

1. INTRODUCCIÓN

Una de las características de la célula bacteriana es la capacidad de adaptarse rápidamente a los frecuentes cambios en las condiciones ambientales del entorno en el que se desarrolla. La capacidad de adaptarse al entorno para obtener la mayor energía posible resulta tan imprescindible como el ahorro de la misma. Por este motivo, en el mundo bacteriano la regulación de la expresión génica constituye una herramienta fundamental para adaptarse a un hábitat cambiante. La maquinaria bioquímica se ajusta así en función de variaciones en las características físico-químicas del exterior, variaciones que las bacterias han aprendido a detectar.

Las bacterias detectan tanto parámetros químicos, relacionados con los aspectos nutricionales del medio (presencia o ausencia de iones, fuentes de energía, aceptores finales de energía, aceptores finales de electrones, etc.), como parámetros físicos (la temperatura, la presión hidrostática o la presión osmótica). Algunos de estos factores ambientales, sirven de señales para regular la expresión génica y adaptarse al nuevo ambiente con éxito.

Los estímulos ambientales son detectados directa o indirectamente dentro de la célula mediante componentes solubles del citoplasma. Ya sea porque el estímulo sea capaz de penetrar dentro de la célula (por ejemplo, un substrato del catabolismo) o bien por la detección del efecto del estímulo sobre componentes celulares o en su metabolismo (por ejemplo, una elevada temperatura puede ser detectada a través de su efecto en el plegamiento de determinadas proteínas). Algunos estímulos, entre los que se encuentran el pH y la osmolaridad, pueden ser detectados a través de receptores transmembrana específicos que actúan como sensores de la señal, a nivel de la superficie celular (Ninfa *et al.*, 1996).

Los mecanismos de reconocimiento de estos estímulos y la transmisión de esta información pueden ser más o menos complejos dependiendo del sistema particular y de si la respuesta que conlleva es específica o global.

1.1 REGULACIÓN DE LA EXPRESIÓN GÉNICA A NIVEL TRANSCRIPCIONAL

La necesidad de regulación de la expresión génica en bacterias implica la existencia de mecanismos moleculares, tanto o más complejos que los de la célula eucariota, que pueden afectar a las distintas etapas del proceso: inicio, elongación o terminación de la transcripción, procesamiento y estabilidad del ARN mensajero, a nivel de la traducción y, finalmente, a nivel post-traducciona mediante la modificación de la proteína (Rojo, 1999). La mayor parte de mecanismos descritos en bacterias que regulan la expresión génica en función de factores ambientales ejercen un control en el proceso de transcripción, en particular en el inicio de la transcripción. La razón es sencilla: interviniendo a nivel transcripcional, el ahorro energético es muy superior a cuando se actúa a nivel de la degradación o inactivación de una proteína ya sintetizada. La transcripción puede ser activada o reprimida en su inicio por mecanismos que implican modificaciones en la ARN polimerasa, la acción de proteínas reguladoras de unión al ADN, o cambios en el estado de superenrollamiento del ADN.

1.1.1 FACTORES SIGMA

El holoenzima ARN polimerasa (RNAP) está formado por dos componentes: el núcleo enzimático y la subunidad σ correspondiente. El núcleo enzimático, formado por las subunidades $\alpha_2\beta\beta\omega$ contiene toda la maquinaria catalítica necesaria para la síntesis del ARN. No obstante, sólo el holoenzima puede iniciar la transcripción en un promotor determinado. La subunidad σ tiene tres funciones principales: asegurar el reconocimiento de una secuencia promotora específica, posicionar el holoenzima en un promotor diana, y facilitar la apertura de la doble hélice de ADN cerca del inicio de transcripción.

Normalmente, la mayoría de inicios de la transcripción en células en crecimiento exponencial son catalizados por el holoenzima RNAP ($E\sigma$), que está formado por un factor σ “housekeeping” similar al factor σ^{70} de *Escherichia coli*. Factores sigma alternativos controlan regulones específicos que se activan en determinadas condiciones de estrés, de crecimiento y durante cambios morfológicos. Además, los factores sigma interactúan

con los factores de transcripción, ya sean activadores o represores (Ebright y Busby, 1995; Hochschild y Dove, 1998; Dove *et al.*, 2003).

Con la importante excepción de la pequeña familia de factores σ^{54} , todos los factores sigma comparten características comunes. Son proteínas formadas por hasta cuatro dominios diferentes unidos por conectores. Los dominios 2, 3 y 4 están implicados en el reconocimiento del promotor, pero la función del dominio 1 es aún desconocida. Este dominio está incluso ausente en muchos factores σ (Gruber y Gross, 2003; Browning y Busby, 2004).

Escherichia coli contiene un factor σ^{54} y seis factores σ de la familia de σ^{70} , σ^{70} (RpoD), σ^s (RpoS), σ^{32} (RpoH), σ^F (FliA), σ^E (RpoE), y FceI (Gruber y Gross, 2003). Cada uno de ellos dirige la transcripción de grupos de genes específicos. En *E. coli* σ^s transcribe más de 100 genes implicados en supervivencia celular, protección contra diversas situaciones de estrés (estrés oxidativo, radiación UV, choque térmico, hiperosmolaridad, pH ácido) y expresión de la virulencia (Ishihama, 2000). Los niveles celulares de σ^s están cuidadosamente controlados, aumentando significativamente en la entrada en la fase estacionaria de crecimiento, momento en el que alcanza el 30% del nivel del factor “housekeeping” σ^{70} , aumentando así su habilidad de competir por el núcleo enzimático con otros factores σ disponibles en la célula (Jishage *et al.*, 1996). Los estudios sobre los niveles de σ^s han revelado uno de los sistemas más complejos de regulación en bacterias, ya que tiene lugar a nivel de transcripción, de traducción y de estabilidad proteica, todo coordinado con la respuesta a diversas situaciones de estrés celular.

Aunque $E\sigma^{70}$ y $E\sigma^s$ reconocen básicamente las mismas secuencias promotoras, determinadas por las denominadas cajas -10 y -35, algunas características determinan la selectividad de $E\sigma^s$. A nivel de secuencia son: utilización total o parcial de una secuencia UP situada aguas arriba de la caja -35 (las secuencias UP son un tercer elemento de reconocimiento en determinadas regiones promotoras localizado aguas arriba de la caja -35 y que interacciona con el dominio C-terminal de la subunidad α de la RNAP); una citosina en la posición -13 de la caja -10; una zona rica en A-T aguas abajo de la caja -10; una caja -35 degenerada; y, finalmente, un cambio en la distancia entre las cajas -10 y -35. Finalmente, la acción sobre una región promotora de determinados factores, como H-NS,

Lrp o IHF o la topología del ADN en esta misma región, juegan en ocasiones un papel decisivo en la especificidad de $E\sigma^S$ (Hengge-Aronis, 2002).

Como se ha mencionado anteriormente en *E. coli*, la mayoría de los genes expresados durante la fase exponencial de crecimiento se transcriben por $E\sigma^{70}$, mientras que $E\sigma^S$ es esencial en la transcripción de varios genes específicos de fase estacionaria. Algunos genes de respuesta al estrés se transcriben con el holoenzima formado por otras subunidades σ minoritarias. $E\sigma^{54}$ transcribe genes activados por una deficiencia de nitrógeno y otros genes de respuesta al estrés. $E\sigma^{32}$ transcribe genes que codifican para proteínas del choque térmico, y se induce por un número excesivo de proteínas citoplasmáticas que no se pliegan correctamente, resultado de un choque térmico y otros estreses. $E\sigma^F$ activa los genes tardíos implicados en el ensamblaje del flagelo. $E\sigma^E$ se induce por proteínas desplegadas de la envoltura celular y dirige la expresión de los genes para restaurar la integridad de las envueltas. Y, para finalizar, FceI, responsable de la transcripción de la maquinaria necesaria para el transporte por translocación de citrato férrico ante una situación de limitación de hierro y presencia de citrato férrico en el ambiente. FceI forma parte de la subfamilia de factores sigma extracitoplasmáticos (EFC) o sigmas del grupo 4, pertenecientes a la familia de σ^{70} pero caracterizadas por tener una secuencia bastante divergente de otros grupos de la familia σ^{70} . Originalmente se les denominó factores sigma extracitoplasmáticos porque muchos miembros de este grupo identificados inicialmente regulaban aspectos relacionados con la superficie celular o el transporte, y se cotranscribían muy a menudo con un factor anti-sigma transmembrana. Sin embargo, se está demostrando que proteínas de función más amplia pertenecen a este subgrupo. Muchas bacterias, sobre todo las de genomas más complejos, contienen múltiples factores σ EFC que a menudo superan el número de otros tipos de factores sigma. Varios ejemplos son *Bacillus subtilis* (7 factores sigma EFC), *Mycobacterium tuberculosis* (10), *Caulobacter crescentus* (13), *Pseudomonas aeruginosa* (aproximadamente 19), y *Streptomyces coelicolor* (aproximadamente 50) (Helmann, 2002).

En muchos casos, la actividad de un factor σ está controlada por un factor anti-sigma, que lo mantiene secuestrado alejándolo de la RNAP (Hughes y Mathee, 1998). Por ejemplo FlgM está unido a σ^F en la fase temprana de formación del flagelo, pero se secreta fuera de la célula después de la formación del cuerpo basal del flagelo, permitiendo la transcripción de los genes de expresión tardía por $E\sigma^F$. El producto del gen *rseA*,

localizado a continuación del gen *rpoE* en el mismo operón, regula negativamente a σ^E , ya que el extremo amino de RseA interacciona directamente con este factor σ . Una característica común para los EFC es que se cotranscriben con uno o más reguladores negativos, a menudo una proteína transmembrana (Helmann, 2002). Otro ejemplo de factor anti-sigma es Rsd (regulador de sigma D), específico de σ^{70} . Rsd no está presente en células en crecimiento exponencial, pero al entrar en fase estacionaria se empieza a sintetizar llegando a sus niveles máximos en fase estacionaria temprana.

