de activación.

Como se ha descrito anteriormente, los contactos directos entre los activadores y la ARNP pueden favorecer uno o más pasos del proceso de inicio de la transcripción, como la unión de la ARNP al promotor debido a un efecto de cooperatividad en la unión al ADN, la isomerización o el escape del promotor al provocar un cambio conformacional en la ARNP (Rhodius y Busby, 1998). Los tipos de contactos entre activadores y la ARNP dependen fundamentalmente de la localización de los sitios de unión del activador en el promotor. Por otro lado, la mayoría de los activadores al unirse al ADN provocan un cambio conformacional en el mismo, curvándolo, y permitiendo superar la barrera energético-cinética que supone el inicio de la transcripción. Ambos mecanismos están interconectados y frecuentemente, la curvatura provocada en el ADN por la unión del activador favorece el establecimiento de contactos con la ARNP, así como la formación de un mayor número de interacciones entre la ARNP y el ADN en la región promotora.

### 2.1.4. Familias de reguladores transcripcionales.

Los factores de transcripción se agrupan en diferentes familias atendiendo a su similitud de secuencia y estructura, al motivo de unión al ADN que presentan y a los motivos de unión al efector y de oligomerización (Pérez-Rueda y Collado-Vides, 2000). En procariotas existen numerosas familias de reguladores transcripcionales como por ejemplo, la familia EBP (enhancer binding proteins) a la que pertenecen reguladores como NtrC, XylR y DmpR, y las familias CRP, LuxR, OmpR, AsnC, IclR, AraC, LysR, GalR, TetR y GntR, que se nombran en base al miembro más estudiado dentro de cada familia. En la mayoría de los casos la región que define a la familia contiene un motivo HTH, mientras que el resto de la proteína, generalmente un segundo dominio implicado en la respuesta a señales diversas, está menos conservado. La excepción a esta regla la constituyen las familias de reguladores transcripcionales IclR y EBP en las que los dominios conservados son el dominio de unión a efector en el primer caso y los dominios ATPasa y de contacto con la subunidad sigma en la familia EBP. En el apartado 4 de esta introducción se expondrán con más detalle las características de la familia AraC, de la que forma parte el regulador XylS de la ruta meta del plásmido TOL, objeto de esta tesis.

## 2.2. SUPERENROLLAMIENTO DEL ADN. TOPOLOGÍA DEL ADN.

La coordinación global en los niveles de expresión génica permite a una célula adaptarse a cambios constantes en las condiciones de crecimiento. Esta coordinación requiere la integración de señales nutricionales y ambientales para generar una respuesta que optimice el crecimiento y la supervivencia de las células. El superenrollamiento del ADN es un nivel de regulación global que responde a cambios en el estado energético de la célula, el cual está modulado a su vez por señales nutricionales y ambientales (Travers y Muskhelishvili, 2005). El ADN procariota suele presentar superenrollamiento negativo. Este superenrollamiento está regulado por la combinación de varios factores, como proteínas asociadas al ADN, procesos de transcripción y replicación, y la actividad de topoisomerasas (Drlica, 1992; Luttinger, 1995). La molécula reguladora ppGpp (ver apartado siguiente) y los reguladores globales FIS y CRP también juegan un papel en el control del superenrollamiento negativo. En el caso de FIS y CRP ejercen tanto un efecto local directo, provocando una curvatura en el ADN al unirse a algunos promotores (Aiyar *et al*., 2002; Muskhelishvili *et al*., 1997; Rochman *et al*., 2002), como un efecto indirecto al regular la expresión de genes que codifican topoisomeraras o proteínas asociadas al nucleoide, como por ejemplo HU, H-NS o Dps (Claret y Rouviere-Yaniv, 1996; Falconi *et al*., 1996).

El nivel energético de una célula se mantiene homeostáticamente en torno a un valor durante la fase de crecimiento exponencial, independientemente de las

condiciones nutricionales, gracias al balance regulado de la actividad de proteínas que sintetizan y proteínas que consumen ATP, entre las que se incluyen las topoisomerasas (Atkison 65,69). Sin embargo, cuando la célula alcanza la fase de crecimiento estacionaria el estado de energía se mantiene a un nivel más bajo (Jensen y Pedersen, 1990) y en consecuencia el grado de superenrollamiento negativo es menor (Drlica, 1992).

La iniciación de la transcripción requiere el desenrollamiento del ADN, proceso que se ve influenciado por el grado de superenrollamiento negativo del mismo. El efecto que ejerce el grado de superenrollamiento sobre cada promotor depende, entre otros factores de la secuencia de mismo, de modo que no todos los promotores ven afectada su expresión de igual modo. Valores altos de superenrollamiento negativo estimulan la transcripción de algunos genes, mientras que la expresión de otros está favorecida por valores bajos del mismo (Bordes *et al*., 2003; Kusano *et al*., 1996; Ohlsen y Gralla, 1992; Peter *et al*., 2004).

### 2.3. REGULACIÓN POR MOLÉCULAS PEQUEÑAS.

La actividad de la ARNP está regulada en respuesta a cambios en las condiciones ambientales por la concentración de algunas moléculas pequeñas (Ej. (p)ppGpp, ATP, Apppp). La molécula efectora mejor estudiada es el tetrafosfato de guanosina (ppGpp). Este nucleótido se sintetiza en respuesta a una limitación de aminoácidos (respuesta estricta) u otros nutrientes, así como en respuesta a condiciones que provocan una parada del crecimiento celular. Se conocen dos proteínas implicadas en la síntesis de ppGpp, RelA y SpoT. RelA está asociada a los ribosomas y produce ppGpp en condiciones de hambre de aminoácidos, detectada por la bacteria cuando un tRNA descargado ocupa el sitio A del ribosoma (Cashel, 1996). SpoT produce ppGpp en respuesta a condiciones de estrés o de limitación de nutrientes (Magnusson *et al*., 2005). ppGpp se une directamente a la ARNP, interaccionando con las subunidades β y β', cerca del centro activo (Artsimovitch *et al*., 2004; Chatterji *et al*., 1998). La unión de ppGpp a la ARNP provoca la desestabilización del complejo promotor abierto, especialmente en aquellos promotores que forman complejos abiertos inestables, como por ejemplo, los genes de ARNr (ver más adelante). El efecto de ppGpp se ve amplificado por la unión de la proteína DksA a la ARNP, que dada su similitud con el factor de eloganción GreA, se propone que podría unirse a la ARNP a través del canal secundario de entrada de NTPs al centro activo, estabilizando el complejo ppGpp-ARNP (Paul *et al*., 2004). Por otro lado, se han descrito mecanismos alternativos de acción de la molécula de ppGpp. ppGpp altera la tasa de formación del complejo abierto, el escape del promotor, y las pausas de la transcripción durante la elongación. Se han propuesto dos posibles mecanismos por los que

ppGpp podría ejercer dichos efectos: competencia con los sustratos (NTPs) por la unión al centro activo o el emparejamiento con citosinas de la cadena no codificante del ADN. Como consecuencia de todos estos efectos, en presencia de ppGpp aumenta la cantidad de ARNP libre disponible, de manera que aquellos genes cuyos promotores presentan normalmente bajos niveles de ARNP unida a ellos aumentarían su expresión (Magnusson *et al*., 2005).

Como se describe en el apartado 1.3.5, los niveles de un determinado factor σ pueden afectar a la expresión de genes dependientes de otro factor σ distinto, debido a su competencia por la existencia de una cantidad limitante de ARNP. Resultados recientes sugieren que, además aumentar el nivel de ARNP libre, ppGpp regula la capacidad del factor sigma para unirse al núcleo de la ARNP (Jishage *et al*., 2002), ayudando a factores sigma alternativos como $\sigma^{38}$, $\sigma^{54}$ y $\sigma^{32}$ a competir con el factor sigma principal, $\sigma^{70}$, por la unión al núcleo de la ARNP (Bernardo *et al*., 2006; Grossman *et al*., 1985; Kvint *et al*., 2000). De hecho, la fracción de ARNP asociada a factores sigma alternativos se reduce en células que no producen ppGpp.

Los genes que codifican los ARNs ribosómicos constituyen un caso especial de regulación transcripcional (Figura 13). Dada la importancia del proceso en el que intervienen, están regulados de forma precisa por la velocidad de crecimiento, que se traduce en cambios en las concentraciones de ppGpp. Estos genes suelen presentar promotores dependientes de $\sigma^{70}$, a los que se encuentra asociada la mayor parte de la ARNP de la célula, y forman complejos promotor abierto inestables especialmente sensibles a la desestabilización provocada por la unión de ppGpp a la ARNP (Gralla, 2005; Kajitani y Ishihama, 1984). La acumulación de altos niveles de ppGpp como consecuencia de un estrés o de la entrada en la fase estacionaria de crecimiento, desestabiliza la asociación ARNP/ $\sigma^{70}$ e incrementa los niveles de núcleo de ARNP disponible para asociarse a factores sigma alternativos. De esta manera, la molécula de ppGpp también juega un papel como regulador positivo de la expresión de genes dependientes de ARNP asociadas a factores sigma alternativos (Kvint *et al*., 2000; Laurie *et al*., 2003) y de determinados genes dependientes de la ARNP/$\sigma^{70}$, que codifican proteínas relacionadas con la respuesta a estrés (Kvint *et al*., 2003), o con la biosíntesis o el transporte de aminoácidos (Cashel, 1996), que presentan promotores menos sensibles a la concentración de ppGpp.

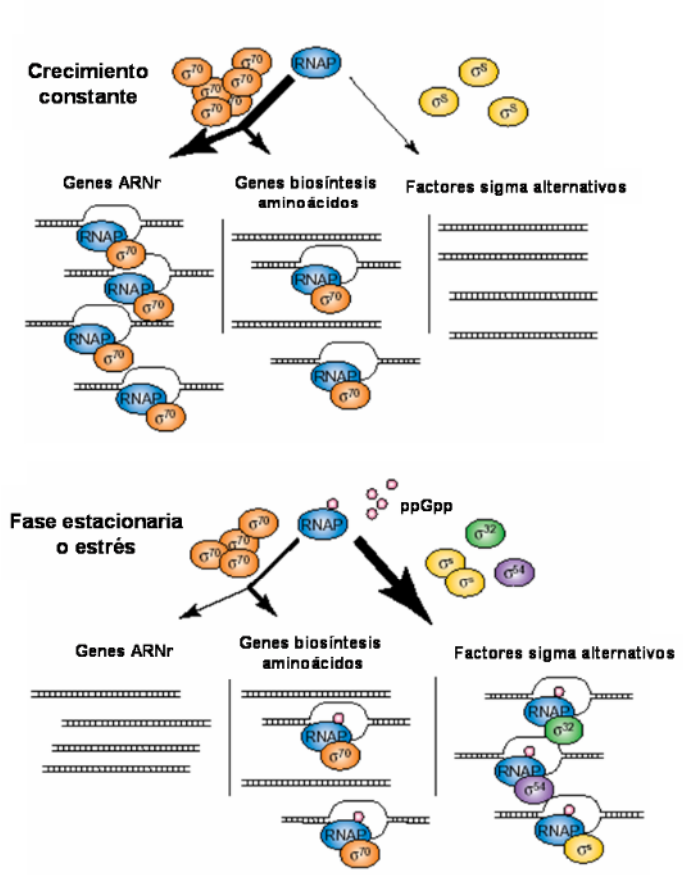



**Figura 13. Modelo de regulación transcripcional mediado por ppGpp.** Durante la fase de crecimiento exponencial la mayor parte de la ARNP se encuentra asociada al factor σ70, transcribiendo genes de ARNr (ARN ribosómicos). Con la entrada en la fase de crecimiento estacionaria o en condiciones de estrés, los niveles de ppGpp aumentan, lo que provoca un aumento de la cantidad de núcleo de ARNP disponible para su asociación con factores sigma alternativos. Obtenida de Magnusson *et al*., 2005.

## 3. INTEGRACIÓN DE RESPUESTAS REGULATORIAS.

En su entorno natural, los microorganismos están sometidos a una dura competencia por los nutrientes disponibles. Las bacterias son capaces de adaptarse rápidamente a cambios en las condiciones ambientales, percibiendo e integrando diferentes señales, como las relacionadas con la limitación de nutrientes, los estreses físico-químicos (tensión de agua u oxígeno, cambios en el pH o en la temperatura ambiente y, en algunos casos, la toxicidad de determinados compuestos). Estas capacidades son consecuencia de la diversidad y plasticidad genéticas propias de las bacterias, que les confieren un amplio arsenal de actividades bioquímicas y de mecanismos de regulación (Ellis, 2000).

Las bacterias del género *Pseudomonas* son un claro ejemplo de integración de señales reguladoras. Estos microorganismos se caracterizan por presentar una gran proporción de genes dedicados al catabolismo, transporte y eflujo de compuestos orgánicos (Nelson *et al*., 2002). Para que la bacteria pueda utilizar la fuente de

carbono más favorable energéticamente disponible en una situación dada, las rutas catabólicas, incluidas las de compuestos aromáticos, deben funcionar eficientemente dentro del contexto del metabolismo central de la bacteria, de manera que ésta utilice en cada situación la fuente de carbono más favorable energéticamente. Por ello, la síntesis de las enzimas que conforman una ruta catabólica especializada sólo debe producirse cuando el compuesto que degradan esté presente, estando además su expresión sometida a mecanismos de represión catabólica que responden al estado nutricional y energético de la bacteria (Aranda-Olmedo *et al*., 2005; Cases y de Lorenzo, 1998; Díaz y Prieto, 2000; Duetz *et al*., 1996). Por otro lado, el estrés originado por la presencia del compuesto aromático provoca en la bacteria otro tipo de respuestas que están interconectadas con la respuesta de degradación del compuesto (Sze y Shingler, 1999; Velázquez *et al*., 2005). Y por último, también se han descrito respuestas quimiotácticas a la presencia de compuestos aromáticos, mediadas por distintos sistemas receptores de membrana que detectan el compuesto, alterando el patrón de movilidad de la bacteria (Parales y Harwood, 2002). Para la integración de cada una de estas señales la bacteria emplea una red de reguladores que responden específicamente al compuesto activando grupos concretos de genes, y una serie de reguladores globales, que supeditan la activación de grupos específicos de genes al metabolismo central de la bacteria. De forma paralela, la bacteria altera los niveles de factores sigma alternativos y sus afinidades relativas por la ARNP para generar una respuesta que permita la supervivencia bacteriana y el crecimiento celular utilizando de forma óptima los recursos disponibles (Shingler, 2003). Un ejemplo bien estudiado de esta integración entre regulación global y específica es el operón TOL de *Pseudomonas putida*. En esta tesis doctoral hemos analizado en detalle la regulación de la ruta *meta* de dicho operón, que está bajo el control de XylS, un regulador específico perteneciente a la familia AraC.

## 4. LA FAMILIA DE REGULADORES TRANSCRIPCIONALES AraC.

La familia de reguladores transcripcionales AraC engloba a más de 1000 proteínas, la mayoría de ellas activadores de la transcripción (Gallegos *et al*., 1997; Tobes y Ramos, 2002). Las proteínas de esta familia comparten un dominio conservado de unos 99 aminoácidos de longitud en el que se encuentran los determinantes de unión al ADN. Además, la mayoría también presenta un segundo dominio encargado del reconocimiento de las moléculas efectoras y que porta los determinantes de dimerización. En aquellos miembros de la familia que presentan sólo el dominio de unión a ADN, y que se comportan como monómeros en solución, la respuesta a la señal inductora ocurre a través de un segundo regulador que controla la expresión del primero. Las proteínas de la familia AraC están implicadas en la regulación de procesos celulares tan diversos como el metabolismo del

carbono, la respuesta a estrés y numerosos mecanismos de virulencia (Gallegos *et al*., 1997).

### 4.1. Estructura del dominio de unión a ADN.

A pesar de haberse identificado más de 1000 miembros de la familia (Egan02), sólo se dispone de la estructura cristalina de dos de ellos, MarA (Rhee *et al*., 1998) y Rob de *E. coli* (Kwon *et al*., 2000). Ambas proteínas pertenecen al subgrupo de reguladores implicados en la regulación de procesos de estrés. MarA, con una longitud de 129 aminoácidos, consta de un único dominio encargado del reconocimiento de secuencias específicas en los promotores. El análisis por difracción de rayos X de los cristales del complejo MarA-ADN revela que MarA se une al ADN como monómero mediante dos motivos hélice-giro-hélice (HTH) que se insertan en surcos mayores consecutivos del ADN (Rhee *et al*., 1998). La estructura secundaria de MarA consta de siete α–hélices que se pliegan formando dos subdominios estructurales similares, cada uno de los cuales presenta un motivo HTH de unión al ADN. Las dos hélices de reconocimiento de los motivos HTH (α–hélices 3 y 6) se proyectan hacia la misma cara de la proteína, de modo que MarA se une a una sola cara del ADN (Figura 14). La distancia entre las hélices de reconocimiento (27 Å) es menor que la distancia entre surcos mayores consecutivos del ADN en su conformación B, de manera que la unión de MarA a su sitio de unión provoca una curvatura del ADN de aproximadamente 35º. MarA establece interacciones de tipo puente de hidrógeno y fuerzas de van der Waals entre los residuos de las hélices de reconocimiento y las bases y el esqueleto de fosfato localizados en los surcos mayores del ADN. Ambos tipos de interacciones son responsables de la estabilidad de la unión de MarA al ADN (Gillette *et al*., 2000).

La familia de reguladores AraC es el primer caso de reguladores transcripcionales procariotas con dos motivos HTH, frecuentes en eucariotas. Sin embargo, la estructura de los reguladores de la familia difiere de los reguladores eucariotas en que la disposición de la hélice 4 que conecta los dos motivos HTH impone la orientación y la distancia entre ellos, determinando la unión a la misma cara del ADN y la curvatura del mismo. En los reguladores eucariotas, por el contrario, los dos motivos HTH están conectados por lazos flexibles, de manera que pueden unirse al ADN en disposiciones muy variadas (Rhee *et al*., 1998).

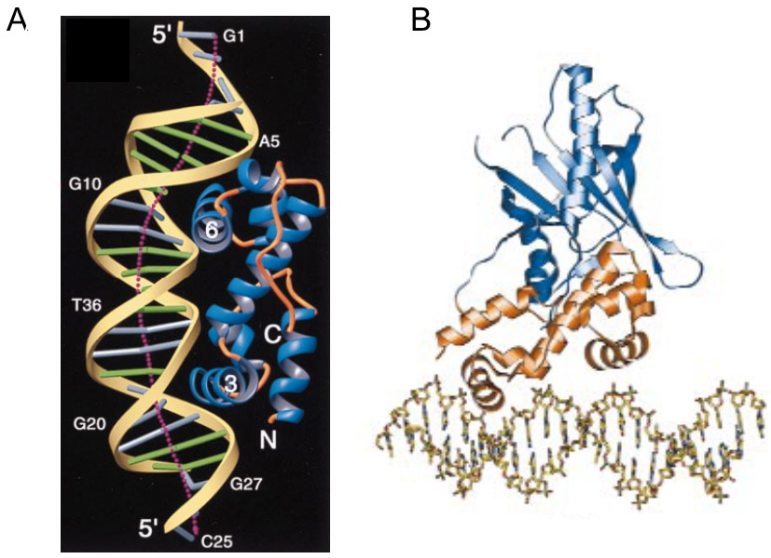



**Figura 14. Estructura tridimensional de MarA y Rob en complejo con el ADN. A.** La estructura tridimensional de MarA en complejo con el ADN revela la presencia de dos motivos HTH que establecen interacciones con surcos mayores adyacentes en el ADN. En la figura se observa la curvatura de 35º en el ADN inducida por la unión de MarA. **B.** La estructura cristalina de Rob en complejo con el ADN muestra el primer motivo HTH interaccionando con un surco mayor del ADN, mientras que el segundo motivo HTH interacciona con el ADN fuera del surco mayor.

Rob, a diferencia de MarA, posee dos dominios funcionales con la particularidad adicional de que el dominio de unión al ADN, al contrario de lo que ocurre en la mayoría de miembros de la familia, se localiza en el extremo amino terminal de la proteína. Según la estructura de Rob en complejo con el ADN, el activador interacciona con el ADN a través de sus dos HTH, aunque en el caso del segundo HTH, situado hacia el extremo carboxilo terminal de la proteína, éste no parece interaccionar con el surco mayor del ADN (Kwon *et al*., 2000). Sin embargo, datos experimentales adicionales apoyan la unión de Rob a dos surcos mayores adyacentes provocando una curvatura en el ADN similar a la que provoca la unión de MarA, lo que sugiere que la estructura Rob-ADN descrita podría no ser correcta (Li y Demple, 1996).

### 4.2. Estructura del dominio de reconocimiento del efector en AraC.

El dominio de reconocimiento del efector no está conservado entre los miembros de la familia, y sólo en el caso de la proteína AraC se ha determinado su estructura tridimensional. El dominio amino terminal de AraC ha sido cristalizado tanto en su forma apo (o libre de efector), como formado complejo con su efector, una molécula de arabinosa (Soisson *et al*., 1997). La estructura muestra que este dominio amino contiene 8 láminas β antiparalelas que se pliegan formando un barril β con topología “jelly-roll”, conectadas por dos vueltas de hélice de tipo $3_{10}$ a una

novena lámina beta que también forma parte del barril. Por último, el extremo carboxilo del dominio amino contiene dos α-hélices de unos 20 amino ácidos de longitud unidas a la última lámina β, que se empaquetan contactando con la superficie externa del barril. La molécula de arabinosa se localiza en un bolsillo de unión, dentro del barril β, completamente enterrada en la estructura. El dominio amino terminal de AraC presenta además un brazo amino terminal, formado por los aminoácidos 7 al 18, que cierra el bolsillo de unión de arabinosa. La estructura del brazo amino terminal no se ha podido determinar en ausencia de arabinosa, por lo que se supone que podría constituir una zona flexible y, por lo tanto, desorganizada en la estructura cristalina. En lo que se refiere al resto de la estructura, no se observan cambios significativos entre presencia y ausencia del efector, pero sí ocurren cambios significativos en el modo de dimerización entre ambas formas (Soisson *et al*., 1997). En presencia de arabinosa, el dominio amino dimeriza en una conformación denominada "coiled-coil", en la que los contactos entre los dos monómeros ocurren a través de sus correspondientes α-hélices en el extremo carboxilo dispuestas en orientación antiparalela, cuyos extremos están anclados por la interacción entre aminoácidos de leucina de ambos monómeros, conformación que se ha denominado "knobs-into-holes" (Figura 15) (Soisson *et al*., 1997).

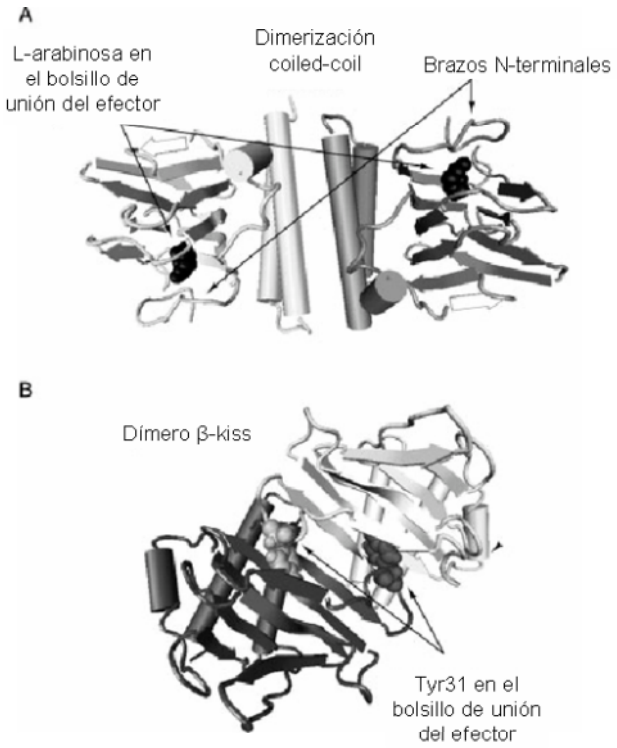



**Figura 15. Estructura tridimensional de los dímeros formados por dos dominios amino terminales de AraC.** A. En presencia de arabinosa, AraC dimeriza estableciendo interacciones a través de las α-hélices de la región "coiled-coil". El brazo amino terminal establece interacciones con la molécula de arabinosa, cerrando el bolsillo de unión del efector. B. En ausencia de arabinosa y a concentraciones no fisiológicas de proteína AraC dimeriza a través de un residuo de Tyr localizado en el bolsillo de unión del efector. Obtenido de Weldon *et al*. 2007.

En ausencia de arabinosa, el dominio amino terminal de AraC dimeriza formando una estructura denominada "β–kiss", en la cual la interacción entre monómeros se produce por contacto entre sendos aminoácidos Tyr31, que se localizan en el interior de los bolsillos de unión de arabinosa. Este tipo de dimerización da lugar a una oligomerización ilimitada de monómeros del dominio amino de AraC que tendría lugar sólo a concentraciones no fisiológicas de la proteína. Estos resultados sugieren que la asociación a través de interacciones entre los residuos Tyr31 no es relevante desde un punto de vista fisiológico (Weldon *et al*., 2007). Por lo tanto, el mecanismo de activación de AraC y, en concreto, la respuesta de la proteína a la presencia de arabinosa, viene determinada por las interacciones alternativas que establece el brazo amino-terminal de AraC con el bolsillo de unión del efector o con el dominio de unión a ADN (Harmer *et al*., 2001; Saviola *et al*., 1998; Wu y Schleif, 2001).

### 4.3. Mecanismo de acción de los reguladores de la familia AraC.

Sólo se han estudiado en detalle algunos reguladores de la familia AraC implicados en la regulación del metabolismo del carbono y en la respuesta a estrés, como por ejemplo: AraC, MarA, MelR, XylS, Rob, SoxS, Ada, RhaR y RhaS (revisado en (Martin y Rosner, 2001). En cuanto al subgrupo implicado en procesos de patogénesis se han realizado avances considerables en el estudio de algunos reguladores como Rns, ToxT, VirF, PerA e InvF. Los mecanismos por los que los reguladores de la familia controlan la expresión desde sus promotores específicos son variados y sólo se conocen en detalle en algunos casos. A continuación se describen algunos ejemplos de los mecanismos de activación mejor estudiados.

#### 4.3.1. Regulación mediante lazo en el ADN.

Las proteínas reguladoras AraC y MelR actúan alternativamente como represores en ausencia de efector, o como activadores en presencia del mismo. En ausencia de efector la conformación de la proteína mantiene los dominios carboxilo terminales del dímero en una orientación que permite su unión a sitios distales en el promotor, provocando la formación de un lazo en el DNA (Figura 16 Aa y Ba).

Cuando AraC y MelR se unen a sus respectivos efectores a través del extremo amino terminal de la proteína, se produce un cambio de conformación que reorienta los dominios carboxilo terminales, permitiendo su unión a sitios adyacentes en el promotor. Por lo tanto la presencia del efector tiene como consecuencia la desaparición del lazo en el DNA (Figura 16 Ab y Bb), lo que permite el reclutamiento de la ARNP al promotor y la estimulación de la formación del complejo abierto (Belyaeva *et al*., 2000; Kahramanoglou *et al*., 2006; Lobell y Schleif, 1990; Soisson *et al*., 1997; Wade *et al*., 2000).

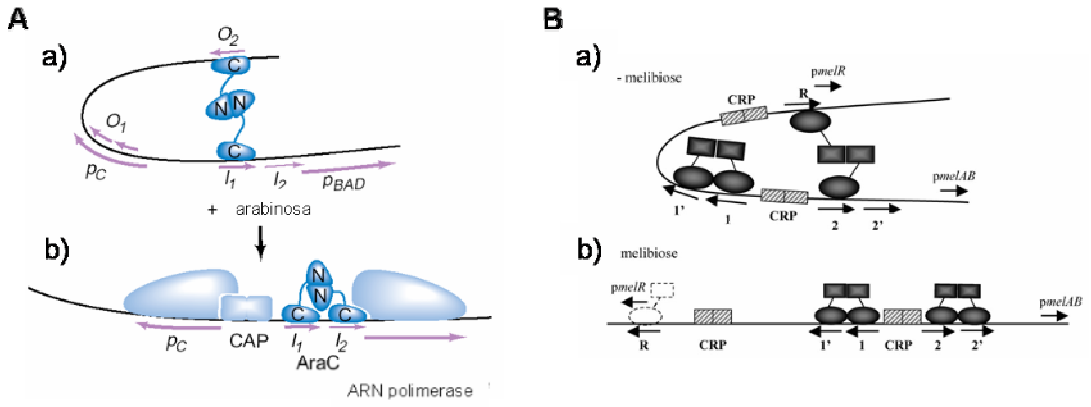



**Figura 16. Modelos de regulación de AraC y MelR sobre sus respectivos promotores pBAD y pMelA.** A. El sitio de unión de AraC $I_2$ está situado solapando con la región -35 de promotor pBAD. Los sitios $O_1$ y $O_2$ están situados en posiciones -125 y -285 con respecto al punto de inicio de la transcripción. N: dominio amino terminal; C: dominio carboxilo terminal. Obtenido de Soisson *et al*., 1997. B. En ausencia de melibiosa, MelR no es capaz de unirse al sitio 2' que solapa con la región -35 del promotor pmelA. La unión de MelR a los sitios 2 y R genera un lazo en el ADN que provoca una represión de la transcripción desde pmelA y pmelR. En presencia de melibiosa, MelR cambia de conformación de manera que reconoce el sitio 2', activando la transcripción desde pmelA. Obtenido de Kahramanoglou *et al*., 2006.

### 4.3.2. Activación dependiente de la concentración del regulador.

Una característica que comparten todos los reguladores de la familia estudiados hasta el momento es la estrecha regulación a la que está sometida su síntesis (Belyaeva *et al*., 2000; Egan y Schleif, 1993; González-Pérez *et al*., 2004; Schleif, 1992). En el caso de las proteínas MarA, Rob y SoxS, la regulación de su síntesis resulta crucial dado que activan los mismos promotores atendiendo a la concentración intracelular de cada una de ellas y a sus afinidades relativas por los sitios de unión (Li y Demple, 1996; Martin *et al*., 2000). En el caso de MarA y SoxS su concentración intracelular es baja en condiciones normales y aumenta unas 25 veces en respuesta a un estrés oxidativo (Griffith *et al*., 2002; Martin *et al*., 2002). Cuando el estrés cesa, los niveles de expresión bajan y dada la corta vida media de estas proteínas, sus niveles intracelulares vuelven rápidamente a su estado basal (Griffith *et al*., 2004). Tanto MarA como SoxS parecen activar la transcripción mediante un mecanismo alternativo conocido como pre-reclutamiento, de manera que se asocian a la ARNP en solución para después reclutarla a sus promotores específicos (Griffith y Wolf, 2004; Martin *et al*., 2002).

### 4.4. Contactos de los reguladores de la familia AraC con la ARNP.

Entre los activadores de la familia AraC encontramos tanto activadores de clase I como de clase II. La localización de los sitios de unión del activador en el promotor determina su orientación con respecto a la ARNP y, por lo tanto, también los contactos que puede establecer con ella. MarA, Rob y SoxS se unen a sus promotores como monómeros en dos orientaciones distintas dependiendo del promotor. En los casos en los que el sitio de unión se localiza a una distancia de más de 30 pb corriente arriba de la región -10 del promotor, es decir en promotores de clase I, la orientación del regulador puede ser tanto directa (α-hélice 6 de reconocimiento interaccionando con la región del operador más próxima al promotor) como reversa (α-hélice 3 más cerca del promotor), y los posibles contactos se establecen con el dominio carboxilo terminal de la subunidad α de la ARNP. Cuando la unión se produce a una distancia menor de 19 pb respecto a la región -10, correspondiente a promotores de clase II, la orientación directa es la única que se produce y en este caso los contactos con la ARNP se podrían establecer a través de la región 4.2 de la subunidad σ (Gillette *et al*., 2000; Martin *et al*., 1999; Wood *et al*., 1999).

El grupo de activadores de la familia AraC relacionado con la regulación del metabolismo del carbono regula mayoritariamente promotores de clase II, en los que la orientación del monómero proximal respecto de la ARNP es directa, de manera que los contactos con la subunidad sigma se establecen entre aminoácidos cargados negativamente localizados en la hélice 5 del dominio carboxilo terminal y aminoácidos cargados positivamente situados en la región 4.2 de la subunidad σ de la ARNP (Bhende y Egan, 2000; Grainger *et al*., 2004b; Landini y Busby, 1999). En algunos casos, se ha demostrado que el dominio carboxilo terminal de la subunidad α de la ARNP establece contactos con este tipo de reguladores cumpliendo una función en el proceso de activación (Grainger *et al*., 2004a; Holcroft y Egan, 2000; Ruíz *et al*., 2001).

### 4.5. La proteína reguladora XylS.

La proteína XylS, objeto de estudio de esta tesis doctoral, es el activador transcripcional de la ruta *meta* codificada en el plásmido TOL pWW0 de *P. putida*. El megaplásmido TOL pWW0 le confiere a la cepa *P. putida* KT2440 la capacidad de degradar tolueno y alquilbenzoatos hasta intermediarios del ciclo de Krebs (Ramos *et al*., 1997). La oxidación secuencial del grupo metilo del tolueno y sus derivados hasta el correspondiente ácido carboxílico se produce en una serie de reacciones que constituyen la llamada ruta *upper* del plásmido TOL pWW0. El catabolismo de los correspondientes ácidos carboxílicos hasta intermediarios del ciclo de Krebs tiene lugar a través de la llamada ruta de rotura en *meta*, en adelante *meta*. Las

enzimas que catabolizan las distintas reacciones de oxidación se encuentran codificadas en dos operones denominados de forma análoga *upper* y *meta* (Figura 17 A).

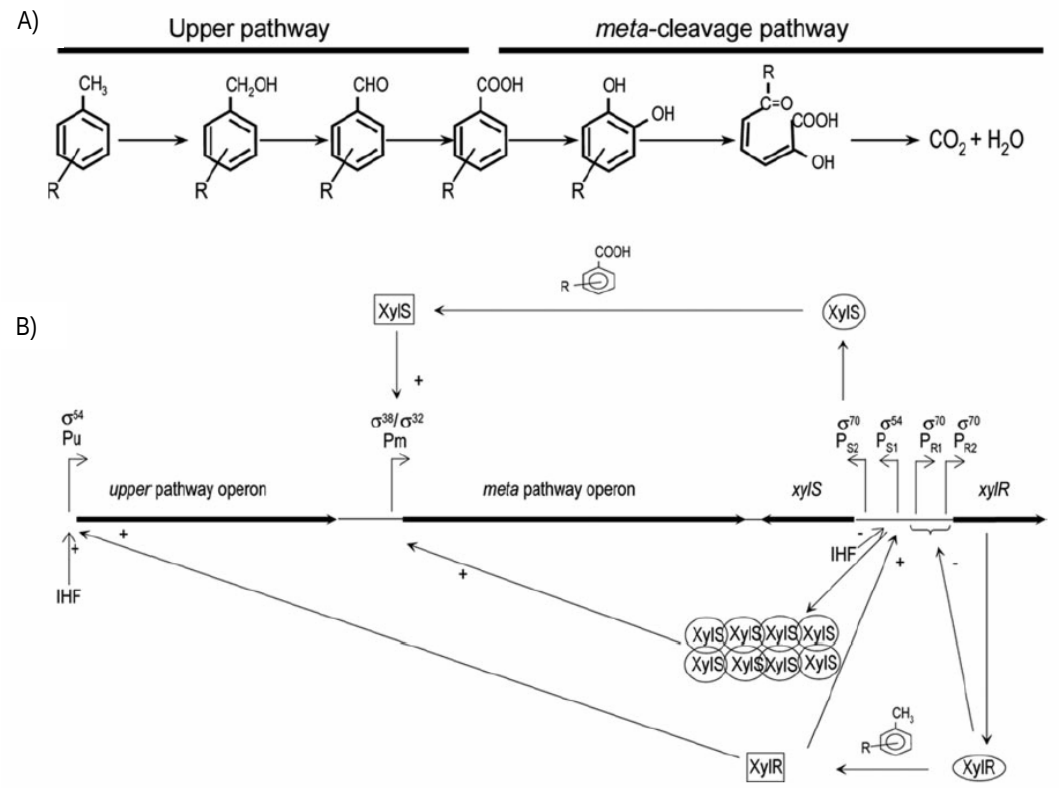



**Figura 17. Organización de los operones *upper* y *meta* del plásmido TOL pWW0 y de sus respectivos genes reguladores *xylS* y *xylR*.** A. Esquema de las reacciones oxidativas que forman parte de las rutas *upper* y *meta* y que permiten la degradación de tolueno y derivados hasta intermediarios del ciclo de Krebs. B. Esquema de la regulación transcripcional de los operones upper y meta y de los genes reguladores xylS y xylR. La participación de distintos factores sigma y del regulador IHF se indica para cada promotor. Extraído de Domínguez-Cuevas, *et al*. 2006.

Corriente abajo del operón *meta* se localizan dos genes reguladores, *xylS* y *xylR*, que se transcriben de forma divergente entre sí y que codifican las proteínas reguladoras de dichas rutas. La proteína XylR pertenece a la familia de reguladores transcripcionales EBP. La transcripción del gen *xylR* se produce a partir de dos promotores en tandem, $P_{R1}$ y $P_{R2}$, dependientes del factor $\sigma^{70}$, cuya regulación depende de la propia proteína XylR. En presencia de un efector de XylR, como el tolueno, la proteína adquiere su forma activa y estimula la transcripción desde dos promotores, el promotor de la ruta *upper* Pu y uno de los promotores del gen *xylS*, $P_{S1}$. Dicha activación depende de la ARNP asociada al factor $\sigma^{54}$ y del regulador global IHF entre otros factores. A su vez, la unión de XylR al promotor $P_{S1}$, provoca

la represión de los promotores $P_{R1}$ y $P_{R2}$, de manera que XylR reprime su propia expresión (Marqués *et al*., 1998). La transcripción del gen *xylS* a partir del promotor $P_{S1}$ produce la sobreexpresión de la proteína XylS, que incluso en ausencia de sus efectores adquiere una conformación activa, estimulando la transcripción desde el promotor Pm en lo que se conoce como cascada de activación (González-Pérez *et al*., 2004; Ramos *et al*., 2002).

La proteína XylS pertenece a la familia de reguladores transcripcionales AraC. En ausencia de efector, el gen *xylS* se expresa a bajo nivel a partir del promotor $P_{S2}$ dependiente de $\sigma^{70}$, lo que produce un nivel bajo de XylS en la célula. En estas condiciones la proteína XylS se encuentra en forma inactiva; en presencia de uno de sus efectores, como por ejemplo el 3-metilbenzoato (3MB), adquiere su forma activa estimulando la transcripción desde el promotor Pm de la ruta *meta* (Figura 17 B) y manteniendo constante la expresión de los genes de la ruta *meta* a lo largo de la curva de crecimiento. La transcripción desde el promotor Pm depende de dos ARN polimerasas asociadas a los factores sigma alternativos, $\sigma^{32}$ y $\sigma^{38}$, de manera que la ARNP/$\sigma^{32}$ permite la expresión al comienzo de la fase exponencial de crecimiento, mientras que la ARNP/$\sigma^{38}$ es la responsable de la transcripción durante las fases exponencial tardía y estacionaria (Marqués *et al*., 1995; Marques *et al*., 1999). El promotor Pm contiene dos repeticiones directas imperfectas de 15 pb que constituyen los sitios de unión del regulador XylS y que se localizan entre las posiciones -70 y -34 del promotor, solapando con la región -35 del mismo (González-Pérez *et al*., 2002; González-Pérez *et al*., 1999). Cada motivo de 15 pb está compuesto por dos cajas de secuencia conservada, denominadas cajas A y B.

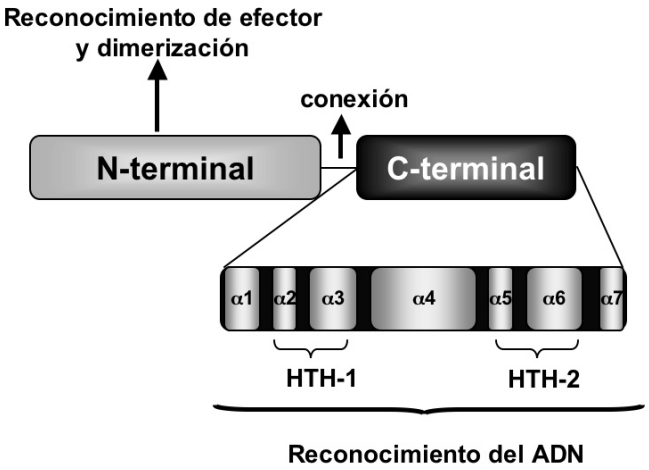



**Figura 18. Esquema de los dominios funcionales de la proteína XylS.** El dominio C-terminal, conservado en la familia AraC/XylS, está formado por 7 α-hélices que se pliegan formando dos motivos HTH de unión al ADN. El dominio N-terminal está implicado en los procesos de dimerización y reconocimiento del efector. Ambos dominios están conectados por una zona flexible de la proteína, con una longitud de 9 aminoácidos.

XylS es una proteína de 321 aminoácidos, dividida en dos dominios funcionales. El dominio carboxilo terminal presenta homología con el dominio conservado de la familia AraC y contiene todos los determinantes necesarios para la activación de la transcripción aunque no responde a la presencia de 3MB (Kaldalu *et al*., 2000). El dominio amino terminal presenta homología en su extremo carboxilo con las α-hélices implicadas en dimerización de la proteína AraC. De acuerdo con esta homología, mutantes en residuos de leucina localizados en posiciones homólogas dentro de dichas hélices resultan en ambos casos en proteínas que presentan defectos en la formación de dímeros (Ruíz *et al*., 2003; Wu y Schleif, 2001). Además se ha demostrado la implicación del dominio amino terminal de XylS en el reconocimiento del efector y en el establecimiento de contactos con la subunidad α de la ARNP (Michán *et al*., 1992; Ramos *et al*., 1990; Ruíz y Ramos, 2001; Ruíz y Ramos, 2002). Ambos dominios están conectados por una secuencia de nueve aminoácidos. Se ha estudiado el papel del 3MB en la dimerización de XylS y, de acuerdo con los resultados obtenidos por Ruíz y colaboradores, la de 3MB favorece la formación de dímeros, aunque el proceso también puede ocurrir en ausencia del efector cuando se alcanza una concentración de proteína umbral.

# OBJETIVOS

A lo largo de esta Tesis Doctoral se han abordado distintos aspectos de la regulación de la expresión de la ruta *meta* del plásmido TOL de *Pseudomonas putida* KT2440 por la proteína activadora XylS. En trabajos anteriores del grupo se determinaron las secuencias de unión de XylS en el promotor Pm del operón *meta* (González-Pérez *et al*., 1999), así como la expresión dependiente de dos ARN polimerasas alternativas atendiendo a la fase de crecimiento (Marqués *et al*., 1999). Con el fin de profundizar en los mecanismos específicos y globales de regulación a que está sujeta la ruta *meta* se plantearon los siguientes objetivos concretos:

1. Establecer las bases moleculares que determinan el reconocimiento del promotor Pm por dos ARN polimerasas alternativas.

2. Determinar el papel de los procesos de dimerización y unión al efector en la activación de la expresión por la proteína XylS.

3. Identificar los contactos que establece la proteína XylS con sus sitios de unión en el ADN.

4. Caracterización bioquímica y funcional del dominio de unión al ADN en XylS.

5. Estudio de la respuesta global de la cepa *Pseudomonas putida* KT2440 (pWW0) a la presencia de compuestos aromáticos.

# RESULTADOS

CAPÍTULO 1

# Dos ARN polimerasas pueden compartir un único sitio de inicio de la transcripción en el promotor Pm

La secuencia de nucleótidos en la región -7 a -18 del promotor determina el reconocimiento por dos ARN polimerasas alternativas asociadas a los factores $\sigma^{32}$ y $\sigma^{38}$

**Patricia Domínguez Cuevas, Patricia Marín, Juan L. Ramos, y Silvia Marqués**

La transcripción desde el promotor Pm de la ruta de rotura en *meta* para la degradación de alquilbenzoatos se produce gracias a dos ARNP alternativas, $E\sigma^{32}$ y $E\sigma^{38}$ dependiendo de la fase de crecimiento, que comparten un único punto de inicio de la transcripción. Dicha transcripción se activa mediante la unión de la proteína reguladora XylS al promotor Pm en respuesta a su efector, el 3-metilbenzoato (3MB). Para dilucidar los determinantes en la secuencia que permiten la interacción de las dos ARN polimerasas con Pm, generamos todos los mutantes posibles a lo largo de la secuencia de Pm entre las posiciones -7 a -18. La actividad transcripcional de dichos mutantes se determinó en las fases exponencial y estacionaria de crecimiento. Los resultados obtenidos nos han permitido delimitar el elemento -10 de Pm a las posiciones comprendidas entre -7 y -12, definiendo un promotor que comparte características de los consensos de $\sigma^{32}$ y $\sigma^{38}$. Las dos primeras y la última posición dentro del hexámero centrado en la posición -10 son cruciales para el reconocimiento por ambas polimerasas. La posición -10 es la única reconocida de manera específica por $E\sigma^{38}$, mientras que las posiciones -8, -9 y la serie de cuatro Cs en posiciones -14 a -17 son importantes para el reconocimiento específico por $E\sigma^{32}$.

El análisis de los niveles de los distintos factores sigma reveló que $\sigma^{32}$ sólo se encontraba presente durante la fase exponencial de crecimiento y que $\sigma^{38}$ podía ser detectado en la fase estacionaria temprana. En un fondo genético carente del factor $\sigma^{32}$ (*E. coli* KY1429), el factor $\sigma^{38}$ también estaba presente durante la fase exponencial de crecimiento, tanto en su forma libre como asociado al núcleo de la ARN polimerasa. El promotor Pm parece estar optimizado para el reconocimiento por $\sigma^{32}$ en respuesta al estrés provocado por la adición del efector (3MB) al medio de cultivo, y permite la unión de una ARNP asociada al factor $\sigma^{38}$, más adaptable a cambios con respecto a su consenso, en la fase estacionaria. En conclusión, el mecanismo que controla la expresión desde Pm depende de la secuencia específica de nucleótido en la región -10, del cambio en los niveles de los factores $\sigma^{32}$ y $\sigma^{38}$ libres y asociados al núcleo de la ARN polimerasa a lo largo de la curva de crecimiento, y de la presencia del regulador XylS activado por un efector.

*The Journal of Biological Chemistry*
Vol. **280**, Nº. 50, pp. 41315-41323, December 16, 2005

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 50, pp. 41315–41323, December 16, 2005


# RNA Polymerase Holoenzymes Can Share a Single Transcription Start Site for the Pm Promoter

## *CRITICAL NUCLEOTIDES IN THE −7 TO −18 REGION ARE NEEDED TO SELECT BETWEEN RNA POLYMERASE WITH $\sigma^{38}$ OR $\sigma^{32}$**

Received for publication, May 17, 2005, and in revised form, September 28, 2005 Published, JBC Papers in Press, October 17, 2005, DOI 10.1074/jbc.M505415200

**Patricia Domínguez-Cuevas**[1], **Patricia Marín, Juan L. Ramos, and Silvia Marqués**[2]

*From the Department of Biochemistry and Molecular and Cellular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, Apartado de Correos 419, E-18008 Granada, Spain*

**The Pm promoter of the benzoate *meta*-cleavage pathway is transcribed with E$\sigma^{32}$ or E$\sigma^{38}$ according to the growth phase, with an identical transcriptional start site. To investigate sequence determinants in the interaction between either of the two RNA polymerases and Pm, all possible single mutants between positions −7 and −18 were generated, and the activity in the exponential and stationary phases of the resulting mutant promoters was compared. The results precisely delimited a −10 element between positions −7 and −12 (TAGGCT), which defined a promoter sharing nucleotides with both $\sigma^{38}$ and $\sigma^{32}$ consensus. The first two and the last positions of this hexamer were crucial for recognition by both polymerases. Position −10 was the only one specifically recognized by E$\sigma^{38}$, whereas positions −8, −9, and the C-track between positions −14 and −17 were important for specific E$\sigma^{32}$ recognition. Western blots showed that $\sigma^{32}$ was only detectable in the exponential phase, and $\sigma^{38}$ appeared in the early stationary phase. In the *rpoH* mutant KY1429, $\sigma^{38}$ was already present in the exponential growth phase both free and bound to the RNA polymerase core, in good correlation with the transcription levels found. Pm seems to be optimized for recognition by $\sigma^{32}$ as an initial response to the addition of effector to the medium and allows binding of the adaptable $\sigma^{38}$-dependent RNA polymerase in the stationary phase. XylS is always required for Pm transcription. Therefore, the mechanism that controls Pm expression involves specific nucleotide sequences, the abundance of free and core-bound $\sigma^{32}$ and $\sigma^{38}$ factors during growth, and the presence of the regulator activated by an effector.**

The TOL plasmid pWW0 of *Pseudomonas putida* specifies a *meta*-cleavage pathway for the oxidative catabolism of benzoates and toluates. Genes encoding the TOL *meta*-cleavage pathway are grouped in a single operon under the control of the Pm promoter. Expression from Pm is positively regulated at the level of transcription by substrate-activated XylS, a regulator belonging to the AraC family (1–6). Genetic analyses established that XylS recognizes two 15-bp imperfect direct repeated motifs (5′-TGCAAPuAAPyGGNTA-3′) extending from −69 to −55 and from −48 to −34 in the Pm region (Fig. 1). Single point mutations in the binding site revealed that nucleotides located at −48 to −45 and at −58, −59, −61, and −69 are the most critical bases for appropriate XylS-Pm interactions (7–9). *In vitro* footprints obtained with a tagged XylS protein immunoadsorbed onto glass beads supported this organization (10). The arrangement of the two motifs is such that the proximal XylS binding site overlaps by 1 bp the RNA polymerase binding site at −35 (11) (Fig. 1).

The Pm promoter seems to be a Class II promoter in which complex interactions occur between the RNA-polymerase and its cognate transcriptional regulator (12). *In vivo* and *in vitro* methylation assays of Pm show extensive methylation of T at position −41 in the bottom strand, suggesting the presence of a key distortion point that may favor XylS/RNA polymerase interactions (8). In this connection, it has been shown that XylS contacts residues 291 and 289 of the α-subunit of RNA polymerase (13). The Pm promoter is unique in that *in vivo* transcription is mediated by RNA polymerase with different alternative σ factors. Transcription from the Pm promoter in the early exponential growth phase is mediated by RNA polymerase with $\sigma^{32}$, but a switch to $\sigma^{38}$ takes place in the late exponential and early stationary phases. Regardless of the gr[54] se, expression from Pm remains dependent on 3MB[3]-activated XylS, and the transcription initiation point is unchanged (14, 15).

The earliest evidence of the involvement of $\sigma^{32}$ in Pm transcription came from the observation that, in an *rpoH* background, no expression of Pm took place in the exponential phase after induction, whereas expression increased during the stationary phase. Dependence on $\sigma^{38}$ was supported by the reduced transcription from Pm in the stationary phase in an *rpoS* mutant (14). In an *rpoH-rpoS* double mutant, only basal activity from the Pm promoter was detected along the growth curve. Analyses of transcription using combinations of mutant Pm promoters and mutant XylS proteins confirmed that the alternative σ factors interacted directly with the Pm promoter (14, 16). The increase in $\sigma^{32}$ activity required for transcription in the exponential phase was provided through induction of the heat-shock response by the presence of the effector 3MB, which is also required for activation of the positive regulator XylS (17). Microarray experiments with 3MB-induced *P. putida* (pWW0) cells further confirmed the fast (15 min) heat-shock response to this effector.[4] From these findings, it follows that the Pm promoter should be recognized by two different RNA polymerases, and therefore it should accommodate the essential elements for recognition by both $\sigma^{38}$ and $\sigma^{32}$ in the same sequence stretch, although they are recognized at different moments during growth.

It is well established that σ factors play an essential role in programming gene expression, where the alternative σ subunits direct transcription toward specific gene sets according to growth and environmental

* This work was supported by Grants BMC2001-0515 and QLK3-2002-01923 from the Spanish Ministry of Science and Technology (MCYT). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Recipient of an I3P contract from the European Social Funds and a fellowship from the Junta de Andalucía (Andalusian Regional Government, Spain).

[2] To whom correspondence should be addressed: Estación Experimental del Zaidín, CSIC, C/o Profesor Albareda 1, E-18008 Granada, Spain. Tel.: 34-958-181600; Fax 34-958-129600; E-mail: silvia@eez.csic.es.

[3] The abbreviations used are: 3MB, 3-methylbenzoate; MU, Miller units.

[4] P. Domínguez-Cuevas, J. E. González-Pastor, S. Marqués, J. L. Ramos, and V. deLorenzo, submitted for publication.

**TABLE ONE**

**Strains and plasmids used to study RNA polymerases and transcription from the Pm promoter in *E. coli* strains**

| | Relevant characteristics | Source or reference |
|---|---|---|
| Strain | | |
| *E. coli* MC4100 | F$^-$, *araD139*, (*argF-lac*)*U169*, *rpsL150*, *relA1*, *flbB5301*, *deoC1*, *ptsF25*, *rpsR*. Sm$^R$ | Ref. 30 |
| *E. coli* RH90 | MC4100 *rpoS59*::Tn*10*. Sm$^R$Tc$^R$ | Ref. 30 |
| *E. coli* KY1429 | MC4100 *rpoH6* [Am] *zhf-50*::Tn*10*. Sm$^R$Tc$^R$ | Ref. 31 |
| *E. coli* SM25 | MC4100 *rpoH6* [Am] *zhf-50*::Tn*10 rpoS*::*kan*$^r$ Sm$^R$. Tc$^R$Km$^R$ | Ref. 15 |
| *E. coli* DH5$\alpha$ | *supE44 lacU169*(*Ø80lacZΔM15*) *hsdR17* ($r^-_K m^-_K$) *recA1 endA1 gyrA96 thi-1 relA1* | Ref. 48 |
| Plasmid | | |
| pJLR100 | Pm cloned in pEMBL9, Ap$^R$ | Ref. 4 |
| pMD1405 | Promoterless *lacZ*, Ap$^R$ | M. Drummond |
| pJLR107 | Pm::*lacZ* in pMD1405, Ap$^R$ | Ref. 4 |
| pERD103 | *xylS*, IncQ, Km$^R$ | Ref. 23 |
| pDCXylS | Gentamicin resistant pACYC177 derivative, *xylS*, Gm | P. Domínguez-Cuevas |
| pMD::Pm*x*−*zy* | Mutants cloned in pMD1405 where *x* is the original base in Pm, *z* is the position in Pm with respect to the +1, and *y* is the new base in the mutant Pm promoter. | This study |

conditions. The set of $\sigma$ factors that forms the $\sigma^{70}$ family shares regions of extensive sequence homology and is organized in similar domains and subdomains. However, the DNA sequence recognized by the holoenzyme bearing each of the subunits is different. A number of $\sigma^{38}$-dependent promoters have been analyzed in an attempt to derive a consensus sequence. No clear −35 sequence has been defined, but the −10 region exhibits distinctive features (*i.e.* a C at position −7 in the −10 hexamer and CT at positions −13/−12). The region downstream from the −10 sequence in $\sigma^{38}$-recognized promoters is rich in adenines and thymines (18–21). On the other hand, a consensus sequence for $\sigma^{32}$-dependent promoters has been defined from the compilation of 18 *Escherichia coli* heat-shock promoters (22) (Fig. 1). As far as we know, none of these eighteen promoters required specific transcriptional activators, but heat-shock stabilization of $\sigma^{32}$ was the only requirement for activation in *E. coli*.

In this study, we analyzed the Pm promoter sequence to decipher the critical features that allow its unique response to RNA polymerases with either $\sigma^{32}$ or $\sigma^{38}$. We first precisely defined the −10 hexamer from mutant phenotypes, and we showed that, depending on the $\sigma$ factor used by the RNA polymerase and despite the sole transcription start point, there are two different overlapping promoters at Pm. Our analyses made it possible to identify promoter bases that are critical for either one or both $\sigma$ factors. We also analyzed the role and influence of relative $\sigma$ factor abundance on Pm activity. A model on the functioning of Pm is proposed based on the analysis of the mutant promoters determined in this study.

## EXPERIMENTAL PROCEDURES

*Bacterial Strains, Culture Media, and Plasmids*—The strains used are listed in TABLE ONE. *E. coli* MC4100, RH90, KY1429, and SM25 were grown at 30 °C in Luria-Bertani medium supplemented, when required, with 100 $\mu$g/ml ampicillin, 25 $\mu$g/ml kanamycin, 50 $\mu$g/ml streptomycin, or 10 $\mu$g/ml tetracycline. Growth was determined turbidometrically at 660 nm. TABLE ONE also shows the plasmids used in this study and constructed previously. pERD103 is an IncQ plasmid encoding kanamycin resistance that bears the *xylS* gene (23), pJLR100 is a pEMBL9 derivative bearing the Pm promoter cloned between the EcoRI and HindIII sites (4), pMD1405 carries a promoterless *lacZ* gene and encodes resistance to ampicillin, and pJLR107 is a pMD1405 derivative bearing the Pm promoter in front of *lacZ* (4).

*DNA Techniques*—DNA preparation, digestion with restriction enzymes, analysis by agarose gel electrophoresis, isolation of fragments, ligation, transformation, and sequencing were done according to standard procedures (24) or to the manufacturer's recommendations.

*Construction of Mutant Pm Promoters*—The Pm mutant promoters were generated by overlap extension polymerase chain reaction mutagenesis as described previously (25). The internal oligonucleotide primers used for mutagenesis exhibited one mismatch with respect to the wild-type sequence. The external oligonucleotides were the so-called M13 reverse primer (5′-CAGGAAACAGCTATGACCATG-3′) or a universal primer (5′-GTTGTAAAACGACGGCCAGTG-3′). The template for each mutagenesis was 200 ng of pJLR100, and amplification conditions were as described by Higushi (26). After DNA amplification, the resulting DNA was digested with EcoRI and HindIII, and the 401-bp EcoRI-HindIII fragments containing mutant Pm sequences were inserted between the EcoRI-HindIII sites of pMD1405 to yield plasmids pMD::*Pmx-zy*, which carry in-frame Pm*::*lacZ* fusions, where *x* is the original base in Pm, located at position −*z* with respect to the transcription start site and mutated to *y*. All of the mutant Pm promoter sequences generated in this study were confirmed by DNA sequencing.

*RNA Extraction and Analysis*—Culture samples of 15 ml were harvested by centrifugation in disposable plastic tubes pre-cooled in liquid $N_2$ and were kept at −80 °C until use. RNA was extracted using the TRI Reagent® from Molecular Research Center (Madrid, Spain) and modified as follows. The lysis step was carried out at 60 °C, and a final digestion step with RNase-free DNase was added at the end of the process. The RNA concentration was determined spectrophotometrically at 260 nm. Hybridization of the single-stranded $^{32}$P end-labeled DNA primer XylX (5′-GGGTCGGTGAACATCTCGCGCTTGC-3′) ($10^5$ cpm/assay) complementary to the mRNA transcript originated from Pm and primer extension with avian myoblastasis virus reverse transcriptase were carried out as described previously (27). In all assays, 10 $\mu$g of total RNA was used as a template. cDNA products were analyzed in a urea-polyacrylamide sequencing gel, and gels were exposed to a phosphor screen (Fuji Photo Film Co, Ltd.) for 24–48 h. Phosphor screens were scanned using a phosphorimaging instrument (Molecular Imager FX, Bio-Rad). Data were quantified with Quantity One software (Bio-Rad).

*β-Galactosidase Assays*—*E. coli* strains bearing the wild-type Pm::*lacZ* or mutant Pm*::*lacZ* fusions in pMD1405 plus a compatible plasmid bearing the wild-type XylS (pERD103, TABLE ONE) were grown overnight on LB medium containing the appropriate antibiotics. Three independent clones of each strain were used. Duplicate cultures were prepared by diluting cells from overnight cultures to 1/100. After 1 h at 30 °C, one of the duplicates was supplemented with 1 mM 3MB,

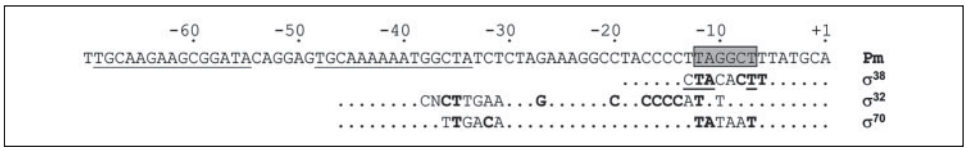



FIGURE 1. **Sequence of the Pm promoter.** The −10 hexamer as defined in this work is *boxed*, the XylS binding site is *underlined*, and the transcription start point is indicated with *+1* (2). Consensus sequences proposed for $\sigma^{38}$ (21), $\sigma^{32}$ (22), and $\sigma^{70}$ (47) are *aligned* with Pm. *Bold letters* indicate bases in each consensus that are conserved in Pm. *Underlined* bases in the $\sigma^{38}$ consensus are nucleotides present in >70% of the aligned promoters (21).

while the other one was kept as an uninduced control. Samples for β-galactosidase activity assays in the exponential phase were taken 30 min after induction when cultures had reached an $OD_{660}$ = 0.1–0.3. Five hours after induction, the cultures had reached the stationary phase ($OD_{660}$ = 2–2.5), and samples for β-galactosidase activity were taken as above. β-Galactosidase activity was determined in permeabilized whole cells according to Miller (28).

*Plasmid Copy Number*—The plasmid copy number in different strains was compared by determining plasmid molecule:chromosome molecule ratio using real-time PCR (29). The *bla gene* and the *rpoD* gene were used as the plasmid and chromosomal marker, respectively. PCR primers 5′-CGAGCGTGACACCACGATG-3′ and 5′-GCGCAGA-AGTGGTCCTG-3′ were used to amplify the *bla* gene, and 5′-GAT-GGCGATGACGACAGC-3′ and 5′-GCGTGACTGCGACCTTTC-3′ were used to amplify the *rpoD* gene. *E. coli* MC4100, RH90, and KY1429 bearing wild-type Pm::*lacZ* fusion in pMD1405 plus a compatible plasmid bearing the wild-type XylS (pERD103, TABLE ONE) were cultured as for β-galactosidase assays (see previous paragraph), and samples were taken in exponential and stationary phases. Pellets were resuspended in 1 ml of distilled water, cells were lysed by sonication, and the crude lysate was spun for 2 min at 12,000 × *g*. Supernatant aliquots of 100 μl were heated at 99 °C for 10 min to inactivate DNases. Four serial dilutions were prepared for each crude DNA sample. Real-time PCR was performed on an iCycler iQ detection system instrument according to the manufacturer's instructions. The PCR reaction (20 μl) was set up using the following reagents: 10 μl of 2× SYBR Green Supermix (Bio-Rad), 0.2 μl each of the forward and reverse primers for each gene (100 μM), 7.6 μl of water, and 2 μl of diluted DNA sample. The thermal cycling conditions used were: 3 min at 95 °C, followed by 35 cycles of 95 °C for 30 s, 60 °C for 20 s, and 72 °C for 30 s. A final melt curve was carried out to check the specific amplification of both genes. All reactions were run in triplicate. For each strain and condition, the two genes were analyzed.

*Western Blots*—Overnight cultures of *E. coli* strains were grown on LB medium containing appropriate antibiotics. Cultures were diluted 100-fold in the same medium, grown for 1 h at 30 °C and then supplemented or not with 1 mM 3MB. After 30 min of incubation at 30 °C, the cells (100 ml) were harvested by centrifugation, and the pellet was resuspended and washed once in 1× M9 buffer (48 mM sodium phosphate, 22 mM potassium phosphate, 19 mM ammonium chloride, and 8.5 mM sodium chloride, pH 7). The resulting paste was stored at −20 °C until use. The same procedure was followed in parallel cultures after 5 h of induction, except that 10-ml samples were harvested. The cell pellets were resuspended in 2 ml of lysis buffer containing Tris (50 mM) pH 7.5, 50 mM NaCl, 2 mM EDTA, 4 mM β-mercaptoethanol, and 1× Complete™ protease inhibitor mixture (Roche Applied Science). The cells were lysed by sonication, and the insoluble fraction was discarded by centrifugation at 18 000 × *g* for 20 min. Care was taken to process all samples in parallel to avoid differences in protein dissociation. Aliquot fractions were analyzed on SDS-PAGE (12.5%) and transferred to a nitrocellulose membrane. To determine the relative fraction of each σ factor bound to the RNA polymerase core, the same samples were run in native 7% (w/v) PAGE and transferred as described above. The membranes were blocked for 3 h at room temperature with 5% nonfat dry milk in phosphate-buffered saline. Blots were incubated at 4 °C overnight with monoclonal antibodies against each *E. coli* σ subunit (purchased from Neoclone). Antibodies were diluted 1:3000 in the case of $\sigma^{70}$, $\sigma^{38}$, and β antibodies and 1:1000 for $\sigma^{32}$ antibodies. Blots were washed with phosphate-buffered saline solution and incubated with goat anti-mouse immunoglobulin G + L conjugated with horseradish peroxidase (1:1000 dilution) for 1 h (Caltag Laboratories). The blots were developed with the SuperSignal® West Dura-extended duration substrate (Pierce). Chemiluminescent blots were exposed to autoradiographic film for 30 s to 2 min. Time course experiments were carried out as described above, except that samples for extract preparation were taken 10, 30 and 60 min and 5 h after 3MB was added.

## RESULTS

*Pm Transcription and Levels of the Different σ Factors along the Growth Curve*—σ factors bind to core RNA polymerase and allow the recognition of short nucleotide motifs located 10 and 35 bp upstream from the transcription start point of the promoters. Depending on the σ factor and the promoter, the so-called −10 box can be located at different positions between −5 and −14. In the Pm promoter, these sequences diverged considerably from consensus ones (Fig. 1). In this promoter, high expression levels were maintained throughout the growth curve by a switch from $E\sigma^{32}$ in the exponential phase to $E\sigma^{38}$ in the stationary phase. As a result, no significant decrease in activity was observed during growth, reflecting optimized coupling of the two RNA polymerases. The limits defining Pm transcription by each RNA polymerase can only be defined in σ mutant backgrounds. In our analysis, we used *E. coli* MC4100 as the wild-type strain and its isogenic derivatives RH90 (*rpoS* mutant) (30) and KY1429 (*rpoH* mutant) (31).

