

00591

7



UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO

INSTITUTO DE BIOTECNOLOGIA

“Diseño y Construcción de cepas de  
*Bacillus subtilis* sobreproductoras  
de proteínas heterólogas”

T E S I S

QUE PARA OBTENER EL GRADO DE:  
DOCTOR EN CIENCIAS BIOQUIMICAS

P R E S E N T A:  
Q.F.B. JANET JAN ROBLERO



TESIS CON  
FALLA DE ORIGEN

276860

CUERNAVACA, MORELOS

2000



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

El presente trabajo se llevó a cabo en el Departamento de Microbiología Molecular del Instituto de Biotecnología de la UNAM, bajo la dirección del Dr. Enrique Merino Pérez. Para la realización de este trabajo se contó con el apoyo económico del Consejo Nacional de Ciencia y Tecnología a través de la beca 91903 y de la Dirección General de Estudios de Posgrado (DGEP) UNAM-México a través del proyecto 030524.

## **MIEMBROS DEL JURADO**

**Presidente:** Dr. Edmundo Calva Mercado

**Secretario:** Dr. Enrique Merino Pérez

**Vocal:** Dra. Amelia Farrés González S.

**Vocal:** Dr. Octavio Tonatiuh Ramírez Reivich

**Vocal:** Dr. Baltazar Becerril Luján

**Suplente:** Dra. Alejandra Bravo de la Parra

**Suplente:** Dr. Jaime Padilla Acero

## AGRADECIMIENTOS

Agradezco de todo corazón a todas las personas que ha lo largo de mi vida me han brindado su apoyo; particularmente en momentos difíciles, sin duda alguna.....los que nos hacen ser más fuertes: a mi queridísimos amigos Graciela y Ramón, por permitirme entrar en su vida, por la amistad, cariño y apoyo incondicional; a Verito R. y Paola J. por ser de las amigas que uno siempre tiene presente; a Raúl F. por ser el mejor e incondicional de los amigos; a Armando por todo la cariño y amistad; Argelia L. por la amistad y solidaridad; a Paulina B., por todo el cariño y palabras de constante motivación; a Barbus por esa noble amistad; a mis amigas Power: Elisa, Viole, MariCarmen, Rosana, Raquel, Laura e Irma, por los momentos maravillosos y divertidos; a Martha y Adriana, por sus consejos sabios y valiosa amistad; a Noemí F., Chava F., Pancho y Lauris, por todos los momentos compartidos y muestras de genuina amistad; a Don Pablo por todas las muestras de cariño; a Pedro S., Cipriano, Freddy, Fer y Abel, siempre dispuestos a ayudarme; a Alfredo M. y Male, por toda la amistad y apoyo incondicional. Agradezco también a cada uno de los compañeros de laboratorio, de manera especial a: Ramón, Alfredo II, Joel, el Rober y Guillermo G.; a Mechita, porque con su trabajo diario nos ayuda a salir adelante; al Dr. Bolivar por su apoyo y motivación constante; de manera especial.....a Enrique Merino, por su paciencia y dirección durante la realización de este trabajo, debo parte fundamental de mi formación; a los miembros de mi comité tutorial: Edgardo E., Guadalupe E., Ale B. y Enrique Morett, por su interés y apoyo en mi formación profesional; a los integrantes del jurado, por las críticas excelentes y constructivas a mi trabajo. Por último, a todas aquellas personas que me han brindado su amistad y cariño desinteresada e incondicionalmente, pero que lamentablemente ya no están presente.....su recuerdo vive por siempre en mi mente y en mi corazón.

## DEDICATORIAS

A **Don Emilio Jan**, mi padre, quien sembró en mí el deseo de superación constante, la búsqueda del éxito, sin dejar a un lado la humanidad y pequeñez que cada ser persona tiene dentro de su corazón. Su recuerdo siempre me motivará e impulsará.

A mi madre, **Doña Rosita**, un ejemplo de lucha constante, que ha lo largo de la vida me ha brindado su infinita paciencia, amor, aliento y solidaridad. Mil gracias por ser más que una madre.....realmente eres una amiga a todo dar.

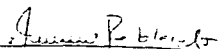
A mis hermanos, **Emilio y Javier**, por el amor incondicional, motivaciones, apoyo y ánimo me han proporcionado a lo que ha lo largo de este camino. Por contagiarme su alegría de vivir y ver que las cosas, por difíciles que parezcan, tienen solución..... simplemente hay que verlas desde otra perspectiva.

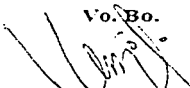
# TESIS: DISEÑO Y CONSTRUCCION DE CEPAS DE *Bacillus subtilis* SOBREPDUCTORAS DE PROTEINAS HETEROLOGAS

## ABSTRACT

The *aprE* gene of *Bacillus subtilis* codes for the mayor alkaline protease known as subtilisin. It is expressed during the transition state and transcribed by the  $\sigma^A$  form of the RNA polymerase (RNAP). Although the basic and applied importance of the *aprE* gene is widely accepted, and its Transcription Initiation Site (TIS) has been determined, there is no clear description of its promoter sequence. Part of this work was involved in the characterization of the regulatory region of the *aprE* gene (*rraprE*) from *B. subtilis*. By computer analysis and site-directed mutagenesis, we localized the *aprE* promoter sequence seven base pairs (bp) upstream from its TIS. The promoter sequence functional in vivo is: TACTAA—17 bp—TACAAT, at least under the growth conditions assayed in this study. We also characterized the static curvature properties of the *rraprE* DNA and found two different areas of DNA-bending, within the first four hundred bp upstream of its TIS. They are located beside DegU-binding site (around 100 and 200 bp upstream the *aprE* TIS). We postulate that these particular curved DNA regions could play a role in the interaction with some regulatory proteins and discuss possible implications related to *aprE* transcription regulation.

We evaluated the effect of several genetic factors that had been reported to play a role in the expression of recombinant protein in *B. subtilis*. We have utilized the  $\beta$ -galactosidase reporter protein from *Escherichia coli* as our model for the overproduction of heterologous proteins in *B. subtilis*. The *lacZ* gene was expressed using the regulatory region of the *aprE* gene. The *aprE* expression is highly controlled by negative and positive regulators. Our efforts were focused to improve factors that modulate the transcription and translation initiation rates, genetic and mRNA stability and the effect of specific genetic backgrounds such as *degL32* and *hpr2*, that until now, had been studied independently in the construction of protein overproducer strains of *B. subtilis*. The most significant result corresponded to the change of the native -35 promoter box to the consensus TTGACA sequence. This modification yielded an increase of 100-fold the  $\beta$ -Gal activity, in relation to the strain considered as the wild type. Interestingly, this new *aprE* promoter still responded to temporal regulation as the one observed in the wild type *aprE* promoter; therefore, its expression begins at the onset of the transition state.

  
Q.F.B. Janer Jan Róblero  
Estudiante de Doctorado en  
Ciencias Bioquímicas

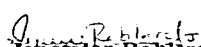
Vo.Bo.  
  
Dr. Enrique Merino Pérez  
Tutor  
Investigador Titular A "T.C"  
Depto. de Microbiología Molecular  
Instituto de Biotecnología

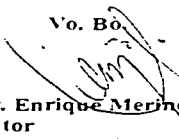
# TESIS: DISEÑO Y CONSTRUCCION DE CEPAS DE *Bacillus subtilis* SOBREPDUCTORAS DE PROTEINAS HETEROLOGAS

## RESUMEN

El gen *aprE* de *Bacillus subtilis* codifica para la proteasa alcalina mayoritaria conocida como subtilisina. Esta enzima se expresa durante el estado de transición y es transcrito por el factor  $\sigma^A$  de la RNA polimerasa (RNAP). Pese a que la importancia básica y aplicada del gen *aprE* es ampliamente aceptada, y que su sitio de inicio de la transcripción (SIT) ha sido determinado, no había una descripción clara de su secuencia promotora. Parte de este trabajo se enfocó en la caracterización de la región de regulación del gen *aprE* ( $\pi aprE$ ) de *B. subtilis*. Mediante análisis de cómputo y mutagénesis sitio dirigida, localizamos la secuencia promotora de *aprE* a siete pares de bases (pb) hacia arriba su SIT. La secuencia promotora funcional *in vivo* es: TACTAA— 17 pb —TACAAT, bajo las condiciones ensayadas en este estudio. Así mismo, caracterizamos las propiedades de curvatura estática del DNA en la  $\pi aprE$  y encontramos dos regiones de DNA-curvo dentro de las primeras cuatrocientos pb hacia arriba de su SIT, localizándose a los lados de la secuencia de unión de la proteína DegU (alrededor de -100 y -200 pb en relación al SIT). Postulamos que estas regiones de DNA curvadas podrían jugar un papel en la interacción con algunas proteínas reguladoras y discutimos las posibles implicaciones relacionadas con la regulación de la transcripción de *aprE*.