En resumen, las células bacterianas tienen un factor sigma “housekeeping” y un número variable de factores σ que poseen diferentes propiedades de reconocimiento de secuencias promotoras. La célula presenta un repertorio de factores sigma que permiten alterar su patrón de expresión génica en respuesta a distintos estímulos, entre ellos los que implican una situación de estrés celular.

1.1.2 FACTORES DE TRANSCRIPCIÓN

Se considera que el genoma de *E. coli* contiene más de 300 genes, que en base al análisis de la secuencia de aminoácidos del producto codificado, codifican para proteínas que, mediante su unión a la región promotora de otros genes son capaces de regular su transcripción (Pérez-Rueda y Collado-Vides, 2000; Babu y Teichmann, 2003). Solamente en la mitad de los casos la función de estas proteínas, denominadas factores transcripcionales, ha sido verificada experimentalmente. La mayoría de los factores de transcripción, aproximadamente tres cuartas partes, son proteínas formadas estructuralmente por dos dominios: un dominio de unión al ADN más un dominio regulador (Babu y Teichmann, 2003). El dominio de unión al ADN predominante es el dominio HTH (“helix-turn-helix”), presente en 248 de los reguladores transcripcionales conocidos y predichos (Pérez-Rueda y Collado-Vides, 2000). Los factores de transcripción se pueden agrupar en base al análisis de su secuencia. De este modo se han identificado toda una serie de familias (Pérez-Rueda y Collado-Vides, 2000) mostradas en la tabla 1, de entre ellas siendo LacI, AraC, LysR, CRP y OmpR las mejor caracterizadas.

Tabla 1. Familias de proteínas reguladoras que actúan como factores transcripcionales en *E. coli*.

Familia	Factores Conocidos	Factores Predichos
CRP	2	0
AsnC	2	1
IclR	2	6
Cold	8	1
TetR/AcrR	4	5
DeoR	7	7
EBP	9	5
GalR/LacI	12	2
OmpR	13	4
LuxR/UhpA	11	6
GntR	8	12
AraC/XylS	14	13
LysR	18	27
Otras	49	66

La primera columna describe el nombre de la familia, la segunda columna el número de reguladores transcripcionales conocidos y la tercera el número de proteínas predichas como reguladores transcripcionales.

Los factores de transcripción acoplan la expresión de los genes a las señales ambientales, y ellos mismos deben ser regulados, bien controlando su expresión, bien su actividad. Para a ello se han descrito diversos mecanismos. En primer lugar, la afinidad de unión de estos reguladores transcripcionales al ADN puede ser modulada mediante su unión a moléculas, en concreto ligandos de bajo peso molecular, cuya concentración cambia en respuesta a la señal ambiental, como la disponibilidad de un nutriente o una situación de estrés. Teniendo en cuenta que en el 44% de los factores de transcripción el dominio regulador es un dominio de unión a moléculas de bajo peso molecular, se podría deducir que la mitad de los factores transcripcionales están directamente regulados por este

mecanismo (Anantharaman *et al.*, 2001; Babu y Teichmann, 2003). En segundo lugar, la actividad de algunos factores transcripcionales está modulada por modificación covalente, como ocurre en los sistemas de transducción de señal de dos componentes. Un 10% de los factores de transcripción tiene un dominio tipo Che-Y (dominio receptor de la fosforilación por parte de una histidina quinasa sensora) presente en los reguladores de respuesta, como por ejemplo NarL (implicado en la regulación de los genes de respiración de nitrato), que se unen a su ADN diana solamente cuando han sido fosforilados por su correspondiente sensor quinasa. Los sensores quinasa se localizan usualmente en la membrana citoplasmática y son regulados por señales extracelulares. NarL está controlado por las quinasas NarX y NarQ, que se activan mediante su unión a iones nitrito o nitrato del medio (Chiang *et al.*, 1992; Cavicchioli *et al.*, 1995). En tercer lugar, para algunos factores transcripcionales, es su concentración en la célula lo que controla su actividad. En estos casos su concentración está regulada a través de su expresión o por proteólisis. Por ejemplo, una respuesta celular al estrés oxidativo está controlada mediante la concentración de la proteína SoxS (Dempfle, 1996). La transcripción del gen que codifica para SoxS es controlada por SoxR, que a su vez está regulado directamente mediante interacciones con ligando oxidativos. Otro mecanismo menos común, es el secuestro del factor de transcripción mediante la unión a una proteína reguladora, lo que disminuye su concentración. Por ejemplo, la transcripción de MalT, activador transcripcional de los genes del regulón maltosa en *E. coli*, está inhibido por Mlc, regulador global que es secuestrado por el transportador PtsG (transportador de glucosa del sistema de la fosfotransferasa) cuando entra glucosa en la célula (Böhm y Boss, 2004). Finalmente, también hay que considerar a los cambios en la topología local del ADN en los que puedan estar implicadas proteínas asociadas al nucleoide, las cuales pueden variar la accesibilidad o la afinidad del factor de transcripción por su sitio de unión al ADN.

Cuando un factor de transcripción se une a un promotor puede activar o reprimir el inicio de transcripción. De todas las proteínas reguladoras descritas o predichas un 34% son activadores, un 43% son represores y un 22% son proteínas duales, es decir, que pueden actuar como activadores o represores según sus promotores diana (Pérez-Rueda y Collado-Vides, 2000). El 25% de estas proteínas reguladoras duales pertenecen a la familia LysR, ya que muchos de sus miembros son activadores de varios genes y represores de su propia expresión (Schell, 1993). Según la localización del dominio HTH podemos deducir la potencialidad de un factor de transcripción como activador o represor: en un 96% de las

proteínas represoras el dominio se encuentra en el extremo N-terminal, mientras que en el 78% de las proteínas activadoras se encuentra en el extremo C-terminal. Una de las características de los factores de transcripción es su autoregulación negativa (Thieffry *et al.*, 1998). Unos pocos se autorregulan positivamente como son PhoB, GutM, TdcA, CadC, RhaS y RhaR, mientras que todavía son menos numerosos los reguladores descritos que ejercen regulación positiva y negativa sobre su propia expresión (Ada, CRP y NtrC) (Magasanik y Neidhart, 1987).

1.1.3 SUPERENROLLAMIENTO DEL ADN

En las células procariotas, el genoma bacteriano se encuentra condensado en una estructura denominada nucleoide. Esta compleja estructura es capaz de conciliar la necesidad de empaquetar la gran cantidad de ADN con el requerimiento de la célula de replicarse y regular la expresión génica en respuesta a cambios rápidos de las condiciones nutricionales y ambientales.

A excepción del ADN plasmídico aislado de algunas arqueobacterias, por norma general las preparaciones de ADN obtenidas a partir de células procariotas presentan un déficit en el índice de enlace, que se manifiesta como un superenrollamiento negativo (Drlica, 1992). Este nivel de superenrollamiento negativo está estrictamente controlado por la influencia combinada de varios factores, que incluyen la transcripción, la replicación, las proteínas de unión al ADN y las actividades de las topoisomerasas.

El nivel de superenrollamiento mantenido por la acción coordinada de los enzimas ADN girasa y ADN topoisomerasa I es el denominado superenrollamiento no restringido. La acción de la ADN girasa provoca un incremento del superenrollamiento negativo, mediante una reacción que requiere la hidrólisis de ATP. La acción de la topoisomerasa I, en cambio, provoca una relajación del ADN, en una reacción independiente de ATP.

La acción de la ADN girasa, que requiere la hidrólisis de ATP, no responde a la concentración de ATP libre, sino a la relación $[ATP]/[ADP]$ intracelular indicativa de la carga energética de la célula. Muchos estudios demuestran que el nivel global de superenrollamiento en una célula está controlado por la carga energética. Una situación de

estrés altera tanto la carga energética como el nivel de superenrollamiento del ADN. Por ejemplo, un choque osmótico hace aumentar aproximadamente 4 veces la relación [ATP/ADP] intracelular, y al mismo tiempo el nivel de superenrollamiento global del cromosoma bacteriano: si el valor típico del superenrollamiento durante un crecimiento balanceado es de -0,05, en este caso aumenta a -0,09 (Higgins *et al.*, 1988). Durante transiciones de crecimiento de un ambiente aerobio a uno anaerobio, la relación [ATP/ADP] disminuye, y el nivel de superenrollamiento global cae momentáneamente de -0,05 a -0,038 (Cortassa y Aon, 1993). En una escala más amplia, mutaciones en las topoisomerasas cambian de tal modo el superenrollamiento que afectan los niveles de expresión de un número muy elevado de proteínas. De las 88 proteínas analizadas en un gel de dos dimensiones, el 39% de las mismas mostraron cambios en el nivel de expresión (Steck *et al.*, 1993).

Se ha propuesto que a nivel global, el grado de superenrollamiento del ADN sería el factor más importante afectando la regulación (Schaechter, 2001; Hatfield y Benham, 2002). Como el nivel de superenrollamiento está influido por la carga energética celular puede cambiar rápidamente en respuesta a un amplio abanico de condiciones ambientales y nutricionales. Ésta es una alteración global que afecta al cromosoma bacteriano al completo y por lo tanto a los niveles de expresión de todos los promotores sensibles al superenrollamiento. En el caso de los operones del regulón *ilv* en *E. coli*, que contienen los genes que codifican para los enzimas de la biosíntesis de aminoácidos de cadenas ramificadas (L-isoleucina, L-valina y L-leucina), se ha demostrado que sus niveles de expresión basal están ajustados al nivel de superenrollamiento del ADN (Hatfield y Benham, 2002), muy por encima del nivel específico de regulación del operón. Así sus niveles de expresión basal son bajos cuando el superenrollamiento también lo es, y altos cuando aumenta el superenrollamiento.