It is known that the promoter sequence is not the only factor that determines *in vivo* promoter selectivity by different RNA polymerases. Parameters, such as σ factor levels and availability or σ factor competition for the polymerase core, also play important roles (32). To pinpoint the conditions when the switch between σ factors occurred in the Pm promoter, it became necessary to determine their relative abundance in the wild-type and σ-deficient backgrounds under different growth conditions. To track the presence of factors $\sigma^{38}$, $\sigma^{32}$, and $\sigma^{70}$ in cells growing on LB medium in the presence or absence of 3MB, samples for Western blot analysis were taken in the early exponential or stationary phase as described under "Experimental Procedures." Similar levels of $\sigma^{70}$ were found in all three strains in both the exponential and stationary phases. These levels were independent of the presence or absence of 3MB (not shown).

In the exponential phase, $\sigma^{38}$ was only detectable in KY1429 (Fig. 2*A*). In the stationary phase, $\sigma^{38}$ was detectable in the parental strain MC4100 and in the *rpoH* mutant KY1429, whereas it was absent in the *rpoS* mutant RH90 (Fig. 2*B*). Interestingly, $\sigma^{38}$ stability was low in KY1429, as denoted by the multiple degradation bands that appeared in Western blots, particularly in the stationary phase. In this strain, in the presence of 3MB, $\sigma^{38}$ concentration greatly increased both in the expo-

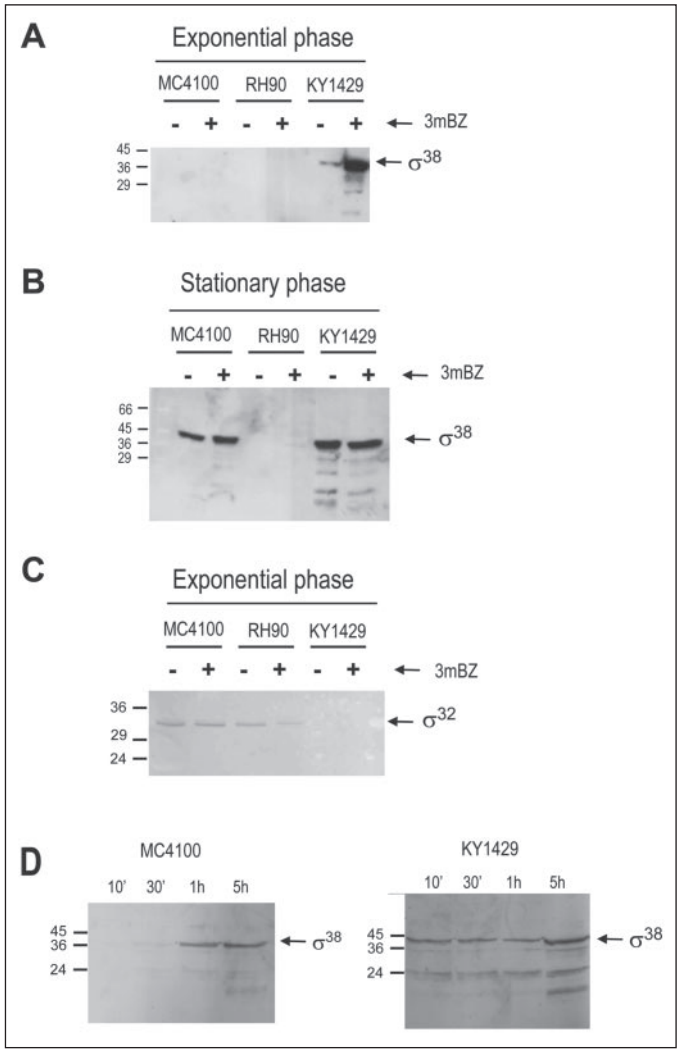



FIGURE 2. **Growth phase-dependent levels of $\sigma^{38}$ and $\sigma^{32}$ proteins.** Either 50 $\mu$g (for $\sigma^{38}$) or 100 $\mu$g (for $\sigma^{32}$) of total proteins from cell lysates of the indicated strains prepared in the exponential (*A* and *C*) or stationary (*B*) phase or at the indicated times after 3MB induction (*D*), were subjected to SDS-PAGE, and the blots were probed with anti-$\sigma^{38}$ or anti-$\sigma^{32}$. *A*, *B*, and *D*, $\sigma^{38}$ antiserum; *C*, $\sigma^{32}$ antiserum.

nential and stationary phases (Fig. 2, *A* and *B*). This could be due to impairment of the heat-shock response in the *rpoH* strain and, hence, to reduced induction of heat-shock proteins, such as ClpXP and DnaK (33, 34), which are responsible for the degradation and stability, respectively, of $\sigma^{38}$. To detect $\sigma^{32}$, doubled amounts of total protein were loaded into the gels. In the wild-type and *rpoS* mutant, $\sigma^{32}$ was detectable only in the exponential phase, regardless of the presence of 3MB, and was absent in the knock-out mutant KY1429 (Fig. 2*C*).

The time course of the appearance of $\sigma^{38}$ in response to the addition of 3MB was followed along a growth curve (Fig. 2*D*). In the wild-type strain, where $\sigma^{38}$ was absent in the exponential phase, the $\sigma$ factor appeared 1 h after the addition of 3MB. In KY1429, $\sigma^{38}$ was detected as promptly as 10 min after induction. Heat-shock or stress induction of $\sigma^{38}$ in the exponential phase has been reported, and unlike $\sigma^{32}$, this response is not transient (33).

In an attempt to estimate changes in the fraction of each $\sigma$ factor bound to the RNA polymerase core under our experimental conditions, samples from all three strains were taken in the exponential and stationary phases from cells grown in the presence and absence of 3MB, electrophoresed under nondenaturing (native) conditions, and blotted onto nitrocellulose membranes in quadruplicate. The membranes were then incubated with antibodies against the $\beta$, $\sigma^{70}$, $\sigma^{38}$, or $\sigma^{32}$ subunits of RNA polymerase (Fig. 3). Free subunits and RNA polymerase holoenzymes could be identified when membranes incubated with antibodies against $\beta$ and $\sigma^{70}$ subunits were compared. Fig. 3, *A* and *B*, shows a slow migrating band that could be detected with antibodies against $\sigma^{70}$ and $\beta$ and was thus identified as the holoenzyme. It is interesting to note that, in the exponential phase, a portion of the $\beta$ subunit was found as a free subunit, whereas in the stationary phase, most of it appeared bound to the core. In contrast, $\sigma^{70}$ factor was mainly found as part of the holoenzyme.

Incubation with antibodies against $\sigma^{38}$ revealed two bands, one corresponding to free $\sigma^{38}$ and a second one migrating slowly (as described for the holoenzyme band) and thus corresponding to core-bound $\sigma^{38}$. The data shown in Fig. 3*C* confirmed that $\sigma^{38}$ was most abundant in KY1429 as a free subunit and also as part of the holoenzyme. In this strain, the addition of 3MB increased the fraction of $\sigma^{38}$ bound to the core. In the stationary phase, this $\sigma$ factor was mostly found as a free subunit based on our analysis of Western blots. Our experimental conditions did not allow the detection of $\sigma^{32}$ in the native electrophoresis.

These results indicate a clear-cut difference between $\sigma^{32}$ and $\sigma^{38}$ in MC4100 in the exponential and stationary phases. In the *rpoH* mutant KY1429, $\sigma^{38}$ appears to become available earlier during growth, both free and as part of the holoenzyme. As a consequence, E$\sigma^{38}$ would be able to transcribe Pm earlier during the growth period than in the wild-type strain MC4100 (see "Expression from Pm in an $\sigma^{32}$-deficient Background").

*Scanning Mutagenesis of the RNA Polymerase Binding Region at Pm*—To identify the critical nucleotides able to confer specific recognition of Pm by either $\sigma^{32}$ or $\sigma^{38}$, we generated all possible single point mutations between positions −7 and −18. The region under analysis was extended to −18, because the Pm sequence resembles the $\sigma^{32}$ consensus sequence between positions −14 and −17 (CCCC) (35) (Fig. 1). The mutant promoters were fused to *lacZ* in plasmid pMD1405, and the level of expression in the early exponential phase of growth (culture turbidity at 660 nm between 0.1 and 0.3) and the stationary phase (culture turbidity at 660 nm of about 2–2.5) was determined in the wild-type strain MC4100 (*xylS* gene in pERD103) and in the presence of the effector molecule 3MB. In this strain, 3MB-dependent $\beta$-galactosidase levels for the wild-type Pm promoter were 1,300 MU in the early exponential phase and 15,000 MU in the stationary phase (Fig. 4). Fig. 4*A* shows $\beta$-galactosidase activity levels obtained with each single mutant promoter in the exponential phase. Pm nucleotide positions in the mutated RNA polymerase binding region could be grouped according to the expression pattern of the corresponding mutant: 1) positions −7, −11, and −12, where any change was critical regardless of the specific base that was introduced and completely abolished Pm activity; 2) positions −8 to −10, where most mutations had a severe effect upon activity; and 3) positions between −13 and −18, where mutations had little effect on Pm activity in this growth phase.

The situation was different when Pm-dependent $\beta$-galactosidase activity was determined in the stationary phase (Fig. 4*B*). We found low levels of transcription with the Pm mutant promoters that had point mutations at −7, −11, and −12, and a modest level of activity was detected only in the −12T → C Pm mutant. On the other hand, activity recorded with mutation −8C → A was almost double the level found in the wild type. One of the possible changes at positions −9 and −16 led to a 65% decrease in activity, whereas all mutations at positions −17 and −18 produced decreases in activity between 25 and 50%. The remaining mutants behaved like the wild type, showing a higher level of expression

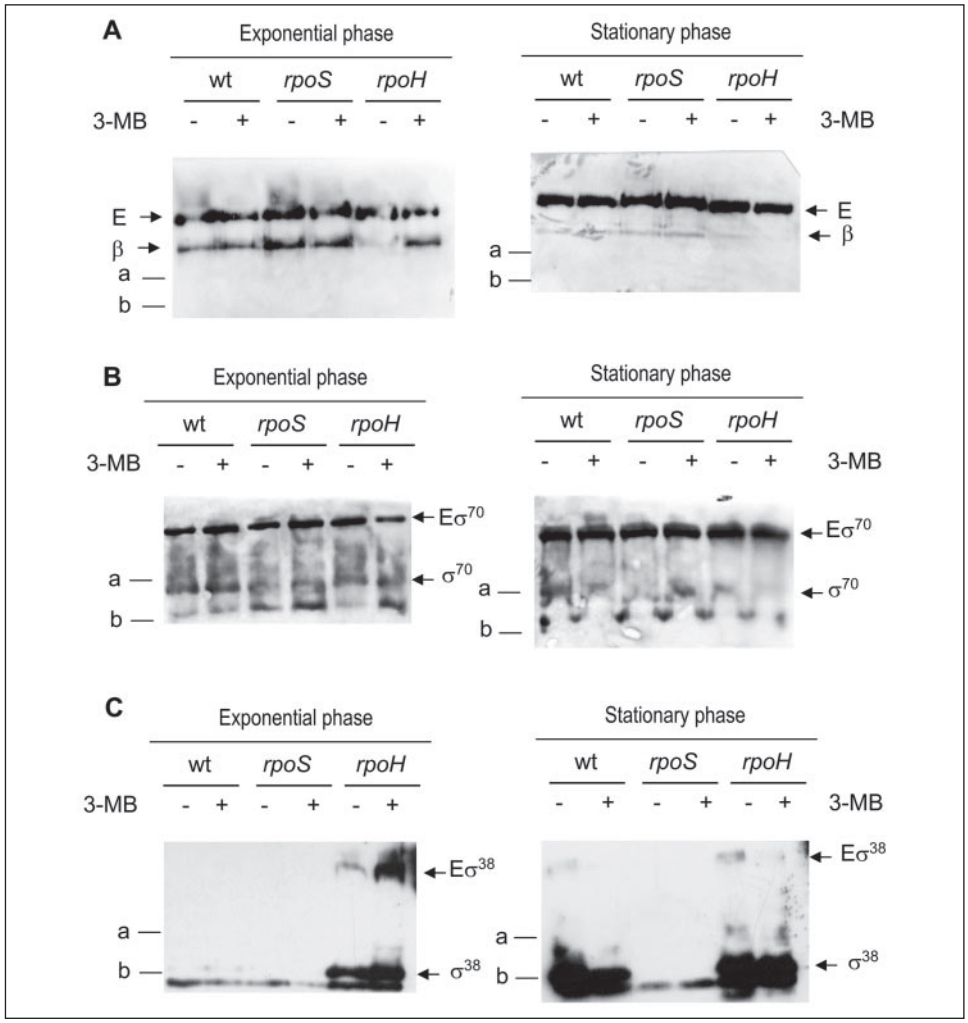



FIGURE 3. **Growth phase-dependent levels of $\sigma^{38}$ and $\sigma^{32}$ bound to core RNA polymerase.** 75 $\mu$g of total protein from cell lysates of the indicated strains prepared in the exponential or stationary phase were subjected to nondenaturing PAGE, and the blots were probed with anti-$\beta$ (*A*), anti-$\sigma^{70}$ (*B*), or anti-$\sigma^{38}$ (*C*) antiserum. *a* and *b* indicate migration of the standard proteins bovine serum albumin and ovalbumin, respectively. *wt*, wild type. *E*, RNA polymerase holoenzyme.

than in the early exponential phase. It is worth noting that the $\beta$-galactosidase levels obtained in the absence of the effector 3MB were negligible.

These results were interpreted as evidence that the RNA polymerases acting in the exponential and stationary phases recognize the same nucleotides at $-7$, $-11$, and $-12$ but recognize additional positions differently. In particular, the low level of activity with certain mutants at $-8$ and $-9$ in the early exponential phase, when $\sigma^{32}$ was used to transcribe Pm, indicates that these two nucleotides are critical for the $\sigma^{32}$ factor. The $-8C \rightarrow A$ change rendered a mutant promoter that was fully active in the exponential phase and exhibited a 2-fold higher expression level than the wild-type promoter in the stationary phase (Fig. 4). This increased transcription from this mutant promoter in the stationary phase was initially attributed to E$\sigma^{70}$, because the mutation improved the $-10$ region to those recognized by E$\sigma^{70}$. To confirm this possibility, all Pm mutants were analyzed in the *rpoS-rpoH* double mutant SM25 (15), where *xylS* was provided in the gentamicin-resistant vector pDCXylS. All Pm mutants showed activity levels below wild-type Pm in the absence of both $\sigma$ subunits (not shown), thus suggesting that $\sigma^{70}$ was not responsible for these transcription levels. We therefore concluded that the sequence in the $-8C \rightarrow A$ mutant was significantly improved for $\sigma^{32}$ and especially for $\sigma^{38}$. The results shown in Fig. 4 suggest that, although wild-type G at position $-9$ is important for recognition by $\sigma^{32}$, a change to A enhances $\sigma^{38}$ recognition of the promoter (Fig. 4*B*). Overall, changes in the promoter sequence toward optimization of recognition by $\sigma^{38}$ seem to be unfavorable for E$\sigma^{32}$. The actual wild-type sequence may therefore be a compromise between the ability to respond to $\sigma^{38}$ in the stationary phase and the requirement to respond to $\sigma^{32}$ in the early exponential phase.

*Sequence Determinants of Pm Expression in the Exponential Phase—E. coli rpoS* strain RH90 was transformed with plasmid pERD103 (*xylS*) and pJLR107 (*Pm:lacZ*) or its derivatives containing mutant promoters, and we determined $\beta$-galactosidase activity in the early exponential growth phase as indicated under "Experimental Procedures." We had previously shown that activity from the wild-type Pm promoter in the exponential phase was almost unaffected in a mutant lacking $\sigma^{38}$ (14), because under these conditions, Pm activity relies mainly on $\sigma^{32}$, which expression is maintained (Fig. 2*C*). Wild-type Pm activity in RH90 in the exponential phase was 1500 MU (115% in Fig. 5*A*). When we examined activity from the mutant promoters in the early exponential phase when only $\sigma^{32}$ was available for transcription from Pm, very low activity was found for all of the mutant promoters with mutations at positions $-7$, $-9$, $-11$, $-12$, $-15$, and $-17$. When mutations were introduced at position $-8$, $-10$, $-13$, $-14$, $-16$, and $-18$, $\beta$-galactosidase activity varied depending on the specific mutation that was introduced. These results indicate that not only the $-10$ hexamer but also the bases upstream from the $-10$ motif (namely Cs at positions $-14$ to $-17$) are crucial for $\sigma^{32}$ recognition. The presence of an A at position $-8$ allowed the levels of expression to approach those in the wild type, thus confirming the results obtained with strain MC4100. In contrast, G at posi-

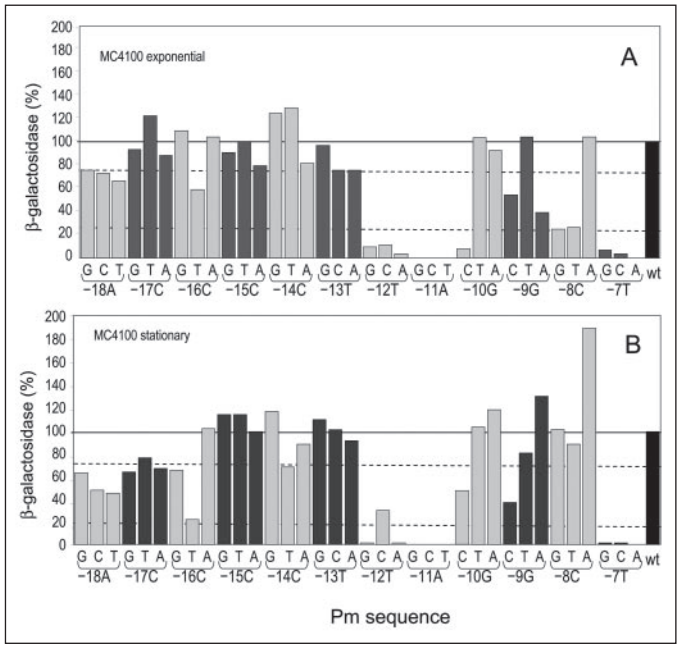



FIGURE 4. **β-Galactosidase expression level from Pm promoter single point mutants in a $\sigma^{32}$- and $\sigma^{38}$-proficient background.** *E. coli* MC4100 bearing the wild-type or the indicated single base mutant Pm:*lacZ* fusion in pJLR107 and *xylS* in pERD103 was grown on LB medium with 3MB for 1 h (early exponential phase) (*A*) or for 6 h (early stationary phase) (*B*). The base present at each indicated position in the wild-type promoter is shown along with the base present at the corresponding position in each mutant promoter assayed. Data are the average of at least six independent determinations, with S.D.<20% of the given values. Wild-type activity (*wt*; *black bar*) was 1,300 units in the early log phase and 20,000 units in the stationary phase. The *dotted lines* indicate 25 and 75% β-galactosidase activity. For clarity, activity of mutants in adjacent positions is depicted in different gray tones.

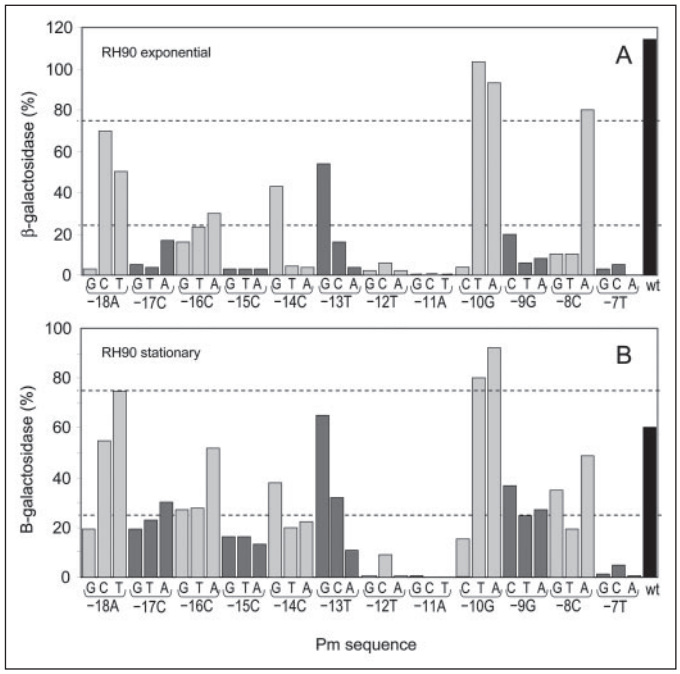



FIGURE 5. **β-galactosidase expression level from single point mutations at the Pm promoter in a $\sigma^{38}$-deficient, $\sigma^{32}$-proficient background.** Conditions are as described in the legend for Fig. 4, except that the strain used was RH90. Wild-type (*wt*; *black bar*) Pm activity was 1,500 MU in the early log phase (115% of Pm activity in MC4100) and 7,000 MU in the stationary phase (60% of the corresponding activity in the wild-type MC4100). The vector pMD1405 copy number, as determined by real-time PCR (see "Experimental Procedures"), both in exponential and stationary phases was the same in MC4100 and RH90.

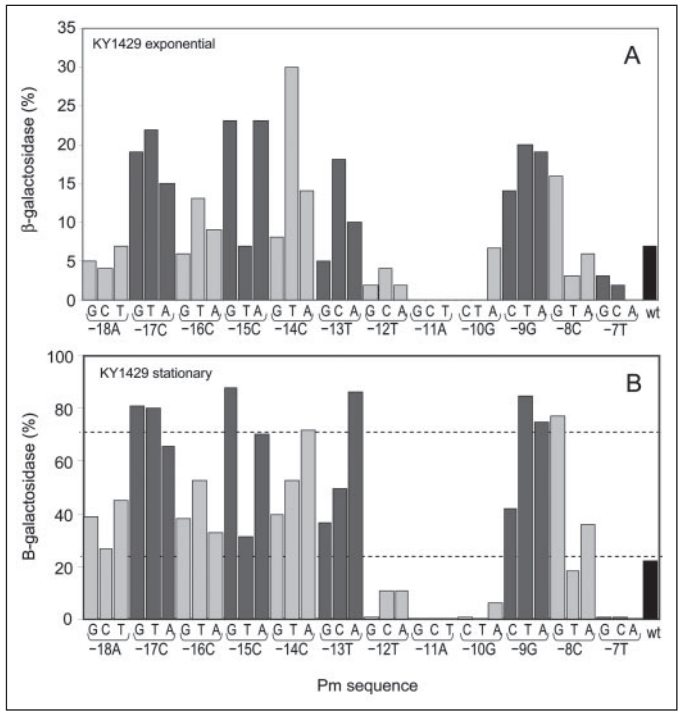



FIGURE 6. **β-Galactosidase expression level from single point mutations at the Pm promoter in a $\sigma^{32}$-deficient, $\sigma^{38}$-proficient background.** Conditions were as described in the legend for Fig. 2, except that the strain used was KY1490. Wild-type (*wt*) Pm activity was 75 MU in the early log phase (8% of the activity in MC4100) and 2700 MU in the stationary phase (~20% of the activity in MC4100). Vector pMD1405 copy number, as determined by real-time PCR (see "Experimental Procedures"), both in exponential and stationary phases was the same in MC4100 and KY1429.

tion −13 or −14 produced β-galactosidase levels of up to 50% of those in the wild-type Pm. Interestingly, two mutations at position −10 led to wild-type activity levels, indicating that the changes were irrelevant for $\sigma^{32}$ recognition in Pm. However, a change to C prevented recognition by $\sigma^{32}$.

From the results in Fig. 5*A*, we were able to define a recognition sequence for $\sigma^{32}$ at Pm as 5′-(H)CCC(S)(K)TA(D)G(M)T-3′ (positions −18 to −7), where T at positions −7 and −12, G at position −9, A at position −11, and C at positions −14 to −17 are pivotal for E$\sigma^{32}$ transcription. Although similar to the functional $\sigma^{32}$ binding sequence deduced from the *E. coli groE* promoter (36), this sequence shows a number of differences, particularly and most importantly the identification of position −7 as relevant for recognition by $\sigma^{32}$.

Wild-type Pm activity in the *rpoS* mutant background was 60% of its value in MC4100 (7,000 MU) in the stationary phase. In general, in this mutant background, most of the mutants tested exhibited low activity levels. As mentioned above, expression in the −8C → A mutant remained at the wild-type level (Fig. 5), and mutations −10G → T or −10G → A led to promoters with higher levels of expression than wild type.

*Expression from Pm in an $\sigma^{32}$-deficient Background*—To identify the promoter positions that were important for E$\sigma^{38}$ recognition, we decided to follow Pm expression in a mutant lacking the E$\sigma^{32}$ polymerase involved in expression in the exponential phase. In this $\sigma^{32}$-deficient background, most of the transcriptional activity measured at Pm is presumably due to E$\sigma^{38}$, which is already present in stationary phase (Fig. 2). Pm expression in the MC4100 *rpoH* derivative KY1429 was very low for the wild-type promoter in the exponential phase (75 MU) and increased in the stationary phase (2700 MU) (Fig. 2) to levels below those seen in wild-type MC4100. When we measured mutant Pm

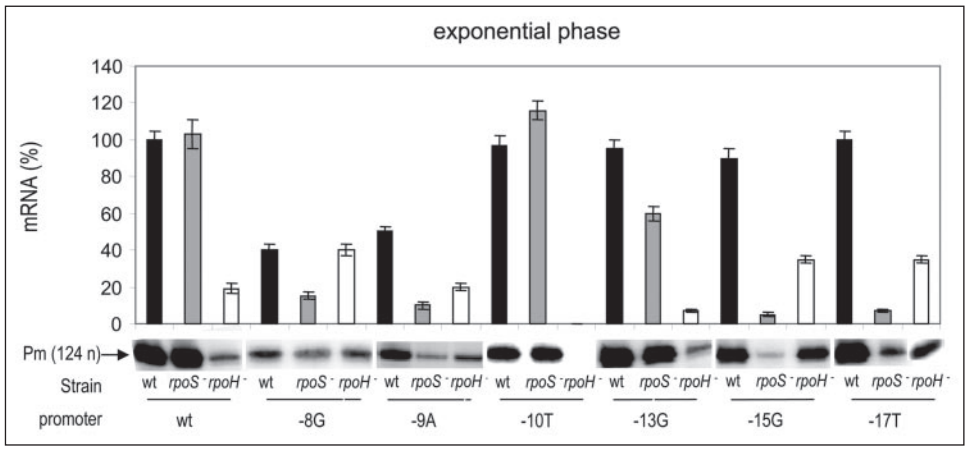



FIGURE 7. **Primer extension analysis of wild-type and mutant Pm promoters.** MC4100, RH90, or KY1429 strains bearing pERD103 (*xylS*) and pMD1405::Pm* were grown until the exponential phase and total RNA was isolated. Equal amounts of total RNA (10 $\mu$g) were subjected to primer extension analysis, as described under "Experimental Procedures," and the extension products were analyzed in urea-PAGE. The phosphorimaging scan of duplicate experiments was quantified as described under "Experimental Procedures." A typical phosphorimaging scan is shown.

expression in this strain, we observed that mutations at positions −7, −11, and −12 in Pm resulted in similar basal levels of expression to those observed in the wild-type MC4100 and RH90 strains (Fig. 6). In addition, the mutations at position −10 in Pm, which behaved like the wild type in RH90, were expressed at a lower level than those in the wild-type strain, particularly in the stationary phase. This was also the case for mutant −10G → C. These results suggest that position −10 is important for $\sigma^{38}$ recognition. However, when mutated, most of the remaining positions rendered a promoter with increased expression in the exponential phase. This was the case for all mutations at position −9, the mutation at −8G, and at least one of the mutations at positions −14 to −17 (Fig. 6*A*). These differences were maintained in the stationary phase, when all mutations except those at positions −7, −11, and −12 led to levels of expression equal to or higher than those induced by the wild-type Pm. This general increase in activity with respect to the wild type both in the exponential phase, when $\sigma^{38}$ was already present in this strain, and in the stationary phase, when the highest levels of $\sigma^{38}$ were seen in KY1429 (Fig. 2), indicates that the wild-type Pm promoter was far from optimal for recognition by E$\sigma^{38}$ and that the absence of $\sigma^{32}$ competition for both the RNA polymerase core and DNA binding would allow transcription by E$\sigma^{38}$ to increase. The previously mentioned lack of transcription in a double mutant *rpoS-rpoH* rules that these levels were because of residual $\sigma^{70}$ activity.

*Mutant Pm Promoters Maintain the Same Transcription Initiation Site*—Changing the −10 region sequence of a promoter could generate new artificial promoters showing different transcription initiation sites that the $\beta$-galactosidase assays we used were not able to detect. To verify that the results shown in Figs. 4–6 were actually derived from Pm activity and not from any accidental promoter created as a consequence of the mutagenesis, the transcription start site of the most relevant mutant promoters in our analyses was determined in the wild-type and $\sigma$ mutant strains. Fig. 7 shows that, in all mutants tested, the transcription start site was the canonical Pm start site, and activity levels correlated well with $\beta$-galactosidase activity. Transcription always required activation by 3MB-activated XylS, which probably guarantees the proper positioning of RNA polymerase for transcription activation.

## DISCUSSION

The expression of the *meta*-cleavage pathway operon both in the exponential and stationary phases is driven by the Pm region. This region exhibits all the necessary features for the binding of the XylS regulator and for its recognition by two different RNA polymerases in a 70-bp region. The XylS binding sites cover the −70 to −34 region and overlap −35 by 1 bp. As a probable consequence of this overlap, the −35 region shows poor homology with any canonical consensus sequence recognized by different RNA polymerases. Therefore, the −10 region must accommodate overlying sequences for the binding of the two polymerases, E$\sigma^{32}$ and E$\sigma^{38}$, involved in Pm expression.

A number of $\sigma^{38}$-dependent promoters have been reported to be also transcribed *in vivo* by a second RNA polymerase depending on physiological conditions. However, in all cases described so far, E$\sigma^{70}$ is responsible for this alternative transcription. In the *csgBA* promoter, H-NS influences the use of either E$\sigma^{38}$ or E$\sigma^{70}$ to transcribe this promoter (37), whereas in the *dps* promoter E$\sigma^{70}$ directs transcription in the exponential phase and requires OxyR. In the stationary phase, however, transcription from the *dps* promoter is $\sigma^{38}$-dependent and OxyR-independent (38). In contrast with these promoters, which use alternative regulators depending on the growth phase, Pm always requires XylS activated by 3MB to mediate transcription with either E$\sigma^{32}$ or E$\sigma^{38}$ regardless of the growth phase.

Our results with the Pm promoter in the wild-type and $\sigma$ mutant strains show that E$\sigma^{32}$ is the major polymerase involved in transcription of Pm. In the *rpoH* mutant KY1429, expression levels were only 7% of those in the wild type. Furthermore, in KY1429, the levels of expression from Pm in the stationary phase were relatively low, which suggests that the E$\sigma^{38}$ polymerase transcribed poorly the Pm promoter (Fig. 6). In fact, expression was low despite the presence of E$\sigma^{38}$ in the early phase of growth (Fig. 2).

*Defining the Downstream Sequences Recognized by $\sigma^{32}$ and $\sigma^{38}$*—To define the downstream sequences recognized by $\sigma^{32}$ and $\sigma^{38}$, we used a series of Pm mutant promoters and tested expression in different genetic backgrounds deficient in either of the alternative $\sigma$ factors. The rationale behind this approach was the following. If Pm is transcribed with E$\sigma^{32}$ in the exponential phase and E$\sigma^{38}$ in the stationary phase, one would expect the nucleotides at Pm that are important for transcription by E$\sigma^{32}$ to be detected in the exponential phase as inactive mutants in MC4100 (wild-type background). In contrast, nucleotides involved in Pm recognition by E$\sigma^{38}$ would be inactive in the stationary phase. Along these same lines, mutations at the Pm promoter involved in $\sigma^{32}$ recognition should also be inactive in RH90 ($\sigma^{32}$-proficient, $\sigma^{38}$-deficient) in the exponential phase when E$\sigma^{32}$ plays an important role in Pm transcription. Moreover, mutations in Pm at positions relevant for E$\sigma^{38}$ recognition should be affected in KY1429 ($\sigma^{38}$-proficient, $\sigma^{32}$-deficient) in the stationary phase. In addition to the above considerations, in our analysis, we took into account the levels of $\sigma^{38}$ and $\sigma^{32}$ in the different genetic backgrounds used in this study that would influence the availability of $\sigma^{32}$ and $\sigma^{38}$ for binding to the core RNA polymerase (Fig. 3). No less important is the fact that mutations in the −10 element could

generate recognition sites for $\sigma^{70}$. These issues were taken into consideration in our analysis of each mutant.

The precise definition of the −10 hexamer in promoters transcribed by RNA polymerase with the $\sigma^{70}$ family of alternative $\sigma$ factors is elusive, because the distance to the transcription start site can vary from 4 to 8 bp and also because this region is usually rich in adenines and thymines; so different hexamers can be proposed (18–21, 39, 40). In the case of Pm, because of the profusion of adenines and thymines between positions −5 and −13 (Fig. 1), several putative −10 hexamers can be hypothesized. In our analysis, any mutation at positions −7T, −11A, and −12T dramatically decreased Pm activity in all the genetic backgrounds we tested. This allowed us to precisely define the −10 recognition element of Pm as TAGGCT (positions −12 to −7). The presence of T in the first and last position of the hexamer is in agreement with observations reported for other promoters transcribed by $\sigma^{38}$ (35, 41, 42).

*Mutations within the −10 Element*—The consensus sequence shown for $\sigma^{32}$ in Fig. 1 was first defined some twenty years ago. Recently, scanning mutagenesis of the *E. coli groE* promoter confirmed the earlier consensus sequence of the $\sigma^{32}$ −10 element (36) (Fig. 1). The alignment of the Pm promoter sequence with the $\sigma^{32}$ consensus sequence supports the recognition of Pm by E$\sigma^{32}$. Two mutations at positions −8 (−8C→G and −8C→T) and one at −9 (−9G→A) showed the expected phenotype for a relevant $\sigma^{32}$ binding position, low activity in the exponential phase and unaffected in the stationary phase in the wild-type MC4100. These Pm mutants were also inactive in RH90 in the exponential phase. However, the behavior of two of these mutants in the $\sigma^{32}$-deficient KY1429 background merited further investigation. The −8C→G and −9G→A mutants showed high activity levels in this strain, suggesting that E$\sigma^{38}$ was able to transcribe from this promoter in the absence of $\sigma^{32}$. Published reports of the changing levels of $\sigma^{38}$ throughout the growth phase or at different growth temperatures are discrepant and seem to vary from strain to strain (43, 44). Our results in Fig. 2*A* showed $\sigma^{38}$ to be present at significant levels in the early growth stages both in MC4100 and KY1429, although in the latter, $\sigma^{38}$ was always present and its levels were higher and detectable even in the absence of 3MB (Figs. 2 and 3). The $\sigma^{38}$-dependent transcription of Pm mutants −8C→G and −9G→A was not observed in MC4100, presumably because the presence of $\sigma^{32}$ decreased RNA polymerase core availability for $\sigma^{38}$.

Specific features common to $\sigma^{38}$-dependent promoters have been established from compilations of an increasing number of these promoters (18, 21) from mutagenic analyses both *in vivo* and *in vitro* and from the analysis of artificial promoters containing putative $\sigma^{38}$-dependent sequences (45, 46). These promoters usually lack a defined −35 element, exhibit four highly conserved nucleotides in the −10 box (21) (Fig. 1, *underlined*), and generally have a TC at positions −14/−13 (47). The Pm promoter shares five of eight positions with the recently updated $\sigma^{38}$ consensus sequence (21) (Fig. 1). However, the promoter lacks the main features described by Lacour *et al.* (20) for promoters that are exclusively recognized by E$\sigma^{38}$, *i.e.* a C in the first position in the −10 hexamer.

In the stationary phase, the set of mutations in −8 and −9 allowed activity levels that approached the wild-type levels in MC4100, and that surpassed wild-type levels in KY1429, suggesting that almost any change in Pm would favor E$\sigma^{38}$ recognition. However this is not the case with mutations −9G→C or −8C→T, which do not improve E$\sigma^{38}$ recognition. In RH90, the levels were not maximal, as a consequence of the negligible activity in the exponential phase, which curtailed $\beta$-galactosidase accumulation. Mutations in the −10 position produced a similar expression pattern in MC4100 but exhibited negligible activity in the KY1429 background, although two of the mutants yielded promoters that were fully active in RH90. These findings suggest that this position was not very important for $\sigma^{32}$ but was necessary for $\sigma^{38}$ recognition. In fact, this was the only position where mutants showed a $\sigma^{38}$-dependent phenotype. This is striking, because guanosine was the least common nucleotide in the −10 hexamer in a recent compilation of $\sigma^{38}$-dependent promoters, although the substitution of a canonical consensus nucleotide for a G at this position in an artificial $\sigma^{38}$-dependent promoter had no effect on $\sigma^{38}$ DNA binding (21).

We showed previously that, in a double *rpoS*,*rpoH* mutant, transcription from Pm was almost negligible, which was taken as an indication that, in the wild-type Pm promoter, E$\sigma^{70}$ played a minor role (15). Mutations at −7 and −11 yielded a nonfunctional Pm promoter in all of the backgrounds tested in this study. This was also true for position −12, except that when the −12T position in Pm was mutated to a C, the resulting promoter retained 25% of the activity in the stationary phase in all backgrounds (Figs. 4*B* and 5*B*). Position −12 is known to be crucial for RNA polymerase recognition. The −12T→C mutation gives rise to a mutant Pm promoter with a low activity, which however, was still recognized to a certain extent by both polymerases. We also found that the −8C→A change resulted in a high level of Pm expression in the stationary phase in the wild-type MC4100 (Fig. 3*B*) and in the $\sigma^{38}$-deficient mutant. As for mutant −12T→C, this mutant is transcribed by both polymerases but with a high efficiency, especially by E$\sigma^{38}$. In this case, the Pm sequence was improved for both holoenzymes.

Taken together, these results show that, in addition to the pivotal positions −7, −11, and −12, which are crucial for the transcription of Pm with any RNA polymerase, positions −8 and −9 are critical for $\sigma^{32}$ recognition but are much less important for $\sigma^{38}$ recognition, whereas the −10 position is key for $\sigma^{38}$ recognition.

*Mutations Upstream from the −10 Element*—When we applied the same rationale as above to mutants at positions −12 to −18, we observed that no mutation had a significant effect on activity in the exponential phase in MC4100. However, in RH90, which lacks $\sigma^{38}$, all Pm mutant promoters bearing mutations between positions −13 and −17 showed between 10 and 50% activity of the wild-type Pm in the same background. Moreover, the Cs present at positions −14 to −17 could be aligned with the $\sigma^{32}$ consensus (Fig. 1). The high activity levels of these mutants in MC4100 suggest that, in this wild-type strain, E$\sigma^{38}$ takes over transcription from these promoters. This hypothesis is further supported by the considerable levels of $\beta$-galactosidase obtained in KY1429 in the exponential phase.

These results suggest that mutations between −13 and −17 resulted in mutant promoters better recognized by E$\sigma^{38}$ in both the exponential phase and stationary phase (*e.g.* all changes at positions −17, −15A, −15G, −14T, −13C). However, this was only observed in the absence of $\sigma^{32}$ (strain KY1429) and was confirmed by the lower activities obtained from these mutant promoters in RH90 in the stationary phase. In MC4100, the activity of these promoters is similar to wild-type Pm, probably because $\sigma^{38}$ must compete with $\sigma^{32}$ for binding to the RNA polymerase core.

We showed that Pm is transcribed by two RNA polymerases, E$\sigma^{38}$ and E$\sigma^{32}$, from the same transcriptional start point in a process that is strictly dependent on XylS activated by 3MB. Our results suggest that deviations from the $\sigma^{38}$ consensus are required to adapt Pm to the demands of its function, *i.e.* recognition by E$\sigma^{32}$ as a rapid response to the presence of 3MB. This response is strong but transient, so the promoter must be optimized for recognition by this polymerase. In fact, the Pm sequence is closer to the $\sigma^{32}$ consensus, so changes in this sequence

are not tolerated. As a result, Pm is adjusted to ensure almost optimal recognition by this less abundant polymerase. After this initial reaction of the cells to the presence of the aromatic effector, activity with E$\sigma^{38}$ ensues and is subsequently maintained. This polymerase does not rely so strictly on the promoter sequence and is able to promote transcription from the modestly conserved $\sigma^{38}$ binding sequence present in Pm.

The Pm promoter always requires the presence of activated XylS to be transcribed. It has been shown that the nature of the regulator that activates Pm strongly influences the $\sigma$ factor used in transcription. Wild-type XylS seems to be designed to prevent E$\sigma^{70}$ accessibility to Pm (16). This implies an additional benefit for Pm transcription control, namely XylS would restrict transcription to situations in which the heat shock response and, hence, E$\sigma^{32}$ abundance are triggered by the presence of aromatic pathway substrates.

*Acknowledgments—We thank Carmen Lorente for valuable assistance and K. Shashok for help with the use of English in the manuscript.*

## REFERENCES


1. Franklin, F. C. H., Bagdasarian, M., Bagdasarian, M. M., and Timmis, K. N. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 7458–7462
2. Inouye, S., Nakazawa, A., and Nakazawa, T. (1984) *Gene* **29,** 323–330
3. Egan, S. M. (2002) *J. Bacteriol.* **184,** 5529–5532
4. Ramos, J. L., Stolz, A., Reineke, W., and Timmis, K. N. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 8467–8471
5. Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J. L. (1997) *Microbiol. Mol. Biol. Rev.* **61,** 393–410
6. Tobes, R., and Ramos, J. L. (2002) *Nucleic Acids Res.* **30,** 318–321
7. Gallegos, M. T., Marqués, S., and Ramos, J. L. (1996) *J. Bacteriol.* **178,** 6427–6434
8. González-Pérez, M. M., Ramos, J. L., Gallegos, M. T., and Marqués, S. (1999) *J. Biol. Chem.* **274,** 2286–2290
9. Kessler, B., de Lorenzo, V., and Timmis, K. (1993) *J. Mol. Biol.* **230,** 699–703
10. Kaldalu, N., Mandel, T., and Ustav, M. (1996) *Mol. Microbiol.* **20,** 569–579
11. González-Pérez, M. M., Marqués, S., Domínguez-Cuevas, P., and Ramos, J. L. (2002) *FEBS Lett.* **519,** 117–122
12. Busby, S., and Ebright, R. H. (1997) *Mol. Microbiol.* **23,** 853–859
13. Ruíz, R., Ramos, J. L., and Egan, S. M. (2001) *FEBS Lett.* **491,** 207–211
14. Marqués, S., Gallegos, M. T., and Ramos, J. L. (1995) *Mol. Microbiol.* **18,** 851–857
15. Marqués, S., Manzanera, M., González-Pérez, M. M., Gallegos, M. T., and Ramos, J. L. (1999) *Mol. Microbiol.* **31,** 1105–1113
16. Ruíz, R., and Ramos, J. L. (2002) *J. Biol. Chem.* **277,** 7282–7286
17. Marqués, S., Manzanera, M., González-Pérez, M. M., Ruiz, R., and Ramos, J. L. (1999) *Environ. Microbiol.* **1,** 103–104
18. Espinosa-Urgel, M., Chamizo, C., and Tormo, A. (1996) *Mol. Microbiol.* **21,** 657–659
19. Gaal, T., Ross, W., Estrem, S. T., Nguyen, L. H., Burgess, R. R., and Gourse, R. L. (2001) *Mol. Microbiol.* **42,** 939–954
20. Lacour, S., Kolb, A., and Landini, P. (2003) *J. Biol. Chem.* **278,** 37160–37168
21. Lee, S. J., and Gralla, J. D. (2001) *J. Biol. Chem.* **276,** 30064–30071
22. Gross, C. (1996) in *Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K.B., Magasanic, B., Reznikoff, W. S., Riley, M., Schaecter, M., and Umbarger, H. E., eds) pp. 1382–1399, ASM Press, Washington, D. C.
23. Ramos, J. L., Mermod, N., and Timmis, K. N. (1987) *Mol. Microbiol.* **1,** 293–300
24. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (2005) *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc., New York
25. Gallegos, M. T., Marqués, S., and Ramos, J. L. (1996) *J. Bacteriol.* **178,** 2356–2361
26. Higushi, R. (1990) in *PCR Protocols. A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds) pp. 177–183, Academic Press, San Diego, CA
27. Manzanera, M., Aranda-Olmedo, I., Ramos, J. L., and Marques, S. (2001) *Microbiology* **147,** 1323–1330
28. Miller, J. (1972) *Experiments in Molecular Genetics*, pp. 352–355, Cold Spring Harbor, New York
29. Tao, L., Jackson, R. E., and Cheng, Q. (2005) *Metab. Eng.* **7,** 10–17
30. Lange, R., and Hengge-Aronis, R. (1991) *Mol. Microbiol.* **5,** 49–59
31. Zhou, Y. N., Kusukawa, N., Erickson, J. W., Gross, C. A., and Yura, T. (1988) *J. Bacteriol.* **170,** 3640–3649
32. Ishihama, A. (2000) *Annu. Rev. Microbiol.* **54,** 499–518
33. Muffler, A., Barth, M., Marschall, C., and Hengge-Aronis, R. (1997) *J. Bacteriol.* **179,** 445–452
34. Zhou, Y., and Gottesman, S. (1998) *J. Bacteriol.* **180,** 1154–1158
35. Gross, C. A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., and Young, B. (1998) *Cold Spring Harb. Symp. Quant. Biol.* **63,** 141–155, Cold Spring Harbor Laboratory Press
36. Wang, Y., and de Haseth, P. L. (2003) *J. Bacteriol.* **185,** 5800–5806
37. Arnqvist, A., Olsén, A., and Normark, S. (1994) *Mol. Microbiol.* **13,** 1021–1032
38. Altuvia, S., Almirón, M., Huisman, G., Kolter, R., and Storz, G. (1994) *Mol. Microbiol.* **13,** 265–272
39. Cowing, D. W., Bardwell, J. C., Craig, E. A., Woolford, C., Hendrix, R. W., and Gross, C. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 2679–2683
40. Domínguez-Cuevas, P., and Marqués, S. (2004) in *The Pseudomonads, Virulence and Gene Regulation* (Ramos, J. L., ed) Vol. II, pp. 319–343, Kluwer Academic/Plenum Publishers, New York
41. Dombroski, A. J. (1997) *J. Biol. Chem.* **272,** 3487–3494
42. Bordes, P., Repoila, F., Kolb, A., and Gutierrez, C. (2000) *Mol. Microbiol.* **35,** 845–853
43. Jishage, M., and Ishihama, A. (1997) *J. Bacteriol.* **179,** 959–963
44. Maeda, H., Fujita, N., and Ishihama, A. (2000) *Nucleic Acids Res.* **28,** 3497–3503
45. Hengge-Aronis, R. (2002) *Curr. Opin. Microbiol.* **5,** 591–595
46. Typas, A., and Hengge, R. (2005) *Mol. Microbiol.* **55,** 250–260
47. Lisser, S., and Margalit, H. (1993) *Nucleic Acids Res.* **21,** 1507–1516
48. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580

CAPÍTULO 2

# Doble papel del efector en la dimerización y unión de XylS al ADN. Diseccionando los pasos del mecanismo de activación.

**Patricia Domínguez Cuevas, Patricia Marín, Stephen Busby, Juan L. Ramos, y Silvia Marqués**

La proteína XylS, que pertenece a la familia de reguladores transcripcionales AraC, activa la transcripción desde el promotor Pm, controlando la degradación de benzoatos. XylS consta de dos dominios funcionales conectados por una secuencia flexible de aminoácidos. El dominio N-terminal de XylS se encarga del reconocimiento del efector y de la dimerización de la proteína. El dominio C-terminal posee los determinantes necesarios para la unión al ADN. Debido a la insolubilidad inherente a XylS, los estudios sobre la activación del promotor Pm se han limitado a aproximaciones genéticas. En este trabajo hemos desarrollado un procedimiento de obtención de preparaciones semi-purificadas de la proteína XylS para su empleo en ensayos de retardo en gel, que nos han permitido desvelar los pasos clave en el mecanismo de activación. Nuestros resultados muestran que la unión de XylS a sus secuencias operadoras es independiente de la presencia de 3MB, y se torna estrictamente dependiente del efector en un mutante incapaz de dimerizar. El extremo N-terminal de XylS establece contactos directos con el extremo C-terminal, de forma independiente de la conexión entre ambos dominios, provocando la represión de la capacidad del dominio C-terminal para unirse al ADN. Las interacciones entre ambos dominios se ven impedidas en presencia de 3MB. En conjunto nuestros resultados nos llevan a proponer dos funciones para el 3MB en la activación de XylS, en primer lugar favoreciendo el proceso de dimerización y en segundo, alterando la conformación de XylS de manera que se elimina la represión ejercida por el dominio N-terminal sobre la unión al ADN del dominio C-terminal. Ensayos de inmunoprecipitación de cromatina mostraron que la proteína XylS es capaz de reclutar a la ARNP al promotor Pm. Los ensayos de transcripción *in vitro* y de formación de complejo promotor abierto empleando ARN polimerasas reconstituidas con los factores $\sigma^{32}$ y $\sigma^{38}$ de *P. putida* confirmaron nuestros resultados previos acerca de la dependencia de la transcripción desde Pm de estas dos ARN polimerasas alternativas (Marqués, *et al*., 1999).

Enviado a *Molecular Microbiology.*

Molecular Microbiology (2007) *Submitted*

# Dual role of effector in XylS dimerization and DNA binding. Dissecting the steps in the mechanism of activation

**Patricia Domínguez-Cuevas[1], Patricia Marín[1], Stephen Busby[2], Juan L. Ramos[1] and Silvia Marqués[1*]**

*[1]Consejo Superior de Investigaciones Científicas, Estación Experimental del Zaidín, Department of Environmental Protection, Granada, Spain.*

*[2]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.*

**Summary**

**XylS, the Pm promoter activator protein that controls the degradation of benzoates, belongs to the AraC family of transcription regulators. Owing to the persistent insolubility of the protein, XylS activation has generally been studied *in vivo* with genetic approaches. We developed a procedure to obtain purified preparations of XylS which can be used to unravel key steps in its mechanism of activation. Our gel shift assays showed specific 3–methylbenzoate-independent XylS binding to its target operator, which became strictly 3MB-dependent in a dimerization-defective mutant. We demonstrated direct, linker-independent contacts between the N- and C-terminal domains of the protein, which repressed the capacity of the C–terminal domain to bind DNA. Interactions between the two domains were hampered in the presence of 3MB. These and other findings led us to propose two roles for 3MB in XylS activation: favoring protein dimerization, and modifying XylS conformation to trigger the terminal domain. Chromatin immunoprecipitation results indicated that the mechanism of XylS activation involved RNA polymerase recruitment to the Pm promoter. All $\sigma^{32}$-, $\sigma^{38}$- and $\sigma^{70}$-reconstituted RNA polymerase holoenzymes were able to support Pm transcription in a rigorous XylS-dependent manner, as demonstrated by the formation of open complexes. Our *in vitro* transcription assays document RNA polymerase switching in Pm transcription, and confirm previous results *in vivo* (Marqués *et al.*, 1999).**

## Introduction

The XylS protein is the transcriptional activator of Pm, the promoter that controls expression of the *meta*-cleavage pathway for benzoate and alkylbenzoate catabolism coded by the *Pseudomonas putida* TOL plasmid (Marqués and Ramos, 1993; Ramos *et al.*, 1987). Two factors modulate transcription from Pm: the presence of effector molecules such as the pathway substrate 3–methylbenzoate (3MB), which triggers XylS activation, and intracellular levels of the regulatory protein. In fact, in the presence of the aromatic effector, XylS is synthesized at low levels from the $\sigma^{70}$-dependent $P_{S2}$ promoter, and becomes activated by the effector to promote expression from Pm. When XylS is overproduced from the inducible $P_{S1}$ promoter, e.g., in the presence of toluene, it activates Pm transcription through a cascade regulatory system in the absence of 3MB (Gallegos *et al.*, 1996a; González-Pérez *et al.*, 2004; Ramos *et al.*, 1987).

XylS belongs to the AraC/XylS family of transcription regulators (Ramos *et al.*, 1990). This family includes a large number of activators that regulate expression of genes involved in virulence, stress and metabolism. All members of the family share significant homology over a 100 amino acid stretch constituting the DNA binding domain (Gallegos *et al.*, 1997; Tobes and Ramos, 2002). Most family members also contain an effector-binding domain which modulates their activity and carries dimerization determinants (Bustos and Schleif, 1993; Michán *et al.*, 1995; Poore *et al.*, 2001; Prouty *et al.*, 2005; Ramos *et al.*, 1990). Early work based on structure prediction analysis allowed the definition of two HTH DNA binding motifs within the conserved DNA-binding domain (Gallegos *et al.*, 1997). Direct structural analysis of most proteins of the family has been hampered by protein insolubility. However, MarA and Rob proteins were crystallized in complex with their cognate binding sites in the DNA. The X-ray crystallographic data confirmed that the conserved DNA binding domain is composed of seven α-helices folding in two HTH motifs (Figure 1A). The two HTH recognition helices, α3 and α6, bind two adjacent segments of the major groove (Kwon *et al.*, 2000; Rhee *et al.*, 1998). Interestingly, both proteins showed a monomeric structure.

As in other members of the family, the DNA binding domain of XylS is located in the C-terminal end of the protein, connected by a linker domain to the 200 amino acid N-terminal end. Genetic evidence suggests this domain is involved in effector recognition (Michán *et al.*, 1992b; Ruíz and Ramos, 2002). In AraC, effector binding has major consequences on AraC structure and function. Scheilf and coworkers established that AraC operates as a dimer in which a flexible arm formed by the first 20 amino acids of the N-terminal domain is essential for the protein switch between two conformations with opposite

*For correspondence. E-mail silvia@eez.csic.es; Tel. (+34) 958 181608, Ext. 285; Fax (+34) 958 129600.

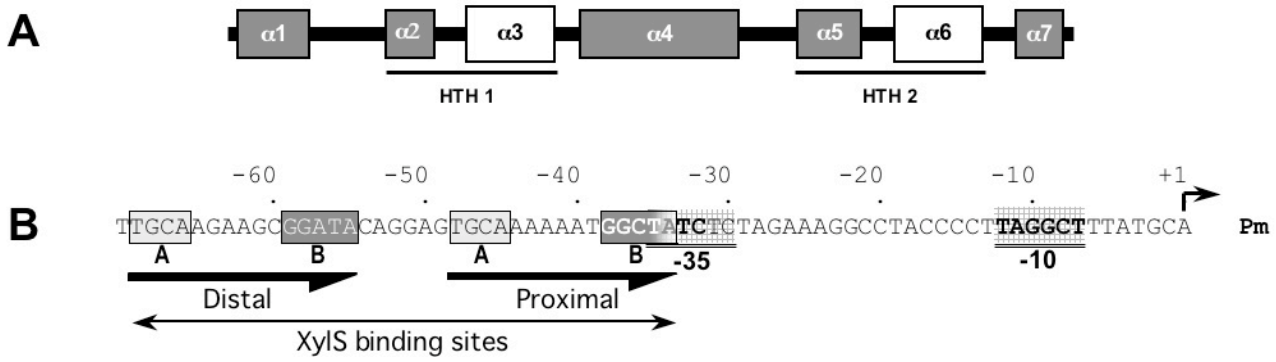



**Fig. 1.** Schematic structure of XylS protein C-terminal domain and Pm promoter.
A. Schematic representation of XylS C-terminal domain structure. The predicted (α-helixes in the sequence are depicted as grey boxes. The relative sizes of the different helixes are drawn to scale.
B. Pm promoter sequence organization. The rightward arrows indicate the two XylS binding sites (proximal and distal), each composed of conserved A (white-framed) and B (grey-framed) boxes. The –10 and –35 hexamers are in bold and double-underlined. A square arrow indicates the transcription initiation site.

functions (Lobell and Schleif, 1990; Soisson *et al*., 1997b) (see below). The molecular consequences of effector binding in XylS are unknown, and the oligomeric state of the XylS protein has not been determined directly. This protein, like many AraC/XylS family members, is insoluble at high concentrations, and accurate biochemical determinations have proved unattainable. However indirect evidence suggests that the XylS N-terminal domain was able to dimerize in a process that, at low protein concentrations, depended on the presence of the effector (Ruíz *et al*., 2003).

XylS recognizes two 15-bp direct repeats (TGCA-$N_6$-GGNTA), each consisting of two half-sites: a 5'-box A (TGCA) and a 3'-box B (GGNTA) (Figure 1B). The arrangement of the two repeats is such that the proximal XylS binding site overlaps the RNA polymerase -35 binding box by 2 bp (González-Pérez *et al*., 2002) (Figure 1B). The Pm promoter can be considered a class-II promoter according to Busby and Ebright's (1997) classification. In these promoters, regulators activate transcription by establishing multiple interactions with the RNAP α and σ subunits. Contacts with the C–terminal domain of the RNAP α subunit have been implicated in transcription activation by a number of AraC/XylS family members (Holcroft and Egan, 2000; Jair *et al*., 1995; Jair *et al*., 1996; Landini *et al*., 1997; Ruíz *et al*., 2001; Shah and Wolf, 2004). The mechanism of transcription activation used by some AraC/XylS family proteins also requires specific amino acid interactions with RNAP $\sigma^{70}$ subunit region 4 (Bhende and Egan, 2000; Grainger *et al*., 2004b; Landini and Busby, 1999; Ruíz *et al*., 2001).

Despite our substantial knowledge of XylS and other members of the AraC family of transcriptional activators, the molecular events underlying XylS activation remain unsolved. To address this question, we investigated the processes of XylS binding to DNA and protein dimerization, and we analyzed the role of the effector 3MB in these processes. Establishing the connections between these processes should shed light on the mechanism of XylS activation.

Although alternative activation mechanisms used by AraC family members are not well understood, some points have been clarified. MarA and SoxS proteins only include the DNA-binding domain, and both lack the effector responsiveness domain. They are synthesized *de novo* in response to their inducing signals, where upon their intracellular levels increase 10-fold and they activate their target promoters. Upon removal of the inducing signal, activator levels decrease through protein degradation and return to basal levels. Some effector-responsive AraC family members containing two functional domains influence DNA topology. In the absence of effectors, AraC and MelR bind two distant sites in the operator, giving rise to DNA loops that repress gene expression. Effector binding leads to a conformational change involving occupancy of adjacent DNA binding sites; the DNA loops are thus dismantled and gene expression is induced from the corresponding promoters (Kahramanoglou *et al*., 2006; Lobell and Schleif, 1990; Wade *et al*., 2000). Both AraC and MelR have been shown to recruit the RNA polymerase to their promoters, where they induce isomerization to form an open complex and thus transcription initiation (Grainger *et al*., 2004b; Zhang *et al*., 1996).

In this study we used purified preparations of XylS to analyze the activation mechanism. We were able to describe the different molecular events leading to promoter activation by the XylS protein, where the effector played a dual role: to favor XylS dimerization and to release intramolecular repression.

## Results

### *XylS specifically binds two direct repeats in Pm*

XylS, like most members of the AraC family, is poorly soluble and tends to aggregate when its concentration exceeds a threshold value, making it reluctant to biochemical analysis (Bhende and Egan, 1999; Jair *et al.*, 1995; Michán *et al.*, 1995; Munson and Scott, 1999; Timmes *et al.*, 2004). Although XylS binding sites have been determined from thorough genetic analyses (Gallegos *et al.*, 1996b; González-Pérez *et al.*, 1999; Kaldalu *et al.*, 1996), precise characterization of the binding process itself has been unfeasible. To overcome this obstacle we decided to use XylS concentrations below the aggregation threshold value. We devised a simple purification method in which concentration values were kept close to optimal, starting from cell extracts as described under "Experimental Procedures". Briefly, crude extracts prepared from *E. coli* CC118 cultures expressing XylS protein or mutant derivatives at moderately high levels were centrifuged and subjected to affinity chromatography through a heparin column. Bound and unbound fractions were tested with EMSA for their ability to bind a 100-bp DNA fragment containing the wild-type Pm promoter region from positions –10 to –110 with respect to the transcriptional start point. Only the heparin-unbound fraction was able to form a DNA–protein complex which appeared as a clearly shifted band (Figure 2A), suggesting that XylS binding to the heparin affinity matrix was weak. In the retardation assay, only an excess of unlabeled competitor DNA was able to titrate XylS out from its target DNA (Figure 2A, lane 8). The addition of nonspecific DNA did not reduce the shifted DNA-protein complex (Figure 2A, lane 9). To rule out the presence of unexpected Pm-DNA-binding proteins in the purified extracts, the labeled Pm target was incubated with increasing concentrations of similarly processed extracts containing or not the XylS protein. Only those extracts containing XylS produced retardation in DNA electrophoresis mobility assays (not shown).

The two direct repeats (TGCA-$N_6$-GGNTA) identified as XylS binding sites are located between positions –35 to –49, and –56 to –70 (Figure 1B), where the proximal binding site located between –35 and –49 was critical for XylS-dependent transcription from Pm (González-Pérez *et al.*, 1999). The Pm245 mutant promoter accumulates six point mutations in the proximal binding site (Figure 2B), and accordingly produced no detectable XylS–dependent transcriptional activity *in vivo* (González-Pérez *et al.*, 1999). We decided to investigate XylS binding to this promoter with EMSA. The results in Figure 2B show that XylS binding to Pm245 mutant DNA was significantly hampered, supporting the notion that the XylS protein present in cell extracts specifically bound Pm in the previous EMSA, and further confirming

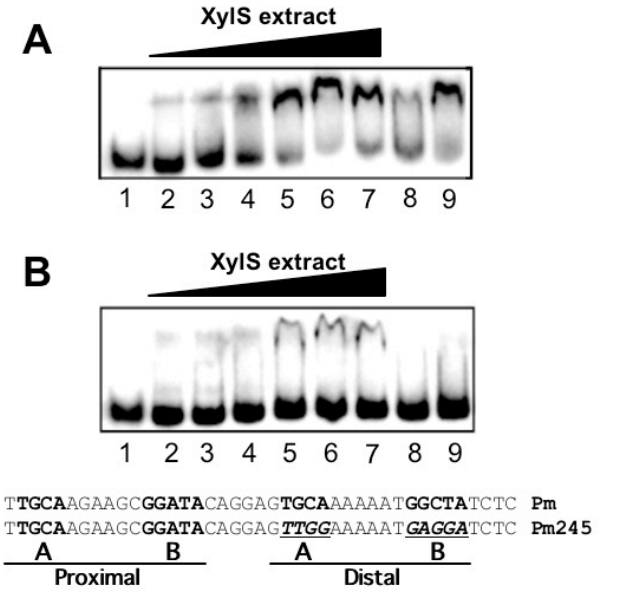



**Fig. 2.** DNA binding of XylS to Pm promoter. EMSA for binding of $^{32}$P-labeled (A) wild-type Pm DNA fragment or (B) a mutant Pm DNA fragment by purified extracts containing wild-type XylS protein. EMSA was performed as described in Experimental Procedures with either no protein added (Lane 1) of increasing amounts of XylS-containing extracts (lanes 2-7). An excess of specific (lane 8) or nonspecific competitor (lane 9) was added to reactions that also contained crude extract as in lane 7. Wild–type and mutant Pm sequences used in A or B, respectively, are shown. A and B box sequences are shown in bold, and mutations in Pm245 nucleotides are underlined.

the relevance of the B box in XylS binding. This finding is consistent with the absence of activity detected *in vivo* with this mutant.

The above series of EMSA results were identical in the presence and in the absence of 3MB, indicating that under these conditions the wild-type regulator was able to bind the target promoter *in vitro* in the absence of effectors.

### *Two roles for 3MB in XylS binding to the Pm promoter*

A general picture emerging from direct and indirect analyses of the AraC protein family suggests that those members which lack a defined N-terminal domain bind DNA as monomers, whereas the majority of two-domain proteins are dimers in solution and most probably bind DNA in a dimer conformation (Caswell *et al.*, 1992; Poore *et al.*, 2001; Prouty *et al.*, 2005; Soisson *et al.*, 1997a). It has been suggested that effector binding to the XylS N–terminal moiety causes a conformational change in this domain which stabilizes XylS dimers (Ruíz *et al.*, 2003). The mutant allele that encodes XylSL193A,L194A,I205A (hereafter XylS3L), in which alanine replaces three residues located in the antiparallel coiled-coil region connecting the N- and C-terminal domains, was unable to establish stable dimers *in vivo* in double-hybrid assays, and was unable to stimulate transcription either in the presence or in the absence of 3MB (Ruíz *et al.*, 2003). To

investigate the influence of dimerization on XylS DNA binding and the possible role of 3MB in this process, we prepared cell-free extracts from *E. coli* cultures expressing the XylS3L mutant. Incubation in the absence of 3MB of the wild-type Pm fragment with increasing amounts of extracts that contained XylS3L did not yield shifted bands, indicating that under these conditions the three-leucine dimerization mutant was unable to bind DNA. However, when 3MB was present, two well-defined shifted bands were observed: with low XylS3L concentrations a rapidly migrating band was observed; a second band with higher weigh appeared at higher protein concentrations (Figure 3). Migration of the latter band coincided with the migration previously observed with wild-type XylS cell extracts (not shown).

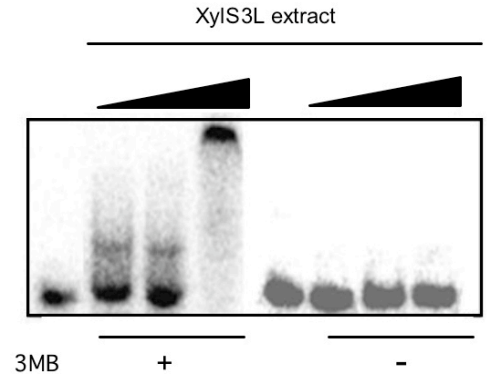



**Fig. 3.** DNA binding of mutant XylS3L to wild-type Pm promoter. EMSA for binding of $^{32}$P-labeled wild-type Pm DNA fragment in the presence (+) or absence (–) of 3MB by purified extracts containing mutant XylS3L protein. Mobility shift assays were performed as described in Experimental Procedures with either no protein added (first and fifth lane from the left) of increasing amounts of extracts that contained XylS3L (second to fourth lane and sixth to eighth lane).

