

1.- INTRODUCCIÓN

1.1.- FACTORES AMBIENTALES Y MODULADORES GLOBALES DE LA EXPRESIÓN GÉNICA

Las células bacterianas son organismos procariotas que se desarrollan en un entorno que, en la mayoría de ocasiones, varía constantemente. Para la bacteria las condiciones nutricionales y ambientales pueden cambiar de manera drástica y rápida, lo que tiene como consecuencia que estas células hayan desarrollado una gran capacidad de adaptación a los frecuentes cambios que se producen en el medio en el cual se desarrollan.

La adaptación de los microorganismos implica, por una parte, el reconocimiento de ciertas señales captadas del medio externo y, por otra, la elaboración de una respuesta que generalmente se encuentra relacionada con un cambio en su patrón de expresión génica. La adaptación de las células bacterianas al hábitat a través de un cambio en el patrón de expresión génica representa una estrategia muy útil para conseguir un aprovechamiento óptimo de la energía disponible (Mahan *et al.*, 1996).

Las señales tanto físicas como químicas pueden ser detectadas de muy diferentes formas: el estímulo puede penetrar al interior celular (por ejemplo, el sustrato de una vía metabólica), puede provocar cambios en algún componente celular (por ejemplo, modificación del superenrollamiento del ADN) o incluso ser detectado a través de receptores específicos transmembrana, que reconocen la señal a nivel de superficie celular (por ejemplo la proteína EnvZ de *Escherichia coli* y *Salmonella* (Csonka y Hanson, 1991)). Todas estas señales que la célula bacteriana detecta influirán sobre la maquinaria bioquímica y, como se ha dicho anteriormente, esto sucederá a través de un cambio en el patrón de expresión génica de estas células.

Todo el proceso de regulación de la expresión génica requiere la presencia en la célula procariota de unos mecanismos moleculares que pueden ser tan o más complejos que los existentes en la célula eucariota. La regulación de la expresión génica de la célula bacteriana implica diferentes niveles: a nivel de transcripción (inicio, elongación y terminación), a nivel de procesamiento y estabilidad del ARN mensajero, a nivel de traducción y finalmente a nivel de modificaciones post-traduccionales (activación o inactivación de la proteína y degradación de la misma) (Rojo, 1999). Cuando se analizan los diferentes mecanismos de regulación de la expresión génica presentes en bacterias se hace aparente que la mayor parte se encuentran actuando a nivel de transcripción y más concretamente, del inicio de la transcripción.

1.1.1.- REGULACIÓN EN RESPUESTA A ESTÍMULOS AMBIENTALES DE FACTORES DE VIRULENCIA

Dentro de los genes que se están regulados a través de factores ambientales se encuentran aquellos que codifican para factores de virulencia. Los microorganismos patógenos, al igual que el resto de microorganismos, se desarrollan en un ambiente con constantes variaciones, pero además, las bacterias patógenas para su supervivencia y multiplicación necesitan detectar aquellas condiciones que sean más favorables para el establecimiento de un proceso de colonización (por ejemplo detectar señales de los propios tejidos del hospedador para poder interactuar con él, adherirse y colonizarlo) y responder de manera adecuada a través de, la expresión de entre otros genes, los factores de virulencia. Clásicamente, la patogenicidad bacteriana se ha definido como el conjunto de mecanismos bioquímicos por los cuales un microorganismo es capaz de causar enfermedad (Smith, 1968, revisado en Wilson *et al.*, 2002).

Muchos de los parámetros ambientales que afectan a la expresión de los factores de virulencia han sido ya identificados y mayoritariamente se corresponden a los mismos parámetros físico-químicos que la regulación génica utiliza para la adaptación a los cambios ambientales fuera de la célula hospedadora. Algunos de estos parámetros serían: temperatura, osmolaridad, pH, niveles de oxígeno, disponibilidad de hierro, concentración de iones y nutrientes, estrés, etc.

En general, los factores de virulencia propiamente dichos (toxinas, factores de adherencia, invasinas, cápsulas, etc.), no son componentes de la estructura celular, ni participan en las funciones biológicas de la célula cuando ésta se encuentra fuera del hospedador. Es decir, en estas condiciones no son esenciales. Por este motivo es importante el control de la regulación de los factores de virulencia, para que únicamente se expresen en el momento y ambiente adecuados. Además, raramente estos factores se expresan de forma constitutiva, ya que esto significaría un gasto energético demasiado grande (Mahan *et al.*, 1996).

En función de la fase de infección o del entorno anatómico del hospedador donde se produzca la infección, puede ser más importante un tipo de factor de virulencia u otro. Por esta razón es importante que la regulación de los factores de virulencia esté controlada no sólo por las señales que indiquen el paso de una forma de vida libre a estar asociado a un hospedador, sino también por las diferentes señales que vayan indicando el punto dentro

del proceso de infección en el que el microorganismo patógeno se encuentra en cada momento.

1.1.2.- REGULACIÓN GÉNICA A NIVEL TRANSCRIPCIONAL

Como se ha comentado anteriormente, el principal mecanismo de control de la expresión génica lo encontramos a nivel transcripcional. Esta regulación actúa a diferentes niveles, pero principalmente en la primera etapa: el inicio de la transcripción. Al actuar el proceso de regulación en un punto tan inicial, se produce en las células bacterianas un ahorro energético muy importante. De esta forma no se sintetizan proteínas que no sean necesarias y el ahorro es mayor que actuando a nivel de la inactivación o degradación de las proteínas ya sintetizadas.

En su inicio, la transcripción puede ser activada o reprimida por diferentes mecanismos: cambios o modificaciones en la ARN polimerasa, factores de transcripción (proteínas reguladoras de unión al ADN) y cambios en el grado de superenrollamiento del ADN.

En muchas ocasiones la adaptación a los cambios ambientales que se producen en células procariontas requiere de la modificación de diferentes genes, operones o regulones de forma conjunta (denominamos regulones al grupo de operones coordinadamente regulados). Estas redes reguladoras también tienen como objetivo responder a un cambio ambiental y pueden presentar un elevado grado de complejidad (Neidhardt y Savageau, 1996).

1.1.2.1.- ARN polimerasa

La ARN polimerasa bacteriana (RNAP) es un holoenzima dependiente de ADN y concretamente es el componente central en la regulación transcripcional y el responsable de toda la síntesis celular de ARN (Ebright, 2000).

El núcleo enzimático está formado por las subunidades $\alpha_2\beta\beta'\omega$, y contiene toda la maquinaria necesaria para la síntesis del ARN, pero previamente es necesaria la formación del holoenzima, es decir, la unión con el factor σ correspondiente, para que se produzca el

inicio de la transcripción. La subunidad σ tiene 3 funciones principales: asegurar el reconocimiento de las secuencias específicas promotoras, posicionar el holoenzima en el promotor diana y facilitar la apertura de la doble hélice de ADN cerca del inicio de transcripción (Wösten, 1998; revisado por Browning y Busby, 2004).

Los diferentes factores σ comparten una serie de características comunes con excepción de la familia σ^{54} . Son proteínas multiméricas que tienen 4 dominios funcionales unidos por conectores. En el caso de los dominios 2, 3 y 4, se sabe que están implicados en el reconocimiento de la región promotora, pero la función del dominio 1 todavía no es conocida y de hecho este dominio se encuentra ausente en algunos factores σ (Gruber y Gross, 2003; Browning y Busby, 2004). Las secuencias más importantes en el reconocimiento de la región promotora son las cajas -10 y -35 , pero también intervienen el motivo localizado inmediatamente antes de la caja -10 y el elemento UP (elemento medio-lejano, localizado aguas arriba del promotor).

El inicio de transcripción en bacterias que se encuentran en fase de crecimiento exponencial es catalizado por el holoenzima RNAP ($E\sigma$) formado por un factor σ “housekeeping”, similar al factor σ^{70} de *E. coli*. Existen factores σ alternativos que pueden competir por la ARN polimerasa para la formación del holoenzima y que se activan en determinadas condiciones de estrés, crecimiento y durante cambios morfológicos. Los factores σ pueden interactuar con los factores de transcripción, ya sean activadores o represores (Ebright y Busby, 1995; Hochschild y Dove, 1998; Dove *et al.*, 2003). Además de los factores de transcripción y los factores σ , otros mecanismos pueden afectar a la distribución de la RNAP en los diferentes promotores: la secuencia promotora, metabolitos de bajo peso molecular (por ejemplo, el ppGpp), efectos en la topología del ADN y proteínas de unión al ADN como H-NS, Lrp y IHF que se unen en las regiones promotoras (Browning y Busby, 2004).

E. coli, además del factor σ^{70} (factor “housekeeping”), presenta otros 6 factores sigma, 5 de ellos pertenecientes a la misma familia que σ^{70} (σ^s , σ^{32} , σ^F , σ^E y $\sigma^{F_{cel}}$), siendo el sexto el factor σ^{54} (Ishihama, 2000). Los diferentes factores σ permiten a la célula bacteriana alterar su patrón de expresión en respuesta a diferentes estímulos, respondiendo con la expresión de un determinado grupo de genes a diferentes situaciones de estrés (estrés oxidativo, radiación UV, choque térmico, hiperosmolaridad, pH ácido, deficiencia de nitrógeno, hierro, etc.) (Ishihama, 2000). Por ejemplo, $E\sigma^{32}$ transcribe genes que

codifican para proteínas del choque térmico, y se induce por un número excesivo de proteínas citoplasmáticas que no se pliegan correctamente (Browning y Busby, 2004).

Finalmente, en algunos casos los factores σ tienen su actividad controlada a través de un factor anti-sigma. Este factor los mantiene secuestrados, impidiendo su interacción con la RNAP bacteriana (Hughes y Mathee, 1998).

1.1.2.2.- Factores de transcripción

En base al análisis de secuencias de proteínas se han descrito más de 300 genes en *E. coli* que se corresponderían con factores de transcripción, es decir, proteínas capaces de unirse a las regiones promotoras de determinados genes y regular su transcripción (Pérez-Rueda y Collado-Vides, 2000; Babu y Teichmann, 2003). De todos estos genes, únicamente en la mitad de ellos se han comprobado experimentalmente sus funciones.

Babu y Teichmann (2003) describen que tres cuartas partes de los factores de transcripción se componen de dos dominios estructurales, un dominio de unión al ADN (predominantemente el dominio hélice-giro-hélice) y un dominio regulador (enzimático, de unión a pequeñas moléculas, etc.). Además, los factores de transcripción se pueden agrupar en familias en función de su secuencia (Pérez-Rueda y Collado-Vides, 2000), de las cuales las mejor caracterizadas son LacI, AraC, LysR, CRP y OmpR.

Los factores de transcripción responden a diferentes estímulos ambientales o señales internas que impliquen cambios a través del control de la expresión de diferentes genes. Además, ellos mismos han de ser regulados (ya sea en su actividad o en su expresión) y existen diferentes mecanismos para ello. En primer lugar, existen pequeños ligandos de bajo peso molecular que pueden modificar la afinidad de los factores de transcripción por el ADN y la concentración de estas moléculas puede variar en función de la disponibilidad de nutrientes o de situaciones de estrés (Browning y Busby, 2004). Es importante recordar que uno de los dominios reguladores de los factores de transcripción se ha definido como dominio de unión a pequeñas moléculas, y se encuentra presente en el 44 % de los factores de transcripción (Babu y Teichmann, 2003). En segundo lugar, los factores de transcripción pueden ver modificada su actividad mediante modificación covalente a través de sensores quinasa que generalmente se encuentran en la membrana citoplasmática. En tercer lugar, la concentración de los factores de transcripción en la célula puede controlar su propia actividad a través de la regulación de su expresión o por

proteólisis. Finalmente, otro mecanismo menos común es el secuestro del factor de transcripción por otra proteína reguladora.

Los factores de transcripción pueden 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).