Por otra parte, evaluamos el efecto de algunos factores genéticos que han sido reportados en la inducción de la expresión de proteínas recombinantes en *B. subtilis*. Utilizamos a la proteína reportera  $\beta$ -galactosidasa como nuestro modelo para la sobreproducción de proteínas heterólogas en *B. subtilis*. El gen *lacZ* fue expresado usando la región de regulación del gen *aprE*. La expresión de *aprE* es altamente controlada por reguladores tanto negativos como positivos. Nuestros esfuerzos fueron dirigidos hacia factores que modulan la tasa de iniciación de la transcripción y la traducción, la estabilidad del material genético y del mRNA, el efecto de fondos genéticos específicos tales como *degL*<sup>32</sup> and *hpr2*; que hasta ahora, han sido estudiados de manera independiente en la construcción de cepas sobreproductoras de proteínas de *B. subtilis*. De los factores analizados, el cambio más importante correspondió a la formación de la caja -35 consenso de los promotores dependientes de  $\sigma^A$  en la  $\pi aprE$ . Este cambio produjo un incremento de 100 veces la actividad con respecto a la cepa considerada como silvestre. Es de importancia destacar que la regulación de esta nueva cepa es similar a la que lleva el promotor *aprE* silvestre; es decir, su expresión se presenta al inicio de la fase de transición.

  
Q. E. B. Jancelián Robliero  
Estudiante de Doctorado en  
Ciencias Bioquímicas

Vo. Bó.  
  
Dr. Enrique Merino Pérez  
Tutor  
Investigador Titular A "T.C"  
Depto. de Microbiología Molecular  
Instituto de Biotecnología



# INDICE GENERAL

	Página
I.1 Resumen.....	1
I.2 Abstract.....	4
II Introducción.....	6
II.1 <i>Bacillus subtilis</i> en la producción de proteínas recombinantes. ....	7
II.2 Características del gen <i>aprE</i> . ....	11
III Generalidades.....	14
III.1 Sistema de expresión. ....	15
III.2 Estabilidad del material genético. ....	16
III.3 Diseño y construcción de una unidad de expresión. ....	16
III.4 Factores a nivel transcripcional.....	17
III.5 Factores a nivel traduccional. ....	20
III.6 Uso preferencial de codones. ....	20
III.7 Degradación proteica. ....	20
III.8 Ciclo celular.....	21
III.8.1 Crecimiento exponencial. ....	21
III.8.2 Estado de transición. ....	21
III.8.3 Proceso de esporulación. ....	29
IV Justificación.....	33
V Objetivos.....	35
VI Presentación de manuscritos	
VI.1 Characterization of the 5' subtilisin ( <i>aprE</i> ) regulatory region from <i>Bacillus subtilis</i> . ....	38

VI.2	Construction of protein overproducer strains of <i>Bacillus subtilis</i> by an integrative approach. ....	45
VII	Experimentos no incluidos en los artículos	
VII.1	Adición de 20 pb entre los sitios de unión 3 y 4 de la proteína Hpr en la región de regulación del gen <i>aprE</i> .....	71
VIII	Discusión general.....	79
IX	Conclusiones.....	91
X	Perspectivas.....	94
XI	Referencias.....	96

## ABREVIATURAS

AMPC	AMP cíclico
<i>amyE</i>	Gen que codifica para la enzima amilasa
<i>aprE</i>	Gen que codifica para la proteasa alcalina de <i>Bacillus subtilis</i>
Ap <sup>R</sup>	Resistencia a ampicilina
<i>B</i>	<i>Bacillus</i>
β-Gal	Proteína β-Galactosidasa
CAP	Proteína activadora catabólica
<i>cat</i>	Gen que codifica para cloranfenicol acetyltransferasa
Cm <sup>R</sup>	Resistencia a cloranfenicol
<i>hisA</i>	Gen participante en la biosíntesis de histidina. Fosforribosil-fomimino 5-aminoimidazol carboxamida ribótido isomerasa
g	gramo
<i>glyB</i>	Gen involucrado en la biosíntesis de glicina. Fosfoserin aminotransferasa
GRAS	Generalmente reconocida como segura
<i>lacZ</i>	Gen que codifica para la proteína β-Galactosidasa
NTP	Nucleótido trifosfato
ORF	Open reading frame
pb	pares de bases
plk	poli-linker
RNAP	RNA polimerasa
<i>rraprE</i>	Región de regulación del gen <i>aprE</i> de <i>B. subtilis</i>
<i>rraprE</i> (WT)	Región de regulación silvestre del gen <i>aprE</i>
<i>rraprE</i> (TTGACA)	Región de regulación del gen <i>aprE</i> que lleva la caja -35 consenso
Met	Metionina
<i>npr</i>	Gen que codifica para la proteasa neutra de <i>B. subtilis</i>

mRNA	RNA mensajero
rRNA	RNA ribosomal
tRNA	RNA de transferencia
R	Resistencia
SD	Shine-Dalgarno
SIT	Sitio de inicio de la transcripción
U/mgprot	Unidades sobre microgramos de proteína
X-Gal	5-bromo-4-cloro-3-indolyl- $\beta$ -D-galactopyranoside
$\Omega$ 20pb	Adición de 20 pb
$\Delta$	Remoción

**En los artículos:**

<i>amy</i> -front	5' end of the <i>amyE</i> gene
<i>amy</i> -back	3' end of the <i>amyE</i> gene
bp	base pair
IPTG	isopropyl $\beta$ -D-thiogalactopyranoside
nt	nucleotide
PAGE	Polyacrylamide-gel electrophoresis
PCR	Polymerase chain reaction
RBS	Ribosome binding site
TIS	Transcription initiation site
wt	wild type

## **I.1 RESUMEN**

El gen *aprE* de *Bacillus subtilis* codifica para la proteasa alcalina mayoritaria conocida como subtilisina. Esta enzima se expresa durante el estado de transición y es transcrito por el factor  $\sigma^A$  de la RNA polimerasa (RNAP). Pese a que la importancia básica y aplicada del gen *aprE* es ampliamente aceptada, y que su sitio de inicio de la transcripción (SIT) ha sido determinado, no había una descripción clara de su secuencia promotora. Parte de este trabajo se enfocó en la caracterización de la región de regulación del gen *aprE* (*rraprE*) de *B. subtilis*. Mediante análisis de cómputo y mutagénesis sitio dirigida, localizamos la secuencia promotora de *aprE* a siete pares de bases (pb) hacia arriba su SIT. La secuencia promotora funcional *in vivo* es: TACTAA— 17 pb —TACAAT, bajo las condiciones ensayadas en este estudio. Así mismo, caracterizamos las propiedades de curvatura estática del DNA en la *rraprE* y encontramos dos regiones de DNA-curvo dentro de las primeras cuatrocientos pb hacia arriba de su SIT, localizándose a los lados de la secuencia de unión de la proteína DegU (alrededor de -100 y -200 pb en relación al SIT). Postulamos que estas regiones de DNA curvadas podrían jugar un papel en la interacción con algunas proteínas reguladoras y discutimos las posibles implicaciones relacionadas con la regulación de la transcripción de *aprE* (Jan *et al.*, 2000).

Por otra parte, evaluamos el efecto de algunos factores genéticos que han sido reportados en la inducción de la expresión de proteínas recombinantes en *B. subtilis*. Utilizamos a la proteína reportera  $\beta$ -galactosidasa como nuestro modelo para la sobreproducción de proteínas heterólogas en *B. subtilis*. El gen *lacZ* fue expresado usando la región de regulación del gen *aprE*. La expresión de *aprE* es altamente controlada por reguladores tanto negativos como positivos. Nuestros esfuerzos fueron dirigidos hacia factores que modulan la tasa de iniciación de la transcripción y la traducción, la estabilidad del material genético y del mRNA, el efecto de fondos genéticos específicos tales como *degU32* and *hpr2*; que hasta ahora, han sido estudiados de manera independiente en la construcción de cepas sobreproductoras de proteínas de *B. subtilis*. De los factores analizados, el cambio más importante correspondió a la formación de la caja -35 consenso de los promotores dependientes de  $\sigma^A$  en la *rraprE*. Este cambio produjo un incremento de 100 veces la actividad con respecto a la cepa considerada como silvestre. Es de importancia destacar que la

regulación de esta nueva cepa es similar a la que lleva el promotor *aprE* silvestre; es decir, su expresión se presenta al inicio de la fase de transición (Jan *et al.* sometido).