Como se ha mencionado anteriormente el nivel de superenrollamiento del cromosoma bacteriano está controlado además de por la acción de las topoisomerasas por proteínas de unión al ADN. Entre ellas cabe destacar las proteínas asociadas al nucleóide, responsables de compensar aproximadamente la mitad del superenrollamiento global del cromosoma bacteriano, el denominado superenrollamiento restringido. Se caracterizan por ser un grupo de proteínas de bajo peso molecular ricas en aminoácidos cargados, que a diferencia de las histonas eucariotas forman uniones inestables con el ADN. Estas

proteínas pueden afectar profundamente a la estructura local o global del cromosoma, pero aún no se conoce el papel fundamental de cada proteína en la condensación del genoma bacteriano. Además de esta función de organización del nucleoide estas proteínas influyen en la regulación de la expresión génica, así como en otros procesos en los que está implicado el ADN, como la replicación, recombinación, reparación o transposición.

En *E.coli* se han descrito 12 proteínas asociadas al nucleoide o NAP (Nucleoid Associated Proteins) (revisado en Ali Azam y Ishihama, 1999). Algunas de las más estudiadas son HU (Heat-Unstable Nucleoid Protein), IHF (Integration Host Factor), H-NS (Histone-like Nucleoid-Structuring protein), StpA (Supresor of *td*-mutant phenotype A), FIS (Factor for Inversion Stimulation), y Dps (DNA-binding protein from starved cells). En general puede considerarse que ninguna de estas proteínas es esencial para la viabilidad de la bacteria, aunque ciertas combinaciones de mutantes pueden ser muy perjudiciales para la célula y en algunos casos resultan letales (McLeod y Johnson, 2001).

Todas las NAP se unen al ADN de forma inespecífica, aunque algunas muestran preferencia por una secuencia o estructura del ADN particular. Por ejemplo H-NS se une preferentemente al ADN curvado (Dame *et al.*, 2001; Rimsky *et al.*, 2001).

La concentración de estas proteínas en el citoplasma varía dependiendo de las condiciones de crecimiento celular y la fase de crecimiento. Fis, HU, IHF, StpA, y H-NS son las proteínas más abundantes en fase exponencial. Sin embargo, durante la fase estacionaria las proteínas mayoritarias son Dps, IHF, y HU (Ali Azam *et al.*, 1999). Este cambio en la composición del conjunto de las proteínas asociadas al nucleoide sería el responsable de cambios globales en la estructura y actividad del nucleoide según la fase de crecimiento.

Las NAP actúan directa o indirectamente, y muy a menudo en combinación, para controlar la expresión de una gran variedad de genes que son fundamentales para la viabilidad de la célula, como los implicados en la síntesis proteica y el metabolismo. También intervienen en la regulación de otros genes implicados en respuesta a estímulos ambientales, como por ejemplo, los que responden a cambios en la osmolaridad o a la temperatura, así como también genes que afectan a la virulencia (McLeod y Johnson, 2001).

Las NAP mejor caracterizadas son HU, FIS, IHF y H-NS. De las mismas cabe resaltar algunas características:

- HU: la proteína HU es muy abundante (15.0000-30.000 dímeros por célula en fase exponencial). Es un heterodímero formado por dos subunidades intercambiables, HU- α y HU- β , de 9,2 y 9,5 kDa respectivamente, muy similares entre ellas (70%) en *Escherichia coli*, *Salmonella enterica* serovar Typhimurium, *Erwinia chrysanthemi* y *Shigella flexneri*. En otras especies se trata de un homodímero (Oberto y Rouvière-Yaniv, 1996). *E. coli* puede variar la composición de HU (α^2 o $\alpha\beta$ o β^2) a lo largo de la fase de crecimiento. Estos cambios en la composición de HU parecen tener consecuencias en su habilidad para organizar la cromatina bacteriana e influir en la expresión génica (revisado en Dorman *et al.*, 1999). Aunque hay muchas evidencias de la unión de HU a regiones promotoras, no hay detalles moleculares, excepto en el caso de la represión dependiente de GalR del promotor de *gal P2* en *E. coli*. En este caso, HU y el superenrollamiento del ADN colaboran para asistir a la unión de un tetrámero de GalR para formar un bucle de represión en el promotor (Kar y Adhya, 2001). Se ha propuesto que los papeles estructurales de las proteínas H-NS y HU son opuestos, y como consecuencia también lo serán sus contribuciones a la regulación de la expresión génica (Dame *et al.*, 2002).

- FIS: es una proteína básica formada por dos subunidades iguales codificadas por el gen *fis* y está altamente conservada en bacterias entéricas. Se trata de la proteína asociada al nucleoide más abundante en células de *E. coli* en fase de crecimiento exponencial (20.00-40.000 copias por célula). FIS regula la expresión de un número muy elevado de genes, tanto por control directo a nivel del inicio de transcripción, como por efecto indirecto (Finkel y Johnson, 1992; Xu y Johnson, 1995; González-Gil *et al.*, 1996; Schneider *et al.*, 1999). Actualmente se está considerando a FIS como un regulador importante de genes de virulencia de patógenos bacterianos: promueve la transcripción de los genes de invasión de *Shigella flexneri* (Falconi *et al.*, 2001), y contribuye en la regulación de los genes de virulencia en *S. Typhimurium* (Wilson *et al.*, 2001) y en cepas de *E. coli* enteropatógenas (Goldberg *et al.*, 2001).

- IHF: está implicada en la integración y en la escisión de la proteína integrasa λ tanto *in vivo* como *in vitro*. Es una proteína básica, formada por heterodímeros de las subunidades IHF α y IHF β codificadas por los genes *himA* y *himD* respectivamente, y abundante en fase estacionaria temprana. IHF afecta, directa o indirectamente, la transcripción de unos 100 genes de una amplia gama de funciones (Freundlinch *et al.*, 1992; Goosen y van de Putte, 1995; revisado en McLeod y Johnson, 2001). Entre ellos hay ejemplos de genes relacionados con la virulencia bacteriana en *Shigella flexneri* (Porter y Dorman, 1997) y *S. Typhimurium* (Marshall *et al.*, 1999).

- H-NS: interviene en el control de un número elevado de genes en bacterias Gram-negativas. Está considerado como un regulador global. Actúa usualmente reprimiendo la transcripción, y los escasos ejemplos en los que regula la expresión positivamente son efectos indirectos o postranscripcionales (Schröder y Wagner, 2002). En *E. coli*, hasta un 5% de los genes presentan una expresión dependiente de la presencia o ausencia de H-NS (Schröder y Wagner, 2002; Hommais *et al.*, 2001). Es interesante resaltar que muchos de los genes reprimidos por H-NS necesitan activadores para superar la represión (Schröder y Wagner, 2002). En algunos casos, las señales ambientales contribuyen a la inactivación de estos complejos de represión. Entre estos factores se encuentran la temperatura y la osmolaridad, dos factores que alteran la topología del ADN. También son señales del proceso de infección, y resulta interesante que muchos genes de virulencia bajo el control de H-NS responden a estas señales (Dorman y Deighan, 2003). En *S. flexneri*, el regulador más importante de la expresión de los genes de invasión, *virB*, está reprimido por H-NS y activado por Vir-F, un factor de transcripción de la familia AraC, en colaboración con un cambio en el superenrollamiento del ADN (Dorman *et al.*, 2001). En otros ejemplos H-NS actúa en colaboración con otra proteína asociada al nucleóide, la proteína Hha (Madrid *et al.*, 2002).

1.2 REDES REGULADORAS. MODULACIÓN GLOBAL

Muchas de las actividades bacterianas asociadas a la adaptación a los cambios ambientales implican la modificación de la expresión génica de diferentes genes u operones simultáneamente, de manera que requieren una coordinación a niveles de organización superiores al nivel de una unidad transcripcional individual. En la mayoría de los casos, sobre los sistemas de regulación específicos de determinados operones, se superponen sistemas reguladores globales que afectan y controlan la expresión de numerosos genes u operones implicados en diversas funciones celulares, constituyendo las denominadas redes reguladoras. La coordinación de estas redes es tan o más compleja que la regulación de los operones individuales (Neidhart y Savageau, 1996), y tiene como objetivo dar una respuesta global frente a un cambio ambiental.

1.2.1 LOS MODULADORES GLOBALES DENTRO DE LAS FAMILIAS DE PROTEÍNAS REGULADORAS

Típicamente los reguladores globales han sido reconocidos en base a su fenotipo pleiotrópico y su habilidad de regular operones implicados en diferentes funciones celulares (Gottesman, 1984). Recientemente varios autores han identificado en *E. coli* los factores de transcripción globales (tabla 2) siguiendo diferentes criterios. Babu y Teichmann (2003) basan su estudio en el análisis teórico de las secuencia proteicas y en la comparación con estructuras tridimensionales de proteínas ya descritas, que les permiten establecer la comparación de dominios de los factores de transcripción, y consideran a los factores de transcripción como reguladores globales cuando regulan más de 15 genes, ya sea por influencia directa o indirecta. Estos autores, hacen un estudio de las redes reguladoras y buscan dominios propios de red en los factores de transcripción, analizando los datos de interacciones transcripcionales directas entre los factores de transcripción y los operones que regulan. Shen-Orr *et al.* (2002), definen a los moduladores globales como factores de transcripción que regulan 10 o más operones. Martínez-Antonio y Collado-Vides (2003) han descrito toda una serie de características que permiten distinguir de entre los factores de transcripción aquellos que actúan como reguladores globales:

- **Regulan** directamente la expresión de **un número elevado de genes**: tan sólo siete proteínas reguladoras (CRP, FNR, IHF, FIS, ArcA, NarL y Lrp) pueden afectar el 51% de genes en *E. coli*.