Algunos factores de transcripción pueden ser considerados reguladores globales. Tradicionalmente se conoce como reguladores globales aquellos que causan un fenotipo pleiotrópico y que son capaces de regular operones que pertenecen a diferentes vías (Gottesman, 1984). Existen diferentes criterios para considerar si un factor de transcripción es o no un regulador global. Para Babu y Teichmann (2003), es necesario que regulen más de 15 genes (ya sea de forma directa o indirecta), mientras que para Shen-Orr *et al.* (2002), han de regular 10 o más operones. Martínez-Antonio y Collado-Vides (2003) han definido una serie de características para decidir si un factor de transcripción puede ser considerado un regulador global:

- Regulan un elevado número de genes. El 51 % de los genes regulados en *E. coli* están controlados únicamente por siete proteínas reguladoras que, por tanto, controlan decenas o centenas de genes (CRP, FNR, IHF, FIS, ArcA, NarL y LrP). En cambio, muchos factores de transcripción sólo controlan un promotor individual o un número limitado de genes.

- Actúan conjuntamente con otros factores de transcripción. El 49 % de los genes se encuentran regulados por múltiples factores, interaccionando juntos reguladores globales con reguladores específicos.

- Pueden regular a otros factores transcripcionales y de esta forma regular diferentes genes de forma indirecta. Por ejemplo, CRP regula 23 factores de transcripción incluyéndose a sí mismo.

- Un factor de transcripción global puede regular genes que son transcritos por diferentes factores σ y, una determinada subunidad σ puede transcribir genes regulados por diferentes factores de transcripción. De esta forma, la célula tiene varias alternativas para determinar que gen o genes van a ser regulados o transcritos frente a un determinado estímulo.

- Los factores de transcripción globales responden a un mayor número de señales ambientales que los reguladores específicos.

- La capacidad de autoregulación es una característica común en los factores de transcripción, pero es predominante en los factores de transcripción globales (Thieffry *et al*, 1998).

- Los factores transcripcionales globales se transcriben desacoplados de los genes que regulan a diferencia de lo que sucede con los factores específicos. Esta característica es una consecuencia de la capacidad de regular muchos genes.

- La mayoría de los factores de transcripción globales pertenecen a familias que contienen un número de genes bajo. Por ejemplo, CRP, IHF, FNR, Lrp y H-NS están clasificados en familias con 3 o menos miembros dentro de una misma especie.

Los factores de transcripción globales se pueden clasificar en diferentes niveles jerárquicos según estas características que se han definido y en función del grado de interconexiones transversales que puedan presentar. Además, en las redes reguladoras también intervienen otros niveles de regulación, como la metilación Dam (Oshima *et al*, 2002) o el grado de superenrollamiento en relación con el nivel energético de la célula como se comentará posteriormente en esta introducción. Todo este escenario supone un complejo modelo de regulación que permite a la célula establecer el patrón más óptimo de expresión.

1.1.2.3.- Superenrollamiento del ADN: proteínas asociadas al nucleoide

Las moléculas de ADN tienen una extraordinaria longitud comparadas con el tamaño de las células bacterianas. Por tanto, estos organismos necesitan diferentes mecanismos de compactación y empaquetamiento de su información genética, formando una estructura denominada nucleoide bacteriano. Esta estructura compleja permite, además de la compactación del ADN, que la célula pueda replicarse y regular la expresión génica como respuesta a los diferentes cambios en las condiciones ambientales y nutricionales.

El nucleoide puede estar organizado en dominios de tal forma que cambios en uno de los dominios no afecten a la totalidad del cromosoma. En general, el ADN extraído de células procariotas tiene un superenrollamiento negativo y este superenrollamiento está fuertemente controlado por varios factores combinados, entre ellos, la replicación, la

transcripción, las proteínas de unión al ADN y la actividad de las topoisomerasas (Drlica, 1992).

Principalmente el superenrollamiento bacteriano es mantenido por la actividad coordinada de dos enzimas: la ADN girasa y la ADN topoisomerasa I. La topoisomerasa I provoca una disminución del superenrollamiento del ADN, modificando el índice de enlace de la doble hélice en una reacción independiente de ATP. En cambio, la ADN girasa provoca un aumento del grado de superenrollamiento negativo. En este caso la reacción es dependiente de la hidrólisis de ATP y responde a la relación $[ATP]/[ADP]$ intracelular, que es indicativa de la carga energética de la célula. Hsieh *et al.*, (1991, a, b) han descrito que esta actividad ADN girasa aumentaría en condiciones ambientales de baja presión parcial de oxígeno o de aumento de osmolaridad del medio, aumentando de este modo el superenrollamiento negativo.

Las diferentes condiciones ambientales y nutricionales pueden provocar un cambio en la carga energética de la célula y tal y como se acaba de comentar en relación con la ADN girasa, esta carga energética puede afectar al grado de superenrollamiento. Por esta razón, algunos autores han propuesto que el nivel de superenrollamiento del ADN sería un elemento muy importante que afecta a la regulación (Hatfeild y Benham, 2002; Schaechter, 2001). Un ejemplo sobre la importancia del superenrollamiento en la expresión génica es el de los operones del regulón *ilv* de *E. coli*. Estos genes codifican para los enzimas implicados en la biosíntesis de aminoácidos de cadenas ramificadas (L-valina, L-Leucina y L-isoleucina). Un incremento en el nivel de superenrollamiento se traduce en un incremento en la expresión del regulón (Hatfeild y Benham, 2002).

Como se ha comentado anteriormente, para la compactación del ADN cromosómico son muy importantes las proteínas de unión al ADN o proteínas asociadas al nucleoide. La unión de estas proteínas al ADN puede modificar el nivel de superenrollamiento del mismo. Tienen cierta similitud con las histonas eucariotas, pero no en su secuencia de aminoácidos, sino en que son proteínas de bajo peso molecular, de bajo número de copias, por su carga electrostática y su capacidad para unirse al ADN, aunque a diferencia de las histonas eucariotas forman uniones inestables con el ADN.

Las proteínas asociadas al nucleoide, tal y como se ha explicado en el apartado 1.1.2.2., están consideradas como reguladores globales de la transcripción, ya que aunque no se conoce el papel exacto de cada una de estas proteínas en la compactación del genoma bacteriano, pueden afectar profundamente a la estructura global o local del cromosoma y

de esta forma a la expresión génica. También están implicadas en otros procesos del ADN como la recombinación, la replicación, la reparación o la transposición.

En *E. coli* se han descrito al menos 12 proteínas asociadas al nucleoide (NAP; Nucleoid Associated Proteins) (revisado en Ali Azam y Ishihama, 1999). Algunas de las proteínas asociadas al nucleoide más estudiadas son HU (Heat-Unstable Nucleoid Protein), IHF (Integration Host Factor), H-NS (Histone-like Nucleoid-Structuring protein), StpA (Supressor of *td*-mutant phenotype A), FIS (Factor for Inversion Stimulation), y Dps (DNA-binding protein from starved cells). La concentración de estas proteínas varía en función de las condiciones de crecimiento celular y de la fase de crecimiento en la que se encuentre la célula. Fis, HU, IHF y H-NS son las proteínas más abundantes en fase exponencial, mientras que Dps, IHF y HU lo son durante la fase estacionaria (Ali Azam *et al.*, 1999). Estos cambios en la composición del “pool” de NAP influiría en los cambios que se producen en las diferentes fases de crecimiento, expresando los genes adecuados en cada momento. Mutaciones en estas proteínas pueden resultar muy perjudiciales para la célula (incluso letales) ya que directa, indirectamente o en combinación, afectan a genes que son fundamentales para la viabilidad como lo son los implicados en la síntesis proteica y en el metabolismo. También intervienen en la regulación de genes que están implicados en la respuesta a factores ambientales como la osmolaridad o la temperatura y, por tanto, también en genes implicados en la virulencia (McLeod y Johnson, 2001).

Las proteínas asociadas al nucleoide no suelen presentar una elevada especificidad de unión al ADN. Alguna de ellas tiene preferencia por una secuencia o estructura. Por ejemplo H-NS o StpA se unen preferencialmente a regiones curvadas del ADN (Rimsky *et al.*, 2001).

A continuación se refieren algunas características de las NAP mejor estudiadas:

- **HU**: HU es una proteína muy abundante (15.000-30.000 dímeros por célula en fase exponencial). Es un heterodímero formado por dos subunidades intercambiables α y β de 9'2 y 9'5 kDa respectivamente, muy similares entre ellas (70%) en *E. coli*, *S. enterica* serovar *typhimurium*, *Erwinia chrysanthemi* y *Shigella flexneri*. En otras especies, se forma un homodímero (Oberto y Rouvière-Yaniv, 1996). *E. coli* puede modificar la composición de estas subunidades en función de la fase de crecimiento. Estos cambios parecen influir en su capacidad para organizar la cromatina bacteriana y de esta forma modificar la expresión génica (revisado en Dorman *et al.*, 1999). Existen bastantes evidencias de la unión de la proteína HU a las regiones promotoras de los genes, preferentemente en zonas curvadas o

cruciformes, pero los detalles moleculares no se conocen excepto para la regulación del operón *gal* (Kar y Adhya, 2001). Además, se ha propuesto que las proteínas HU y H-NS tienen papeles estructurales antagónicos y, por tanto, también diferentes contribuciones a la regulación génica (Dame y Goosen, 2002).

- **FIS**: FIS es una proteína básica formada por dos subunidades idénticas codificadas por el gen *fis* y está altamente conservada en bacterias entéricas. Es una de las proteínas asociadas al nucleóide más abundante en *E. coli* (20.000-40.000 copias por célula). Inicialmente la proteína FIS fue descubierta como cofactor de la recombinación de lugar específico, pero cada vez está más claro que juega un papel importante en la regulación de la expresión génica y en la modulación del superenrollamiento local (revisado por Dorman y Deighan, 2003). De esta forma, FIS regula un número muy elevado de genes, tanto de forma directa (inicio de la transcripción) como indirecta (Schneider *et al.*, 1999; Finkel y Johnson, 1992; Xu y Johnson, 1995; González-Gil *et al.*, 1996). Resultados recientes ponen de manifiesto la importancia que la proteína FIS está adquiriendo como regulador de genes de virulencia en patógenos bacterianos: promueve la transcripción de genes de invasión en *S. flexneri* (Falconi, *et al.*, 2001) y contribuye en la regulación de genes de virulencia en *S. enterica* serovar *typhimurium* (Wilson *et al.*, 2001) y cepas enteropatógenicas de *E. coli* (Goldberg *et al.*, 2001).

- **IHF**: La proteína IHF está implicada en la integración y escisión de la integrasa del fago 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 (25.000-30.000 dímeros por célula). IHF se une al ADN existiendo secuencias consenso de unión (Ali *et al.*, 2001; Wagner, 2000). Afecta directa o indirectamente a 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 *S. flexneri* (Porter y Dorman, 1997) y *S. enterica* serovar *typhimurium* (Marshall *et al.*, 1999).

Debido a la importancia de la proteína H-NS en este trabajo, se ha dedicado un capítulo de esta introducción a dicha proteína.

1.2.- LA FAMILIA DE PROTEÍNAS H-NS

1.2.1.- LA PROTEÍNA H-NS

1.2.1.1.- Características generales

La proteína H-NS (Histone-like Nucleoid Structuring protein) fue descrita por primera vez como el mayor componente del nucleoide bacteriano de *E. coli* (Varshavsky *et al.*, 1977). El gen que codifica para la proteína H-NS, denominado *hns*, se encuentra localizado en el minuto 27'8 del cromosoma de *E. coli* y es de copia única, aunque presenta un gen parálogo, el gen *stpA*.

Diferentes grupos de investigación han descrito mutaciones en el gen *hns* (que también ha recibido el nombre de H1, *osmZ*, *virR*, *drdX*, *bglY*, *pilG*) en diferentes especies de bacterias Gram-negativas (no se ha descrito en bacterias Gram-positivas). Estas mutaciones fueron inicialmente descritas debido a la alteración que producen en la expresión de diferentes genes que están regulados por temperatura y osmolaridad. Las mutaciones en el gen *hns* son altamente pleiotrópicas, afectando a procesos de recombinación ilegítimos, a frecuencias de transposición, a deleciones cromosómicas, etc. (Lejeune y Danchin, 1990; Falconi *et al.*, 1991). También interviene en diferentes procesos celulares como la replicación y el ciclo celular.