## **I.2 ABSTRACT**



The *aprE* gene of *Bacillus subtilis* codes for the mayor alkaline protease known as subtilisin. It is expressed during the transition state and transcribed by the  $\sigma^A$  form of the RNA polymerase (RNAP). Although the basic and applied importance of the *aprE* gene is widely accepted, and its Transcription Initiation Site (TIS) has been determined, there is no clear description of its promoter sequence. Part of this work was involved in the characterization of the regulatory region of the *aprE* gene (*rprprE*) from *B. subtilis*. By computer analysis and site-directed mutagenesis, we localized the *aprE* promoter sequence seven base pairs (bp) upstream from its TIS. The promoter sequence functional in vivo is: TACTAA—17 bp—TACAAT, at least under the growth conditions assayed in this study. We also characterized the static curvature properties of the *rprprE* DNA and found two different areas of DNA-bending, within the first four hundred bp upstream of its TIS. They are located beside DegU-binding site (around 100 and 200 bp upstream the *aprE* TIS). We postulate that these particular curved DNA regions could play a role in the interaction with some regulatory proteins and discuss possible implications related to *aprE* transcription regulation (Jan *et al.* 2000).

We evaluated the effect of several genetic factors that had been reported to play a role in the expression of recombinant protein in *B. subtilis*. We have utilized the  $\beta$ -galactosidase reporter protein from *Escherichia coli* as our model for the overproduction of heterologous proteins in *B. subtilis*. The *lacZ* gene was expressed using the regulatory region of the *aprE* gene. The *aprE* expression is highly controlled by negative and positive regulators. Our efforts were focused to improve factors that modulate the transcription and translation initiation rates, genetic and mRNA stability and the effect of specific genetic backgrounds such as *degU32* and *hpr2*, that until now, had been studied independently in the construction of protein overproducer strains of *B. subtilis*. The most significant result corresponded to the change of the native -35 promoter box to the consensus TTGACA sequence. This modification yielded an increase of 100-fold the  $\beta$ -Gal activity, in relation to the strain considered as the wild type. Interestingly, this new *aprE* promoter still responded to temporal regulation as the one observed in the wild type *aprE* promoter; therefore, its expression begins at the onset of the transition state (Jan *et al.* submitted).

## **II INTRODUCCION**

## II.1 *Bacillus subtilis* en la producción de proteínas recombinantes

Durante las dos décadas pasadas, la investigación basada en tecnología de DNA recombinante ha permitido el desarrollo de una variedad de productos con aplicaciones tan diversas como el tratamiento de cáncer o la producción de detergentes biológicos (Glick y Pasternak, 1994). La mayoría de estos productos han sido obtenidos a través de sistemas de expresión genética; como consecuencia, la optimización y estudio de dichos sistemas ha tenido importantes avances.

Diversos sistemas hospederos han sido empleados en la producción de proteínas y de metabolitos de interés e incluyen el empleo de fagos, bacterias, levaduras, plantas, hongos, insectos y células de mamífero en cultivo (Marino, 1989). Cada uno de estos sistemas representa características especiales, por lo que la selección de alguno de ellos requiere evaluar las ventajas y desventajas de un sistema sobre otro. En ocasiones la decisión se encuentra sujeta a consideraciones de índole económica, principalmente en la selección del proceso y diseño, requerimientos en equipo, instalaciones y personal (Ver Tabla 1).

De manera general, cuando se inicie la producción de cualquier proteína recombinante se consideran algunas cuestiones tales como la selección del microorganismo huésped y del sistema de expresión a utilizar. Además, se deben tomar en cuenta características relacionadas con el crecimiento del microorganismo hospedero y aquellas vinculadas con la proteína de interés, como son: integridad biológica, grado de pureza, cantidad de proteína que se obtiene, posibles efectos tóxicos para la célula huésped, localización del producto (intracelular o extracelular) y finalmente, la posibilidad de modificaciones post-traduccionales.

De los microorganismos utilizados a nivel industrial, algunas especies del género *Bacillus* son un grupo de hospederos particularmente atractivos para la síntesis de cantidades abundantes de proteínas heterólogas; convirtiéndolo en uno de los géneros más valiosos para la industria biotecnológica (Ferrari *et al.*, 1993; Harwood, 1992; Zukowski, 1992).

**Tabla 1 Diferentes hospederos para la producción de proteínas de interés comercial**

Sistema hospedero	Ventaja	Desventaja
<b><i>Escherichia coli</i></b> (Bacteria)	Amplia variedad de vectores de clonación	No es una bacteria grado GRAS
	Expresión genética fácil de controlar	Las bacterias Gram-negativas poseen alto contenido de endotoxinas
	Experiencia en el área de fermentaciones; ha sido utilizada ampliamente para la producción de proteínas recombinantes y de metabolitos de interés	Las proteínas se agregan como cuerpos de inclusión cuando son sobreexpresadas; haciendo más difícil su purificación y disminuyendo su actividad biológica
<b><i>Bacillus subtilis</i></b> (Bacteria)	La proteína recombinante no sufre modificaciones post-traduccionales	La proteína recombinante no sufre modificaciones post-traduccionales
	Bacteria grado GRAS (Generalmente reconocida como segura por la FDA de los Estados Unidos)	La actividad biológica e inmunogénica de la proteína recombinante puede diferir de la proteína natural
	Una de las bacterias más estudiadas, después de <i>E. coli</i>	<i>B. subtilis</i> produce proteasas que podrían interferir con la integridad de la proteína recombinante
	Secreta proteínas al medio de cultivo, facilitando su purificación	
	<i>B. subtilis</i> ha sido utilizado para la producción de proteínas de interés comercial, por lo que existe experiencia en el área de fermentaciones.	
	Vectores de clonación disponibles	
<b><i>Saccharomyces cerevisiae</i></b> (Levadura)	Crecimiento rápido	Genética poco entendida
	Organismo considerado como GRAS	El control de la expresión genética no es fácil
	Los parámetros para la obtención de productos de interés a gran escala ya han sido establecidos	La manipulación genética de los plásmidos de levadura es difícil
	Las fermentaciones de levaduras son relativamente baratas	Pocos vectores de clonación son disponibles
	Este sistema permite la glicosilación y la formación de puentes disulfuro	

	Solo el 0.5% de las proteínas de levadura nativa son secretadas, simplificando el aislamiento de los productos recombinantes secretados	La glicosilación de levaduras no es idéntica a la de mamíferos
<b><i>Aspergillus sp</i> (Hongo)</b>	Exite experiencia en el área de fermentaciones empleando como hospedero hongos filamentosos  Costos reducidos para el crecimiento del organismo  Es fuente natural de muchas enzimas y en algunos casos secreta las proteínas al medio de cultivo	No se obtiene un nivel elevado de proteína recombinante  Su genética no está bien caracterizada  Pocos vectores de clonación disponibles
<b>Células cultivadas de insecto Vector <i>Baculovirus</i></b>	Posee mecanismos para procesar proteínas de origen eucariótico  Sistema seguro, ya que pocos artrópodos son utilizados como su hospedero  Las propiedades funcionales y antigénicas de los productos recombinantes difieren muy poco de los naturales  La infección viral de las células cultivadas detiene la producción de la proteína del hospedero, pero el gen heterólogo es expresado a altas concentraciones	Pérdida de información durante los mecanismos de glicosilación  La proteína recombinante no siempre es 100% activa
<b>Células de mamífero</b>	Actividad biológica similar a la proteína natural  Los vectores de expresión de mamíferos están comercialmente disponibles	El crecimiento celular es difícil y lento  Proceso caro  Las células manipuladas pueden ser genéticamente inestables  La productividad de la proteína recombinante es baja

Dentro de sus características más destacadas de *B. subtilis*, se encuentra su capacidad para crecer a altas temperaturas, la ausencia de productos tóxicos y su habilidad para producir y secretar grandes cantidades de proteínas. Sin embargo, al igual que otros procariotes, las bacterias de este género no son capaces de glicosilar proteínas y consecuentemente no pueden producir proteínas funcionales de origen eucariótico (Harwood 1992). El género *Bacillus* constituye uno de los grupos más diversos de bacterias Gram-positivas, entre los cuales se

encuentran microorganismos aerobios, aerobios facultativos y anaerobios. Tienen una morfología tipo barra y son formadoras de esporas (Priest, 1993). Estas esporas pueden sobrevivir mucho tiempo, ya que son resistentes al calor, a la luz ultravioleta, la desecación y a solventes orgánicos (Slepecky, 1992).