- Actúan **co-regulando** con más de un factor de transcripción. La regulación por múltiples factores de transcripción es frecuente y en la mayoría de los casos un regulador global trabaja conjuntamente con más de un regulador específico. La co-regulación se puede dar también a nivel de la coordinación de dos o más reguladores globales. Por ejemplo H-NS actúa con Lrp y CRP (Lange *et al.*, 1993; Marshall *et al.*, 1998).

- Una forma de amplificar su rango de control es la de **regular** otros **factores transcripcionales**, de forma que muchos genes pueden estar regulados de forma indirecta. CRP, por ejemplo, regula 23 factores de transcripción, incluyéndose a si mismo.

- El mismo factor de transcripción global puede regular la expresión de genes cuya transcripción está dirigida por diferentes factores σ . CRP, por ejemplo, regula promotores transcritos bajo cuatro factores σ diferentes. Asimismo, una subunidad σ determinada puede transcribir un grupo de genes regulados por diferentes factores de transcripción. De esta forma la célula tiene más de una alternativa para determinar qué grupo de genes van a ser regulados o transcritos frente a un determinado estímulo. Por ejemplo, la entrada en fase estacionaria implica además de la respuesta global por los factores sigma, la regulación y la expresión de los genes implicados en la fuente de carbono precisa, y la disponibilidad de nitrógeno y oxígeno. La subunidad σ^{70} transcribe la mayoría de los genes que codifican para factores de transcripción, lo que sugiere que la maquinaria de toma de decisiones es capaz de inducir cambios en cualquier dirección cuando se adapta a un cambio en las condiciones ambientales. De hecho, sólo 6 de los 75 reguladores transcripcionales conocidos se transcriben exclusivamente por un factor σ alternativo, y ninguno de ellos está considerado como un regulador global.

- Son **capaces de detectar un gran número de cambios ambientales** a través de metabolitos específicos. La mitad de los factores de transcripción de función conocida y los predichos como tal, tienen dominios de unión a pequeños metabolitos, mientras que un 10% tiene un dominio tipo Che-Y típico de los sistemas de transducción de señal de dos componentes. La función tanto de los reguladores globales como de los específicos es

mediar una activación o represión precisas frente los cambios en metabolitos específicos. La única diferencia es que los moduladores globales captan un gran número de señales indicadoras de distintas condiciones de crecimiento.

- La mayoría de los reguladores globales descritos en la tabla 2 pertenecen a **familias** de proteínas evolutivas que contienen un **número de miembros muy bajo**. CRP, IHF, FNR, Lrp se clasifican en familias con tres o menos miembros. FIS, ArcA, NarL, SoxS, CpxR, OmPR y PhoB pertenecen a familias de 14 a 27 miembros.

- Aunque una de las características generales de los factores de transcripción es su capacidad de **autoregulación** (Thieffry *et al.*, 1998), es predominante en los factores de transcripción de gran conectividad.

- Los moduladores globales **se transcriben desacoplados de los genes que regulan**, que no es el caso de los factores de transcripción específicos. Este hecho resulta una consecuencia inevitable de la capacidad de regular muchos genes. CRP, FNR, ArcA, Lrp y H-NS se encuentran en genes aislados con un promotor único. NarL, OmpR, CpxR y PhoB, forman parte de un sistema de transducción de señal de dos componentes, y están codificados en unidades transcripcionales con sus respectivas proteínas sensoras.

En la tabla 2 se incluyen los factores de transcripción descritos como moduladores globales según los tres autores citados anteriormente (Babu y Teichmann, 2003; Shen-Orr *et al.*, 2002; Martínez-Antonio y Collado-Vides, 2003). Todos ellos coinciden sin lugar a dudas en el carácter de reguladores globales de CRP, IHF, FNR, FIS, ArcA, Lrp y H-NS. CRP es la proteína reguladora central sensora del nivel energético de la célula mediante los niveles de AMPc (Busby y Kolb, 1996; Hatfeild y Benham, 2002). FNR y ArcA están directamente relacionados con la producción de energía regulando los tipos de respiración. Lrp controla el estado nutricional general al percibir la concentración de leucina en la célula y ajustar el metabolismo a las cambiantes condiciones nutricionales. FIS, IHF, y H-NS son proteínas asociadas al nucleóide, como ya se ha comentado anteriormente en otro apartado de esta introducción.

En un segundo nivel en la jerarquía reguladora, Martínez-Antonio y Collado-Vides (2003) identifican NarL, Fur, Mlc, y CspA, también sugeridos por Babu y Teichmann (2003), Rob y PurR, descritos como tal por Shen-Orr *et al.* (2002), así como PhoB, CpxR y SoxR. En la tabla 2 se incluyen las familias más representativas de reguladores globales de la transcripción en *E. coli*, el número de factores de transcripción con los que co-regulan, y el número de genes que regulan y codifican para factores de transcripción.

Tabla 2. Moduladores globales de la transcripción (Martínez-Antonio y Collado-Vides, 2003). Los siete factores de transcripción marcados en negrita son los que cumplen las características citadas en el texto. En los miembros de la familia se incluye tanto los factores de transcripción conocidos como los predicho como tal.

Familia (n° miembros)	Factor transcripción	N° genes regulados	Co-reguladores	FT regulados
CRP (2)	CRP	197	47	22
	FNR	111	20	5
HI-HNS (2)	IHF	101	28	9
EBP (14)	FIS	76	15	4
AsnC (3)	Lrp	53	14	3
Histone-like (1)	H-NS	26	14	5
OmpR (14)	ArcA	63	18	2
	OmpR	10	9	3
	PhoB	26	1	3
	CpxR	9	2	1
LuxR/UhpA (17)	NarL	65	10	1
Fur (2)	Fur	26	8	2
AraC/XylS (27)	SoxRS	9	10	3
	Rob	7	8	2
NagC/XylR (7)	Mlc	5	3	1
Cold (9)	CspA	2	2	1
GallR/LacI (13)	PurR	28	7	1

La información sobre la red de regulación de *Escherichia coli* indica que la mayoría de los reguladores transcripcionales modulando finamente su expresión controlan un grupo de genes limitado, específico, mientras que un grupo pequeño de factores de transcripción

son reguladores globales que controlan directa o indirectamente decenas o centenas de genes (Babu y Teichmann, 2003).

Si bien las redes reguladoras están jerarquizadas, el grado de interconexiones transversales es elevado en todos los niveles de la pirámide. A este escenario se suma la acción de otros mecanismos globales de regulación como la metilación Dam (Oshima *et al.*, 2002), o el superenrollamiento en asociación con el nivel energético de la célula, del que ya se ha hablado anteriormente en esta introducción. Todo ello configura un modelo muy complejo de regulación que lleva a seleccionar qué grupo de genes van a ser modulados en cada caso, con el objetivo de obtener el perfil de expresión más óptimo para la bacteria.

1.3 LA FAMILIA LYSR DE REGULADORES TRANSCRIPCIONALES

La familia LysR de reguladores transcripcionales fue descrita inicialmente por Henikoff *et al.* (1988) en base a la homología entre un grupo de proteínas reguladoras. En 1990 esta familia ya contaba con unos 50 miembros y hoy en día se reconoce como uno de los grupos de reguladores transcripcionales más comunes en procariontes. Se han descrito representantes tanto en eubacterias, en Gram positivos y Gram negativos, como en arqueobacterias (revisado en Schell, 1993). Recientes estudios han determinado que el repertorio de 314 proteínas de unión al ADN codificadas por el genoma de *E.coli*, contiene 45 pertenecientes a esta familia (18 verificadas experimentalmente y 27 hipotéticas) (Pérez-Rueda y Collado-Vides, 2000). Durante los últimos 10 años, el número de secuencias proteicas reconocidas como tipo LysR ha aumentado de unas 50 (Schell, 1993) a unas 800, de acuerdo con la base de datos PFAM de familias de proteínas.

Define a esta familia la presencia de una secuencia de aminoácidos en el extremo N-terminal de la molécula muy conservada entre sus miembros, que incluye un motivo “helix-turn-helix” de unión al ADN, además de una serie de características comunes. Son proteínas de un tamaño similar, de unos 300 aminoácidos (+/- 20), generalmente activadores de la transcripción en presencia de una pequeña molécula inductora. En

muchas ocasiones los genes que regulan se encuentran codificados de forma adyacente al gen que codifica para el propio regulador tipo LysR y se transcriben de forma divergente, por lo que sus regiones promotoras se solapan. Este hecho facilita el control de la transcripción simultánea. La mayoría de reguladores tipo LysR reprimen su propia transcripción. Varios reguladores tipo LysR no presentan esta estructura divergente de los promotores. Ejemplos son, SpvR que está codificado adyacentemente pero en la misma dirección que el operón que regula, y OxyR que controla la expresión de un conjunto de operones repartidos por el cromosoma. De entre los sistemas sin estructura divergente de sus operones, algunos mantienen su autorregulación negativa (OxyR, Nac, CysB) mientras que otros aparentemente no se autorregulan negativamente (MleR, AlsR). En algunos casos se ha observado autorregulación positiva: SpvR y Phac son capaces de incrementar su propia transcripción.