La proteína H-NS es considerada como un regulador global de la transcripción, ya que controla la expresión de aproximadamente el 5 % de los genes en *E. coli* (Hommais *et al.*, 2001) y está implicada en la regulación de la expresión de factores de virulencia de cepas uropatógenicas (Göransson *et al.*, 1990; Madrid *et al.*, 2002a) y enterotoxigénicas (Trachman y Maas, 1998) de *E. coli*, en *S. flexneri* (Maurelli y Sansonetti, 1998), en *S. enterica* serovar *typhimurium* (Higgins *et al.*, 1988) y *Vibrio cholerae* (Nye *et al.*, 2000; Tendeng *et al.*, 2000). Sin embargo en otras especies bacterianas no ha podido caracterizarse el papel regulador de H-NS, ya que no ha sido posible obtener cepas mutantes para el gen *hns* como, por ejemplo, en *Proteus mirabilis* (Coker *et al.*, 2000), *Bordetella bronchiseptica* (Goyard y Bertin, 1997) o *Y. pseudotuberculosis* (Heroven *et al.*, 2004).

Las proteínas H-NS identificadas hasta ahora forman una familia de proteínas (Bertin *et al.*, 1999; Bertin *et al.*, 2001) y están ampliamente distribuidas entre las bacterias Gram-negativas (“Phylum” ‘Proteobacteria’) (revisado Tendeng y Bertin, 2003). Dentro de esta familia, las proteínas H-NS de *E. coli* y *S. enterica* serovar *typhimurium* han sido las mejor caracterizadas. H-NS es una proteína de bajo peso molecular (aproximadamente 15,6 kDa y 137 aminoácidos) y muy abundante en el nucleóide bacteriano, ya que presenta aproximadamente 20.000 copias por célula. A diferencia de la mayor parte de proteínas asociadas al nucleóide y de las histonas eucariotas, la proteína H-NS presenta un carácter global neutro, a pesar de la presencia de algunas zonas con aminoácidos cargados (revisado por Rimsky, 2004). Se han detectado tres isoformas de la proteína con diferentes puntos isoeléctricos 5’1, 6’5 y 7’5, siendo las dos últimas las más predominantes (Spassky *et al.*, 1984). La cantidad relativa de las diferentes isoformas presentes en la célula depende de la fase de crecimiento (Spassky *et al.*, 1984). Seguramente este hecho es debido a modificaciones post-traduccionales, aunque no ha sido definitivamente confirmada la naturaleza biológica de las mismas (Atlung y Ingmer, 1997). Se ha sugerido una posible modificación post-traducciona de H-NS con poli-(R)-3-hidroxi-butirato de cadena corta (Reusch *et al.*, 2002).

Los niveles de la proteína H-NS dentro de la célula son muy importantes para su función. La expresión de H-NS se mantiene relativamente constante a lo largo de la fase de crecimiento, pero se ha comprobado que experimenta un ligero incremento en el inicio de la fase estacionaria. En relación con la temperatura, H-NS se expresa de igual forma a 37 °C y a 26 °C, pero afecta a la expresión de los genes que controla de diferente forma (Göransson *et al.*, 1990).

El gen *hns* se transcribe independientemente de los genes que lo flanquean (como ya se ha explicado anteriormente en esta introducción, esta es una característica típica de los reguladores globales como consecuencia de la regulación de muchos genes) y es capaz de autorregular su propia expresión inhibiéndola a nivel de transcripción (Hulton *et al.*, 1990; Dersch *et al.*, 1993; Ueguchi *et al.*, 1993; Falconi *et al.*, 1993). Las proteínas asociadas al nucleóide FIS y Csp pueden activar la expresión de *hns* (Falconi *et al.*, 1996). Además, H-NS y su proteína paróloga StpA son capaces de activar la transcripción de FIS (Johansson *et al.*, 2000).

La proteína H-NS realiza su función reguladora a través de su unión a los ácidos nucleicos, tanto ADN como ARN (Friedrich *et al.*, 1988). De todas formas, H-NS se une preferencialmente a regiones curvadas de ADN de doble cadena y ricas en secuencias A-T (Bracco *et al.*, 1989, Yamada *et al.*, 1990; Owen-Hughes *et al.*, 1992) e incluso puede provocar la curvatura de la región a la cual se une (Spurio *et al.*, 1997). Es decir, reconoce estructuras locales más que secuencias específicas. Existe un gran número de regiones en el genoma de *E. coli* que presentan dichas zonas curvadas (Bracco *et al.*, 1989), pero además en algunos casos como el del gen *proU*, estas regiones se encuentran tras el promotor (Owen-Hughes *et al.*, 1992) y se denominan DRE (Downstream Regulatory Element).

1.2.1.2.- Dominios estructurales de la proteína H-NS

La proteína H-NS presenta dos dominios estructurales independientes que se corresponden con sus dos dominios funcionales. Por una parte presenta un dominio N-terminal en el que recae la capacidad de oligomerizar de la proteína H-NS, y por otra parte un dominio C-terminal responsable de la capacidad de dicha proteína para unirse al ADN. Ambos dominios estructurales se encuentran unidos a través de una pequeña región flexible que permite que ambos puedan actuar de manera independiente (Shindo *et al.*, 1995; Dorman *et al.*, 1999; Smyth *et al.*, 2000; Esposito *et al.*, 2002; Bloch *et al.*, 2003). En la figura 1.2.1. se muestra el esquema de la proteína H-NS.



Figura 1.2.1. Representación esquemática de los dominios funcionales de la proteína H-NS. Se muestra el dominio N-terminal de oligomerización, dividido en el dominio de dimerización (residuos aminoácidos del 1-64) y la región flexible (residuos aminoácidos del 56-89) y finalmente el dominio C-terminal de unión al ADN. Extraído de Dorman, 2004.

1.2.1.2.1.- Dominio N-terminal: dominio de oligomerización

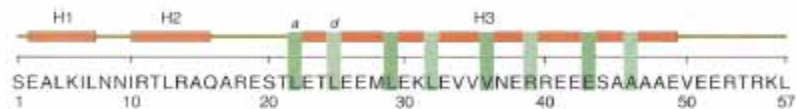
El dominio de oligomerización de la proteína H-NS se encuentra como se ha comentado anteriormente en el extremo N-terminal de la proteína y se extiende hasta el residuo aminoacídico 64 (Smyth *et al.*, 2000).

La proteína H-NS en general y el dominio de oligomerización en particular, se encuentran altamente conservados entre los diferentes miembros de la familia de proteínas H-NS que existen en las diferentes especies bacterianas. Este hecho facilita las interacciones heteroméricas entre las diferentes proteínas (Dorman *et al.*, 1999; Rimsky, 2004).

Para poder realizar sus funciones biológicas, la proteína H-NS requiere de la formación de estructuras de elevado orden de oligomerización (Spurio *et al.*, 1997; Esposito *et al.*, 2002). Si bien los primeros 64 aminoácidos de la proteína son los que constituyen el dominio de oligomerización, recientes estudios de resonancia magnética nuclear (RMN) muestran que el mínimo dominio de dimerización lo constituyen únicamente los primeros 46 aminoácidos (Bloch *et al.*, 2003). Otros estudios (Smyth *et al.*, 2000) han demostrado que para la formación de unidades de oligomerización mayores al dímero, además del propio dominio de oligomerización, es necesaria (al menos en parte) la presencia de la región conectora (aminoácidos del 65 al 89), ya que proteínas truncadas de H-NS (H-NS₁₋₆₄) no son capaces de formar oligómeros mayores.

La estructura “coiled-coil” ha sido identificada como elemento de oligomerización en muchos factores de transcripción (Lupas, 1996) y podría funcionar como sensor de temperatura. La proteína H-NS también posee esta estructura. El dominio de oligomerización de dicha proteína presenta 3 hélices α . Las terceras hélices α (H3, residuos 23-50) de cada monómero se unen entre ellas formando el núcleo de la estructura que se encuentra estabilizada a través de los residuos de las hélices H1 (3-8) y H2 (12-18). Para cada monómero, las hélices H2, H3 y la región conectora formarían una estructura en forma de U en la que la hélice H1 se sitúa en perpendicular (Bloch *et al.*, 2003). Los residuos hidrofóbicos de las posiciones claves se unen en el interior (Figura 1.2.2.). Además esta región rica en hélices α es similar a proteínas eucariotas con estructura “coiled-coil” como la miosina o la distrofina (Atlung y Ingmer, 1997).

A)



B)

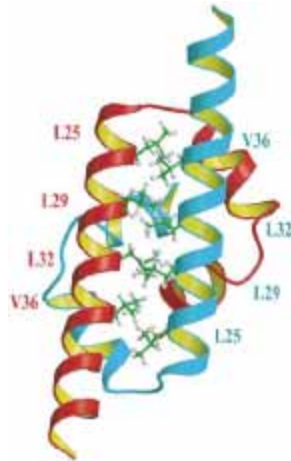


Figura 1.2.2. Esquema del dominio de oligomerización de la proteína H-NS.

- A) Secuencia de aminoácidos del dominio de oligomerización (N-terminal) de H-NS. También se muestra la situación de las 3 hélices α presentes en el dominio. Extraída de Esposito *et al.*, 2002.
- B) Estructura tridimensional del dominio N-terminal (de oligomerización) de un dímero de H-NS. Los dos protómeros se muestran en rojo y azul respectivamente. Se detallan los aminoácidos envueltos en la interacción y su localización. Extraída de Rimsky, 2004.

Para la oligomerización, en primer lugar se formaría la unidad estructural (el dímero) que es dependiente de la concentración, la temperatura y los tipos de cationes del medio (Smyth *et al.*, 2000). A continuación se forman los oligómeros, en los que el número de unidades que lo forman es proporcional a la concentración, facilitando de este modo la compactación del ADN. El estado de oligomerización que presenta H-NS en la célula es incierto y existe bastante controversia, aunque parece que se encontraría formando predominantemente dímeros o estructuras oligoméricas siempre de números pares (Ueguchi *et al.*, 1996; Smyth *et al.*, 2000).

Recientemente Esposito *et al.*, 2002 han definido un modelo por el cual se formarían los oligómeros de la proteína a través de la unión de diferentes dímeros (homodímeros) mediante una estructura de cabeza-cola (“head-to-tail”). En la figura 1.2.3. se puede observar un esquema sobre el modelo. La hélice H3 de uno de los homodímeros

(región C-terminal del dominio de oligomerización) interacciona con la hélice H2 de otro homodímero (región N-terminal del dominio de oligomerización). La integridad de la estructura y la orientación de la hélice H2 también es necesaria para la formación de los oligómeros. Mutaciones en la hélice H2 implican un defecto de la proteína H-NS para producir oligómeros, sugiriendo que se produce una interacción entre este elemento y la región conectora.

Los primeros 20 aminoácidos de dicho dominio son responsables de la represión/silenciamiento génico (mutaciones en estos aminoácidos evitan la formación de oligómeros de elevado orden). Además esta zona participaría también junto con el dominio de C-terminal en el reconocimiento de las regiones curvadas y por tanto de la unión al ADN (Dorman, 2004). Trabajos de Bloch *et al.*, 2003 se muestran de acuerdo con esta hipótesis.

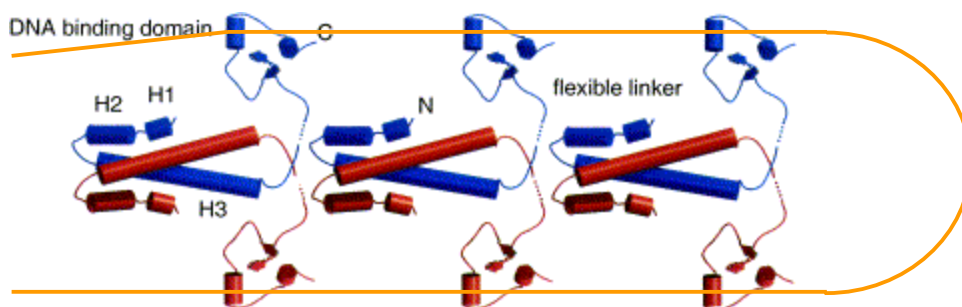


Figura 1.2.3. Representación esquemática del modelo cabeza-cola que tiene lugar para la asociación de dímeros de H-NS y la formación de oligómeros. Se muestra que residuos del extremo C-terminal del dominio de oligomerización interaccionan con residuos del extremo N-terminal del mismo dominio del siguiente homodímero formando de esta forma un filamento proteico (también se incluyen las 3 hélices α del dominio). También se muestra la región flexible que conecta con el dominio de unión al ADN. Extraído de Esposito *et al.*, 2002.