Uno de los miembros más estudiados del género *Bacillus* es *B. subtilis*, el cual se encuentra ampliamente distribuido, habitando principalmente el suelo, aunque también se le halla en el aire, ríos y estuarios (Priest, 1989). *B. subtilis* es una bacteria grado GRAS (Generalmente Reconocida Como Segura) por la FDA de los EUA por no considerarse potencialmente patógena, en gran parte debido a que no posee lipopolisacáridos sobre su superficie celular. Posee un sistema natural de excreción muy eficiente, secretando al medio de cultivo grandes cantidades de enzimas degradativas, entre las cuales se encuentra la subtilisina (Zukowski, 1992; Pierce, 1992). Aunado a ello, en los últimos años *B. subtilis* ha sido utilizado ampliamente como un microorganismo hospedero para la producción de proteínas heterólogas, en parte debido al conocimiento genético y bioquímico que se tiene sobre él. El progreso en los procesos de fermentación ha permitido la producción de diversas proteínas heterólogas en cantidades que por otras vías, tales como fuentes naturales o microorganismos silvestres, hubiese sido más difícil de obtener (Lee, 1996). En 1989 se estimó que el mercado mundial de enzimas industriales de gran volumen era de 600 millones de dólares, de los cuales 33% eran producidas con cepas recombinantes de *B. subtilis* (Argige y Pitcher, 1989). Las principales enzimas producidas por *B. subtilis* son:  $\alpha$ -amilasa,  $\beta$ -amilasa, proteasa alcalina (subtilisina),  $\beta$ -galactosidasa,  $\beta$ -glucanasa, pentosanasas y penicilasa. Además de la producción de enzimas industriales, *B. subtilis* ha sido extensamente empleado para la producción de metabolitos primarios; tales como, hipoxantina, inosina y guanosina, que por fosforilación química son convertidos a ácido xantanílico, ácido inosínico y ácido guanílico, respectivamente. Estos últimos son de gran valor comercial en la industria alimenticia, debido a que son utilizados para mejorar el sabor (Harwood, 1992; Zukowski, 1992).

Por otro lado, *B. subtilis* ha sido también utilizado como un modelo para entender los mecanismos de diferenciación celular, ya que en este microorganismo culminan con la formación de una spora. La producción de enzimas degradativas se lleva a cabo al inicio de la esporulación; se considera que este proceso no está asociado al crecimiento de la bacteria, por lo tanto no es constitutivo (Schaeffer, 1965; Ferrari *et al.*, 1993). De manera que la inducción que desencadena la producción de enzimas degradativas, se produce de manera natural al principio de la fase estacionaria; lo cual representa una ventaja importante ya que no se requiere la adición de algún inductor exógeno; además, el cultivo en ese momento ha generado una alta cantidad de biomasa.

## II.2 Características del gen *aprE*

Entre los genes que codifican enzimas degradativas se encuentra el gen *aprE*, el cual codifica la enzima subtilisina, la proteasa mayoritaria que produce *B. subtilis* (Jacobs, 1995; Nagarajan *et al.*, 1992; Nagarajan *et al.*, 1993; Strauch y Hoch, 1992a; Valle y Ferrari, 1989; Wang *et al.*, 1988a; Wong, 1995). Esta enzima es utilizada ampliamente en la elaboración de detergentes biológicos. Para producirla a nivel industrial se han utilizado cepas sobreproductoras, reportándose concentraciones de hasta 20 g de subtilisina por litro (Zukowski, 1992). Sin embargo, dichas cepas tienen el inconveniente de presentar efectos pleiotrópicos; entre ellos, la sobreproducción de otras enzimas degradativas tales como levansacarasa, proteasa neutra,  $\beta$ -glucanasa y serin-proteasa intracelular; defectos en la síntesis de flagelo y del desarrollo de competencia; defectos en la represión catabólica normal de la esporulación por glucosa (Kunst *et al.*, 1974; Valle y Ferrari, 1989; Strauch y Hoch, 1992b).

El gen *aprE* es transcrito por  $\sigma^A$  (Park *et al.*, 1989). Estudios sobre su regulación transcripcional sugieren que se lleva a cabo a través de la acción concertada de múltiples elementos de control, como son las proteínas AbrB, Hpr, DegU/DegS y SinR, entre otras. También existe un conjunto de reguladores que afectan en menor grado la expresión de *aprE*, entre los cuales se encuentran las proteínas Pai, SenS, DegQ, DegR. Se sabe que estos factores

actúan a nivel de la transcripción del gen *aprE* y que 400 pb de su región 5' de regulación son suficientes para permitir su regulación normal (Fig. 1). Aunado al efecto de las proteínas reguladoras antes mencionadas, se sugirió que la geometría curva de esta región pudiera también afectar la expresión de *aprE* (Strauch y Ayazifar, 1995).





### **III GENERALIDADES**

Los sistemas de expresión de proteínas homólogas y heterólogas son ampliamente empleados en la obtención de proteínas recombinantes por la conveniencia en su manipulación y habilidad para alcanzar altos niveles de producción. En la actualidad el campo en la producción de proteínas recombinantes está en una etapa de continuo avance. Se han desarrollado cepas sobreproductoras en diversos microorganismos a través de mutaciones químicas y de manipulación genética. Sin embargo, en muchos casos el vasto conocimiento de la genética del microorganismo huésped ha sido explotado solo parcialmente.

En nuestro laboratorio hemos considerado a *B. subtilis* como un organismo modelo, con un potencial teórico y aplicado que aún no ha sido explorado en su totalidad. Se conocen varios de los elementos que intervienen en la sobreproducción de proteínas en *B. subtilis*. A continuación se mencionan algunos de ellos:

### **III.1 Sistema de expresión.**

Existe un gran número de sistemas para la expresión de genes heterólogos en *B. subtilis*, la selección de alguno de ellos depende de varios factores. Entiéndase por sistema de expresión aquel conjunto de elementos que intervienen en la obtención final de un producto biológico de interés. En este sentido el promotor afecta en el inicio de la transcripción; a nivel de la traducción, la secuencia SD, la disponibilidad de tRNAs, el uso preferencial de codones, la secuencia SD.

Cuando se considera la configuración de la construcción recombinante, la expresión de genes heterólogos en *B. subtilis* se clasifica como directa o a través de fusiones. Además, de acuerdo a la localización del producto recombinante puede dividirse en expresión intracelular o secreción.

La expresión directa se requiere cuando los productos tienen funciones que dependen absolutamente de una estructura primaria no-fusionada, por ejemplo, productos farmacéuticos. A diferencia de las proteínas de fusión, las cuales son fusiones indeseables, por

ejemplo aquellas generadas por falta de sitios de clonación adecuados; o bien, planeadas, lo que facilita la localización del material clonado.

### **III.2 Estabilidad del material genético.**

En *B. subtilis* el material genético que se desea expresar puede ser clonado en plásmidos de alto número de copias. Sin embargo, existen limitaciones que disminuyen su uso. Las clonas recombinantes frecuentemente tienen problemas de inestabilidad, resultando en la pérdida total o bien rearrreglos del gen de interés (Erhlinch *et al.*, 1986). Alternativamente se pueden elegir plásmidos de bajo número de copias, uno o dos copias por cromosoma (Imanaka *et al.*, 1987; Bron *et al.*, 1987). Con la finalidad de disminuir estos problemas, en *B. subtilis* se han construido cepas donde el material genético es integrado en el cromosoma. Sin embargo, en algunas ocasiones una sola copia dentro del cromosoma no alcanza los niveles transcripcionales requeridos para ser utilizados en aplicaciones biotecnológicas. Como opción para aumentar el nivel de expresión, la dosis genética se incrementa por la amplificación del gen de interés (Fahnestock y Fisher, 1986; Fahnestock *et al.*, 1986; Petit *et al.*, 1990; Kallio *et al.*, 1986).

Es interesante notar que el nivel de expresión en algunos genes de *B. subtilis* no depende estrictamente del número de copias (Mountain, 1989). Se han reportado casos donde se obtienen niveles de expresión similar cuando *B. subtilis* cuenta con dos copias de  $\alpha$ -amilasa de *B. amyloliquefaciens* en el cromosoma, en lugar de 40 copias del mismo gen clonadas en plásmido (Kallio *et al.*, 1986). Para cada gen blanco el efecto del número de copias debe determinarse experimentalmente.

### **III.3 Diseño y construcción de una unidad de expresión**

Una unidad básica para la expresión de genes heterólogos en *B. subtilis* está compuesta usualmente de los siguientes componentes: promotor, sitio de unión a ribosoma o secuencia de Shine-Dalgarno (SD), péptido señal cuando se requiera exportar la proteína recombinante, secuencia codificante del gen de interés. En algunos casos para facilitar la localización de la

proteína se inserta una secuencia en el extremo amino o carboxilo terminal del gen heterólogo; ejemplos de ello es la localización de proteínas en: cultivos celulares (Ludin *et al.* 1996), cromosoma de levaduras (Bahler *et al.*) y plantas (Lindsey *et al.* 1993). En ocasiones también se incluye una secuencia de término de la transcripción localizada río abajo de la secuencia codificante. El terminador transcripcional sirve como señal de término y como elemento protector evitando la degradación del mRNA (Gold y Stormo, 1990; Hall *et al.*, 1982). Los elementos citados anteriormente tienen impacto directo sobre la eficiencia de la expresión genética.

### III.4 Factores a nivel transcripcional

La transcripción es el paso primario en la expresión genética. Puede ser dividida en tres etapas: (1) iniciación, (2) elongación y (3) terminación.