Tal y como se ha referido anteriormente la región con mayor grado de identidad de secuencia de aminoácidos entre los miembros de la familia LysR es el extremo N-terminal de la molécula (fig. 1.3.1). La porción central (residuos 23-42) es idéntica en un 40 % en todos los miembros de la familia. Estudios de la estructura secundaria de esta región predicen la presencia del motivo “helix-turn-helix” de unión al ADN. Experimentos de mutagénesis han confirmado la presencia de este motivo HTH entre los residuos 23 y 42, el cual, junto con 15 residuos a cada lado mediatiza la unión específica de las proteínas tipo LysR a sus promotores regulados (fig. 1.3.1). Pérez-Rueda y Collado-Vides (2000) han puesto de manifiesto la existencia de una correlación entre la localización del motivo HTH en la molécula y su papel como activador o represor. Así las proteínas activadoras presentan usualmente el motivo HTH cercano al extremo C-terminal, mientras que las proteínas represoras lo tienen en el extremo N-terminal. Estas posiciones se conservan a través de las distintas familias evolutivas de proteínas (Pérez-Rueda *et al.*, 1998) e incluso se observa una tendencia entre los factores de transcripción eucariotas. Sin embargo, la familia de reguladores transcripcionales LysR es una excepción puesto que presenta el dominio de unión al ADN con el motivo HTH en el extremo N-terminal de la molécula, y sus miembros son mayoritariamente activadores de la transcripción. Esto se debe al carácter dual de la molécula, en el sentido de su papel activador de la transcripción de los genes que regula y al mismo tiempo de represor de su propia transcripción.

Como ya se ha mencionado anteriormente la mayoría de reguladores LysR son activadores de la transcripción en presencia de una molécula inductora. Por lo tanto, otro de los dominios de los reguladores LysR es el de reconocimiento del coinductor y generación de la respuesta. Estudios de mutantes con una función de reconocimiento inductor/respuesta alterada de NodD, NahR, AmpR, OxyR, AlsR, y CysB indican la presencia de dos subdominios. Los residuos comprendidos entre los 95 y 173 son importantes en la interacción entre el inductor y el regulador. Mutaciones en esta región no eliminan la capacidad de unión al ADN y pueden provocar una activación de la transcripción independiente del coinductor, probablemente mimetizando el cambio conformacional que sufre el regulador LysR tras la unión al inductor. El otro subdominio está localizado entre los residuos 196-206, y mutaciones en esta región eliminan la activación de la transcripción. Finalmente en el extremo carboxi-terminal, entre los residuos 227-253 se encuentra un dominio conservado con similitud a nivel de secuencia entre los reguladores LysR, requerido para la respuesta al inductor y en algunos casos para la unión al ADN y multimerización (fig. 1.3.1).

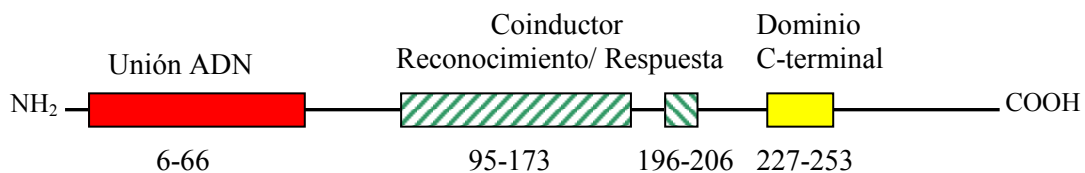


Fig. 1.3.1. Modelo de organización de dominios de un regulador LysR (Schell, 1993).

Estudios de cristalización de la proteína CysB de *Klebsiella aerogenes* (Tyrrell *et al.*, 1997) demostraron que su estructura terciaria (residuos 88-324) es muy similar a la observada en la familia de proteínas periplásmicas de unión a sustratos (PBPs), que sirven como receptores para sistemas de transporte responsables de la entrada de nutrientes en bacterias Gram negativas. Muchas de estas proteínas también sirven como receptores de quimiotaxis. Este mismo plegamiento también se observa en el dominio de unión al cofactor del represor Lac, lo que podría implicar una relación estructural entre la familia LacR de represores y la familia LysR.

Las proteínas LysR reconocen una secuencia de 15 pb de estructura T-N₁₁-A (con la T y la A formando parte de una repetición invertida) y posición (cerca de -65) comunes

en los promotores de los genes que regulan. La unión al ADN a nivel de esta secuencia no depende de la presencia del coinductor. Sin embargo cuando éste está presente se producen interacciones adicionales de los reguladores LysR con secuencias cercanas a la posición -35 de unión de la RNA polimerasa o/y un cambio de conformación del ADN que resulta en la activación de la transcripción. De forma similar a otros tipos de factores transcripcionales bacterianos, los tipo LysR actúan como homodímeros u homotetrámeros.

El modelo de activación transcripcional propuesto que seguirían los reguladores tipo LysR presenta dos fases (Schell, 1993; Sheehan y Dorman, 1998). Primero el activador se uniría en la secuencia de reconocimiento, que presenta el motivo de unión T-N₁₁-A. La unión a esta secuencia sería independiente del co-inductor. La segunda fase de activación transcripcional incluiría la interacción del regulador con otra secuencia distinta (sitio de activación) localizada aguas abajo de la secuencia de reconocimiento. Para la mayoría de activadores tipo LysR la unión a esta región requiere la presencia del inductor y es un requisito esencial para la activación de la transcripción. Además, la unión al sitio de activación depende de la unión previa de la proteína reguladora al sitio de reconocimiento. Alternativamente, ha sido demostrado para algunos sistemas que el co-inductor no es requerido para la unión de la proteína reguladora al sitio de activación sino para provocar un posterior cambio conformacional en el ADN que favorezca la transcripción (Carmel *et al.*, 1997; Rhee *et al.*, 1998). En algunos modelos regulatorios descritos el co-inductor no existe o no ha sido identificado.

Otra característica de esta familia es que sus miembros presentan un cociente Lys/Arg bajo. Posiblemente como consecuencia del elevado contenido G+C de los genes: los codones Lys contienen sólo A+T, mientras que los de Arg contienen sólo G+C. Esto seguramente lleva a la selección de Arg frente a Lys en las proteínas codificadas por ADN con elevada G+C (Viale *et al.*, 1991).

Las proteínas tipo LysR regulan genes que intervienen en funciones muy diversas, como el catabolismo de diversos substratos (Neidle *et al.*, 1989; Von Lintig *et al.*, 1991; Seeger *et al.*, 1995; Jourlin-Castelli *et al.*, 2000), biosíntesis de aminoácidos (Stragier *et al.*, 1983; Wek y Hatfield, 1986; Plamann y Stauffer, 1987; Otrowsky *et al.*, 1987; Guillouard *et al.*, 2002), utilización de fuentes de nitrógeno y azufre (Bender, 1991; Kredich, 1996; Aichi y Omata, 1997; Fernández *et al.*, 2002; Bykowski *et al.*, 2002),

respuesta al estrés oxidativo (Christman *et al.*, 1989), expresión de factores de virulencia (Caldwell y Gulig, 1991; Krause *et al.*, 1991; Goldberg *et al.*, 1991; Kovacicova y Skorupski, 2001; Deghmane *et al.*, 2002) o quórum sensing (Cao *et al.*, 2001; Sperandio *et al.*, 2002; Kim *et al.*, 2004).

Aunque muchas de las proteínas reguladoras de la familia LysR parecen ejercer su acción circunscritas a los genes u operones que le rodean, otras están inmersas en redes de regulación más complejas. SpvR regula positivamente la transcripción del operón de virulencia *spv* de *Salmonella* en coordinación con RpoS (Heiskanen *et al.*, 1994; Abe *et al.*, 1994). OxyR modula la expresión de un regulón implicado en la respuesta celular frente a condiciones de estrés oxidativo (Christman *et al.*, 1989). Recientemente se ha propuesto que la proteína LrhA de *E. coli*, implicada en el control de la estabilidad o la concentración de RpoS a través de la vía proteolítica de SprE/ClpXP (Gibson y Silhavy, 1999), tendría una función mucho más amplia en la célula. Así una de sus funciones más importantes sería la regulación transcripcional de los genes *flhDC* y del regulón *flhDC*. El regulón *flhDC* codifica para un gran número de genes implicados en la biosíntesis y en la regulación del flagelo, la motilidad y la quimiotaxis. LhrA actúa al más alto nivel jerárquico regulando la transcripción del operón *flhDC* que codifica para el regulador FlhD2C2, un heterotetrámero que controla positivamente la expresión del resto de genes del regulón (Lehnen *et al.*, 2002).

A continuación se muestran tres ejemplos de reguladores LysR estudiados en profundidad, implicados en funciones celulares diversas: la virulencia, la respuesta al estrés osmótico y la biosíntesis de aminoácidos. Con estos ejemplos se podrá observar también la compleja interacción de mecanismos de regulación global y específicos.

SpvR

Los genes plasmídicos de virulencia *spv* de *S. Typhimurium* se activan cuando la bacteria entra en fase estacionaria y contribuyen al establecimiento de una infección sistémica en el huésped. El sistema *spv* incluye a *spvR*, gen que codifica para un factor transcripcional de la familia LysR, y los cuatro genes estructurales adyacentes que forman el operón *spvABCD*. El gen *spvR* se autorregula positivamente y su producto también activa la transcripción del operón *spvABCD*. La expresión de los genes *spv* está ligada al estado fisiológico de la célula bacteriana mediante la dependencia del factor sigma RpoS

para su activación máxima. El sistema de regulación de SpvR posee algunas características poco frecuentes en los factores transcripcionales de la familia LysR al combinar la regulación con el factor sigma RpoS, al transcribirse los genes estructurales que activa a partir de la misma cadena de ADN (normalmente se transcriben de forma divergente) y al no requerir la presencia de un co-inductor. El modelo de activación transcripcional de *spvA* por SpvR responde a un mecanismo en dos pasos en el que intervienen dos regiones reguladoras (Sheehan y Dorman, 1998). El activador SpvR se une en primer lugar a la región I o región de reconocimiento formada por la repetición invertida TGTGC-N₇-GCACA. A continuación se produce la unión de SpvR a la región II o secuencia activadora, indispensable para que se produzca la activación del gen *spvA*. Sin embargo la unión del regulador LysR a la región activadora no es posible si no se ha producido anteriormente la unión a la región de reconocimiento. Otros operadores con doble secuencia de unión han sido descritos para otros reguladores LysR, por ejemplo la proteína TrpI de *Pseudomonas aeruginosa* (Piñeiro *et al.*, 1997).