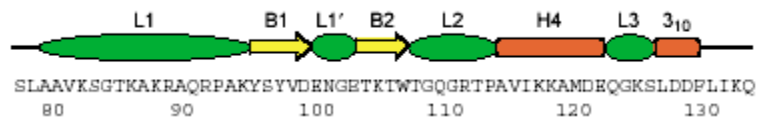
1.2.1.2.2.- Dominio C-terminal: dominio de unión al ADN

El dominio de unión al ADN de H-NS se encuentra situado en el dominio C-terminal de la proteína e incluiría del aminoácido 90 al 137 (Shindo *et al.*, 1995).

Al hablar de las características generales de la proteína H-NS ya se ha comentado que dicha proteína no reconoce una secuencia consenso de unión al ADN, no obstante se une preferencialmente a regiones del ADN curvadas y ricas en secuencias A-T (Yamada *et al.*, 1990; Owen-Hughes *et al.*, 1992; Ueguchi *et al.*, 1993). Además los promotores regulados por H-NS siempre presentan más de un lugar de unión para la proteína (Rimsky, 2004) y la situación espacial de los diferentes lugares de unión es muy importante para que H-NS actúe sobre la conformación del ADN.

La estructura del dominio C-terminal de la proteína H-NS fue identificada por primera vez por Shindo *et al.*, 1995 (Figura 1.2.4.B). Se cree que las regiones que estarían directamente implicadas en la unión con el ADN serían las que forman el bucle 2 y que se encuentran entre unas estructuras secundarias rígidas (hélice α y lámina β). Al observar dicha estructura tridimensional se puede apreciar una protuberancia (que comprende del aminoácido 108 al 116) precisamente en la región del bucle 2 que apoyaría la idea de que esta región es la implicada en la interacción con el ADN (Ueguchi *et al.*, 1996). Además el motivo TWTGXGRXP (donde X puede ser cualquier aminoácido) presente en esta región se encuentra altamente conservado en la familia de proteínas H-NS (Dorman *et al.*, 1999). En la revisión de Dorman, 2004 se indica que el aminoácido 109 (Triptófano) estaría íntimamente relacionado con la interacción con el ADN (Tippner y Wagner, 1995) y el aminoácido 116 (Prolina) está relacionado con el reconocimiento de regiones curvadas o no curvadas y producir curvaturas en el ADN (fuertemente dependiente del estado de oligomerización). También se ha observado que después de la asociación con el ADN se pueden inducir estructuras secundarias en el bucle 1 (Shindo *et al.*, 1995).

A)



B)

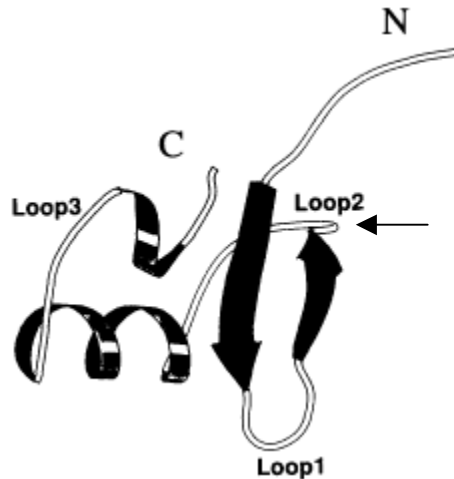


Figura 1.2.4. Esquema del dominio de unión al ADN de la proteína H-NS.

A) Secuencia de aminoácidos del dominio de unión al ADN (C-terminal) de H-NS. También se muestra la situación de los diferentes bucles, las láminas β y la hélice α presentes en el dominio. Extraído de Tendeng y Bertin, 2003.

B) Estructura tridimensional del dominio C-terminal (de unión al ADN) de la proteína H-NS que comprende del residuo 90 al 137 obtenido por resonancia magnética nuclear. Se muestran los bucles existentes en la estructura y la flecha sobre el bucle 2 indica la posición de los residuos T108-P116. Las flechas negras muestran la situación de las láminas β presentes y las regiones negras y curvadas señalan la situación de las hélices α . Extraído de Shindo *et al.*, 1995.

La presencia de la región conectora flexible (rica en Arg y Lys) en la proteína H-NS permite que el dominio C-terminal exista como un dominio independiente y que pueda rotar libremente (Smyth *et al.*, 2000). Aunque el dominio no presenta ninguna función en la oligomerización de la proteína, la unión de la proteína al ADN no es suficiente para producir el silenciamiento o represión transcripcional, sino que también son importantes los primeros 20 aminoácidos del dominio N-terminal como ya se ha explicado anteriormente (revisión Dorman, 2004).

El hecho de que el dominio C-terminal tenga una moderada afinidad por el ADN (Shindo *et al.*, 1995) sugiere que es importante la cooperación de la proteína y por tanto la formación de oligómeros de elevado orden para mejorar la afinidad. Por otra parte la interacción con el ADN se ve favorecida por la estructura “coiled-coil” del dominio de oligomerización que expone cargas negativas que tienen tendencia a oponerse al ADN (Esposito *et al.*, 2002).

Algunos modelos de interacción sugieren que se formaría un filamento de tipo bastón donde los residuos responsables de la unión al ADN quedarían expuestos para facilitar su unión al mismo. De esta forma H-NS podría unir dos hebras de ADN y actuar como cremallera formando una estructura en el ADN curvada (Figura 1.2.3.) (Dame *et al.*, 2002).

1.2.1.3.- Modulación de la expresión génica por H-NS

Tal y como se ha comentado anteriormente, H-NS es considerado como un modulador global de la expresión génica en respuesta a cambios ambientales que afecta a la transcripción, ya sea de forma directa o indirecta, de un 5 % de los genes de *E. coli* (Hommais *et al.*, 2001).

Diversos autores han propuesto dos mecanismos moleculares a través de los cuales se podría producir la regulación de la transcripción por parte de la proteína H-NS. Estos dos mecanismos se conocen como silenciamiento transcripcional (Göransson *et al.*, 1990) y represión vía topología del ADN (Hulton *et al.*, 1990; Tupper, *et al.*, 1994). Los dos modelos no son necesariamente excluyentes.

Ambos modelos tienen una característica en común: la proteína H-NS se une a unas determinadas regiones o lugares, normalmente en la vecindad de determinados promotores. Existe más de un lugar de unión para H-NS en los diferentes promotores tanto antes como después de los mismos. A continuación se produce un recubrimiento a lo largo de la secuencia del ADN a través de la nucleación y polimerización de moléculas de H-NS (Williams y Rimsky, 1997). Este recubrimiento es el responsable del control de la actividad del promotor.

En cambio, los dos modelos difieren en como el complejo H-NS/ADN puede afectar a la actividad de la ARN polimerasa. En el caso del modelo denominado silenciamiento transcripcional H-NS actuaría directamente como represor transcripcional

debido a su capacidad para polimerizar linealmente sobre el ADN e impidiendo la correcta unión de la polimerasa, ya sea por captura o por oclusión del lugar de unión (Williams y Rimsky, 1997). Esta represión de la transcripción por parte de H-NS no se produce de forma indefinida, sino que determinadas señales ambientales (como la temperatura o la osmolaridad) o diferentes factores de transcripción (como la proteína Fis) pueden contribuir a la disrupción de los complejos represores (Dorman y Deighan, 2003). En la figura 1.2.5. se muestra un esquema del posible modelo de silenciamiento transcripcional. Además este modelo es el propuesto para la regulación del gen *virF* en *S. flexneri* (Falconi *et al.*, 1998) y para el operón hemolítico de *E. coli* (Madrid *et al.*, 2002a).

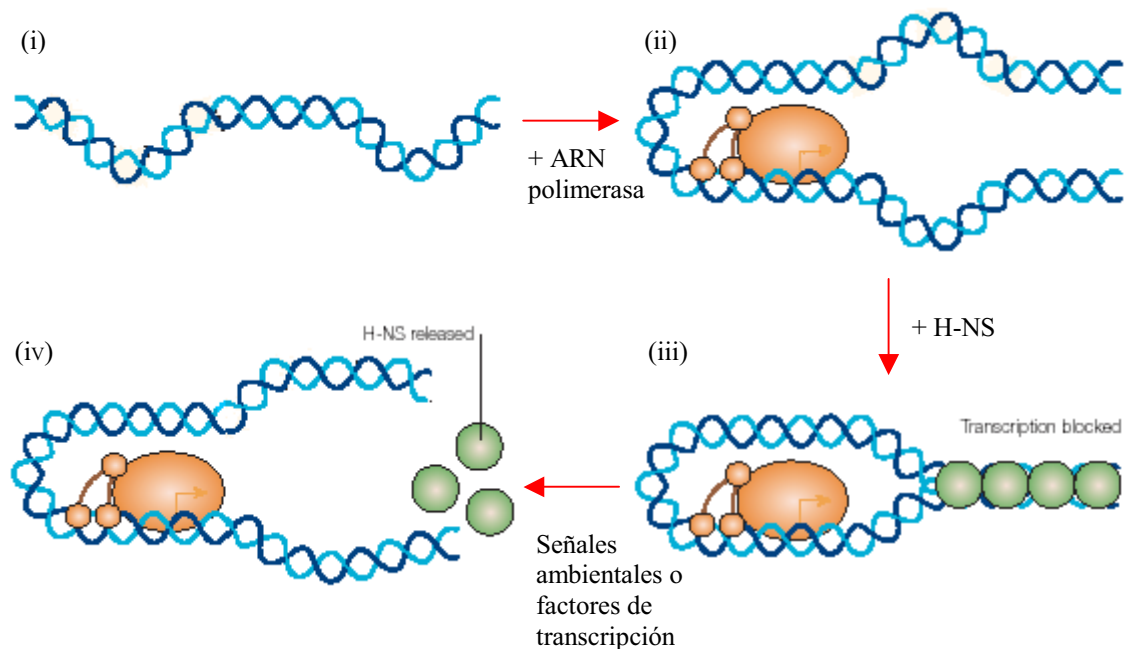


Figura 1.2.5. Represión de la transcripción por H-NS. Figura extraída de Dorman, 2004.

(i) La región de unión en el promotor de la ARN polimerasa se encuentra flanqueada por curvaturas intrínsecas. (ii) El ADN se enrolla alrededor de la ARN polimerasa permitiendo la interacción entre zonas curvadas próximas (anterior y posterior al promotor). (iii) H-NS se une a las regiones curvadas dando lugar a un complejo nucleoproteico que atrapa a la ARN polimerasa en su interior en el estadio inicial de la transcripción, reprimiendo el promotor. (iv) El complejo nucleoproteico puede deshacerse a través de determinadas señales ambientales o de la unión de algún factor de transcripción.

El modelo de represión vía topología del ADN viene de la observación que H-NS puede actuar directamente sobre la topología del ADN, produciendo cambios conformacionales a nivel local. Las alteraciones topológicas modificarían la interacción con el promotor de otras proteínas, factores de transcripción y sobre todo de la ARN polimerasa (Hulton *et al.*, 1990). Está descrito que muchos promotores son sensibles al superenrollamiento del ADN y que una variación del estado de superenrollamiento puede modificar la actividad de un promotor (Tupper *et al.*, 1994). En este caso, la formación del complejo nucleoproteico H-NS/ADN podría modificar la topología local, alterando la capacidad de la ARN polimerasa de formar el complejo de transcripción (Spassky *et al.*, 1984; Schröder y Wagner, 2000; Rimsky *et al.*, 2001).

De todos los genes que se encuentran directa o indirectamente afectados por la proteína H-NS, un 35 % están regulados por osmolaridad, temperatura, pH o disponibilidad de oxígeno (Hommais *et al.*, 2001), es decir, están involucrados en la adaptación de la bacteria a los cambios ambientales o forman parte de la envoltura celular. Un 20 % están implicados en la fase o tasa de crecimiento y codifican para proteínas implicadas en el aparato transcripcional o traduccional. Por otra parte no se conoce la función que realizan un 34 % de los genes afectados por la proteína H-NS, indicando que todavía restan muchos aspectos de la regulación por parte de dicha proteína que no se conocen.