En la mayoría de los casos la expresión genética en procariontes es controlada en el inicio de la transcripción por la holoenzima, formada por un factor sigma y el complejo de la RNA polimerasa (RNAP). Este proceso es la etapa limitante en la síntesis de mRNA (deHaseth *et al.*, 1998). El inicio de la transcripción es controlada por secuencias promotoras, las cuales pueden ser moduladas por una variedad de mecanismos, principalmente la interacción de una o más proteínas reguladoras con secuencias específicas en la vecindad del promotor. En *B. subtilis* la interacción holoenzima-DNA es muy específica y en algunos casos la transcripción a partir de un promotor heterólogo por la holoenzima de *B. subtilis* es ineficiente (Doi, 1984; Doi y Wang, 1986). Una forma de expresar eficientemente un gen heterólogo en *B. subtilis* es reemplazar su región promotora por un promotor de *B. subtilis* o un promotor compatible a la RNAP de *B. subtilis*. Se han empleado varios promotores nativos junto con promotores heterólogos para la expresión de genes heterólogos en *B. subtilis*; entre ellos: promotores fuertes de la fase vegetativa: *vegII* (Peschke *et al.*, 1985), *spoI* (Yansura y Henner, 1984), promotores regulados temporalmente: el de subtilisina (*aprE*) y  $\alpha$ -amilasa (Park, 1989); promotores inducibles: el de triptófano, arginina (Shimotsu and Henner, 1986), sacarosa, xilosa, arabinosa y gluconato (Crutz *et al.*, 1990; Gartner *et al.*, 1988), levansacarasa (Crutz *et al.*, 1990). Por tanto, es necesario decidir bajo que condiciones el gen homólogo o

heterólogo quiere ser expresado, antes de la selección del promotor (Moran, 1989). Las secuencias consenso de los promotores dependientes de cada tipo de holoenzima de *B. subtilis* se muestran en la Tabla 2

**Tabla 2. Secuencias consenso de los promotores de *B. subtilis***

Factor Sigma	Funciones	Secuencias Promotores Consenso	
		-35	-10
$\sigma^A$	Generales Básicas	TTGACA	TATAAT
$\sigma^B$	Respuesta a condiciones de estrés	AGGNTT	GGNATTGNT
$\sigma^C$	Desconocidas	AAATC	TANTGNTNTA
$\sigma^D$	Síntesis de flagelo	CTAAA	CCGATAT
$\sigma^E$	Fase de esporulación II Específico de la madre	GAANAANT	CATATT
$\sigma^F$	Fase de esporulación II Específico de la ferospora	Similar a $\sigma^G$	
$\sigma^G$	Fase de esporulación III Específico de la ferospora	YGHATR	CAHWHTAH
$\sigma^H$	Fase de esporulación 0	CAGGA	GAATTWWT
$\sigma^K$	Fase de esporulación Específico de la célula madre	Desconocido	

Se exhibe la cadena que no transcribe. 5' a 3' de izquierda a derecha. H=A, C o T; R=A o G; W=A o T; Y=C o T. Referencia de las secuencias consensos:  $\sigma^A$ , Moran *et al.* (1982);  $\sigma^B$ , Tatti y Moran (1984);  $\sigma^C$ , Johnson *et al.* (1983);  $\sigma^D$ , Gilman *et al.* (1981);  $\sigma^E$ , Roel *et al.* (1992);  $\sigma^F$ , Sun *et al.* (1991);  $\sigma^G$ , Nicholson *et al.* (1989);  $\sigma^H$ , Tatti (1989);  $\sigma^K$  Foulger y Errington (1991);  $\sigma^I$  Debarbouille *et al.* (1991).

Los factores sigma hacen contactos de tipo secuencia-específico con las regiones -10 y -35 de los promotores de *B. subtilis*. Además, se ha establecido que además de las regiones -10 y -35 de los promotores reconocidos por  $\sigma^{70}$  de *E. coli* y  $\sigma^A$  de *B. subtilis*, existe en algunos

casos una región rica en As y Ts localizada aproximadamente a 20 pb hacia arriba de la región -35. Este tercer elemento se considera importante para la interacción de la subunidad  $\alpha$  de la RNAP. Se puede considerar que la fuerza del promotor es una función de los tres elementos: la región -10, la -35 y la secuencia rica en As y Ts. Al conjunto de ellas se le ha llamado promotor extendido (Busby y Ebright, 1994; Ross *et al.*, 1993; Helmann, 1995). Estudios posteriores proponen en *E. coli* que la secuencia rica en As y Ts tiene el consenso: -59 nnAAA(A/T)(A/T)T(A/T)TTTTnnAAAAnnn-38 (Ross *et al.*, 1998; Estrem *et al.*, 1998).

### **Término de la transcripción**

Una vez que la transcripción se inicia a partir del promotor, el mRNA se sintetiza y concluye en sitios específicos. El término eficiente de la transcripción determina una expresión genética favorable; la transcripción no controlada puede ser nociva, especialmente cuando el gen a expresar está presente en un plásmido de alto número de copias o cuando se expresa a partir de un promotor fuerte. Así mismo se asegura que hayan suficientes moléculas de RNAP para la expresión de los genes de mantenimiento de la bacteria, previniendo el secuestro de todas las moléculas de RNAP para la expresión del gen heterólogo (Wong y Chang, 1986). Dependiendo de la localización cromosomal y orientación de la integración, la transcripción de escape podría también tener un papel central en determinar la estabilidad y nivel de expresión de los genes recombinantes integrados en el cromosoma del hospedero. La incorporación de un terminador transcripcional eficiente es, por tanto, altamente recomendada en la construcción y diseño de sistemas de expresión.

Una señal de término de la transcripción parecida a la señal de terminación p-independiente de *E. coli* se ha observado en *B. subtilis*. Esta señal es usualmente una región del tipo "tallo y asa" seguida por una corrida de seis residuos de timina. Estas señales han sido observadas en los extremos de los genes y operones, y también en señales de atenuación localizadas entre promotores y marcos abiertos de lectura de genes u operones. Aunado a lo antes mencionado, la secuencia terminadora aumenta la vida media del mRNA, evitando la degradación por exonucleasas en el extremo 3' (Carter *et al.*, 1988). Wang *et al.* han

identificado terminadores fuertes de la transcripción en *B. subtilis*, dentro de ellos se encuentran el del gen *senS* (Wang *et al.*, 1988b) y en *Bacillus thuringiensis* el del gen *cryIIIA* (Wong y Chang, 1986)

### **III.5 Factores a nivel traduccional**

La expresión de un gen depende de un nivel elevado de transcripción, así como también de una traducción eficiente. La maquinaria traduccional de *B. subtilis* es altamente específica, lo que representa una barrera para la mayoría de genes heterólogos expresados en este microorganismo. En términos generales, los elementos del mRNA que condicionan el inicio de la traducción son: 1) codón de inicio preferencial AUG, aunque con menor frecuencia GUG, UUG, AUU y AUA; 2) la secuencia de unión al ribosoma o secuencia Shine-Dalgarno (SD) tiene al menos 4 nucleótidos de la secuencia AAGGAGG; 3) el espacio entre la secuencia SD y el AUG es de 9+/-3 nucleótidos; 4) al lado de la secuencia SD, los nucleótidos localizados en el extremo 5' del AUG deben ser As y Us; 5) la secuencia GCAU o AAAA localizada después del AUG mejora la traducción; 6) la región alrededor del sitio de inicio debe ser poco estructurada (Revisado en Vellanoweth, 1993).

### **III.6 Uso preferencial de codones**

En términos generales el uso preferencial de codones es otro parámetro a considerar. Los genes expresados débilmente se caracterizan por la presencia de codones reconocidos por especies raras de tRNA. Sin embargo *B. subtilis* presenta la ventaja de no tener una tendencia marcada en su uso de codones (Kunisawa, 1995).

### **III.7 Degradación proteica**

Pese a las ventajas que ofrece *B. subtilis*, su empleo en la producción de proteínas recombinantes aún presenta un problema, el causado por el gran número de proteasas producidas por el microorganismo que degradan al producto de interés. Dentro de estas proteasas se encuentran las de tipo extracelular como son AprE, Npr, Epr, Bpf y Mpr y de



tipo intracelular, siendo la más importante IspI. Las proteasas AprE y Npr son producidas mayoritariamente, representan cerca del 70% de la concentración total de proteasas extracelulares producidas por *B. subtilis*. Con la finalidad de evitar la degradación proteica del producto de interés se han empleado mutantes que producen cantidades mínimas de proteasas (He *et al.*, 1991).

### **III.8 Ciclo celular**

Por ser de particular interés el periodo en el cual se lleva a cabo la expresión del gen *aprE* durante el crecimiento de la bacteria, a continuación se presenta un breve resumen de su ciclo celular y los cambios que sufre a lo largo del mismo.

El ciclo celular de *B. subtilis* consiste en tres etapas fácilmente discernibles: crecimiento exponencial, estado de transición y fase estacionaria.