Our EMSA results show that high concentrations of wild-type XylS were able to bind the Pm promoter even in the absence of 3MB (Figure 2A). A similar 3MB-independent activation was observed *in vivo* when XylS was overproduced in the cell, either through the natural cascade regulatory network (Ramos *et al.*, 1987; Ramos *et al.*, 1990) or via the control of an artificial strong promoter in a high-copy-number vector (Mermod *et al.*, 1987; Michán *et al.*, 1992a). The shift of the wild-type protein to the active form under EMSA conditions may result from high relative XylS protein concentrations, which would favor dimer formation (Ruíz *et al.*, 2003). In fact, previous studies showed that dimer formation was 3MB-independent *in vitro*, but 3MB-dependent *in vivo* (Ruíz *et al.*, 2003). If dimer formation triggered DNA binding, we would expect a XylS mutant unable to dimerize to also be unable to bind Pm. As expected, EMSA assays with mutant XylS3L clearly showed that high levels of the mutant protein used did not suffice for DNA binding. However, the presence of 3MB allowed mutant protein binding to DNA, despite its inability to activate transcription (Ruíz *et al.*, 2003). The results with wild-type and mutant XylS3L proteins *in vivo* and *in vitro* are strong evidence that DNA binding and protein dimerization are two connected but self-reliant processes, both of which are favored by the presence of 3MB. The absence of activity of XylS3L even in the presence of 3MB suggests that not only DNA binding but also dimerization, which does not occur in this mutant, are essential for the activation of Pm. We thus suggest a double role for 3MB in XylS activation under physiological conditions. On the one hand, 3MB-activated conformation would favor and stabilize dimerization, and in addition, the 3MB-induced change would be required to make the DNA binding surface accessible, a change which could also be achieved through protein dimerization. Each function probably entails a different protein determinant. It is worth noting that standard protein/DNA concentration ratios used for *in vitro* assays are generally more than one order of magnitude higher than the ratio found *in vivo* for a specific protein (Zimmerman and Trach, 1991), thus our *in vitro* assay conditions reflect a physiological situation of overproduction.

### *The XylS N-terminal domain represses C-terminal domain activity*

Both C- and N-terminal domains of XylS can act independently, and it has been suggested that interaction between the two domains allows the N-terminal domain to operate as an intramolecular repressor (Kaldalu *et al.*, 2000). The C-terminal domain, devoid of its N-terminal counterpart, shows constitutive activity, i.e., it activates the Pm promoter in the absence of effector (Domínguez-Cuevas *et al.*, 2007). This was taken as evidence of a repressive effect of the N-terminal domain on XylS-C binding activity. If this were the case, we would predict that both domains would establish stable interactions in solution, and these interactions would negatively influence XylS-C DNA binding. 3-Methylbenzoate would act as a modulator in this interaction, with consequences for DNA binding. To test our prediction and determine whether the linker between the two regions is required for interaction, we purified both N- and C-terminal domains expressed independently, with established protocols (Domínguez-Cuevas *et al.*, 2007) and analyzed XylS-C-DNA complex formation when a fixed concentration of the XylS-C domain (500 nM) was preincubated with increasing amounts of the purified and refolded XylS-N domain (from 500 nM to 4 μM) prior to the addition of labeled Pm promoter probe. Quantification of the complexes formed in each assay showed that the presence of the XylS-N domain in the mixtures repressed DNA binding by the XylS-C domain (Figure 4), suggesting that the two domains establish direct protein-protein interactions even when they are not connected by a linker. When 3MB was added to the preincubation mixture, the repression exerted by the N-terminal domain was significantly

reduced (Figure 4). Taken together, these results clearly suggest that the N-terminal region controls, through a direct interaction, the ability of the C-terminal domain to bind DNA, and that the effector 3MB acts to release repression. According to this set of results, the effects of 3MB on XylS can be visualized as a direct interaction that triggers a crucial conformational change, which simultaneously releases the protein constraints on both dimerization and DNA binding.

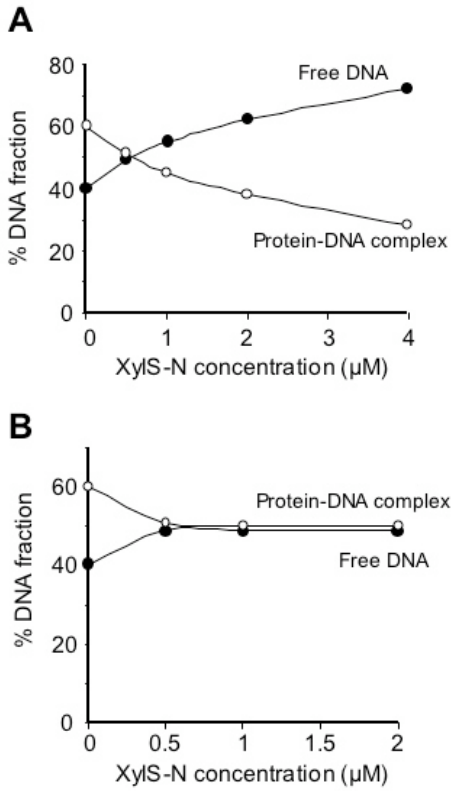



**Fig. 4.** N-terminal domain repression of XylS-C binding to Pm DNA.
A. EMSA for binding of $^{32}$P-labeled wild-type Pm DNA fragment by XylS-C in the presence and absence of increasing concentrations of purified XylS N-terminal domain. EMSA was performed as described in Experimental Procedures with either no protein added (black circles, free DNA) of a fixed amount of purified XylS-C (750 nM) in the presence of increasing concentrations of N-terminal XylS (from 0.5 to 4 μM, empty circles). The fraction of total radiolabel in each band was quantified in each lane and plotted as a function of XylS N–terminal domain concentration.
B. Effect of 3MB on XylS N-terminal domain repression of XylS–C DNA binding capacity. The experiment was carried out as in A) except that mixtures contained 1 mM 3MB. Black circles, free DNA; empty circles, increasing amounts of XylS N–terminal domain (from 0.5 to 2 μM).

### *XylS recruits RNA polymerase to the Pm promoter in response to 3MB*

XylS activates transcription from the Pm promoter in response to 3MB, and establishes direct interactions with the RNA polymerase α–CTD subunit (Ruíz *et al.*, 2001). In addition, the 1-bp overlap between the proximal XylS site and the σ subunit –35 binding box suggests direct contacts between XylS and the σ factor, as has been shown for other members of the family (Grainger *et al.*, 2004b; Wickstrum and Egan, 2004). To investigate the steps leading to this interaction and the influence of 3MB on the process, we carried out ChIP assays to measure XylS and RNA polymerase binding to the Pm promoter in the presence or absence of 3MB. *In vivo* cross-linked nucleoprotein samples are immunoprecipitated with antibodies directed against either XylS or the α-subunit of RNA polymerase. Then immunoprecipitated protein-bound DNA fragments (500 bp to 1 kb) are analyzed and quantified with real-time PCR using Pm-specific primers.

*Escherichia coli* strain CC118λPm::*lacZ* carrying a Pm::*lacZ* fusion with or without a plasmid bearing the wild–type XylS (pERD103) was grown in the presence or absence of 3MB. Cells were treated with formaldehyde to freeze protein-DNA bonds formed *in vivo*, and DNA was extracted from lysed cells and sheared by sonication. Either polyclonal anti-XylS antiserum or monoclonal anti-σ subunit antibody were used to immunoprecipitate DNA fragments. Figure 5 shows the fold enrichments obtained in real-time PCR analysis with primers specific for the Pm promoter. A pair of primers designed to amplify the *mtlA* promoter was used as a reference. ChIP with XylS antibody resulted in a two-fold enrichment of XylS-bound Pm in response to 3MB (Figure 5A), suggesting that the aromatic effector favored XylS binding to DNA. To further analyze the influence of XylS on RNA polymerase binding to Pm, immunoprecipitates were prepared with antibodies directed against the RNA polymerase σ subunit. Figure 5B shows that enrichment of the Pm promoter DNA was greatest in cells grown in the presence of both XylS activator and 3MB. This supports the hypothesis that XylS enhances the association of RNA polymerase to the Pm promoter in the presence of 3MB; in other words, XylS recruits RNA polymerase to Pm. Figure 5B also shows that RNA polymerase was able to bind the Pm promoter in response to 3MB even in the absence of the XylS activator. We have previously shown that the addition of an aromatic compound such as 3MB to cells growing at 30ºC triggers the heat-shock response (Marqués *et al.*, 1999), subsequently stabilizing the $\sigma^{32}$ subunit (Domínguez-Cuevas *et al.*, 2006). A 42ºC heat-shock produced a low but consistent increase in basal Pm activity in the absence of effector, presumably as a consequence of the increase in $\sigma^{32}$ concentration (Marqués *et al.*, 1999). Thus, the two-fold enrichment in anti-σ-captured Pm observed in the presence of 3MB and the absence of XylS is in agreement with the previously observed increase in Pm transcription during the heat–shock response (Figure 5B). However, from the results in Figure 5 we could not determine whether XylS recruited RNA polymerase before (i.e., in solution) or after binding DNA.

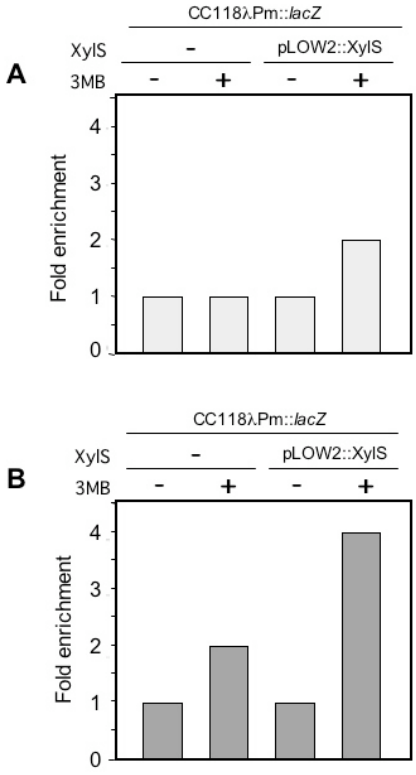



**Fig. 5.** Chromatin immunoprecipitation (ChIP) analysis of RNA polymerase and XylS binding to the Pm promoter. ChIP was carried out as described in Experimental Procedures. (A) XylS or (B) RNA polymerase binding to the Pm promoter in the presence or absence of 3MB. The strains used were *E. coli* CC118Pm::*lacZ* bearing or not pERD103, grown in the presence or absence of 3MB. Immunoprecipitated samples obtained in the presence of antibodies and control samples obtained without antibodies were analyzed by RT-PCR with two pairs of primers for Pm promoter and *mtlA* gene, used as reference. The ratio between quantitation of Pm and *mtlA* PCR products in each sample is shown in panel A and B. Pm/*mtlA* ratio between samples obtained in the absence of antibody remained unaltered in all growth conditions.

### *XylS promotes open complex formation at the Pm promoter*

Previous studies have shown that Pm transcription *in vivo* required two distinct RNA polymerases: E$\sigma^{32}$ in the exponential phase and E$\sigma^{38}$ in the stationary phase (Domínguez-Cuevas *et al*., 2005; Marqués *et al*., 1999) To confirm this unusual  requirement, we used $KMnO_4$ footprinting to analyze the formation of open complexes with the Pm promoter by E$\sigma^{32}$, E$\sigma^{38}$ and E$\sigma^{70}$. To assess the dependence of complex formation on the XylS protein, we used XylS-C, which was shown to faithfully reproduce the wild-type XylS activity except that, as expected, XylS-C was active even in the absence of effector (Domínguez-Cuevas *et al*., 2007). An end–labeled Pm fragment was treated with potassium permanganate after preincubation with each of the three different reconstituted RNA polymerases, in the presence and absence of XylS-C. The results in Figure 6 show that in the absence of XylS-C, none of the reconstituted alternative RNA polymerases produced significant changes in the reactivity pattern with respect to controls (Figure 6, lanes 2 to 4). However, in the presence of XylS-C, all three RNA polymerases induced clear reactivity of the thymines between positions -11 and +2 of the Pm promoter, a reaction distinctive of the formation of a transcription bubble (Figure 6, lanes 5 to 7). The reactivity pattern was identical with each of the three RNA polymerases, except that intensity was lower with E$\sigma^{32}$. When mutant promoter Pm-12C (described as totally inactive) was analyzed in permanganate footprinting assays, the reactivity pattern for all three RNA polymerases was different and incompatible with transcription bubble formation. A hyperactive band at -8 not observed in the wild-type Pm was clearly visible with all three RNA polymerases. These findings are the first biochemical evidence of two alternative RNA polymerases that recognize the Pm promoter at the same initiation site, in a consistently XylS-dependent manner, and confirm previous results obtained *in vivo* (Gallegos *et al*., 1996b; Marqués *et al*., 1999).

Previous findings showed that in an *rpoS*, *rpoH* double mutant, Pm transcription remained at basal levels, suggesting that activity of E$\sigma^{70}$ *in vivo* was low (Marqués *et al*., 1999). This apparent contradiction with the high open complex formation rate found *in vitro* with this holoenzyme (Figure 6A, lane 7) has been reported previously for several promoters, which may be transcribed *in vitro* by both E$\sigma^{38}$ and E$\sigma^{70}$, but which *in vivo* are strictly dependent on E$\sigma^{38}$ (Tanaka *et al*., 1993). It is worth noting that the RNA polymerase holoenzyme with the alternative sigma factors $\sigma^{32}$ and $\sigma^{38}$ was reconstituted for this experiment with commercial core and purified His-tagged $\sigma^{32}$ or $\sigma^{38}$ factors, whereas E$\sigma^{70}$ was purchased directly as a holoenzyme. The low open

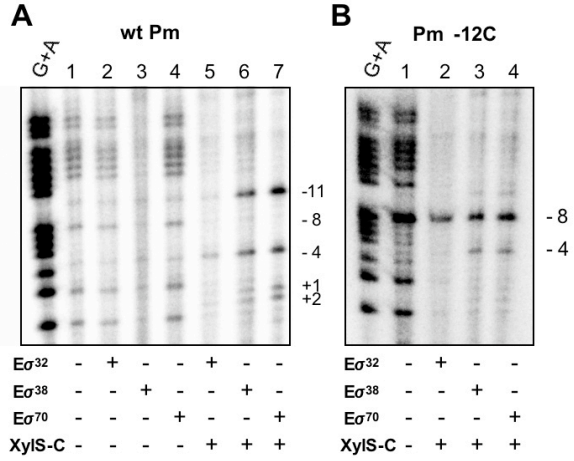



**Fig. 6.** Potassium permanganate footprinting of the Pm promoter. The DNA (bottom strand) in the presence of either no protein (lane 1) or RNA polymerase with different sigma factors in the presence (lanes 5-7, 2-4) or absence (lanes 2-4) of XylS-C was modified with potassium permanganate (10 mM) and cleaved with piperidine. The numbers ( -11, -8, -4, +1 and +2) indicate the cleavage sites. A) wt Pm promoter; B: mutant Pm -12C promoter.

complex formation rate with E$\sigma^{32}$ can be ascribed to the lower amounts of this holoenzyme, since our reconstitution conditions did not compensate for possible differences in $\sigma$ factor affinity for the core (Maeda et al., 2000).

### *XylS-C promotes* in vitro *transcription from the Pm promoter*

To confirm the essential nature of the role of XylS in Pm transcription, we explored its ability to activate transcription with different RNA polymerase holoenzymes in an *in vitro* transcription assay. Single-round transcription assays were performed on both supercoiled plasmids and linear DNA templates carrying the Pm wild-type promoter between positions -113 and +35. *In vitro* transcription reactions were carried out with RNAP in the presence or absence of the constitutive truncated XylS derivative XylS-C, which did not require 3MB for activation (Domínguez-Cuevas *et al*., 2007). The results in Figure 7 show a strong signal corresponding to the expected transcription product with E$\sigma^{38}$, which was considerably weaker with E$\sigma^{38}$, though detectable. In both cases, transcription was dependent on the presence of the XylS-C regulator.

The relative transcription rates observed *in vitro* with both RNAPs were not consistent with previous data obtained *in vivo* (Domínguez-Cuevas *et al*., 2005). We should take into account, however, that *in vitro* transcription assays were not influenced by physiological

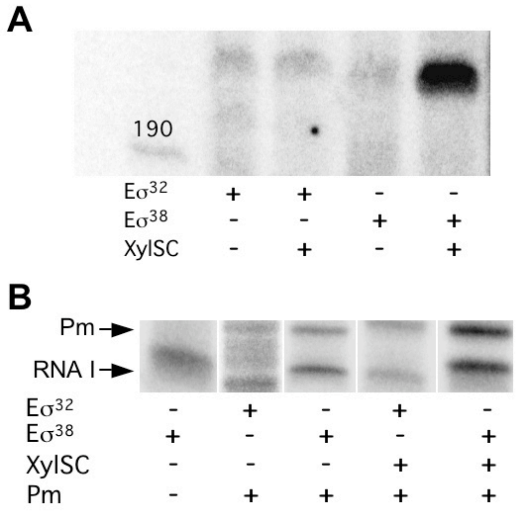



**Fig. 7.** Effect of XylS on Pm transcription. (A) Single-round *in vitro* transcription assay of a linear DNA template using reconstituted RNA polymerase with either $\sigma^{32}$ or $\sigma^{38}$ and XylS-C in the presence of $\alpha$[$^{32}$P]-UTP-labeled nucleotide. (B) Single-round *in vitro* transcription assay of a supercoiled DNA template. Migration of transcripts corresponding to either Pm or the reference transcription product RNAI is indicated with arrows. The differences in size of RNA I transcripts obtained with E$\sigma^{32}$ or E$\sigma^{38}$ are due to different origins of the plasmid transcripts depending on the RNA polymerase used.

modulators such as ppGpp concentration, the differential effect of supercoiling on alternative RNAPs, sigma factor levels, competition and affinity for the RNAP core, which are known to play an essential role in Pm transcription regulation *in vivo* (Domínguez-Cuevas *et al*., 2005).

## Discussion

The nature of XylS functioning in its cell host requires this regulator to activate transcription from the Pm promoter in response to 3MB; and that its overproduction through a toluene-induced cascade regulatory circuit leads to transcriptional activation in a 3MB-independent manner (Marqués and Ramos, 1993; Ramos *et al*., 1986; Ramos *et al*., 1987). Our findings provide a likely explanation for this apparent contradiction, and we present evidence supporting a molecular model to explain the mechanism of XylS activation when produced at low and high concentrations.

XylS belongs to the AraC family of transcriptional regulators defined by a conserved 100-amino acid region, although most family proteins contain an additional receiver domain (Gallegos *et al*., 1997; Tobes and Ramos, 2002). Activation mechanisms used by AraC family members have been studied in detail in some proteins, and processes such as interaction with RNA polymerase $\alpha$ and $\sigma$ subunits, N-terminal domain modulation, DNA looping or overproduction have been implicated in transcriptional activation (Griffith and Wolf, 2004; Kahramanoglou *et al*., 2006; Lobell and Schleif, 1990; Martin and Rosner, 2001). AraC, the best known protein in the family, controls gene expression through DNA looping by means of an N-terminal arm (residues 7 to 18) shown to establish alternative interactions with either the DNA binding domain or the arabinose binding pocket in a ligand-dependent manner (Weldon and Schleif, 2006; Weldon *et al*., 2007). XylS-AraC sequence alignment did not reveal the presence of a structure homologous to this N-terminal arm in XylS, therefore the mechanism is likely to be different. MelR, an AraC family protein regulating melibiose degradation, also seems to use DNA looping to control transcription from *melA* and *melR* promoters in response to melibiose, through an as yet unknown mechanism (Kahramanoglou *et al*., 2006; Wade *et al*., 2000). In XylS it has been established that the N-terminal and C-terminal domains are functionally independent, and the N-terminal domain is involved in effector recognition and dimerization (Kaldalu *et al*., 2000; Ruíz and Ramos, 2002; Ruíz *et al*., 2003). In contrast, the XylS C-terminal domain lacks effector responsiveness and activates transcription constitutively (Domínguez-Cuevas *et al*., 2007). It is worth noting that this is a unique feature among two-domain AraC family proteins, which are totally inactive when their N-terminal domain is deleted. XylS N-terminal domain modelling against AraC N-terminal domain showed that the coiled-coil region involved in AraC dimerization was highly conserved in

XylS. This is consistent with previous results pointing to three conserved leucine residues in the XylS antiparallel coiled-coil region being involved in XylS dimerization.

XylS presents a unique feature not shared by AraC or MelR: in the absence of effector it is able to reach an active conformational state when its concentration exceeds a threshold value. Interestingly, this behavior is similar to the activation mechanism in family proteins such as MarA and SoxS containing one single domain, where induction of the pathway in response to specific signals increases intracellular levels of the regulator and triggers activation of the target promoter (Griffith *et al.*, 2004; Griffith and Wolf, 2004). XylS shares with MarA and Rob that they all respond to signal molecules that induce stress response; in fact, both 3MB (the effector of XylS) and toluene (which triggers the cascade circuit leading to XylS overproduction) were shown to induce the heat-shock response (Marqués *et al.*, 1999; Van Dyk *et al.*, 1995). Furthermore, XylS activation of the Pm promoter was functional with $E\sigma^{32}$ RNA polymerase (Figures 6 and 7), consistent with the requirement of this RNA polymerase for transcription *in vivo*. These observations point to a strategy based on regulator overproduction in AraC family-dependent regulation processes responsive to stress conditions. The evolutionary implications of this finding are interesting: considering the distinct two-domain structure of XylS among the stress-responsive AraC family regulators, one is tempted to suggest that XylS may have originated from a fusion protein that joined two functional domains: a MarA-like DNA-binding element, and an aromatic responsive element. Interestingly in *E. coli* a n untranslated sequence related to the XylS N-terminal domain was found directly upstream from MarA (Hachler *et al.*, 1996). This untranslated sequence could be superimposed on the unrelated MarR element, although in a different frame.

To address the molecular function of 3MB in Pm activation by XylS, we reasoned *a priori* that 3MB could play two different roles: 1) directly favoring XylS DNA binding, or 2) promoting protein dimerization, thus indirectly favoring DNA binding. In addition, we examined the mechanism of XylS activation in the absence of 3MB, a physiological process known as a "cascade circuit" which depends on increased XylS concentrations in the cell (Gallegos *et al.*, 1996a; Ramos *et al.*, 1987).

To overcome the difficulty imposed by the inherent insolubility of XylS, we developed a gentle purification procedure with heparin-affinity chromatography to eliminate DNA-binding proteins in the extracts. This step greatly improve the quality of the retardation assays despite the use of crude extracts (Figure 2 and 3), and turned out to be especially suitable for the extensive analysis of a number of XylS mutants. Appropriate controls ensured specific XylS binding to its target DNA. Our results with wild-type XylS-containing extracts showed that XylS was able to bind DNA in a 3MB–independent manner, in contrast to what we observed in ChIP experiments *in vivo*, where the addition of the effector specifically increased the amount of XylS bound to its target promoter (Figure 5B). On the other hand, a XylS mutant defective in dimerization was completely dependent on the presence of 3MB to bind Pm in gel retardation assays. Previous results suggested that XylS dimerization depended on 3MB *in vivo* when XylS was present at low levels in the cell, whereas the high protein levels used in *in vitro* assays led to XylS dimerization independently of the effector (Ruíz *et al.*, 2003). Irrespective of the pathway used in dimer formation, XylS dimerization was essential for Pm activation. Considering the suggested dynamic equilibrium between inactive and active XylS conformations in the cell (Mermod *et al.*, 1987), our results point to a role for 3MB that favors a shift towards the active form, a process that could also be modulated by XylS protein concentration. The switch from the "off" state to the "on" state of the regulator mediated by effector binding or overproduction was also described for regulators belonging to other families of transcriptional regulators (Raibaud and Schwartz, 1984)

It was established that recognition of 3MB by the XylS N-terminal domain favored the dimeric state of the protein (Ruíz *et al.*, 2003). In fact, when protein oligomerization was impeded, i.e., in the XylS3L mutant, 3MB became essential for XylS to bind DNA. Thus, as regards DNA binding, the conformational change resulting from dimerization at high protein concentrations may unmask DNA binding determinants in a manner similar to the way in which 3MB activates XylS when it binds the regulator at low protein concentration. The fact that the XylS C–terminal domain was able to activate Pm in the absence of 3MB, together with the observation that binding of 3MB to XylS led to a conformational change essential for activation when dimerization was blocked indicate that the N-terminal domain acts as an intramolecular repressor. Intramolecular repression of the DNA binding domain has been described for members of the NtrC family of regulators such as DmpR and XylR, in which deletion of the N-terminal effector binding domain resulted in an activator that mediated transcription constitutively (Fernández *et al.*, 1995; Ng *et al.*, 1996).

To test the possibility that the XylS N-terminal domain could repress the C-terminal domain through direct interaction and not just by signal transduction of the 3MB–triggered conformational change, we performed gel shift assays with the XylS C-terminal domain in the presence of increasing concentrations of the N-terminal domain. The decrease in DNA bound by the XylS DNA binding domain after pre-incubation with the N-terminal domain demonstrated that intramolecular repression was the result of a direct interaction between independently functional domains, and not of the connection of these domains by the flexible linker.

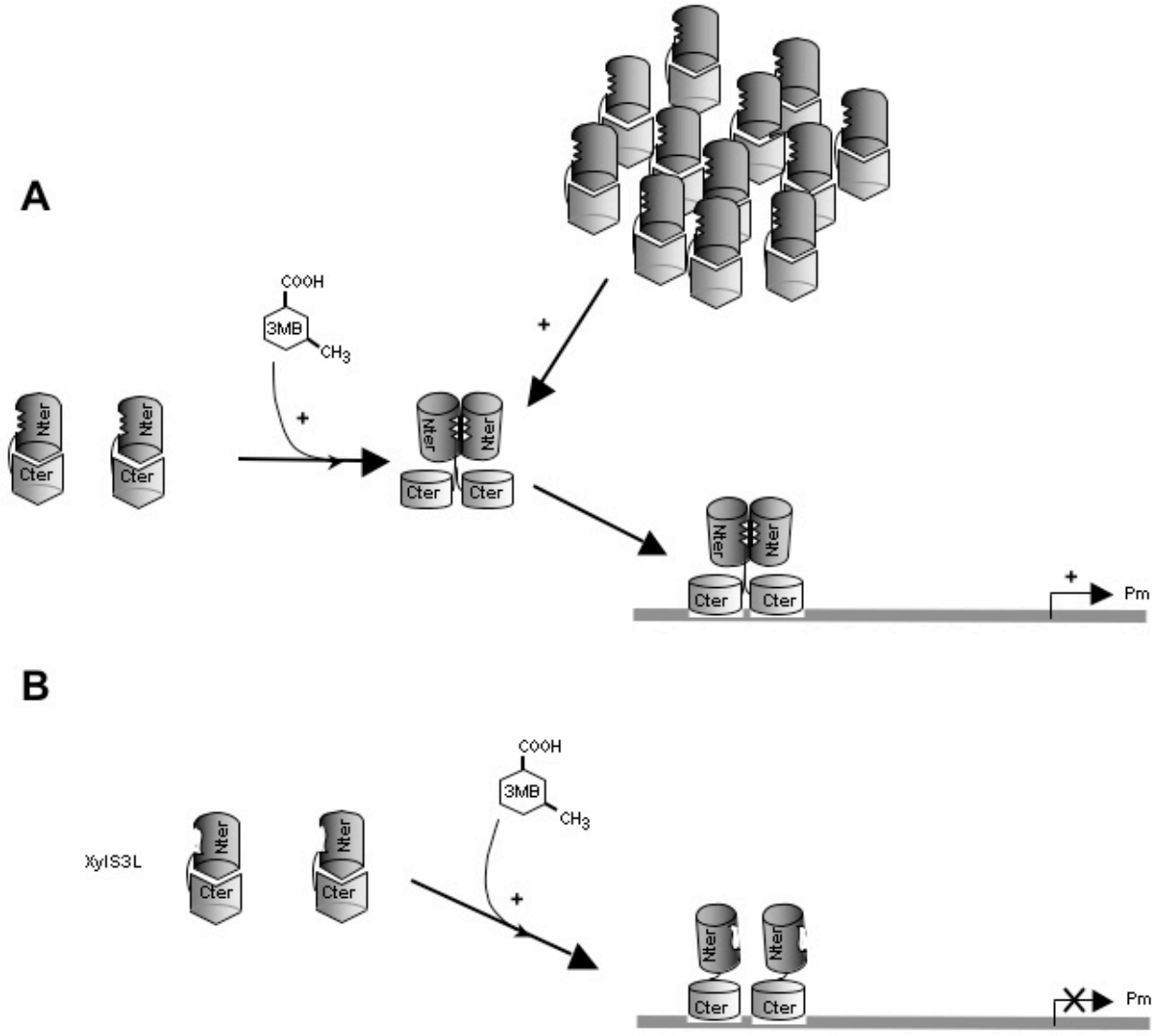



**Fig. 8.** Model for XylS activation of the Pm promoter.
A Under basal conditions, XylS DNA-binding domains (light grey barrels) are unable to make contact with DNA. The addition of 3MB produces a conformational change with two important consequences: 1) dimerization interactions are favored between monomers, probably because monomer dimerization regions are exposed, and 2) the regulator DNA-binding domains are opened, which favors contacts with DNA. Under basal physiological expression conditions (i.e., *in vivo*), XylS does not dimerize unless 3MB is present. When overexpressed or in purified preparations, high concentrations of XylS favor dimerization in the absence of effector. Dimerization might also lead to conformational changes which make DNA-binding domains more available for interactions with target DNA sites. In both cases, XylS binding to Pm in the presence of RNA polymerase activates the promoter
B. In XylS-3L, 3MB generates the corresponding conformational change, but the result in this case is only the opening of DNA-binding sites, since this mutant is unable to dimerize. Accordingly, the bound protein is not able to promote transcription.

XylS binding to its operator sequences at Pm, and binding of RNA polymerase to its specific recognition sequence. Further sequential events involve isomerization of the promoter from a closed to an open complex, and transcription initiation. We investigated the specific role(s) of XylS binding and 3MB in this process. ChIP assays allowed us to measure the capacity of 3MB to recruit RNAP to the Pm promoter. Using an antibody directed against the RNA polymerase  subunit, we observed that the Pm promoter fraction bound to RNA polymerase was higher in cells grown with 3MB (Figure 5B). This enrichment was more significant when XylS protein was present, demonstrating that XylS recruited RNA polymerase to the Pm promoter in response to 3MB. This was further confirmed by the data obtained with anti-XylS antiserum, which showed an increase in XylS-bound Pm in the presence of effector (Figure 5A).

Marqués *et al*. (1999) showed that transcription from the Pm promoter is mediated by two alternative RNA polymerases: E$\sigma^{32}$ or E$\sigma^{38}$ (Domínguez-Cuevas *et al.*,

2005; Marqués *et al*., 1999). Alternation of the RNA polymerases at Pm was further confirmed in permanganate footprinting assays of the XylS DNA binding domain, which showed a distinct pattern in the presence of each of the three RNA polymerases when the XylS activator was present. This indicates that the interaction between XylS and RNA polymerase leads to an increase in the rate of isomerization from a closed to an open complex. This was further confirmed in *in vitro* assays showing a clear XylS-C requirement to obtain full transcription activity with both $\sigma^{32}$ and $\sigma^{38}$.

In this work we have deciphered the molecular role of 3MB in Pm activation by XylS. Two central processes in 3MB functioning are established (Figure 8). First, 3MB favors XylS DNA binding by altering direct interactions between XylS N-terminal and C-terminal domains. Second, the conformational change concomitant to dimerization which occurs either after 3MB activation or through an increase in XylS concentration unmask the DNA binding domain from the constraint imposed by the N-terminal domain. Dimer formation, favored at high protein concentrations, allows XylS to bind DNA and activate transcription from Pm in the absence of 3MB. It is worth noting that although 3MB allows XylS-3L to bind DNA, this mutant remained unable to activate transcription. This led us to conclude that XylS dimerization is essential to activate transcription from Pm. Finally, our results show that XylS recruits RNA polymerase to the Pm promoter in response to 3MB, and subsequently increases the rate of isomerization of RNA polymerase from closed to open complexes.

## Experimental procedures

### *Bacterial strains and plasmids*

The bacterial strains and plasmids used in this work are listed in Table 1. Bacterial strains were routinely grown at 30ºC and 200 rpm in Luria-Bertani medium as described before (Domínguez-Cuevas *et al*., 2005). Growth was determined turbidometrically at 660 nm. The plasmids pCMX2, pLOW2, pJLR100, pJLR107, pERD103 and pMD::Pm245 have been described previously (Table 1).

### *DNA techniques*

All DNA manipulations were done according to standard procedures (Ausubel *et al*., 2007) or to the manufacturer's recommendations.

### *Crude extract preparation*

*Escherichia coli* CC118 bearing pCMX2 or pLOW2::xylS was grown overnight in LB medium with tetracycline (10 µg/ml) or kanamycin (25 µg/ml) at 30ºC and 200 rpm. Flasks containing 500 ml fresh medium were inoculated with an aliquot fraction of these precultures and were incubated at 30ºC. When OD660 reached 0.6, 1 mM 3MB was added and incubation was continued for 3 h. The cultures were then harvested by centrifugation and cell pellets were stored at –70ºC. Crude extracts were prepared by resuspending the pellet in 3 ml lysis buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 4 mM β–mercaptoethanol, 2 mM EDTA, 1x Complete™ protease inhibitor mixture (Roche Applied Science) followed by cell disruption by sonication. The clear supernatant was filtered through a 0.45-µm pore-size nylon membrane and loaded onto a 1-ml Heparin column (HiTrap Heparin HP, Amersham Biosciences) and pre-equilibrated with buffer A (10 mM sodium phosphate pH 7, 4 mM β-mercaptoethanol) in the presence or absence of 1 mM 3MB. Unbound material was collected and the column was washed with buffer A until nonspecifically bound material had been removed. Specifically bound proteins were eluted from the column with buffer A supplemented with 1 M NaCl. The protein concentration in the fractions was determined according to Bradford (1976).

### *Overexpression and purification of His-tagged XylS C-terminal and N-terminal domains*

The 400-bp DNA fragment covering the XylS C-terminal domain (XylS residues 196-321) was amplified by PCR from pERD103 plasmid using primers XylSCNdeI (5'-GGAATTCCATATGCTGGGCAGCAATGTCAGC-3') and XylSXhoI (5'-CCGCTCGAGTCAAGCCACTTCCTTTTTG C-3'). The PCR product was digested with *Nde*I and *Xho*I enzymes and subsequently cloned into the pET16b vector (Novagen). The histidine-tagged XylS C-terminal domain (hereafter XylS-C) was overexpressed and purified as described (Domínguez-Cuevas *et al*., 2007).

The 643-bp DNA fragment covering the XylS N-terminal domain (XylS residues 1 to 207) was amplified by PCR from pERD103 plasmid using primers XylSNdeI (5'-GGAATTCCATATGGATTTTTGCTTATTGAACGAG -3') and XylSNterXhoI (5'-CCGCTCGAGTCAGCTGAAAATT TCACGGCTGAC-3'). The PCR product was digested with *Nde*I and *Xho*I enzymes and subsequently cloned into the pET16b vector (Novagen). Freshly transformed BL21(DE3) cells harboring the pET16b::xylSN plasmid were grown in 500 ml 2x YT medium (Sambrook *et al*., 1989) at 30ºC until turbidity at 660 nm reached 0.5. At this point the culture was transferred to 16ºC, and the addition of 0.1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) allowed the induction of fusion proteins during the subsequent 3-4 h of incubation, after which cells were pelleted and frozen. The cells were resuspended in 50 ml lysis buffer (30 mM TrisHCl, pH 8.8, 300 mM NaCl, 0.1 mM EDTA, 2.5 mM 2-mercaptoethanol, 10% [v/v] glycerol, 10 mM imidazole, 0.5% Triton X-100 and 1 mM Complete™ protease inhibitor mixture (Roche Applied Science) and disrupted in a French pressure cell. After centrifugation at 22 500 *g* for 45 min, the inclusion body was resuspended in 60 ml solubilization buffer (30 mM

**Table 1.** Bacterial strains and plasmids used in this work

| Strain / plasmid | Relevant characteristics[a] | Source or reference |
|---|---|---|
| Strains: | | |
| *E. coli* CC118 | Δ(*ara-leu*) *araD* Δ*lacX74 galE galK phoA20 thi-1 rpsE rpoB argE recA1* | Manoil and Beckwith (1985) |
| *E. coli* BL21(DE3) | Carries T7 RNA polymerase under the control of *lacUV5* promoter | Novagen |
| *E. coli* DH5α | *supE44 lacU169*(*Ø80lacZΔM15*) *hsdR17 recA1 endA1 gyrA96 thi–1 relA1* | Hanahan (1983) |
| *E. coli* MC4100 | F⁻, *araD139* (*argF-lac*)*U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rpsR* | Lange and Hengge-Aronis (1991) |
| Plasmids: | | |
| pJLR100 | $Ap^R$, Pm cloned in pEMBL9 | Ramos *et al.* (1986) |
| pCMX2 | $Tc^R$, pSELECT derivative containing the wild-type *xylS* gene | Manzanera *et al.* (2000) |
| pERD103 | $Km^R$, derivative containing the wild-type *xylS* gene | Ramos *et al.* (1987) |
| pLOW2 | $Km^R$, pACYC177 derivative, low-copy-number cloning vector | Hansen *et al.* (1997) |
| pJLR107 | $Ap^R$, Pm::*lacZ* in pMD1405 | Ramos *et al.* (1986) |
| pMD::Pm245 | $Ap^R$, Pm245::*lacZ* in pMD1405 | González-Pérez *et al.* (1999) |
| pET16b::xylSC | $Ap^R$, pET16b derivative used to produce His-tagged XylS C-terminal domain | Domínguez-Cuevas *et al*. (2007) |
| pET16b::rpoS | $Ap^R$, pET16b derivative used to produce His-tagged *P. putida* $\sigma^{38}$ | This work |
| pET16b::rpoH | $Ap^R$, pET16b derivative used to produce His-tagged *P. putida* $\sigma^{32}$ | This work |
| pET16b::xylSN | $Ap^R$, pET16b derivative used to produce His-tagged XylS N-terminal domain | This work |

**a.** $Ap^R$, $Km^R$ and $Tc^R$ stand for resistance to ampicillin, kanamycin and tetracycline, respectively.

Tris HCl, pH 8.8, 500 mM NaCl, 0.1 mM EDTA, pH 8, 2.5 mM 2-mercaptoethanol, 10% [v/v] glycerol, 10 mM imidazole, 0.05% Triton X-100 and 6 M guanidine-HCl). After centrifugation at 30 000 *g* for 60 min, the supernatant fraction was passed through a 0.45-μm pore-size filter and loaded onto a 5-ml Ni-agarose column (Amersham Biosciences) pre-equilibrated with solubilization buffer. After nonspecifically bound material had been removed, the His-tagged XylS N-terminal domain (hereafter XylS-N) was eluted from the column with about 400 mM imidazole in a 30-ml elution buffer gradient from 0-1 M imidazole. The purified XylS-N was refolded by sequential dialysis against dialysis buffers (30 mM Tris HCl, pH 8.8, 300 mM KCl, 0.1 mM EDTA, 2 mM DTT, 10% [v/v] glycerol, 1 mM GSH/0.2 mM GSSG) and decreasing urea concentrations from 4 M to 0 M.

*Electrophoretic mobility shift assays*

The 100-bp DNA fragments containing the wild-type or mutant xylS binding site (positions -110 to -10 of the Pm promoter) were amplified by PCR with primers Pm3 (5'-CTGCAGTGTCCGGTTTGATAGGG-3') and Pm4 (5'-CCTAAGGGGTAGGCCTTTCTAG-3'). The PCR products were isolated from agarose gels and end-labeled with [γ-$^{32}$P]ATP as described (Terán *et al.*, 2003). Cell extracts and end-labeled DNA fragments were mixed and incubated at 30ºC for 15 min in 10 μl binding buffer (Tris-glycine buffer [25 mM Tris-HCl, 200 mM glycine, pH 8.6], 200 mM NaCl, 4 mM β-mercaptoethanol, 4 mM $MgCl_2$, 4 mM EDTA), supplemented or not with 1 mM 3MB. Samples were loaded onto a 4.5% nondenaturing polyacrylamide gel and eletrophoresed at 50 V in Tris-glycine buffer (25 mM Tris-HCl, 200 mM glycine, pH 8.6) for 2 h at 4ºC. The gels were dried and visualized by exposure to Phosphorimager screens. The results were analyzed with Molecular Imager FX equipment and QuantityOne software (Bio-Rad, Madrid, Spain).

*Overexpression and purification of* Pseudomonas putida *sigma factors*

The 874-bp and 1029-bp DNA fragments containing the *rpoH* and *rpoS* genes respectively were amplified by PCR from *P. putida* KT2440 chromosomal DNA using primers rpoHNdeI (5'-GGAATTCCATATGACCACATCGTTGCAACC-3'), rpoHXhoI (5'-CCGCTCGAGTCAGGCAGCGATCAGTGCC-3'), rpoSNdeI (5'-GGAATTCCATATGGCTCTCAGTAAAGAAGT-3') and rpoSXhoI (5'-CCGCTCGAGAGCTACTGGAACAATGACTCG-3'). The PCR products were digested with *Nde*I and *Xho*I enzymes and subsequently cloned into the pET16b vector (Novagen).

Freshly transformed BL21 (DE3) cells harboring the pET16b::*rpoH* plasmid were grown in 500 ml 2x YT medium (Sambrook *et al.*, 1989) at 30ºC until turbidity at 660 nm reached 0.4. Cultures were then incubated at 16ºC and incubated with 0.25 mM IPTG for 3 h, after which cells were pelleted and frozen. The cell pellet was resuspended in 80 ml lysis buffer (50 mM TrisHCl, pH 8, 50 mM NaCl, 0.5 mM EDTA, pH 8, 5% [v/v] glycerol, 10 mM imidazole, 10 μl benzonase (Roche) and 1 mM Complete™ protease inhibitor mixture (Roche Applied Science) and disrupted in a French pressure cell. After centrifugation at 30 000 *g* for 1 h, the soluble extract was passed through a 0.45-μm pore-size filter and loaded onto a 5-ml Ni-agarose column (Amersham Biosciences) pre-equilibrated with buffer A-RpoH (50 mM TrisHCl, pH 8, 50 mM NaCl, 0.5 mM EDTA, pH 8, 5% glycerol, 10 mM imidazole). The column was washed with buffer A until

nonspecifically bound material had been removed. RpoH eluted from the column at about 400 mM imidazole in a 30-ml imidazole gradient in buffer B-RpoH (buffer A-RpoH + 750 mM imidazole) from 55 mM to 750 mM imidazole. Eluted fractions containing RpoH protein were dialyzed against storage buffer (50 mM Tris HCl, pH 8, 50 mM NaCl, 0.5 mM EDTA, pH 8, 20% glycerol) and stored at -70ºC until use.

Freshly transformed BL21 (DE3) cells harboring the pET16b::*rpoS* plasmid were grown in 500 ml 2 □ YT medium (Sambrook *et al.*, 1989) at 30ºC until turbidity at 660 nm reached 0.4. Cultures were then incubated at 16ºC and incubated with 0.25 mM IPTG for 3 h, after which cells were pelleted and frozen. The cell pellet was resuspended in 80 ml lysis buffer (20 mM TrisHCl, pH 8, 150 mM NaCl, 0.5 mM EDTA, 2 mM □-mercaptoethanol, 10% [v/v] glycerol, 10 mM imidazole, 10 µl benzonase (Roche) and 1 mM Complete™ protease inhibitor mixture (Roche Applied Science) and disrupted in a French pressure cell. After centrifugation at 30 000 *g* for 1 h, the soluble extract was passed through a 0.45-µm pore-size filter and loaded onto a 5-ml Ni-agarose column (Amersham Biosciences) pre-equilibrated with buffer A-RpoS (20 mM Tris HCl pH 8, 150 mM NaCl, 2 mM □–mercaptoethanol, 10% glycerol and 10 mM imidazole). The column was washed with buffer A until nonspecifically bound material had been removed. RpoS eluted from the column at about 350 mM imidazole in a 30-ml imidazole gradient in buffer B-RpoS (buffer A-RpoS + 1 M imidazole). Eluted fractions containing RpoS protein were dialyzed against storage buffer (20 mM Tris HCl, pH 8, 150 mM NaCl, 10% glycerol, 1 mM DTT) and stored at -70ºC until use.

*Single-round* in vitro *transcription assays*

Single-round *in vitro* transcription assays were performed on derivatives of the supercoiled pSR plasmid carrying the □oop terminator (Kolb *et al.*, 1995) and a replication repressor RNA I transcript used as an internal control to aid quantification of XylS-dependent transcripts. The 148-bp *Eco*RI-*Hin*dIII fragment containing the Pm promoter region (positions -113 to +35) was cloned into the pSR plasmid. Active holoenzymes were reconstituted by incubation of core RNA polymerase (Epicentre) with $\sigma^{32}$ or $\sigma^{38}$ sigma factors (1:5 ratio) in transcription buffer S (40 mM Hepes pH 8, 100 mM potassium glutamate, 10 mM $MgCl_2$, 2 mM DTT, 100 µg/ml BSA) for 15 min at 30°C. Plasmids (5 nM) and reconstituted RNA polymerase holoenzymes (50 nM) were incubated for 10 min at 30°C in transcription buffer S (supercoiled) to allow RNA polymerase-promoter complexes to form. Elongation was started by the addition of a prewarmed mixture containing nucleotides and heparin (final concentrations were 250 µM GTP, and CTP; 50 µM UTP; 2 µCi [$\alpha^{32}$ P]UTP at 3000 Ci/mmol; and 100 µg/ml heparin) to the template-polymerase mix, and allowed to proceed for 10 min at 30°C. Reactions were stopped by the addition of 10 µl loading buffer (formamide containing 20 mM EDTA, xylene cyanol, and bromophenol blue). Samples were electrophoresed in a 5.5% (w/v) polyacrylamide denaturing sequencing gel. The results were analyzed with Molecular Imager FX equipment (Bio-Rad, Madrid, Spain).

For the linear Pm template, a 310-bp DNA fragment (positions -114 to +195 of the Pm promoter) prepared by PCR with primers Pm300H (5'-GCCAAGCTTGGGCGAG ATAAATCCAGTTGC-3') and Pm141E (5'-GGAATTCG GCTGCAGTGTCCGGTTTG-3') was used. Active holoenzymes were reconstituted by incubation of core RNA polymerase (Epicentre) with $\sigma^{32}$ or $\sigma^{38}$ sigma factors (1:5 ratio) in transcription buffer L (50 mM Tris HCl pH 8, 200 mM potassium glutamate, 10 mM magnesium-acetate, 2 mM DTT, 100 µg/ml BSA) for 15 min at 30°C. Linear DNA template (5 nM) and reconstituted RNA polymerase holoenzymes (50 nM) were incubated for 10 min at 30°C in transcription buffer L (linear) to allow RNA polymerase-promoter complexes to form. Elongation was started by the addition of a prewarmed mixture containing nucleotides and heparin (final concentrations were 250 µM GTP and CTP; 50 µM UTP; 2 µCi [$\alpha^{32}$P]UTP at 3000 Ci/mmol; and 100 µg/ml heparin) to the template-polymerase mix, and allowed to proceed for 10 min at 30°C. Reactions were stopped by the addition of 10 µl of loading buffer (formamide containing 20 mM EDTA, xylene cyanol, and bromophenol blue). Samples were electrophoresed in a 5.5% (w/v) polyacrylamide denaturing sequencing gel. The results were analyzed with Molecular Imager FX equipment (Bio-Rad, Madrid, Spain).

In vitro *$KMnO_4$ footprinting experiments*

Potassium permanganate footprinting is based on the hyper-reactivity of single-stranded thymines to this compound, a reaction that makes it possible to probe open complex formation by RNA polymerase (Sasse-Dwight and Gralla, 1989). A labeled 158-bp DNA fragment containing the Pm promoter (positions -113 to +35) was generated by PCR using pJLR100 as template with the primers Pm141E (5'GGAATTCGGCTGCAG TGTCCGGTTTG-3') and Pm141H (5'CCCAAGCTTGT CATGGTCATGACTCC-3') as described (Higushi, 1990). Reconstituted RNA polymerases (50 nM, RNA polymerase core from Epicentre) and the Pm promoter fragment were incubated in transcription buffer (40 mM Hepes pH 8, 100 mM potassium glutamate, 10 mM $MgCl_2$, 2 mM DTT, 100 µg/ml BSA) for 15 min at 37ºC in a final volume of 20 µl; incubations also contained 500 nM XylS C-terminal domain (XylS-C, see above) when indicated. Freshly prepared $KMnO_4$ was added to a final concentration of 10 mM, and the reaction was stopped after 30 s by adding 50 µl of a solution containing 1.5 M □–mercaptoethanol, ammonium acetate and 0.1 mM EDTA. The samples were phenol-extracted, glycogen was added to a final concentration of 0.1 mg/ml, and samples were precipitated with 100% ethanol, washed with 70%

ethanol and resuspended in 40 µl piperidine 1 M. After 30 min at 90ºC, 80 µl of a solution containing 0.3 M sodium acetate pH 5.2 and 250 µg/ml glycogen was added, samples were ethanol-precipitated, washed with 70% ethanol and resuspended in 8 µl loading buffer. Urea-polyacrylamide sequencing gels were calibrated with Maxam-Gilbert 'G+A' sequencing reactions of the labeled fragment and quantified with a Bio-Rad Molecular Imager FX and Quantity One software.

*Chromatin immunoprecipitation (ChIP)*

*Escherichia coli* CC118□Pm::*lacZ* strain carrying a Pm::*lacZ* fusion with or without a plasmid bearing the wild-type XylS (pERD103) was grown overnight in LB medium with or without kanamycin (25 µg/ml) respectively at 30ºC and 200 rpm. Flasks containing 25 ml fresh medium were inoculated with an aliquot fraction of these precultures, and were then incubated at 30ºC. When $OD_{660}$ reached 0.3-0.5, 1 mM 3MB was added and incubation was continued for 20 min in the presence or absence of 3MB. Cells were treated *in vivo* with formaldehyde cross-linking agent (1% final concentration). After 20 min, cross-linking was quenched by the addition of glycine (0.5 M final concentration) and DNA was extracted from lysed cells and sheared by sonication as described by Grainger *et al*. (2004). Monoclonal mouse antibodies against the □ subunit of RNA polymerase were obtained from Neoclone (Madison, WI), and rabbit polyclonal anti-XylS was produced by Eurogentec (Belgium, EGT group). Immunoprecipitations using antibodies against XylS or subunits of RNA polymerase were carried out as previously described (Grainger *et al.*, 2004a), except that antibody-nucleoprotein incubations were done at 4°C overnight and all subsequent washing steps were performed at 4°C. Before analysis, DNA was purified from the immunoprecipitate with a PCR purification kit (QIAGEN) and resuspended in 200 µl water. After purification, real-time PCR was used to analyze immunoprecipitated DNA; PCR primers *mtlA1* and *mtlA2* for mannitol permease were used as the reference for quantitation; PCR primers Pm3 and Pm4 were used to amplify the Pm promoter. Real-time PCR was performed on an iCycler iQ detection system according to the manufacturer's instructions. The PCR reactions (10 µl) were set up with the following reagents: 5 µl 2□ SYBR Green Supermix (Bio-Rad), 2.5 µl immunoprecipitated DNA samples, and 1 µM of each oligonucleotide primer. The thermal cycling conditions used were: 10 min at 95ºC followed by 35 cycles of 95ºC for 30 s, 52ºC for 30 s, and 72ºC for 30 s. A final melt curve was plotted to check the specific amplification of both probes. All reactions were run in triplicate.

**Acknowledgments**

This work was supported by MEC grants BCM2001-0515 and BIO2006-05668 and SYSMO EU grant GEN 2006-27750-C5-J-E/SYS. P. Domínguez-Cuevas is the recip-ient of a Junta de Andalucía Predoctoral Fellowship in Spain and an EMBO Short Term Fellowship (work at SB's laboratory in Birmingham, UK). The authors wish to thank David Grainger for help with the ChIP experiments and M. Trinidad Gallegos for her thorough reading of the manuscript and constructive suggestions.

**References**

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (2007). *Current Protocols in Molecular Biology,* Wiley, New York.

Bhende, P.M., and Egan, S.M. (1999) Amino acid-DNA contacts by RhaS: an AraC family transcription activator. *J Bacteriol* **181:** 5185-5192.

Bhende, P.M., and Egan, S.M. (2000) Genetic evidence that transcription activation by RhaS involves specific amino acid contacts with sigma 70. *J Bacteriol* **182:** 4959-4969.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:** 248-254.

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

Bustos, S.A., and Schleif, R.F. (1993) Functional domains of the AraC protein. *Proc Natl Acad Sci USA* **90:** 5638-5642.

Caswell, R., Webster, C., and Busby, S. (1992) Studies on the binding of the *Escherichia coli* MelR transcription activator protein to operator sequences at the MelAB promoter. *Biochem J* **287:** 501-508.

Domínguez-Cuevas, P., Marín, P., Ramos, J.L., and Marqués, S. (2005) RNA polymerase holoenzymes can share a single transcription start site for the Pm promoter. Critical nucleotides in the -7 to -18 region are needed to select between RNA polymerase with sigma38 or sigma32. *J Biol Chem* **280:** 41315-41323.

Domínguez-Cuevas, P., González-Pastor, J.E., Marqués, S., Ramos, J.L., and de Lorenzo, V. (2006) Transcriptional tradeoff between metabolic and stress-response programs in *Pseudomonas putida* KT2440 cells exposed to toluene. *J Biol Chem* **281:** 11981-11991.

Domínguez-Cuevas, P., Ramos, J. L., and Marqués, S. (2007) Sequential and cooperative XylS-C domain binding to the Pm promoter provokes DNA bending essential for activation. In *J Biol Chem, submitted*.

Fernández, S., de Lorenzo, V., and Pérez-Martín, J. (1995) Activation of the transcriptional regulator XylR of *Pseudomonas putida* by release of repression between functional domains. *Mol Microbiol* **16:** 205-213.

Gallegos, M.T., Marqués, S., and Ramos, J.L. (1996a) Expression of the TOL plasmid *xylS* gene in *Pseudomonas putida* occurs from a sigma 70-dependent promoter or from sigma 70- and sigma 54-dependent tandem promoters according to the compound used for growth. *J Bacteriol* **178:** 2356-2361.

Gallegos, M.T., Marqués, S., and Ramos, J.L. (1996b) The $TACAN_4TGCA$ motif upstream from the -35 region in the sigma70-sigmaS-dependent Pm promoter of the TOL plasmid is the minimum DNA segment required for transcription stimulation by XylS regulators. *J Bacteriol* **178:** 6427-6434.

Gallegos, M.T., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J.L. (1997) AraC/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* **61:** 393-410.

González-Pérez, M.M., Ramos, J.L., Gallegos, M.T., and Marqués, S. (1999) Critical nucleotides in the upstream region of the XylS-dependent TOL meta-cleavage pathway operon promoter as deduced from analysis of mutants. *J Biol Chem* **274:** 2286-2290.

González-Pérez, M.M., Marqués, S., Domínguez-Cuevas, P., and Ramos, J.L. (2002) XylS activator and RNA polymerase binding sites at the Pm promoter overlap. *FEBS Lett* **519:** 117-122.

González-Pérez, M.M., Ramos, J.L., and Marqués, S. (2004) Cellular XylS levels are a function of transcription of *xylS* from two independent promoters and the differential efficiency of translation of the two mRNAs. *J Bacteriol* **186:** 1898-1901.

Grainger, D.C., Overton, T.W., Reppas, N., Wade, J.T., Tamai, E., Hobman, J.L., Constantinidou, C., Struhl, K., Church, G., and Busby, S.J. (2004a) Genomic studies with *Escherichia coli* MelR protein: applications of chromatin immunoprecipitation and microarrays. *J Bacteriol* **186:** 6938-6943.

Grainger, D.C., Webster, C.L., Belyaeva, T.A., Hyde, E.I., and Busby, S.J. (2004b) Transcription activation at the *Escherichia coli melAB* promoter: interactions of MelR with its DNA target site and with domain 4 of the RNA polymerase sigma subunit. *Mol Microbiol* **51:** 1297-1309.

Griffith, K.L., Shah, I.M., and Wolf, R.E., Jr. (2004) Proteolytic degradation of *Escherichia coli* transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol Microbiol* **51:** 1801-1816.

Griffith, K.L., and Wolf, R.E., Jr. (2004) Genetic evidence for pre-recruitment as the mechanism of transcription activation by SoxS of *Escherichia coli*: the dominance of DNA binding mutations of SoxS. *J Mol Biol* **344:** 1-10.

Hachler, H., Cohen, S.P., and Levy, S.B. (1996) Untranslated sequence upstream of MarA in the multiple antibiotic resistance locus of *Escherichia coli* is related to the effector-binding domain of the XylS transcriptional activator. *J Mol Evol* **42:** 409-413.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166:** 557-580.

Hansen, L.H., Sorensen, S.J., and Jensen, L.B. (1997) Chromosomal insertion of the entire *Escherichia coli* lactose operon, into two strains of *Pseudomonas*, using a modified mini-Tn*5* delivery system. *Gene* **186:** 167-173.

Higushi, R. (1990) PCR protocols. A guide to methods and applications. In *Recombinant PCR*. Academic Press, San Diego, pp. 177-183.

Holcroft, C.C., and Egan, S.M. (2000) Interdependence of activation at *rhaSR* by cyclic AMP receptor protein, the RNA polymerase alpha subunit C-terminal domain, and RhaR. *J Bacteriol* **182:** 6774-6782.

Jair, K.W., Martin, R.G., Rosner, J.L., Fujita, N., Ishihama, A., and Wolf, R.E., Jr. (1995) Purification and regulatory properties of MarA protein, a transcriptional activator of *Escherichia coli* multiple antibiotic and superoxide resistance promoters. *J Bacteriol* **177:** 7100-7104.

Jair, K.W., Yu, X., Skarstad, K., Thony, B., Fujita, N., Ishihama, A., and Wolf, R.E., Jr. (1996) Transcriptional activation of promoters of the superoxide and multiple antibiotic resistance regulons by Rob, a binding protein of the *Escherichia coli* origin of chromosomal replication. *J Bacteriol* **178:** 2507-2513.

Kahramanoglou, C., Webster, C.L., El-Robh, M.S., Belyaeva, T.A., and Busby, S.J. (2006) Mutational analysis of the *Escherichia coli melR* gene suggests a two-state concerted model to explain transcriptional activation and repression in the melibiose operon. *J Bacteriol* **188:** 3199-3207.

Kaldalu, N., Mandel, T., and Ustav, M. (1996) TOL plasmid transcription factor XylS binds specifically to the Pm operator sequence. *Mol Microbiol* 20: 569-579.

Kaldalu, N., Toots, U., de Lorenzo, V., and Ustav, M. (2000) Functional domains of the TOL plasmid transcription factor XylS. *J Bacteriol* **182:** 1118-1126.

Kolb, A., Kotlarz, D., Kusano, S., and Ishihama, A. (1995) Selectivity of the *Escherichia coli* RNA polymerase E sigma 38 for overlapping promoters and ability to support CRP activation. *Nucleic Acids Res* **23:** 819-826.

Kwon, H.J., Bennik, M.H., Demple, B., and Ellenberger, T. (2000) Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nat Struct Biol* **7:** 424-430.

Landini, P., Gaal, T., Ross, W., and Volkert, M.R. (1997) The RNA polymerase alpha subunit carboxyl-terminal domain is required for both basal and activated transcription from the *alkA* promoter. *J Biol Chem* **272:** 15914-15919.

Landini, P., and Busby, S.J. (1999) The *Escherichia coli* Ada protein can interact with two distinct determinants in the sigma70 subunit of RNA polymerase according to promoter architecture: identification of the target of Ada activation at the *alkA* promoter. *J Bacteriol* **181:** 1524-1529.

Lange, R., and Hengge-Aronis, R. (1991) Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. *Mol Microbiol* **5:** 49-59.

Lobell, R.B., and Schleif, R.F. (1990) DNA looping and unlooping by AraC protein. *Science* 250: 528-532.

Maeda, H., Fujita, N., and Ishihama, A. (2000) Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res* **28:** 3497-3503.

Manoil, C., and Beckwith, J. (1985) Tn*phoA*: a transposon probe for protein export signals. *Proc Natl Acad Sci USA* **82:** 8129-8133.

Manzanera, M., Marqués, S., and Ramos, J.L. (2000) Mutational analysis of the highly conserved C-terminal residues of the XylS protein, a member of the AraC

family of transcriptional regulators. *FEBS Lett* **476:** 312-317.

Marqués, S., and Ramos, J.L. (1993) Transcriptional control of the *Pseudomonas putida* TOL plasmid catabolic pathways. *Mol Microbiol* **9:** 923-929.

Marqués, S., Manzanera, M., González-Pérez, M.M., Gallegos, M.T., and Ramos, J.L. (1999) The XylS-dependent Pm promoter is transcribed *in vivo* by RNA polymerase with sigma 32 or sigma 38 depending on the growth phase. *Mol Microbiol* **31:** 1105-1113.

Martin, R.G., and Rosner, J.L. (2001) The AraC transcriptional activators. *Curr Opin Microbiol* **4:** 132-137.

Mermod, N., Ramos, J.L., Bairoch, A., and Timmis, K.N. (1987) The xylS gene positive regulator of TOL plasmid pWWO: identification, sequence analysis and overproduction leading to constitutive expression of meta cleavage operon. *Mol Gen Genet* **207:** 349-354.

Michán, C., Kessler, B., de Lorenzo, V., Timmis, K.N., and Ramos, J.L. (1992a) XylS domain interactions can be deduced from intraallelic dominance in double mutants of *Pseudomonas putida*. *Mol Gen Genet* **235:** 406-412.

Michán, C., Zhou, L., Gallegos, M.T., Timmis, K.N., and Ramos, J.L. (1992b) Identification of critical amino-terminal regions of XylS. The positive regulator encoded by the TOL plasmid. *J Biol Chem* **267:** 22897-22901.

Michán, C.M., Busby, S.J., and Hyde, E.I. (1995) The *Escherichia coli* MelR transcription activator: production of a stable fragment containing the DNA-binding domain. *Nucleic Acids Res* **23:** 1518-1523.

Munson, G.P., and Scott, J.R. (1999) Binding site recognition by Rns, a virulence regulator in the AraC family. *J Bacteriol* **181:** 2110-2117.

Ng, L.C., O'Neill, E., and Shingler, V. (1996) Genetic evidence for interdomain regulation of the phenol-responsive final sigma54-dependent activator DmpR. *J Biol Chem* **271:** 17281-17286.

Poore, C.A., Coker, C., Dattelbaum, J.D., and Mobley, H.L. (2001) Identification of the domains of UreR, an AraC-like transcriptional regulator of the urease gene cluster in *Proteus mirabilis*. *J Bacteriol* **183:** 4526-4535.

Prouty, M.G., Osorio, C.R., and Klose, K.E. (2005) Characterization of functional domains of the *Vibrio cholerae* virulence regulator ToxT. *Mol Microbiol* **58:** 1143-1156.

Raibaud, O., and Schwartz, M. (1984) Positive control of transcription initiation in bacteria. *Annu Rev Genet* **18:** 173-206.

Ramos, J.L., Stolz, A., Reineke, W., and Timmis, K.N. (1986) Altered effector specificities in regulators of gene expression: TOL plasmid *xylS* mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. *Proc Natl Acad Sci USA* **83:** 8467-8471.

Ramos, J.L., Mermod, N., and Timmis, K.N. (1987) Regulatory circuits controlling transcription of TOL plasmid operon encoding meta-cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. *Mol Microbiol* **1:** 293-300.

Ramos, J.L., Rojo, F., Zhou, L., and Timmis, K.N. (1990) A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and the *Escherichia coli* AraC activators. *Nucleic Acids Res* **18:** 2149-2152.

Rhee, S., Martin, R.G., Rosner, J.L., and Davies, D.R. (1998) A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc Natl Acad Sci USA* **95:** 10413-10418.

Ruíz, R., Ramos, J.L., and Egan, S. (2001) Interactions of the XylS regulators with the C-terminal domain of the RNA polymerase alpha subunit influence the expression level from the cognate Pm promoter. *FEBS Lett* **491:** 207-211.

Ruíz, R., and Ramos, J.L. (2002) Residues 137 and 153 at the N terminus of the XylS protein influence the effector profile of this transcriptional regulator and the sigma factor used by RNA polymerase to stimulate transcription from its cognate promoter. *J Biol Chem* **277:** 7282-7286.

Ruíz, R., Marqués, S., and Ramos, J.L. (2003) Leucines 193 and 194 at the N-terminal domain of the XylS protein, the positive transcriptional regulator of the TOL *meta*-cleavage pathway, are involved in dimerization. *J Bacteriol* **185:** 3036-3041.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual.* Cold spring Harbor, N.Y., Cold spring Harbor Laboratory.

Sasse-Dwight, S., and Gralla, J.D. (1989) $KMnO_4$ as a probe for *lac* promoter DNA melting and mechanism *in vivo*. *J Biol Chem* **264:** 8074-8081.

Shah, I.M., and Wolf, R.E., Jr. (2004) Novel protein--protein interaction between *Escherichia coli* SoxS and the DNA binding determinant of the RNA polymerase alpha subunit: SoxS functions as a co-sigma factor and redeploys RNA polymerase from UP-element-containing promoters to SoxS-dependent promoters during oxidative stress. *J Mol Biol* **343:** 513-532.

Soisson, S.M., MacDougall-Shackleton, B., Schleif, R., and Wolberger, C. (1997a) The 1.6 Å crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose. *J Mol Biol* **273:** 226-237.