Muchas bacterias patógenas, cuando aumenta la temperatura, responden con la expresión de factores de virulencia. Una propiedad de H-NS, como se ha comentado anteriormente, es su capacidad para regular genes en respuesta a diferentes temperaturas. Este es el caso de los genes del operón hemolítico o del promotor *virF*.

En general, H-NS está considerada como un regulador negativo de la transcripción, pero se ha mostrado que influye positivamente sobre la motilidad bacteriana. Este hecho no es necesariamente un reflejo de un papel directo de esta proteína como activador transcripcional. Por una parte, H-NS es represor del gen *hdfR* que codifica para una proteína tipo LysR que regula negativamente el promotor del operón flagelar *flhDC*, es decir, en ausencia de H-NS aumenta el producto de *hdfR* y por tanto se reduce la expresión del operón flagelar. Además, el promotor de *flhDC* también está sujeto a una represión directa por parte de H-NS. Por otra parte parece que H-NS es capaz de interactuar directamente con la proteína motor flagelar, FliG, incrementando su función (revisado por Dorman, 2004).

Algunos de los ejemplos clásicos de genes regulados por H-NS corresponden al gen *virF* (Cornelis *et al.*, 1991), al operón *proa* (Higgins *et al.*, 1988) y al operón *bgl* (Mahadevan *et al.*, 1987; Schnetz, 1995).

El operón críptico *blg* codifica para un sistema de utilización de β -glucósidos aromáticos (aryl- β ,D-glucoside) de *E. coli.*, implicado en la captación de salicina y arbutina (Schaefer, 1967). El operón *bgl* codifica para un regulador transcripcional (BglG), para una permeasa específica de β -glucósidos (dependiente de fosfoenolpiruvato, sistema PTS) denominada EII^{Bgl} (o BglF) y para una fosfo- β ,D-glucosidasa (BglB) (Mahadevan *et al.*, 1987; Schnetz *et al.*, 1987). La regulación específica de dicho operón se lleva a cabo a través de las proteínas BglG y BglF (Mahadevan *et al.*, 1987). Cuando los β -glucósidos están presentes, BglG es fosforilada por BglF, resultando en una inactivación que causa una prematura terminación de la transcripción y una disminución de la expresión de los genes situados a continuación (Amster-Choder *et al.*, 1989; Amster-Choder y Wright, 1990; Schnetz y Rak, 1988). En la represión del operón *bgl* es importante el efecto de la represión por catabolito, es decir, por parte de la glucosa (Gulati y Mahadevan, 2000). La proteína H-NS se encuentra silenciando el operón *bgl* (Higgins *et al.*, 1988), pero H-NS sola no es suficiente para la represión del operón, sino que necesita elementos del ADN que se encuentran localizados antes y después del promotor (Schnetz, 1995). La proteína H-NS reprime el inicio de la transcripción del promotor *bgl* uniéndose a lugares anteriores al promotor en unas 3 veces (Dole *et al.*, 2004a). Además la unión de H-NS a regiones situadas tras el promotor (600-700 pb tras el inicio de transcripción), dentro de la región codificante del gen *bglG*, reprime unas 7 veces la expresión y es dependiente del factor Rho (Dole *et al.*, 2004b). Otros factores pleiotrópicos actúan modulando la acción de H-NS como son la proteína Hfq y la proteasa Lon (Dole *et al.*, 2004a).

Existen otros factores celulares que también ayudan a mantener el promotor del operón *bgl* inaccesible para la ARN polimerasa y la proteína CAP (Catabolite Activator Protein), como la proteína RpoS (Schnetz, 2002).

Una vez activado, el operón es inducible por salicina y arbutina y la transcripción regulada por anti-terminación, con la participación de una proteína de unión a ARNm modulada por fosforilación (Houman *et al.*, 1990). Diversas mutaciones espontáneas desreprimen el silenciamiento del promotor *bgl*, como en la ADN girasa, en el receptor AMPc y en H-NS, probablemente igual que esta última afectando a la formación del complejo nucleoproteico (Mukerji y Mahadevan, 1997).

1.2.2.- LA PROTEÍNA StpA: PARÁLOGA DE H-NS

La proteína StpA (Supresor of *td* mutant Phenotype A) fue identificada por Zhang y Belfort en 1992. Inicialmente fue aislada como supresor multicopia del fenotipo Td⁻ de los bacteriófagos T4 defectivos en el procesamiento de intrones. Posteriormente StpA fue también identificada por su capacidad, cuando se encuentra sobreexpresada, de complementar algunos fenotipos de un mutante *hns* (Shi y Bennett, 1994).

El gen *stpA* se localiza en el minuto 60'24 del mapa genético de *E. coli* (Berlyn *et al.*, 1996) y codifica para una proteína de 134 aminoácidos y 15'3 kDa (Zhang *et al.*, 1995). La proteína StpA es bastante más básica que la proteína H-NS ya que su pI teórico es de 9'08 mientras que el de H-NS es de 5'25. StpA presenta un 58 % de identidad y un 67 % de similitud a nivel de secuencia con H-NS. La identidad es mayor en el extremo C-terminal (73 % hasta el residuo 91) que en el extremo N-terminal (51 % hasta el residuo 90) (Cusick y Belfort, 1998). Las dos proteínas presentan una organización estructural similar.

Dada la elevada similitud entre ambas proteínas, no es de extrañar que presenten muchos paralelismos funcionales, aunque también presentan algunas características diferentes. Ambas proteínas se encuentran localizadas por todo el nucleoide (Ali Azam *et al.*, 1999), actúan como represores transcripcionales, son capaces de compactar el ADN, de unirse a regiones curvadas del mismo e inhibir la transcripción de un promotor (Zhang *et al.*, 1996). Ambas proteínas tienen actividad chaperona del ARN *in vitro*. De hecho la proteína StpA fue inicialmente descrita precisamente como chaperona, ya que facilita la correcta conformación de las moléculas de ARN (Zhang *et al.*, 1995). Esta actividad reside en el extremo C-terminal de la proteína (Cusick y Belfort, 1998) y es unas 10 veces más eficiente que la proteína H-NS (Zhang *et al.*, 1996).

Una diferencia entre ambas proteínas se encuentra en sus patrones de expresión. Mientras que la expresión de StpA es dependiente de la fase de crecimiento, los niveles de H-NS se mantienen más o menos constantes. Los niveles de StpA se mantienen bajos al inicio de la fase exponencial en medio rico pero aumentan un 50 % en fases posteriores (Free y Dorman, 1997).

La transcripción del gen *stpA* está muy condicionada por los cambios en los factores ambientales. El promotor de dicho gen es muy sensible a los cambios que se producen en la topología del ADN por dichos factores ambientales. StpA se encuentra fuertemente inducida en condiciones de estrés osmótico, aumento de temperatura (Free y Dorman, 1997) y crecimiento en medio mínimo. En este último caso, esta inducción es dependiente de la proteína Lrp (Leucine-Responsive Regulatory Protein), ya que la presencia de leucina en el medio inhibe su expresión (Sondén y Uhlin, 1996).

StpA y H-NS regulan su propia expresión y al mismo tiempo regulan la expresión de la otra, es decir, mantienen una regulación cruzada y simultánea que permite mantener la concentración adecuada de cada proteína (Zhang *et al.*, 1996). En mutantes *hns* la expresión de *stpA* se encuentra significativamente incrementada. En algunos casos, como en el operón *bgl*, la sobreexpresión de StpA puede compensar las funciones de H-NS (Free *et al.*, 1998). De todas formas, los niveles de StpA en mutante *hns* nunca llegan a los niveles de H-NS existentes en una cepa salvaje (Sonnenfield, 2001).

Los niveles de StpA también se encuentran regulados a través de su degradación por la proteasa Lon. La proteasa Lon tiene la capacidad de degradar a la proteína StpA cuando esta no se encuentra unida a H-NS (Johansson y Uhlin, 1999). La protección por parte de H-NS se produce por interacción directa entre ambas proteínas (Johansson *et al.*, 2001), lo que indica que StpA se encuentra de forma predominante formando heterodímeros con H-NS.

En función de la información disponible hasta el momento, los efectos moduladores de la proteína StpA han sido caracterizados en mutantes *hns*, en los que los niveles de StpA se incrementan. Queda, por tanto, por caracterizar el papel modulador de StpA en células *hns*⁺, y no existen ejemplos claros de una modulación por parte de StpA sola (revisado en Dorman, 2004).

Aunque StpA puede unirse con más especificidad al ADN, de 4 a 6 veces (Zhang *et al.*, 1996; Sonnenfield *et al.*, 2001), H-NS tiene un papel dominante, quizá debido a sus niveles de expresión más que a su actividad intrínseca (Zhang *et al.*, 1996).

1.2.3.- H-NS: INTERACCIÓN CON OTRAS PROTEÍNAS

H-NS presenta una gran versatilidad en su función reguladora y es capaz de controlar una gran variedad de genes. Para ello es importante su capacidad de unirse al ADN de diferentes formas, pero también la interacción con una gran variedad de proteínas reguladoras.

Entre las proteínas que interactúan con H-NS se encuentran sus proteínas parálogas StpA (Williams *et al.*, 1996) y Sfh, recientemente identificada en *S. flexneri* y con un 59 % similitud con H-NS (Deighan *et al.*, 2003). Ambas pueden formar heterodímeros con H-NS y/o compensar parcialmente las funciones de H-NS en un mutante *hns* (Shi y Bennett, 1994; Deighan *et al.*, 2003). Éste es el caso del operón *bgl*. Tal y como ya se ha explicado en el apartado anterior, la proteína StpA puede complementar las funciones de mutantes en H-NS (Free *et al.*, 1998). Se ha descrito que la interacción entre StpA y H-NS se produce a través del extremo N-terminal de ambas proteínas (Williams *et al.*, 1996). Otros autores han propuesto unos dominios de interacción diferentes que incluirían la parte C-terminal de la proteína y que plantearían la existencia de diferencias entre la homodimerización de H-NS y la heterodimerización entre H-NS y StpA. Este hecho podría alterar la habilidad de las dos proteínas para unirse al ADN (Johansson *et al.*, 2001).

H-NS también interactúa con el regulador post-transcripcional Hfq (proteína de unión al ARN) y que juntamente con DrsA participarían en la regulación de *rpoS* (Muffler *et al.*, 1996). La interacción entre la proteína H-NS y el producto génico 5.5 del fago T7 suprime la actividad inhibitoria de H-NS en la transcripción tanto *in vivo* como *in vitro* (Liu y Richardson, 1993).

Aparte de la regulación a nivel transcripcional, H-NS también interactúa directamente con la proteína motora flagelar FliG y esta interacción es importante para la actividad funcional de esta proteína motora (revisado Dorman, 2004).

Finalmente miembros de la familia H-NS interactúan con miembros de la familia Hha, aspecto que se refiere en el próximo apartado.

1.3.- LA FAMILIA DE PROTEÍNAS HHA/YMOA

La familia de proteínas Hha/YmoA incluye proteínas de bajo peso molecular (alrededor de 8 kDa) y moderadamente básicas que participan, entre otros procesos, en la termo y osmoregulación de la expresión génica (Cornelis *et al.*, 1991; Nieto *et al.*, 1991; Mouriño *et al.*, 1996). Hha e YmoA han sido las más estudiadas de esta familia y han sido consideradas de forma independiente como miembros de una nueva familia de factores moduladores de la expresión génica (Mikulskis y Cornelis, 1994; de la Cruz *et al.*, 1992).

1.3.1.- LA PROTEÍNA HHA

1.3.1.1.- Características generales

La proteína Hha (High Hemolytic Activity) fue identificada en *E. coli* como modulador de la expresión de la toxina α -hemolisina (Nieto *et al.*, 1991). Presenta un peso molecular de 8,6 kDa, tiene 72 aminoácidos (muchos de ellos cargados lo que le proporciona características hidrofílicas) y se encuentra codificada por el gen *hha*, que se localiza en el minuto 10'5 del genoma de *E. coli*.