#### **III.8.1 Crecimiento Exponencial**

Cuando las bacterias se encuentran en condiciones medio-ambientales óptimas y en un medio de cultivo que les provea de todos los nutrientes necesarios para reproducirse, tienden a dividirse a una velocidad máxima. A este tipo de desarrollo se le llama crecimiento exponencial y es característico de cada bacteria. *B. subtilis* es capaz de dividirse cada 20 minutos bajo condiciones controladas; mientras que en vida libre su tiempo de generación varía entre 50 y 100 horas (Harwood y Archibald, 1990).

#### **III.8.2 Estado de transición**

El estado de transición es la etapa entre el crecimiento exponencial y el inicio de la fase estacionaria. En este periodo la célula monitorea condiciones de estrés. En *B. subtilis*, así como en algunas otras bacterias, hongos y eucariotes inferiores, una respuesta a dicho estrés acumulado es la formación de una espora latente (Setlow, 1994; Strauch y Hoch, 1992b). Además de las proteínas necesarias para el inicio de este complejo programa de desarrollo; en

este momento, se sintetizan otras proteínas no esenciales para la esporulación (Kunst *et al.*, 1994). Muchas de estas proteínas son enzimas degradativas, tales como levansacarasa, alfa-amilasa y proteasas, que permiten a la célula utilizar todas las fuentes disponibles de energía y metabolitos en el medio circundante. Concomitante a la síntesis de enzimas degradativas, otras enzimas se expresan en este intervalo e incluyen aquéllas relacionadas con el desarrollo de competencia para la asimilación de moléculas de DNA, la producción de antibióticos, la síntesis de flagelo y la inducción total del ciclo de Krebs, el cual normalmente no funciona en su totalidad durante el crecimiento vegetativo de este microorganismo (Strauch y Hoch, 1992a).

En *B. subtilis* se conocen al menos nueve genes, *abrB*, *hpr*, *sinR*, *degU*, *degS*, *degR*, *degQ*, *senS* y *pai*, cuyos productos peptídicos producen efectos pleiotrópicos sobre la síntesis de muchas de las proteínas y de las funciones antes mencionadas (Kunst *et al.*, 1994; Strauch y Hoch, 1992a). Debido que dichas proteínas controlan funciones que son expresadas durante el estado de transición se ha acuñado el término de **reguladores del estado de transición** para describirlos, Fig. 2 (Perego *et al.*, 1988).

Ninguna de las mutaciones reportadas en los reguladores del estado de transición conducen a defectos significativos en la esporulación. Sin embargo, existe evidencia que algunos de ellos son controlados por genes esenciales para la esporulación. Es decir, el control de la síntesis de enzimas degradativas y otros procesos del estado de transición se sobrelapan con la red específica para la esporulación y puede ser, de hecho, una parte de ella. Con objeto de comprender completamente cómo operan estos reguladores y cómo son integrados con la vía de esporulación, se ha vuelto necesario definir las funciones bioquímicas de los productos de los genes reguladores y como interactúan uno con otro. Dado que la mayoría de los reguladores del estado de transición que se conocen afectan la expresión del gen *aprE*, en este apartado se describirá las interacciones regulatorias involucradas en su expresión genética.

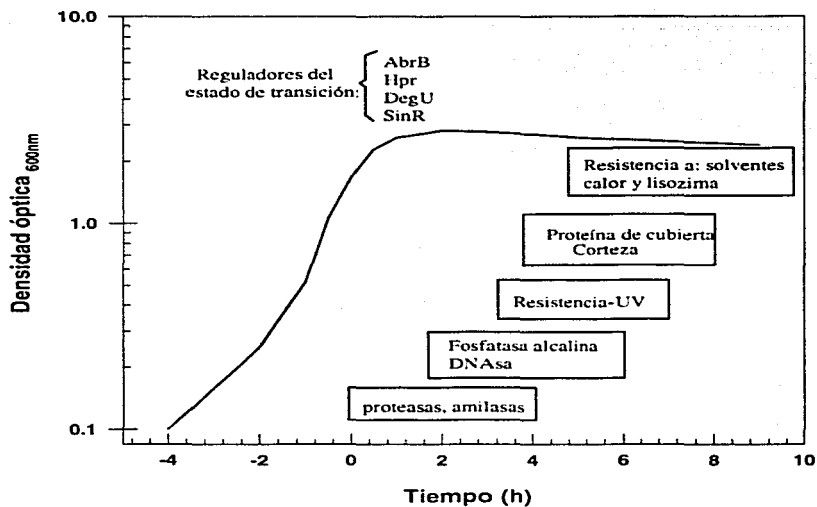


Fig. 2 Eventos que se producen durante el crecimiento de *Bacillus subtilis*. El tiempo está expresado en horas. T0 corresponde al inio de la fase estacionaria en cultivos utilizando medio Shaeffer (Shaeffer, 1965).

## Proteína AbrB

El producto del gen *abrB* funciona como un regulador, con efectos positivos y negativos sobre ciertos genes que son expresados durante el estado de transición. En el caso del gen *aprE*, su efecto es el de reprimir su transcripción. El gen *abrB* codifica para una proteína hexamérica y posee un peso molecular de 10,500 Da (FürbaB y Marahiel, 1991). Mediante un estudio de protección de AbrB al DNA se ha sugerido que la proteína puede interferir de alguna manera con la interacción entre la RNA polimerasa y los promotores que afecta (Robertson *et al.*, 1989). No se ha podido detectar alguna secuencia consenso para la unión de AbrB; sin embargo cada una de las regiones protegidas es relativamente rica en AT, por lo que se cree que la unión se lleva a cabo sobre una estructura secundaria específica de DNA (Fig. 4). Se ha señalado un efecto regulador positivo de AbrB sobre el gen *hpr*, ver Fig. 3 (Strauch *et al.*, 1989).

## Proteína Hpr

El locus *hpr* fue definido por mutaciones que causaban sobreproducción de las proteasas neutra y subtilisina (Higerd *et al.*, 1972). El gen *hpr* ha sido clonado y secuenciado, y la proteína purificada a homogeneidad. Mediante experimentos de protección con DNasa I a la región 5' de control de *aprE* se encontró que Hpr se une a cuatro regiones separadas, dos sitios muy cercanos al promotor y otros dos sitios en la región que lo precede, aproximadamente 200 pb hacia arriba. Dos de estos sitios están localizados muy lejos del promotor y cercanos uno del otro: -324 a -295 y -292 a -267 (Fig. 4). Los otros dos sitios están también cercanos uno del otro, pero localizados muy cerca y superlapano con el promotor: -79 a -59 y -35 a -14. El análisis de todas las regiones protegidas ha revelado lo que parece ser la secuencia consenso de Hpr: ATANTATT. También se ha revelado que uno de los sitios de unión de Hpr, el que comprende la región de -35 a -14, está ubicado completamente dentro de la región a la cual se une la proteína AbrB (-59 a +15). Por la técnica de "footprinting" y en experimentos donde se mezclan Hpr y AbrB, se determinó que

ambas moléculas se unen a la misma región de DNA (Strauch y Hoch, 1992a). La presencia de dos series de sitios de unión de Hpr separados por distancias relativamente grandes (183 pb) sugiere que Hpr pudiera funcionar a través de la formación de una horquilla y que la represión ocurra por la existencia de esa estructura. Si bien Hpr es claramente un regulador negativo de la transcripción de *aprE*, éste no es un represor clásico ya que en su ausencia la subtilisina no es expresada constitutivamente (Dod y Balassa, 1973; Ito y Spizizen, 1971).

### **Proteína SinR**

El gen *sinR* clonado en un plásmido multicopia inhibe la esporulación y la producción de proteasas extracelulares (Gaur *et al.*, 1986). Este gen forma parte de un operón dicistrónico y es precedido por el gen *sinI*, el cual codifica para la proteína SinI constituida por 57 aminoácidos. La función de SinI es inactivar a la proteína SinR a través de interacciones proteína-proteína durante la etapa final de T<sub>2</sub> (Bai *et al.*, 1992). El sitio de unión de SinR a la región de regulación de *aprE* abarca de la posición -263 a -216 respecto al inicio de la transcripción, ver Fig. 4 (Gaur *et al.*, 1991; Strauch y Hoch, 1992a).

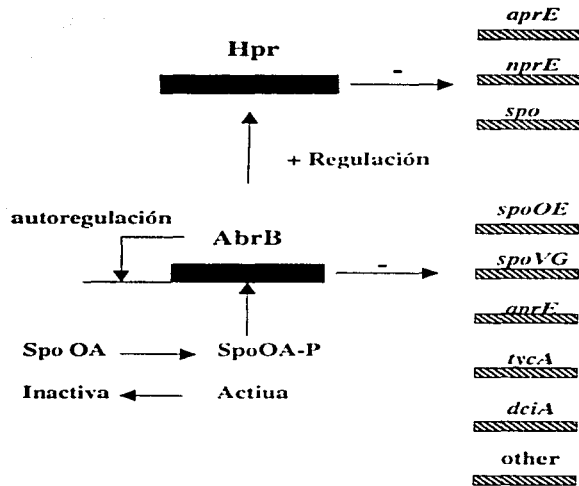


Fig. 3 Regulación de la proteína AbrB en el estado de transición. La regulación positiva se indica con (+) y la regulación negativa con (-). Además, *abrB* se autoregula. La conversión de la proteína SpoOA a su forma fosforilada (activa), reprime su transcripción y ocurre al final del crecimiento exponencial.