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

Tanaka, K., Takayanagi, Y., Fujita, N., Ishihama, A., and Takahashi, H. (1993) Heterogeneity of the principal sigma factor in *Escherichia coli*: the *rpoS* gene product, sigma 38, is a second principal sigma factor of RNA polymerase in stationary-phase *Escherichia coli*. *Proc Natl Acad Sci USA* **90:** 8303.

Terán, W., Felipe, A., Segura, A., Rojas, A., Ramos, J.L., and Gallegos, M.T. (2003) Antibiotic- dependent induction of *Pseudomonas putida* DOT-T1E TtgABC efflux pump is mediated by the drug binding repressor TtgR. *Antimicrob Agents Chemother* **47:** 3067-3072.

Timmes, A., Rodgers, M., and Schleif, R. (2004) Biochemical and physiological properties of the DNA binding domain of AraC protein. *J Mol Biol* 340: 731-738.

Tobes, R., and Ramos, J.L. (2002) AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucleic Acids Res* **30:** 318-321.

Van Dyk, T.K., Reed, T.R., Vollmer, A.C., and LaRossa, R.A. (1995) Synergistic induction of the heat shock

response in *Escherichia coli* by simultaneous treatment with chemical inducers. *J Bacteriol* **177:** 6001-6004.

Wade, J.T., Belyaeva, T.A., Hyde, E.I., and Busby, S.J. (2000) Repression of the *Escherichia coli melR* promoter by MelR: evidence that efficient repression requires the formation of a repression loop. *Mol Microbiol* **36:** 223-229.

Weldon, J.E., and Schleif, R.F. (2006) Specific interactions by the N-terminal arm inhibit self-association of the AraC dimerization domain. *Protein Sci* **15:** 2828-2835.

Weldon, J.E., Rodgers, M.E., Larkin, C., and Schleif, R.F. (2007) Structure and properties of a truly apo form of AraC dimerization domain. *Proteins* **66:** 646-654.

Wickstrum, J.R., and Egan, S.M. (2004) Amino acid contacts between sigma 70 domain 4 and the transcription activators RhaS and RhaR. *J Bacteriol* **186:** 6277-6285.

Zhang, X., Reeder, T., and Schleif, R. (1996) Transcription activation parameters at *ara* $p_{BAD}$. *J Mol Biol* **258:** 14-24.

Zimmerman, S.B., and Trach, S.O. (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J Mol Biol* **222:** 599-620.

CAPÍTULO 3

# Interacciones XylS-Pm mediadas por dos HTHs: Identificación de los residuos de XylS importantes para la unión al ADN y la activación.

**Patricia Domínguez Cuevas, Patricia Marín, Silvia Marqués y Juan L. Ramos**

El activador transcripcional de la ruta *meta* del plásmido TOL para la degradación de alquilbenzoatos, XylS, pertenece a la familia de reguladores transcripcionales de AraC y consta de un dominio N-terminal encargado del reconocimiento del efector, y de un dominio C-terminal compuesto de siete α-hélices que se pliegan formando dos motivos hélice-giro-hélice (HTH) de unión al ADN. Las α-hélices 3 y 6 constituyen las hélices de reconocimiento de ambos motivos HTH. De acuerdo con la organización estructural de XylS, Pm presenta un motivo de unión bipartito que consta de dos cajas de secuencia conservada llamadas A y B, cuyas secuencias son respectivamente TGCA y GGNTA. Este motivo de unión bipartito está repetido en Pm, de manera que se une un monómero de XylS por motivo. El análisis exhaustivo de las interacciones entre XylS y Pm mediante experimentos de epistasis genética en los que combinamos mutantes del promotor y alelos de XylS con mutaciones puntuales, reveló que los residuos de la α-hélice 3 establecían contactos con los nucleótidos de las cajas A, mientras que la α-hélice 6 contacta con las cajas B. Entre los aminoácidos de la α-hélice 3, la Asn246 y la Arg242 están implicadas en el contacto específico con el dinucleótido TG de la caja A. La Arg296 y el Glu299 contactan con los nucleótidos G y T de la caja B, respectivamente. Basándonos en nuestros resultados y en el modelo tridimensional del dominio C-terminal de XylS, proponemos que los residuos Ser243, Glu249, Lys250 en α3 y Asn299 y Arg302 en α6 interaccionan con el esqueleto de fosfato del ADN. La substitución de los residuos predichos de contacto con el esqueleto de fosfato a alanina rindió alelos de XylS con actividad transcripcional muy baja, lo que sugiere que la especificidad de unión entre XylS y Pm no sólo viene determinada por interacciones específicas entre bases y aminoácidos, sino por una complementariedad en la forma que adopta el ADN alrededor de la proteína que aumenta el número e importancia de los contactos con el esqueleto de fosfato. Finalmente proponemos un modelo en el que un dímero de XylS reconoce las repeticiones directas en Pm orientándose en una conformación cabeza-cola que determina la región del monómero proximal de XylS implicada probablemente en la interacción con la subunidad sigma de la ARN polimerasa.

Enviado a *Journal of Molecular Biology*.

*J. Mol. Biol. (2007), submitted*

JMB

ELSEVIER

# XylS-Pm Promoter Interactions through two Helix-Turn-Helix Motifs: Identifying XylS Residues Important for DNA Binding and Activation

**Patricia Domínguez-Cuevas, Patricia Marín, Silvia Marqués and Juan Luis Ramos***

*Department of Environmental Protection, Estación Experimental del Zaidín, C.S.I.C., Granada, Spain*

The XylS protein is the positive transcription regulator of the TOL plasmid *meta*-cleavage pathway operon. XylS belongs to the AraC family of transcriptional regulators and exhibits an N-terminal domain involved in effector recognition, and a C-terminal domain made up of seven α-helices conforming two helix-turn-helix DNA binding domains. α-Helix 3 and α-helix 6 are the recognition helices. In consonance with XylS structural organization, Pm exhibits a bipartite DNA binding motif consisting of two boxes, named A and B, whose sequences are TGCA and GGNTA, respectively. This bipartite motif is repeated at the Pm promoter so that one XylS monomer binds to each of the repeats. An extensive series of genetic epistasis assays combining mutant Pm promoters and XylS single-point mutants revealed that α-helix 3 contacts A box nucleotides, whereas α-helix 6 residues contact B box nucleotides. In α-helix 3, Asn246 and Arg242 are involved in specific contacts with the TG dinucleotide at box A, whereas Arg296 and Glu299 contact the second G and T nucleotides at box B. On the basis of our results and of the three-dimensional model of XylS C-terminal domain we propose that Ser243, Glu249 and Lys250 in α-helix 3, and Asn299 and Arg302 in α-helix 6 contact the phosphate backbones. Alanine substitutions at predicted phosphate backbone contacting residues yielded mutants with very low levels of activity, suggesting that XylS-Pm binding specificity does not only involve specific amino acid-base interactions, but also relies on "shape complementarity" determined by broad interactions between α-helices and DNA phosphate backbone. We propose a model in which a XylS dimer binds the direct repeats in Pm in a head-to-tail conformation that allows the direct interaction of the XylS proximal subunit with the RNA polymerase sigma factor.

*Keywords:* XylS, binding orientation, DNA contacts, AraC family, transcription activation.

**Corresponding author*

## Introduction

The AraC family of transcriptional regulators includes a large number of activators that regulate cellular processes involved in carbon metabolism, virulence and stress responses[1]. Members of this family are characterized by significant homology over a ~100-amino acid stretch containing two helix-turn-helix (HTH) motifs (Figure 1A) that recognize 15-20-bp operator sequences at target promoters[2-5]. Many members of the AraC family also contain a non-conserved domain which is involved in effector binding and dimerization processes[1,6-10].

Abbreviations used: bp, base pair(s); HTH, helix-turn-helix; 3MB, 3-methylbenzoate; EMSA, electroforetic mobility shift assays.

E-mail address of the corresponding author: jlramos@eez.csic.es

The *Pseudomonas putida* XylS protein belongs to the AraC family of transcriptional regulators[1,11]. In response to the presence of benzoates, XylS activates expression from the Pm promoter that controls expression of the *meta*-cleavage pathway for the oxidative catabolism of benzoates and toluates present in *Pseudomonas putida* TOL plasmid pWW0[12]. XylS is a 321-aminoacid protein with a molecular mass of 36 kDa. Like most AraC family members, XylS appears to consist of two domains: the N-terminal domain is involved in effector recognition and protein dimerization[8,9,13], whereas the C-terminal domain seems to be responsible for DNA binding[14]. The XylS oligomeric state in solution has not been determined directly since this protein, like many AraC family members, is difficult to study *in vitro* due to its intrinsic insolubility[15-18]. However, Ruíz and coworkers found that the XylS N-terminal domain was able to dimerize[8], suggesting that XylS, like most AraC family proteins, can form dimers in solution[6,19,20].

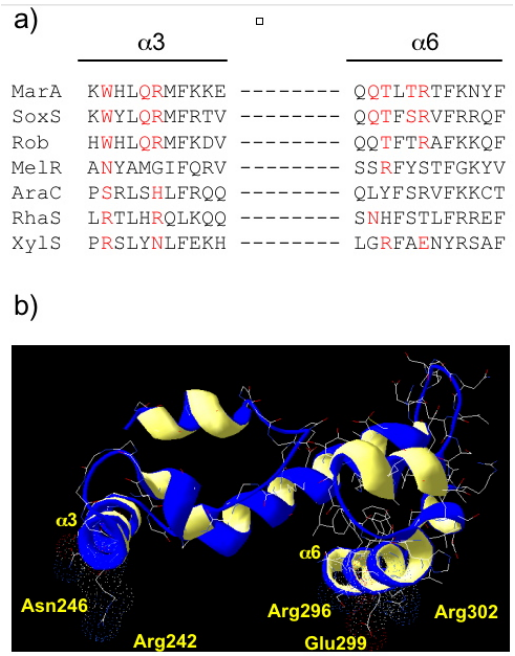



**Figure 1.** A. Partial alignment of members of the AraC family covering the recognition helices α3 and α6. Bases in red represent positions of base-specific contacts. B. Structure-based model of the XylS DNA binding domain. The XylS C-terminal domain was modeled on the structure of the MarA protein. Residues Arg242 and Asn246 in α3 and residues Arg296 and Glu299 in α6 are highlighted in yellow. We propose that these residues establish base-specific contacts with XylS binding sites at the Pm promoter.

The XylS C-terminal domain is the most conserved part in the protein family, and is predicted to adopt the same tertiary structure as MarA; seven α-helices which fold to form two HTH DNA-binding motifs (Figure 1B)[1,21]. The XylS C-terminal domain is able to bind and activate transcription from the Pm promoter independently of the presence of 3-methylbenzoate (3MB)[22,23]. Genetic analyses established that XylS recognizes two 15-bp direct repeats (TGCA-$N_6$-GGNTA) extending from –69 to –55 and from –48 to –34. Furthermore, XylS binding site architecture was defined as two repeats each consisting of a 5′-box A (TGCA) and a 3′-box B (GGNTA). The arrangement of the two repeats is such that the proximal XylS binding site overlaps by 1 bp the RNA polymerase binding site at –35[24]. Despite the fact that AraC family members share great homology at the DNA binding domain, the organization of target DNA binding sites at the different promoters and the specific activation mechanism vary from one promoter to the other. Hence, for any regulator, it is important to define the specific contacts established between the protein and the DNA, and the relative orientation of protein subunits. In class II promoters the binding orientation determines the protein region exposed to interact with the RNA polymerase sigma subunit, which binds at the –35 promoter element[25].

In this study we used genetic analysis to identify XylS residues important for DNA recognition and binding, focusing on the orientation of XylS binding at its tandem sites. We used genetic epistasis experiments to identify XylS amino acids that make base-specific contacts at Pm. Several XylS mutants were tested *in vitro* for their ability to bind Pm promoter DNA. Our results suggest that XylS binds Pm direct repeats in a head-to-tail organization, with α-helix 3 in the first HTH motif recognizing the A boxes, and α-helix 6 in the second HTH motif interacting with the B boxes. Both HTH motifs contribute to DNA binding, and are critical to establish contacts with target DNA sequences at Pm.

## Results

### Alanine-scanning mutagenesis of the XylS HTH binding domains

MarA was the first AraC family member whose high-resolution structure was determined. It is composed of seven α-helices which fold in two HTH subdomains connected by a long α-helix[21]. The recognition helices (α3 and α6) of the two HTHs interact with two adjacent major grooves. The conserved DNA-binding domains of other AraC family members, which share more than 20% similarity, are expected to adopt the same tertiary structure[1,13].

Based on the multiple alignment of the AraC family proteins, residues Glu231 to His251 in XylS are predicted to form an HTH motif (corresponding to the α2-turn-α3 in MarA), and residues Ile281 to Phe305 in XylS are anticipated to form the second HTH corresponding to α5-turn-α6 in MarA (See Figure 1A for α3 and α6 sequences)[1,11]. When the XylS C-terminal domain was modeled on the structure of the MarA protein, the XylS recognition helices in the HTH motifs appeared protruding from the same face of the protein, suggesting that XylS contacts the same face of the DNA (Figure 1B).

On combining information from the planar projection of XylS recognition helices (constructed using the conventional values for helical wheel predictions and number of residues per turn [Figure 2A]) and the XylS structure model based on the MarA structure, we observed that the polar residues in α3 and α6 emerged from the external surface of the protein, so that their side chains faced the DNA they made contact with, whereas most of the nonpolar residues were positioned towards the inside of the protein (Figure 2A).

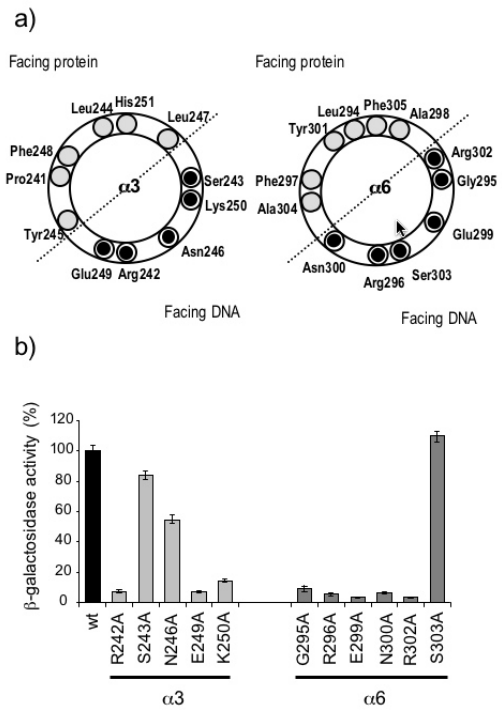



**Figure 2.** A. Planar projection of XylS recognition helices. All polar residues in α3 and α6 are positioned on the external surface of the protein. B. Percentage of wild-type activation of alanine-substituted XylS mutants in α3 and α6 measured as the level of β-galactosidase expression from the Pm promoter. *Escherichia coli* MC4100 (pJLR107) bearing a wild-type or mutant xylS gen in pLOW2 was cultured as indicated in Materials and Methods. The X-axis shows the amino acid substitution at the indicated XylS position within α-helix 3 or α-helix 6. 100% activity corresponded to 5500 Miller Units.

As a first step to identify XylS residues required for DNA binding and transcriptional activation at the Pm promoter, we performed an alanine scanning mutagenesis of all polar residues in recognition helices α3 and α6. The rationale of this approach is that since alanine is a strongly helix-favoring amino acid[26], alanine changes in surface-exposed residues will have little or no consequences on protein folding or stability. Thus, any effect detected on ala-mutant activity can be assigned to changes in the contacts established between the protein residue and its target DNA.

These ala-mutant proteins were tested *in vivo* for their ability to activate transcription from the Pm promoter (Figure 2B). We found that all alanine mutants in α6, except for Ser303Ala, had lost more than 95% of the wild-type activity, whereas the situation in α3 was slightly different. Mutants Lys250Ala, Arg242Ala and Glu249Ala retained less than 10% activity, whereas alanine changes in Ser243 and Asn246 produced mutants that retained more than 50% of the wild-type activity. It is worth noting that in the absence of 3MB, activity of the different mutants was negligible.

These results suggest that XylS is more sensitive to disruption of interactions involving α-helix 6 than those involving α-helix 3. To further explore this hypothesis we decided to generate a second series of single-point XylS mutants in the same residues by changing their chemical properties.

## Mutants with altered chemical properties of surface-exposed residues

Three types of mutations were generated: changes in amino acid charge (positive for negatively charged amino acids), changes in the length of the lateral chain (e.g., Ser for Thr) or changes in the length of the chain without changing the charge (e.g., Arg for Lys). Activation of Pm by mutants at the different positions in α3 and α6 was determined *in vivo* as above (Figure 3). Although certain conservative amino acid changes (such as Lys250Arg, Asn300Gln and Arg302Lys) were tolerated without significant loss of activity, the situation was different for the mutants Arg242Lys, Ser243Thr, Asn246Gln, Glu249Asp and Arg296Lys, which showed complete loss of activity with respect to the wild type. The Ser303Thr mutant was found to retain 15% of the activity. Surprisingly, when glutamic residues at positions 249 and 299 were replaced by lysine, the resulting mutants conserved a high level of activity. When these Glu residues were replaced by Gln, they maintained between 20% and 60% of the wild-type activity. We found that nonconservative changes such as Ser242Glu, Ser243Lys, Asn246Arg or Asn246Asp in α3 and Gly295Ser, Arg296Gln or Arg296Glu,

Asn300Asp, Arg302Asn or Arg302Glu and Ser303Asn in α6 rendered XylS mutants defective in activation.

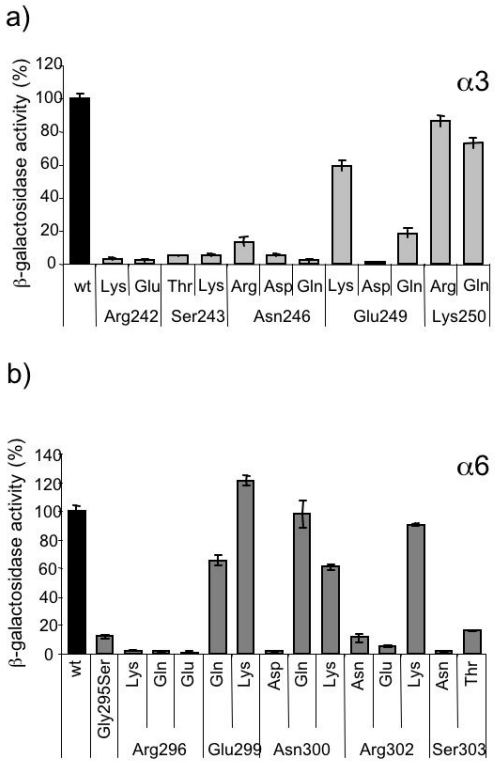



**Figure 3.** A. Level of β-galactosidase expression from the Pm promoter induced by XylS mutants with alterations in α-helix 3. B. Level of β-galactosidase expression from the Pm promoter induced by XylS mutants in α-helix 6. The X-axis shows the amino acid substitution at each XylS position within α-helix 3 or α-helix 6. Conditions are as in the legend for Figure 2B. Data are the average of at least four independent assays.

## Effects of surface-exposed amino acid substitutions on DNA binding

Failure of the above series of XylS* mutants to transcribe could have been due to the inability of the mutant protein to bind target DNA, or to nonproductive binding to target sequences from a transcriptional point of view. To distinguish between these two possibilities we decided to determine XylS* DNA binding capacity in gel shift assays. We expressed relevant mutant variants in *E. coli*, purified them as described under "Material and Methods" section, and used them in electrophoretic mobility gel shift assays (EMSA) with the wild-type Pm DNA fragment. As a positive control in these assays, we tested XylSLys250→Arg (mutant in α3, see Figure 4A) and XylSGlu299→Lys (mutant in α6) (Figure 4) regulators, which are able to activate expression from Pm. As expected, the XylSLyS250Arg and XylSGlu299Lys mutants bound and retarded the Pm promoter like the wild type (Figure 4). When we assayed mutants unable to stimulate transcription from Pm, we found two different patterns: most mutant proteins, e.g., Ser242→Ala, Glu249→Ala, Arg296→Ala, were not able to bind to target DNA, whereas some XylS mutants that were inactive in the transcription assays partially retained their binding capacity and retarded Pm, e.g., Glu249→Asp and Arg296→ Lys (Figure 4A and B).

a)

| | XylS Mutants | Activity | DNA Binding Ability |
|---|---|---|---|
| | XylS wt | ++ | +++ |
| α3 | Arg242Ala | - | - |
| | Asn246Ala | ++ | +++ |
| | Asn246Gln | - | - |
| | Glu249Ala | - | - |
| | Glu249Asp | - | - |
| | Glu249Gln | -/+ | + |
| | Lys250Arg | ++ | +++ |
| α6 | Arg296Ala | - | - |
| | Arg296Lys | - | -/+ |
| | Glu299Ala | - | - |
| | Glu299Lys | ++ | +++ |
| | Arg302Glu | - | - |
| | Arg302Asn | - | + |

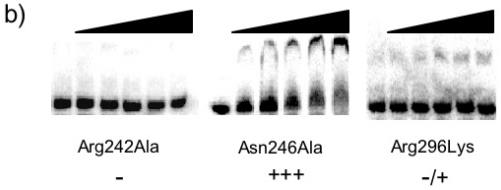



**Figure 4.** XylS mutant ability to bind DNA *in vitro* and activate Pm transcription *in vivo*. A) *In vivo* assay activities refer to β-galactosidase levels. ++ stands for β-galactosidase levels >2000 Miller's units, -/+ means β-galactosidase levels between 500 and 1000 Miller Units and - means no detectable induction. *In vitro* DNA binding ability refers to the ability of the protein to shift DNA in EMSA. +++, > 50% of the DNA was retarded; +, 30% of the DNA was retarded; -/+, a faint visible retarded band; - no retardation. B) EMSA assays for Arg242Ala, Asn246Aln and Arg296Lys are shown as example. The degree of retardation in each case is reflected in panel A.

## Specific contacts between recognition helices 3 and 6 and the XylS operator: orientation of XylS HTHs at their DNA binding domains

To test whether XylS orientation corresponded to a head-to-tail conformation, we used a classical genetic loss-of-contact approach[27,28]. The rationale of this approach was that a protein with a single amino acid substitution will lose the ability to discriminate

between the wild-type and mutant base pairs only at the position(s) in the DNA that are normally contacted by the substituted amino acid. As a result, a mutation in any DNA position normally contacted by the substituted amino acid would have relatively little effect on regulator activity, whereas mutations at other DNA positions would have large effects. This strategy makes it possible to determine which recognition helix binds to the A or B boxes at the DNA target. This analysis also determines the base pair positions contacted by the surface-exposed amino acids that we identified as important in DNA binding.

Activity of XylS mutants at positions important for DNA binding and displaying strongly decreased activity were tested with 15 Pm derivatives using Pm*:*lacZ* fusions (Figures 5 and 6). The Pm derivatives contained previously characterized single-point mutations in the distal XylS binding site (A or B boxes) and in the A box of the proximal XylS binding site[2] (Figures 5 and 6). Note that mutants in the proximal B box would overlap the -35 hexamer of Pm, making it impossible to determine whether the effect was due to the alteration of the XylS recognition site or to the inability of RNA polymerase to interact with the -35 promoter element.

Figure 5 shows that XylS mutants in the first HTH motif (α3) showed a considerable decrease in activation of Pm mutants at B box. In addition, Ser243Lys, Glu249Asp (data not shown) and Lys250Ala (Figure 5) all presented significantly lower activities with all mutant Pm promoters. This suggests that these amino acids do not establish base-specific contacts with the XylS binding sites. Point mutations at positions -69, -48 and -47 at boxes A had little effect on activation by Arg242Ala. These data suggest that Arg242 does contact these base pairs. In contrast, we found that Arg242Ala showed lower activity with mutants at positions -70, -67, -49 and -46, indicating that this amino acid does not make specific contacts at the T and A base pairs of A boxes. XylS mutant Asn246Gln had little activity with mutants at positions -69 and -48, whereas point mutations at -70 and -47 showed no significant decrease in activation by Asn246Gln. Our results indicate that Asn246 can make specific contacts with T and C bases at A boxes, whereas Asn246 does not seem to establish specific contacts with the G base at A boxes. Pm mutants at the proximal A motif showed a more marked effect than Pm mutants at the distal A motif with this regulator. These results are in agreement with the previous finding that the XylS proximal binding site is more important for Pm transcriptional activation than the distal site[2].

XylS mutants in the second HTH motif (α6) were also tested at Pm mutant promoters. Gly295Ala, Asn300Gln (not shown) and Arg302Ala (Figure 6) all presented significantly lower activities from all mutant Pm promoters, indicating that these residues did not establish specific contacts with any mutanted Pm position.

The Arg296Ala XylS mutant showed decreased transcriptional activation of Pm mutant promoters at positions -59, -57 and -56 (G-TA at the distal B box), whereas activation of a mutant promoter at position -60 (first G at the B boxes) was little affected. These results indicate that Arg296 makes specific contacts with the first guanine residue at B boxes. On the other hand,

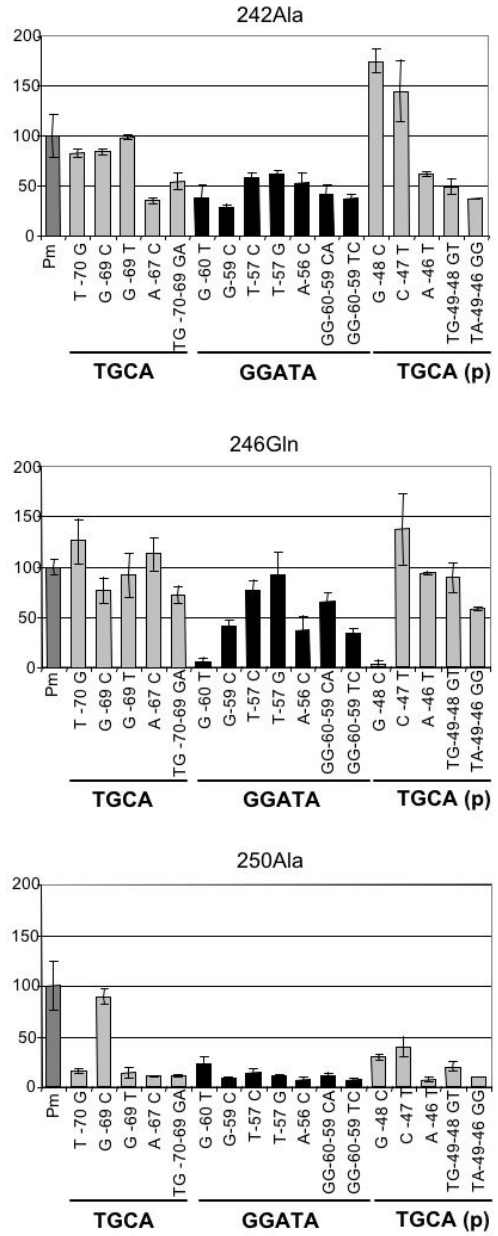



**Figure 5.** Activity of XylS single-point mutants in recognition helix 3 analyzed at the mutant Pm promoters. The X-axis represents the mutant Pm promoters assayed. The Y-axis represents the percentage of β-galactosidase activity for the corresponding XylS single-point mutant at each mutant promoter compared to activity at the wild-type promoter. Data are the average of at least four independent assays.

the XylSGlu299Ala mutant showed higher (T at B boxes) and a marked defect in mutants at positions -60, -59 and -56 ( GG - -A at B motifs). These data support the hypothesis that Glu299 establishes specific contacts with the T base at B boxes. XylS mutants in the second HTH motif (α6) showed activities below 100% in most Pm mutants in A motifs.

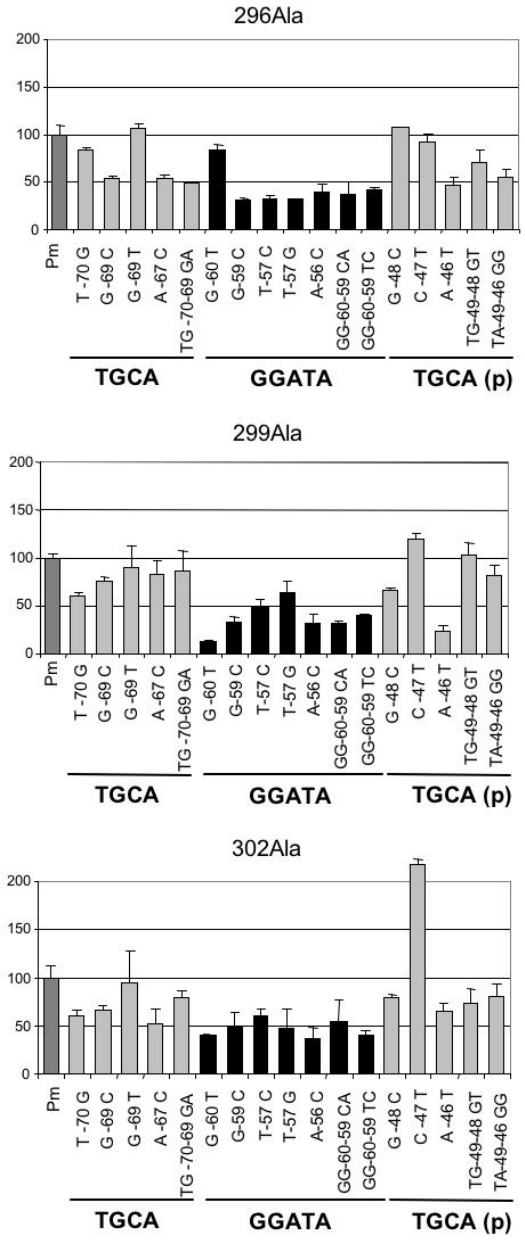



**Figure 6.** Activity of XylS single-point mutants in recognition helix 6 analyzed at the mutant Pm promoters. The X-axis represents the mutant Pm promoters assayed. The Y-axis represents the percentage of β-galactosidase activity for the corresponding XylS variant at each mutant promoter compared to activity at the wild-type promoter. Data are the average of at least four independent assays.

## Discussion

Alignment of almost 1000 members of the AraC family of transcriptional regulators reveals that they exhibit a significant conserved sequence of about 100 amino acid residues at their C-terminal domain, constituting the DNA-binding domain[11]. Most AraC family members also contain a non-conserved domain involved in effector binding and dimerization[8-10,13,14,23,29,30]. The three-dimensional structure of MarA and Rob, two AraC family members, in complex with DNA is available. On the basis of the high primary sequence homology shared by family members in the DNA binding domain[21,31], we constructed a structural model of the XylS C-terminal conserved domain. This model predicts that this domain is made up of seven α-helices folding in two HTHs motifs (Figure 1B). The helices predicted to interact with target DNA are α3 and α6, which protrude from the same face of the protein. Potential surface-exposed amino acids of importance for binding to DNA target sequences were identified based on the planar projection of these helices (Figure 2). To gain insights into the role of these amino acids, we performed a scanning mutagenesis of the two recognition helices. The results revealed the importance of the two helices in DNA binding, and showed that substitutions in the α6 recognition helix impaired transcription activation by XylS to a greater extent than did substitutions at the α3 recognition helix. These results are consistent with findings reported for mutants in other members of the family. For example, AraC mutants in its DNA binding motif failed to recognize the corresponding target sequences[5,32].

XylS binding sites in Pm were defined as two 15-bp direct repeats, each consisting of a 5′ box A (TGCA) and a 3′ box B (GGNTA), where base substitutions at either the A or B box rendered mutant promoters severely impaired in transcription activation[2]. Point mutation analysis of the XylS operator at Pm defined the TG bases as the most relevant within TGCA box A, and the GGNT sequence as critical within GGNTA box B[2].

In this study we carried out a series of genetic epistasis experiments that combined a number of XylS point mutant derivatives with an available library of Pm mutant promoters[2]. This series of assays allowed us to identify four XylS residues, Arg242 and Asn246 in α-helix 3 and Arg296 and Glu299 in α-helix 6, as involved in the specific recognition of A and B boxes, respectively. XylSArg242Ala and XylSAsn246Gln mutants showed a severe decrease in transcription activation of Pm promoters bearing mutations at the B box, whereas a moderate decrease in transcription activation of Pm mutants in the box A was observed with Arg296Ala and Glu299Ala XylS mutants. XylS mutation in α-helix 6 caused

a greater impairment in transcription activation than the mutation in α-helix 3, consistent with the greater sensitivity of the B box to substitutions. This finding is in agreement with the fact that, when α-helix 6 XylS mutants are combined with Pm mutants in the A boxes, transcription activation decreased only slightly, reflecting the greater importance of contacts between α-helix 6 and B boxes. In addition, specific interactions between Arg296 and Glu299 and bases at the B boxes were inferred from the suppression genetics experiments. This set of results supports that the first and second XylS HTH motifs interact with the A and B box, respectively, and allowed us to conclude that XylS dimers bind Pm direct repeats in a head-to-tail conformation.

logues, as was the B box sequence at the binding sites of their respective target promoters. However, deviation observed between XylS homologues in α3 residues located at base-specific contact positions are in accordance with the poor sequence similarity between the respective A boxes (Figure 7).

Alignment of the α-helix 3 and α-helix 6 of MarA, Rob and SoxS showed that the amino acid residues interacting with the sugar-phosphate backbone, as deduced from DNA/regulator co-crystals, are not conserved in this set of proteins (Gillete, 2000). It has been shown that the disruption of certain protein-DNA phosphate contacts in MarA and SoxS often affects DNA binding and transcription activation more severely than the substitution of base-specific contacts[33,35].

**Table 1.** Identified amino acid-DNA interactions within AraC family members[1].

| | Aa | C | Aa | C | Aa | C | Aa | C | Aa | C | Aa | C | Aa | C | Aa | C | Aa | C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MarA | Trp42 | C/T | Gln45 | T/G | Arg46 | G/C | Lys49 | P | Ser90 | - | Gln92 | T | Thr93 | C/T | Thr95 | T | Arg96 | G/A/T | Lys99 | P |
| SoxS | Trp36 | - | Gln39 | - | Arg40 | - | Arg43 | P | Ser84 | - | Gln86 | - | Thr87 | - | Ser89 | - | Arg90 | - | Arg93 | P |
| Rob | Trp46 | C | Gln39 | G | Arg40 | G | Lys43 | - | Ser84 | - | Gln86 | - | Thr87 | - | Thr89 | - | Arg90 | P | Lys93 | - |
| RhaS | Arg202 | G/C | His205 | - | Arg206 | G/C | Lys209 | - | Asp250 | C | Asn252 | G/A | His253 | - | Ser255 | - | Thr256 | - | Arg259 | - |
| AraC | Ser208 | G/A | Ser211 | - | His212 | G | Arg215 | - | Asp256 | C | Leu258 | - | Tyr259 | - | Ser261 | - | Arg262 | - | Lys263 | - |
| MelR | Asn222 | G | Met225 | - | Gly226 | - | Gln229 | - | Asp270 | - | Ser272 | - | Arg273 | G | Tyr275 | - | Ser276 | - | Gly279 | - |
| XylS | Arg242 | G | Tyr245 | - | Asn246 | T/A | Glu249 | P | His293 | - | Gly295 | P | Arg296 | G | Ala298 | - | Glu299 | T | Arg302 | P |

[1]Aa, amino acid/position; C, type of contact: C, T, G, A, specific base contacts; -, unknown; P, phosphate backbone contacts

In light of the alignment in Figure 1A and the amino acid-base contacts established for some of the AraC family members (Table 1), we conclude that there is a correlation between the amino acids located at base contact positions and the nucleotide present at the corresponding cognate DNA binding motif. MarA, Rob and SoxS share a Trp residue that recognizes C or T nucleotides in the binding site, and Arg residues present in RhaS, MelR, MarA, Rob and SoxS preferentially establish contacts with G nucleotides[15,21,31,33,34]. On the basis of the data summarized in Table 1 and our present results, we propose that Arg242 and Arg296 contact the G nucleotides at the A and B boxes, respectively, whereas Asn246 and Glu299 recognize the T nucleotides in box A and B, respectively.

In close relation to this hypothesis, the alignment of the primary amino acid sequence of XylS homologues involved in aromatic degradation (BenR, IpbR, EbdR and BphS. See http://www. bactregulators.org) and of their corresponding target promoter sequences (Figure 7) supports the predicted contacts of α3 and α6 helices with A and B boxes respectively. Base-recognition residues identified at XylS α6 recognition helix (Arg296 and Glu299) were conserved in these homo-

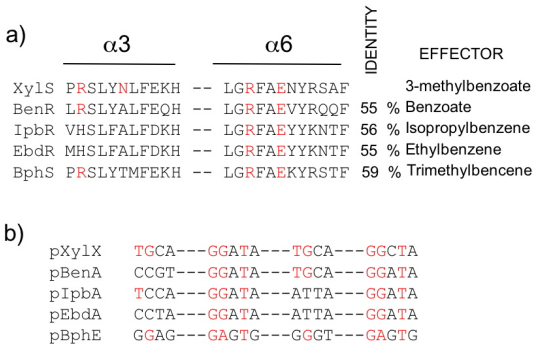



**Figure 7.** A) Alignment of XylS, BenR, IpbR, EbdR and BphS covering recognition helices α3 and α6. Residues in red denote positions of base-specific contacts in XylS. B) Alignment of proposed activator binding sequences at the corresponding regulated promoters.

Alanine substitution of XylS residues predicted to establish phosphate backbone contacts (i.e., Ser243, Glu249, Lys250 in α-helix 3 and Asn299, Arg302 in α-helix 6) led to an equal or greater defect in transcription activation than did mutations in the XylS residues proposed to make base-specific contacts. This suggests that XylS-Pm binding specificity may also be derived from "shape complementarity," i.e., different types of

interactions may be established inside the major grooves, with phosphate contacts making a large contribution to the binding strength[21,33]. In fact, we have shown that the binding of the XylS C-terminal domain to DNA provokes a curvature of about 100º in Pm[22]. This curvature would increase the probability of protein-DNA contacts taking place at major grooves, consequently enhancing binding specificity and affinity of the XylS protein for the Pm promoter[36,37].

The positioning of activators with respect to RNA polymerase (RNAP) on DNA helix depends on the orientation and spacing between the activator binding sites and the -35 element in the corresponding promoter[25]. The binding orientation of different AraC family members has been determined experimentally by a variety of methods, including genetic epistasis experiments for RhaS[15] and biochemical assays for MelR and AraC[5,32,34]. These three regulators, as well as XylS, activate class II promoters and share the same orientation, known as "forward orientation", in which the C-terminal HTH-2 is adjacent to, or overlaps by 1 or 2 bp, the –35 region of their respective regulated promoters. As for other AraC family members, this asymmetric binding of XylS reflects both the asymmetry of its binding sites and of its C-terminal domain, i.e., one XylS monomer bound per single binding site with two different HTH motifs recognizing two adjacent major grooves with dissimilar sequence (boxes A and B). Activators of class II promoters have been reported to establish close contact with the sigma subunit of RNA polymerase. In this context, previous studies in several AraC family members defined an amino acid region in the regulator which was required to establish interactions with the sigma factor and therefore for activation[38-40]. This region determines a surface-exposed patch, which includes some acidic residues (located in α5 in the HTH2), which is conserved among AraC family proteins involved in transcription activation at class II promoters. Preliminary results with XylS alanine mutants at positions predicted to contact the sigma subunit of RNAP (Glu284, Thr283) have rendered XylS variants impaired in transcription activation. Further work to identify specific amino acid contacts with the sigma subunit is underway.

In summary, our results support that a XylS dimer binds to bipartite target sequences called box A and box B by making specific contacts with α-helix 3 and α-helix 6 at the C-terminal end of XylS, respectively. Contacts established between Arg242 and Asn246 residues in α-helix 3 and the GT dinucleotide at box A, and contacts between Arg296 and Glu299 residues in α-helix 6 and G and T in box B, are essential for XylS specific binding to Pm. This binding conformation will define XylS interactions with RNAP leading to transcription activation.

## Materials and Methods

### Bacterial strains and plasmids

The bacterial strains and plasmids used in this work are listed in Table 2. Bacterial strains were grown with routine methods at 30 ºC and 200 rpm in Luria-Bertani medium as described before[41]. Growth was determined turbidometrically at 660 nm. The following plasmids have been previously described: pCMX2 is a tetracycline resistance derivative of pSELECT containing the entire *xylS* gene[42]; pLOW2 is a pACYC177 derivative, low-copy-number cloning vector encoding kanamycin[43]; pJLR100 is a pEMBL9 derivative bearing the Pm promoter cloned between the *Eco*RI and *Hin*dIII sites[44]; pJLR107 is a pMD1405 derivative bearing the Pm promoter in front of *lac* [44,45]

### DNA techniques

All DNA manipulations were done according to standard procedures[48] or to the manufacturer's recommendations.

### Generation of xylS single point mutants

The *xylS* mutants were generated by overlap extension polymerase chain reaction mutagenesis as described[42]. The internal oligonucleotide primers used for mutagenesis had one mismatch with respect to the wild-type sequence. The external oligonucleotide primers were complementary to plasmid flanking region (5'-GTTGTAAA ACGACGGCCAGTGAAT-3') and (5'-GTTATC ATCTGCAAATAATACTCAAAGG-3'). The template for each mutagenesis was 200 ng pCMX2, and amplification conditions were as described[49]. After DNA amplification, the resulting DNA was digested with *Eco*RI and *Xho*I to isolate the *xylS* N-terminal DNA fragment or with *Xho*I and *Nco*I to obtain *xylS* C-terminal fragment, and the resulting fragment containing the mutant *xylS* sequence was inserted between the same sites of pCMX2 to restore the entire *xylS* gene in plasmids pCMX2::*xylS** (where * refers to the amino acid change in the *xylS* mutant). Mutations in all *xylS* mutants sequences generated in this study were confirmed by DNA sequencing. Finally, *xylS** mutants present in pCMX2::*xylS** were cloned between the *Eco*RI and *Xba*I sites of the low-copy-number plasmid pLOW2 to generate the entire set of plasmids pLOW2::xylS* (Table 2) to be used in the induction experiments.

**Table 2**. Strains and plasmids used in this study.

| Strain | Relevant characteristics[1] | References |
|---|---|---|
| *E. coli* DH5α | *supE44 lacU169*(*Ø80lacZΔ M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1* | 46 |
| *E. coli* MC4100 | F⁻, *araD139*, (*argF-lac*)*U169*, *rpsL150*, *relA1*, *flbB5301*, *deoC1*, *ptsF25*, *rpsR*. $Sm^R$ | 47 |
| **Plasmid** | | |
| pJLR100 | $Ap^R$, Pm cloned in pEMBL9 | 44 |
| pCMX2 | $Tc^R$, pSELECT derivative containing the wild-type xylS gene | 42 |
| pLOW2 | KmR. pACYC177 derivative, low copy number cloning vector | 43 |
| pJLR107 | $Ap^R$, Pm::*lacZ* in pMD1405 | 44,45 |
| pMD::Pm* | $Ap^R$, pMD1405 derivatives bearing Pm promoter mutants | 2 |

[1] $Ap^R$, $Km^R$ and $Tc^R$ stand for resistance to ampicillin, kanamycin and tetracycline, respectively

#### β-galactosidase expression assays

*Escherichia coli* MC4100 bearing the wild-type and *xylS* mutants in pLOW2 plus the Pm::*lacZ* fusion in pJLR107 were grown overnight on Luria-Bertani medium containing the appropriate antibiotics. Cultures were diluted 100-fold in the same medium and grown at 30 ºC, 200 rpm for 1 h. Then each culture was divided in two fractions, and 1 mM 3MB was added to one of them. *β*-galactosidase activity was measured 4 h after addition, as described by Miller (1972)[50].

#### Crude extract preparation

*Escherichia coli* DH5α bearing a wild-type or mutant *xylS* in pLOW2 was grown overnight in LB medium with tetracycline (10 µg/ml) or kanamycin (25 µg/ml) respectively at 30 ºC and 200 rpm. An aliquot fraction of these precultures was used to inoculate 500 ml fresh medium, and the culture was incubated at 30 ºC. When the $OD_{660}$ reached 0.6, 1 mM 3MB was added and incubation was continued for 3 h. The cultures were then harvested by centrifugation, and cell pellets were stored at -70 ºC. Crude extracts were prepared by resuspending the pellet in 3 ml lysis buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 4 mM β-mercaptoethanol, 2 mM EDTA, 1 × Complete™ protease inhibitor mixture (Roche Applied Science) followed by cell disruption by sonication. The clear supernatant was filtered through a 0.45 µm-pore-size nylon membrane and loaded onto a 1-ml Heparin column (HiTrap Heparin HP, Amersham Biosciences) pre-equilibrated with buffer A (10 mM sodium phosphate pH 7, 4 mM β-mercaptoethanol, ± 3MB). Unbound material was collected and the column was washed with buffer A until nonspecifically bound material had been removed. Specifically bound proteins were eluted from the column with buffer A supplemented with 1 M NaCl. The protein concentration in the fractions was determined according to Bradford (1976)[51].

#### Electrophoretic mobility shift assays

The 100-bp DNA fragments containing the wild-type or mutant *xylS* binding site (positions from -110 to -10 of Pm promoter) were amplified by PCR with primers Pm3 (5′-CTGCAGTGTCCGGTTTGATAGGG-3′) and Pm4 (5′-CCTAAGGGGTAGGCCTTTCTAG-′). The PCR products were isolated from agarose gels and end-labeled with [γ-$^{32}$P]ATP as described[52]. Cell extracts and end-labeled DNA fragments were mixed and incubated at 30 ºC for 15 min in 10 µl binding buffer (Tris-glycine buffer [25 mM Tris-HCl, 200 mM glycine pH 8.6], 200 mM NaCl, 4 mM β-mercaptoethanol, 4 mM $MgCl_2$, 4 mM EDTA), supplemented or not with 1 mM 3MB. Samples were loaded onto a 4.5 % (wt/vol) non-denaturing polyacrylamide gel and eletrophoresed at 50 V in Tris-glycine buffer (25 mM Tris-HCl, 200 mM glycine pH 8.6) for 2 h at 4 ºC. The gels were dried and visualized by exposure to phosphorimager (Bio-Rad, Madrid, Spain) screens. The results were analyzed with Molecular Imager FX equipment and QuantityOne software (Bio-Rad, Madrid, Spain).

#### Semi-automated microtiter plate β-galactosidase assays

*Escherichia coli* MC4100 bearing the wild-type and *xylS* mutants in pLOW2 plus the *lacZ* fusions of wild-type Pm or mutant promoters in pMD1405, were grown at 30 ºC overnight in microtiter plates on Luria-Bertani medium containing the appropriate antibiotics, and in the presence or absence of 3MB. Then 15 µl of each culture was transferred into a new microtiter plate and β-galactosidase activity was measured as described by Miller (1972). Absorbance was measured at 660, 420 and 550

nm with a microplate spectrophotometer (TECAN systems).

## Acknowledgments

This study was supported by Grant BIO2006-05668 from the CICYT, and by Project of Excelencia from the Junta de Andalucía. P. Domínguez-Cuevas is the recipient of a Junta de Andalucía Predoctoral Fellowship in Spain. We thank Ana Hurtado for DNA sequencing, M. Mar Fandila and Carmen Lorente for secretarial assistance and for improving the use of English in the manuscript.

# References

1. Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K. & Ramos, J. L. (1997). Arac/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* **61**, 393-410.
2. González-Pérez, M. M., Ramos, J. L., Gallegos, M. T. & Marqués, S. (1999). Critical nucleotides in the upstream region of the XylS-dependent TOL *meta*-cleavage pathway operon promoter as deduced from analysis of mutants. *J Biol Chem* **274**, 2286-90.
3. Martin, R. G., Jair, K. W., Wolf, R. E., Jr. & Rosner, J. L. (1996). Autoactivation of the *marRAB* multiple antibiotic resistance operon by the MarA transcriptional activator in *Escherichia coli*. *J Bacteriol* **178**, 2216-23.
4. Webster, C., Gardner, L. & Busby, S. (1989). The *Escherichia coli melR* gene encodes a DNA-binding protein with affinity for specific sequences located in the melibiose-operon regulatory region. *Gene* **83**, 207-13.
5. Brunelle, A. & Schleif, R. (1989). Determining residue-base interactions between AraC protein and araI DNA. *J Mol Biol* **209**, 607-22.
6. Poore, C. A., Coker, C., Dattelbaum, J. D. & Mobley, H. L. (2001). Identification of the domains of UreR, an AraC-like transcriptional regulator of the urease gene cluster in *Proteus mirabilis*. *J Bacteriol* **183**, 4526-35.
7. Prouty, M. G., Osorio, C. R. & Klose, K. E. (2005). Characterization of functional domains of the *Vibrio cholerae* virulence regulator ToxT. *Mol Microbiol* **58**, 1143-56.
8. Ruíz, R., Marqués, S. & Ramos, J. L. (2003). Leucines 193 and 194 at the N-terminal domain of the XylS protein, the positive transcriptional regulator of the TOL *meta*-cleavage pathway, are involved in dimerization. *J Bacteriol* **185**, 3036-41.
9. Ruíz, R. & Ramos, J. L. (2002). Residues 137 and 153 at the N terminus of the XylS protein influence the effector profile of this transcriptional regulator and the sigma factor used by RNA polymerase to stimulate transcription from its cognate promoter. *J Biol Chem* **277**, 7282-6.
10. Soisson, S. M., MacDougall-Shackleton, B., Schleif, R. & Wolberger, C. (1997). Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**, 421-5.
11. Tobes, R. & Ramos, J. L. (2002). AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucleic Acids Res* **30**, 318-21.
12. Ramos, J. L., Marqués, S. & Timmis, K. N. (1997). Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. *Annu Rev Microbiol* **51**, 341-73.
13. Ramos, J. L., Rojo, F., Zhou, L. & Timmis, K. N. (1990). A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and the *Escherichia coli* AraC activators. *Nucleic Acids Res* **18**, 2149-52.
14. Zhou, L. M., Timmis, K. N. & Ramos, J. L. (1990). Mutations leading to constitutive expression from the TOL plasmid *meta*-cleavage pathway operon are located at the C-terminal end of the positive regulator protein XylS. *J Bacteriol* **172**, 3707-10.
15. Bhende, P. M. & Egan, S. M. (1999). Amino acid-DNA contacts by RhaS: an AraC family transcription activator. *J Bacteriol* **181**, 5185-92.
16. Jair, K. W., Martin, R. G., Rosner, J. L., Fujita, N., Ishihama, A. & Wolf, R. E., Jr. (1995). Purification and regulatory properties of MarA protein, a transcriptional activator of *Escherichia coli* multiple antibiotic and superoxide resistance promoters. *J Bacteriol* **177**, 7100-4.
17. Michán, C. M., Busby, S. J. & Hyde, E. I. (1995). The *Escherichia coli* MelR transcription activator: production of a stable fragment containing the DNA-binding domain. *Nucleic Acids Res* **23**, 1518-23.
18. Munson, G. P. & Scott, J. R. (1999). Binding site recognition by Rns, a virulence regulator in the AraC family. *J Bacteriol* **181**, 2110-7.
19. Bustos, S. A. & Schleif, R. F. (1993). Functional domains of the AraC protein. *Proc Natl Acad Sci USA* **90**, 5638-42.
20. Caswell, R., Williams, J., Lyddiatt, A. & Busby, S. (1992). Overexpression, purification and characterization of the

*Escherichia coli* MelR transcription activator protein. *Biochem J* **287 (Pt 2)**, 493-9.
21. Rhee, S., Martin, R. G., Rosner, J. L. & Davies, D. R. (1998). A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc Natl Acad Sci USA* **95**, 10413-8.
22. Domínguez-Cuevas, P., Ramos, J. L., and Marqués, S. (2007). Sequential and cooperative XylS-C domain binding to the Pm promoter provokes DNA bending essential for activation. In *J Biol Chem, submitted*.
23. Kaldalu, N., Toots, U., de Lorenzo, V. & Ustav, M. (2000). Functional domains of the TOL plasmid transcription factor XylS. *J Bacteriol* **182**, 1118-26.
24. González-Pérez, M. M., Marqués, S., Domínguez-Cuevas, P. & Ramos, J. L. (2002). XylS activator and RNA polymerase binding sites at the Pm promoter overlap. *FEBS Lett* **519**, 117-22.
25. Martin, R. G. & Rosner, J. L. (2001). The AraC transcriptional activators. *Curr Opin Microbiol* **4**, 132-7.
26. Baase, W. A., Eriksson, A. E., Zhang, X. J., Heinz, D. W., Sauer, U., Blaber, M., Baldwin, E. P., Wozniak, J. A. & Matthews, B. W. (1992). Dissection of protein structure and folding by directed mutagenesis. *Faraday Discuss*, 173-81.
27. Ebright, R. H. (1991). Identification of amino acid-base pair contacts by genetic methods. *Methods Enzymol* **208**, 620-40.
28. Ebright, R. H. (1985). Use of "loss-of-contact" substitutions to identify residues involved in an amino acid-base pair contact: effect of substitution of Gln18 of lac repressor by Gly, Ser, and Leu. *J Biomol Struct Dyn* **3**, 281-97.
29. Kahramanoglou, C., Webster, C. L., El-Robh, M. S., Belyaeva, T. A. & Busby, S. J. (2006). Mutational analysis of the *Escherichia coli melR* gene suggests a two-state concerted model to explain transcriptional activation and repression in the melibiose operon. *J Bacteriol* **188**, 3199-207.
30. Wickstrum, J. R., Skredenske, J. M., Kolin, A., Jin, D. J., Fang, J. & Egan, S. M. (2007). Transcription Activation by the DNA-Binding Domain of the AraC Family Protein RhaS in the Absence of its Effector-Binding Domain. *J Bacteriol. In press.*
31. Kwon, H. J., Bennik, M. H., Demple, B. & Ellenberger, T. (2000). Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nat Struct Biol* **7**, 424-30.
32. Niland, P., Huhne, R. & Muller-Hill, B. (1996). How AraC interacts specifically with its target DNAs. *J Mol Biol* **264**, 667-74.
33. Griffith, K. L. & Wolf, R. E., Jr. (2002). A comprehensive alanine scanning mutagenesis of the *Escherichia coli* transcriptional activator SoxS: identifying amino acids important for DNA binding and transcription activation. *J Mol Biol* **322**, 237-57.
34. Grainger, D. C., Belyaeva, T. A., Lee, D. J., Hyde, E. I. & Busby, S. J. (2003). Binding of the *Escherichia coli* MelR protein to the *melAB* promoter: orientation of MelR subunits and investigation of MelR-DNA contacts. *Mol Microbiol* **48**, 335-48.
35. Gillette, W. K., Martin, R. G. & Rosner, J. L. (2000). Probing the Escherichia coli transcriptional activator MarA using alanine-scanning mutagenesis: residues important for DNA binding and activation. *J Mol Biol* **299**, 1245-55.
36. Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M. & Zhurkin, V. B. (1998). DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci USA* **95**, 11163-8.
37. Sarai, A. & Kono, H. (2005). Protein-DNA recognition patterns and predictions. *Annu Rev Biophys Biomol Struct* **34**, 379-98.
38. Wickstrum, J. R. & Egan, S. M. (2004). Amino acid contacts between sigma 70 domain 4 and the transcription activators RhaS and RhaR. *J Bacteriol* **186**, 6277-85.
39. Landini, P. & Busby, S. J. (1999). The *Escherichia coli* Ada protein can interact with two distinct determinants in the sigma70 subunit of RNA polymerase according to promoter architecture: identification of the target of Ada activation at the *alkA* promoter. *J Bacteriol* **181**, 1524-9.
40. Grainger, D. C., Webster, C. L., Belyaeva, T. A., Hyde, E. I. & Busby, S. J. (2004). Transcription activation at the *Escherichia coli melAB* promoter: interactions of MelR with its DNA target site and with domain 4 of the RNA polymerase sigma subunit. *Mol Microbiol* **51**, 1297-309.
41. Domínguez-Cuevas, P., Marín, P., Ramos, J. L. & Marqués, S. (2005). RNA polymerase holoenzymes can share a single transcription start site for the Pm promoter. Critical nucleotides in the -7 to -18 region are needed to select between RNA polymerase with sigma$^{38}$ or sigma$^{32}$. *J Biol Chem* **280**, 41315-23.
42. Manzanera, M., Marqués, S. & Ramos, J. L. (2000). Mutational analysis of the highly conserved C-terminal residues of the XylS protein, a member of the AraC family of transcriptional regulators. *FEBS Lett* **476**, 312-7.
43. Hansen, L. H., Sorensen, S. J. & Jensen, L. B. (1997). Chromosomal insertion of the entire

*Escherichia coli* lactose operon, into two strains of *Pseudomonas*, using a modified mini-Tn5 delivery system. *Gene* **186**, 167-73.
44. Ramos, J. L., Stolz, A., Reineke, W. & Timmis, K. N. (1986). Altered effector specificities in regulators of gene expression: TOL plasmid *xylS* mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. *Proc Natl Acad Sci USA* **83**, 8467-71.
45. Ramos, J. L., Mermod, N. & Timmis, K. N. (1987). Regulatory circuits controlling transcription of TOL plasmid operon encoding *meta*-cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. *Mol Microbiol* **1**, 293-300.
46. Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-80.
47. Lange, R. & Hengge-Aronis, R. (1991). Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. *Mol Microbiol* **5**, 49-59.
48. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (2007). *Current Protocols in Molecular Biology, Wiley, New York*.
49. Higushi, R. (1990). PCR protocols. A guide to methods and applications. In *Recombinant PCR*, pp. 177-183. Academic Press, San Diego, CA.
50. Miller, J. (1972). Experiments in molecular genetics, pp. 352-355. Cold Spring Harbor, N.Y., USA.
51. Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-54.
52. Terán, W., Felipe, A., Segura, A., Rojas, A., Ramos, J. L. & Gallegos, M. T. (2003). Antibiotic-dependent induction of *Pseudomonas putida* DOT-T1E TtgABC efflux pump is mediated by the drug binding repressor TtgR. *Antimicrob Agents Chemother* **47**, 3067-72.

CAPÍTULO 4

# La unión secuencial y cooperativa de XylS-C al promotor Pm induce una curvatura en el ADN necesaria para la activación.

**Patricia Domínguez Cuevas, Juan L. Ramos y Silvia Marqués**

La proteína XylS, miembro de la familia de reguladores transcripcionales AraC, consta de dos dominios funcionales, un dominio C-terminal de unión al ADN, y un dominio N-terminal que participa en el reconocimiento del efector y en el proceso de dimerización. Hasta la fecha, la insolubilidad inherente a la proteína ha dificultado el análisis bioquímico de la misma, limitando su estudio a aproximaciones genéticas *in vivo*. En este trabajo hemos empleado una versión truncada de XylS, carente del extremo N-terminal, para determinar las consecuencias moleculares de la unión a sus secuencias operadoras. El extremo C-terminal de XylS (XylS-C) activa la transcripción desde el promotor Pm de manera constitutiva, es decir, independientemente de la presencia del efector. Mediante ensayos de retardo en gel demostramos la unión secuencial de XylS-C al promotor Pm, de forma que se observa la aparición de dos complejos proteína-ADN, correspondientes a la unión de 1 ó 2 monómeros de XylS respectivamente. El promotor Pm presenta dos repeticiones directas de la secuencia $TGCAN_6GGNTA$, localizadas entre las posiciones -69 a -55 y -48 a -34, que constituyen los sitios de unión de XylS. El análisis de la unión de XylS-C a Pm reveló su asociación en primer lugar al motivo proximal (localizado entre -48 y -34), solapando con la región -35 de reconocimiento de la ARN polimerasa. Dicha unión provoca un aumento de la curvatura en Pm y un desplazamiento del punto de máxima curvatura desde el motivo proximal de reconocimiento de XylS, hacia la región de ADN entre los dos sitios de unión. Esto favorece la unión cooperativa del segundo monómero de XylS al motivo distal, aumentando el ángulo de curvatura hasta 98º. De acuerdo con los resultados obtenidos proponemos un modelo en el que la estructura del ADN y las secuencias de unión determinan la sucesión de eventos en la asociación de XylS a Pm, provocando finalmente la activación de la transcripción.

Enviado a *Journal of Biological Chemistry*

# Sequential and Cooperative XylS-C Domain Binding to the Pm Promoter Induces DNA Bending Essential for Activation*

**Patricia Domínguez-Cuevas, Juan-Luis Ramos and Silvia Marqués[1].**

*From the Department of Environmental Protection, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, Apartado de Correos 419, E-18080-Granada, Spain.*

**XylS protein, a member of the AraC family of transcriptional regulators, comprises two domains, a C-terminal domain involved in DNA binding, and an N-terminal domain required for effector binding and protein dimerization. In the absence of benzoate effectors the N-terminal domain behaves as an intramolecular repressor of DNA binding domain. To date insolubility of the whole protein had limited protein analysis to genetic approaches *in vivo*. To characterize the molecular consequences of XylS binding to its operator sequences, we used a recombinant XylS-C version devoided of the N-terminal domain. The truncated protein activated transcription from its cognate promoter in an effector-independent manner and proved to be soluble. EMSA showed sequential binding of XylS-C to Pm, forming two protein-DNA complexes which corresponded to DNA bound to one and two XylS-C monomers, respectively. The Pm promoter exhibits two direct repeats TGCAN$_6$ GGNTA that extends from -69 to -55 and from -48 to -34. Analysis of XylS-C binding to the Pm promoter revealed the protein binds first to the -48 /-34 site overlapping RNAP binding sequence, which induces the intrinsic Pm bending to increase, concomitant with a shift in the bending angle from the proximal site to the region between the two binding sites. This favors cooperative binding of the second monomer, which provokes a further increase in the bending angle to 98º. We propose a model in which DNA structure and binding sequences strongly influence XylS binding events leading to transcription activation.**

XylS, the transcriptional regulator of the *meta*-cleavage pathway for alkylbenzoates catabolism in *Pseudomonas putida* TOL plasmid, senses the presence of 3-methylbenzoate (3MB)[2] and activates transcription from the *meta*-cleavage operon Pm promoter (1). XylS overproduction through a toluene-induced cascade regulatory circuit also leads to transcription activation in a 3MB independent manner (2,3). XylS protein is a member of the AraC family of transcriptional regulators (4-6). The proteins of this family family share significant homology over a 100 amino-acid segment, which constitutes the DNA binding domain. Most family members contain an additional domain, involved in effector recognition and oligomerization, which modulates their transcriptional activity (7-9).

[1]This work was supported by MEC Grant BCM 2001-0515 and Consolider BIO2006-05668. PD-C was the recipient of a Junta de Andalucía predoctoral fellowship.
[1]To whom correspondence should be addressed. Tel. + 34 958 181600; Fax + 34 958 129600; E-mail: silvia@eez.csic.es
[2]The abbreviations used are: 3MB, 3-methylbenzoate; EMSA, electrophoresis mobility shift assay; HTH, helix-turn-helix

AraC family members studied in some detail so far use different mechanisms to regulate transcription. Most AraC class proteins related to stress response are active as monomers, e.g., MarA and SoxS. These two proteins are composed of only the C-terminal DNA binding domain (10,11). They activate transcription from both class I and class II promoters exclusively depending upon their intracellular concentrations (12,13). A number of AraC proteins involved in carbon metabolism have been extensively characterized. They are usually active as dimers and appear to activate mainly class-II promoters (5). Activation by this second group of AraC family proteins, e.g. AraC, Rhas and MelR, responds to the presence of their specific effectors. Effector recognition by the N-terminal domain triggers a conformational change that leads to the correct positioning of the regulator at its binding sites in the promoters (7,14-16). Irrespective of the mechanism used for activation, a common characteristic shared by most family members is that their synthesis is strictly regulated (17-20).

As in other members of the AraC family, XylS is composed of two separate and independently functional domains. The DNA binding domain of XylS is located in the C-terminal end of the protein, connected by a short (amino acids 205 to 213) linker domain to the 204 amino-acid N-terminal end involved in 3MB recognition and dimerization (21-24). XylS C-terminal domain devoid of the N-terminal domain of the protein is able to activate Pm promoter in a 3MB independent manner (21). Functional separation in XylS protein comes to such an extent that XylS DNA binding domain, unlike most AraC family proteins, is able to activate transcription without the intervening of any additional determinant present in the N-terminal domain (21,25-28). In fact, we have shown that the XylS N-terminal domain repressed XylS-C DNA binding capacity by establishing direct interactions with this moiety, and one of 3MB functions would be to release this repression (24).

X-ray crystallography of MarA-DNA complex showed that the conserved DNA binding domain was composed of seven α-helices folding into two HTH motifs. The recognition helix of each motif was inserted into adjacent major grooves of the DNA (29). In XylS, a thorough genetic analysis allowed us to determine XylS binding geometry. This protein recognizes two 15-bp direct repeats (TGCA-N6-GGNTA) in the Pm promoter each consisting of two highly conserved sequences, the 5'-box A (TGCA) and the 3'-box B (GGNTA), whose centres are separated by 10 bp, corresponding to one DNA helical turn (30,31) (Figure 1). Our results showed that XylS bound Pm direct repeats in a head-to-tail organization, with helix 3 from the first HTH motif (residues Arg242 and Asn246) recognizing specific bases of A boxes and with the second HTH motif interacting with the B boxes through recognition helix 6 (residues Arg296 and Glu299) (32).