Los genes de virulencia *spv* están sujetos a un complejo control a nivel transcripcional, mostrado en la figura 1.3.2. Están regulados positivamente por el factor sigma RpoS y la proteína SpvR. Están sujetos a modulación por el nivel de superenrollamiento y la proteína asociada al nucleoide H-NS (Robbe-Saule *et al.*, 1997). IHF se une aguas arriba al promotor de *spvR*, estimulando la expresión del gen *spvR* y por lo tanto la expresión del operón *spvABCD* (Marshall *et al.*, 1999). Y el regulador Lrp compete con SpvR por la región II o activadora del promotor de *spvA* cuando no hay leucina, impidiendo la activación de la transcripción de los genes estructurales (Marshall *et al.*, 1999).

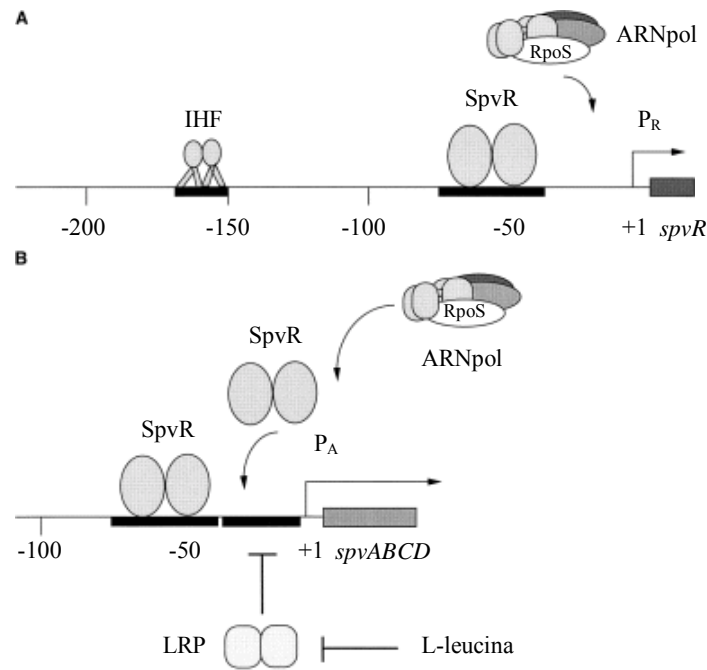


Fig. 1.3.2. Resumen de la regulación de los promotores *spv* (Marshall *et al.*, 1999). En la figura A se puede observar cómo se une SpvR y captura a la ARN polimerasa. El sitio de unión de IHF se encuentra en la posición -160. En la figura B se muestra la unión de SpvR al sitio de unión del promotor más lejano, que captura a más proteína SpvR en el sitio de unión más cercano, necesario para la activación de la transcripción del operón *spvABCD*. El regulador global Lrp en ausencia de leucina ocupa el sitio de unión de SpvR más cercano, impidiendo la activación de la transcripción

NhaR

En *Escherichia coli* existen dos antiporters, NhaA y NhaB, que intercambian específicamente iones Na^+ y Li^+ por H^+ . El gen *nhaA* es indispensable para superar la toxicidad del Li^+ y para crecer a pH alcalino en presencia de Na^+ . La expresión de *nhaA* está regulada positivamente por *nhaR*, gen que codifica para un polipéptido de 34,2 KDa (NhaR) miembro de la familia LysR de reguladores transcripcionales. El gen *nhaR* se encuentra aguas abajo de *nhaA*, y en la secuencia aguas arriba de *nhaR* que lo separa de *nhaA* no se ha encontrado ningún promotor. NhaR es un regulador positivo de la expresión de *nhaA* sensible al Na^+ (Rahav-Manor *et al.*, 1992). La regulación de *nhaA* mediada por NhaR depende de dos procesos (Dover *et al.*, 1996): inducción de la transcripción de *nhaA* al aumentar la concentración de Na^+ intracelular, y un control de “feedback” negativo, establecido por NhaA en las condiciones en las que este antiporter, y no NhaB, determina los niveles del inductor, la concentración de Na^+ intracelular. Se ha comprobado que H-NS es un represor de la transcripción de *nhaA*, esté presente o no NhaR, pero en presencia de Na^+ NhaR es capaz de vencer la represión mediada por H-NS y activar la transcripción de

nhaA (Dover *et al.*, 1996). Ensayos de protección contra la digestión por DNasaI demostraron que la región de unión de NhaR en la región promotora de *nhaA* se extiende en 92 pb (Carmel *et al.*, 1997), sugiriendo que se une al ADN de forma multimérica. Se han encontrado tres motivos de unión designados como I, II, y III, diferentes entre ellos pero que mantienen la secuencia consenso de unión de los reguladores LysR (T-N₁₁-A). NhaR se une a la región promotora de *nhaA* aún en ausencia de su inductor, característica compartida con otros reguladores LysR (Toledano *et al.*, 1994), pero sólo en presencia de Na⁺ actúa como sensor y transductor de la señal. El Na⁺ provoca un cambio conformacional en NhaR que le permite activar la transcripción de *nhaA* (Carmel *et al.*, 1997).

Recientemente se ha comprobado que NhaR es responsable de la inducción osmótica de *osmC*_{p1} (Toesca *et al.*, 2001), el promotor del gen inducible por estrés *osmC*. El gen *osmC* codifica para una hipotética proteína de envuelta necesaria para la resistencia a peróxidos orgánicos y supervivencia a largo plazo en fase estacionaria, que se transcribe a partir de dos promotores solapados *osmC*_{p1} y *osmC*_{p2}. NhaR es responsable de la estimulación de *osmC*_{p1} por NaCl, LiCl, y en menor grado por una solución de sacarosa. El sitio de unión protegido contra la Dnasa I en *osmC*_{p1} es de aproximadamente 30 pb, menor que en el promotor de *nhaA* (92 pb), y probablemente también es menor la afinidad de unión (Sturny *et al.*, 2003).

IlvY

La proteína IlvY es un miembro prototipo de la familia LysR de reguladores transcripcionales, necesaria para la regulación del gen *ilvC* en *Escherichia coli* (Watson *et al.*, 1979; Wek *et al.*, 1986; Henikof *et al.*, 1988). El gen *ilvC* codifica para el enzima ácido acetohidroxi isomeroreductasa, implicado en la biosíntesis de aminoácidos de cadena ramificada: L-isoleucina, L-valina, y L-leucina. La proteína IlvY, junto a cualquiera de los sustratos de este enzima, el α -acetolactato o α -acetohiroxibutarato, induce la activación de la transcripción del gen *ilvC*. Este regulador LysR se une de modo altamente cooperativo a las dos secuencias operadoras en tandem de la región promotora solapada y divergente de los genes *ilvY* e *ilvC*, independientemente de la presencia del inductor y autorregula negativamente su propia expresión por un mecanismo de oclusión de la ARN polimerasa (Rhee *et al.*, 1998). Sin embargo, la unión de la proteína IlvY no es suficiente para activar la transcripción del gen *ilvC*, siendo indispensable la unión del inductor (los sustratos del

enzima), que produce un cambio conformacional en la doble hélice de ADN. Este cambio en el complejo proteína-operador facilita el secuestro de la ARN polimerasa en el promotor de *ilvC* sin afectar la ocupación de la proteína IlvY en las dos zonas operadoras en tandem (Rhee *et al.*, 1998).

Por la estructura divergente de los promotores del operón *ilvYC* se estudió la posibilidad de que la actividad transcripcional de un promotor sensible al superenrollamiento podría estar influido por el superenrollamiento negativo local generado por la transcripción desde un promotor cercano (Liu y Wang, 1987). Efectivamente se encontraron evidencias *in vivo* e *in vitro* del acoplamiento transcripcional de los genes *ilvYC* (Rhee *et al.*, 1999): el aumento y la disminución de la actividad transcripcional del promotor de *ilvY* se correlaciona con el aumento y la disminución de la transcripción desde el promotor de *ilvC*. Además, la transcripción de *ilvC* aumenta unas diez veces por encima del rango fisiológico de superenrollamiento negativo en células de *E. coli* (Rhee *et al.*, 1999). Este mecanismo dependiente del grado de superenrollamiento permite ajustar la expresión del gen *ilvC* al superenrollamiento global, que está influido por las condiciones nutricionales y de crecimiento de la célula (Rhee *et al.*, 1999; Opel y Hatfeild, 2001). Opel y Hatfeild (2001) postulan que este tipo de regulación global de la expresión podría explicar la distribución divergente de muchos sistemas LysR.

1.4 OSMOREGULACIÓN

Como se ha mencionado anteriormente, las bacterias son capaces de notar y responder a los cambios en el ambiente. Uno de los parámetros ambientales es la osmolaridad del medio extracelular. Las células bacterianas deben mantener una presión osmótica intracelular mayor a la del medio de crecimiento, lo que genera la llamada turgencia celular, considerada como la fuerza conductora de la extensión, crecimiento y división celular. Así, la habilidad para adaptarse a cambios en la osmolaridad del medio ambiente es fundamental para el crecimiento y la supervivencia.

1.4.1 LA RESPUESTA DE LAS BACTERIAS AL ESTRÉS HIPO-OSMÓTICO

En una situación en la que la osmolaridad del medio disminuye drásticamente, por ejemplo en las bacterias del suelo sobre las que cae una lluvia repentina, el agua entra e inmediatamente aumentan la presión de turgencia y la tensión de membrana, corriendo las células el peligro de explotar. En estas condiciones, la adaptación al choque hipo-osmótico implica la salida de solutos por un sistema de transporte específico (transporte secundario) e inespecífico (canales “mecanosensitivos”), conjuntamente a la salida de agua por las “aquaporinas” (Sleator y Hill, 2002).