-400

GGCGGCCGCA TCTGATGTCT TTGCTTGGCG AATGTTTCATC TTATTTCTTC CTCCTCTCA

HPR (1) -300

ATAATTTTT CATTCTATCC CTTTCTGTA AGTTTATTTCAGAATAC TTTATCATC

HPR (2) SINR

ATG GTGGAA AAATATCAG GATAATATCC AT TCGTTCTATTCAGG TAAAGG TTTAAAG

-200

TAA GGAATTGCC GGGACTCAGG AGCATTTAAC CT AAAAAAGC

DEGU/Q

ATGACATTT CAG CATAATGA ACATTTACTC ATGTCTATTT TCGTTCTTTT CTGTATGAAA

-100 HPR (3)

ATAGTTATT CGAGTCTCTA CGGAAATAGC GAGAGA TGATATACC TAAATAGAGATA AAAA

-35 HPR (4) -10 +1

TCATCT CAAA AAAATGGG TCT TAAATATCCATCTA TTACAATA AATTCACAGA

ABRB

ATAGTCTTTT AAGTAAGTCT ACTCTGAATT TTTTAAAAG GAGAGGCTAA AGAGTGAGAA

Met

**Fig.4** Secuencia nucleotídica de la región de regulación del gen *aprE*. Los sitios de unión de los reguladores transcripcionales están indicados en rectángulos: verde (sitio de unión de Hpr); rojo (sitio de unión de SinR), amarillo (sitio de unión de DegU). La región protegida por AbrB se indica con la línea subrayada. Asimismo, se exhiben las cajas -10 y -35 de la secuencia promotora y el inicio de la traducción (Met). La numeración de la secuencia se realizó en base al inicio de la transcripción (+1).

## Proteínas Deg

Los productos de los genes *degS* y *degU* están involucrados en activar la transcripción de *aprE* durante el estado de transición (Henner *et al.*, 1988a; Msadek *et al.*, 1990). Las mutaciones en *degS* y *degU* afectan pleiotrópicamente la producción de otras enzimas degradativas y otros procesos involucrados en el estado de transición (Strauch y Hoch, 1992a; Msadek *et al.*, 1990). El análisis de la secuencia de aminoácidos de DegS y DegU reveló que forman parte de una familia muy conservada de proteínas procarióticas. Esta familia cuenta con un sistema que transmite señales por medio de dos componentes (Henner *et al.*, 1988b). DegS es una cinasa que se autofosforila en respuesta a señales ambientales específicas y que posteriormente fosforila al segundo componente, la proteína DegU, la cual a su vez actúa como proteína reguladora uniéndose al DNA (Henner *et al.*, 1988a).

De acuerdo a las observaciones mencionadas y analizando la naturaleza molecular de varias mutaciones en *degU* y *degS*, ha sido posible plantear la siguiente hipótesis: la forma fosforilada de DegU tiene dos funciones: a) regulación positiva sobre la expresión de algunas enzimas degradativas y b) regulación negativa en el desarrollo de competencia y movilidad (Msadek *et al.*, 1990). Por otra parte, un análisis de eliminación enfocado a localizar los sitios de unión de DegU en la región de regulación del gen *aprE* ( $\pi aprE$ ), sugirió que el fragmento de DNA involucrado abarcaba desde la posición -164 a la -141 respecto al inicio de la transcripción (Henner *et al.*, 1988a). Sin embargo, en experimentos de protección *in vitro* utilizando DegU y la región de regulación de *aprE* no se localizó el sitio de unión para esta proteína. Contrario a estos resultados, un análisis posterior basado en el alineamiento de regiones promotoras de genes regulados por DegS y DegU, se infirió que la secuencia consenso para el control de DegU es AGAAN<sub>n</sub>TTCAG (Dartois, *et al.* 1998); en *aprE* esta secuencia se ubica entre las posiciones -188 a -169 en relación al inicio de la transcripción (Fig. 4).



## Proteínas *Pai* y *SenS*

Además de los reguladores antes mencionados, existen dos genes que codifican proteínas involucradas en la regulación de la expresión génica en el estado de transición. La presencia del operón *pai* en un plásmido multicopia resulta en la inhibición de  $\alpha$ -amilasa, levansacarasa, proteasas neutra y alcalina; así como inhibición de la esporulación. El operón *pai* codifica para dos proteínas necesarias para el efecto multicopia. La eliminación del ORF2 del operón *pai* es letal; mientras que la eliminación del ORF1 no tiene efecto sobre la producción de enzimas degradativas, pero libera la represión por la acción de la glucosa sobre la esporulación (Honjo *et al.*, 1990).

El gen monocistrónico *senS*, cuando está presente en multicopia, tiene efectos opuestos a los vistos con *pai*: estimula la transcripción de los genes que codifican para subtilisina, proteasa neutra, alfa-amilasa y fosfatasa alcalina (Wang *et al.*, 1988b; Wang y Doi, 1990). Así mientras que *pai* codifica para un regulador negativo de subtilisina, *senS* codifica un activador. El conocimiento que se tiene sobre ellos es aún limitado.

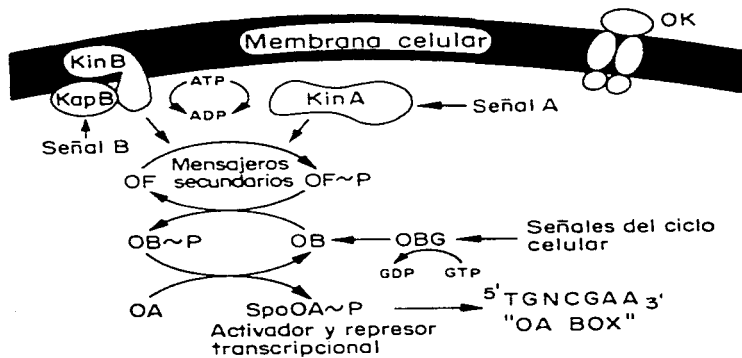
### III.8.3 Proceso de esporulación

La esporulación en *B. subtilis* es parte de un complejo programa de desarrollo y diferenciación celular que tiene como resultado la formación de una estructura altamente resistente denominada espora. Este proceso se ha dividido en diferentes estadios. La formación de esta estructura es un ejemplo extremo de la adaptabilidad de *B. subtilis* al medio e involucra a más de 80 genes (Kunkel, 1991). En el inicio de la esporulación la proteína SpoOA, la cual funciona como represor y activador, juega un papel central. El gen *spoOA* es transcrito a partir de dos promotores, P<sub>v</sub> y P<sub>s</sub>. P<sub>v</sub> es activado por el factor de transcripción E $\sigma^A$ , lo cual proporciona un nivel basal de síntesis de SpoOA en células en crecimiento

vegetativo. Ps es regulado por  $E\sigma^H$ , el cual produce un incremento súbito en la síntesis de SpoOA al inicio de la esporulación (Chibazakura *et al.*, 1991; Yamashita *et al.*, 1989).

La proteína SpoOA pertenece a la familia de reguladores bacterianos de dos componentes (Stock *et al.*, 1989). La fosforilación de SpoOA es la clave en el control del inicio de la esporulación. El fosfato es transferido a SpoOA por un mecanismo complejo que involucra a proteínas histidín-cinasas y al menos a dos acarreadores de fosfato que actúan como intermediarios. Las interacciones entre los distintos participantes han sido reproducidas *in vitro* (Burbulys *et al.*, 1991). Al menos dos proteínas histidín-cinasas, codificadas por los genes *kinA* y *kinB*, pueden transferir el fosfato a la proteína SpoOF, la cual a su vez transfiere el fosfato a un segundo acarreador llamado SpoOB. La proteína SpoOB es el donador directo del fosfato a SpoOA. Esta vía puede permitir que mediante sensores ambientales, tanto internos como externos, se controle el nivel de SpoOA fosforilada de una manera rápida y precisa (Grossman, 1991). Congruente con esta idea, se ha demostrado que las cinasas KinA y KinB pueden responder a diferentes señales, ver Fig. 5 (LeDeaux, 1995).

SpoOA-P regula negativamente algunos genes, siendo el más importante *abrB*, uno de los represores de *aprE* (Perego *et al.*, 1988; Strauch *et al.*, 1990). Por otra parte, la función positiva de SpoOA-P es responsable de la activación de ciertos genes específicos de la esporulación, como son los operones *spolIA* (Perego *et al.*, 1991), *spolIG* (Satola *et al.*, 1991; Satola *et al.*, 1992), *spolIE* (York *et al.*, 1992) y el gen *sinI* (Bai *et al.*, 1993). Se sabe que algunas mutaciones en *spoOA* bloquean la esporulación en su fase inicial y eliminan la producción de proteasas extracelulares, entre ellas subtilisina (Strauch y Hoch, 1992a). Otras mutaciones, como la denominada *spoOA9V*, bloquean la esporulación en el estadio II. Sin embargo, al parecer esta mutación produce niveles normales de subtilisina (Perego *et al.*, 1991).