As most members of the AraC family, XylS is poorly soluble, hampering protein purification and making it reluctant to biochemical analysis (8,25,33-35). In the work reported here,

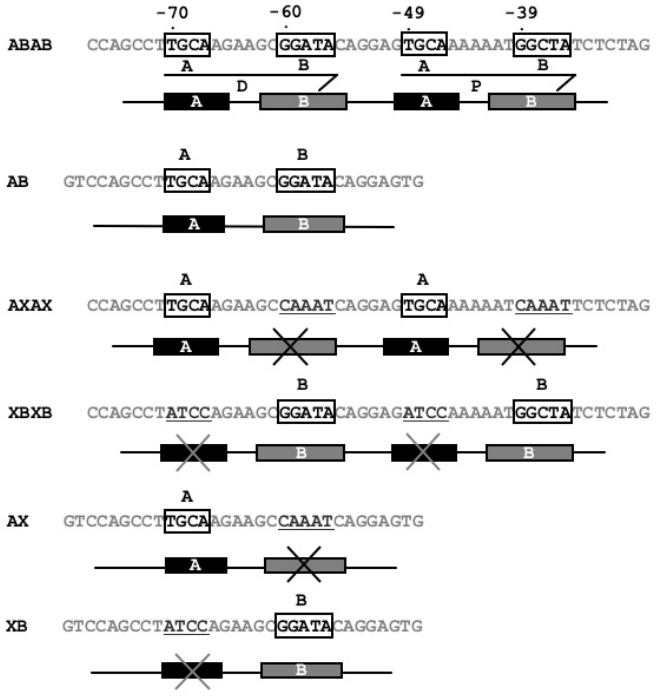



FIGURE 1. **Pm sequence and mutant promoters used in this work.** Nucleotide sequences of annealed oligonucleotides, containing one or two XylS binding sites, used for gel shift experiments. ABAB sequence corresponds to wild-type Pm. Letters correspond to the A and B boxes at each binding site, which is indicated by an arrow below the sequence. Mutant Pm templates AXAX and XBXB show the substituted bases in dark grey and underlined. The specific base substitutions in each knocked-out box were designed based on mutant Pm activities determined by González-Pérez (1999) to select mutations totally impairing Pm activity.

we examined XylS DNA binding properties for which we have overexpressed and purified a XylS C-terminal domain his-tagged fusion protein, obtaining *in vitro* active protein preparations. We have found that XylS-C monomers bind cooperatively to tandem sites and that their sequential binding provokes a gradual DNA bending. Further analysis let us conclude that XylS-C DNA recognition not only directly depended on specific amino-acid base contacts (32) but also indirectly on intrinsic DNA conformation features and mainly on binding-induced DNA bending, that in the end would determine XylS-RNA polymerase contacts leading to transcription activation.

## EXPERIMENTAL PROCEDURES

*Bacterial strains, culture media and plasmids*—Bacterial strains and plasmids used in this study are listed in Table 1. Bacterial strains were grown at 30ºC in Luria-Bertani medium supplemented, when required, with 100 µg/ml ampicillin, 25 µg/ml kanamycin or 20 µg/ml rifampicin. *E. coli* BL21 (DE3) strain was grown on 2xYT medium for protein production. Plasmid pGP1-2 carries an inducible T7 RNA polymerase gene (36), pSR is a pBR322 derivative for cloning promoter fragments upstream of the λoop terminator (Kolb, A et al 1995), pJLR100 is a pEMBL9 derivative bearing the Pm promoter cloned between the EcoRI and HindIII sites (Ramos 1986), pBend2::Pm80 is a pBend2 derivative carrying a 86-bp DNA fragment containing wild-type XylS binding sites (37).

*Overexpression and purification of his-tagged XylS C-terminal domain (XylS-C)*—The 400-bp DNA fragment covering the XylS C-terminal domain (XylS residues 196-321) was PCR-amplified from the TOL plasmid DNA using primers XylS-CNdeI (5'-GGAATTCCATATGCTGGGCAGCAATGTC AGC-3') and XylSXhoI (5'-CCGCTCGAGTCAAGCCACTT CCTTTTTGC-31). The PCR amplification product was digested with NdeI and XhoI enzymes and cloned into the pET16b vector (Novagen). The resulting clones were sequenced to confirm XylS-C sequence. To purify truncated derivative XylS-C (XylS C-terminal domain), freshly transformed BL21 (DE3) cells harboring pET16b::*XylS-C* plasmid were grown in 500 ml 2xYT medium (38) at 30°C until turbidity at 660 nm reached 0.5. Culture was then transferred to 16°C and incubated with 0.1mM IPTG for 3-4 h, cells were then pelleted and frozen. The cell pellet was resuspended in 50 ml of lysis buffer (20 mM sodium phosphate, pH 7.5, 300 mM NaCl, 0.1 mM EDTA, 2.5 mM 2-mercaptoetanol, 10% (v/v) glycerol, 10 mM imidazol and 1 mM Complete™ (Roche)) and disrupted in a French pressure cell. The crude extract was centrifuged at 30.000 x g for 1 hour, filtered through a 0.45 µm pore-size filter and loaded onto a 5 ml Ni-agarose column (Amersham Biosciences) pre-equilibrated with buffer A-XylS-C (lysis buffer without Complete). The column was washed with buffer A-XylS-C until nonspecifically bound material had been removed. A 40 ml imidazol gradient in buffer A-XylS-C (from 0 to 1 M) was then applied. XylS-C eluted at about 500 mM imidazol. Eluted fractions containing XylS-C protein were dialyzed against storage buffer (20 mM sodium phosphate, pH 7.5, 300 mM NaCl, 2 mM DTT, 10% glycerol) and stored at -70°C until use.

*Gel Filtration Chromatography*—XylS-C molecular weight was estimated by gel filtration chromatography on a Superdex 75 HR 10/30 column (Amersham Biosciences) equilibrated with 30 mM Tris-HCl pH 8, 200 mM KCl, 2 mM DTT, 10% (v/v) glycerol at room temperature. Alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome c (12.4 kDa) were used as molecular mass standards. Proteins were detected at 280 nm. Eluted fractions of XylS-C were pooled and analyzed by SDS-PAGE.

*Gel retardation assays (EMSA)*—Complementary oligonucleotides containing one or two intact XylS binding sites and the corresponding whole half-site mutants were annealed and blunt-end ligated into HincII-linearized pUC19. DNA probes, 130 or 150-bp DNA fragments respectively, were obtained by PCR-amplification of these constructs with M13 universal and reverse pair of primers. The PCR products were isolated from agarose gels using the QIAquick Gel Extraction kit (Qiagen) and end-labelled with [γ-$^{32}$P]ATP using T4 polynucleotide kinase. The sequences of annealed oligonucleotides used in this work are shown in figure 1. Single point Pm mutants used in this work were obtained from the corresponding plasmids already described (Gonzalez-Perez 99) as follows: the 100-bp DNA fragments containing the wild-type or mutant *xylS* binding sites (positions from -110 to -10 of Pm promoter) were amplified by PCR with primers Pm3 (5'-CTGCAGTGTCCGGTTTGATAGGG-3') and Pm4 (5'-CCTAAGGGGTAGGCCTTTCTAG-3') from pJLR07 or pMD:Pm* bearing different Pm mutant promoters (30). The PCR products were isolated from agarose gels and end-labelled with [γ-$^{32}$P]ATP as described (39). Labelled probes (1 nM) were incubated with increasing amounts of purified XylS-C domain at 30°C for 15 min in 10 µl of binding buffer (5 mM Tris, 24 mM Hepes pH 8, 50 mM potassium glutamate, 20 mM NaCl, 1.4 mM EDTA, 0.4 mg/ml BSA, 9% (v/v) glycerol, 0.5 µg of poly(dIdC) and 1 mM DTT). Samples were loaded onto 8% (w/v) non-denaturing polyacrylamide gels (Bio-Rad Mini-

**TABLE 1**
Strains and plamids used in this work.

| Strain | Relevant characteristics | Source or reference |
|---|---|---|
| *E. coli* CC118λPm::lacZ | RifR; CC118 with chromosomal mini-Tn*5*::Pm::*lacZ* insertion | (54) |
| *E. coli* BL21 (DE3) | Carries T7 RNA polymerase under the control of *lacUV5* promoter | Novagen |
| *E. coli* DH5α | *supE44 lacU169(Ø80lacZΔ M15) hsdR17* (r –*Km-K) recA1 endA1 gyrA96 thi-1 relA1* | (55) |
| **Plasmid** | | |
| pJLR100 | $Ap^R$, Pm cloned in pEMBL9 | (56) |
| pGP1-2 | $Km^R$, contains an inducible T7 RNA polymerase gene[a] | (36) |
| pET16b | $Ap^R$, protein expression vector | Novagen |
| pET16b::XylS-C | $Ap^R$, pET16b derivative used to produce his-tagged XylS | This work |
| pBend2::Pm80 | $Ap^R$, pBend2 derivative carrying a 86-bp DNA fragment containing the Pm promoter | (37) |

[a]pGP1-2 contains gene *1* (RNA polymerase) of phage T7 under the control of the inducible *PL* promoter of phage l; pGP1-2 also contains the *Plac* promoter and the gene for the heat-sensitive l repressor, *cI857*

Protean III) in Tris-glycine buffer (0.2 M glycine, 0.025 M Tris-HCl [pH 8.6]). The gels were run for 3 h at 50 V and 4°C, vacuum dried and exposed to phosphorimager plates. The results were analyzed using Molecular Imager FX equipment and QuantityOne software (Bio-Rad, Madrid, Spain).

*Determination of XylS-C DNA stoichiometry in native PAGE*—The molecular weight of XylS-C-DNA complexes was determined electrophoretically as described by Orchard and May (1993). Briefly, binding reaction mixtures containing 500 nM to 5 μM) XylS-C and 50 bp target DNA were analyzed on a series of polyacrylamide gels (7 to 10% [w/v] polyacrylamide; 37.5:1) electrophoresed in Tris-glycine buffer (see above) until the bromophenol blue front in flanking sample lanes reached the bottom of the gels. Protein molecular weight markers were used as standards. The gels were stained with Bio-Rad Silver Stain Kit (Bio-Rad). The distance from the loading well to each complex was measured and standardized with the bromophenol blue migration distance to determine the relative mobility ($R_f$). Protein markers run in parallel were used as control. The logarithm of the $R_f$ was plotted against the gel concentration for each complex and protein standards, and best-fit lines were obtained. The negative slopes of these lines were then plotted against the molecular weights of the protein standards on a double-logarithmic scale, and best-line fit was obtained. Interpolation of the line slopes corresponding to protein-DNA complexes was used to deduce the molecular weights of each complex.

*DNase I Footprints*—A 158-bp DNA fragment covering XylS binding sites (positions from -113 to +35 of Pm promoter) was specifically radiolabelled at either strand 5'-end using [γ-$^{32}$P]ATP and T4 polynucleotide kinase. End-labeled DNA probes (10 nM) were preincubated for 15 min at 30°C in binding buffer with different amounts of purified XylS-C domain. DNase I was added and incubated for 5 min at 30°C, and reaction was stopped by addition of EDTA. Footprint patterns were analyzed on polyacrylamide sequencing gels calibrated with Maxam-Gilbert sequence ladders.

*Determining of association constants and cooperativity for XylS-C binding to two tandem sites*—$^{32}$P-labelled DNA fragment (1nM) containing two identical XylS binding sites was incubated with increasing amounts of purified XylS-C domain at 30°C for 15 min in 10 μl of binding buffer (5 mM Tris, 24 mM Hepes pH 8, 50 mM potassium glutamate, 20 mM NaCl, 1.4 mM EDTA, 0.4 mg/ml BSA, 9% [v/v] glycerol, 0.5 μg of poly(dIdC) and 1 mM DTT). Samples were loaded onto 8% (w/v) non-denaturing polyacrylamide gels (Bio-Rad Mini-Protean III) in Tris-glycine buffer (0.2 M glycine, 0.025 M Tris-HCl [pH 8.6]). The gels were run for 3 h at 50 V and 4°C, vacuum dried and exposed to phosphorimager plates. The results were analyzed using Molecular Imager FX equipment and QuantityOne software (Bio-Rad, Madrid, Spain). After electrophoresis gels were dried and the radioactivity in separated bands was quantified with a Phosphorimager as described (39). Interactions between XylS C-terminal domain and DNA fragments containing two binding sites resulted in the formation of stable complexes containing one bound monomer (complex I) and complexes with two bound monomers (complex II). The fraction of free DNA ($E_0$), complex I ($E_1$), or complex II ($E_2$) is given by equations A to C. To estimate $K_I$ and $K_{II}$, data from gel mobility shift assays were fit to equations A to C described by Chen *et al*. (2005) using a global non-linear least square regression method from Sigma plot (Systat software, Inc), where ($P_2$) is the total protein concentration; $K_I$ and $K_{II}$ are the macroscopic association constants for XylS-C binding to ABAB DNA fragment. Hill coefficient (nH) was used to determine the cooperativity in XylS-C monomers binding. Equation (D) was used to obtain values of Hill coefficient from different gel shift assays.

For complex CI: $DNA + mP_2 \Leftrightarrow DNA - (P_2)_m$;

$$K_I = [DNA - (P_2)_m] / [DNA][P_2]^m$$

For complex CII:

$$DNA - (P_2)_m + {}_nP_2 \Leftrightarrow DNA - (P_2)_{m+n};$$

$$K_{II} = [DNA - (P_2)_{m+n}] / K_I [DNA][P_2]^{m+n}$$

(A) $$E_0 = \frac{1}{1 + K_I [P_2]^m + K_I K_{II} [P_2]^{m+n}}$$

(B) $$E_1 = \frac{K_I [P_2]}{1 + K_I [P_2]^m + K_I K_{II} [P_2]^{m+n}}$$

(C) $$E_2 = \frac{K_I K_{II} [P_2]^{m+n}}{1 + K_I [P_2]^m + K_I K_{II} [P_2]^{m+n}}$$

(D) Hill coefficient: $$nH = 2\left[1 + \sqrt{K_{II}/K_I}\right]$$

*DNA-bending assays*—The pBend2 plasmid derivative, pBend2-Pm80 (37) was digested with different enzymes to yield a series of DNA fragments with permuted positions of XylS binding sites. These fragments were end-labelled and gel retardation assays were performed as above. To estimate the apparent bend angle (α) and to locate the centre of the bend (40), changes in the mobility of the protein-bound DNA fragments (relative to the well position) were fitted to a cosine function (E).

(E) $$\cos \alpha/2 = \mu_M/\mu_E;$$

where $\mu_M$ = Complex with the lowest mobility; $\mu_E$ = Complex with the highest mobility.

## RESULTS

*XylS-N terminal domain is responsible of XylS insolubility*—As for many members of the AraC family, previous attempts to purify XylS regulator were unsuccessful and produced protein aggregates forming inclusion bodies (22,25,41,42). To investigate the determinants of this insolubility, we decided to analyze the two protein domains independently. The XylS DNA binding domain was thought to be responsible for the insolubility of the protein (25) but, unexpectedly, it could be overproduced, and directly purified in a soluble form, while XylS N-terminal domain (XylSN) aggregated and formed inclusion bodies whatever the expression conditions used. First, each domain, cloned in a pET16b vector as a 10 His fusion gene, were tested *in vivo* for their ability to promote Pm transcription. *E. coli* CC118 (Pm::lacZ) cells bearing pGP1-2, were transformed with pET-XylS-C or pET-XylSN (Table 1). β-galactosidase activity of the two transformed strains was determined in the absence of inducer, and in the presence of 3MB, IPTG or both, and compared with a construction bearing wild-type XylS (pETXylS) (Figure 2). β-galactosidase activity levels in the absence of any inducer were 1500 Miller Units for XylS-C protein and 2300 Miller Units in the case of full length XylS. These high effector-independent levels corresponded to basal T7 polymerase expression from the *lac* promoter. When *E. coli* CC118 (Pm::*lacZ*) pGP1-2 were transformed with pET16b::XylSN plasmid bearing the XylS N-terminal domain,

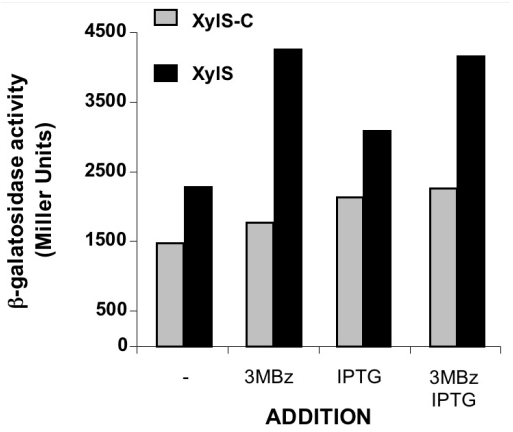



FIGURE 2. **Transcription activation capacity of XylS domains.** β-galactosidase activity of full length XylS and XylS-C proteins in *E. coli* CC118λPm:lacZ (pGP1-2) was determined as described in experimental procedures after 1 h incubation with 1 mM 3MB, 0.5 mM IPTG or both. Data are the average of 3 measurements. Standard deviation was always below 10 %.

activity from the Pm promoter was negligible in all conditions (not shown). In the *E. coli* CC118 (Pm::*lacZ*) pGP1-2 strain carrying the XylS-C fusion protein, no significant increase in Pm activity was detected with any of the inducers tested, whilst in the case of full length XylS, the addition of 3MB promoted a two-fold increase in Pm expression level. These results are in agreement with previously assigned roles for the N and C-terminal domains of XylS. Full length XylS responds to 3MB through its N-terminal region, whilst the DNA binding domain devoid of the N-terminal domain becomes constitutive and activates transcription even in the absence of effector.

*Pm half-sites are necessary and sufficient for XylS binding*—Most AraC family members appear to form dimers in solution, and dimerization determinants have been proposed to reside in the N-terminal domain (9,22,27,28,43). If this were the case, a XylS protein devoid of its N-terminal domain would behave as a monomer in solution. To analyze whether XylS-C quaternary structure was similar or different to whole protein quaternary structure, we determined XylS-C oligomeric state using gel filtration through a Superdex G-75 column as described under Material and methods. The XylS-C elution volume corresponded to a molecular weight of 16 kDa, indicating that it behaved as a monomer under our experimental conditions (not shown).

We next examined the binding of XylS-C to its target sites at the Pm promoter. Previous genetic analysis allowed us to define XylS operator as two binding sites organized as direct repeats, each composed of two boxes, A and B, different in sequence but homologous between sites (figure 1) (30,32). DNA fragments containing one or two XylS binding sites were used in gel mobility shift experiments. Figure 3 shows the binding of XylS-C to either a DNA fragment containing a single XylS binding site composed of A and B half-sites, or to the ABAB DNA fragment, containing the complete two-site wild-type binding region. Titration of XylS-C into the AB DNA fragment resulted in the formation of a single complex CI (Figure 3A), while two shifted bands appeared in titrations of ABAB DNA fragment (Figure 3B). In the later assay, the faster migrating band (CI) was progressively replaced in the gel by the heavier complex (CII) as concentrations of XylS-C increased. It should be noted that the CII complex appeared early after the CI complex was formed. The kinetics of CI and CII formation suggested a positive cooperativity in CII complex formation that required further analysis (see below).

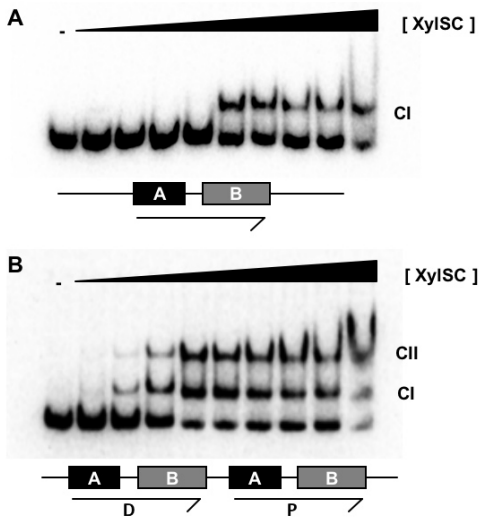



FIGURE 3. **XylS-C binding to Pm operator.** (A) Electroforetic mobility shift assay of XylS-C binding to AB template. Labelled DNA fragment (1 nM) was incubated with 50 nM, 100 nM, 200 nM, 400 nM, 600 nM, 1μM, 2 μM, 4 μM and 8 μM of purified XylS-C (lanes 2-10). Lane 1 free DNA. One XylS-C-DNA complex was formed (CI). The scheme below depicts DNA fragment sequence used in the assay, including only one binding site (arrow) composed of A (black) and B (grey) half-sites. (B) Electroforetic mobility shift assay of XylS-C binding to ABAB template. Labelled DNA fragment was incubated with 50 nM, 100 nM, 200 nM, 400 nM, 600 nM, 1μM, 2 μM, 4 μM and 8 μM of purified XylS-C (lanes 2-10). Lane 1, free DNA. XylS-C-DNA complexes CI and CII were formed. The scheme below depicts DNA fragment sequence used in the assay, including two binding sites (arrows) composed of A (black) and B (grey) half-sites.

*One XylS-C monomer per single binding site*—To investigate whether CI and CII complexes corresponded to the binding of one and two monomers of XylS-C respectively, we determined the molecular weight of the two complexes using the

electrophoretic method developed by Orchard and May (1993) based on protein-DNA complex size determination in gel mobility shift assays of ABAB DNA fragment at different acrylamide concentrations. Mobility of a series of molecular weight markers and of the two complexes was first plotted against polyacrylamide concentration and slopes were calculated in each case. Figure 4A shows a plot of the linear regression slopes obtained for each molecular weight marker against the logarithm of molecular weight. When protein-DNA complexes

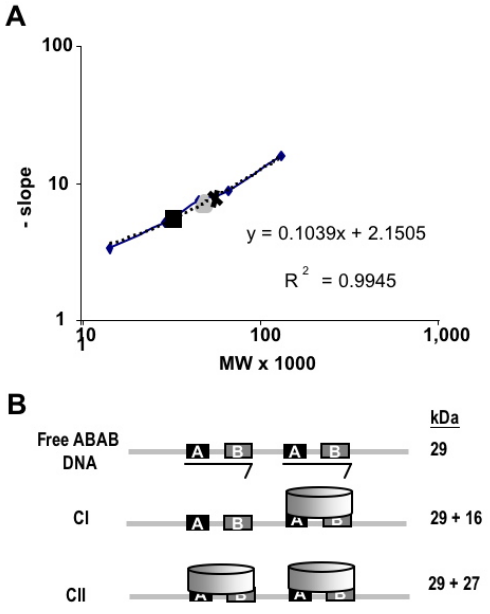



FIGURE 4. **Determination of the stoichiometry of XylS-C-DNA complexes by native PAGE.** Protein markers and XylS bound to ABAB DNA sequence were electrophoresed at different polyacrylamide concentrations. (A) The logarithm of the relative mobilities of XylS-C-DNA complexes and marker proteins with respect to bromophenol blue, were first plotted as a function of polyacrylamide concentration. The slope of the regression equation obtained for each marker is shown as a function of their molecular weight. The slopes of free DNA (■), CI (▲) and CII (**X**) were depicted and their molecular weight was extrapolated according to the equation y= 0.1039x + 2.1505. The marker proteins used were carbonic anhydrase, α-lactalbumin, bovine serum albumin dimer, bovine serum albumin monomer, and ovalbumin. (B) Scheme of XylS binding stoichiometry to either AB or ABAB DNA fragment. XylS binding sites are indicated with an arrow.

CI and CII slope values were fitted to the plot, an apparent molecular mass of 46 kDa was extrapolated for CI XylS-C-DNA complex. We estimated 17 kDa to be the molecular weight of the protein component of the CI complex (after subtracting 29 kDa corresponding to the apparent molecular mass of the ABAB DNA fragment (50 pb)), suggesting binding of a single XylS-C monomer. Likewise, an apparent molecular mass of 57 kDa was obtained for CII complex (28 kDa correspond to the protein component), in close agreement with the expected mass of 32 kDa calculated for two XylS-C monomers. The interpolated molecular weight of the XylS-C-DNA complex II was 4 kDa smaller than the sum of the individual components, a deviation that we attributed to protein-DNA complexes having a different shape than the globular protein standards (see below).

*Pm half-site proximal box upholds XylS binding*—To establish the relative influence of each half-site for XylS binding, we designed mutant DNA fragments lacking either B half-sites (AXAX) or A half-sites (XBXB) or covering only one mutated XylS binding site, AX and XB (Figure 1). The different templates were titrated with increasing XylS-C concentrations. Figure 5 shows the results with the AXAX and XBXB

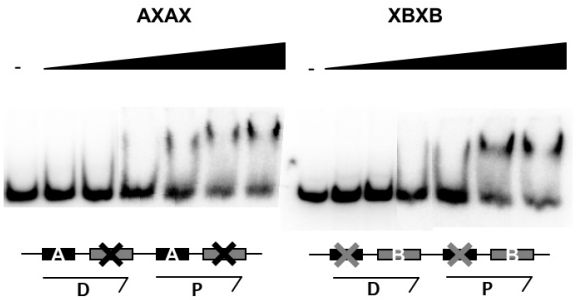



FIGURE 5. **Relevance of XylS binding half-sites for XylS-C binding to DNA.** Electroforetic mobility shift assays of XylS-C binding to a DNA fragment containing the two XylS binding sites with either the B half-sites or the A half-sites mutated in all positions. The XylS-C concentrations used were 0, 1 μM, 2 μM, 4 μM, 6 μM, 8 μM and 10 μM. The scheme below each gel depicts the DNA fragment used in each assay. The actual sequence of the fragments is shown in figure 1.

templates. XylS-C binding to both templates was severely hampered, and retarded bands could only be obtained with protein concentrations above 6 μM. This indicates that both A and B boxes are important for XylS recognition. Incubation of XylS-C with AX and XB DNA fragments did not produce any shifted complex (not shown), suggesting that both binding sites were required to obtain a shifted complex when XylS DNA binding was strongly impaired.

Wild-type Pm and a series of promoter variants bearing point mutations in the different A and B boxes were selected to determine the relevance of each half-site and the roles of the proximal and distal sites in XylS binding to Pm. Figure 6 shows that XylS-C binding to DNA fragments containing point mutations at A or B boxes of distal site resulted in the formation of the CI complex, while the shift into CII complex was severely hampered. However XylS-C bound to a Pm mutant at A proximal box shifted the DNA template forming both CI and CII

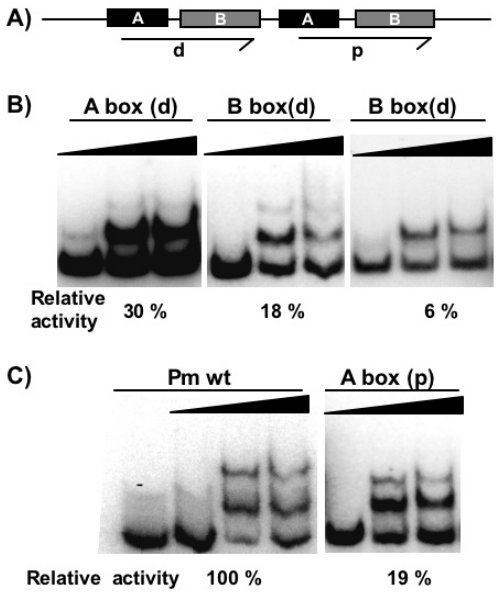



FIGURE 6. **Role of proximal and distal binding sites in XylS binding to Pm.** Gel retardation assays of XylS binding to different DNA fragment containing either wild-type of point mutants of Pm promoter XylS binding sites. A) Scheme of DNA region used in the assays, indicating the distal (d) and proximal (p) sites. B) Gel retardation assays of point mutants in A or B boxes of the distal site. C) Gel retardation assays of wild-type Pm and point mutant in A box of the proximal site. The percentages shown below each gel refers to transcriptional activity previously obtained *in vivo* with these promoters and XylS wild-type protein (30). The XylS-C concentrations used were 50 nM, 500 nM, and 2.5 μM.

complexes. Relative activity from the corresponding mutant promoters with wild-type XylS showed a severe decrease as compared to wild type promoter (30). However when we compared activities and EMSA of Pm A box mutants at proximal and distal sites we observed that proper XylS-C binding to proximal site was more important for transcription activation than CII complex formation, despite the fact XylS-C did not bind the distal site (A box(d) mutant).

*Binding of XylS-C to Pm promoter induces bending*—Previous work showed that Pm promoter DNA region was intrinsically curved, with an apparent bent angle of 35° ±5° (2,37). To explore whether binding of XylS-C induced further distortion of the DNA, we applied the circular permutation assay developed by Kim and coworkers (44) (see experimental procedures). The plasmid pBend2:Pm80 was digested with different enzymes to yield a series of 195-bp DNA fragments with the position of XylS binding sites permuted. These fragments were incubated with XylS-C and mobility of the two complexes formed (CI and CII), was analyzed in native electrophoresis (44) (Figure 7). The intrinsic bend angles (α) and the location of the centre of the bend was estimated according to Thompson and Landy. The Pm promoter intrinsic curvature centered in the A track spanning positions –41 to –46 increased to 50° ±5° after binding of one XylS-C monomer, and to 98°±2° after binding of two monomers. The centre of the new bent was located between positions –49 and –52 in both CI and CII complexes, showing a small shift from the previously determined free DNA bending centre within XylS binding site towards the DNA region between the two binding sites (Figure 7).

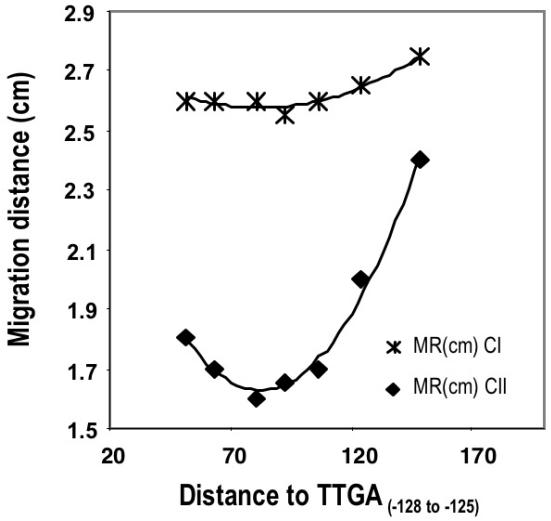



FIGURE 7. **XylS-C binding to DNA provokes DNA bending.** Gel retardation assays of XylS-C on a series of 195 bp DNA fragments, containing XylS binding sites in permutated locations were run in 8 % polyacrylamide. The migration distance of each complex (ж, CI; ♦, CII) was plotted against the distance between XylS binding site 5'-end and fragment 5'-end in each fragment tested. The bending angle for each complex was estimated according to the equation $\cos \alpha/2=\mu_M/\mu_E$. Values are the average of six independent determinations.

*XylS-C binds sequential and cooperatively the two binding sites at Pm*—Our results with half-site deleted promoters suggested that XylS-C monomers binding at Pm proceeded in two consecutive steps. To analyze whether the binding of the first XylS-C monomer occurred preferentially to either the distal or the proximal XylS binding sites, we carried out a DNase I footprinting assay of XylS-C to a Pm fragment carrying both binding sites using increasing concentrations of the regulator.

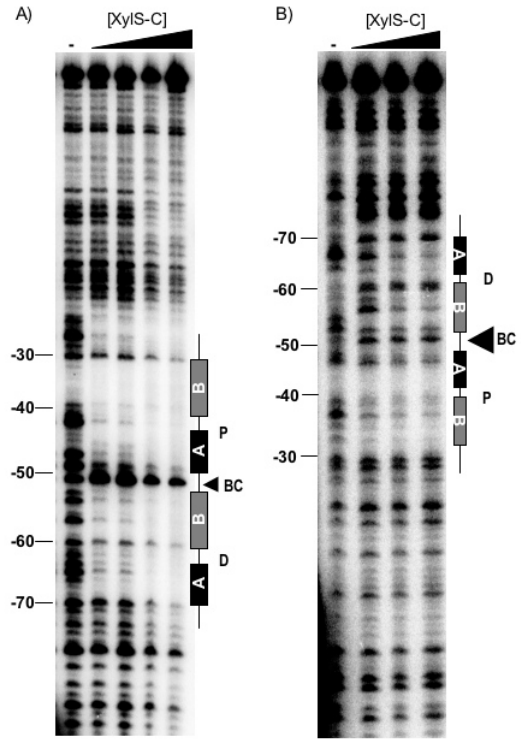



FIGURE 8. **DNase I footprint analysis of XylS/Pm complexes.** A) Upper strand; B) Bottom strand. Lane 1, free DNA; The XylS-C concentrations used were 1 μM (lane 2), 2 μM (lane 3), 5 μM (lane 4) and 10 μM (lane 5) in A) and 500 nM (lane 2), 1 μM (lane 3), 2 μM (lane 4) in B).

Figure 8 shows that XylS-C protected two regions at Pm promoter. It bound two 15-bp direct repeats centred at positions -60 (distal site) and -40 (proximal site). Furthermore, the results showed that at low XylS-C protein concentration, a clear protection appeared at the proximal site, whereas the distal site was only progressively protected by increasing XylS-C concentration. This was particularly especially evident in the bottom strand footprinting. It should be noted that a hyper-reactive band at position -42, visible in free DNA, almost disappeared at the lowest XylS-C concentration. Concomitant with it, the -51 band coinciding with the estimated Pm bending centre increased in intensity after XylS binding at the proximal site (Figure 8A). These results support sequential binding of the two XylS-C monomers at Pm, providing a progressive curvature on the DNA between the two binding sites.

The sequential binding of XylS-C to its operator prompted us to study whether the binding of the first monomer to the Pm proximal site would facilitate the binding of the second monomer. In fact, as it was mentioned above, mobility shift assays of XylS-C bound to ABAB fragment revealed the possibility of a positive cooperative effect in CII complex formation (figure 3). To test this possibility we quantified XylS-C binding to the ABAB DNA fragment. The amount of label in complexes CI and CII and in free DNA was quantified with a phosphorimager. Values were plotted as a function of the concentration of XylS-C (Figure 9) and were fitted to derived Hill equations (Experimental procedures: equations A to C). Because XylS binding sites are not identical in sequence, it is not possible to derive separate estimates of intrinsic binding affinity for distal and proximal XylS binding sites from ABAB gel shift experiments. However, we could apply mathematical modeling to estimate the macroscopic association constants, $K_I$ and $K_{II}$. The estimated value for constants $K_I$ and $K_{II}$ was

$K_I$=1.37 x $10^{-6}$ M and $K_{II}$=1.92 x $10^{-9}$ M, respectively (45). On the other hand, we fitted data from AB EMSA quantifications to derived Hill equation for a single binding site, obtaining an estimation of intrinsic XylS binding affinity for the distal site. Estimated values of intrinsic binding affinity of distal and proximal binding sites were compared ($k_I$distal=2.4 x $10^{-6}$ nM and $k_I$proximal=1.6 x $10^{-6}$ M), leading us to conclude that XylS bound proximal site with higher affinity than the distal one. According to Hill derived equation (nH= 2 / [1+√ $K_{II}$/ $K_I$]), we calculated a Hill coefficient of 1.92, which correspond to the strong positive cooperative effect observed at XylS-C monomers binding.

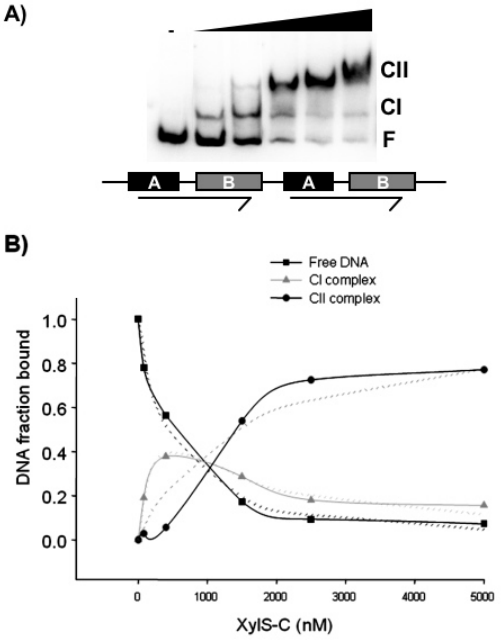



FIGURE 9. **Cooperativity in XylS-C binding to Pm.** A) EMSA of XylS-C binding to the DNA containing the XylS binding sites shown below as a scheme. The XylS-C concentrations were 0, 80 nM, 400 nM, 1.5 µM, 2.5 µM and 5 µM. CI, complex one; CII, complex two; F, free DNA. B). Quantitation of gel shift assay data from Figure 9A. The free DNA (squares), CI complex (triangles) and the CII complex (circles) were quantified in a phosphorimager. Dashed lines represent the best fit of the data to equations A-C (see Materials and methods).

## DISCUSSION

XylS, as other regulators of AraC family, contains two separate and independently functional domains, namely an N-terminal domain involved in effector recognition and dimerization, and a C-terminal domain involved in DNA binding (21-23). XylS C-terminal domain devoid of the N-terminal part of the protein is able to activate Pm promoter in the absence of effector (21) (figure 2), whereas XylS N-terminal domain (XylS-N) acts as an intramolecular repressor of XylS C-terminal domain (XylS-C) to bind DNA. In the presence of 3MB this repression is released, so that XylS-C binds Pm promoter and activates transcription (24). Therefore, XylS-C harbors all the determinants required for DNA binding and transcription activation.

XylS, as most proteins of the AraC family, is intrinsically insoluble, hampering protein purification and subsequent biochemical analysis of its properties. In this work we have studied XylS DNA binding properties based on the facts that i) its DNA binding domain unpredictably proved to be soluble when overproduced and ii) this domain could reproduce all XylS features except for 3MB responsiveness. Gel filtration column chromatography of purified XylS-C showed that it behaved as a monomer in solution. Interestingly, unlike most AraC family proteins, XylS-C was able to activate transcription, although it remained a monomer in solution (Figure 2) (24), in contrast to AraC DNA binding domain which strictly required a leucine zipper or an artificial covalent-linkage mediated dimerization to activate transcription from pBAD promoter, and to MelR C-terminal domain, which did not activate transcription despite the fact it bound DNA and provoked a bend in a similar way to the full-length MelR. Thus, XylS-C situation resembles those AraC subfamily proteins related to stress response such as MarA, Sox and Rob, which activate transcription as monomers critically and exclusively depending upon their intracellular concentrations. In addition, our results showed that XylS-C bound two direct repeats at Pm promoter in two consecutive steps, where first XylS monomer occupied the proximal binding site, leading to the binding of the second monomer to the distal site (figures 3, 4, and 8).

Genetic analysis of Pm mutant promoters revealed the relevance of B boxes for XylS binding (30). On the other hand, XylS single-point mutants at either HTH recognition helices showed that disruption of contacts established by the second HTH α6 were dramatically impaired in activation (32). These results indicated that interactions between the latter and base pairs at B-boxes were more important for transcription activation. Despite the fact that XylS-C binding affinity for DNA targets lacking either A or B boxes was severely diminished, DNA target mutants lacking B-boxes (AX sequence) were slightly more affected than mutants at A-boxes (figure 5 and 6). Besides, XylS does not seem to make direct contacts with sequence between A and B boxes. In fact, this sequence is not conserved in the consensus and it showed a weak or inexistent phenotype when mutated.

Results in Figure 8 shows that proximal and distal sites each play a specific role in XylS binding. Although both are required for optimal binding, disruption of XylS binding to the proximal site impaired complex-II formation to a lower extent than disruption of distal site binding. In contrast, mutants in the proximal site showed a severe decrease in transcription activation. This series of results suggests that while proper contacts between XylS-C and proximal site, probably required for interaction with RNA polymerase, are crucial for activation, distal site is essential to maintain complex stability despite a defective XylS-C binding to proximal site. However, it is well established that non-contacted bases can have important effects on protein affinity by affecting DNA flexibility or intrinsic curvature (46). Pm promoter exhibits a curvature of 35º±5° centered in a series of As between -41 and -46 positions and in vivo methylation showed a hypermethylated T located at -42, indicating that the DNA was distorted at this point (30). We used the cyclic permutation assay to test whether XylS-C monomers binding to Pm distorted target DNA. Our results showed that binding of the first XylS-C monomer increased Pm curvature to 50º±5° (figure 7). When the second XylS-C monomer bound Pm promoter, the apparent bending angle reached 98º±2°. The sequential binding of the two monomers shifted the bending centre towards the region between XylS binding sites (Figure 7). This position matched the band of enhanced DNase sensitivity located at –51 (Figure 8).

Intrinsic curvature of Pm at proximal binding site would cause a compression of the minor groove between A and B boxes, and an increase in the number of protein-DNA contacts

available at the major grooves. AT base pairs have been referred to as 'natural hinges' for protein-induced DNA bending (47). The A-track present at XylS proximal binding site between boxes A and B would act as a flexible motif, increasing XylS-C affinity. In fact, XylS-C affinity for the proximal site was actually higher than the affinity for the distal one, as revealed by the intrinsic binding affinity constants measured for both XylS binding sites. Other AraC family proteins, including AraC, MelR, RhaS, and MarA also tend to have AT-rich centres at their binding sites (11,42,43,48), although in the case of MelR and AraC this feature only occurs at the highest affinity binding site (14,15). It is probable that binding specificity, affinity and complex stability could be enhanced by flexibility in DNA binding site, allowing more extensive protein-DNA contacts and protein-protein interactions. Analysis of the MarA-DNA co-crystals showed that the recognition helices are held in two adjacent major grooves by a rigid long α-helix 4. The two recognition helices are separated by 27 Å whereas the distance between consecutive major groves of canonical DNA B-form is 34 Å. Thus, MarA binding to its target DNA would induce a 35º bending in that region. Due to high homology shared by the DNA binding domains of family members, XylS-C is predicted to adopt a similar arrangement (32).

Our results show that XylS-C binding to proximal site favours binding of a second XylS-C monomer to the distal site (Figure 9). This positive cooperativity could be the consequence of the gradual bending of Pm promoter that, as discussed above, would increase complex stability. On the other hand, the energy cost associated with DNA bending could be compensated for by the increase in protein-DNA contacts. The stability of the DNA-bound complex, and especially of the proximal monomer, was essential for promoter activation (Figure 6). However, from our set of data we cannot exclude an additional effect on complex stability partially caused by protein-protein interactions established between XylS-C monomers bound to DNA. The lack of N-terminal dimerization determinants in our truncated protein XylS-C does not allow further analysis in this direction.

Pm promoter can be classified as a class-II promoter (49), where multiple interactions between XylS and RNA polymerase would be established. In fact, it has previously been reported that XylS establishes interactions with the α-CTD RNA polymerase subunit (50,51). Several family members have been reported to make contacts with sigma subunit region 4 through a conserved acidic region located at C-terminal domain helix 5. Genetic epistasis experiments carried out with some family members, i.e. RhaS, AraC, and MelR identified a surface exposed patch, which included acidic residues at C-terminal helix 5 of the second HTH motif, which contributed to activation at class-II promoters. The previously discussed XylS-induced bending process is likely to facilitate and stabilize XylS-RNAP contacts in vivo to form an active transcription initiation complex. In fact, besides protein-protein interactions induced by DNA curvature, some activators stimulate contacts of DNA sequences with the RNA polymerase, favoring the transition from closed to open complex (52,53). We have shown that XylS recruits RNA polymerase to Pm promoter in response to 3MB, promoting open complex formation (24). Preliminary results where we used site-directed mutagenesis to substitute T283, E284 XylS residues for alanine rendered XylS derivatives severely defective in their ability to activate transcription from Pm promoter (data not shown). Further studies to define specific amino acid contacts between XylS and sigma RNA polymerase subunit are underway.

Figure 10 illustrates our current model of the molecular events underlying XylS-C Pm binding and transcriptional activation. As already mentioned, Pm promoter sequence shows an intrinsic bending of 35º, centered in the proximal XylS binding site overlapping RNAP –35 hexamer binding box. This bending is likely to increase interaction possibilities at the proximal binding site so that a XylS monomer would first recognize and bind this sequence. This first event would induce a shift in DNA bending angle towards the inter-site region, together with a significant increase in the bending angle (50º). This change probably increases binding affinity for the distal site, provoking a fast, cooperative binding of the second monomer. The observed cooperativity, partly consequence of progressive changes in DNA structure, is so strong that transition of CI to CII in the complete wt XylS protein is not detected unless dimerization is impaired (24). Alternatively, the set of results could point to XylS being a dimer in solution and binding Pm sequentially as above but fast enough to escape detection of CI in EMSA. Complex with a single monomer would thus only be detectable when the protein were unable to dimerize in solution.

XylS binds Pm as in class II promoters, establishing contacts with RNAP through α–CTD (50,51) and most probably through σ factor too. Contacts would induce open complex formation (24) and further transcription initiation under the appropriate conditions.

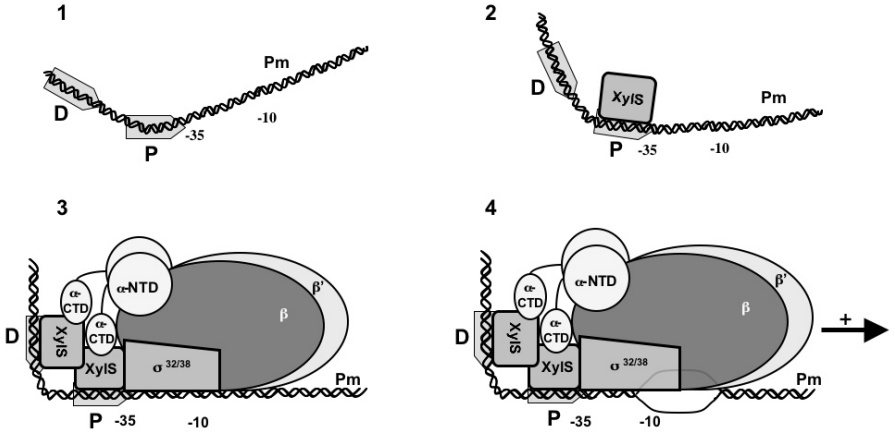



FIGURE 10. **Model of the sequential binding of XylS to Pm**. 1) Free DNA. The -10/-35 RNAP binding site and the two XylS binding sites (D and P) are depicted. The bending angle is 35º, centered in XylS proximal binding site. 2) A first XylS monomer binds Pm at the proximal site, shifts the bent center to DNA sequence between the two sites and increases the bending angle to 50º. 3) This change favors cooperative binding of a second monomer at the distal site. Contacts are established with RNAP through the α C-terminal domain and probably with σ subunit, 4) favoring open complex formation, and transcription initiation (24).

## REFERENCES


1. Ramos, J. L., Marqués, S., and Timmis, K. N. (1997) *Annu Rev Microbiol* **51**, 341-373
2. Gallegos, M. T., Marqués, S., and Ramos, J. L. (1996) *J Bacteriol* **178**, 2356-2361
3. González-Perez, M. M., Ramos, J. L., and Marqués, S. (2004) *J Bacteriol* **186**, 1898-1901
4. Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J. L. (1997) *Microbiol Mol Biol Rev* **61**, 393-410
5. Martin, R. G., and Rosner, J. L. (2001) *Curr Opin Microbiol* **4**, 132-137
6. Tobes, R., and Ramos, J. L. (2002) *Nucleic Acids Res* **30**, 318-321
7. Soisson, S. M., MacDougall-Shackleton, B., Schleif, R., and Wolberger, C. (1997) *J Mol Biol* **273**, 226-237
8. Michán, C. M., Busby, S. J., and Hyde, E. I. (1995) *Nucleic Acids Res* **23**, 1518-1523
9. Bustos, S. A., and Schleif, R. F. (1993) *Proc Natl Acad Sci USA* **90**, 5638-5642
10. Jair, K. W., Yu, X., Skarstad, K., Thony, B., Fujita, N., Ishihama, A., and Wolf, R. E., Jr. (1996) *J Bacteriol* **178**, 2507-2513
11. Martin, R. G., Jair, K. W., Wolf, R. E., Jr., and Rosner, J. L. (1996) *J Bacteriol* **178**, 2216-2223
12. Griffith, K. L., and Wolf, R. E., Jr. (2004) *J Mol Biol* **344**, 1-10
13. Martin, R. G., Gillette, W. K., Martin, N. I., and Rosner, J. L. (2002) *Mol Microbiol* **43**, 355-370
14. Grainger, D. C., Webster, C. L., Belyaeva, T. A., Hyde, E. I., and Busby, S. J. (2004) *Mol Microbiol* **51**, 1297-1309
15. Lobell, R. B., and Schleif, R. F. (1990) *Science* **250**, 528-532
16. Soisson, S. M., MacDougall-Shackleton, B., Schleif, R., and Wolberger, C. (1997) *Science* **276**, 421-425
17. Belyaeva, T. A., Wade, J. T., Webster, C. L., Howard, V. J., Thomas, M. S., Hyde, E. I., and Busby, S. J. (2000) *Mol Microbiol* **36**, 211-222
18. Egan, S. M., and Schleif, R. F. (1993) *J Mol Biol* **234**, 87-98
19. Shah, I. M., and Wolf, R. E., Jr. (2006) *Mol Microbiol* **60**, 199-208
20. Griffith, K. L., Shah, I. M., and Wolf, R. E., Jr. (2004) *Mol Microbiol* **51**, 1801-1816
21. Kaldalu, N., Toots, U., de Lorenzo, V., and Ustav, M. (2000) *J Bacteriol* **182**, 1118-1126
22. Ruíz, R., Marqués, S., and Ramos, J. L. (2003) *J Bacteriol* **185**, 3036-3041
23. Ruíz, R., and Ramos, J. L. (2002) *J Biol Chem* **277**, 7282-7286
24. Domínguez-Cuevas, P., Marín, P., Busby, S.J., Ramos, J. L., and Marqués, S. (2007) *Mol Microbiol, submitted.*
25. Timmes, A., Rodgers, M., and Schleif, R. (2004) *J Mol Biol* **340**, 731-738
26. Howard, V. J., Belyaeva, T. A., Busby, S. J., and Hyde, E. I. (2002) *Nucleic Acids Res* **30**, 2692-2700
27. Prouty, M. G., Osorio, C. R., and Klose, K. E. (2005) *Mol Microbiol* **58**, 1143-1156
28. Poore, C. A., Coker, C., Dattelbaum, J. D., and Mobley, H. L. (2001) *J Bacteriol* **183**, 4526-4535
29. Rhee, S., Martin, R. G., Rosner, J. L., and Davies, D. R. (1998) *Proc Natl Acad Sci USA* **95**, 10413-10418
30. González-Pérez, M. M., Ramos, J. L., Gallegos, M. T., and Marqués, S. (1999) *J Biol Chem* **274**, 2286-2290
31. Kaldalu, N., Mandel, T., and Ustav, M. (1996) *Mol Microbiol* **20**, 569-579
32. Domínguez-Cuevas, P., Marín, P., Marqués, S., and Ramos, J. L. (2007) *J Mol Biol, submitted.*
33. Bhende, P. M., and Egan, S. M. (1999) *J Bacteriol* **181**, 5185-5192
34. Jair, K. W., Martin, R. G., Rosner, J. L., Fujita, N., Ishihama, A., and Wolf, R. E., Jr. (1995) *J Bacteriol* **177**, 7100-7104
35. Munson, G. P., and Scott, J. R. (2000) *Mol Microbiol* **36**, 1391-1402
36. Tabor, S., and Richardson, C. C. (1985) *Proc Natl Acad Sci U S A* **82**, 1074-1078
37. González-Pérez, M. M., Marqués, S., Domínguez-Cuevas, P., and Ramos, J. L. (2002) *FEBS Lett* **519**, 117-122
38. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual.*, Cold spring Harbor, N.Y.
39. Terán, W., Felipe, A., Segura, A., Rojas, A., Ramos, J. L., and Gallegos, M. T. (2003) *Antimicrob Agents Chemother* **47**, 3067-3072
40. Thompson, J. F., and Landy, A. (1988) *Nucleic Acids Res* **16**, 9687-9705
41. Brunker, P., Hils, M., Altenbuchner, J., and Mattes, R. (1998) *Gene* **215**, 19-27
42. Egan, S. M., and Schleif, R. F. (1994) *J Mol Biol* **243**, 821-829
43. Caswell, R., Webster, C., and Busby, S. (1992) *Biochem J* **287 (Pt 2)**, 501-508
44. Kim, J., Zwieb, C., Wu, C., and Adhya, S. (1989) *Gene* **85**, 15-23
45. Chen, S., Iannolo, M., and Calvo, J. M. (2005) *J Mol Biol* **345**, 251-264
46. Sarai, A., and Kono, H. (2005) *Annu Rev Biophys Biomol Struct* **34**, 379-398
47. Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M., and Zhurkin, V. B. (1998) *Proc Natl Acad Sci USA* **95**, 11163-11168
48. Brunelle, A., and Schleif, R. (1989) *J Mol Biol* **209**, 607-622
49. Busby, S., and Ebright, R. H. (1997) *Mol Microbiol* **23**, 853-859
50. Ruíz, R., Ramos, J. L., and Egan, S. M. (2001) *FEBS Lett* **491**, 207-211
51. Ruíz, R., and Ramos, J. L. (2001) *Biochem Biophys Res Commun* **287**, 519-521
52. Straney, D. C., Straney, S. B., and Crothers, D. M. (1989) *J Mol Biol* **206**, 41-57
53. Harrington, R. E. (1992) *Mol Microbiol* **6**, 2549-2555
54. Kessler, B., Marques, S., Kohler, T., Ramos, J. L., Timmis, K. N., and de Lorenzo, V. (1994) *J Bacteriol* **176**, 5578-5582
55. Hanahan, D. (1983) *J Mol Biol* **166**, 557-580
56. Ramos, J. L., Stolz, A., Reineke, W., and Timmis, K. N. (1986) *Proc Natl Acad Sci USA* **83**, 8467-8471

CAPÍTULO 5

# Balance transcripcional entre programas metabólicos y de respuesta a estrés en células de *Pseudomonas putida* KT2440 expuestas a tolueno

**Patricia Domínguez Cuevas, José Eduardo González Pastor, Silvia Marqués, Juan L. Ramos y Víctor de Lorenzo**

Cuando *Pseudomonas putida* percibe tolueno en el medio de cultivo, lo hace simultáneamente como nutriente potencialmente metabolizable, como sustancia tóxica que produce daño en las membranas y que por lo tanto debe ser expulsada, y como agente desnaturalizante de macromoléculas como por ejemplo las proteínas. Cada una de estas señales desencadena una respuesta transcripcional específica que involucra un gran número de genes. En este trabajo hemos empleado microarrays de ADN para investigar la conexión entre estas respuestas en *P. putida* KT2440 tras una exposición de 15 minutos a tolueno. Hemos comparado dichos resultados con los obtenidos tras la exposición a *o*-xileno (análogo no metabolizable del tolueno) y 3MB (sustrato específico de la ruta TOL de rotura en *meta* de *P. putida* codificada en el plásmido pWW0). Los patrones de expresión resultantes sugieren que la mayor parte de la maquinaria transcripcional expresa genes relacionados con la respuesta a estrés, mientras que solamente una pequeña proporción se encarga de transcribir genes relacionados con la degradación de los compuestos aromáticos. De manera específica tanto el tolueno como el *o*-xileno inducen la expresión de las rutas TOL del plásmido pWW0. De forma análoga, el 3MB induce la expresión de la ruta *meta*, además de inducir una ruta adicional del metabolismo de benzoatos no substituidos, lo que podría ocasionar efectos deletéreos. Paralelamente estos compuestos aromáticos inhiben funciones relacionadas con la movilidad bacteriana en respuesta a la toxicidad de los aromáticos. Teniendo en cuenta nuestros resultados proponemos que el tolueno es detectado por *P. putida* KT2440 como un compuesto tóxico más que como nutriente, y que la inhibición de múltiples funciones en respuesta a los aromáticos es consecuencia del balance necesario para activar la resistencia al estrés asumiendo el menor coste energético posible.

*The Journal of Biological Chemistry*
Vol. **281**, Nº. 17, pp. 11981-11991, April 28, 2006

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 17, pp. 11981–11991, April 28, 2006


# Transcriptional Tradeoff between Metabolic and Stress-response Programs in *Pseudomonas putida* KT2440 Cells Exposed to Toluene*[S]

Received for publication, September 7, 2005, and in revised form, January 24, 2006 Published, JBC Papers in Press, February 22, 2006, DOI 10.1074/jbc.M509848200

**Patricia Domínguez-Cuevas[‡], José-Eduardo González-Pastor[§], Silvia Marqués[‡], Juan-Luis Ramos[‡1], and Víctor de Lorenzo[§¶]**

*From the [‡]Department of Biochemistry and Molecular and Cellular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, Profesor Albareda, 1, E-18008 Granada, Spain, [§]Instituto Nacional de Tecnica Aeroespacial-Consejo Superior de Investigaciones Científicas, Centro de Astrobiología, Torrejón de Ardoz, E-28850 Madrid, Spain, and [¶]Department of Microbiology, Centro Nacional de Biotecnología-Consejo Superior de Investigaciones Científicas, Cantoblanco, E-28049 Madrid, Spain*

**When *Pseudomonas putida* KT2440 cells encounter toluene in the growth medium, they perceive it simultaneously as a potential nutrient to be metabolized, as a membrane-damaging toxic drug to be extruded, and as a macromolecule-disrupting agent from which to protect proteins. Each of these inputs requires a dedicated transcriptional response that involves a large number of genes. We used DNA array technology to decipher the interplay between these responses in *P. putida* KT2440 subjected to a short challenge (15 min) with toluene. We then compared the results with those in cells exposed to *o*-xylene (a non-biodegradable toluene counterpart) and 3-methylbenzoate (a specific substrate of the lower TOL pathway of the *P. putida* pWW0 plasmid). The resulting expression profiles suggest that the bulk of the available transcriptional machinery is reassigned to endure general stress, whereas only a small share of the available machinery is redirected to the degradation of the aromatic compounds. Specifically, both toluene and *o*-xylene induce the TOL pathways and a dedicated but not always productive metabolic program. Similarly, 3-methylbenzoate induces the expression not only of the lower *meta* pathway but also of the non-productive and potentially deleterious genes for the metabolism of (nonsubstituted) benzoate. In addition, toluene (and to a lesser extent *o*-xylene) inhibit motility functions as an unequivocal response to aromatic toxicity. We argue that toluene is sensed by *P. putida* KT2440 as a stressor rather than as a nutrient and that the inhibition by the aromatic compounds of many functions we tested is the tradeoff for activating stress tolerance genes at a minimal cost in terms of energy.**

*Pseudomonas putida* KT2440 is a soil microorganism characterized by its metabolic versatility, which enables the strain to degrade a wide variety of natural and man-made aromatic compounds (1, 2). This strain can use toluene and *m*- and *p*-xylene via the pWW0 TOL plasmid-encoded pathways (3, 4). When the strain is confronted with these aromatic compounds, cells face an enticing paradox. On the one hand, these chemical species can be mineralized to yield carbon and energy for growth, allowing the strain to colonize niches refractory to other microbes. On the other hand, toluene and xylenes are toxic for the bacteria above a certain threshold, since they partition in the cell membrane and disorganize it by removing lipids and proteins, which eventually leads to cell death (5).

The mechanism of toluene, *m*-xylene, and *p*-xylene metabolism to $CO_2$ and $H_2O$ in *P. putida* bearing pWW0 involves two sets of reactions and depends on the ability of two cognate regulatory proteins to become activated by certain aromatic compounds and to trigger the expression of specific pathways (summarized in Fig. 1) (4). But activation of metabolic genes is not the only effect of exposing *P. putida* cells to aromatic hydrocarbons, as these compounds also trigger a solvent tolerance response which involves a number of mechanisms that are not yet fully understood (6, 7). However, a number of factors involved in the process have been characterized over the last 10 years (for review, see Ref. 6). Several laboratories have identified efflux pumps belonging to the RND[2] family as being involved in solvent tolerance (8–13), and the TtgABC efflux pump present in all strains that extrude several organic solvents and antibiotics is also present in *Pseudomonas putida* KT2440 (14, 15). Although extrusion of the solvent is probably the most important mechanism of solvent resistance, this phenomenon involves other factors such as membrane rigidification via *cis* to *trans* isomerization of the unsaturated fatty acids (16, 17) or the induction of a number of chaperones (18). Finally, exposure of *P. putida* cells to solvents and chaotropic agents triggers a heat-shock (HS) response that is likely to require a share of the available gene expression machinery of the cell (18, 19).

The issue at stake is, thus, how *P. putida* cells harboring a relatively stable transcriptional machinery are able to process all three different inputs from toluene; that is, a nutritional signal bound to trigger a large metabolic program, a toxic signal that triggers a solvent extrusion and tolerance response, and interference with the protein folding machinery leading to activation of the HS regulon. How the strain is able to manage all three tasks has not been examined because of the dearth of suitable technology (20). In this study we used a genome-wide DNA array setup to inspect in detail the immediate changes in global expression pattern of *P. putida* KT2440 in response to exposure to three different aromatic compounds; toluene, a metabolizable aromatic hydrocarbon (a carbon source and a stressor), *o*-xylene, a gratuitous counterpart known to induce the degradation pathway but which cannot be metabolized (21) (and, thus, acts only as a stressor), and 3MB, a specific inducer of the

* This work was supported by European Commission Grant QLK3-CT-2002-1923 and Plan Nacional de Biotecnología Grant GEN2001-4698-CO5-03 (to J.-L. R.) and by grants from the 5th and 6th Framework Programmes of the European Union (to V. d. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables 1–11.

[1] To whom correspondence should be addressed. Tel.: 34-958-181608 or -181600 (ext. 326); Fax: 34-958-135740; E-mail: jlramos@eez.csic.es.

[2] The abbreviations used are: RND, resistance nodulation cell division; HS, heat shock; 3MB, 3-methylbenzoate; RNAP, RNA polymerase; TIGR, The Institute of Genomic Research.

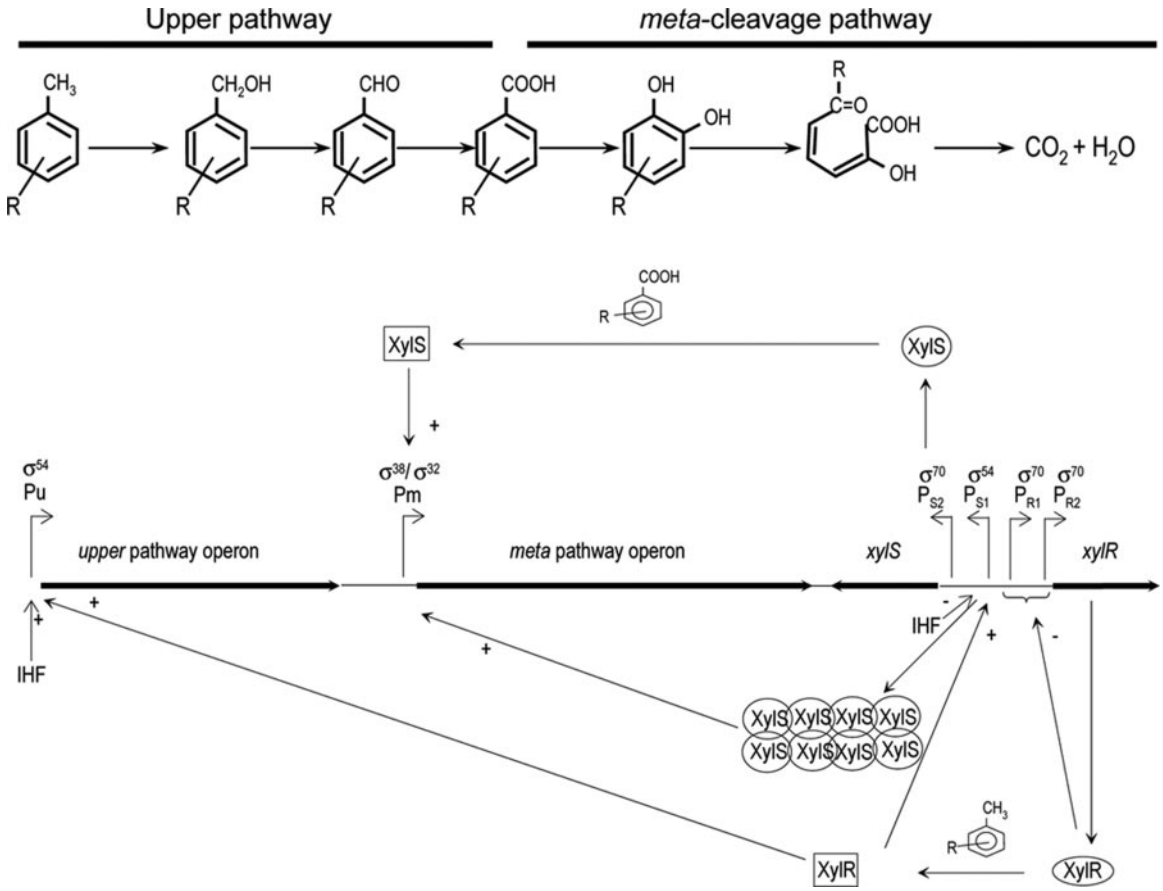



FIGURE 1. **Organization of the TOL operons of plasmid pWW0.** The plasmid encodes enzymatic machinery for the stepwise oxidation of one of the methyl groups of toluene, *m*-xylene and *p*-xylene (but not *o*-xylene), into the corresponding alkylbenzoates (encoded by the so-called upper operon, driven by the $\sigma^{54}$ Pu promoter). Oxidation is followed by the dioxygenation of this carboxylic acid and the *meta*-cleavage of the resulting (methyl)catechol all the way down to tricarboxylic acid cycle intermediates (encoded by the so-called *meta* or lower TOL pathway, driven by the Pm promoter). Expression of each of the gene clusters is subject to an intricate interplay between plasmid-encoded regulators and host factors in which the major players are the pWW0-encoded proteins XylR (an *m*-xylene-responsive $\sigma^{54}$-dependent activator acting on Pu and Ps) and XylS (an *m*-toluate-responsive AraC-type factor acting on Pm). Activation of the *meta* operon requires not only XylS and 3MB but also the concourse of either the heat-shock $\sigma^{32}$ or the stationary stress $\sigma^{38}$ of the host because the RNAP-$\sigma^{70}$ holoenzyme is unable to start transcription in Pm. Overproduction of XylS can also activate Pm in the absence of lower pathway effectors. XylR expression appears to be auto-regulated because the upstream activating sequence of the $\sigma^{54}$ Ps promoter overlaps the Pr promoter of *xylR*. IHF, integration host factor. The scheme summarizes a large number of studies based on enzyme activities, transcriptional fusions of the Pu and Pm promoter to reporter genes, and determination of the expression level of the leader mRNA in primer extension analyses (for review, see Ref. 4).

lower TOL pathway (primarily a C-source with lower toxicity). Our main objectives were to understand how the presence of all three compounds (*i.e.* of two contradictory signals) is differentially sensed, to detect the direct consequences of their presence, and to determine the short-term global outcome on the genetic program. Our results reveal that after exposure to aromatic compounds a strong and specific metabolic response at the level of aromatic pathways is concomitant with a general stress response at several levels, reflected by the induction of genes known to respond to membrane damage, oxidative stress, and misfolding of soluble proteins. The interplay between the HS response, the induction of catabolic pathways, the synthesis of efflux pumps, and the stimulation of amino acid biosynthesis resulted in distinct patterns of transcription under all three conditions. Interestingly, these patterns appear to reflect a shift of resources available to the transcriptional machinery from possibly more dispensable functions and toward these vital tasks. Furthermore, our data suggest that the regulatory network behind this metabolic and genetic interplay is not fortuitous but reflects how priorities in the use of the transcriptional machinery are established in different expression programs.