La estructura de la proteína Hha ha sido resuelta por resonancia magnética nuclear (RMN). Se trata de una proteína con motivos hélice-vuelta-hélice (Yee *et al.*, 2002) (Figura 1.3.1.).



Figura 1.3.1. Estructura tridimensional de la proteína Hha. En rojo se muestran las estructuras hélices α presentes en la proteína. Extraída de Yee *et al.*, 2002.

En la etapa inicial de caracterización del gen *hha* se puso de manifiesto que su inactivación confería un fenotipo pleiotrópico. Como que los mutantes *hha* presentaban características similares a los mutantes *hns* (alteraciones en el nivel de superenrollamiento

del ADN, alteraciones en la regulación de la expresión génica (Carmona *et al.*, 1993; Mouriño *et al.*, 1996), alteraciones en la frecuencia de transposición de elementos de inserción (Mikulskis y Cornelis, 1994; Balsalobre *et al.*, 1996)), se planteó que la proteína Hha podría ser representativa de una nueva familias de proteínas asociadas al nucleóide (de la Cruz *et al.*, 1992).

Por lo que hace referencia a sus niveles de expresión, el gen *hha* presenta unos niveles de transcripción y expresión más elevados en fase exponencial y decrece cuando las células llegan a fase estacionaria, a diferencia de lo que ocurre con el gen *hns*. Precisamente en fase exponencial, la osmolaridad del medio de crecimiento influye en los niveles intracelulares de Hha (Mouriño *et al.*, 1998).

La proteína Hha en respuesta a cambios ambientales, como la temperatura y la osmolaridad, puede afectar a la expresión de diferentes genes además de la expresión del operón hemolítico por el cual fue identificada. A continuación se describen algunos de ellos y sus efectos:

- La termoregulación de la toxina CNF2 y del antígeno de superficie (Plásmido Vir) se modifica en mutantes *hha* (Mouriño *et al.*, 1996).

- Diversos genes presentan reprimida su expresión dependiendo de la osmolaridad. Entre ellos se encuentra el gen que codifica para la proteína de membrana externa OmpA, para el enzima *crr/IIA^{Glc}* del sistema PTS (sistema fosfotransferasa) y para el gen *ahpC* (Alquil superóxido reductasa) miembro del regulón de estrés superóxido (Balsalobre *et al.*, 1999).

- El gen *hha* también se ha descrito en *S. enterica* serovar *typhimurium* (Fahlen *et al.*, 2000) y los mutantes *hha* incrementan la transcripción dependiente de osmolaridad y de disponibilidad de oxígeno del gen *hilA* (activador transcripcional requerido para la invasión y regulado por factores ambientales).

- Trabajos recientes han demostrado la intervención de la proteína Hha en la regulación del gen *ler* en la cepa de *E. coli* enterohemorrágica O157:H7 (Sharma y Zuerner, 2004). La proteína Ler (similar a la familia de proteínas H-NS (Elliot *et al.*, 2000)) es necesaria para la transcripción del operón *esp*, encargado de la formación del contacto con las células epiteliales intestinales (Jarvis *et al.*, 1995) en función de los estímulos ambientales.

1.3.1.2.- Regulación de la toxina α -hemolisina por parte de la proteína Hha

Las cepas uropatogénicas de *E. coli* frecuentemente son productoras de la toxina α -hemolisina. La α -hemolisina es una citotoxina extracelular considerada como factor de virulencia (Welch *et al.*, 1981) y que también ha sido estudiada por ser uno de los pocos ejemplos de proteínas secretadas al medio por *E. coli* (Holland *et al.*, 1986).

La hemolisina de *E. coli* pertenece a una amplia familia de toxinas citolíticas que se encuentran en bacterias Gram-negativas denominada RTX (Repeats in Toxin) (Welch *et al.*, 1981; Menestrina *et al.*, 1994). Esta familia incluye entre otros, la leucotoxina (LktA) de *Pasteurella haemolytica* (Strathdee y Lo, 1987) y la adenilato ciclasa (CyaA) de *Bordetella pertussis* (Glaser *et al.*, 1988). Las toxinas pertenecientes a esta familia comparten diversas características además de la presencia de una secuencia consenso de 8 repeticiones de aminoácidos (repeticiones RTX que dan nombre a la familia), como la capacidad para formar poros, la actividad leucotóxica y la organización génica.

El determinante genético responsable de la síntesis y secreción de la toxina es un operón altamente conservado constituido por cuatro genes *hlyCABD* (Figura 1.3.2.) y que se encuentra localizado tanto a nivel cromosómico como plasmídico (Goebel y Schrempf, 1971; O'Hanley *et al.*, 1993). Uno de los ejemplos mejor caracterizados es el del plásmido pHly152 que pertenece al grupo de incompatibilidad IncJ2 (Noegel *et al.*, 1981).

El operón hemolítico (*hly*) tiene una longitud aproximada de 7 kb y está formado por los genes *hlyC*, *hlyA*, *hlyB*, *hlyD*. El gen *hlyA* codifica para un polipéptido (pre-hemolisina) de 110 kDa que es activado a través del producto génico del gen *hlyC* dando lugar a la toxina madura. HlyC es una proteína de 170 aminoácidos (Juárez *et al.*, 1984) que activa a HlyA mediante acilación (Hardie *et al.*, 1991; Issartel *et al.*, 1991). Dentro de *hlyC* se encuentra una secuencia denominada *hlyM* de unas 200 pb, que participa en la modulación de la expresión de la hemolisina (Jubete *et al.*, 1995). La secreción de la toxina activa HlyA al medio externo es dependiente de las proteínas HlyB y HlyD (Braun *et al.*, 1993). Además, la secreción de hemolisina al medio externo es el prototipo de los sistemas de secreción de Tipo I, que implica a la proteína de membrana externa TolC (Wandersman y Delepelaire, 1990).

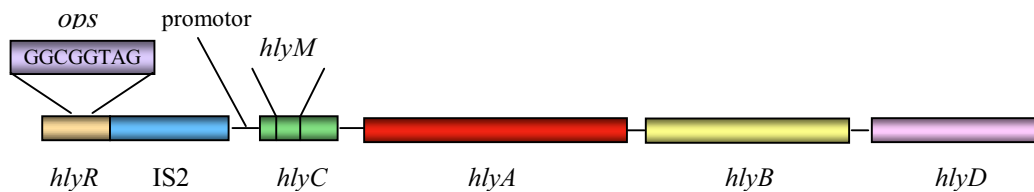


Figura 1.3.2. Estructura del determinante hemolítico *hlyCABD* del plásmido pHly152, donde se puede observar tanto la región reguladora como la codificante del operón.

A nivel transcripcional, la regulación del operón *hly* es compleja y existen, adicionalmente a la región promotora, otras secuencias implicadas en la regulación (revisión Madrid *et al.*, 2002b). Cuando se analizó la secuencia reguladora del operón *hly* presente en el plásmido pHly152, se observó que incluye la secuencia *hlyR* (Figura 1.3.2.) (Vogel *et al.*, 1988). Esta secuencia de 669 pb se encuentra localizada a unas 1'5 kpb antes del inicio del primer gen del operón (*hlyC*) y es esencial para la correcta expresión del mismo. Cuando esta secuencia no se encuentra presente, como en el caso del plásmido pANN202-312 (Godessart *et al.*, 1988), la expresión del operón se encuentra notablemente reducida (sobre todo en el caso de los genes *hlyB* y *hlyD*) a pesar de encontrarse en un elevado número de copias (Vogel *et al.*, 1988). Dentro de la secuencia *hlyR* se localiza la secuencia moduladora *ops* (Operon Polarity Suppressor) de 8 pb (Nieto *et al.*, 1996). Este elemento se encuentra ampliamente distribuido entre las bacterias Gram-negativas (Nieto *et al.*, 1996) y es un anti-terminador que se encuentra en la secuencia no codificante anterior a los operones y que actuaría eliminando la polaridad de la transcripción dentro de los operones.

La transcripción del operón *hlyCABD* produce dos transcritos diferentes, uno mayoritario de 4'0 kb (*hlyCA*) y uno minoritario de 8'0 kb (*hlyCABD*). La existencia de estos dos transcritos explicaría la mayor afectación de los genes *hlyB* y *hlyD* en ausencia de la región *hlyR*. Ambos transcritos se inician en el mismo promotor situado 400 pb antes del gen *hlyC*, pero la transcripción es fuertemente polar, y este fenómeno se encuentra acentuado por la formación de un bucle del ADN (“Stem-loop”) entre los genes *hlyA* y *hlyB* (Welch y Pellet, 1988; Koronakis *et al.*, 1988). Se cree que el transcrito policistrónico *hlyCABD* se origina al evitarse la terminación de la transcripción que habitualmente tiene

lugar en este terminador y el transcrito *hlyCA* es más estable (Hess *et al.*, 1986; Welch y Pellet, 1988; Koronakis *et al.*, 1988).

Se ha demostrado que otros genes no pertenecientes al operón *hly* están implicados en la síntesis y secreción de la hemolisina. El gen *tolC* interviene, como ya se ha comentado, en la secreción de la toxina al medio (Wandersman y Delepelaire, 1990) y el producto del gen *rfaH* (*hlyT* o *srfB*) ha sido caracterizado como activador transcripcional de la síntesis y secreción de la toxina (Bailey *et al.*, 1992).

El gen *hha* fue identificado al obtener mutantes de cepas de *E. coli* portadoras del plásmido pANN202-312 que incrementaba la producción de hemolisina a pesar de la ausencia de la región reguladora *hlyR* (Nieto *et al.*, 1991).

Al igual que sucede con otras toxinas sintetizadas por bacterias entéricas, la expresión de la toxina α -hemolisina de *E. coli* está regulada por diferentes condiciones ambientales. En condiciones de elevada osmolaridad, baja temperatura o en anaerobiosis, la síntesis de la toxina está reprimida (Mouriño *et al.*, 1994). Pudo ponerse de manifiesto que mutantes *hha* desreprimen la producción de hemolisina en condiciones no permisivas (elevada osmolaridad, baja temperatura). En estudios posteriores que se reseñan más adelante han permitido relacionar los efectos de la mutación *hha* incrementando la expresión de la toxina α -hemolisina tanto en ausencia de la secuencia *hlyR* como en condiciones de elevada osmolaridad o baja temperatura.

1.3.2.- INTERACCIÓN DE LA PROTEÍNA HHA CON LA PROTEÍNA H-NS PARA MODULAR LA EXPRESIÓN GÉNICA

Dentro del proyecto que tenía como objetivo conocer cómo ejerce su función la proteína Hha para modular la expresión génica del operón *hly*, un aspecto importante fue determinar la bioactividad de la proteína. Una vez obtenida la proteína purificada se analizó su capacidad para unirse a las regiones reguladoras del operón hemolítico. De forma sorprendente pudo ponerse de manifiesto que Hha se une a la región reguladora del operón *hly* de forma no específica y con baja afinidad, requiriendo grandes cantidades de la proteína para la formación de los complejos nucleoproteicos (Nieto *et al.*, 2000). Ello planteó la hipótesis de que quizás para la regulación del operón *hly* era necesaria la intervención de otros factores que se unieran al ADN, por lo que se realizaron estudios de interacción de Hha con otras proteínas. Estos estudios demostraron que Hha interacciona

con elevada afinidad con la proteína H-NS de *E. coli* (Nieto *et al.*, 2000). La interacción entre ambas proteínas sugirió que estuvieran implicadas de forma conjunta en la modulación de la expresión génica. Cuando se analiza la expresión de hemolisina en mutantes *hha* y *hns*, en ambos casos se produce la desrepresión parcial de la expresión de la toxina (a elevada osmolaridad y baja temperatura) que además es máxima en el doble mutante (Nieto *et al.*, 2000), sugiriendo que el complejo formado por Hha, H-NS y el ADN está implicado en la termo y osmoregulación del operón *hly*.

Estudios posteriores demostraron que la proteína H-NS es capaz de unirse a la región reguladora del operón *hly* y que esta unión es dependiente de la concentración. La cantidad de la proteína H-NS necesaria para formar los complejos nucleoproteicos es mucho menor que la de Hha. Además cuando se utilizan ambas proteínas conjuntamente (Hha y H-NS) los complejos formados son diferentes a los formados por cada una de las proteínas por separado (Nieto *et al.*, 2000).