Los canales mecanosensitivos en bacterias se caracterizan por una elevada velocidad de transporte y conductancia (permeabilidad). Además tienen poca especificidad por determinados iones o solutos. Como resultado, este tipo de canales permiten la adaptación casi al instante y son la principal ruta de salida de solutos citoplasmáticos durante la transición de un medio de alta a baja osmolaridad (Sleator y Hill, 2002; Morbach y Krämer, 2002). *E. coli* posee de tres a cinco canales mecanosensitivos, y los mejor caracterizados son MscL, MscS y MscM. Estos canales se activan por el estiramiento de la membrana, hecho demostrado por su activación observada al intercalar compuestos anfipáticos en la bicapa lipídica (Martinac *et al.*, 1990).

Evidencias recientes sugieren que las bacterias, como las células de plantas y animales, poseen aquaporinas, por ejemplo AqpZ de *Escherichia coli* (Calamita, 2000). Estos canales específicos facilitan el rápido eflujo/influjo de agua aliviando el estrés hídrico sin afectar al potencial transmembrana.

1.4.2 LA RESPUESTA DE LAS BACTERIAS AL ESTRÉS HIPER-OSMÓTICO

Principalmente se conocen dos estrategias de adaptación a la elevada osmolaridad del medio (Sleator y Hill, 2002): la de “sal en el citoplasma” y la del “solute compatible”. La estrategia de acumulación de sal en el citoplasma es típica de miembros de la familia Halobacteriaceae (Martin *et al.*, 1999), y consigue el equilibrio osmótico manteniendo la concentración de KCl intracelular similar a la del medio externo. Como consecuencia, el citoplasma está expuesto a una gran fuerza iónica, haciendo necesarias adaptaciones

estructurales. De este modo, las proteínas de las Halobacterias han sufrido un gran número de sustituciones de aminoácidos, que implican enriquecimiento en aspartato, glutamato y disminución de los residuos hidrofóbicos (Lanyi, 1974). Los organismos que utilizan este sistema de adaptación están confinados a ambientes de elevada osmolaridad.

La estrategia del soluto compatible es una respuesta bifásica. Como respuesta primaria hay un aumento de la concentración de K^+ intracelular, acompañado por la síntesis de glutamato en *Escherichia coli* y *Corynebacterium glutamicum*, y de prolina en *Bacillus subtilis* (revisado en Morbach y Krämer, 2002). Esta primera respuesta rápida de acumulación de K^+ conduce a una elevada concentración de iones que puede inducir la agregación de macromoléculas al facilitar las interacciones hidrofóbicas. Por lo tanto las bacterias intercambian el potasio acumulado inicialmente por solutos compatibles, lo que sería la respuesta secundaria. Tales compuestos son, en general, moléculas altamente solubles sin carga neta a pH fisiológico, que no interaccionan con proteínas. Estas características facilitan su acumulación a altas concentraciones intracelulares sin molestar a procesos celulares vitales como la reparación del ADN, las interacciones ADN-proteína o la maquinaria metabólica celular (Csonka y Epstein, 1996). Los principales solutos compatibles en bacterias por orden de importancia son: glicina-betaína, carnitina y prolina (Beumer *et al.*, 1994). Casi todas las bacterias poseen multitud de sistemas de entrada, que en general tienen diferente especificidad y afinidad de sustrato, lo que permite una óptima adaptación a las cambiantes condiciones ambientales. Hasta ahora se han identificado sistemas de entrada de solutos compatibles regulados osmóticamente entre todas las diferentes clases de sistemas de transporte (Morbach y Krämer, 2002): transportadores ABC dependientes de unión a proteína (ProU de *E. coli* o OpuA de *B. subtilis*), transportadores secundarios dependientes de unión a proteína, o transportadores dependientes de Na^+ o H^+ (BetP o EctP de *C. glutamicum* o ProP de *E. coli*).

1.4.2.1 Regulación del operón *kdp*

La primera respuesta de *Escherichia coli* ante un aumento de la osmolaridad es la rápida acumulación de K^+ a través del sistema de baja afinidad Trk y el de elevada afinidad Kdp. La ATPasa “tipo-P” *kdpFABC* está codificada por el operón *kdpFABCDE*, que incluye el sistema regulador de dos componentes KdpDE. La percepción de la señal induce la autofosforilación de la quinasa sensora KdpD, que transfiere el grupo fosfato al

regulador de respuesta soluble KdpE, el cual se une al promotor de *kdpFABC* y estimula la transcripción. KdpD también es el responsable de la represión del operón *kdpFABC*, ya que incluye una actividad fosfatasa que resulta en la desactivación de KdpE (revisado en Morbach y Krämer, 2002).

La clase de estímulo que percibe KdpD ha generado mucha controversia. Uno de los modelos es el llamado de turgencia (Laimins *et al.*, 1981; Jung *et al.*, 1997), según el cual una disminución en la presión de turgencia induce la autofosforilación de KdpD. Sin embargo, otros estudios han demostrado que la pérdida de turgencia no puede ser la única señal, y han evidenciado que la concentración de K^+ intracelular puede funcionar también como señal (Rhoads *et al.*, 1976; Gowrishankar, 1985). Otros autores han conseguido mutantes en KdpD sensibles a un estrés osmótico y no a una limitación de K^+ , o viceversa, demostrando que KdpD puede activarse por diferentes estímulos (Sugiura *et al.*, 1994; Jung *et al.*, 1997). Más recientemente Jung *et al.* (2001) demostraron que el CsCl (que disminuye la concentración de K^+ intracelular) induce la expresión del operón *kdpFABC* a un nivel más alto que el NaCl, remarcando que además de la osmoregulación el sistema Kdp es fundamental en la homeostasis de K^+ .

1.4.2.2. Regulación de ProP y ProU

En estos sistemas transportadores la función osmosensora y osmoreguladora residen en la misma proteína. Sin embargo, aunque los dos responden a un estrés osmótico aumentando tanto su actividad como su transcripción, en ProP la regulación principal se ejerce a nivel de actividad y en ProU a nivel de transcripción (revisado en Sleator y Hill, 2002; Morbach y Krämer, 2002).

El sistema ProP cataliza el simporte de H^+ / betaína, prolina y ecotoína, con afinidades similares. La transcripción de *proP* se dirige a través de dos promotores, P-1 y P-2, ambos activados por el aumento de la osmolaridad del medio. Mientras que CRP reprime *proP* P-1, la actividad de *proP* P-2 depende de RpoS y la proteína asociada al nucleóide Fis (Mellies *et al.*, 1995; Xu y Johnson, 1997). Respecto al aumento de actividad de ProP a elevada osmolaridad el mecanismo más probable implica una modulación de la formación “ α -hélice coiled-coil” del dominio carboxi-terminal de ProP. A nivel post-

traduccional la proteína citoplasmática ProQ también influye en la activación del sistema transportador ProP a elevada osmolaridad (Culham *et al.*, 2000).

ProU es una ATPasa multicomponente, que transporta glicina betaína con más afinidad que prolina. Los componentes de ProU están codificados en un operón que contiene 3 cistrones: *proV*, *proW* y *proX*, que codifican para dos proteínas de membrana citoplasmática, ProV y ProW, y la proteína de unión periplasmática ProX. Se ha identificado dos promotores: uno osmoregulado dependiente de RpoD, situado aguas abajo del otro promotor, débil y dependiente de RpoS. Más que una señal específica dependiente de K^+ , el aumento de transcripción de *proU* tras un choque osmótico, es un reflejo de la estimulación del ión sobre las reacciones enzimáticas en general (Csonka y Epstein, 1996). El nivel de superenrollamiento funciona también como un regulador de la expresión de *proU*: mutaciones en *topA* aumentan su transcripción, mientras que mutaciones en *gyrA* y *gyrB* en *Salmonella* reducen la expresión del operón a baja osmolaridad. Asimismo, las proteínas asociadas al nucleóide H-NS, IHF, o HU, actúan como moduladores, más que reguladores, de la expresión de *proU* (Kempf y Bremer, 1998). En condiciones normales H-NS se encuentra unida a dos regiones ricas en A-T aguas arriba y abajo del promotor dependiente de RpoD, evitando la unión de la ARN polimerasa y por lo tanto la transcripción. Cuando la osmolaridad del medio aumenta se rompe esta unión ADN-proteína. Mutaciones en IHF disminuyen el nivel de inducción de *proU* aproximadamente dos veces, mientras que mutaciones en HU-B reducen la transcripción tanto en condiciones de expresión basal como en condiciones de inducción (revisado en Sleator y Hill, 2002).

1.4.3 OSMOLARIDAD Y EXPRESIÓN DE PROTEÍNAS ESTRUCTURALES

Los genes osmoregulados más extensamente estudiados que no contribuyen directamente a la acumulación del soluto compatible son *ompC* y *ompF*, los cuales codifican para los dos canales de membrana externa de Gram negativos OmpC y OmpF. La expresión de estas porinas, que facilita la difusión inespecífica de pequeñas moléculas hidrofílicas (<500 Da) a través de la membrana, responde al incremento de la osmolaridad externa: *ompF* se reprime y *ompC* aumenta su expresión. Estas porinas también responden a una gran variedad de parámetros ambientales incluyendo la temperatura, la fuente de carbono, la disponibilidad de oxígeno así como el pH del medio. Otros genes que se

inducen a elevada osmolaridad y codifican para otras proteínas de pared son los que codifican para OsmB, OsmC, OsmE y OsmY (revisado en Wood, 1999). Se cree que OsmB y OsmY son lipoproteínas.