## EXPERIMENTAL PROCEDURES

*Bacterial Strains, Culture, and Growth Conditions—P. putida* KT2440 (pWW0) was grown on M9 minimal medium (21) with 25 mM citrate as a carbon source. 500-ml flasks containing 100 ml of this culture medium were inoculated with *P. putida* KT2440 to reach a turbidity of 0.03 and shaken at 30 °C and 200 rpm until the turbidity of the culture reached 0.7 ± 0.05 at 660 nm. Then the cultures were exposed to a 1 mM concentration of the different organic compounds, a concentration that neither inhibits cell growth nor affects cell viability. After 15 min, culture samples (15 ml) were harvested by centrifugation at 4 °C in tubes precooled in liquid nitrogen, and the cell pellets were immediately immersed in liquid nitrogen and stored at −80 °C. The $P_{benA}$::*lacZ* transcriptional fusion was constructed by creating an XbaI site downstream from the transcription start site of *P. putida* KT2440 *benA* gene and cloning a 0.7-kilobase EcoRV-XbaI fragment containing the $P_{benA}$ promoter between the BglII-XbaI cloning sites of the broad-host range vector pMP220 (22). The resulting reporter plasmid pMQ220EV was transferred to *P. putida* KT2440 with or without pWW0 plasmid by conjugation.

*Genomic DNA Microarrays of P. putida KT2440*—The genome-wide DNA chip used in this work has been described in detail (23). It consists of an array of 5539 oligonucleotides (50-mer) spotted in duplicate onto γ-aminosilane-treated slides and covalently linked with UV light and heat. The oligonucleotides represent the 5350 open reading frames annotated in the *P. putida* KT2440 genome (1) along with the 140 open reading frames defined for the TOL plasmid pWW0 (24) and a suite of commonly used reporter genes and antibiotic resistance markers. The

chips are also endowed with homogeneity controls consisting of oligonucleotides for the *rpoD* and *rpoN* genes spotted at 20 different positions as well as duplicate negative controls at 203 predefined positions (www.progenika.com).

*RNA Isolation*—RNA was extracted from cells with the hot acid-phenol isolation procedure (25). To this end the cell pellet was resuspended in 1.6 ml of LETS buffer prewarmed to 75 °C (100 mM LiCl, 10 mM EDTA, 1% (w/v) SDS, 10 mM Tris-HCl pH 8.0) and transferred to a prewarmed 15-ml tube containing 1 ml of glass beads (106-mm acid washed, Sigma) and 1.2 ml of acid phenol. The mixture was thoroughly vortexed for 3 min to complete cell lysis. Then 1.2 ml of chloroform was added, and tubes were vortexed for 30 s. The solution was centrifuged at 3200 × *g* at 4 °C for 10 min. Samples (1.6 ml) of the aqueous phase were recovered and transferred to a 15-ml tube containing 1.6 ml of acid phenol at 75 °C. The sample was thoroughly mixed for 3 min and centrifuged at 3200 × *g* for 10 min. Samples of the aqueous phase (2 × 0.7 ml) were recovered, precipitated with 0.7 ml of isopropanol at room temperature, and centrifuged at 12,000 × *g* for 25 min at 4 °C. The resulting pellet was washed with 1 ml of 75% (v/v) ethanol (4 °C), dried, and resuspended in 75 μl of diethyl pyrocarbonate/double distilled $H_2O$. Then 1.2 ml of TRIzol (Sigma) was added to each sample, and the mixture was vortexed for 15 s and left to stand for 5 min at room temperature. Then 0.24 ml of chloroform was added, and the mixture was shaken and left to stand for 3 min at room temperature. Finally, samples were centrifuged at 12,000 × *g* for 15 min (4 °C). The aqueous phase (0.7 ml) was recovered and precipitated with 0.7 ml of isopropanol. The pellet was washed with 1 ml of 75% (v/v) ethanol (4 °C), dried, and resuspended in 15–25 μl of diethyl pyrocarbonate/double distilled $H_2O$. Samples were heated to 55–60 °C for 5 min and stored at −80 °C. The integrity and purity of the RNA were checked with agarose gel electrophoresis to rule out any traces of DNA. The concentration of RNA was measured spectrophotometrically at 260 nm.

*Labeling and Hybridization Conditions*—cDNA was generated from RNA samples by direct incorporation of Cy3- or Cy5-labeled dUTP into cDNA. Differentially labeled samples from two different conditions were mixed and hybridized to the *P. putida* microarrays. For each experiment three independent biological replicates were done. For each reaction, RNA (30 μg) was incubated with 0.5 μg of random hexamers (Amersham Biosciences) at 70 °C for 10 min and then chilled on ice for 2 min. A mix containing 1× reverse transcription buffer, 5 mM $MgCl_2$, 10 mM dithiothreitol, deoxynucleotide triphosphates (0.5 mM dATP, 0.5 mM dGTP, 0.5 mM dCTP, and 0.2 mM dTTP), 40 units of RNase OUT (Invitrogen), and either Cy3-dUTP or Cy5-dUTP (Amersham Biosciences) was added to the RNA primer mixture and incubated at 25 °C for 5 min. Superscript II reverse transcriptase (200 units, Invitrogen) was added, and the mixture was incubated at 25 °C for 10 min and then at 42 °C for 120 min. The reaction was stopped by heating the reaction mixture at 70 °C for 15 min. The RNA was then digested with RNase H (2 units, Invitrogen) at 37 °C for 20 min. Unincorporated nucleotides were removed through QiaQuick purification spin columns (Qiagen). The purified probes were mixed, dried, and reconstituted in 30 μl of hybridization buffer (3× SSC (1× SSC = 0.15 M NaCl and 0.015 M sodium citrate), 5× Denhardt's solution, 0.2% (w/v) SDS, 5% dextran sulfate, 50% (v/v) formamide). Before hybridization, the probe was heated at 95 °C for 2 min, applied directly onto the microarray slide, and covered with a glass coverslip (24 × 60 mm). The array was hybridized overnight at 50 °C in a humidified hybridization chamber. After hybridization, the slides were washed for 5 min in 2× SSC and 0.1% (w/v) SDS (final volume 500 ml) followed by two 5-min washes in 50 ml of 1× SSC and 0.1× SSC. Before scanning, the slides were spun at 1600 × *g* for 5 min to remove residual salts. Finally, slides were scanned on a GMS 418 apparatus (Genetic Microsystems, Inc.). Separate images were acquired for Cy3 and Cy5 and then processed and analyzed with GenePix 4.0 software (Axon Instruments, Inc.).

*Data Analysis*—After background subtraction, signal intensities for each replica were normalized and statistically analyzed using the Lowess intensity-dependent normalization method (26) included in the Almazen System software (Alma Bioinformatics S.L.). *p* values were calculated with Student's *t* test algorithm based on the differences between log 2 ratio values for each biological replicate. Genes were considered differentially expressed when they fulfilled the filter parameters of expression ratio ≥1.8 and *p* value ≤0.2. Cluster and Treeview software (27) were used to group and visualize genes whose expression varied in the same direction in response to all three aromatic compounds. Then a hierarchical clustering algorithm based on the average linkage method was used to identify genes that were expressed differentially under all three experimental conditions.

*Western Blots—P. putida* KT2440 (pWW0) was grown as described under "Bacterial Strains, Culture, and Growth Conditions"; after 15 min of challenge with aromatic compounds the cells (100 ml) were harvested by centrifugation, and the pellets were resuspended, washed once in 1× M9 buffer, and resuspended again in 2 ml of lysis buffer containing Tris 50 mM, pH 7.5, 50 mM NaCl, 2 mM EDTA, 4 mM β-mercaptoethanol, and 1× Complete™ protease inhibitor mixture (Roche Applied Science). Cells were lysed by sonication, and the insoluble fraction was discarded by centrifugation at 18,000 × *g* for 20 min. Aliquot fractions of the supernatant containing 70 μg of total protein were analyzed on SDS-PAGE (12.5% (w/v)) and transferred to a nitrocellulose membrane. The membranes were blocked for 3 h at room temperature with 5% nonfat dry milk in phosphate-buffered saline. Membranes were incubated overnight at 4 °C with monoclonal antibodies against *Escherichia coli* $\sigma^{32}$ subunit (Neoclone), diluted 1:1000, washed with phosphate-buffered saline solution, and incubated with goat anti-mouse immunoglobulin G+L conjugated with horseradish peroxidase (1:1000 dilution) for 1 h (Caltag Laboratories). The blots were developed with the Super-Signal® West Dura Extended Duration Substrate (Pierce). Chemiluminescent blots were exposed to autoradiographic film for 30 s to 2 min.

*Swimming Assay*—The medium used for assays was M9 citrate that contained 0.3% (w/v) agar and 0.05% (w/v) tetrazolium red. Swim plates were inoculated with bacteria from an overnight culture in M9 citrate agar (1.2% (w/v)) plates with a sterile toothpick. The plates were kept for 14–16 h at 30 °C in humidified containers to prevent dehydration. Plates were then photographed.

*Hydrogen Peroxide Production—P. putida* KT2440 (pWW0) was grown on M9 with citrate as a carbon source to a turbidity of 0.7 at 660 nm. The cultures were then exposed to 5 mM concentrations of different organic compounds. After 5 and 30 min of exposure, 10 ml of the culture was harvested by centrifugation, and the cells were broken as described above. We used 50 μl of the cell-free extract to determine the hydrogen peroxide concentration as described in Buege and Aust (28).

*β-Galactosidase Assays—P. putida* KT2440 (pMQ220EV) cells bearing or not bearing plasmid pWW0 were grown overnight on M9 citrate and diluted 1/100 in fresh medium, and cultures were divided in four aliquot fractions (10 ml each). After 1 h at 30 °C, one of the fractions was supplemented with 1 mM 3MB, another one with toluene, and a third one with *o*-xylene. The remaining fraction was kept without the addition of aromatics. Samples for β-galactosidase activity assays were taken 3 h after induction. β-Galactosidase activity was determined in permeabilized whole cells according to Miller (29).

*Supplemental Material*—Complete microarray datasets are available as Microsoft Excel files. Supplemental Tables 1–11 show the most relevant changes.

## RESULTS

*P. putida KT2440 Genome-wide Microarrays Reveal a Massive Response to Aromatic Compounds*—The DNA microarrays described previously (23) were used to study the global expression profile of *P. putida* (pWW0) in response to metabolizable (toluene and 3MB) and nonmetabolizable (*o*-xylene) aromatic compounds. *P. putida* (pWW0) grows in M9 minimal medium with citrate as the sole carbon source with a duplication time of about 70 ± 4 min. When cultures reached the mid-exponential growth phase (turbidity at 660 nm 0.7 ± 0.05) they were split into four aliquots. One was kept as a control, and 1 mM toluene, *o*-xylene, or 3MB was added to the other three. This concentration of aromatic compounds was sublethal, and 100% of the cells survived the addition of these chemicals. To analyze the early events in the response to these compounds, samples were collected 15 min after the addition of the tested aromatics, and total RNA was prepared from all cultures.

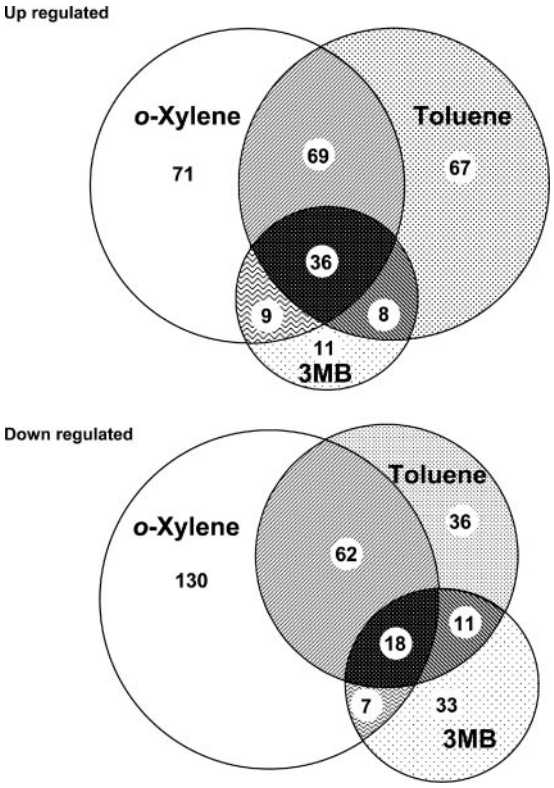



FIGURE 2. **Venn diagram showing the overlap between up-regulated or down-regulated genes in the presence of each aromatic compound.** The area in each *circle* is proportional to the number of up- or down-regulated genes in the corresponding condition(s).

Our statistical analysis of the microarray data indicates low inter-experiment variance between replicates. On the basis of our data, we selected changes of 1.8-fold or more in mRNA levels between aromatic compound-treated and untreated preparations as significant. A total of 180, 185, and 64 genes were significantly up-regulated in response to toluene, *o*-xylene, and 3MB respectively, with *o*-xylene and toluene sharing 69 up-regulated genes (Fig. 2). On the other hand, 127, 217, and 69 genes were significantly down-regulated in toluene-, *o*-xylene-, and 3MB-treated cultures. All three treatments produced a similar down-regulation in 18 genes (Fig. 2), all of them related to membrane functions or energy transduction. Most up-regulated genes shared by all three aromatic treatments were related to the heat-sock response and general metabolism.

These genes were grouped as proposed previously according to the metabolic function of the corresponding products (30), *e.g.* metabolism of aromatic compounds, gene regulation, stress response, energy metabolism, transport of metabolites, chemotaxis, etc. (Fig. 3; see also below and supplemental Tables 1–11). Both up-regulated and down-regulated genes were found in most physiological groups with the exception of those genes grouped under "energy metabolism," "pili-flagella," and "protein metabolism," where the presence of any of the aromatics caused repression in all cases. For most groups, *o*-xylene had the most pronounced effect. These general effects were concomitant with HS response, general stress endurance, and peripheral pathways for aromatic compound metabolism, among others (Fig. 3). These results and their significance are analyzed in detail below.

*Specific Metabolic Response to Aromatics; Induction of Distinct Segments of the Aromatic Degradation Pathways*—Of all the known effects caused by toluene, *o*-xylene, and 3MB on *P. putida* KT2440 (pWW0),

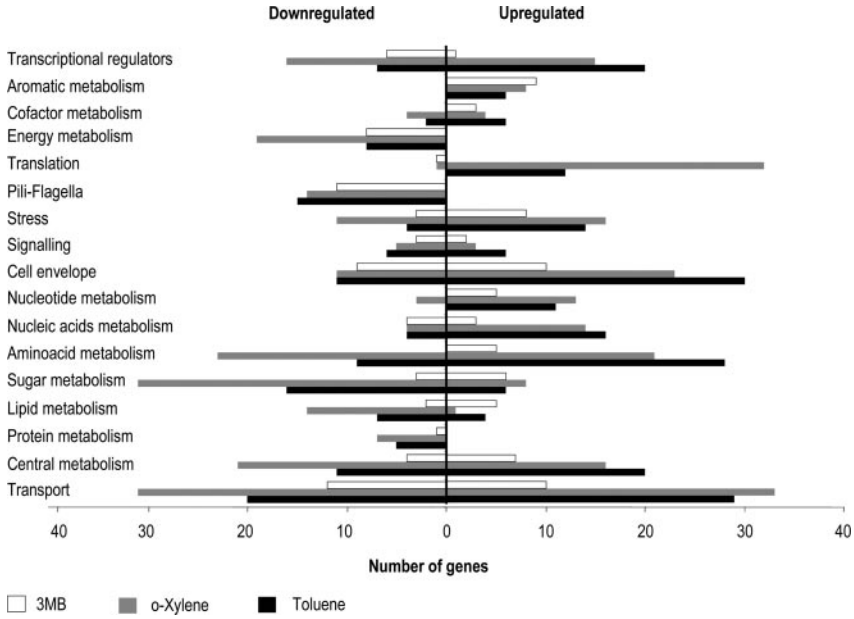



FIGURE 3. **Breakdown of responses of *P. putida* KT2440 to toluene, *o*-xylene, and 3MB according to functional categories.** Each plot indicates the type of physiological roles and the number of genes whose expression increased or decreased in response to exposure to the different chemicals for 15 min (see supplemental Tables 1–11). The overall score can be taken as a descriptor of the engagement of the roaming transcriptional machinery with promoters that express different functions.

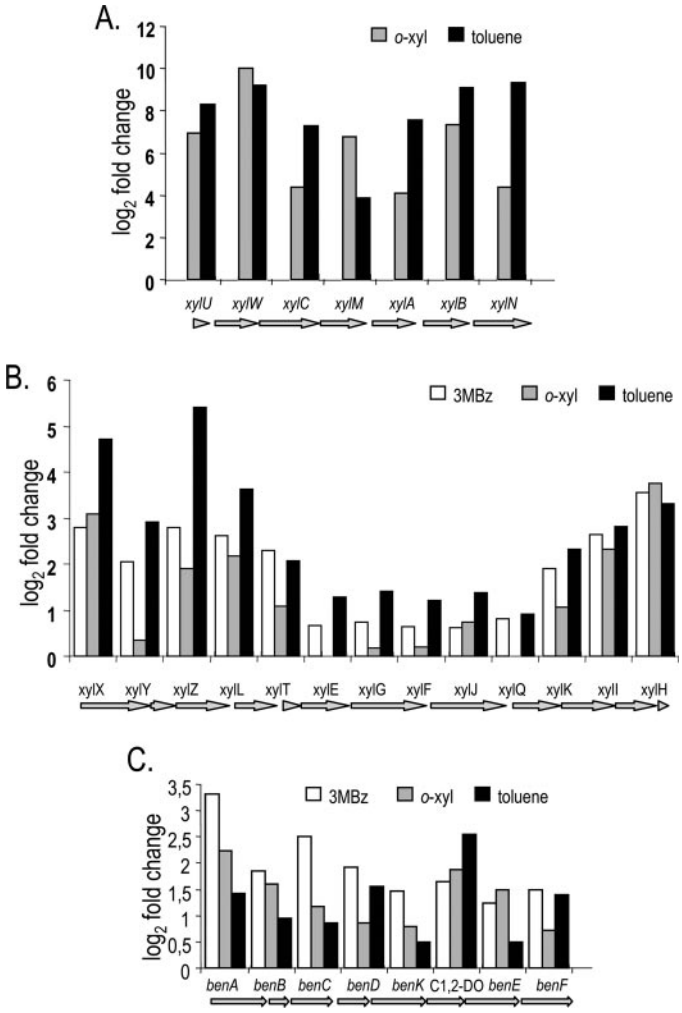



FIGURE 4. **Response of *xyl* and *ben* genes of *P. putida* KT2440 to TOL pathway effectors.** The plots show the induction of *xyl* genes of the upper operon of the TOL plasmid (*A*), *xyl* genes of the *meta* operon of the TOL plasmid (*B*), and chromosomal *ben* genes for benzoate metabolism (*C*). Cells were treated as explained under "Experimental Procedures," and the relative expression profile of the upper pathway genes was determined in cells exposed to each of the inducers indicated *versus* untreated samples.

the best documented is induction of the promoters of TOL operons through two regulatory proteins specifically responding to the presence of aromatics (Fig. 1). We, therefore, used the output of *xyl* genes as a positive control to assess the performance of the array and then as an internal validation control. Fig. 4 highlights the changes in the expression of catabolic genes involved in the reaction of *P. putida* KT2440 (pWW0) to the three aromatic inducers we compared (see also supplemental Table 11). As expected, the whole of the TOL upper pathway was induced in response to both toluene and *o*-xylene. Although not identical, the results with the metabolizable substrate toluene were similar to those with *o*-xylene, which cannot serve as a substrate, although it is able to induce expression of the *xyl* genes by activating the XylR regulatory protein (21). Expression of the genes for the *meta*-cleavage pathway was also detected in response to both aromatic hydrocarbons and to 3MB (Fig. 4*B*). The differences in activation of the *meta*-cleavage pathway in response to toluene and *o*-xylene should be considered in terms of metabolism; whereas toluene yields benzoate which activates the XylS *meta*-cleavage regulator, *o*-xylene cannot be metabolized, and internal production of the cognate XylS effector is prevented (Fig. 1). The low levels of *meta*-cleavage pathway induction with this aromatic

**TABLE 1**

**Effector-induced $P_{benA}$ expression in different genetic backgrounds**

The indicated strains were grown as described under "Experimental Procedures" until the mid-exponential phase was reached. β-Galactosidase activity (given in Miller's units) was determined in permeabilized cells as described previously (29). Values represent the average of three independent assays with an S.D. <10% of the given values.

| *P. putida* strain | Genotype | | β-Galactosidase effector | | | |
|---|---|---|---|---|---|---|
| | *benR* | *xylS* | Control | 3MB | Toluene | *o*-Xylene |
| | | | *Miller Units* | | | |
| KT2440::Δ*benR* | − | − | 35 | 125 | 25 | 25 |
| KT2440 | + | − | 30 | 670 | 30 | 25 |
| KT2440::Δ*benR* (pWW0) | − | + | 50 | 3500 | 2900 | 2100 |
| KT2440 (pWW0) | + | + | 30 | 3850 | 3550 | 2900 |

are most probably due to the so-called cascade activation through XylS overproduction (4) (Fig. 1).

Interestingly, the global analysis allowed us to detect differences in the induction levels of the different regions of the upper and *meta* operons (Fig. 4, *A* and *B*), suggesting additional regulatory steps in TOL gene expression regulation, *e.g.* changes in mRNA stability of different segments of the operon, which is transcribed as a single polycistronic mRNA (31, 32). Yet the results summarized in Fig. 4, *A* and *B*, not only confirmed data reported in the literature in the past years but also new features of this pathway.

The *benABCDKEZF* operon, present in the *P. putida* chromosome, specifies a conserved *ortho*-cleavage route for benzoate metabolism, which coexists with TOL plasmid pathways. Our data showed that 3MB, whose metabolism via the chromosomal *ortho*-cleavage pathway is nonproductive and leads to the accumulation of dead-end products (33), significantly induced this pathway (Fig. 4*C*). Toluene, which gives rise to unsubstituted benzoate through the action of the upper TOL pathway enzymes, induced the chromosomal *ben* route to lower levels than the induction caused by *o*-xylene, which is unable to act as a substrate for the TOL upper pathway enzymes. Our interpretation of this result is that XylS, which is overproduced in the presence of *o*-xylene, is responsible for this induction. The pWW0-encoded XylS regulator is highly homologous to the chromosomally encoded BenR protein, to the point that both regulators show a degree of cross-regulation (34, 35). To further test this hypothesis we constructed a transcriptional $P_{benA}$::′*lacZ* fusion and determined β-galactosidase levels in *benR*-proficient and *benR*-deficient backgrounds in *P. putida* with or without the TOL plasmid, which provides the *xylS* gene. We found that 3MB was recognized by BenR but induced higher levels of β-galactosidase expression from $P_{benA}$ when *xylS* was provided on the TOL plasmid (Table 1). In the presence of either toluene or *o*-xylene, we detected β-galactosidase expression only in the XylS-proficient backgrounds irrespective of the presence of BenR, a finding that supports the cross-regulation hypothesis.

Because both Pm and $P_{benA}$ are induced by 3MB and benzoate, cross-activation is translated into the simultaneous induction of competing pathways. The -fold induction of the TOL *xylXYZ* genes (encoding toluate dioxygenase) by 3MB is similar to that of the chromosomal *benABC* genes (Fig. 4, *B* and *C*). Enzyme affinities and activities not withstanding, the fact is that 3MB can enter either of these two routes, although biodegradation only becomes productive when the plasmid-encoded enzymes are involved.

All three regulators involved in TOL and *ben* pathway expression are constitutively present in the cell at low levels (4, 21, 32, 34). Thus, direct specific sensing of the aromatics may occur as a result of their physical interaction with these regulators to trigger the synthesis of a whole set of new enzymes. Although this would represent a considerable waste of

energy (36), it would also allow an efficient use of the cell resources. Nevertheless, as for other known systems, genes induced for the utilization of new carbon sources constitute only a small fraction (5%) of the entire transcriptome reprogramming.

*The Cell Envelope Is an Early Sensor of Aromatic Toxicity*—The first defense against environmental stresses is the bacterial cell envelope, which tackles and senses toxic aromatics before they contact specific pathway regulators. Accordingly, in terms of the number of genes that change their expression pattern, the strongest response of the cell to aromatics is observed in functions related to the cell barrier; the cell envelope, lipid metabolism, transport, and pili-flagella synthesis accumulate a total of 124 genes that change their expression level (30 up-regulated and 94 down-regulated). Furthermore, 16 of 18 genes that were down-regulated in the presence of each of the three aromatics belong to this group (Fig. 2). We and others have shown before that *P. putida* cells exposed to a sudden shock of toxic chemicals maintain a balance between membrane fluidity and permeability by changing the ratio of *cis* to *trans* unsaturated fatty acids (37). Our genome-wide analysis revealed that genes for enzymes involved in fatty acid and phospholipid metabolism, particularly FadA, FadB, and $\beta$-ketothiolase, were repressed in the presence of toluene and *o*-xylene (see supplemental Table 1). In addition, we found that CfaB (PP5365), the major cyclopropane fatty acid synthetase, behaved differently in response to different inducers. CfaB was recently identified as the major cyclic fatty acid synthase (38). With *o*-xylene, *cfaB* expression was reduced 4-fold, whereas with toluene and benzoate no significant change was found. This suggests that *o*-xylene is detected by the cells as the most toxic compound among those tested in this study. In accordance with this hypothesis we found that certain genes related to fatty acid composition (*phaG*, PP4030, and PP4975) were repressed in response to *o*-xylene but not to toluene or 3MB.

A number of membrane-related cell activities such as the synthesis of porins, peptidoglycans, and flagella (motor and hook proteins) were repressed as a consequence of the stress imposed by each of the three aromatic compounds tested (supplemental Table 1). In fact, unlike the situation in other functional groups, all motility-chemotaxis genes and pili-flagella-related genes influenced by the presence of the aromatics were repressed, and in no case was a change toward activation detected within this group. To confirm the actual consequence of this apparent absolute switch-off of motility on pili and flagella functionality, swimming motility assays were carried out in the presence and absence of each of the three aromatics. Fig. 5 clearly shows a diminished swimming halo when cells were incubated in the presence of the tested compounds. The effects were more pronounced with toluene, although *o*-xylene and 3MB also produced a quantifiable response.

The assembly and functioning of molecular machines for motility consume large amounts of ATP but do not provide traits essential for survival under these conditions, so repression could be a way to save energy for more useful stress endurance programs. By way of comparison, it has been estimated that the flagellum alone can represent as much as 8% of the total cell protein when flagellar operons are expressed optimally. Flagellar genes and chemotaxis are repressed in other organisms in response to a number of stresses, including osmotic or saline shock, alcohol, pH, and temperature upshift (39–43). The repression of a variety of genes involved in chemotaxis, pili biosynthesis, the flagellar basal body and flagellar assembly proteins (supplemental Table 2), thus, has precedents in other systems. The signals that trigger this response could involve the dinucleotide dGMPc, which has been shown to regulate cell transition to a sessile state after a number of stresses (44). The conserved GGDEF motif proteins are key elements in the synthesis of

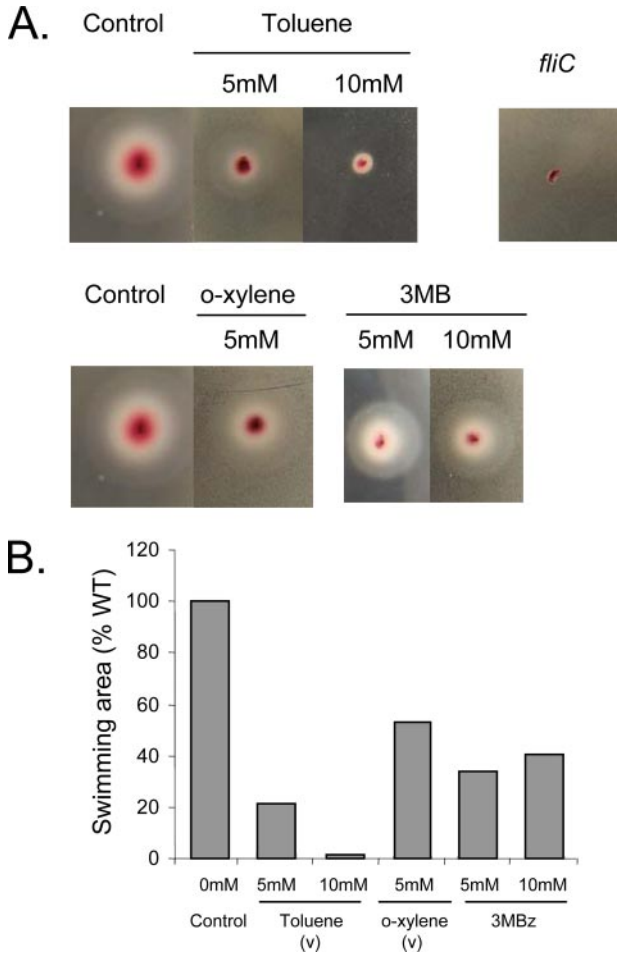



FIGURE 5. **Swimming motility of *P. putida* KT2440 on semisolid agar plates in the presence of toluene, *o*-xylene, or 3MB.** Cells were inoculated with a toothpick from an overnight M9 agar plate onto a swim plate (M9 citrate medium plus 0.3% (w/v) agar) with tetrazolium red and containing the indicated concentration of the corresponding aromatic compound. After 16 h of incubation at 30 °C, plates were photographed. The non-motile *P. putida* KT2440 *fliC* mutant was used as a nonswimming control (*A*). Three independent assays were run, and the average area of the halo of each strain was calculated and presented as a percentage of the control with no aromatics added. S.D. was below 2% of the given values (*B*). *WT*, wild type.

this messenger molecule in response to environmental signals (45). Interestingly, up to six GGDEF proteins (30% of all GGDEF-domain proteins annotated in *P. putida*) were induced more than 2-fold in response to the presence of all three aromatics (supplemental Table 5). We also observed a clear inhibition of the *parA* gene (the product of which mediates chromosome and plasmid segregation; supplemental Table 8) and of various penicillin-binding proteins (*pbp* genes), membrane proteins which participate in cell division. That these functions are all inhibited to some degree in cells exposed to the aromatics we tested could reflect a transient slowdown in cell division rate after sudden exposure to the aromatic chemicals, as is generally observed in response to other oxidative stress conditions (46). Further support for the idea of membrane damage is the finding that the expression of several extracytoplasmic $\sigma$ factors known to sense stress at the membrane and periplasm level clearly changed in the presence of *o*-xylene (supplemental Table 6A).

We observed the induction of a set of genes corresponding to various RND efflux pump operons involved in the extrusion of toxic chemicals (6) as a defense mechanism against the toxicity of toluene and *o*-xylene (see also supplemental Tables 4 and 7). The induction of the *ttgABC* operon (PP1384 through to PP1387) had been previously described in *P.*

**TABLE 2**
**Changes in genes encoding efflux pump transporters**

| TIGR accession no. | Description | -Fold change | | |
|---|---|---|---|---|
| | | Toluene | *o*-Xylene | 3MB |
| PP1516 | RND membrane fusion protein | 2.41 | 3.24 | |
| PP1517 | RND efflux transporter | | 1.73 | |
| PP3342 | NikA-nickel ABC transporter periplasmic nickel-binding protein | 1.84 | 1.82 | |
| PP3349 | Major facilitator family transporter | 9.54 | 4.61 | |
| PP3588 | Multidrug resistance transporter Bcr/CflA family | 4.35 | 5.59 | 5.64 |
| PP1387 | Transcriptional regulator TtgR | 3.00 | 3.48 | |
| PP1386 | TtgA solvent RND membrane fusion protein | 1.83 | | 1.74 |
| PP1385 | TtgB solvent RND transporter TtgB | | | 2.22 |
| PP1384 | TtgC solvent RND outer membrane protein TtgC | 2.25 | 1.65 | 2.47 |
| PP0958 | Ttg2A-toluene tolerance ABC efflux transporter ATP binding | 2.33 | 2.01 | |

**TABLE 3**
**Changes in enzymes related to Krebs cycle feeding and functioning**

| TIGR accession no. | Gene | Description | -Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-Xylene | 3MB |
| PP1022 | *zwf-1* | Glucose-6-phosphate-1-dehydrogenase | 2.13 | 2.11 | 1.67 |
| PP1024 | *eda* | 2-Dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase | | | −2.43 |
| PP0213 | *gabD* | Succinate-semialdehyde dehydrogenase | −2.2 | −2.01 | |
| PP0338 | *aceF* | Pyruvate dehydrogenase dihydrolipoamide acetyltransferase | −2.06 | | |
| PP0339 | *aceE* | Pyruvate dehydrogenase E1 component | −1.84 | | −2.16 |
| PP0356 | *glcB* | Malate synthase | −2.06 | −2.36 | |
| PP0654 | *mdh* | Malate dehydrogenase | −1.88 | −2.01 | |

**TABLE 4**
***P. putida* KT2440 (pWW0) hydrogen peroxide production after a 5- or 30-min challenge with aromatic compounds**
Details for determination of $H_2O_2$ are given under "Experimental Procedures."

| Treatment | $H_2O_2$ accumulation | |
|---|---|---|
| | 5 min | 30 min |
| | $\mu M$ | |
| None | 0.39 | 0.51 |
| Toluene | 1.84 | 2.98 |
| *o*-Xylene | 2.60 | 4.31 |
| 3MB | 2.73 | 4.06 |

*putida* KT2440, but this study extends the finding to other efflux pumps such as TtgA2 (PP0950), PP3349, and PP3588 (Table 2) as well as to the pump encoded by the operon that encodes PP1516 and PP1517. The -fold induction of these newly discovered pumps was found to be much higher than *ttgABC* induction levels.

*Aromatic-induced Membrane Damage Generates Oxidative Stress*—As shown above, a relevant noxious effect of toluene and *o*-xylene occurs at the membrane level. The disruption of membrane functions and the ensuing impairment in electron transfer may lead to the generation of active oxygen species in the respiratory chains. That oxidative damage did occur is revealed by two distinct sets of genes. First, *o*-xylene clearly reduced the numbers of transcripts for components of the membrane-bound electron transfer chain (PP1319, PP4264, PP4651) (see supplemental Table 3). This general repression was concomitant with the onset transcription for the synthesis of glutathione (*gshA*, PP0243) and glutathione *S*-transferase (*gst*, PP2023), which are generally involved in the metabolic detoxification of xenobiotics (19, 47). On the other hand the effects of all three inducers on the core step in aerobic metabolism (the Krebs cycle) showed that a number of enzymes involved in metabolic feeding (pyruvate dehydrogenase complex proteins PP0338 and PP0339, glyceraldehyde-3-phosphate dehydrogenase PP3443, and pyruvate kinase PP1362) were repressed when cells were exposed to toluene or *o*-xylene but not when they were exposed to 3MB (Table 3). Thus, it is likely that the sensing of membrane disruption induced by organic solvents is eventually translated into a slowdown in the synthesis of tricarboxylic acid cycle enzymes, which is only compensated for when the toxic compound can be used as a carbon source. In fact, several tricarboxylic acid enzymes such as succinic semialdehyde dehydrogenase, malate dehydrogenase (*mdh*, PP0654), or isocitrate lyase (PP4116) were also repressed (the latter only in response to *o*-xylene) (Table 3).

To experimentally confirm that oxidative stress was generated by the presence of the aromatic compounds, we followed hydrogen peroxide production after challenging *P. putida* with each of the three aromatic compounds. Table 4 shows a consistent increase in hydrogen peroxide production after brief exposure to each of these chemicals.

Further evidence that exposure to aromatic inducers caused oxidative damage came from the inhibition of iron-acquisition functions. The *fpvA* gene (PP4217), which codes for the iron-pyoverdine complex receptor, was tuned down in the presence of 3MB, like other TonB-dependent receptors (PP3612). It has been suggested that a decrease in iron acquisition may help compensate for the oxidative stress generated by exposure to the aromatics we tested (48).

Finally, we observed that several components of the OxyR regulon, known to be involved in the response to oxidative damage in *P. aeruginosa* and other bacteria, were substantially induced under our conditions. This was the case for the two annotated alkylhydroperoxidases (PP2439 and PP2440) and a new one identified as *ahpB* (PP1084) (Table 5) (49). In this connection it is worth noting that most of the changes we found in DNA metabolism related to genes that are associated with a defense response against oxidative stress (*e.g.* PP1974, PP5284; see supplemental Table 7) (50).

In summary, it appears that membrane damage is translated into oxidative stress, which is checked by slowing down or inhibiting the most productive step of aerobic metabolism, *i.e.* the Krebs cycle. In the short term this will inevitably cause a loss of energy, which must be compensated for somehow.

*Toluene, o-Xylene, and 3MB Induce Different Types of Heat-shock Responses*—Fig. 3 shows that one major change in response to the addition of the aromatic compounds tested here was the induction of a large number of genes involved in HS response, although the number of genes

**TABLE 5**
**Changes in genes encoding proteins related to stress response**

| TIGR accession no. | Gene | Description | -Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-Xylene | 3MB |
| PP0089 | *osmC* | Osmotically inducible protein OsmC | 3.72 | 3.11 | 1 |
| PP0252 | | Chaperonin, 33 kDa | 3.02 | 3.67 | 1.15 |
| PP0636 | | Cold-shock DNA-binding domain protein | 2.21 | 2.01 | −1.4 |
| PP1225 | | Radical activating enzyme | 1.2 | 3.95 | 1 |
| PP1254 | *xenA* | Xenobiotic reductase A | 1.2 | 2.08 | −1.75 |
| PP1360 | *groES* | Chaperonin, 10 kDa | 1 | 1.50 | 1.45 |
| PP1361 | *groEL* | Chaperonin, 60 kDa | 1.1 | 1.83 | 1.24 |
| PP1443 | *lon-1* | ATP-dependent protease La | 1.3 | 3.85 | 1.45 |
| PP1982 | *ibpA* | Heat-shock protein IbpA | 3.77 | 5.63 | 6.46 |
| PP2299 | *tig* | Trigger factor | 2.24 | 2.04 | −1.2 |
| PP2302 | *lon-2* | ATP-dependent protease La | 2.26 | 2.13 | 1.49 |
| PP4178 | | Dienelactone hydrolase family protein | 2.26 | 3.78 | 1.84 |
| PP4179 | *htpG* | Heat-shock protein HtpG | 2.26 | 4.81 | 2.25 |
| PP4180 | | Conserved hypothetical protein | 2.28 | 5.57 | 1.88 |
| PP4726 | *dnaJ* | DnaJ protein | 2.73 | 4.62 | 1.54 |
| PP4727 | *dnaK* | DnaK protein | 1.64 | 6.43 | 1.86 |
| PP4728 | *grpE* | Heat-shock protein GrpE | 2.31 | 5.87 | 2.08 |
| PP5000 | *hslV* | Heat-shock protein HslV | 1.71 | 4.95 | 2.08 |
| PP5001 | *hslU* | Heat-shock protein HslVU ATPase subunit HslU | 4.38 | 12.03 | 2.58 |
| PP1478 | | Xenobiotic reductase putative | −4.12 | −3.82 | |
| PP2187 | | Universal stress protein family | −2.27 | −9.87 | −1.6 |
| PP3070 | *ppiC-1* | Peptidyl-prolyl *cis-trans*-isomerase C | −2.48 | −3.06 | |
| PP3234 | | Heat-shock protein HSP20 family | −47.05 | −14.84 | −1.79 |

and the -fold induction were different with each compound (Fig. 3). In general, the intensity of the HS response (as judged from the number of genes involved and their expression) followed the order *o*-xylene > toluene ≫ 3MB, which may reflect the relative toxicity of each of these chemicals. In fact, *o*-xylene and toluene are known to be more toxic than 3MB, as they can dissolve in the membrane, disrupt it, and release lipids and proteins (5). It, thus, makes sense that chaperones involved in the HS response be massively produced to compensate for this damage. As expected, proteins induced by the aromatics we tested belong to the RpoH regulon family, which comprises GroES, GroEL, GrpE, Lon protease, DraJ, DnaK, IbpA, and other proteins. The strongest induction level (12-fold) was observed for HslU (PP5001) followed by DnaK (PP4727), IbpA (PP1982), and GrpE (PP4728), which were induced about 6-fold (see supplemental Table 7). All these proteins reflect the presence of misfolded proteins in the cytoplasm and constitute a signal in cell aromatic sensing. The HS response entails an increase in the level of $\sigma^{32}$, which is translated in the transcription of $\sigma^{32}$-dependent promoters. The increase in $\sigma^{32}$ protein is not regulated at the transcriptional level and depends on $\sigma^{32}$ protein stabilization and protection against degradation (51). To confirm that the aromatic compounds we tested induced the HS response, cell-free extracts of *P. putida* grown in the presence or absence of each aromatic compound were electrophoretically separated and analyzed by Western blot with an anti-$\sigma^{32}$ antibody. Fig. 6 shows an increase in $\sigma^{32}$ protein in extracts from cells exposed to the aromatics in comparison to control extracts, thus confirming stabilization of this key $\sigma$ factor (52, 53).

That the harm to cells caused by toluene and *o*-xylene differed from that caused by 3MB was evidenced by the smaller –fold change generally found with the latter and by the behavior of *xenA* (PP1254), PP0636 (a member of the cold-shock response), and the tig-trigger factor (PP2299), which were induced by about 2-fold by the aromatic hydrocarbons but repressed by 3MB (Table 5). In the solvent-tolerant strain *P. putida* DOT-T1E, *xenA* together with *cspA* gene products are involved in this trait (18). Their induction in *P. putida* KT2440 (pWW0) upon exposure to *o*-xylene or toluene not only confirms proteomic results for another *P. putida* strain (18) but also corroborates that these gene products are involved in the response to toluene. A few other chaperones such as the universal stress protein (PP2187), XenB (PP1478), and the

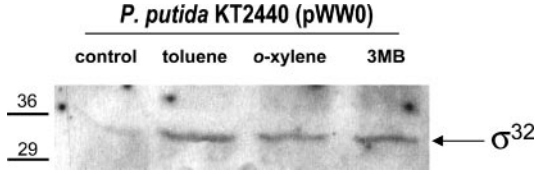



FIGURE 6. **Levels of $\sigma^{32}$ protein in *P. putida* KT2440 (pWW0) cells subjected to a 15-min challenge with aromatics.** Cells cultures grown as indicated under "Experimental Procedures" were treated with toluene, *o*-xylene, or 3MB for 15 min before obtaining crude extracts. For each condition, 70 $\mu$g of total proteins from the corresponding cell lysates were subjected to SDS-PAGE, and the blots were probed with anti-$\sigma^{32}$ antiserum.

HS response protein 20 (PP3234) were repressed regardless of whether the cells were exposed to the aromatic hydrocarbons or to 3MB (Table 5).

*Effect of Aromatic Inducers on Amino Acid and Nucleotide Homeostasis and on Biosynthesis of Proteins and Nucleic Acids*—Any physiological response (whether metabolic or otherwise) to the presence in the medium of *o*-xylene, toluene, or 3MB involves *de novo* protein biosynthesis. It was, therefore, not surprising that major changes in the transcriptome were observed in genes involved in amino acid biosynthesis and in critical functions for protein production (Table 6). The most remarkable finding was the strong induction of methionine biosynthesis genes (every cistron in the pathway was induced) (Table 6). We also found that *gltB* and *gltD*, which encode glutamate synthase, were induced between 1.83- and 5.9-fold (Table 6). The increased production of glutamate may not only offset the possible impairment by this compound of the biosynthesis of other amino acids but may also enhance the selectivity of RNAP with alternative $\sigma$ factors for their cognate promoters (54). But many other pathways of amino acid biosynthesis were also induced, in particular those for leucine, isoleucine, tryptophan, serine, and arginine (Table 6 and supplemental Table 5). We also found that the pathways for the catabolism of tryptophan and arginine were turned down, an effect that should have resulted in a further increase in the availability of these two amino acids.

The increase in amino acids in the short-term response to toluene probably reflects the need of these amino acids in the synthesis of the new proteome found in cells exposed to toluene (18). This hypothesis is supported by the induction of the Tuf-1 elongation factor for protein

**TABLE 6**
**Changes in genes encoding enzymes related to amino acid metabolism**

| TIGR accession no. | Gene | Description | -Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-Xylene | 3MB |
| PP0675 | *gdhA* | Glutamate dehydrogenase | 1.7 | 1.81 | |
| PP0840 | *cysE* | Serine *O*-acetyltransferase | | 1.92 | |
| PP1025 | *leuA* | 2-Isopropylmalate synthase | | | 2.01 |
| PP1079 | *argF* | Ornithine carbamoyltransferase | 1.87 | 2.43 | |
| PP1346 | *argJ* | Glutamate *N*-acetyltransferase/amino acid | 2.57 | | |
| PP1995 | *trpF* | *N*-(5′phosphoribosyl)anthranilate isomerase | 1.7 | 2.26 | |
| PP2375 | *metH* | 5-methyltetrahydrofolate-homocysteine methyltransferase | 3.49 | 1.99 | |
| PP2698 | *metE* | 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase | 5.23 | 3.25 | 3.77 |
| PP2776 | | Homocysteine *S*-methyltransferase family protein | 2.25 | 1.87 | |
| PP3571 | | Putative acetylornithine deacetylase | 2.10 | | |
| PP4680 | *ilvB* | Acetolactate synthase large subunit | | 1.79 | |
| PP4977 | *metF* | 5,10-Methylenetetrahydrofolate reductase | | 1.81 | |
| PP5075 | *gltD* | Glutamate synthase small subunit | 3.10 | 5.90 | |
| PP5076 | *gltB* | Glutamate synthase large subunit | 1.83 | 2.37 | |
| PP5078 | *aroB* | 3-Dehydroquinate synthase | 2.32 | 1.91 | 1.53 |
| PP5079 | *aroK* | Shikimate kinase | 2.32 | 2.02 | |
| PP5097 | *metX* | Homoserine *O*-acetyltransferase | | 1.84 | |
| PP5098 | metW | Methionine biosynthesis protein | 2.40 | | |
| PP5149 | *ilvA* | 2-Threonine dehydratase biosynthetic | | 1.82 | −1.59 |
| PP5185 | *argA* | *N*-Acetylglutamate synthase | 2.93 | | 1.55 |
| PP4725 | *dapB* | Dihydrodipicolinate reductase | 3.58 | 5.64 | 1.35 |
| PP0420 | *trpG* | Anthranilate synthase component II | −1.8 | −2.00 | |
| PP0421 | *trpD* | Anthranilate phosphoribosyltransferase | −1.84 | −2.53 | |
| PP1001 | *arcA* | arginine deiminase | −1.9 | −6.73 | |

synthesis and an inorganic pyrophosphatase (PP0538), which is a major provider of energy for the cognate polypeptide-making reactions.

Exposure to *o*-xylene, toluene, or 3MB increased the level of expression of *zwf-1*, a pivotal enzyme in the synthesis of 6-phosphogluconate necessary to replenish some of the pentose phosphate intermediates needed for *de novo* nucleotide biosynthesis (see supplemental Table 5). This enzyme, which plays a key role in maintaining cellular reducing power, belongs to the SoxR regulon (55); therefore, the fact that it was induced in response to the aromatic compounds we tested is further evidence that oxidative stress is one of the driving forces in transcriptome reprogramming. Induction of the *zwf*-1 gene is also consistent with the fact that genes that encode other functions involved in nucleotide synthesis (*purU*, *ushA*-2, *gmk*-1, and others) were also up-regulated in response to the aromatic compounds we tested (supplemental Table 5).

## DISCUSSION

Bacteria living in hydrocarbon-polluted soil are exposed to multiple environmental inputs that set in motion extensive gene expression programs. Three types of response are possible; they are metabolic programs consisting of the expression of gene sets required for the catabolism or anabolism of nutrients and intermediates. stress-response programs for adaptation to suboptimal growth conditions, and morphological programs related to shape, transport, and surface chemistry of the bacterial cell. These three major programs are intimately connected, and their functioning determines the survival of a given population in specific niches or its displacement by a fitter organism. *P. putida* KT2440 (pWW0) provides an exceptional experimental system to examine how environmental bacteria manage the interplay between different gene expression programs, as some of this organism's growth substrates (in particular toluene and *m*-xylene) also happen to be acute physiological stressors. We used DNA array technology to examine the response of *P. putida* KT2440 (pWW0) to toluene, a chemical that is a carbon and energy source for this strain but that also damages membranes and denatures proteins (for review, see Refs. 4 and 6). When cells face toluene in the medium, both new metabolic programs and stress endurance programs are activated. How such programs interact and influence each other is the subject of the work reported here. For comparison, we examined the response to *o*-xylene, a chemically close relative to toluene that is not a growth substrate for *P. putida* KT2440, and to 3MB, a substrate of the lower TOL pathway (Fig. 1) that lacks the considerable hydrophobicity and toxicity of the aromatic hydrocarbons.

Induction of the whole complement of the *xyl* genes borne in plasmid pWW0 was a useful criterion to establish the reliability of the DNA array technology in light of the availability of data on the regulation of catabolic operons (for review, see Ref. 4). Activation of the expression from the Pu promoter is a process assisted by the plasmid-encoded XylR protein activated by toluene or *o*-xylene, whereas the activation of the Pm promoter by toluene, *o*-xylene, and 3MB is a process mediated by the XylS regulator (see Fig. 1). With some minor refinements, our set of data on the *xyl* genes of the TOL plasmid is generally consistent with the gross expression profiling of the same cistrons reported earlier (21, 31, 56, 57). One remarkable aspect of these experiments is that the noncarbon source *o*-xylene was able to induce both the upper and the lower pathways of the TOL plasmid, albeit to a lesser extent than the true pathway substrates, toluene and 3MB.

However, an interesting finding in this respect was the very significant activation of the *ben* cluster for the metabolism of non-substituted benzoate through a competing chromosomal, *ortho*-cleavage pathway. Because benzoate is produced from toluene in a series of reactions catalyzed by the TOL upper pathway, it comes as no surprise that the *ben* genes are induced by toluene despite the competition between the *meta* and the *ortho* downstream routes. The intriguing thing about our findings is that non-substrates such as *o*-xylene and 3MB also induced the *ben* gene cluster. Although this can be explained by invoking the proven cross-activation of the *benA* promoter by XylS (35), the result of this gratuitous induction is misrouting of 3-methylcatechol into an *ortho*-cleavage route that yield dead-end products.

In terms of the number of regulated genes, the changes in expression observed in the genes of "peripheral" metabolic routes for toluene and 3MB represented only a minor fraction of the total number of genes, whose expression significantly varied upon exposure to each of the aromatic compounds we tested. The significant changes detected with the DNA array used here are summarized in Fig. 3, in which genes are arbitrarily grouped according to 17 general functions (see also supplemental Tables 1–11). Fig. 3 illustrates functions for almost 20% of all

genes in the *P. putida* KT2440 strain, and no significant changes in any of the remaining genes were found. The data in Fig. 3 suggest that the number of genes that were induced upon exposure to each of the three aromatics is roughly the same as the number of genes that were inhibited. Yet this general trend was not identical for the three chemical species. *o*-Xylene, which cannot be metabolized, acted mainly as a chemical stressor, whereas in the response to toluene and 3MB, more metabolic genes were induced than repressed. The same general pattern was observed when, instead of counting the number of genes, we scored the -fold induction of individual genes (not shown).

Our results also support that the gross effect of toluene on cells is much more similar to that of *o*-xylene (which acts as a strong stressor) than to 3MB (predominantly a substrate). In fact, the order of magnitude of both the number and the intensity of the changes was *o*-xylene > toluene ≫ 3MB. The Venn diagram shown in Fig. 2 illustrates that toluene- and *o*-xylene-dependent reprogramming share more than 55% of the up-regulated genes, but only 35% of the genes down-regulated by *o*-xylene were also repressed in the presence of toluene (Fig. 2). Genes down-regulated by all three compounds represent functions involved in primary sensing at the membrane level, whereas about half of the up-regulated genes under all three conditions are part of the HS regulon and metabolic functions. Cluster analysis of the response to each of the three aromatics clearly grouped *o*-xylene with toluene and distinguished the former from 3MB.

As deduced from the magnitude of the responses, it seems that the primary reaction of the cells to the presence of aromatic compounds takes place at the level of the cell envelope (membrane proteins, lipid metabolism, transport, etc.). Injury to the membranes leads to oxidative damage that is observed as a reduction in electron transport chain activity and increases in hydrogen peroxide production, which in turn produces a general response to oxidative stress. Finally, protein misfolding generated by the presence of aromatics triggers the classical HS response leading to induction of the HS regulon. Our results show that the induction of metabolic and stress-related functions by all three compounds was accompanied by the inhibition of motility and repression of enzymes involved in metabolic feeding or part of the Krebs cycle (see supplemental Tables 3 and 5). Although the specific regulatory mechanisms that switch off cell motility are not understood, inhibition of motility is a common response to almost any stress situation and may, thus, represent a general mechanism for saving energy. A group of recently characterized proteins containing the so-called GGDEF motif, involved in biosynthesis of the messenger molecule dGTPc, seems likely to be involved in signal transmission. In fact, six GGDEF motif proteins were induced under our experimental conditions.

Changes in ~20% of *P. putida* KT2440 genes occurred in response to the chemicals we tested within a short period of time (15 min), with no variation in the expression of RNAP components and only minor changes in the expression of transcriptional regulators. A plausible explanation for how the activation of genes occurred in the absence of changes in the transcriptional machinery is that stress caused by aromatic compounds led to a rapid reassignment of available transcriptional elements from dispensable functions and promoters to functions required for stress endurance. Depending on the relative affinities of RNAP for given promoters, one could argue that a fraction of the enzyme pool is permanently engaged in the transcription of basic housekeeping functions. But there may be a pool of roaming RNAP available to conditionally activate other sets of genes. The assignment of this pool to alternative promoters may depend on the presence of alternative $\sigma$ factors, specific transcriptional factors, and specific promoter strength and on an appropriate intracellular environment. This hypothesis is supported by our finding that $\sigma^{32}$ concentration increased in the presence of all three compounds (Fig. 6), reflecting the known mechanism of $\sigma^{32}$ stabilization (51). We entertain the notion that exposure to *o*-xylene and toluene (and to a lesser extent, to 3MB) rapidly activates the HS $\sigma^{32}$ protein (Fig. 6 and Table 5) (52, 53), which may then take over much of the roaming RNAP and redirect the response toward the defense against damage to the cell architecture (54, 58). In this scenario of limited RNAP availability, other metabolic and stress endurance functions can be expressed if other promoters are liberated. In other words, the number of genes that are not expressed in the presence of the aromatic compounds may not be specifically repressed but, rather, deprived of an otherwise engaged transcriptional apparatus which is reassigned to express functions that now become compulsory for survival.

How could this occur? It is possible that changes in general conditions such as DNA supercoiling or an increase in intracellular solutes (such as glutamate) trigger a shift in the binding of the available population of RNAP molecules among DNA sequences for which they have intrinsically low affinity. This possibility will be the subject of further studies but is supported by our finding that the levels of glutamate synthase increased in response to exposure to the toxic chemicals we tested.

Although these hypotheses may provide a retrospective explanation of the general response of *P. putida* to aromatic compounds, the issue is still whether these effects are simply a consequence of transcriptional economy at the single-cell level or whether the response has been selected via evolution as a survival strategy and might be common to other groups of microorganisms or even higher organisms. When *E. coli* cells are faced with a decrease in the growth potential of the environment, they seem to use a risk-prone foraging behavior (40). This strategy consists of increased motility and the massive induction of pathways for the metabolism of unavailable carbon sources (40). In contrast, we observed that stress caused by the aromatic hydrocarbons tested here led to the loss of motility functions and the inhibition of large portions of the basic metabolic machinery. This strategy of minimal energy expenditure, rather than expending energy in the pursuit of a better environment, may reflect a short-term response that forms part of the panoply of mechanisms by which bacteria adapt to hostile environments (59).

*Acknowledgments—We thank Carolina González de Figueras and P. Marín-Quero for technical assistance, Carmen Lorente for secretarial assistance, and K. Shashok for checking the English in the manuscript.*

**REFERENCES**

1. Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., Martins dos Santos, V. A., Fouts, D. E., Gill, S. R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R. T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J. A., Timmis, K. N., Dusterhoft, A., Tummler, B., and Fraser, C. M. (2002) *Environ. Microbiol.* **4,** 799–808
2. Jiménez, J. I., Miñambres, B., and García, J. L., Díaz, E. (2004) in *Pseudomonas* (Ramos, J. L., ed) Vol. 3, pp. 425–462, Kluwer Academic/Plenum Publishers, London
3. Nakazawa, T. (2002) *Environ. Microbiol.* **4,** 782–786
4. Ramos, J. L., Marqués, S., and Timmis, K. N. (1997) *Annu. Rev. Microbiol.* **51,** 341–373
5. Sikkema, J., de Bont, J. A., and Poolman, B. (1995) *Microbiol. Rev.* **59,** 201–222
6. Ramos, J. L., Duque, E., Gallegos, M. T., Godoy, P., Ramos-González, M. I., Rojas, A., Terán, W., and Segura, A. (2002) *Annu. Rev. Microbiol.* **56,** 743–768
7. Segura, A., Terán, W., Guazzaroni, M. E., Rojas, A., Duque, E., Gallegos, M. T., and Ramos, J. L. (2004) in *Pseudomonas* (Ramos, J. L., ed) Vol. 2, pp. 479–508, Kluwer Academic/Plenum Publishers, London
8. Kieboom, J., Dennis, J. J., Zylstra, G. J., and de Bont, J. A. (1998) *J. Bacteriol.* **180,** 6769–6772

9. Kim, K., Lee, S., Lee, K., and Lim, D. (1998) *J. Bacteriol.* **180,** 3692–3696
10. Ramos, J. L., Duque, E., Godoy, P., and Segura, A. (1998) *J. Bacteriol.* **180,** 3323–3329
11. Mosqueda, G., and Ramos, J. L. (2000) *J. Bacteriol.* **182,** 937–943
12. Rojas, A., Duque, E., Mosqueda, G., Golden, G., Hurtado, A., Ramos, J. L., and Segura, A. (2001) *J. Bacteriol.* **183,** 3967–3973
13. Rojas, A., Segura, A., Guazzaroni, M. E., Terán, W., Hurtado, A., Gallegos, M. T., and Ramos, J. L. (2003) *J. Bacteriol.* **185,** 4755–4763
14. Duque, E., Segura, A., Mosqueda, G., and Ramos, J. L. (2001) *Mol. Microbiol.* **39,** 1100–1106
15. Terán, W., Felipe, A., Segura, A., Rojas, A., Ramos, J. L., and Gallegos, M. T. (2003) *Antimicrob. Agents Chemother.* **47,** 3067–3072
16. Junker, F., and Ramos, J. L. (1999) *J. Bacteriol.* **181,** 5693–5700
17. von Wallbrunn, A., Richnow, H. H., Neumann, G., Meinhardt, F., and Heipieper, H. J. (2003) *J. Bacteriol.* **185,** 1730–1733
18. Segura, A., Godoy, P., van Dillewijn, P., Hurtado, A., Arroyo, N., Santacruz, S., and Ramos, J. L. (2005) *J. Bacteriol.* **187,** 5937–5945
19. Santos, P. M., Benndorf, D., and Sa-Correia, I. (2004) *Proteomics* **4,** 2640–2652
20. Cases, I., and de Lorenzo, V. (2005) *Nat. Rev. Microbiol.* **3,** 105–118
21. Abril, M. A., Michán, C., Timmis, K. N., and Ramos, J. L. (1989) *J. Bacteriol.* **171,** 6782–6790
22. Spaink, H. P., Okker, R. J., Wijffelman, C. A., Pees, E., and Lugtenberg, B. J. (1987) *Plant Mol. Biol.* **9,** 27–39
23. Yuste, L., Hervás, A. B., Canosa, I., Tobes, R., Jiménez, J. I., Nogales, J., Pérez-Pérez, M. M., Santero, E., Díaz, E., Ramos, J. L., de Lorenzo, V., and Rojo, F. (2006) *Environ. Microbiol.* **8,** 165–177
24. Greated, A., Lambertsen, L., Williams, P. A., and Thomas, C. M. (2002) *Environ. Microbiol.* **4,** 856–871
25. Fawcett, J., and Pawson, T. (2000) *Science* **290,** 725–726
26. Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. (2002) *Nucleic Acids Res.* **30,** e15
27. Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 14863–14868
28. Buege, J. A., and Aust, S. D. (1978) *Methods Enzymol.* **52,** 302–310
29. Miller, J. (1972) in *Experiments in Molecular Genetics*, pp. 352–355, Cold Spring Harbor, NY
30. Tatusov, R. L., Mushegian, A. R., Bork, P., Brown, N. P., Hayes, W. S., Borodovsky, M., Rudd, K. E., and Koonin, E. V. (1996) *Curr. Biol.* **6,** 279–291
31. Harayama, S., Rekik, M., Wubbolts, M., Rose, K., Leppik, R. A., and Timmis, K. N. (1989) *J. Bacteriol.* **171,** 5048–5055
32. Marqués, S., Ramos, J. L., and Timmis, K. N. (1993) *Biochim. Biophys. Acta* **1216,** 227–236
33. Rojo, F., Pieper, D. H., Engesser, K. H., Knackmuss, H. J., and Timmis, K. N. (1987) *Science* **238,** 1395–1398
34. Kessler, B., Marqués, S., Kohler, T., Ramos, J. L., Timmis, K. N., and de Lorenzo, V. (1994) *J. Bacteriol.* **176,** 5578–5582
35. Cowles, C. E., Nichols, N. N., and Harwood, C. S. (2000) *J. Bacteriol.* **182,** 6339–6346
36. Duetz, W. A., and Van Andel, J. G. (1991) *J. Gen. Microbiol.* **137,** 1369–1374
37. Ramos, J. L., Duque, E., Rodríguez-Herva, J. J., Godoy, P., Haïdour, A., Reyes, F., and Fernández-Barrero, A. (1997) *J. Biol. Chem.* **272,** 3887–3890
38. Muñoz-Rojas, J., Bernal, P., Duque, E., Godoy, P., Segura, A., and Ramos, J. L. (2006) *Appl. Environ. Microbiol.* **72,** 472–477
39. Li, C., Louise, C. J., Shi, W., and Adler, J. (1993) *J. Bacteriol.* **175,** 2229–2235
40. Liu, Y., Gao, W., Wang, Y., Wu, L., Liu, X., Yan, T., Alm, E., Arkin, A., Thompson, D. K., Fields, M. W., and Zhou, J. (2005) *J. Bacteriol.* **187,** 2501–2507
41. Maurer, L. M., Yohannes, E., Bondurant, S. S., Radmacher, M., and Slonczewski, J. L. (2005) *J. Bacteriol.* **187,** 304–319
42. Ruberg, S., Tian, Z. X., Krol, E., Linke, B., Meyer, F., Wang, Y., Puhler, A., Weidner, S., and Becker, A. (2003) *J. Biotechnol.* **106,** 255–268
43. Wen, Y., Marcus, E. A., Matrubutham, U., Gleeson, M. A., Scott, D. R., and Sachs, G. (2003) *Infect. Immun.* **71,** 5921–5939
44. Simm, R., Morr, M., Kader, A., Nimtz, M., and Romling, U. (2004) *Mol. Microbiol.* **53,** 1123–1134
45. Römling, U., Gomelsky, M., and Galperin, M. Y. (2005) *Mol. Microbiol.* **57,** 629–639
46. Gottesman, S., Halpern, E., and Trisler, P. (1981) *J. Bacteriol.* **148,** 265–273
47. Allocati, N., Favaloro, B., Masulli, M., Alexeyev, M. F., and Di Ilio, C. (2003) *Biochem. J.* **373,** 305–311
48. Hassett, R., Dix, D. R., Eide, D. J., and Kosman, D. J. (2000) *Biochem. J.* **351,** 477–484
49. Ochsner, U. A., Vasil, M. L., Alsabbagh, E., Parvatiyar, K., and Hassett, D. J. (2000) *J. Bacteriol.* **182,** 4533–4544
50. Marnett, L. J. (2002) *Toxicology* **181–182,** 219–222
51. Yura, T., and Nakahigashi, K. (1999) *Curr. Opin. Microbiol.* **2,** 153–158
52. Domínguez-Cuevas, P., Marín, P., Ramos, J. L., and Marqués, S. (2005) *J. Biol. Chem.* **280,** 41315–41323
53. Marqués, S., Manzanera, M., González-Pérez, M. M., Gallegos, M. T., and Ramos, J. L. (1999) *Mol. Microbiol.* **31,** 1105–1113
54. Ishihama, A. (2000) *Annu. Rev. Microbiol.* **54,** 499–518
55. Green, J., and Paget, M. S. (2004) *Nat. Rev. Microbiol.* **2,** 954–966
56. González-Pérez, M. M., Ramos, J. L., Gallegos, M. T., and Marqués, S. (1999) *J. Biol. Chem.* **274,** 2286–2290
57. Velázquez, F., Parro, V., and de Lorenzo, V. (2005) *Mol. Microbiol.* **57,** 1557–1569
58. Jishage, M., Kvint, K., Shingler, V., and Nystrom, T. (2002) *Genes Dev.* **16,** 1260–1270
59. Ramos, J. L., Gallegos, M., Marqués, S., Ramos-González, M., Espinosa-Urgel, M., and Segura, A. (2001) *Curr. Opin. Microbiol.* **4,** 166–171