Se han definido los lugares de unión de la proteína H-NS al operón *hly* (Madrid *et al.*, 2002a). Uno de ellos (lugar I) se localiza dentro de la secuencia reguladora *hlyR* y el otro (lugar II) se solapa parcialmente con la región promotora del operón. Ambos lugares se encuentran separados por el elemento de inserción IS2 y una zona de curvatura intrínseca (Figura 1.3.3.).

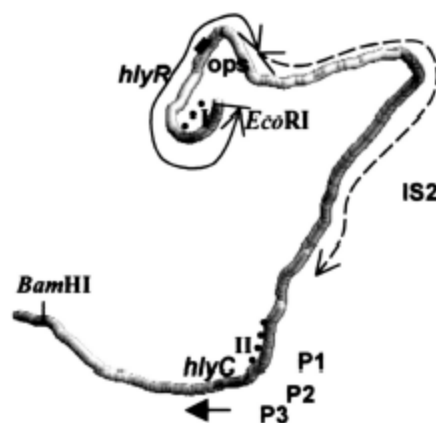


Figura 13.1. Predicción *in silico* de la curvatura de la región reguladora del operón *hly*. I y II indican los lugares de unión de H-NS. Extraído de Madrid *et al.*, 2002a.

La presencia de una región curvada entre los lugares de unión al ADN de H-NS ya se había descrito para otros genes reprimidos por dicha proteína, ya sea de forma dependiente (*virF*) o independiente (*espA*) de la temperatura (Brandi *et al.*, 1999; Falconi *et al.*, 1998). La afinidad de H-NS por el ADN varía en función de la temperatura y es más elevada a baja temperatura. Este hecho concuerda con los datos del fenotipo del mutante *hns* y con la más eficiente represión de la transcripción de H-NS a baja temperatura. La flexibilidad del ADN también aumenta a baja temperatura (aumento del grado de superenrollamiento del ADN) permitiendo el contacto entre moléculas de H-NS para la generación de un complejo nucleoproteico. Este complejo provocaría la oclusión de la secuencia *ops* y de la región promotora, reprimiendo la transcripción y eliminando el efecto anti-terminador del elemento *ops*. Existe un equilibrio entre la afinidad de H-NS por sus lugares de unión y el efecto de la temperatura.

Según el modelo propuesto por Madrid *et al.*, 2002a, la proteína Hha jugaría un papel facilitando la generación de heterooligómeros Hha-H-NS que nuclean la región promotora del operón *hly*, reprimiendo muy eficientemente la transcripción.

Teniendo en cuenta que una de las secuencias de unión de la proteína H-NS se localiza dentro de la secuencia *hlyR*, podemos ahora entender por qué un mutante *hha* incrementa la transcripción de un operón *hly* sin dicha secuencia. Un delicado equilibrio entre la unión de H-NS a ambas secuencias y la formación o no de un bucle determina el nivel de transcripción (Figura 1.3.2., revisado por Dorman 2004). La temperatura modifica la flexibilidad del ADN y desplaza el equilibrio hacia la apertura o cierre del bucle.

En ausencia de una de las dianas de unión de H-NS (dentro de *hlyR*), la topología puede alterarse, no se produce el “secuestro” de parte del complejo H-NS-Hha al separarse las dos dianas y el complejo represor bloquea la transcripción de forma permanente. En estas condiciones, un incremento de la expresión sólo se obtiene eliminando bien Hha, bien H-NS o bien las dos proteínas.

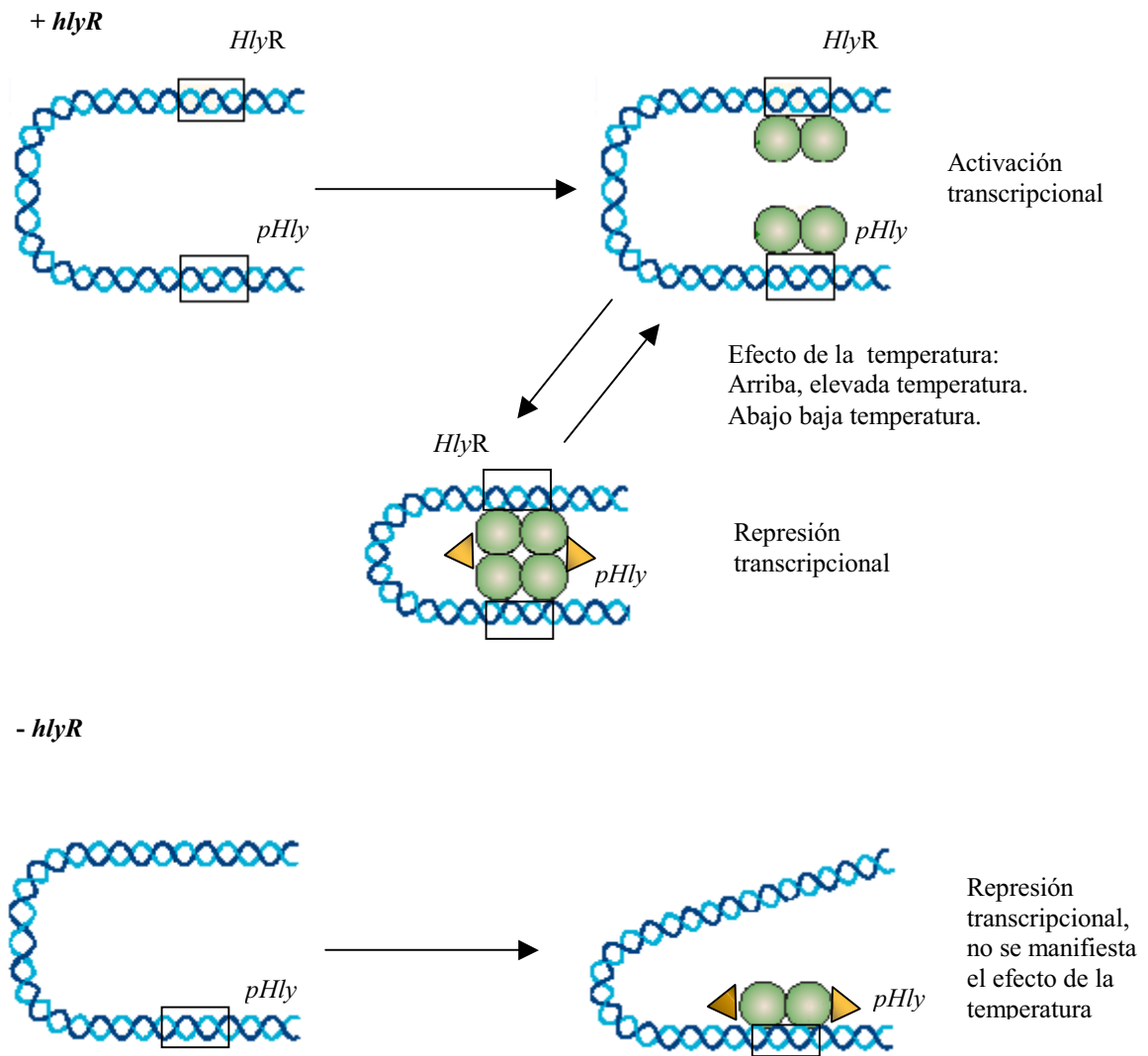


Figura 1.3.2. Esquema de la regulación del operón hemolítico por parte de las proteínas Hha y H-NS, tanto en presencia como en ausencia de la región reguladora *hlyR* y en función de la temperatura. Marcado con un recuadro se marcan los lugares de unión de H-NS que se encuentra representada como redondas de color verde. Hha está representada por los triángulos de color naranja. También se sitúa el inicio del operón hemolítico.

1.3.3. LA PROTEÍNA YmoA

1.3.3.1.- Regulación de la virulencia en *Y. enterocolitica*

Bacterias patógenas del género *Yersinia* (*Y. enterocolitica*, *Y. pseudotuberculosis* y *Y. pestis*) causan enfermedades en humanos con síntomas que pueden ir desde enteritis a septicemia y muerte, principalmente por invasión de los tejidos. En el caso de infecciones por *Y. enterocolitica* las manifestaciones clínicas son principalmente gastroenteritis con dolor abdominal que puede desencadenar en diarreas y apendicitis. *Y. enterocolitica* secreta una enterotoxina termoestable con capacidad de penetrar en la mucosa intestinal y concentrarse en las placas de Peyer y nódulos linfáticos (Cornelis *et al.*, 1991).

Una de las claves de la virulencia de *Y. enterocolitica* (y de las bacterias relacionadas *Y. pseudotuberculosis* y *Y. pestis*) se encuentra en el plásmido de virulencia denominado pYV (Cornelis *et al.*, 1987) de aproximadamente 70 kb. Este plásmido contiene el virulón *yop*, responsable de la síntesis y secreción de aproximadamente 14 proteínas llamadas Yop (*Y*ersinia *o*uter proteins) y de 2 proteínas de membrana externa, la adhesina YadA y la lipoproteína YipA. El plásmido pYV además impone un requerimiento de calcio para crecer a 37 °C y en su ausencia se produce la secreción de las proteínas Yop (Cornelis *et al.*, 1987). Estas proteínas pueden representar el 20 % del total de las proteínas celulares (revisado por Cornelis *et al.*, 1998).

El plásmido pYV presenta una región conservada (en los plásmidos de las diferentes especies) de 20 kb responsable del fenómeno dependiente de calcio y también de la expresión coordinada de los genes *yop* (Cornelis *et al.*, 1987). En esta región se encuentra el gen *virF* (situado entre los locus *virB* y *virC*), que es un activador transcripcional responsable de la transcripción de muchos genes del virulón *yop* (Cornelis *et al.*, 1987). El gen *virF* fue descrito anteriormente en *Y. pestis* por Yother *et al.*, 1986 y se denominó *lcrF*. En el plásmido pYV también se encuentran los genes *ysc* que codifican para el sistema de secreción de tipo III responsable de la secreción de las proteínas Yop (Michiels *et al.*, 1991) y del que se hablará más adelante.

La mayoría de los genes implicados en la síntesis y liberación de las proteínas Yop y YadA están organizados en un único regulón bajo control transcripcional. El activador transcripcional VirF (Cornelis *et al.*, 1989; Lambert de Rouvroit *et al.*, 1992) juega un papel esencial en su regulación. A su vez, la expresión de VirF está termoregulada.

El activador transcripcional VirF es una proteína de 30'8 kDa y 271 aminoácidos perteneciente a la familia de reguladores AraC, con un 23 % de homología en el extremo C-terminal (Cornelis *et al.*, 1989) y que se encuentra formando dímeros en solución (Lambert de Rouvroit *et al.*, 1992). VirF se encuentra autoregulado y fuertemente regulado por temperatura, pero no es el sensor de la presencia o ausencia de calcio. La expresión de las proteínas Yop y YadA es dependiente de la proteína VirF, que se une a una región de 40 pb rica en AT localizada justo antes del lugar de unión de la ARN polimerasa (revisado Cornelis *et al.*, 1998).

En la revisión de Cornelis *et al.*, 1998 se dividen las proteínas Yop (altamente conservadas en el género *Yersinia*) en dos grupos: (i) las efectoras, que son aquellas que se liberan y translocan dentro de la célula eucariota y que no presentan una secuencia señal típica y, (ii) las translocadoras, que son aquellas involucradas en la translocación de los efectores a la célula eucariota.

Dentro de los efectores cabe destacar:

- YopH (51 kDa): Inicialmente llamada Yop51 es una proteína con actividad tirosín-fosfatasa que actúa sobre las proteínas tirosín-fosforiladas de los macrófagos de una forma muy rápida e impidiendo la adhesión.

- YopE (23 kDa): Proteína citotóxica que de forma indirecta causa la despolimerización de la estructura de microfilamentos de actina de las células HeLa y ayuda a evitar la fagocitosis.

- YopO/YpkA (81 kDa): Proteína con actividad serín-treonín quinasa. Se sitúa sobre la superficie interna de la membrana plasmática e interfiere sobre algunas señales de transducción de la célula eucariota.