Los niveles de las porinas OmpC y OmpF son controlados predominantemente a nivel de expresión génica por el sistema regulador de dos componentes OmpR/EnvZ, pero en el caso de *ompF* hay un nivel adicional de control que implica al RNA antisentido MicF: actúa como un regulador negativo de la expresión de *ompF* a nivel post-traducciona (Pratt *et al.*, 1996). EnvZ es la proteína transmembrana osmosensora que activa por fosforilación a la proteína OmpR, la cual actúa como un activador de la transcripción (Csonka y Hanson, 1991). Aunque se desconoce la señal ambiental a la que responde la proteína EnvZ, se ha propuesto que, por su localización y estructura, debe captar una señal existente en el citoplasma o periplasma opuesta al exterior de la célula. En este sentido los oligosacáridos altamente aniónicos del periplasma de Gram negativos (MDO_s), se han considerado como la señal (Fielder y Rottering, 1988). Al contrario de los solutos compatibles los niveles de estos oligosacáridos disminuyen con el aumento de la osmolaridad.

1.4.4 OSMOREGULACIÓN Y VIRULENCIA

Cada vez es más evidente que además de los factores de virulencia clásicos, es decir los que codifican para toxinas, invasinas o adhesinas, existen otras proteínas esenciales para que el patógeno colonice. De los diferentes ejemplos existentes, aquí se menciona uno bien caracterizado: la inducción dependiente de osmolaridad de las proteínas del sistema de secreción de tipo III (TTSSs), presentes en las islas de patogenicidad 1 y 2 de *Salmonella*.

La mutación del gen *ompR* reduce dramáticamente la virulencia de *Shigella flexneri* y *S. Typhimurium*. Las proteínas EnvZ y OmpR intervienen en el control de la virulencia y la osmoregulación de los genes de invasión *vir* en *Shigella* como factores de regulación positiva (Bernardini *et al.*, 1990). En *S. Typhimurium* mientras una mutación en el gen *ompR* no afecta a la expresión del gen de invasión *invA* que se induce en un medio de elevada osmolaridad, si que afecta a la virulencia de esta especie (Dorman *et al.*, 1989).

La pérdida simultánea de las porinas OmpF y OmpC que contribuyen en la virulencia de *S. Typhimurium* y *S. Typhi*, sólo puede explicar parcialmente este efecto, por lo que a la proteína OmpR se le atribuye algún otro papel en la virulencia (Chatfield *et al.*, 1991). La transcripción del operón *mxi* (membrane expression of invasion plasmid antigens) de *S. Typhi* que se induce a elevada osmolaridad, no presenta esta osmoregulación en un mutante *ompR*. Además, Pickard *et al.* (1994) demostraron que la cápsula Vi de *S. Typhi*, cuya síntesis disminuye al aumentar la osmolaridad del medio, estaba también afectada por mutaciones en *ompR*.

En *Salmonella* los sistemas de secreción de tipo III (TTSSs) que se encuentran en la isla de patogenicidad 2 (SPI-2), necesarios para una infección sistémica y la replicación intracelular dentro de los macrófagos y las células epiteliales, se inducen a baja osmolaridad y a pH ácido (Lee *et al.*, 2000; Garmendia *et al.*, 2003). Esta inducción es completamente dependiente del sistema regulador de dos componentes SsrA-B, codificados en SP-2, y parcialmente dependiente de EnvZ/OmpR (Garmendia *et al.*, 2003). La expresión de *ssrAB* está regulada por la unión de OmpR a su región promotora.

En cambio, los sistemas de secreción de tipo III (TTSSs) que se encuentran en la isla de patogenicidad 1 (SP-1) se expresan en las condiciones de crecimiento óptimas que reflejan las del lumen del intestino, incluyendo baja concentración de oxígeno, elevada osmolaridad y una alcalinidad suave (Bajaj *et al.*, 1996). La expresión de los genes de invasión está controlada por HilA, un miembro de la familia de reguladores transcripcionales OmpR/ToxR por similitud de secuencia en su dominio de unión al ADN. La expresión del gen *hila* está regulada por las mismas condiciones ambientales que regulan el fenotipo invasivo. Se han identificado muchas mutaciones en genes que alteran la regulación positiva o negativa de *hila*. Estas incluyen reguladores positivos como *hilD*, *hilC*, *fis*, y *csrAB*. HilD es un regulador de la familia AraC/XylS que se une a la región promotora de *hila*, y es necesario para su completa activación al interactuar con la subunidad α de la ARN polimerasa (Boddicker *et al.*, 2003). HilC tiene una elevada homología a HilD y también se une a la región promotora de *hila*, sin embargo no es necesario para la expresión de *hila*. CsrA parece regular los niveles de ARN mensajero de *hilC* y *hilD*. Otras mutaciones en genes identificados como reguladores negativos son: *hha*, *hns*, *hilE*, *ams*, *pag* y *lon*. Hha y H-NS son proteínas asociadas al nucleóide que reprimen la expresión del gen *hila* a baja osmolaridad (Fahlen *et al.*, 2001; Schechter *et al.*, 2003).

HilE es una proteína específica de *Salmonella* que evita la activación de *hilA* por interacción directa proteína-proteína con HilD (Baxter *et al.*, 2003). Boddicker y Jones (2004) han sugerido que después de la invasión de las células epiteliales, cuando el patógeno crece ya intracelularmente, la proteasa Lon degrada al activador transcripcional HilD, impidiendo la transcripción de *hilA*.

ANTECEDENTES

1. Obtención de la cepa YFER

En el marco del estudio de la regulación de la expresión génica en respuesta a factores ambientales, al iniciar este trabajo se disponía en el grupo de investigación de una serie de cepas de *Salmonella enterica* serovar Typhimurium con mutaciones en genes que presentaban una expresión diferencial dependiendo de la osmolaridad del medio. Estas cepas habían sido obtenidas por el Dr. Prenafeta (1999) utilizando una estrategia distinta a la habitual en el estudio de la osmoregulación: mientras que habitualmente se han identificado y estudiado genes cuya expresión se induce cuando la osmolaridad aumenta, el Dr. Prenafeta buscó genes cuya expresión se reprimiese a elevada osmolaridad. Con esta finalidad se realizó una mutagénesis sobre la cepa TT1704 de *S. Typhimurium*, mediante el transposón defectivo MudJ (Hughes y Roth, 1988). Las inserciones de este transposón generan fusiones génicas con el gen *lacZ* que permiten determinar el grado de expresión del gen mutado mediante la medida de la actividad β -galactosidasa de la cepa mutante. De esta forma se seleccionaron y se aislaron cepas mutantes con fusiones génicas *lacZ* que presentaban un fenotipo Lac⁺ a baja osmolaridad y Lac⁻ a elevada osmolaridad.

2. Identificación de la inserción de MudJ en la cepa YFER como un hipotético regulador transcripcional de la familia LysR

Con el objetivo de identificar alguno de estos genes osmoregulados se clonó la región que contenía la fusión génica, a fin de realizar posteriormente la correspondiente secuenciación. Para ello, el ADN total fue digerido con un enzima que no eliminase la resistencia a Km del fago MudJ, permitiendo así seleccionar los clones transformantes que contenían el ADN cromosómico con la inserción del transposón. La secuencia obtenida a partir de una de las cepas mutantes, denominada OSMT, fue considerada de particular interés, ya que era homóloga en un 92% al gen de *Escherichia coli* YYBE, una hipotética pauta de lectura abierta que codificaría una proteína de 308 aminoácidos. Esta hipotética proteína podía pertenecer a la familia LysR de reguladores transcripcionales. Analizando la secuencia de nucleótidos del gen YYBE y la secuencia obtenida de la cepa OSMT, se determinó que la inserción de MudJ en esta cepa se localizaba justo cuatro nucleótidos antes del primer triplete ATG de inicio de la traducción.

Carolina Polo (2000) continuó la secuenciación de la inserción del fago *MudJ* en la cepa OSMT. Hasta el momento se había obtenido una secuencia parcial del gen, dado que la estrategia utilizada sólo permitió clonar 600 pb desde el extremo 3' de la inserción de *MudJ*, y la traducción de la secuencia nucleotídica a aminoacídica no presentaba un codón de terminación del gen. Para obtener la secuencia completa del gen se utilizó el cósmido pLA2917, que permite clonar fragmentos de aproximadamente 20 kpb. Esto permitió clonar el fago *MudJ* entero (11,3 kpb), más las secuencias que se encontraban a ambos lados de la inserción del transposón. De este modo se pudo continuar la secuenciación del gen, hasta que, analizando las pautas de lectura, se localizó un codón de terminación situado a 924 pb del inicio del gen. La disponibilidad de más datos de secuencia confirmó la homología con la pauta de lectura abierta del gen *YYBE* de *E. coli*, que codificaría como un hipotético regulador transcripcional similar a la familia *LysR*. Además, la secuencia de aminoácidos deducida también presenta homología con otras proteínas *LysR* caracterizadas como tal. Así, la inserción del transposón *MudJ* en la cepa OSMT se encontraba en un gen que codificaría para un hipotético regulador transcripcional tipo *LysR*. Este gen, además, se expresa de forma diferencial según la osmolaridad del medio. En base a estos resultados se planteó el estudio del modelo de regulación que se recoge en esta Memoria.

En la actualidad, la pauta de lectura abierta de *E. coli* *YYBE*, ha cambiado de nombre, denominándose *yfeR*. Esta pauta de lectura abierta está localizada también, recibiendo el mismo nombre, en el genoma de *S. Typhimurium*, secuenciado en su totalidad (McClelland *et al*, 2001) y disponible hoy en día en las bases de datos. En consecuencia, la cepa OSMT será identificada en este trabajo como *YFER*, y la hipotética pauta de lectura abierta donde se encuentra la inserción de *MudJ* como *yfeR*.