SUPPLEMENTAL MATERIAL

Table 1: Changes in membrane composition

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MB |
| **Fatty acid and phospholipid metabolism** | | | | | |
| PP0368 | | acyl-CoA dehydrogenase putative | -1.82 | -2.96 | 2.12 |
| PP0542 | | ethanolamine ammonia-lyase light subunit putative | -4.21 | -3.36 | - |
| PP0543 | eutB | ethanolamine ammonia-lyase heavy subunit | -1.82 | -2.52 | - |
| PP0579 | | conserved hypothetical protein | - | 3.48 | - |
| PP0580 | | MaoC domain protein | -2 | -2.02 | - |
| PP0581 | | 3-oxoacyl-(acyl-carrier- protein) reductase | -2.09 | -2.03 | - |
| PP0763 | | medium-chain-fatty-acid CoA ligase | -2.24 | -2.88 | -1.42 |
| PP1408 | phaG | acyl-transferase | | -2.25 | -1.67 |
| PP2038 | | long-chain-fatty-acid--CoA ligase putative | -4.83 | - | - |
| PP2136 | fadB | fatty oxidation complex alpha subunit | -1.95 | -4.11 | - |
| PP2137 | fadA | 3-oxoacyl-CoA thiolase | -6.95 | -3.77 | - |
| PP2778 | | 3-oxoacyl-(acyl-carrier-protein) synthase II | - | 2.12 | - |
| PP3124 | | transporter short-chain fatty acid transporter family | - | -2.01 | - |
| PP3754 | | beta-ketothiolase | *-5.04* | -3.35 | - |
| PP4030 | | enoyl-CoA hydratase/isomerase family protein | - | -2.32 | 2.1 |
| PP4975 | | long-chain acyl-CoA thioester hydrolase family protein | - | -2.61 | - |
| PP5365 | | cyclopropane-fatty-acyl-phospholipid synthase putative | - | -4.12 | 1.74 |
| PP3278 | phaF | ring-oxidation complex protein 1 | 2.13 | - | - |
| PP3283 | phaB | enoyl-CoA hydratase/isomerase PhaB | 2.48 | - | - |
| PP3540 | mvaB | hydroxymethylglutaryl-CoA lyase | 2.04 | - | 3.02 |
| **Cell envelope** | | | | | |
| PP0504 | oprG | outer membrane protein OprG | -2.54 | -10.46 | - |
| PP0505 | | lipoprotein putative | - | -2.32 | - |
| PP0547 | mpl | UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-d | -2.42 | - | - |
| PP0572 | pbpC | penicillin-binding protein 1C | -1.78 | -2 | - |
| PP0778 | | glycosyl transferase group 1 family protein | -1.96 | - | - |
| PP0933 | mreB | rod shape-determining protein MreB | - | 1.83 | -2.23 |
| PP0934 | mreC | rod shape determining protein MreC | - | - | -3.54 |
| PP0935 | mreD | rod shape-determining protein MreD | - | 1.74 | -2.34 |
| PP1065 | | membrane protein putative | - | -2.27 | - |
| PP1121 | | OmpA family protein | -2.44 | -3.37 | - |
| PP1206 | oprD | porin D | -1.87 | - | -2.04 |
| PP1234 | | permease PerM putative | - | - | -1.9 |
| PP2124 | | glycosyl transferase putative | -3.1 | - | -2.15 |
| PP2244 | | membrane protein putative | - | - | -2.07 |
| PP3069 | | outer membrane autotransporter | -1.63 | -2.05 | - |
| PP3141 | | glycosyl transferase WecB/TagA/CpsF family | -6.33 | - | - |
| PP3612 | tonB | dependent receptor putative | -3.42 | -1.84 | -1.87 |
| PP4217 | fpvA | outer membrane ferripyoverdine receptor | - | - | -14.1 |
| PP5206 | | membrane fusion protein | - | -4.61 | - |
| PP1757 | bolA | bolA protein | | -1.8 | -1.86 |
| PP0883 | | porin putative | - | 1.93 | - |
| PP1039 | tatC-1 | Sec-independent periplasmic protein translocator TatC | 1.97 | - | 3.71 |
| PP1249 | | lipoprotein putative | 1.76 | 2.07 | - |
| PP1284 | algE | outer membrane protein AlgE | 2.34 | - | 1.6 |
| PP1604 | lpxB | lipid A disaccharide synthase | 2.56 | 3.28 | - |
| PP3436 | rarD-2 | rarD protein | - | 2.13 | - |
| PP3954 | | periplasmic binding protein putative | 2.24 | 3.48 | 2.13 |
| PP4592 | | membrane protein putative | 1.67 | 2.52 | - |
| PP5190 | | type II secretion system protein | 2.67 | 1.92 | - |
| PP5256 | | cyclic nucleotide-binding protein | 2.29 | 3.3 | - |

Table 2: Motility and chemotaxis

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MB |
| PP0320 | | methyl-accepting chemotaxis transducer | -1.9 | -2.05 | -2.5 |
| PP2310 | | methyl-accepting chemotaxis transducer | - | -5.11 | - |
| PP2358 | | type 1 pili subunit CsuA/B protein putative | -2.31 | - | - |
| PP2359 | | type 1 pili subunit CsuA/B protein putative | -2.92 | -4.95 | - |
| PP2360 | | type 1 pili subunit CsuA/B protein putative | -3.75 | -2.93 | - |
| PP2362 | csuD | type 1 pili usher protein CsuD | -3.77 | -3.66 | -1.91 |
| PP2363 | csuE | type 1 pili protein CsuE | -11.86 | - | - |
| PP4337 | cheB | protein-glutamate methylesterase CheB | - | - | -2.54 |
| PP4364 | | anti-sigma F factor antagonist putative | -3.99 | -1.91 | -1.7 |
| PP4366 | fliI | flagellum-specific ATP synthase FliI | -2.33 | -2.37 | -1.78 |
| PP4367 | fliH | flagellar assembly protein FliH | -1.89 | - | -1.4 |
| PP4375 | fliS | flagellar biosynthetis protein FliS | - | -1.85 | - |
| PP4380 | flgL | flagellar hook-associated protein FlgL | - | -2.03 | - |
| PP4382 | flgJ | peptidoglycan hydrolase FlgJ | -1.83 | -1.94 | -1.4 |
| PP4384 | flgH | flagellar L-ring protein precursor FlgH | -1.86 | -2.04 | -1.4 |
| PP4385 | flgG | flagellar basal-body rod protein FlgG | - | -2.27 | -1.5 |
| PP4386 | flgF | flagellar basal-body rod protein FlgF | -2.13 | -2.15 | -1.88 |
| PP4389 | flgD | flagellar basal-body rod modification protein | -2.02 | -2.58 | -1.6 |
| PP4390 | flgC | flagellar basal-body rod protein FlgC | -1.82 | -1.93 | -1.48 |
| PP4391 | flgB | flagellar basal-body rod protein FlgB | -1.83 | - | - |

Table 3: Energy metabolism

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| PP0072 | qor-1 | quinone oxidoreductase | -2.67 | -3.40 | - |
| PP0105 | | cytochrome c oxidase assembly protein | - | -3.01 | 2.61 |
| PP0106 | | cytochrome c oxidase subunit III | - | -2.80 | 1.41 |
| PP0110 | cyoE-1 | protoheme IX farnesyltransferase | - | -1.66 | 1.82 |
| PP0125 | | cytochrome c-type protein | -1.99 | -3.41 | -1.79 |
| PP0126 | | cytochrome c4 | - | -3.09 | -1.59 |
| PP0155 | pntB | pyridine nucleotide transhydrogenase beta subunit | -2.78 | -1.66 | - |
| PP0180 | | cytochrome c family protein | 2.84 | - | 1.69 |
| PP0812 | cyoA | cytochrome o ubiquinol oxidase subunit II | -1.66 | -1.77 | 1.20 |
| PP0813 | cyoB | cytochrome o ubiquinol oxidase subunit I | -1.67 | -1.93 | 1.07 |
| PP0814 | cyoC | cytochrome o ubiquinol oxidase subunit III | -1.47 | -1.77 | 1.51 |
| PP0815 | cyoD | cytochrome o ubiquinol oxidase protein CyoD | - | -1.70 | 1.36 |
| PP0816 | cyoE-2 | protoheme IX farnesyltransferase | - | -1.62 | 1.82 |
| PP1317 | petA | ubiquinol--cytochrome c reductase iron-sulfur subunit | -1.48 | -2.90 | - |
| PP1318 | petB | ubiquinol--cytochrome c reductase cytochrome b | -1.79 | -3.40 | - |
| PP1319 | petC | ubiquinol--cytochrome c reductase cytochrome c1 | -1.55 | -2.09 | - |
| PP2204 | | copper resistance protein B putative | - | - | -2.09 |
| PP2737 | | oxidoreductase short-chain dehydrogenase/reductase family | -4.39 | - | - |
| PP3488 | | Sco1/SenC family protein | - | - | -1.90 |
| PP3720 | | NAD(P)H quinone oxidoreductase putative | -2.69 | - | - |
| PP3822 | | cytochrome c family protein | -3.77 | -3.15 | -2.12 |
| PP4250 | ccoN-1 | cytochrome c oxidase cbb3-type subunit I | -4.03 | -1.50 | - |
| PP4251 | ccoO-1 | cytochrome c oxidase cbb3-type subunit II | -4.70 | -1.44 | - |
| PP4252 | ccoQ-1 | cytochrome c oxidase cbb3-type CcoQ subunit | -2.89 | - | - |
| PP4253 | ccoP-1 | cytochrome c oxidase cbb3-type subunit III | -5.60 | - | - |
| PP4255 | ccoN-2 | cytochrome c oxidase cbb3-type subunit I | - | 1.60 | - |
| PP4256 | ccoO-2 | cytochrome c oxidase cbb3-type subunit II | - | - | - |
| PP4257 | ccoQ-2 | cytochrome c oxidase cbb3-type CcoQ subunit | - | 1.77 | |
| PP4258 | ccoP-2 | cytochrome c oxidase cbb3-type subunit III | - | - | - |
| PP4259 | | iron-sulfur cluster-binding protein | - | -2.15 | - |
| PP4262 | ccoS | cytochrome oxidase maturation protein cbb3-type | -3.48 | - | -1.60 |
| PP4264 | hemN | oxygen-independent coproporphyrinogen III oxidase | -7.37 | -1.64 | - |
| PP4651 | cioA | ubiquinol oxidase subunit I cyanide insensitive | -5.10 | - | - |
| PP4870 | | azurin | -2.30 | -5.32 | - |
| PP5378 | | cytochrome c family protein | - | -3.38 | -1.47 |
| PP5379 | copB | copper resistance protein B | -1.98 | -3.27 | -1.37 |

Table 4 : Transporters

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| PP0219 | | ABC transporter permease protein | 2.38 | 3.24 | - |
| PP0652 | | gluconate transporter | 2.33 | 2.21 | - |
| PP0958 | ttg2A | toluene tolerance ABC efflux transporter ATP-binding | 2.33 | 2.01 | - |
| PP1384 | ttgC | multidrug/solvent RND outer membrane protein TtgC | 2.25 | 1.65 | 2.47 |
| PP1385 | ttgB | multidrug/solvent RND transporter TtgB | - | - | 2.22 |
| PP1386 | ttgA | multidrug/solvent RND membrane fusion protein | 1.83 | - | 1.74 |
| PP1387 | ttgR | transcriptional regulator TtgR | 3 | 3.48 | - |
| PP1516 | | RND membrane fusion protein | 2.41 | 2.4 | - |
| PP1517 | | RND efflux transporter | - | 1.73 | - |
| PP3342 | nikA | nickel ABC transporter periplasmic nickel-binding protein | 1.84 | 1.82 | - |
| PP3349 | | major facilitator family transporter | 9.54 | 4.61 | - |
| PP3588 | | multidrug resistance transporter Bcr/CflA family | 4.35 | 5.59 | 5.64 |
| PP4756 | | amino acid permease | 1.77 | 4.03 | - |
| PP0026 | | cobalt/cadmium/zinc transporter CDF family | -1.88 | -2 | - |
| PP0076 | | glycine betaine-binding protein putative | -2.93 | -8.61 | - |
| PP0227 | | cysteine ABC transporter periplasmic cysteine-bindin | -1.92 | -2.1 | - |
| PP0238 | ssuD | organosulfonate monooxygenase | -2.26 | -2.31 | - |
| PP0239 | ssuC | sulfonate ABC transporter permease protein SsuC | - | -2.91 | - |
| PP0296 | | glycine betaine/L-proline ABC transporter periplasmi | -3.39 | -4.97 | - |
| PP0544 | | ethanolamine transporter | -2.3 | -2.22 | - |
| PP0803 | | protein secretion ABC efflux system membrane fusion | -2.32 | -2.32 | - |
| PP0881 | dppB | dipeptide ABC transporter permease protein | -1.88 | -2.62 | - |
| PP0882 | dppA | dipeptide ABC transporter periplasmic dipeptide-bind | - | -4.6 | - |
| PP0885 | | dipeptide ABC transporter periplasmic peptide-bindin | -1.88 | -1.8 | - |
| PP1002 | arcD | arginine/ornithine antiporter | - | -5.62 | - |
| PP1139 | braE | high-affinity branched-chain amino acid transport pro | -1.54 | - | -1.94 |
| PP1373 | | low-affinity inorganic phosphate transporter | - | 2.44 | -1.97 |
| PP1689 | | long-chain fatty acid transporter putative | -2.1 | -1.62 | -2.17 |
| PP1724 | | ABC transporter permease protein | - | -2.62 | -2.04 |
| PP3069 | | outer membrane autotransporter | -1.63 | -2.05 | - |
| PP3612 | TonB | dependent receptor putative | -3.42 | -1.84 | -1.87 |
| PP4103 | | low-affinity inorganic phosphate transporter | 2.83 | -6.12 | - |
| PP5175 | HlyD | family secretion protein | - | -3.02 | -1.8 |

Table 5: Metabolism

Aminoacids

| TIGR ID | | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| PP0243 | gshA | glutamate--cysteine ligase | 1.8 | 2.61 | 1.7 |
| PP0420 | trpG | anthranilate synthase component II | -1.8 | -2 | - |
| PP0421 | trpD | anthranilate phosphoribosyltransferase | -1.84 | -2.53 | - |
| PP0675 | gdhA | glutamate dehydrogenase | 1.7 | 1.81 | - |
| PP0840 | cysE | serine O-acetyltransferase | - | 1.92 | - |
| PP0864 | | ornithine decarboxylase putative | - | 1.93 | - |
| PP0987 | sda-2 | L-serine dehydratase iron-sulfur-dependent | - | 2.48 | - |
| PP0988 | gcvP-1 | glycine cleavage system P protein | - | 5.3 | - |
| PP0989 | gcvH-1 | glycine cleavage system H protein | - | 7.1 | - |
| PP1025 | leuA-2 | isopropylmalate synthase | - | - | 2.01 |
| PP1079 | argF | ornithine carbamoyltransferase | 1.87 | 2.43 | - |
| PP1255 | | D-amino acid dehydrogenase small subunit family | - | 1.88 | - |
| PP1346 | argJ | glutamate N-acetyltransferase/amino-acid | 2.57 | - | - |
| PP1995 | trpF | N-(5'phosphoribosyl)anthranilate isomerase | 1.7 | 2.26 | - |
| PP2375 | metH | 5-methyltetrahydrofolate--homocysteine | 3.49 | 1.99 | - |
| PP2698 | metE | 5-methyltetrahydropteroyltriglutamate-homocysteine me | 5.23 | 3.25 | 3.77 |
| PP2776 | | homocysteine S-methyltransferase family protein | 2.25 | 1.87 | - |
| PP2929 | nspC | carboxynorspermidine decarboxylase | - | 1.82 | - |
| PP3515 | hyuA | hydantoin utilization protein A | 2.08 | - | 1.53 |
| PP3571 | | acetylornithine deacetylase putative | 2.1 | - | - |
| PP3596 | | D-amino acid dehydrogenase small subunit family | 2.21 | - | 2.31 |
| PP4680 | ilvB | acetolactate synthase large subunit | - | 1.79 | - |
| PP4725 | dapB | dihydrodipicolinate reductase | 3.58 | 5.64 | 1.35 |
| PP4752 | | aminopeptidase putative | - | 3.03 | - |
| PP4757 | | glucarate dehydratase | 1.78 | 3.22 | - |
| PP4977 | metF | 5 10-methylenetetrahydrofolate reductase | - | 1.81 | - |
| PP5075 | gltD | glutamate synthase small subunit | 3.1 | 5.9 | - |
| PP5076 | gltB | glutamate synthase large subunit | 1.83 | 2.37 | - |
| PP5078 | aroB | 3-dehydroquinate synthase | 2.32 | 1.91 | 1.53 |
| PP5079 | aroK | shikimate kinase | 2.32 | 2.02 | - |
| PP5097 | metX | homoserine O-acetyltransferase | - | 1.84 | - |
| PP5098 | metW | methionine biosynthesis protein | 2.4 | - | - |
| PP5340 | | acetylpolyamine aminohydrolase | 2.01 | - | - |
| PP5149 | ilvA-2 | threonine dehydratase biosynthetic | - | 1.82 | -1.59 |
| PP5185 | argA | N-acetylglutamate synthase | 2.93 | | 1.55 |
| PP5269 | dadX | alanine racemase catabolic | 2.43 | 2.42 | 1.96 |
| | | | | | |
| PP0420 | trpG | anthranilate synthase component II | -1.8 | -2 | - |
| PP0421 | trpD | anthranilate phosphoribosyltransferase | -1.84 | -2.53 | - |
| PP0567 | speA | biosynthetic arginine decarboxylase | -1.7 | -2.12 | - |
| PP0999 | arcC | carbamate kinase | - | -9.23 | - |
| PP1000 | argI | ornithine carbamoyltransferase catabolic | -21.4 | -13.02 | - |
| PP1001 | arcA | arginine deiminase | -1.9 | -6.73 | - |

Table 5: Metabolism

Sugars

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3-MBz |
| PP0058 | | 1-acyl-sn-glycerol-3-phosphate acyltransferase putative | 2.02 | - | - |
| PP0339 | aceE | pyruvate dehydrogenase E1 component | -2.02 | - | 2.16 |
| PP1009 | gap-1 | glyceraldehyde 3-phosphate dehydrogenase | -1.52 | - | 2.64 |
| PP1010 | edd | 6-phosphogluconate dehydratase | - | - | 1.9 |
| PP1022 | zwf-1 | glucose-6-phosphate 1-dehydrogenase | 2.13 | 2.11 | 1.67 |
| PP1024 | eda | 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase | - | - | 2.43 |
| PP1256 | | ketoglutarate semialdehyde dehydrogenase | - | 3.2 | -1.7 |
| PP1720 | | alcohol dehydrogenase zinc-containing | 2.05 | - | - |
| PP1943 | purU | 3-formyltetrahydrofolate deformylase | 1.7 | 2.33 | 11.73 |
| PP2637 | | endo-1 4-beta-D-glucanase | 2.17 | - | - |
| PP3746 | glcE | glycolate oxidase subunit GlcE | 2,25 | 1,98 | - |
| PP4301 | pykF | pyruvate kinase I | 1.7 | 1.9 | - |
| PP4736 | lldD | L-lactate dehydrogenase | - | 1.79 | |
| PP5210 | | alcohol dehydrogenase zinc-containing | 1.74 | 2.3 | - |
| PP0154 | | acetyl-CoA hydrolase/transferase family protein | - | -4.2 | - |
| PP0213 | gabD | succinate-semialdehyde dehydrogenase | -2.2 | -2.01 | - |
| PP0338 | aceF | pyruvate dehydrogenase dihydrolipoamide acetyltransferase | -2.06 | - | - |
| PP0339 | aceE | pyruvate dehydrogenase E1 component | -1.84 | - | 2.16 |
| PP0356 | glcB | malate synthase | -2.06 | -2.36 | - |
| PP0368 | | acyl-CoA dehydrogenase putative | -1.82 | -2.96 | 2.12 |
| PP0654 | mdh | malate dehydrogenase | -1.88 | -2.01 | - |
| PP1290 | | polysaccharide deacetylase family protein | -4.2 | -2.57 | - |
| PP1290 | | polysaccharide deacetylase family protein | -2.64 | - | - |
| PP1362 | pykA | pyruvate kinase II | -1.8 | -2.87 | - |
| PP1444 | gcd | glucose dehydrogenase (pyrroloquinoline-quinone) | -2.89 | -4.73 | - |
| PP1505 | PPc | phosphoenolpyruvate carboxylase | -3.5 | -2.52 | - |
| PP1649 | ldhA | D-lactate dehydrogenase | -2.31 | - | - |
| PP1755 | fumC-2 | fumarate hydratase class II | -1.8 | -2.22 | - |
| PP1893 | | acyl-CoA dehydrogenase putative | - | -3.11 | - |
| PP3443 | | glyceraldehyde-3-phosphate dehydrogenase putative | - | -3.02 | - |
| PP3638 | | acyl-CoA dehydrogenase putative | - | -2.01 | - |
| PP3662 | | decarboxylase family protein | - | -2.02 | - |
| PP4011 | icd | isocitrate dehydrogenase NADP-dependent prokaryotic | - | -4.74 | - |
| PP4050 | glgA | glycogen synthase | -3.88 | -5.43 | - |
| PP4051 | | alpha-amylase family protein | - | -4.24 | - |
| PP4052 | malQ | 4-alpha-glucanotransferase | - | -5.72 | - |
| PP4053 | | glycosyl hydrolase putative | - | -4.21 | - |
| PP4054 | | conserved hypothetical protein | - | -6.57 | - |
| PP4055 | glgX | glycogen operon protein GlgX | -3.03 | -4.12 | - |
| PP4060 | | alpha-amylase family protein | -9.05 | -6.43 | - |
| PP4116 | aceA | isocitrate lyase | - | -3.36 | - |
| PP5004 | | poly(3-hydroxyalkanoate) depolymerase | - | -2.66 | - |
| PP5007 | | polyhydroxyalkanoate granule-associated protein GA2 | - | -2.4 | - |
| PP5008 | | polyhydroxyalkanoate granule-associated protein GA1 | - | -2.07 | - |
| PP5056 | pgm | 2 3-biphosphoglycerate-independent phosphoglycerate mutase | - | -2.06 | - |
| PP5153 | | fumarylacetoacetate hydrolase family protein | - | -2.09 | - |
| PP5248 | | hydrolase isochorismatase family | - | -2.15 | - |
| PP5409 | | glmS-glucosamine--fructose-6-phosphate aminotransferase isomerizing | - | -2.36 | - |

Table 5: Metabolism

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| **Cofactors and prosthetic groups** | | | | | |
| PP0362 | bioB | biotin synthetase | 1.8 | 2.63 | - |
| PP0365 | bioC | biotin biosynthesis protein BioC | - | 2.07 | - |
| PP4826 | cobJ | precorrin-3B C17-methyltransferase | - | 2.05 | - |
| PP3999 | coba2 | precorrin-3B C17-methyltransferase | 1.8 | - | - |
| PP5285 | coaBC | phosphopantothenoylcysteine decarboxylase | - | 2.15 | - |
| PP0376 | pqqE | coenzyme PQQ synthesis protein E | -2.52 | -2.7 | - |
| PP0377 | pqqD | coenzyme PQQ synthesis protein D | -1.91 | -2.6 | - |
| PP0530 | ribB | 3 4-dihydroxy-2-butanone 4-phosphate synthase | - | -2.32 | - |
| PP2090 | cobA | 1-uroporphyrin-III C-methyltransferase | - | -1.93 | - |
| **Nucleotides** | | | | | |
| PP0798 | | GGDEF domain protein | - | - | 3.27 |
| PP1144 | | GGDEF domain protein | 2.07 | - | - |
| PP1411 | | GGDEF domain protein | 2.83 | 2.12 | 1.79 |
| PP1414 | ushA | 5`-nucleotidase | - | 4.72 | - |
| PP1964 | | deoxynucleotide monophosphate kinase putative | 1.94 | 1.86 | - |
| PP1967 | | deoxyribonuclease TatD family | - | 2.14 | 1.86 |
| PP2860 | | nicotinamide mononucleotide transporter putative | 2.23 | - | - |
| PP3348 | | GGDEF domain protein | 2.24 | 2.68 | 1.74 |
| PP3452 | | GGDEF domain protein | 1.81 | - | 1.53 |
| PP4469 | gmk-1 | guanylate kinase | 1.4 | 2.78 | 1.64 |
| PP5045 | thiI | thiamine biosynthesis protein ThiI | 1.6 | 3.39 | -1.97 |
| PP5256 | | cyclic nucleotide-binding protein | 2.29 | 3.29 | - |
| PP5256 | | cyclic nucleotide-binding protein | - | 3.29 | - |
| PP5263 | | GGDEF domain protein | 2.19 | - | - |
| PP5286 | dut | deoxyuridine 5-triphosphate nucleotidohydrolase | 1.7 | 2.01 | - |
| PP5296 | gmk-2 | guanylate kinase | 1.68 | 2.06 | - |
| PP1177 | nrdB | ribonucleoside reductase beta subunit | - | -2.32 | - |
| PP1179 | nrdA | ribonucleoside reductase alpha subunit | - | -2.56 | - |
| PP1761 | | sensory box protein/GGDEF family protein | - | -2.51 | - |
| PP2151 | sthA | soluble pyridine nucleotide transhydrogenase | -1.9 | -2.22 | - |

Table 5: Metabolism

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | o-xylene | 3-MBz |
| PP0077 | betC | choline sulfatase | -2.16 | -3.05 | - |
| PP0100 | cynT | carbonic anhydrase | -2.68 | -5.4 | - |
| PP0541 | | acetyltransferase GNAT family | -2.66 | -2.55 | - |
| PP0541 | | acetyltransferase GNAT family | -2.65 | -2.55 | - |
| PP1661 | | dehydrogenase subunit putative | -2.09 | -2.67 | - |
| PP1821 | | glutathione S-transferase family protein | - | -2.22 | -2.32 |
| PP1894 | | glutathione S-transferase family protein | - | -1.98 | - |
| PP2023 | | glutathione S-transferase family protein | 1.8 | 1.83 | - |
| PP3151 | | aldehyde dehydrogenase family protein | - | -3.36 | -2.26 |
| PP0210 | | phycobiliprotein putative | 5.88 | 4.51 | 4.58 |
| PP0651 | | acetyltransferase GNAT family | - | 1.86 | 1.5 |
| PP0939 | | carbon-nitrogen hydrolase family protein | - | 1.91 | - |
| PP1440 | | methyltransferase putative | - | 3.98 | - |
| PP1700 | | acyltransferase | - | 1.87 | - |
| PP1781 | | O-acyltransferase putative | - | 2.43 | - |
| PP1944 | | aminomethyltransferase putative | - | - | 6.24 |
| PP1992 | | semialdehyde dehydrogenase family protein | - | 2.23 | -1.7 |
| PP2377 | | acetyltransferase Act putative | 1.52 | 3.39 | - |
| PP2415 | | acetyltransferase GNAT family | 2.43 | 2.7 | - |
| PP2805 | | monooxygenase flavin-binding family | 1.81 | 1.81 | 2.01 |
| PP3243 | | acetyltransferase GNAT family | 3.5 | 2.11 | - |
| PP3436 | rarD-2 | rarD protein | - | 2.13 | - |
| PP3646 | | aldehyde dehydrogenase family protein | 2.15 | - | 3.14 |
| PP4558 | | acetyltransferase GNAT family | 1.8 | 2.45 | - |

Table 6 A. Transcriptional regulators

| TIGR ID | Gene name | Description(*) | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| PP0667 | | RNA polymerase sigma-70 factor ECF subfamily | - | 1.81 | -2.51 |
| PP1438 | czrR-2 | DNA-binding response regulator CzrR | 1.93 | 2.36 | - |
| PP2457 | rbsR | ribose operon repressor | 1.99 | 2.01 | - |
| PP2547 | | transcriptional regulator LysR family | 5.11 | - | 1.82 |
| PP3577 | | RNA polymerase sigma-70 factor ECF subfamily | - | 3.36 | -5.8 |
| PP3750 | | transcriptional regulator GntR family | 5.01 | 1.79 | - |
| PP4208 | | RNA polymerase sigma-70 factor ECF subfamily | - | 2.07 | - |
| PP4464 | | transcriptional regulator LysR family | 2.00 | 3.83 | - |
| PP4553 | | RNA polymerase sigma-70 factor ECF subfamily | - | 3.42 | - |
| PP5375 | | transcriptional regulator LysR family | 1.92 | 1.81 | - |
| PP0173 | | transcriptional factor-related protein | - | -2.25 | -3.55 |
| PP0271 | gltR-1 | DNA-binding response regulator GltR | -2.43 | - | -1.82 |
| PP0563 | | response regulator | -1.60 | -2.18 | - |
| PP2210 | | transcriptional regulator LysR family | -1.87 | -3.05 | -1.77 |
| PP3419 | | sigma-54 dependent transcriptional regulator/response | -4.00 | -3.27 | - |

* Table 6 B. Transcriptional regulators homology comparison

| TIGR ID | Predicted Roles |
|---|---|
| PP1438 | Response regulator of heavy metal sensor two component system. Subrole : detoxication |
| PP2457 | Salvage of nucleotides |
| PP2547 | Similar structure to periplasmic binding proteins |
| PP3750 | MocR family. Catabolism of rhizopine, symbiotic specific compound. |
| PP4464 | Similar structure to periplasmic binding proteins |
| PP5375 | - |
| PP0173 | Arg-repressor domain |
| PP0271 | Response regulator of heavy metal sensor two component system/ Phosphate regulon |
| PP0563 | Signal receiver domain/ GGDEF domain/ PleD homology E= 3e-70 |
| PP2210 | - |
| PP3419 | Sensor histidine kinase/ ATPase/HTH domains |

Table 7: Stress

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| | | Heat-shock | | | |
| PP0089 | osmC | osmotically inducible protein OsmC | 3.72 | 3.11 | 1 |
| PP0252 | | chaperonin 33 kDa | 3.02 | 3.67 | 1.15 |
| PP0636 | | cold shock DNA-binding domain protein | 2.21 | 2.01 | -1.4 |
| PP1225 | | radical activating enzyme | 1.2 | 3.95 | 1 |
| PP1254 | | xenA-xenobiotic reductase A | 1.2 | 2.08 | -1.75 |
| PP1360 | groES | chaperonin 10 kDa | 1 | 1.5 | 1.45 |
| PP1361 | groEL | chaperonin 60 kDa | 1.1 | 1.83 | 1.24 |
| PP1443 | lon-1 | ATP-dependent protease La | 1.3 | 3.85 | 1.45 |
| PP1982 | ibpA | heat-shock protein IbpA | 3.77 | 5.63 | 6.46 |
| PP2299 | tig | trigger factor | 2.24 | 2.04 | -1.2 |
| PP2302 | lon-2 | ATP-dependent protease La | 2.26 | 2.13 | 1.49 |
| PP4178 | | dienelactone hydrolase family protein | 2.26 | 3.78 | 1.84 |
| PP4179 | htpG | heat shock protein HtpG | 2.26 | 4.81 | 2.25 |
| PP4180 | | conserved hypothetical protein | 2.28 | 5.57 | 1.88 |
| PP4726 | dnaJ | dnaJ protein | 2.73 | 4.62 | 1.54 |
| PP4727 | dnaK | dnaK protein | 1.64 | 6.43 | 1.86 |
| PP4728 | grpE | heat shock protein GrpE | 2.31 | 5.87 | 2.08 |
| PP5000 | hslV | heat shock protein HslV | 1.71 | 4.95 | 2.08 |
| PP5001 | hslU | heat shock protein HslVU ATPase subunit HslU | 4.38 | 12.03 | 2.58 |
| PP1478 | xenB | xenobiotic reductase putative | -4.12 | -3.82 | - |
| PP2187 | | universal stress protein family | -2.27 | -9.87 | -1.6 |
| PP3070 | PPiC-1 | peptidyl-prolyl cis-trans isomerase C | -2.48 | -3.06 | - |
| PP3234 | | heat shock protein HSP20 family | -47.05 | -14.84 | -1.79 |
| | | Oxidative stress | | | |
| PP0483 | uvrA | excinuclease ABC A subunit | 1.65 | - | - |
| PP1217 | ruvB | holliday junction DNA helicase RuvB | 1.7 | - | -1.79 |
| PP1821 | | glutathione S-transferase family protein | - | -2.22 | -2.32 |
| PP1894 | | glutathione S-transferase family protein | - | -1.98 | - |
| PP1974 | uvrB | Excinuclease ABC B subunit | 2.29 | 1.49 | - |
| PP2023 | | glutathione S-transferase family protein | 1.8 | 1.83 | - |
| PP3116 | lexA | lexA-2-LexA repressor | - | 2.48 | - |
| PP3117 | sulA | cell division inhibitor SulA putative | - | 2.32 | -5.22 |
| PP3639 | | alkylhydroperoxidase AhpD domain protein | - | -2.33 | - |
| PP5284 | radC | DNA repair protein RadC | 2.37 | 2.94 | - |

Table 8: DNA metabolism

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| PP1868 | | ATP-dependent RNA helicase DEAD box family | 2.46 | 3.26 | - |
| PP2025 | sbcD | exonuclease SbcD | 2.31 | 2.23 | - |
| PP3114 | | ISPpu13 transposase Orf2 | 2.25 | - | - |
| PP3375 | endA-2 | endonuclease I | - | 3.13 | 2.1 |
| PP3893 | | phage DNA helicase putative | 1.85 | 1.89 | 1.76 |
| PP3965 | | ISPpu14 transposase Orf2 | 3.52 | - | - |
| PP4267 | recR | recombination protein RecR | 2.67 | 3.21 | 1.88 |
| PP4269 | dnaX | DNA polymerase III gamma and tau subunits | - | 1.9 | - |
| PP4276 | | chromosome segregation SMC protein | 2.2 | 2.24 | - |
| PP4796 | holA | DNA polymerase III delta subunit | 2.45 | 1.86 | - |
| PP4813 | | PAP2 family protein/DedA family protein | - | 2.04 | - |
| PP4873 | dnaB | replicative DNA helicase | 1.7 | 1.93 | - |
| PP4952 | dbpA | ATP-dependent RNA helicase DbpA | 3.55 | - | - |
| PP5125 | mutM | formamidopyrimidine-DNA glycosylase | 2.96 | - | - |
| PP5214 | rho | transcription termination factor Rho | 1.78 | 2.27 | - |
| PP5218 | | DedA family protein | 2.4 | 4.16 | - |
| PP5264 | rep | ATP-dependent DNA helicase Rep | - | 1.7 | - |
| PP5407 | | transposition protein TnsD-related protein | 1.78 | 2.30 | - |
| | | | - | - | - |
| PP0334 | | ISPpu11 transposase | -1.94 | - | - |
| PP1250 | | group II intron-encoding maturase | - | - | -2.32 |
| PP1533 | | excisionase putative | - | -2.04 | - |
| PP2218 | | ISPpu8 transposase | - | - | -2.37 |
| PP3032 | | DNA-binding protein Roi-related protein | -2 | -2.58 | - |
| PP3172 | | group II intron-encoding maturase | - | - | -2.48 |
| PP3868 | | group II intron-encoding maturase | - | - | -2.27 |
| PP4409 | | site-specific recombinase phage integrase family | -1.6 | -2.17 | - |
| PP5070 | | ParA family protein | -2.02 | -2.09 | - |

Table 9: Translation

| TIGR ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MBz |
| PP0389 | rpsU | ribosomal protein S21 | 2 | - | - |
| PP0436 | tyrS | tyrosyl-tRNA synthetase | | 2 | - |
| PP0441 | secE | preprotein translocase SecE subunit | - | 2.7 | - |
| PP0442 | nusG | transcription antitermination protein NusG | - | 3.45 | - |
| PP0443 | rplK | ribosomal protein L11 | - | 2.26 | - |
| PP0444 | rplA | ribosomal protein L1 | - | 2.32 | - |
| PP0465 | rplX | ribosomal protein L24 | - | 2.32 | - |
| PP0466 | rplE | ribosomal protein L5 | 1.91 | 2.41 | - |
| PP0467 | rpsN | ribosomal protein S14 | - | 4.85 | - |
| PP0688 | rplU | ribosomal protein L21 | - | 3.09 | - |
| PP0689 | rpmA | ribosomal protein L27 | - | 2.4 | - |
| PP0690 | | GTP-binding protein GTP1/Obg family | - | 3.54 | - |
| PP0720 | pth | peptidyl-tRNA hydrolase | - | 1.98 | - |
| PP0721 | | ribosomal 5S rRNA E-loop binding protein Ctc/L25/TL5 | - | 2.64 | - |
| PP0722 | prsA | ribose-phosphate pyrophosphokinase | - | 2.37 | - |
| PP0832 | queA | S-adenosylmethionine:tRNA ribosyltransferase-isomerase | - | 2.6 | - |
| PP1315 | rplM | ribosomal protein L13 | - | 2.04 | - |
| PP1316 | rpsI | ribosomal protein S9 | - | 2.12 | - |
| PP1591 | rpsB | ribosomal protein S2 | 1.93 | 2.65 | - |
| PP1772 | rpsA | ribosomal protein S1 | 1.83 | - | - |
| PP1858 | efp | translation elongation factor P | - | - | -2.05 |
| PP2466 | infC | translation initiation factor IF-3 | 1.93 | - | - |
| PP2467 | rpmI | ribosomal protein L35 | - | 3.7 | - |
| PP4111 | fusA-2 | translation elongation factor G | - | -2.14 | - |
| PP4221 | | non-ribosomal peptide synthetase | - | 2.31 | - |
| PP4709 | rpsO | ribosomal protein S15 | 2.28 | - | - |
| PP4711 | rbfA | ribosome-binding factor A | - | 2.36 | - |
| PP4794 | leuS | leucyl-tRNA synthetase | - | 2.47 | - |
| PP4874 | rplI | ribosomal protein L9 | 2.54 | 3.47 | - |
| PP4876 | rpsR | ribosomal protein S18 | 2.09 | 3.06 | - |
| PP4877 | rpsF | ribosomal protein S6 | 1.97 | 2.92 | - |
| PP4893 | hflX | GTP-binding protein HflX | - | 2.41 | - |
| PP4895 | miaA | tRNA delta(2)-isopentenylpyrophosphate transferase | 2.16 | - | - |
| PP5281 | rpmG | ribosomal protein L33 | 3 | 2.5 | - |
| PP5282 | rpmB | ribosomal protein L28 | 2.66 | 2.85 | - |

Table 10A. Conserved hypothetical proteins up-regulated

| TIGR ID | Toluene | *o*-xylene | 3MBz | Assigned function by homology | E value |
|---|---|---|---|---|---|
| PP0550 | - | - | 3.2 | Semialdehide dehydrogenase ArgC/ WcaG epimerase (cell envelope biogenesis) | 9e-04/ 9e-14 |
| PP0891 | 2.38 | - | - | YceI-base induced periplasmic protein | 4e-14 |
| PP1359 | - | 2.93 | - | Fxa cytoplasmic membrane protein | 6e-09 |
| PP1410 | - | 2.51 | 2.22 | conserved hypothetical protein | |
| PP1415 | - | 2.83 | - | AmoA putative. Amonia monooxigenase | 4e-72 |
| PP1442 | - | 3.46 | - | Secreted protein | 3e-18 |
| PP1556 | - | - | 2.52 | conserved hypothetical protein | |
| PP1910 | 1.74 | 2.88 | - | DUF177 | 2e-10 |
| PP1947 | - | - | 4.21 | AmpS. Leucyl aminopeptidase (amino acid transport and metabolism) | 4e-22 |
| PP1971 | - | - | 2.18 | DUF1285 pfam06938 | 7e-78 |
| PP2222 | -2.56 | - | 2.62 | conserved hypothetical protein | |
| PP2617 | - | 3.26 | - | DUF876 | 3e-110 |
| PP3096 | - | 2.97 | - | DUF1305 | 1e-93 |
| PP3353 | 2.79 | 11.11 | 5.87 | DUF323. In some cases contain a kinase domain of unknown function | 6e-47 |
| PP3494 | - | - | 35.92 | conserved hypothetical protein | |
| PP3945 | - | 3.48 | - | AmpS. Leucyl aminopeptidase (amino acid transport and metabolism) | 9e-07 |
| PP4180 | 2.28 | 5.57 | 1.87 | 4HBT domain/PaaI domain* | 3e-06/ 5e-16 |
| PP4488 | - | 1.8 | 2.96 | conserved hypothetical protein | |
| PP4528 | - | 2.95 | - | conserved hypothetical protein | |
| PP4529 | - | 3.1 | - | conserved hypothetical protein | |
| PP4554 | - | 3.54 | - | DUF704-Aha1 domain | 4e-20 |
| PP4769 | 1.8 | 2.64 | - | DUF88. Predicted function endonuclease or glycosyl hydrolase | 1e-42 |
| PP4770 | 2.84 | 4.04 | 2.28 | COG3416 | 2e-07 |
| PP4810 | 2.34 | 2.06 | - | NadD, Nicotinic acid mononucleotide adenylyltransferase (Coenzyme metabolism) | 3e-50 |
| PP4872 | 1.73 | 4.69 | - | Elp3, Elongator protein 3, MiaB family, Radical SAM/ Fe-S oxidoreductase [Energy production and conversion]. | 2e-12 / 3e-38 |
| PP4875 | 2.08 | 4.25 | - | Membrane protein putative*. Pseudomonas syringae* | 1e-64 |
| PP4916 | 3.04 | 2.66 | - | Predicted esterase | 1e-56 |
| PP5051 | - | 3.74 | - | conserved hypothetical protein | - |
| PP5362 | 1.96 | 6.26 | 1.97 | hypothetical protein PSPTO5527 (*Pseudomonas syringae* ) | 3e-28 |

Table 10B. Conserved hypothetical protein down-regulated

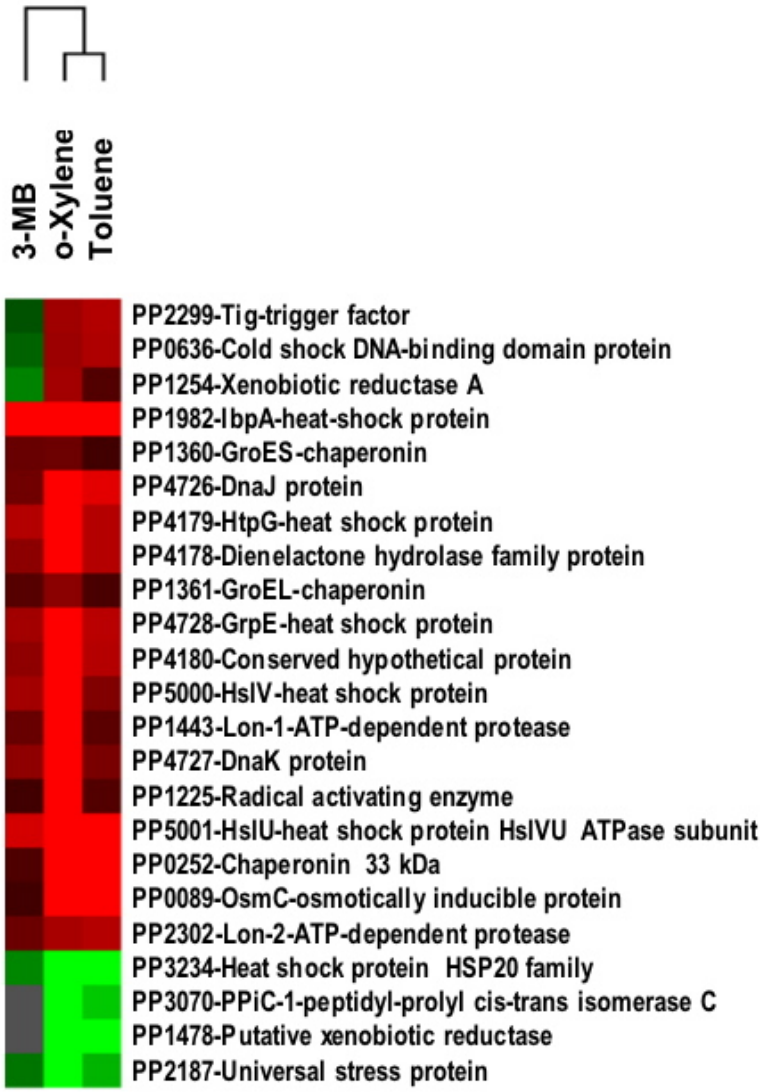
| TIGR ID | Toluene | *o*-xylene | 3MBz | Assigned function by homology | E value |
|---|---|---|---|---|---|
| PP0153 | -1.92 | -5.11 | - | Conserved hyphotetical protein | - |
| PP0174 | - | - | -3.23 | COG0790 | 2e-12 |
| PP0273 | - | -14.53 | - | Conserved hyphotetical protein | - |
| PP0395 | -1.92 | -3.88 | - | SpoVR. Spore cortex formation in Bacillus subtilis (σE dependent) | 0.0 |
| PP0396 | -1.87 | -4.35 | - | DUF444 | 3e-140 |
| PP0397 | -2.03 | -4.35 | - | PrkA serine protein kinase. Possesses the A-motif of nucleotide-binding proteins | 0.0 |
| PP0568 | - | -3.63 | - | COG1943/ Transposases and inactive derivates | 3e-07 |
| PP0579 | - | -3.48 | - | MazG. Predicted pyrophosphatase | 9e-06 |
| PP0710 | - | - | -2.23 | COG4319. Ketosteroid isomerase homolog | 1e-09 |
| PP0728 | -21.23 | - | - | PEP-utilizers, mobile domain/ PtsA, Phosphoenolpyruvate-protein kinase (PTS system EI component in bacteria) | 3e07/ 2e06 |
| PP0760 | - | - | -2.66 | UPF0016. Integral membrane proteins of unknown function | 8e-35 |
| PP0765 | -2.66 | -3.11 | - | DUF1302. *Pseudomonas* | 0.0 |
| PP0766 | -2.27 | -2.79 | - | DUF1329. *Pseudomonas* | 0.0 |
| PP0889 | - | -3.65 | - | AmpC/ Beta-lactamase | 1e-23/ 1e-16 |
| PP0890 | - | -3.04 | - | Conserved hyphotetical protein | - |
| PP0913 | - | - | -2.82 | IrpA- iron-regulated protein | 7e-155 |
| PP0982 | - | - | -2.66 | YjgP_YjgQ, Predicted permease | 7e-31 |
| PP0998 | -2.14 | -8.77 | - | Acetyl-CoA carboxylase alpha subunit (*P.fluorescens* PfO-1) | 3e-24 |
| PP1218 | - | - | -2.01 | 4HBT. Thioesterase superfamily/ FcbC. Predicted thioesterase | 1e-10/ 6e-34 |
| PP1364 | -1.98 | -3.7 | - | Conserved hyphotetical protein | - |
| PP1404 | - | - | -1.97 | DUF477 | 3e-12 |
| PP1514 | - | -7.21 | - | Conserved hyphotetical protein | - |
| PP1659 | -2.4 | -3.06 | - | Cytochrome c (*Synechococcus elongatus* PCC 7942) | 4e-39 |
| PP1828 | - | - | -4.78 | Conserved hyphotetical protein | - |
| PP2012 | - | - | -1.81 | NAD-kinase/ Predicted sugar kinase | 3e-52/ 1e-72 |
| PP2059 | - | -3.63 | 2.27 | CsbD like. Stress protein σB dependent | 0.004 |
| PP2161 | -2.18 | -4.58 | - | Conserved hyphotetical protein | - |
| PP2172 | -1.63 | - | -2.06 | Conserved hyphotetical protein | - |
| PP2309 | - | -3.42 | - | XtmA, Phage terminase, small subunit | 7e-06 |
| PP2314 | -4.2 | - | -1.97 | Conserved hyphotetical protein | - |
| PP2345 | - | - | -1.9 | Predicted membrane protein(*Pseudomonas fluorescens* Pf0-1 ) | 5e-24 |
| PP2346 | - | - | -1.98 | Propeptide, PepSY and peptidase M4 (*Pseudomonas syringae* pv. syringae B728a) | 6e-23 |
| PP2839 | - | - | -2.45 | DEAD-like helicases superfamily (1) / DinG, Rad3-related DNA helicases (Transcription / DNA replication, recombination, and repair) | 7e-09/ 1e-51 |
| PP3091 | - | - | -2.46 | ImcF-related (2) | 2e-64 |
| PP3134 | - | -3.88 | - | LbetaH, Left-Handed Parallel beta-Helix/ WbbJ, Acetyltransferase | 0.001/ 3e-17 |
| PP3145 | - | -4.43 | - | Ribosomal protein L7/L12 (*Pseudomonas aeruginosa* UCBPP-PA14) | 2e-61 |
| PP3231 | - | -4.36 | - | Erythromycin esterase | 2e-48 |
| PP3235 | -2.45 | -16.32 | -1.93 | Predicted signal-transduction protein containing cAMP-binding and CBS domains (3) | 9e-07 |

| | | | | | |
|---|---|---|---|---|---|
| PP3799 | - | -1.77 | -2.28 | Enolase (*Rhodospirillum rubrum*) | 7e-29 |
| PP4054 | - | -9.73 | - | Conserved hyphotetical protein | - |
| PP4086 | -2.34 | -4.59 | - | Conserved hyphotetical protein | - |
| PP4102 | - | -3.52 | - | DUF488 | 5e-27 |
| PP4138 | - | -3.21 | - | NADPH-dependent FMN reductase/ Predicted flavoprotein | 2e-19/ 5e-29 |
| PP4227 | - | -4.2 | - | sensor histidine kinase (*Pseudomonas putida* KT2440) | 6e-49 |
| PP4260 | - | -3.84 | - | FixH (4)/ Predicted integral membrane protein linked to a cation pump | 7e-08/ 1e-04 |
| PP4331 | -2.05 | -1.57 | -2.09 | Methyl-accepting chemotaxis protein (*Pseudomonas fluorescens* PfO-1) | 4e-44 |
| PP4362 | - | -2.25 | -2.39 | HPT, Histidine Phosphotransfer domain (5)/ ArcB, FOG: HPt domain | 2e-04/ 6e-06 |
| PP4593 | - | -3.54 | - | Conserved hyphotetical protein | - |
| PP4609 | - | - | -2.31 | Conserved hyphotetical protein | - |
| PP4703 | -1.71 | -1.81 | -2.19 | COG3662 | 4e-59 |
| PP5156 | -3.17 | -3.33 | - | ATPases of the AAA+ class (*Pseudomonas aeruginosa* UCBPP-PA14) | 1e-32 |
| PP5392 | - | -3.24 | - | COG2706, 3-carboxymuconate cyclase/ Cytochrome D1 heme domain(6) | 2e-09/ 3e-08 |

Table 11 : Aromatic metabolism

| TIGR or GeneBank(*) ID | Gene name | Description | Fold change | | |
|---|---|---|---|---|---|
| | | | Toluene | *o*-xylene | 3MB |
| PP1948 | | benzaldehyde dehydrogenase | 1.3 | 1.4 | 25.62 |
| PP2329 | pabB | para-aminobenzoate synthase component I | 1.87 | 3.26 | 1.19 |
| PP2679 | | quinoprotein ethanol dehydrogenase putative | 3.26 | 2.56 | - |
| PP3161 | benA | benzoate dioxygenase alpha subunit | 2.7 | 5.78 | 9.96 |
| PP3162 | benB | benzoate dioxygenase beta subunit | 1.93 | 3,042 | 3.62 |
| PP3163 | benC | benzoate dioxygenase ferredoxin reductase | 2.2 | 2.26 | 6.92 |
| PP3164 | benD | cis-diol dehydrogenase | 2.93 | 1.82 | 3.81 |
| PP3165 | benK | benzoate MFS transporter BenK | 1.4 | 1.73 | 3.86 |
| PP3166 | | catechol 1 2-dioxygenase | 3.1 | 3.66 | 3.13 |
| PP3167 | benE-2 | benzoate transport protein | 1.4 | 2.8 | 2.2 |
| PP3168 | benF | benzoate-specific porin | 2.73 | - | 2.8 |
| U20269* | xylU | unkown | 342.24 | 119.78 | 32.23 |
| U20269* | xylW | benzyl alcohol dehydrogenase II | 663.07 | 1050.46 | 1.58 |
| D66341* | xylC | benzaldehyde dehydrogenase | 151.73 | 20.18 | 3.2 |
| D66341* | xylM | toluene monooxygenase | 13.36 | 108.67 | 1.2 |
| D66341* | xylA | toluene monooxygenase | 410.64 | 15.96 | 19.9 |
| D66341* | xylB | benzyl alcohol dehydrogenase | 519.02 | 145.88 | 8.03 |
| D66341* | xylN | XylN outer membrane protein | 417.62 | 20.03 | 12.01 |
| M64747* | xylX | toluate 1,2-dioxygenase alpha subunit | 29.02 | 8.69 | 7.76 |
| M64747* | xylY | toluate 1,2-dioxygenase beta subunit | 6.66 | 1.32 | 4 |
| M64747* | xylZ | toluate 1,2-dioxygenase electron transfer component | 51.34 | 3.64 | 7.12 |
| M64747* | xylL | cis-1,2-dihydroxycyclohexa-3,4-diene carboxylate | 23.85 | 4.08 | 6.44 |
| M64747* | xylT | chloroplast-type ferredoxin | 4.64 | 2.16 | 3.76 |
| M64747* | xylE | catechol 2,3-dioxygenase | 2.24 | 1 | 1.59 |
| M64747* | xylG | 2-hydroximuconic semialdehyde dehydrogenase | 2.39 | 1.13 | 1.66 |
| M64747* | xylF | 2-hydroximuconic semialdehyde hydrolase | 2.1 | 1.18 | 1.48 |
| M64747* | xylJ | 2-oxo-4-pentenoato hydratase | 2.96 | 1.6 | 1.5 |
| M94186* | xylQ | acetaldehyde dehydrogenase | 1.9 | 1.01 | 1.69 |
| M94186* | xylK | 4-hydroxy-2-oxovalerate aldolase | 4.75 | 2.12 | 3.73 |
| M94186* | xylI | 4-oxalocrotonate decarboxylase | 6.65 | 9.62 | 6.31 |
| M94186* | xylH | 4-oxalcrotonate tautomerase | 9.47 | 14.09 | 11.81 |
| M15739* | xylS | regulatory protein | 5.77 | 5.58 | 2.73 |
| M15740* | xylR | regulatory protein | -1.6 | -1.59 | 2.15 |

**Figure 1.** Cluster analysis of heat-shock response



# GENERAL DISCUSSION

Microorganisms in nature are exposed to a series of variable stimuli, among which the presence of toxic aromatic compounds, originated from decomposition of green plants or resulting from human industrial activities, can make their environment far from optimal (Ramos *et al*., 2001). Many aromatic compounds are transformed by microorganisms, which may completely mineralize these carbon sources through different metabolic pathways (Sikkema *et al*., 1995). However, aromatic compounds are also extremely toxic to microbial cells at very low concentrations. Pseudomonads are able to metabolize a wide range of noxious organic compounds, which act as chaotropes, destabilizing cellular macromolecules and inhibiting growth (Hallsworth *et al*., 2003). The bacterial cell envelope is the first barrier against environmental stresses. Bacterial cytoplasmic membrane senses and copes with toxic aromatics before they exert damage on cytoplasmic proteins or contact specific pathway regulators (Ramos *et al*., 1997; Ramos *et al*., 2001). Aromatics are known to accumulate in and disrupt the bacterial cell membrane affecting the cell integrity and its functionality. Their interaction with biological membranes leads to changes in membrane structure and function, causing the “expansion” of the membrane surface area, the inhibition of membrane-embedded enzymes, such as ion pumps, and the increase in proton permeability (Sikkema *et al*., 1992). The overall consequence of these changes leads to a decrease in cell viability. The advances in high-throughput analytical technologies including proteomics, transcriptomics and metabolomics provide new insights into the global response of microorganisms to these toxic carbon sources and on how the microorganisms tackle with their noxious effects to avoid cell death (Del Castillo *et al*., 2007; Santos *et al*., 2004; Segura *et al*., 2005; Velázquez *et al*., 2005).

In the past decades, research on aromatic biodegradation has focused on the specific degradation pathways and their regulation. Despite the enormous amount of detailed information available on the regulatory processes and mechanisms, little is known about the initial response of *Pseudomonas* to the presence of toluene, i.e., on how the cell copes with the presence of these compounds. In this work, we have addressed for the first time *Pseudomonas* KT2440 the initial response to toluene from a global point of view. A general conclusion of our work establishes that appropriate induction of a general stress response is pivotal in the cell reaction mechanism triggered by the aromatic, and that the alternative RNA polymerase $E\sigma^{32}$ is essential in this response. Interestingly, the *meta*-cleavage pathway for the degradation of the aromatic 3MB present in this strain has been shown to require this alternative RNA polymerase for transcription. It seems thus that integration of the plasmid encoded TOL pathway within host metabolism has taken advantage of the ability of *Pseudomonas* to face aromatics and of the strategy used in its defense. As second objective of this work, we have analyzed in detail the interactions of RNA polymerase and XylS regulator with the Pm promoter. Our results support a general activation mechanism for this promoter in which 3MB would play an essential role.

## 1. Stress vs. catabolic responses in *P. putida* KT2440 cells exposed to organic solvents.

When *Pseudomonas putida* encounters toluene in the medium, this aromatic compound is sensed both as carbon source and as toxic drug that causes membrane damages and protein misfolding, finally leading to cell death. The transcriptional metabolic response to this aromatic in *P. putida* bearing pWW0 TOL plasmid involves two catabolic pathways: the so-called *upper* and *meta*-cleavage operons (Chapter 5, Fig. 1 and 4). Transcriptional regulation of these two clusters involves both TOL plasmid-encoded regulators (XylR and XylS), and host factors such as IHF global regulator and several alternative sigma factors, all of which are also subject to distinct global regulation mechanisms including catabolite repression mediated by alternative carbon sources or sigma factor competition (Aranda-Olmedo *et al*., 2005; Duetz *et al*., 1996; Velázquez *et al*., 2004). However, the activation of catabolism is not the only effect of aromatic compounds on *P. putida* cells, but the exposure to these compounds has many others deleterious effects that trigger a solvent tolerance response involving a number of processes. The majority of the mechanisms involved are not yet fully understood; nevertheless, several effects of toluene on the physiology of cells had been observed. The changes in the integrity of the membrane as a result of the accumulation of lipophilic compounds originated the leakage of macromolecules and an increase in permeability of the cytoplasmic membrane (de Smet *et al*., 1978; Sikkema *et al*., 1994). Solvent-tolerant bacteria have developed tolerance mechanisms, including efflux pumps or *cis* to *trans* isomerization of the unsaturated fatty acids or changes in the phospholipids head-group composition, or *cis* to *trans* isomerization of the unsaturated fatty acids or changes in the phospholipids head-group composition (Kieboom *et al*., 1998; Kim *et al*., 1998; Mosqueda and Ramos, 2000; Ramos *et al*., 1998; Rojas *et al*., 2001).

Using a genome-wide DNA array setup to identify changes in gene expression pattern in *P. putida* KT2440 (pWW0) in the presence of toluene, we observed that in addition to the specific metabolic response at the level of degradation pathways, exposure to aromatic compounds unleashed a general stress response, as denoted by the induction of a set of genes known to respond to membrane damage, oxidative stress, damaged DNA and misfolding of soluble proteins. To distinguish between metabolic and stress responses, we compared the effect of toluene with *o*-xylene and 3MB. 3MB is a metabolizable aromatic of low toxicity, while *o*-xylene cannot be metabolized through the TOL pathway and shows a higher degree of toxicity. According to logP values (toluene [$logP_{o/w}$ = 2.5], *o*-xylene [$logP_{o/w}$ = 2.77], and 3MB [$logP_{o/w}$=2.12], one would expect that the relative toxicity of these chemicals followed the order *o*-xylene> toluene>> 3MB (Sikkema *et al*., 1995). Actually, cluster analysis of the response to each of the three aromatic compounds showed that the number and intensity of the reprogrammed functions shared by *o*-xylene and toluene clearly grouped these two compounds together, and distinguished the response to 3MB (Chapter 5, Table 1 and Supplemental material Fig. 1).

The first bacterial barrier against the organic solvents is the cell envelope. In fact, exposure to the above mentioned organic solvents injured the membranes leading to the generation of active oxygen species in the respiratory chains (Chapter 5, Table 4). In response to this oxidative stress, several genes in the OxyR and SoxR regulons, considered to be involved in cell protection against reactive oxygen species (ROS), were induced (Ferrante *et al*., 1995; Greenberg *et al*., 1991; Michán *et al*., 1999). In this context, inhibition of transcription of iron-receptor protein genes (outer membrane ferric receptors) was concomitant with transcriptional induction of genes coding for iron-chelating proteins. It is known that iron reacts with hydrogen peroxide to generate ROS. In a second step*, P. putida* cells increased the general synthesis of proteins and compounds that allow the bacteria to neutralize the reactive oxygen species, such as glucose-6-phosphate-1-dehydrogenase, which increases the reducing power, and the synthesis of glutathione, essential in ROS detoxification (Chapter 5, Table 3 and Supplemental material Tables 1, 4, 5 and 7). Finally, we have found transcriptional changes in DNA repair systems and alkylhydroperoxidases (AhpCF) implicated in repair of nucleic acid, lipids and protein damages generated by ROS. Several genes involved in cell envelope biosynthesis were induced: *LpxAB*, which catalyzes lipid A biosynthesis and *murD* and *murG*, linked to cell wall biosynthesis (Chapter 5, Supplemental material Table 1). Induction of several RND efflux pumps associated with the extrusion of toxic chemicals had been previously reported for different *P. putida* strains (Kieboom *et al*., 1998; Mosqueda and Ramos, 2000; Ramos *et al*., 1998; Rojas *et al*., 2001). Our results confirmed the induction of *ttgABC* operon and we found an additional efflux pumps induced by these organic solvents, which could probably play a direct role in KT2440 strain solvent tolerance (Chapter 5, Table 2). Current projects to investigate the role of these efflux systems in *Pseudomonas putida* tolerance to a toluene shock are underway.

We have also found a strongly reduced expression of flagellum components, which was evidenced by a reduction in swimming motility (Chapter 5, Fig. 5 and Supplemental material Table 2). Although this seems to be in contradiction with the positive chemotatic response of *P. putida* towards toluene, this chemotatic response is known to be concentration-dependent, and only takes place until optimal toluene concentration is reached, probably below the concentrations used in our analysis (Parales and Harwood, 2002; Vardar *et al*., 2005). Flagellar and chemotaxis genes repression also occurs in response to a number of different stresses (Li *et al*., 1993; Liu *et al*., 2005; Maurer *et al*., 2005; Ruberg *et al*., 2003). This fact could be reflecting a general mechanism to save energy and amino acid resources to synthesize more useful components for stress survival. In fact, the extra-amino acid pool requirement of cells that induced a number of short-term adaptation mechanisms in order to survive to a solvent shock, is reflected in the up-regulation of amino acid and ribosomal proteins biosynthesis.

Concomitant with the global responses described above, *P. putida* strongly induced a specific pathway as response to stress: the so-called heat shock response.

As previously mentioned, *o*-xylene and toluene are known to be more toxic than 3MB. As expected, all three aromatics up-regulated genes belonging to the RpoH regulon and involved in cell protection and repair mechanisms against a variety of stresses (i.e. GroESL, IbpA, DnaK, DnaJ, HslUV) (Chapter 5, Table 5). As a consequence, $\sigma^{32}$ protein was stabilized and its concentration increased in the cell, allowing a more favorable competition for RNAP core (Chapter 5, Fig. 6). The amount of RNAP available in cells is limited. According to this limitation, changes in transcription activation patterns resulting from toxic chemical exposure require the reassignment of the transcription machinery from stress-survival dispensable functions to a set of genes involved in stress endurance. This reassignment could be responding to specific regulators that lead the RNAP to activate a specific set of genes or could be reached by changing the levels of alternative sigma factor such as $\sigma^{32}$ and $\sigma^{38}$ stress factors. In fact, this has been proved for $\sigma^{32}$. According to this line of argument, it seems logical to relate Pm promoter dependence on the two alternative RNA polymerases to a host strategy to link aromatic metabolism expression with general heat shock response of *Pseudomonas putida* cells.

In the past 5 years different works using high-throughput technologies such as proteomics and transcriptomics have reached not always coincident conclusions about *Pseudomonas putida* KT2440 response to organic solvents. A significant agreement is observed with *P. putida* response to the presence of phenol (Santos *et al*., 2004). The disagreements are probably due to combined minor differences in experimental design, such as variations in culture medium, growth conditions, oxygen availability and toxic compound concentration. The interpretation of microarray data also reflects differences in DNA chip features and in statistical methods used to compare gene transcription levels. In conclusion, data from DNA microarrays and proteomic analysis have to be cautiously interpreted and verified using alternative experimental approaches.

**2. <u>Two-alternative RNA polymerases share a single transcriptional site at Pm promoter.</u>**

Expression from Pm promoter is maintained at high level along the growth curve by a switch between two stress sigma factors: $\sigma^{32}$ in the early exponential phase and $\sigma^{38}$ in the late exponential and stationary phases (Marqués *et al*., 1999). The particular characteristics of $\sigma^{32}$ synthesis pathway imply that Pm transcription requires heat shock response induction in order to increase $\sigma^{32}$ levels. The increase in $\sigma^{32}$ levels observed after exposure to 3MB, toluene or *o*-xylene in microarray experiments with *P. putida* KT2440 (pWW0) cells (Chapter 5, Fig. 6), confirmed that these three compounds triggered a fast heat shock response, allowing in this way Pm transcription induction. In addition, 3MB would be required for XylS protein activation.