- YopM (41'6 kDa): Proteína quinasa ácida (presenta motivos LRR) que hace posible la unión a la trombina e inhibe la agregación plaquetaria y que se dirige al núcleo celular.

Estos últimos efectores forman parte de la familia Rho de GTPasas monoméricas.

Dentro de los translocadores cabe resaltar:

- YopB (41'8 kDa) y Yop D (33'3 kDa): Ambas implicadas en la formación del poro a través del cual el efector Yop puede ser translocado.

- YopN (32'6 kDa): Es el elemento controlador de la liberación del aparato translocador y de los efectores. Además esta implicada en el control de la secreción de Yop por calcio. Cuando falta calcio se desregula la función de YopN y por tanto la liberación de

Yops. Mutantes en el gen que codifica para esta proteína permiten a las células secretar otras Yops a 37 °C incluso en presencia de iones calcio.

La proteína YadA (*Yersinia adhesin A*) es una proteína de membrana externa involucrada en la interacción entre *Yersinia* y la célula huésped (Cornelis *et al.*, 1991). YadA tiene un peso molecular de 41 a 47 kDa dependiendo de la especie que la produzca, forma estructuras homodiméricas que funcionan como adhesina y son responsables de la aglutinación (revisado por Cornelis *et al.*, 1998). La proteína YadA puede unirse a una gran variedad de moléculas de superficie celular eucariotas y por tanto está implicada en la colonización, además tiene función protectora de la bacteria sobre la acción bactericida del suero humano. Esta adhesión a la célula huésped permite la liberación de las proteínas Yop.

El plásmido pYV también codifica para la lipoproteína YipA que se encuentra regulada por VirF (Lambert de Rouvroit *et al.*, 1992).

Otros factores de virulencia importantes pero de codificación cromosómica son: (i) la invasina (*inv*), una proteína de membrana externa que promueve la invasión de los tejidos. Se expresa en fase exponencial tardía en respuesta a cambios en las condiciones de crecimiento y está influenciada por la proteína YmoA (Ellison *et al.*, 2003) y, (ii) también es importante la enterotoxina Yst, solo producida a baja temperatura, en fase tardía de crecimiento.

Finalmente en la virulencia también es importante el sistema de secreción codificado en el virulón Yop y que es considerado como el arquetipo del sistema de secreción de tipo III o también denominado inyectosoma (TTSS). La actividad del sistema de secreción se encuentra influenciada por diferentes señales ambientales tanto de forma positiva como negativa (DeBord *et al.*, 2003). El inyectosoma completo está formado por dos partes, una parte interna que contiene 10 proteínas (similares a las proteínas del cuerpo basal del flagelo) y una parte externa. Esta estructura tiene un poro central de aproximadamente 50 Å, una longitud de 60-80 nm y 6-7 nm de ancho (revisado en Cornelis, 2002). La formación de los inyectosomas se produce cuando la bacteria detecta que la temperatura ha aumentado a 37 °C. La secreción de las proteínas Yop también requiere de chaperonas de la familia Syc. En la figura 1.3.3. se muestra la formación del inyectosoma y la translocación de las proteínas Yop.

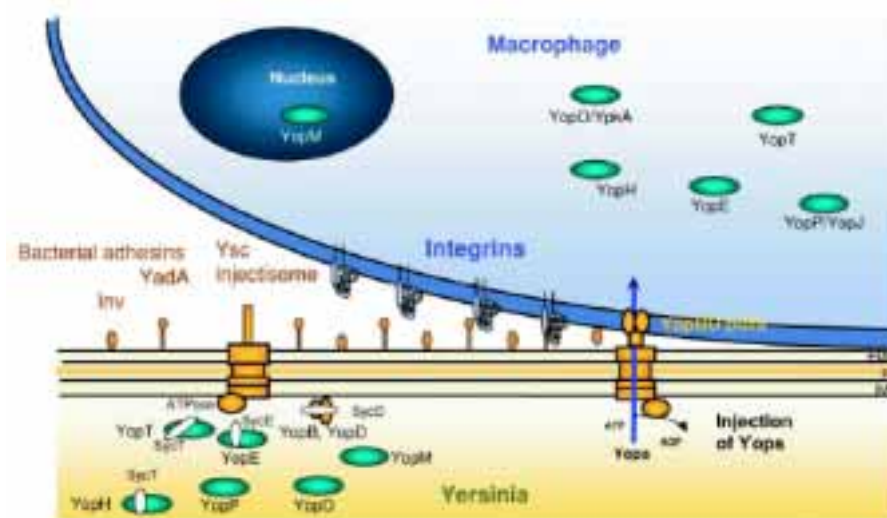


Figura 1.3.3. Secreción de las proteínas Yop mediante el inyectosoma Ysc y translocación a través de la membrana celular señalizada. Se produce el contacto con la célula eucariota a través de las adhesinas bacterianas y las integrinas de la superficie celular. Cuando aumenta la temperatura se produce la formación del inyectosoma y el contacto entre las proteínas Yop y sus correspondientes chaperonas Syc. Finalmente se produce la secreción a través del canal en el que intervienen las proteínas translocadoras YopB y D. Figura extraída de Cornelis, 2002.

1.3.3.2.- Características generales de la proteína YmoA

La proteína YmoA (*Yersinia modulator*) fue caracterizada en *Y. enterocolitica* como modulador de la expresión de diferentes factores de virulencia (proteínas Yop y adhesina YadA) dependiente de temperatura (Cornelis *et al.*, 1991) tal y como se ha comentado anteriormente. La expresión de dichos factores se encuentra reprimida en condiciones de vida libre, a baja temperatura (25° C) y se ve incrementada cuando *Yersinia* interacciona con el hospedador, aumentando la temperatura (37° C). YmoA también está implicada en la regulación de la expresión génica a través de otros factores ambientales como el estrés osmótico (Lambert de Rouvroit *et al.*, 1992).

YmoA es una proteína pequeña, de aproximadamente 8 kDa y extremadamente rica en residuos cargados (36 %). Además presenta un alto grado de homología en su secuencia de aminoácidos (82 %) con la proteína Hha de *E. coli* (De la Cruz *et al.*, 1992). Ambas proteínas pueden ser funcionalmente intercambiables (Mikulskis y Cornelis, 1994;

Balsalobre *et al.*, 1996), sugiriendo ello que forman parte de una nueva familia de moduladores (Cornelis *et al.*, 1991; Carmona *et al.*, 1993).

Los mutantes *ymoA* presentan desregulada la expresión del regulador VirF y en consecuencia la del operón de virulencia *yop* (proteínas Yop y YadA) provocando un aumento en la expresión de estos factores de virulencia (Cornelis *et al.*, 1991). A baja temperatura el nivel de transcripción de estos factores está desreprimido y a elevada temperatura están incrementados respecto a la cepa salvaje.

YmoA también ha sido descrita como un regulador negativo de la expresión de la invasina (*inv*) de *Y. enterocolitica* necesaria para el proceso de invasión (Ellison *et al.*, 2003).

1.3.4.- OTRAS PROTEÍNAS DE LA FAMILIA HHA/YMOA

A lo largo de los últimos años se han ido describiendo nuevos miembros pertenecientes a la familia Hha/YmoA (revisado por Madrid *et al.*, 2002b).

Uno de esos miembros es la proteína RmoA. RmoA se encuentra codificada en el plásmido R100 de *E. coli* (cerca al promotor del operón *tra*) y fue identificada mediante la búsqueda de homologías con las proteínas Hha y YmoA (Nieto y Juárez, 1996). La proteína RmoA podría actuar como moduladora de la transferencia plasmídica en respuesta a determinados factores ambientales como la osmolaridad. Este hecho viene sugerido por el aumento de la frecuencia de conjugación que se produce en un mutante *rmoA* en condiciones de baja osmolaridad (Nieto *et al.*, 1998).

Otro miembro identificado de esta familia de proteínas es la proteína YdgT. YdgT fue identificada a través de un análisis del genoma completo de *E. coli* y presenta un 38 % de identidad y un 67 % de similitud a nivel de aminoácidos con la proteína Hha (Paytubi *et al.*, 2004). La proteína YdgT tiene 71 aminoácidos y cuando es sobreexpresada atenúa el fenotipo de un mutante *hha* como en el caso del operón hemolítico, en el que se incrementa la expresión de hemolisina en ausencia de YdgT (Paytubi *et al.*, 2004). Al igual que Hha, YdgT también es capaz de unirse a H-NS (de forma estable) y a StpA.

El plásmido R446 de *Proteus morgani*, perteneciente al grupo de incompatibilidad IncM, presenta una ORF denominada Orf5 que está implicada en la expresión dependiente de temperatura del pili conjugativo (Tietze y Tschäpe, 1994). La proteína codificada por esta Orf5 presenta homología con la proteína Hha (Nieto y Juárez, 1996). Además este

plásmido también presenta justo antes la Orf4 que es un miembro de la familia de proteínas H-NS (Nieto y Juárez, 1999).

En el plásmido de virulencia pO157 de *E. coli* O157:H7 también se ha identificado un miembro de la familia de proteínas Hha/YmoA (Burland *et al.*, 1998).

Recientemente en el plásmido de virulencia pWR501 de *S. flexnerii* se ha identificado una proteína homóloga a la proteína Hha. La proteína ha sido descrita como un hipotético regulador (Venkatesan *et al.*, 2001).

Otro miembro de dicha familia ha sido identificado en el plásmido R27 (Sherburne *et al.*, 2000), la ORF 182. R27 es un plásmido de *S. enterica* serovar *typhimurium* perteneciente al grupo de incompatibilidad HI y responsable de la resistencia múltiple a antibióticos. Este plásmido también presenta una ORF homóloga a la proteína H-NS (ORF 164) (Sherburne *et al.*, 2000). Los plásmidos pertenecientes al grupo de incompatibilidad IncHI tienen una transferencia termosensible (Maher y Taylor, 1993) y la frecuencia de conjugación es mayor a baja temperatura (25-30 °C). Tanto la proteína homóloga a Hha como la proteína homóloga de H-NS codificadas en el plásmido R27 están implicadas en el proceso de transferencia termoregulado, como se demuestra con la construcción de diferentes mutantes que presentan una conjugación desreprimida a 33 °C (Forns *et al.*, 2005).

1.4.- OBJETIVOS DE ESTE TRABAJO

En este trabajo se planteó como primer objetivo analizar los dominios funcionales de la proteína Hha. Ya se conocía que la unión de esta proteína a la proteína H-NS es importante para la regulación del operón hemolítico (revisado Madrid *et al.*, 2002a) y por ello se intenta ver cuáles son los residuos importantes de Hha para la unión. Con dicha finalidad se obtuvieron diferentes construcciones de proteínas truncadas de Hha y de proteínas con diferentes mutaciones puntuales que fueron analizadas en su capacidad para retener la proteína H-NS.

Relacionado con este objetivo y ante la hipótesis de que proteínas de la familia Hha pudiesen ser equivalentes al dominio N-terminal de proteínas de la familia H-NS, se procedió a comprobar si una proteína híbrida entre Hha y el dominio de unión al ADN de H-NS era capaz de complementar los fenotipos de un mutante *hms*, es decir, ver si la proteína Hha era capaz de sustituir al dominio de oligomerización de H-NS. También se quería comprobar si esta proteína híbrida era capaz de unirse al ADN con la misma afinidad y especificidad con la que lo hace la proteína H-NS.

Por otra parte, estudios previos habían demostrado la existencia de miembros de la familia H-NS en diferentes especies de *Enterobacteriaceas* (apartado 1.2.1), pero no había sido descrita todavía en *Y. enterocolitica*. Por ello, en esta memoria se planteó como objetivo identificar la secuencia del gen *hms* en *Y. enterocolitica* y posteriormente demostrar que la interacción entre las proteínas Hha y H-NS es común a otros miembros de estas dos familias (YmoA y H-NS).

Finalmente, como último objetivo se planteó la obtención de un mutante para el gen *hms* de *Y. enterocolitica*. La obtención de dicho mutante permitiría ver que genes tienen alterada su expresión e intentar conocer algunos detalles más sobre la regulación de la virulencia en *Y. enterocolitica*.