In the Pm promoter, the -10 and the -35 sequence regions diverged considerably from the consensus sequences defined for $\sigma^{32}$, $\sigma^{38}$ and $\sigma^{70}$ factors (Chapter 1, Fig. 1). This divergence in the -35 region could be partially explained by the fact that XylS

binding sites overlap with the RNA polymerase recognition site at -35 (González-Pérez *et al*., 2002). Considering the unique dependence on two different alternative sigma factors, it is derived that the -10 region of Pm should contain the essential determinants for recognition by the two polymerases involved, $E\sigma^{32}$ and $E\sigma^{38}$. One of the main objectives at the beginning of this work was to decipher the critical features that allow this specific recognition of Pm promoter sequence by two alternative RNA polymerases. We rationalized that Pm nucleotides important for transcription by $E\sigma^{32}$ would be essential in the exponential phase in a wild-type mutant background (MC4100) and also in a $\sigma^{32}$-proficient, $\sigma^{38}$-deficient background (RH90). On the other hand, nucleotides important to Pm recognition by $E\sigma^{38}$ would be important in the stationary phase in MC4100 and also in a $\sigma^{38}$-proficient, $\sigma^{32}$-deficient background (KY1429). We analyzed transcription activity from all possible Pm single mutants between positions -7 and -18 along the growth curve in different sigma mutant backgrounds (Chapter 1, Fig. 4 and 6). In the analysis of our data, we also took into account differences in $\sigma^{32}$ and $\sigma^{38}$ levels between the different genetic backgrounds that would influence the proportion of core RNA polymerase associated to each sigma factor (Chapter 1, Fig. 2 and 3). In addition to a precise definition of the -10 element, our results confirmed that $\sigma^{32}$ and $\sigma^{38}$ shared some recognition elements in the promoter, while some positions were crucial for the discrimination between the two RNAP. In fact Pm sequence at positions -9 and -8 (GC) allows recognition by $E\sigma^{32}$, preventing $E\sigma^{38}$ competition for Pm binding, whilst -10 position was the only Pm position key for $\sigma^{38}$ recognition. Four nucleotides upstream from the -10 element, a four C sequence matching the $\sigma^{32}$ consensus resulted crucial for $E\sigma^{32}$ recognition. The general increase in activity observed in KY1429 strain with most tested Pm mutant promoters, indicates that Pm promoter sequence is far from optimal for $E\sigma^{38}$ recognition. In that situation, both the absence of $\sigma^{32}$ competition for core and changes in promoter sequence allow transcription by $E\sigma^{38}$ to increase. Pm promoter sequence seems to have evolved to adapt the *meta*-cleavage pathway expression to the heat shock response triggered by the presence of toluene or 3MB. The increase in $\sigma^{32}$ level is strong but transient, so Pm promoter sequence is optimized for $E\sigma^{32}$ recognition. After this initial heat shock response, Pm transcription is maintained through activity of a second alternative RNAP also associated with a stress sigma factor, $\sigma^{38}$, which appears latter after the heat shock. Despite the fact that *in vivo* Pm transcription dependence on $E\sigma^{32}$ and $E\sigma^{38}$ has not been determined in a *P. putida* genetic background, our results have shown that RNAP holoenzymes reconstituted with *P. putida* $\sigma^{32}$ and $\sigma^{38}$ factors, were both able to promote Pm transcription *in vitro*, always remaining dependent on the presence of the XylS regulator (Chapter 1, Fig. 7).

In conclusion, Pm promoter activation mechanism depends on 3MB presence to activate XylS protein and to increase stress sigma factor level, which would efficiently compete for core binding and promoter recognition.

## 3. XylS transcription activation mechanism.

In *P. putida* XylS protein is constitutively expressed at low level and becomes transcriptionally active after binding a benzoate effector such as 3MB. In addition, when XylS is overproduced through a toluene-induced regulatory circuit, it stimulates Pm transcription even in the absence of benzoate effectors. XylS belongs to the AraC family of transcriptional regulators (Gallegos *et al*., 1997; Tobes and Ramos, 2002) and is composed of two separate functional domains. The N-terminal domain is involved in effector recognition, as deduced from the isolation of XylS mutants with altered effector specifity (Michán *et al*., 1992; Ramos *et al*., 1990; Ruíz and Ramos, 2002). It has also been shown that this domain is involved in dimerization. Two critical leucine residues (Leu193 and Leu194) plus the Ile205 residue were identified as essentials for XylS dimerization and in the end also for XylS transcriptional activation. The DNA binding domain of XylS is located in the C-terminal end of the protein. X-ray crystallographic structures of MarA and Rob show that the DNA binding domain is composed of seven α-helices folding in two HTH domains, that bind two adjacent segments of the major groove (Kwon *et al*., 2000; Rhee *et al*., 1998). XylS domains are functionally independent. In fact, XylS DNA binding domain, unlike most AraC family proteins except RhaS, is able to activate transcription despite the absence of XylS N-terminal domain determinants (Chapter 2, Fig. 7 and Chapter 4, Fig. 2) (Wickstrum *et al*., 2007).

Most AraC proteins recognize specific sequences of 15-20 bp in length at their target promoters (Brunelle and Schleif, 1989; Caswell *et al*., 1992; Egan and Schleif, 1994; Martin *et al*., 1996). Previous work defined XylS binding sites at Pm as two 15-bp direct repeats (TGCA-$N_6$-GGNTA), where each binding site can be divided in two sequence boxes A (TGCA) and B (GGNTA). The arrangement of the two repeats is such that the proximal XylS binding site overlaps by 1-bp the RNA polymerase -35 recognition element (González-Pérez *et al*., 2002). The GG-T sequence within the GGNTA box B was the most critical for XylS-Pm interactions (González-Pérez *et al*., 1999). Interestingly, Pm promoter exhibits an intrinsic curvature centred in the A-track located between proximal boxes A and B, with an apparent bent angle of 35º±5º, also observed *in vivo* (Gallegos *et al*., 1996; González-Pérez *et al*., 2002).
Because of XylS extremely poor solubility, we decided to address XylS activation mechanism analyzing the two domains independently. We overexpressed and purified truncated versions of both domains (XylS-N and XylS-C). As mentioned above, XylS-C was able to activate transcription irrespective of 3MB. XylS-C binding to DNA fragments including one or two XylS binding sites revealed the formation of one or two complexes (CI and CII), respectively (Chapter 4, Fig. 3). These two nucleoprotein complexes corresponded to DNA bound to one or two XylS-C monomers (Chapter 4, Fig. 4). On the other hand, DNA targets lacking either A or B sequence boxes displayed severely diminished XylS-C binding affinities, suggesting that both sequence motifs are important for XylS recognition (Chapter 4, Fig. 5). XylS-C monomers bound sequentially to Pm promoter, as denoted by progressive

protection of proximal and distal sites in DNA footprint assays: at low protein concentration only the XylS proximal site was occupied, whereas XylS distal site appeared protected with increasing XylS-C concentration (Chapter 4, Fig. 8). The sequential binding of XylS-C monomers probably reflects the existing difference between binding affinities to proximal and distal sites (Chapter 4, Fig. 9). Additionally, our results using different Pm mutant promoters showed that XylS binding to proximal site is more critical for transcription activation (Chapter 4, Fig. 6 and Chapter 3, Fig. 2). This fact could be revealing the importance of contacts established between XylS proximal monomer and the RNAP, which would be hampered in the case of improper proximal XylS-C positioning.

Our EMSA assays with a two-binding site DNA fragment pointed to a possible cooperative effect in XylS-C monomers binding (Chapter 4, Fig. 9). When the first XylS-C monomer bound to Pm proximal site, Pm intrinsic curvature raised from 35º to 50º. Concomitantly with this increase, the bent centre shifted towards the DNA region between XylS binding sites (Chapter 4, Fig. 7). Therefore, binding of the second XylS-C monomer could be facilitated either by changes in DNA structure or by contacts between XylS-C monomers, being these two possibilities not mutually exclusive. This positive cooperative effect, probably consequence of the gradual increase in Pm bending, would raise complex stability. Considering that full-length XylS protein behaves as a dimer in its active form, the cooperativity would be reflected in a higher binding affinity as compared to XylS-C monomers, in a similar way to what happens with AraC proteins (Timmes *et al*., 2004). Pm curvature increased to 98º with the binding of the second XylS-C monomer, process that is likely to further contribute to the establishment of XylS-RNAP contacts, which in the end would lead to transcription activation. Changes in Pm structure provoked by XylS induced bending could favor the formation of a larger number of protein-DNA contacts at the major grooves. Actually, our results indicate that sugar-phosphate backbone contacts greatly contribute to XylS binding strength. Pm curvature around XylS monomers would provide “shape complementarity”, enhancing nucleoprotein stability. In addition, XylS establishes base specific contacts that are on the basis of unambiguous recognition of Pm direct repeats (Chapter 3, Fig. 5 and 6). Related to these base-specific contacting residues located at conserved positions, several aromatic-related XylS homologues, including BenR, IpbR, EbdR and BphS, share this amino acid-promoter sequence correlation with XylS (Chapter 3, Fig. 7), which points to co-evolution of protein and promoters. In fact, we and others have detected cross-regulation between XylS and BenR proteins and their respective promoters (Cowles *et al*., 2000) (Chapter 5, Table 1).

XylS binding orientation corresponds to the previously defined “forward orientation”, in which the XylS C-terminal HTH binds adjacent to the -35 region of Pm promoter. This arrangement defines the protein regions exposed to make contacts with RNAP. Actually, MarA, Rob and SoxS establish contacts with RNAP through different protein regions depending on the binding orientation they adopt at various

target promoters (Martin *et al*., 1999). XylS proximal binding site overlaps by 1-bp the -35 region at Pm promoter and is thus classified as a class-II promoter activator. Several AraC family class-II activators have been reported to make contacts with sigma subunit region 4 through a conserved region located at the C-terminal domain α-helix 5 (Bhende and Egan, 2000; Grainger *et al*., 2004; Wickstrum and Egan, 2004). Preliminary results using XylS mutants at equivalent positions in α-helix 5 showed a severe impairment of transcription activation.

Along this work we have also tried to decipher XylS activation mechanism in response to 3MB. Previous works indicated that XylS dimer formation was enhanced *in vivo* by the presence of 3MB, but became an effector independent process at high protein concentration *in vitro* (Ruíz *et al*., 2003). XylS DNA binding capacity followed the same behavior in response to 3MB: DNA binding was favored *in vivo* by the effector (Chapter 2, Fig. 5), but it turned 3MB independent in our *in vitro* assays (Chapter 2, Fig. 2). In this connection, a XylS mutant unable both to dimerize and to activate transcription required the presence of 3MB to bind DNA (Chapter 2, Fig. 3). On the other hand, XylS-C was able to activate transcription when overproduced in the absence of 3MB, suggesting a possible intramolecular repression of XylS-N upon XylS-C DNA binding (Chapter 4, Fig. 2). Gel shift assays where XylS independently purified domains were pre-incubated together, confirmed intramolecular repression and made it clear it relayed on direct contacts between domains (Chapter 2, Fig. 4). Furthermore, intramolecular repression was released in the presence of 3MB. Based on all this set of results, propose a model for XylS activation in response to 3MB (Chapter 2, Fig. 8). In the absence of 3MB, direct interaction between N- and C-terminal domains in XylS maintains the protein in an inactive state. The addition of 3MB releases N-terminal domain repression upon XylS-C, allowing XylS to bind DNA and to activate transcription. 3MB binding to XylS to derepress DNA binding domain also triggers the conformational change concomitant to dimerization, which can also occur in the absence of effector as a result of high XylS concentration. In the absence of 3MB, XylS overproduction would favor XylS C-terminal domain derepression, unmasking dimerization surface and DNA binding determinants, leading to Pm recognition. However, a XylS dimerization mutant was able to bind DNA in the presence of 3MB, but remained inactive in transcription. Therefore, dimerization stands as an essential process in transcription activation. Finally, the XylS in the presence of 3MB recruits RNAP to Pm promoter (Chapter 2, Fig. 5). As mentioned above, the presence of 3MB increases $\sigma^{32}$ amounts in the cell giving rise to a higher proportion of $E\sigma^{32}$ (Chapter 1, Fig. 2 and Chapter 5, Fig. 6). Thus, 3MB role in RNA polymerase recruitment would be exerted at several levels: activating XylS to bind DNA and recruit RNAP, and increasing the amount of RNAP available for transcription. RNA recruitment to Pm promoter is followed open complex formation and transcription initiation (Chapter 2, Fig. 6 and 7).

# CONCLUSIONES

# CONCLUSIONES

El conjunto de resultados obtenidos a lo largo de este trabajo de Tesis Doctoral nos ha permitido establecer las siguientes conclusiones:

1. La activación de la transcripción desde el promotor Pm requiere un aumento en los niveles de los factores sigma relacionados con estrés, $\sigma^{32}$ y $\sigma^{38}$, para incrementar la proporción de ARNP asociada a los mismos. Este aumento es consecuencia de la presencia del inductor, que desencadena la respuesta a estrés, además de activar al regulador XylS.

2. Las dos ARNP alternativas responsables de la transcripción desde el promotor Pm reconocen una única región -10 que aúna los determinantes necesarios para la unión de los factores $\sigma^{32}$ y $\sigma^{38}$. La secuencia en la región -10 presenta nucleótidos críticos para el reconocimiento por ambas ARN polimerasas, además de características que discriminan entre los factores $\sigma^{32}$ y $\sigma^{38}$. La secuencia del promotor Pm está optimizada para el reconocimiento por $E\sigma^{32}$, siendo $E\sigma^{38}$ una ARNP más permisiva a desviaciones con respecto a su secuencia consenso.

3. La proteína XylS consta de dos dominios independientes desde un punto de vista funcional, que interaccionan entre sí de manera que el dominio amino terminal reprime la capacidad del dominio carboxilo terminal de XylS de unirse al ADN. La unión de 3MB al dominio N-terminal libera la represión intramolecular, permitiendo al dominio C-terminal reconocer secuencias específicas en Pm y activar la transcripción.

4. El 3MB ejerce un segundo papel sobre el regulador XylS favoreciendo su dimerización, proceso esencial para la activación de la transcripción desde el promotor Pm. La sobreproducción del regulador XylS permite su dimerización de forma independiente a la presencia de 3MB, hecho que elimina la represión sobre la unión al ADN del dominio C-terminal de XylS.

5. En respuesta a la presencia de 3MB el regulador XylS recluta a la ARNP al promotor Pm, favoreciendo la formación de un complejo promotor abierto e induciendo el inicio de la transcripción.

6. La especificidad en la unión de XylS a sus secuencias operadoras en Pm viene determinada por contactos específicos de base. Los contactos que se establecen entre ambas hélices de reconocimiento y el esqueleto de fosfato del ADN proporcionan estabilidad a la unión. XylS se orienta en Pm adoptando una disposición cabeza-cola, en la que el primer HTH de cada monómero establece interacciones con el extremo distal de los sitios de unión mientras que el segundo HTH contacta con el extremo proximal. La orientación cabeza-cola determina la región de XylS más próxima a la ARNP.

7. El dominio carboxilo terminal de XylS se comporta como monómero en solución. La unión secuencial y cooperativa de dos monómeros del dominio carboxilo terminal a los sitios proximal y distal de Pm, respectivamente, provoca la curvatura gradual del ADN necesaria para el establecimiento de contactos entre activador y ARNP, críticos para la activación de la transcripción.

8. La respuesta inicial de *Pseudomonas putida* a la presencia de compuestos aromáticos tóxicos se detecta a nivel de la envuelta celular. Las lesiones provocadas en las membranas generan daño oxidativo a nivel celular. Finalmente, la desnaturalización de proteínas ocasionada por la presencia de compuestos aromáticos desencadena la respuesta clásica de choque térmico.

9. La presencia de compuestos aromáticos tóxicos y mineralizables por *Pseudomonas putida* tiene como consecuencia cambios en los patrones de transcripción globales, de manera que disminuye la expresión de genes asociados a funciones dispensables para la supervivencia celular, mientras que aumenta la transcripción de genes relacionados con el catabolismo de dichos compuestos y con la respuesta defensiva de la bacteria a los mismos.

# BIBLIOGRAFÍA

**Aiyar, S.E., McLeod, S.M., Ross, W., Hirvonen, C.A., Thomas, M.S., Johnson, R.C. and Gourse, R.L.** (2002) Architecture of Fis-activated transcription complexes at the *Escherichia coli rrnB* P1 and *rrnE* P1 promoters. *J Mol Biol*, **316**, 501-516.

**Aranda-Olmedo, I., Ramos, J.L. and Marqués, S.** (2005) Integration of signals through Crc and PtsN in catabolite repression of *Pseudomonas putida* TOL plasmid pWW0. *Appl Environ Microbiol*, **71**, 4191-4198.

**Arfin, S.M., Long, A.D., Ito, E.T., Tolleri, L., Riehle, M.M., Paegle, E.S. and Hatfield, G.W.** (2000) Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor. *J Biol Chem*, **275**, 29672-29684.

**Artsimovitch, I., Patlan, V., Sekine, S., Vassylyeva, M.N., Hosaka, T., Ochi, K., Yokoyama, S. and Vassylyev, D.G.** (2004) Structural basis for transcription regulation by alarmone ppGpp. *Cell*, **117**, 299-310.

**Bar-Nahum, G. and Nudler, E.** (2001) Isolation and characterization of $\sigma^{70}$-retaining transcription elongation complexes from *Escherichia coli*. *Cell*, **106**, 443-451.

**Barne, K.A., Bown, J.A., Busby, S.J. and Minchin, S.D.** (1997) Region 2.5 of the *Escherichia coli* RNA polymerase $\sigma^{70}$ subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J*, **16**, 4034-4040.

**Barrios, H., Valderrama, B. and Morett, E.** (1999) Compilation and analysis of $\sigma^{54}$-dependent promoter sequences. *Nucleic Acids Res*, **27**, 4305-4313.

**Belyaeva, T.A., Wade, J.T., Webster, C.L., Howard, V.J., Thomas, M.S., Hyde, E.I. and Busby, S.J.** (2000) Transcription activation at the *Escherichia coli melAB* promoter: the role of MelR and the cyclic AMP receptor protein. *Mol Microbiol*, **36**, 211-222.

**Benoff, B., Yang, H., Lawson, C.L., Parkinson, G., Liu, J., Blatter, E., Ebright, Y.W., Berman, H.M. and Ebright, R.H.** (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science*, **297**, 1562-1566.

**Berghofer-Hochheimer, Y., Lu, C.Z. and Gross, C.A.** (2005) Altering the interaction between $\sigma^{70}$ and RNA polymerase generates complexes with distinct transcription-elongation properties. *Proc Natl Acad Sci USA*, **102**, 1157-1162.

**Bernardo, L.M., Johansson, L.U., Solera, D., Skarfstad, E. and Shingler, V.** (2006) The guanosine tetraphosphate (ppGpp) alarmone, DksA and promoter affinity for RNA polymerase in regulation of sigma-dependent transcription. *Mol Microbiol*, **60**, 749-764.

**Bhende, P.M. and Egan, S.M.** (2000) Genetic evidence that transcription activation by RhaS involves specific amino acid contacts with $\sigma^{70}$. *J Bacteriol*, **182**, 4959-4969.

**Blaszczak, A., Georgopoulos, C. and Liberek, K.** (1999) On the mechanism of FtsH-dependent degradation of the $\sigma^{32}$ transcriptional regulator of *Escherichia coli* and the role of the DnaK chaperone machine. *Mol Microbiol*, **31**, 157-166.

**Bordes, P., Conter, A., Morales, V., Bouvier, J., Kolb, A. and Gutierrez, C.** (2003) DNA supercoiling contributes to disconnect $\sigma^{S}$ accumulation from $\sigma^{S}$-dependent transcription in *Escherichia coli*. *Mol Microbiol*, **48**, 561-571.

**Borukhov, S. and Nudler, E.** (2003) RNA polymerase holoenzyme: structure, function and biological implications. *Curr Opin Microbiol*, **6**, 93-100.

**Borukhov, S., Sagitov, V. and Goldfarb, A.** (1993) Transcript cleavage factors from *E. coli. Cell*, **72**, 459-466.

**Bougdour, A., Wickner, S. and Gottesman, S.** (2006) Modulating RssB activity: IraP, a novel regulator of $\sigma^{S}$ stability in *Escherichia coli*. *Genes Dev*, **20**, 884-897.

**Brown, N.L., Stoyanov, J.V., Kidd, S.P. and Hobman, J.L.** (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev*, **27**, 145-163.

**Browning, D.F., Beatty, C.M., Wolfe, A.J., Cole, J.A. and Busby, S.J.** (2002) Independent regulation of the divergent *Escherichia coli nrfA* and *acsP1* promoters by a nucleoprotein assembly at a shared regulatory region. *Mol Microbiol*, **43**, 687-701.

**Brunelle, A. and Schleif, R.** (1989) Determining residue-base interactions between AraC protein and *araI* DNA. *J Mol Biol*, **209**, 607-622.

**Buck, M., Gallegos, M.T., Studholme, D.J., Guo, Y. and Gralla, J.D.** (2000) The bacterial enhancer-dependent $\sigma^{54}$ ($\sigma^{N}$) transcription factor. *J Bacteriol*, **182**, 4129-4136.

**Burgess, R.R.** (1971) RNA polymerase. *Annu Rev Biochem*, **40**, 711-740.

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

**Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L. and Darst, S.A.** (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell*, **9**, 527-539.

**Cases, I. and de Lorenzo, V.** (1998) Expression systems and physiological control of promoter activity in bacteria. *Curr Opin Microbiol*, **1**, 303-310.

**Cashel, M., D. R. Gentry, V. J. Hernandez, and D. Vinella.** (1996) The stringent response, p. 1458-1495. In Neidhardt, F.C., R. Curtiss, J. L. Ingraham., E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington DC.

**Caswell, R., Webster, C. and Busby, S.** (1992) Studies on the binding of the *Escherichia coli* MelR transcription activator protein to operator sequences at the MelAB promoter. *Biochem J*, **287**, 501-508.

**Claret, L. and Rouviere-Yaniv, J.** (1996) Regulation of HU alpha and HU beta by CRP and FIS in *Escherichia coli*. *J Mol Biol*, **263**, 126-139.

**Constantinidou, C., Hobman, J.L., Griffiths, L., Patel, M.D., Penn, C.W., Cole, J.A. and Overton, T.W.** (2006) A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as *Escherichia coli* K12 adapts from aerobic to anaerobic growth. *J Biol Chem*, **281**, 4802-4815.

**Cowing, D.W., Bardwell, J.C., Craig, E.A., Woolford, C., Hendrix, R.W. and Gross, C.A.** (1985) Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc Natl Acad Sci USA*, **82**, 2679-2683.

**Cowles, C.E., Nichols, N.N. and Harwood, C.S.** (2000) BenR, a XylS homologue, regulates three different pathways of aromatic acid degradation in *Pseudomonas putida*. *J Bacteriol*, **182**, 6339-6346.

**Craig, M.L., Suh, W.C. and Record, M.T., Jr.** (1995) HO. and DNase I probing of $E\sigma^{70}$ RNA polymerase-lambda $P_R$ promoter open complexes: $Mg^{2+}$ binding and its structural consequences at the transcription start site. *Biochemistry*, **34**, 15624-15632.

**Chadsey, M.S. and Hughes, K.T.** (2001) A multipartite interaction between *Salmonella* transcription factor $\sigma^{28}$ and its anti-sigma factor FlgM: implications for $\sigma^{28}$ holoenzyme destabilization through stepwise binding. *J Mol Biol*, **306**, 915-929.

**Chatterji, D., Fujita, N. and Ishihama, A.** (1998) The mediator for stringent control, ppGpp, binds to the beta-subunit of *Escherichia coli* RNA polymerase. *Genes Cells*, **3**, 279-287.

**Choy, H., and S. Adhya.** (1996) Negative control, p. 1287-1299. In Neidhardt, F.C., R. Curtiss, J. L. Ingraham., E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington DC.

**de Smet, M.J., Kingma, J. and Witholt, B.** (1978) The effect of toluene on the structure and permeability of the outer and cytoplasmic membranes of *Escherichia coli*. *Biochim Biophys Acta*, **506**, 64-80.

**Decatur, A.L. and Losick, R.** (1996) Three sites of contact between the *Bacillus subtilis* transcription factor $\sigma^{F}$ and its antisigma factor SpoIIAB. *Genes Dev*, **10**, 2348-2358.

**Del Castillo, T., Ramos, J.L., Rodríguez-Herva, J.J., Fuhrer, T., Sauer, U. and Duque, E.** (2007) Convergent peripheral pathways catalyze initial glucose catabolism in *Pseudomonas putida*: Genomic and flux analysis. *J Bacteriol*. doi: 10.1128/JB.00203-07.

**Díaz, E. and Prieto, M.A.** (2000) Bacterial promoters triggering biodegradation of aromatic pollutants. *Curr Opin Biotechnol*, **11**, 467-475.

**Ding, Q., Kusano, S., Villarejo, M. and Ishihama, A.** (1995) Promoter selectivity control of *Escherichia coli* RNA polymerase by ionic strength: differential recognition of osmoregulated promoters by $E\sigma^D$ and $E\sigma^S$ holoenzymes. *Mol Microbiol*, **16**, 649-656.

**Dombroski, A.J., Johnson, B.D., Lonetto, M. and Gross, C.A.** (1996) The sigma subunit of *Escherichia coli* RNA polymerase senses promoter spacing. *Proc Natl Acad Sci USA*, **93**, 8858-8862.

**Domínguez-Cuevas, P. and Marqués, S.** (2004) Compiling sigma-70 dependent promoters. In Ramos, J.L. (ed.), *The Pseudomonads. Virulence and gene regulation*. Kluwer Academic/Plenum Publishers, New York, Vol. II, pp. 319-343.

**Dove, S.L., Darst, S.A. and Hochschild, A.** (2003) Region 4 of sigma as a target for transcription regulation. *Mol Microbiol*, **48**, 863-874.

**Drlica, K.** (1992) Control of bacterial DNA supercoiling. *Mol Microbiol*, **6**, 425-433.

**Duetz, W.A., Marqués, S., Wind, B., Ramos, J.L. and van Andel, J.G.** (1996) Catabolite repression of the toluene degradation pathway in *Pseudomonas putida* harboring pWW0 under various conditions of nutrient limitation in chemostat culture. *Appl Environ Microbiol*, **62**, 601-606.

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

**Egan, S.M. and Schleif, R.F.** (1993) A regulatory cascade in the induction of *rhaBAD*. *J Mol Biol*, **234**, 87-98.

**Egan, S.M. and Schleif, R.F.** (1994) DNA-dependent renaturation of an insoluble DNA binding protein. Identification of the RhaS binding site at *rhaBAD*. *J Mol Biol*, **243**, 821-829.

**El-Samad, H., Kurata, H., Doyle, J.C., Gross, C.A. and Khammash, M.** (2005) Surviving heat shock: control strategies for robustness and performance. *Proc Natl Acad Sci USA*, **102**, 2736-2741.

**Ellis, L.B.** (2000) Environmental biotechnology informatics. *Curr Opin Biotechnol*, **11**, 232-235.

**Enz, S., Mahren, S., Stroeher, U.H. and Braun, V.** (2000) Surface signaling in ferric citrate transport gene induction: interaction of the FecA, FecR, and FecI regulatory proteins. *J Bacteriol*, **182**, 637-646.

**Falconi, M., Brandi, A., La Teana, A., Gualerzi, C.O. and Pon, C.L.** (1996) Antagonistic involvement of FIS and H-NS proteins in the transcriptional control of *hns* expression. *Mol Microbiol*, **19**, 965-975.

**Ferrante, A.A., Augliera, J., Lewis, K. and Klibanov, A.M.** (1995) Cloning of an organic solvent-resistance gene in *Escherichia coli*: the unexpected role of alkylhydroperoxide reductase. *Proc Natl Acad Sci USA*, **92**, 7617-7621.

**Fish, R.N. and Kane, C.M.** (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochim Biophys Acta*, **1577**, 287-307.

**Fredriksson, A., Ballesteros, M., Peterson, C.N., Persson, O., Silhavy, T.J. and Nystrom, T.** (2007) Decline in ribosomal fidelity contributes to the accumulation and stabilization of the master stress response regulator $\sigma^S$ upon carbon starvation. *Genes Dev*, **21**, 862-874.

**Fu, J., Gnatt, A.L., Bushnell, D.A., Jensen, G.J., Thompson, N.E., Burgess, R.R., David, P.R. and Kornberg, R.D.** (1999) Yeast RNA polymerase II at 5 Å resolution. *Cell*, **98**, 799-810.

**Gaal, T., Ross, W., Estrem, S.T., Nguyen, L.H., Burgess, R.R. and Gourse, R.L.** (2001) Promoter recognition and discrimination by $E\sigma^S$ RNA polymerase. *Mol Microbiol*, **42**, 939-954.

**Gallegos, M.T., Marqués, S. and Ramos, J.L.** (1996) The $TACAN_4TGCA$ motif upstream from the -35 region in the $\sigma^{70}$-$\sigma^S$-dependent Pm promoter of the TOL plasmid is the minimum DNA segment required for transcription stimulation by XylS regulators. *J Bacteriol*, **178**, 6427-6434.

**Gallegos, M.T., Schleif, R., Bairoch, A., Hofmann, K. and Ramos, J.L.** (1997) Arac/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev*, **61**, 393-410.

**Ghosh, P., Ishihama, A. and Chatterji, D.** (2001) *Escherichia coli* RNA polymerase subunit ω and its N-terminal domain bind full-length β' to facilitate incorporation into the $\alpha_2\beta$ subassembly. *Eur J Biochem*, **268**, 4621-4627.

**Gillette, W.K., Rhee, S., Rosner, J.L. and Martin, R.G.** (2000) Structural homology between MarA of the AraC family of transcriptional activators and the integrase family of site-specific recombinases. *Mol Microbiol*, **35**, 1582-1583.

**González-Pérez, M.M., Marqués, S., Domínguez-Cuevas, P. and Ramos, J.L.** (2002) XylS activator and RNA polymerase binding sites at the Pm promoter overlap. *FEBS Lett*, **519**, 117-122.

**González-Pérez, M.M., Ramos, J.L., Gallegos, M.T. and Marqués, S.** (1999) Critical nucleotides in the upstream region of the XylS-dependent TOL *meta*-cleavage pathway operon promoter as deduced from analysis of mutants. *J Biol Chem*, **274**, 2286-2290.

**González-Pérez, M.M., Ramos, J.L. and Marqués, S.** (2004) Cellular XylS levels are a function of transcription of *xylS* from two independent promoters and the differential efficiency of translation of the two mRNAs. *J Bacteriol*, **186**, 1898-1901.

**Grainger, D.C., Aiba, H., Hurd, D., Browning, D.F. and Busby, S.J.** (2007) Transcription factor distribution in *Escherichia coli*: studies with FNR protein. *Nucleic Acids Res*, **35**, 269-278.

**Grainger, D.C., Belyaeva, T.A., Lee, D.J., Hyde, E.I. and Busby, S.J.** (2004a) Transcription activation at the *Escherichia coli melAB* promoter: interactions of MelR with the C-terminal domain of the RNA polymerase α subunit. *Mol Microbiol*, **51**, 1311-1320.

**Grainger, D.C., Hurd, D., Harrison, M., Holdstock, J. and Busby, S.J.** (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci USA*, **102**, 17693-17698.

**Grainger, D.C., Webster, C.L., Belyaeva, T.A., Hyde, E.I. and Busby, S.J.** (2004b) Transcription activation at the *Escherichia coli melAB* promoter: interactions of MelR with its DNA target site and with domain 4 of the RNA polymerase sigma subunit. *Mol Microbiol*, **51**, 1297-1309.

**Gralla, J.D.** (2005) *Escherichia coli* ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein. *Mol Microbiol*, **55**, 973-977.

**Greenberg, J.T., Chou, J.H., Monach, P.A. and Demple, B.** (1991) Activation of oxidative stress genes by mutations at the *soxQ/cfxB/marA* locus of *Escherichia coli*. *J Bacteriol*, **173**, 4433-4439.

**Griffith, K.L., Shah, I.M., Myers, T.E., O'Neill, M.C. and Wolf, R.E., Jr.** (2002) Evidence for "pre-recruitment" as a new mechanism of transcription activation in *Escherichia coli*: the large excess of SoxS binding sites per cell relative to the number of SoxS molecules per cell. *Biochem Biophys Res Commun*, **291**, 979-986.

**Griffith, K.L., Shah, I.M. and Wolf, R.E., Jr.** (2004) Proteolytic degradation of *Escherichia coli* transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol Microbiol*, **51**, 1801-1816.

**Griffith, K.L. and Wolf, R.E., Jr.** (2004) Genetic evidence for pre-recruitment as the mechanism of transcription activation by SoxS of *Escherichia coli*: the dominance of DNA binding mutations of SoxS. *J Mol Biol*, **344**, 1-10.

**Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J. and Young, B.** (1998) The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb Symp Quant Biol*, **63**, 141-155.

**Grossman, A.D., Taylor, W.E., Burton, Z.F., Burgess, R.R. and Gross, C.A.** (1985) Stringent response in *Escherichia coli* induces expression of heat shock proteins. *J Mol Biol*, **186**, 357-365.

**Gruber, T.M. and Bryant, D.A.** (1998) Characterization of the alternative sigma-factors SigD and SigE in *Synechococcus sp.* strain PCC 7002. SigE is implicated in transcription of post-exponential-phase-specific genes. *Arch Microbiol*, **169**, 211-219.

**Guisbert, E., Herman, C., Lu, C.Z. and Gross, C.A.** (2004) A chaperone network controls the heat shock response in *E. coli. Genes Dev*, **18**, 2812-2821.

**Hallsworth, J.E., Heim, S. and Timmis, K.N.** (2003) Chaotropic solutes cause water stress in *Pseudomonas putida*. *Environ Microbiol*, **5**, 1270-1280.

**Harmer, T., Wu, M. and Schleif, R.** (2001) The role of rigidity in DNA looping-unlooping by AraC. *Proc Natl Acad Sci USA*, **98**, 427-431.

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

**Heil, A. and Zillig, W.** (1970) Reconstitution of bacterial DNA-dependent RNA-polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. *FEBS Lett*, **11**, 165-168.

**Helmann, J.D.** (1995) Compilation and analysis of *Bacillus subtilis* $\sigma^A$-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res*, **23**, 2351-2360.

**Helmann, J.D.** (1999) Anti-sigma factors. *Curr Opin Microbiol*, **2**, 135-141.

**Helmann, J.D.** (2002) The extracytoplasmic function (ECF) sigma factors. *Adv Microb Physiol*, **46**, 47-110.

**Hengge-Aronis, R.** (2002) Signal transduction and regulatory mechanisms involved in control of the $\sigma^S$ (RpoS) subunit of RNA polymerase. *Microbiol Mol Biol Rev*, **66**, 373-395.

**Herman, C., Thevenet, D., D'Ari, R. and Bouloc, P.** (1995) Degradation of $\sigma^{32}$, the heat shock regulator in *Escherichia coli*, is governed by HflB. *Proc Natl Acad Sci USA*, **92**, 3516-3520.

**Hinton, D.M. and Vuthoori, S.** (2000) Efficient inhibition of *Escherichia coli* RNA polymerase by the bacteriophage T4 AsiA protein requires that AsiA binds first to free $\sigma^{70}$. *J Mol Biol*, **304**, 731-739.

**Holcroft, C.C. and Egan, S.M.** (2000) Roles of cyclic AMP receptor protein and the carboxyl-terminal domain of the alpha subunit in transcription activation of the *Escherichia coli rhaBAD* operon. *J Bacteriol*, **182**, 3529-3535.

**Hung, S.P., Baldi, P. and Hatfield, G.W.** (2002) Global gene expression profiling in *Escherichia coli* K12. The effects of leucine-responsive regulatory protein. *J Biol Chem*, **277**, 40309-40323.

**Ishihama, A.** (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol*, **54**, 499-518.

**Jacob, F. and Monod, J.** (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, **3**, 318-356.

**Jensen, K.F. and Pedersen, S.** (1990) Metabolic growth rate control in *Escherichia coli* may be a consequence of subsaturation of the macromolecular biosynthetic apparatus with substrates and catalytic components. *Microbiol Rev*, **54**, 89-100.

**Jishage, M. and Ishihama, A.** (1995) Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of sigma 70 and sigma 38. *J Bacteriol*, **177**, 6832-6835.

**Jishage, M., Iwata, A., Ueda, S. and Ishihama, A.** (1996) Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. *J Bacteriol*, **178**, 5447-5451.

**Jishage, M., Kvint, K., Shingler, V. and Nystrom, T.** (2002) Regulation of sigma factor competition by the alarmone ppGpp. *Genes Dev*, **16**, 1260-1270.

**Kahramanoglou, C., Webster, C.L., El-Robh, M.S., Belyaeva, T.A. and Busby, S.J.** (2006) Mutational analysis of the *Escherichia coli melR* gene suggests a two-state concerted model to explain transcriptional activation and repression in the melibiose operon. *J Bacteriol*, **188**, 3199-3207.

**Kajitani, M. and Ishihama, A.** (1984) Promoter selectivity of *Escherichia coli* RNA polymerase. Differential stringent control of the multiple promoters from ribosomal RNA and protein operons. *J Biol Chem*, **259**, 1951-1957.

**Kaldalu, N., Toots, U., de Lorenzo, V. and Ustav, M.** (2000) Functional domains of the TOL plasmid transcription factor XylS. *J Bacteriol*, **182**, 1118-1126.

**Kapanidis, A.N., Margeat, E., Laurence, T.A., Doose, S., Ho, S.O., Mukhopadhyay, J., Kortkhonjia, E., Mekler, V., Ebright, R.H. and Weiss, S.** (2005) Retention of transcription initiation factor $\sigma^{70}$ in transcription elongation: single-molecule analysis. *Mol Cell*, **20**, 347-356.

**Kawakami, K., Saitoh, T. and Ishihama, A.** (1979) Biosynthesis of RNA polymerase in *Escherichia coli*. IX. Growth-dependent variations in the synthesis rate, content and distribution of RNA polymerase. *Mol Gen Genet*, **174**, 107-116.

**Kieboom, J., Dennis, J.J., de Bont, J.A. and Zylstra, G.J.** (1998) Identification and molecular characterization of an efflux pump involved in *Pseudomonas putida* S12 solvent tolerance. *J Biol Chem*, **273**, 85-91.

**Kim, K., Lee, S., Lee, K. and Lim, D.** (1998) Isolation and characterization of toluene-sensitive mutants from the toluene-resistant bacterium *Pseudomonas putida* GM73. *J Bacteriol*, **180**, 3692-3696.

**Kingsford, C.L., Ayanbule, K. and Salzberg, S.L.** (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol*, **8**, R22.

**Kojic, M. and Venturi, V.** (2001) Regulation of *rpoS* gene expression in *Pseudomonas*: involvement of a TetR family regulator. *J Bacteriol*, **183**, 3712-3720.

**Kusano, S., Ding, Q., Fujita, N. and Ishihama, A.** (1996) Promoter selectivity of *Escherichia coli* RNA polymerase $E\sigma^{70}$ and $E\sigma^{38}$ holoenzymes. Effect of DNA supercoiling. *J Biol Chem*, **271**, 1998-2004.

**Kusano, S. and Ishihama, A.** (1997a) Functional interaction of *Escherichia coli* RNA polymerase with inorganic polyphosphate. *Genes Cells*, **2**, 433-441.

**Kusano, S. and Ishihama, A.** (1997b) Stimulatory effect of trehalose on formation and activity of *Escherichia coli* RNA polymerase $E\sigma^{38}$ holoenzyme. *J Bacteriol*, **179**, 3649-3654.

**Kvint, K., Farewell, A. and Nystrom, T.** (2000) RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of sigma(s). *J Biol Chem*, **275**, 14795-14798.

**Kvint, K., Nachin, L., Díez, A. and Nystrom, T.** (2003) The bacterial universal stress protein: function and regulation. *Curr Opin Microbiol*, **6**, 140-145.

**Kwon, H.J., Bennik, M.H., Demple, B. and Ellenberger, T.** (2000) Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nat Struct Biol*, **7**, 424-430.

**Lacour, S., Kolb, A. and Landini, P.** (2003) Nucleotides from -16 to -12 determine specific promoter recognition by bacterial $\sigma^{S}$-RNA polymerase. *J Biol Chem*, **278**, 37160-37168.

**Landini, P. and Busby, S.J.** (1999) Expression of the *Escherichia coli ada* regulon in stationary phase: evidence for *rpoS*-dependent negative regulation of *alkA* transcription. *J Bacteriol*, **181**, 6836-6839.

**Lange, R., Fischer, D. and Hengge-Aronis, R.** (1995) Identification of transcriptional start sites and the role of ppGpp in the expression of *rpoS*, the structural gene for the sigma S subunit of RNA polymerase in *Escherichia coli*. *J Bacteriol*, **177**, 4676-4680.

**Lange, R. and Hengge-Aronis, R.** (1994) The *nlpD* gene is located in an operon with *rpoS* on the *Escherichia coli* chromosome and encodes a novel lipoprotein with a potential function in cell wall formation. *Mol Microbiol*, **13**, 733-743.

**Laurie, A.D., Bernardo, L.M., Sze, C.C., Skarfstad, E., Szalewska-Palasz, A., Nystrom, T. and Shingler, V.** (2003) The role of the alarmone (p)ppGpp in sigma N competition for core RNA polymerase. *J Biol Chem*, **278**, 1494-1503.

**Lee, S.J. and Gralla, J.D.** (2001) $\sigma^{38}$ (*rpoS*) RNA polymerase promoter engagement via -10 region nucleotides. *J Biol Chem*, **276**, 30064-30071.

**Li, C., Louise, C.J., Shi, W. and Adler, J.** (1993) Adverse conditions which cause lack of flagella in *Escherichia coli*. *J Bacteriol*, **175**, 2229-2235.

**Li, Z. and Demple, B.** (1996) Sequence specificity for DNA binding by *Escherichia coli* SoxS and Rob proteins. *Mol Microbiol*, **20**, 937-945.

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

**Liu, X. and De Wulf, P.** (2004) Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling. *J Biol Chem*, **279**, 12588-12597.

**Liu, Y., Gao, W., Wang, Y., Wu, L., Liu, X., Yan, T., Alm, E., Arkin, A., Thompson, D.K., Fields, M.W. and Zhou, J.** (2005) Transcriptome analysis of *Shewanella oneidensis* MR-1 in response to elevated salt conditions. *J Bacteriol*, **187**, 2501-2507.

**Lobell, R.B. and Schleif, R.F.** (1990) DNA looping and unlooping by AraC protein. *Science*, **250**, 528-532.

**Lonetto, M., Gribskov, M. and Gross, C.A.** (1992) The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol*, **174**, 3843-3849.

**Luttinger, A.** (1995) The twisted 'life' of DNA in the cell: bacterial topoisomerases. *Mol Microbiol*, **15**, 601-606.

**Maeda, H., Fujita, N. and Ishihama, A.** (2000) Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res*, **28**, 3497-3503.

**Magnusson, L.U., Farewell, A. and Nystrom, T.** (2005) ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol*, **13**, 236-242.

**Marqués, S., Gallegos, M.T., Manzanera, M., Holtel, A., Timmis, K.N. and Ramos, J.L.** (1998) Activation and repression of transcription at the double tandem divergent promoters for the *xylR* and *xylS* genes of the TOL plasmid of *Pseudomonas putida*. *J Bacteriol*, **180**, 2889-2894.

**Marqués, S., Gallegos, M.T. and Ramos, J.L.** (1995) Role of sigma S in transcription from the positively controlled Pm promoter of the TOL plasmid of *Pseudomonas putida*. *Mol Microbiol*, **18**, 851-857.

**Marqués, S., Manzanera, M., González-Pérez, M.M., Gallegos, M.T. and Ramos, J.L.** (1999) The XylS-dependent Pm promoter is transcribed *in vivo* by RNA polymerase with sigma 32 or sigma 38 depending on the growth phase. *Mol Microbiol*, **31**, 1105-1113.

**Martin, R.G., Gillette, W.K., Martin, N.I. and Rosner, J.L.** (2002) Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in *Escherichia coli*. *Mol Microbiol*, **43**, 355-370.

**Martin, R.G., Gillette, W.K., Rhee, S. and Rosner, J.L.** (1999) Structural requirements for marbox function in transcriptional activation of *mar/sox/rob* regulon promoters in *Escherichia coli*: sequence, orientation and spatial relationship to the core promoter. *Mol Microbiol*, **34**, 431-441.

**Martin, R.G., Gillette, W.K. and Rosner, J.L.** (2000) Promoter discrimination by the related transcriptional activators MarA and SoxS: differential regulation by differential binding. *Mol Microbiol*, **35**, 623-634.

**Martin, R.G., Jair, K.W., Wolf, R.E., Jr. and Rosner, J.L.** (1996) Autoactivation of the *marRAB* multiple antibiotic resistance operon by the MarA transcriptional activator in *Escherichia coli*. *J Bacteriol*, **178**, 2216-2223.

**Martin, R.G. and Rosner, J.L.** (2001) The AraC transcriptional activators. *Curr Opin Microbiol*, **4**, 132-137.

**Mathew, R. and Chatterji, D.** (2006) The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol*, **14**, 450-455.

**Mathew, R., Ramakanth, M. and Chatterji, D.** (2005) Deletion of the gene *rpoZ*, encoding the omega subunit of RNA polymerase, in *Mycobacterium smegmatis* results in fragmentation of the β' subunit in the enzyme assembly. *J Bacteriol*, **187**, 6565-6570.

**Maurer, L.M., Yohannes, E., Bondurant, S.S., Radmacher, M. and Slonczewski, J.L.** (2005) pH regulates genes for flagellar motility, catabolism, and oxidative stress in *Escherichia coli* K-12. *J Bacteriol*, **187**, 304-319.

**McLean, B.W., Wiseman, S.L. and Kropinski, A.M.** (1997) Functional analysis of $\sigma^{70}$ consensus promoters in *Pseudomonas aeruginosa* and *Escherichia coli*. *Can J Microbiol*, **43**, 981-985.

**McLeod, S.M. and Johnson, R.C.** (2001) Control of transcription by nucleoid proteins. *Curr Opin Microbiol*, **4**, 152-159.

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

**Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R. and Ebright, R.H.** (2002) Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, **108**, 599-614.

**Michán, C., Manchado, M., Dorado, G. and Pueyo, C.** (1999) *In vivo* transcription of the *Escherichia coli oxyR* regulon as a function of growth phase and in response to oxidative stress. *J Bacteriol*, **181**, 2759-2764.

**Michán, C., Zhou, L., Gallegos, M.T., Timmis, K.N. and Ramos, J.L.** (1992) Identification of critical amino-terminal regions of XylS. The positive regulator encoded by the TOL plasmid. *J Biol Chem*, **267**, 22897-22901.

**Mika, F. and Hengge, R.** (2005) A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of $\sigma^{S}$ (RpoS) in *E. coli*. *Genes Dev*, **19**, 2770-2781.

**Missiakas, D., Mayer, M.P., Lemaire, M., Georgopoulos, C. and Raina, S.** (1997) Modulation of the *Escherichia coli* $\sigma^{E}$ (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. *Mol Microbiol*, **24**, 355-371.

**Mooney, R.A., Darst, S.A. and Landick, R.** (2005) Sigma and RNA polymerase: an on-again, off-again relationship? *Mol Cell*, **20**, 335-345.

**Mooney, R.A. and Landick, R.** (2003) Tethering $\sigma^{70}$ to RNA polymerase reveals high *in vivo* activity of sigma factors and $\sigma^{70}$-dependent pausing at promoter-distal locations. *Genes Dev*, **17**, 2839-2851.

**Morita, M.T., Kanemori, M., Yanagi, H. and Yura, T.** (2000) Dynamic interplay between antagonistic pathways controlling the sigma 32 level in *Escherichia coli*. *Proc Natl Acad Sci USA*, **97**, 5860-5865.

**Mosqueda, G. and Ramos, J.L.** (2000) A set of genes encoding a second toluene efflux system in *Pseudomonas putida* DOT-T1E is linked to the tod genes for toluene metabolism. *J Bacteriol*, **182**, 937-943.

**Mukhopadhyay, J., Kapanidis, A.N., Mekler, V., Kortkhonjia, E., Ebright, Y.W. and Ebright, R.H.** (2001) Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell*, **106**, 453-463.

**Müller-Hill, B.** (1996) The Lac Operon. de Gruyter, Berlin.

**Müller-Hill, B.** (1998) Some repressors of bacterial transcription. *Curr Opin Microbiol*, **1**, 145-151.

**Murakami, K.S. and Darst, S.A.** (2003) Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol*, **13**, 31-39.

**Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. and Darst, S.A.** (2002a) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, **296**, 1285-1290.

**Murakami, K.S., Masuda, S. and Darst, S.A.** (2002b) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science*, **296**, 1280-1284.

**Muskhelishvili, G., Buckle, M., Heumann, H., Kahmann, R. and Travers, A.A.** (1997) FIS activates sequential steps during transcription initiation at a stable RNA promoter. *EMBO J*, **16**, 3655-3665.

**Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A., Fouts, D.E., Gill, S.R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R.T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J.A., Timmis, K.N., Dusterhoft, A., Tummler, B. and Fraser, C.M.** (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol*, **4**, 799-808.

**Nickels, B.E., Mukhopadhyay, J., Garrity, S.J., Ebright, R.H. and Hochschild, A.** (2004) The sigma 70 subunit of RNA polymerase mediates a promoter-proximal pause at the *lac* promoter. *Nat Struct Mol Biol*, **11**, 544-550.

**Nonaka, G., Blankschien, M., Herman, C., Gross, C.A. and Rhodius, V.A.** (2006) Regulon and promoter analysis of the *E. coli* heat-shock factor, $\sigma^{32}$, reveals a multifaceted cellular response to heat stress. *Genes Dev*, **20**, 1776-1789.

**Nudler, E. and Gottesman, M.E.** (2002) Transcription termination and anti-termination in *E. coli. Genes Cells*, **7**, 755-768.

**Nystrom, T.** (2004) Growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition? *Mol Microbiol*, **54**, 855-862.

**Ohlsen, K.L. and Gralla, J.D.** (1992) DNA melting within stable closed complexes at the *Escherichia coli rrnB* $P_1$ promoter. *J Biol Chem*, **267**, 19813-19818.

**Overton, T.W., Griffiths, L., Patel, M.D., Hobman, J.L., Penn, C.W., Cole, J.A. and Constantinidou, C.** (2006) Microarray analysis of gene regulation by oxygen, nitrate, nitrite, FNR, NarL and NarP during anaerobic growth of *Escherichia coli*: new insights into microbial physiology. *Biochem Soc Trans*, **34**, 104-107.

**Parales, R.E. and Harwood, C.S.** (2002) Bacterial chemotaxis to pollutants and plant-derived aromatic molecules. *Curr Opin Microbiol*, **5**, 266-273.

**Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W. and Gourse, R.L.** (2004) DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell*, **118**, 311-322.

**Pérez-Rueda, E. and Collado-Vides, J.** (2000) The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. *Nucleic Acids Res*, **28**, 1838-1847.

**Peter, B.J., Arsuaga, J., Breier, A.M., Khodursky, A.B., Brown, P.O. and Cozzarelli, N.R.** (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol*, **5**, R87.

**Raffaelle, M., Kanin, E.I., Vogt, J., Burgess, R.R. and Ansari, A.Z.** (2005) Holoenzyme switching and stochastic release of sigma factors from RNA polymerase *in vivo*. *Mol Cell*, **20**, 357-366.

**Ramos, J.L., Duque, E., Gallegos, M.T., Godoy, P., Ramos-González, M.I., Rojas, A., Terán, W. and Segura, A.** (2002) Mechanisms of solvent tolerance in gram-negative bacteria. *Annu Rev Microbiol*, **56**, 743-768.

**Ramos, J.L., Duque, E., Godoy, P. and Segura, A.** (1998) Efflux pumps involved in toluene tolerance in *Pseudomonas putida* DOT-T1E. *J Bacteriol*, **180**, 3323-3329.

**Ramos, J.L., Duque, E., Rodríguez-Herva, J.J., Godoy, P., Haidour, A., Reyes, F. and Fernández-Barrero, A.** (1997a) Mechanisms for solvent tolerance in bacteria. *J Biol Chem*, **272**, 3887-3890.

**Ramos, J.L., Gallegos, M.T., Marqués, S., Ramos-González, M.I., Espinosa-Urgel, M. and Segura, A.** (2001) Responses of Gram-negative bacteria to certain environmental stressors. *Curr Opin Microbiol*, **4**, 166-171.

**Ramos, J.L., Marqués, S. and Timmis, K.N.** (1997b) Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. *Annu Rev Microbiol*, **51**, 341-373.

**Ramos, J.L., Michán, C., Rojo, F., Dwyer, D. and Timmis, K.** (1990) Signal-regulator interactions. Genetic analysis of the effector binding site of XylS, the benzoate-activated positive regulator of *Pseudomonas* TOL plasmid *meta*-cleavage pathway operon. *J Mol Biol*, **211**, 373-382.

**Reynolds, R., Bermudez-Cruz, R.M. and Chamberlin, M.J.** (1992) Parameters affecting transcription termination by *Escherichia coli* RNA polymerase. I. Analysis of 13 rho-independent terminators. *J Mol Biol*, **224**, 31-51.

**Rhee, S., Martin, R.G., Rosner, J.L. and Davies, D.R.** (1998) A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc Natl Acad Sci USA*, **95**, 10413-10418.

**Rhodius, V.A. and Busby, S.J.** (1998) Positive activation of gene expression. *Curr Opin Microbiol*, **1**, 152-159.

**Richardson, J.P.** (1990) Rho-dependent transcription termination. *Biochim Biophys Acta*, **1048**, 127-138.

**Richardson, J.P.** (2002) Rho-dependent termination and ATPases in transcript termination. *Biochim Biophys Acta*, **1577**, 251-260.

**Richardson, J.P.** (2003) Loading Rho to terminate transcription. *Cell*, **114**, 157-159.

**Richardson, J.P., and J. Greenblatt.** (1996) Control of RNA chain elongation and termination, p.822-844. In Neidhardt, F.C., R. Curtiss, J. L. Ingraham., E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington DC.

**Richardson, L.V. and Richardson, J.P.** (1996) Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (rut) sequences of a terminator. *J Biol Chem*, **271**, 21597-21603.

**Richet, E., Vidal-Ingigliardi, D. and Raibaud, O.** (1991) A new mechanism for coactivation of transcription initiation: repositioning of an activator triggered by the binding of a second activator. *Cell*, **66**, 1185-1195.

**Ring, B.Z., Yarnell, W.S. and Roberts, J.W.** (1996) Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell*, **86**, 485-493.

**Roberts, J. and Park, J.S.** (2004) Mfd, the bacterial transcription repair coupling factor: translocation, repair and termination. *Curr Opin Microbiol*, **7**, 120-125.

**Rochman, M., Aviv, M., Glaser, G. and Muskhelishvili, G.** (2002) Promoter protection by a transcription factor acting as a local topological homeostat. *EMBO Rep*, **3**, 355-360.

**Rojas, A., Duque, E., Mosqueda, G., Golden, G., Hurtado, A., Ramos, J.L. and Segura, A.** (2001) Three efflux pumps are required to provide efficient tolerance to toluene in *Pseudomonas putida* DOT-T1E. *J Bacteriol*, **183**, 3967-3973.

**Ross, W., Ernst, A. and Gourse, R.L.** (2001) Fine structure of *E. coli* RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev*, **15**, 491-506.

**Ruberg, S., Tian, Z.X., Krol, E., Linke, B., Meyer, F., Wang, Y., Puhler, A., Weidner, S. and Becker, A.** (2003) Construction and validation of a *Sinorhizobium meliloti* whole genome DNA microarray: genome-wide profiling of osmoadaptive gene expression. *J Biotechnol*, **106**, 255-268.

**Ruíz, R., Marqués, S. and Ramos, J.L.** (2003) Leucines 193 and 194 at the N-terminal domain of the XylS protein, the positive transcriptional regulator of the TOL *meta*-cleavage pathway, are involved in dimerization. *J Bacteriol*, **185**, 3036-3041.

**Ruíz, R. and Ramos, J.L.** (2001) Residues 137 and 153 of XylS influence contacts with the C-terminal domain of the RNA polymerase alpha subunit. *Biochem Biophys Res Commun*, **287**, 519-521.

**Ruíz, R. and Ramos, J.L.** (2002) Residues 137 and 153 at the N terminus of the XylS protein influence the effector profile of this transcriptional regulator and the sigma factor used by RNA polymerase to stimulate transcription from its cognate promoter. *J Biol Chem*, **277**, 7282-7286.

**Ruíz, R., Ramos, J.L. and Egan, S.M.** (2001) Interactions of the XylS regulators with the C-terminal domain of the RNA polymerase alpha subunit influence the expression level from the cognate Pm promoter. *FEBS Lett*, **491**, 207-211.

**Sanderson, A., Mitchell, J.E., Minchin, S.D. and Busby, S.J.** (2003) Substitutions in the *Escherichia coli* RNA polymerase $\sigma^{70}$ factor that affect recognition of extended -10 elements at promoters. *FEBS Lett*, **544**, 199-205.

**Santos, P.M., Benndorf, D. and Sa-Correia, I.** (2004) Insights into *Pseudomonas putida* KT2440 response to phenol-induced stress by quantitative proteomics. *Proteomics*, **4**, 2640-2652.

**Saviola, B., Seabold, R. and Schleif, R.F.** (1998) Arm-domain interactions in AraC. *J Mol Biol*, **278**, 539-548.

**Schleif, R.** (1992) DNA looping. *Annu Rev Biochem*, **61**, 199-223.

**Schmitt, C.K., Darnell, S.C. and O'Brien, A.D.** (1996) The attenuated phenotype of a *Salmonella typhimurium flgM* mutant is related to expression of FliC flagellin. *J Bacteriol*, **178**, 2911-2915.

**Schneider, R., Travers, A. and Muskhelishvili, G.** (2000) The expression of the *Escherichia coli fis* gene is strongly dependent on the superhelical density of DNA. *Mol Microbiol*, **38**, 167-175.

**Segura, A., Godoy, P., van Dillewijn, P., Hurtado, A., Arroyo, N., Santacruz, S. and Ramos, J.L.** (2005) Proteomic analysis reveals the participation of energy- and stress-related proteins in the response of *Pseudomonas putida* DOT-T1E to toluene. *J Bacteriol*, **187**, 5937-5945.

**Shin, M., Kang, S., Hyun, S.J., Fujita, N., Ishihama, A., Valentin-Hansen, P. and Choy, H.E.** (2001) Repression of *deoP2* in *Escherichia coli* by CytR: conversion of a transcription activator into a repressor. *EMBO J*, **20**, 5392-5399.

**Shingler, V.** (2003) Integrated regulation in response to aromatic compounds: from signal sensing to attractive behaviour. *Environ Microbiol*, **5**, 1226-1241.

**Sikkema, J., de Bont, J.A. and Poolman, B.** (1994) Interactions of cyclic hydrocarbons with biological membranes. *J Biol Chem*, **269**, 8022-8028.

**Sikkema, J., de Bont, J.A. and Poolman, B.** (1995) Mechanisms of membrane toxicity of hydrocarbons. *Microbiol Rev*, **59**, 201-222.

**Sikkema, J., Poolman, B., Konings, W.N. and de Bont, J.A.** (1992) Effects of the membrane action of tetralin on the functional and structural properties of artificial and bacterial membranes. *J Bacteriol*, **174**, 2986-2992.

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

**Sosunov, V., Zorov, S., Sosunova, E., Nikolaev, A., Zakeyeva, I., Bass, I., Goldfarb, A., Nikiforov, V., Severinov, K. and Mustaev, A.** (2005) The involvement of the aspartate triad of the active center in all catalytic activities of multisubunit RNA polymerase. *Nucleic Acids Res*, **33**, 4202-4211.

**Studemann, A., Noirclerc-Savoye, M., Klauck, E., Becker, G., Schneider, D. and Hengge, R.** (2003) Sequential recognition of two distinct sites in $\sigma^S$ by the proteolytic targeting factor RssB and ClpX. *EMBO J*, **22**, 4111-4120.

**Studholme, D.J. and Buck, M.** (2000) The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS Microbiol Lett*, **186**, 1-9.

**Sze, C.C. and Shingler, V.** (1999) The alarmone (p)ppGpp mediates physiological-responsive control at the sigma 54-dependent Po promoter. *Mol Microbiol*, **31**, 1217-1228.

**Tatsuta, T., Joob, D.M., Calendar, R., Akiyama, Y. and Ogura, T.** (2000) Evidence for an active role of the DnaK chaperone system in the degradation of sigma(32). *FEBS Lett*, **478**, 271-275.

**Tebbutt, J., Rhodius, V.A., Webster, C.L. and Busby, S.J.** (2002) Architectural requirements for optimal activation by tandem CRP molecules at a class I CRP-dependent promoter. *FEMS Microbiol Lett*, **210**, 55-60.

**Timmes, A., Rodgers, M. and Schleif, R.** (2004) Biochemical and physiological properties of the DNA binding domain of AraC protein. *J Mol Biol*, **340**, 731-738.

**Tobes, R. and Ramos, J.L.** (2002) AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucleic Acids Res*, **30**, 318-321.

**Tomsic, M., Tsujikawa, L., Panaghie, G., Wang, Y., Azok, J. and deHaseth, P.L.** (2001) Different roles for basic and aromatic amino acids in conserved region 2 of *Escherichia coli* $\sigma^{70}$ in the nucleation and maintenance of the single-stranded DNA bubble in open RNA polymerase-promoter complexes. *J Biol Chem*, **276**, 31891-31896.

**Toulokhonov, II, Shulgina, I. and Hernandez, V.J.** (2001) Binding of the transcription effector ppGpp to *Escherichia coli* RNA polymerase is allosteric, modular, and occurs near the N terminus of the β'-subunit. *J Biol Chem*, **276**, 1220-1225.

**Travers, A. and Muskhelishvili, G.** (2005) DNA supercoiling - a global transcriptional regulator for enterobacterial growth? *Nat Rev Microbiol*, **3**, 157-169.

**Tsujikawa, L., Tsodikov, O.V. and deHaseth, P.L.** (2002) Interaction of RNA polymerase with forked DNA: evidence for two kinetically significant intermediates on the pathway to the final complex. *Proc Natl Acad Sci USA*, **99**, 3493-3498.

**Typas, A., Barembruch, C., Possling, A. and Hengge, R.** (2007) Stationary phase reorganisation of the *Escherichia coli* transcription machinery by Crl protein, a fine-tuner of sigma(s) activity and levels. *EMBO J*, **26**, 1569-1578.

**Typas, A. and Hengge, R.** (2005) Differential ability of sigma(s) and $\sigma^{70}$ of *Escherichia coli* to utilize promoters containing half or full UP-element sites. *Mol Microbiol*, **55**, 250-260.

**Typas, A. and Hengge, R.** (2006) Role of the spacer between the -35 and -10 regions in sigmas promoter selectivity in *Escherichia coli*. *Mol Microbiol*, **59**, 1037-1051.

**Vardar, G., Barbieri, P. and Wood, T.K.** (2005) Chemotaxis of *Pseudomonas stutzeri* OX1 and *Burkholderia cepacia* G4 toward chlorinated ethenes. *Appl Microbiol Biotechnol*, **66**, 696-701.

**Vassylyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S. and Yokoyama, S.** (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*, **417**, 712-719.

**Velázquez, F., di Bartolo, I. and de Lorenzo, V.** (2004) Genetic evidence that catabolites of the Entner-Doudoroff pathway signal C-source repression of the $\sigma^{54}$ Pu promoter of *Pseudomonas putida*. *J Bacteriol*, **186**, 8267-8275.

**Velázquez, F., Parro, V. and de Lorenzo, V.** (2005) Inferring the genetic network of *m*-xylene metabolism through expression profiling of the *xyl* genes of *Pseudomonas putida* mt-2. *Mol Microbiol*, **57**, 1557-1569.

**von Hippel, P.H.** (1998) An integrated model of the transcription complex in elongation, termination, and editing. *Science*, **281**, 660-665.

**Vrentas, C.E., Gaal, T., Ross, W., Ebright, R.H. and Gourse, R.L.** (2005) Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA. *Genes Dev*, **19**, 2378-2387.

**Wade, J.T., Belyaeva, T.A., Hyde, E.I. and Busby, S.J.** (2000) Repression of the *Escherichia coli melR* promoter by MelR: evidence that efficient repression requires the formation of a repression loop. *Mol Microbiol*, **36**, 223-229.

**Wade, J.T., Belyaeva, T.A., Hyde, E.I. and Busby, S.J.** (2001) A simple mechanism for co-dependence on two activators at an *Escherichia coli* promoter. *EMBO J*, **20**, 7160-7167.

**Wade, J.T., Roa, D.C., Grainger, D.C., Hurd, D., Busby, S.J., Struhl, K. and Nudler, E.** (2006) Extensive functional overlap between sigma factors in *Escherichia coli*. *Nat Struct Mol Biol*, **13**, 806-814.

**Watson, J.D. and Crick, F.H.** (1953) The structure of DNA. *Cold Spring Harb Symp Quant Biol*, **18**, 123-131.

**Weber, H., Polen, T., Heuveling, J., Wendisch, V.F. and Hengge, R.** (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: $\sigma^{S}$-dependent genes, promoters, and sigma factor selectivity. *J Bacteriol*, **187**, 1591-1603.

**Weldon, J.E., Rodgers, M.E., Larkin, C. and Schleif, R.F.** (2007) Structure and properties of a truely apo form of AraC dimerization domain. *Proteins*, **66**, 646-654.

**Westblade, L.F., Ilag, L.L., Powell, A.K., Kolb, A., Robinson, C.V. and Busby, S.J.** (2004) Studies of the *Escherichia coli* Rsd- $\sigma^{70}$ complex. *J Mol Biol*, **335**, 685-692.

**Whiteley, M., Parsek, M.R. and Greenberg, E.P.** (2000) Regulation of quorum sensing by RpoS in *Pseudomonas aeruginosa*. *J Bacteriol*, **182**, 4356-4360.

**Wickstrum, J.R. and Egan, S.M.** (2004) Amino acid contacts between sigma 70 domain 4 and the transcription activators RhaS and RhaR. *J Bacteriol*, **186**, 6277-6285.

**Wickstrum, J.R., Skredenske, J.M., Kolin, A., Jin, D.J., Fang, J. and Egan, S.M.** (2007) Transcription Activation by the DNA-Binding Domain of the AraC Family Protein RhaS in the Absence of its Effector-Binding Domain. *J Bacteriol*. In press.

**Wood, T.I., Griffith, K.L., Fawcett, W.P., Jair, K.W., Schneider, T.D. and Wolf, R.E., Jr.** (1999) Interdependence of the position and orientation of SoxS binding sites in the transcriptional activation of the class I subset of *Escherichia coli* superoxide-inducible promoters. *Mol Microbiol*, **34**, 414-430.

**Wu, M. and Schleif, R.** (2001) Mapping arm-DNA-binding domain interactions in AraC. *J Mol Biol*, **307**, 1001-1009.

**Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K. and Darst, S.A.** (1999) Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 Å resolution. *Cell*, **98**, 811-824.

**Zheng, D., Constantinidou, C., Hobman, J.L. and Minchin, S.D.** (2004) Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res*, **32**, 5874-5893.

**Zhou, Y. and Gottesman, S.** (2006) Modes of regulation of RpoS by H-NS. *J Bacteriol*, **188**, 7022-7025.