**Fig. 5 Transducción del fósforo.** Diversas señales metabólicas, ambientales y de densidad celular desencadenan la transmisión del grupo fosfato, para dar inicio a la esporulación. Al final de este proceso, la proteína SpoOA queda fosforilada (activa), con el potencial para activar o reprimir diversos genes.

El inicio de la esporulación es complejo y se conocen algunas señales que se interrelacionan y dan inicio a ésta:

**a) Señales nutricionales**

La limitación de fuentes de carbono, nitrógeno o fósforo puede inducir la esporulación (Freese, 1981; Sonenshein, 1989). A su vez, buenas fuentes de carbono, como la glucosa, la reprimen (López *et al.*, 1980). Se ha demostrado que la concentración intracelular de GTP (o GDP) probablemente representa el efector clave de la señal nutricional (López *et al.*, 1979; López *et al.*, 1981). Sin embargo, aún se desconoce cómo la célula monitorea la concentración de GTP y cómo la señal se transmite al aparato transcripcional.

**b) Densidad de población**

Grossman y Losick (1988) observaron que la esporulación no puede inducirse eficientemente en células mantenidas a baja densidad de población. Sin embargo, la esporulación eficiente se obtiene cuando las células se resuspenden a baja densidad en un medio estandarizado para el crecimiento de células a alta densidad. El efecto no se debió al agotamiento de un factor esencial de crecimiento sino a la producción de sustancias, posiblemente oligopéptidos, que son necesarios para inducir una esporulación eficiente (Waldburger *et al.*, 1993). Se conoce poco acerca de la naturaleza de dichos factores y, menos aún, de cómo su presencia es monitoreada e interpretada.

## **IV JUSTIFICACION**

Pese a la importancia básica y aplicada del gen *aprE* y que su sitio de inicio de la transcripción ha sido determinado (Ferrari *et al.*, 1988), no hay una descripción precisa de su secuencia promotora. Consideramos que para posteriores estudios es necesario conocer la posición exacta de los elementos de su promotor, por lo que parte de este trabajo fue enfocado a este fin.

Aunado a ello, en nuestro laboratorio se han construido cepas de *B. subtilis* que presentan un nivel elevado de producción de  $\beta$ -galactosidasa recombinante, cuya expresión es controlada por la región de regulación del gen *aprE*, codificada cromosomalmente y en la que no requiere de la adición de inductores exógenos para su expresión (Bolaños, 1994). Sin embargo, consideramos que estas cepas pueden ser mejoradas si se toma en cuenta de manera conjunta el conocimiento genético relevante de procesos en *B. subtilis*, como son: la transcripción, la traducción, la estabilidad del material genético y la estabilidad del mRNA, que habían sido estudiadas previamente de manera aislada.

## **V. OBJETIVOS**

## OBJETIVOS GENERALES

1. Definir y caracterizar la secuencia promotora del gen *aprE*.
2. Diseñar y construir integralmente cepas sobreproductoras de proteínas heterólogas de *Bacillus subtilis*, empleando el conocimiento de los procesos genéticos que intervienen en la sobreproducción de proteínas.

## OBJETIVOS ESPECIFICOS

1. **Caracterización del promotor funcional del gen *aprE* mediante:**
  - 1.1 Programas de cómputo.
  - 1.2 Mutagénesis sitio específico sobre bases potencialmente relevantes.
2. **Diseño y construcción de cepas sobreproductoras de *B. subtilis***

Evaluar distintas condiciones que modulan el nivel de producción de proteínas heterólogas, usando como modelo la fusión traduccional *raprE'-lacZ*. Los factores considerados son:

### 2.1 Nivel transcripcional

- 2.1.1 Construcción del promotor canónico, que incluya:
  - 2.1.1.1 La caja -35 (TTGACA) canónica.
  - 2.1.1.2 La separación entre las cajas -10 y -35 de 17 pb.
  - 2.1.1.3 La región de reconocimiento para la subunidad  $\alpha$  de la RNAP.
- 2.1.2 Adición del terminador de la transcripción del gen *cryIIIA* de *Bacillus thuringiensis*.



## 2.2 Nivel traduccional

- 2.2.1 Construcción de la secuencia Shine-Dalgarno canónica AAGGAGG
- 2.2.1 Mutación del triplete GTG (codón de inicio de la traducción de *aprE*) por ATG.

## 2.3 Estabilidad del material genético:

- 2.3.1 Integración en el cromosoma de las construcciones realizadas en la *r<sub>aprE</sub>*

## 2.4 Efecto de los fondos genéticos

- 2.4.1 Transferir los genotipos *hpr<sup>-</sup>*, *degU<sup>32</sup>* a las cepas con las construcciones realizadas en la *r<sub>aprE</sub>*.

## **VI.1 MANUSCRITO 1:**

**“ Characterization of the 5' subtilisin (*aprE*) regulatory  
region from *Bacillus subtilis* ”**

**FEMS Microbiology Letters. (2000) 183: 9-14**

## Characterization of the 5' subtilisin (*aprE*) regulatory region from *Bacillus subtilis*

Janet Jan <sup>a</sup>, Fernando Valle <sup>b</sup>, Francisco Bolivar <sup>a</sup>, Enrique Merino <sup>a,\*</sup>

<sup>a</sup> Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62210, México  
<sup>b</sup> Genética Internacional Inc., Fairfax, VA 22034, USA

Received 14 October 1999; received in revised form 17 November 1999; accepted 22 November 1999

### Abstract

The *aprE* gene of *Bacillus subtilis* encodes the major serine alkaline protease known as subtilisin. It is expressed during the transition state and transcribed by the  $\sigma^A$  form of the RNA polymerase (RNAP). In this work, we characterized the regulatory region of the *aprE* gene (*rraprE*) from *B. subtilis*. By computer analysis and site-directed mutagenesis, we localized the *aprE* promoter sequence 7 bp upstream from its transcription initiation site (TIS). We also characterized the static curvature properties of the *rraprE* DNA and found two different areas of DNA bending, within the first 400 bp upstream of its TIS. We postulate that these particular curved DNA regions could play a role in the interaction with some regulatory proteins and discuss possible implications related to *aprE* transcription regulation. © 2000 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

**Keywords:** *aprE*; Curved DNA; Site-directed mutagenesis; Promoter sequence; *Bacillus subtilis*

### 1. Introduction

The *aprE* gene encodes the extracellular alkaline protease subtilisin from *Bacillus subtilis*. It is expressed during the stationary phase and reaches its highest expression level approximately 2 h after the onset. It has been established that AprE contributes to more than 70% of the total extracellular protease activity [22]. The transcription of *aprE* gene has been used as a model to understand the regulatory network that controls the gene expression at the transition state, when the cell ceases exponential growth and sporulation is initiated [16]. Alkaline protease production, as well as other extracellular proteases, are subject to regulation by a complex group of positive and negative regulators, such as: DegU/DegS, AbrB, Hpr and SinR. These proteins act at the transcriptional level and require at least 400 bp upstream of the *aprE* transcription initiation site (TIS) to present their normal mode of control [20].

Although the basic and applied importance of the *aprE* gene is widely accepted, and its TIS has been determined [4], there is no clear description of its promoter sequence. Previous studies by Park et al. [11] reported a putative *aprE* promoter 7 bp upstream of its TIS, that had 19 bp between their -10 and -35 boxes. Conflicting with these data, Strauch and Hoch [18] suggested a different promoter region, 12 bp upstream of its TIS and 16 bp between their -10 and -35 boxes. Taking into account these conflicting data, and the lack of experimental results to evaluate the relevant nucleotides (nt) for RNA polymerase (RNAP) recognition, in this work, we performed site-directed mutagenesis in the regulatory region of the *aprE* gene (*rraprE*) to determine the functional promoter sequence. Our results indicate that the functional *aprE* promoter does not correspond to any of the previous reported promoter sequences, at least under the growth conditions assayed in this study.

In addition, the curved DNA geometry of the *rraprE* has been proposed to be involved in *aprE* gene expression. The presence of a static curvature within this region has previously been reported [17]; however, there are no experimental studies that characterize the exact position of the region(s) involved in such curvature. Using electrophoretic analyses of circular permuted DNA fragments containing *rraprE*, as well as curvature analysis based on com-

\* Corresponding author: Apdo. Postal 510-3, Cuernavaca, Morelos CP 62250, Mexico. Tel.: +52 (73) 29-16-01; Fax: +52 (73) 17-23-88; E-mail: merino@ibt.unam.mx

puter programs, we have found that the major determinants of anomalous gel migration are located around 100 and 200 bp away from its TIS. The results of these studies are discussed in the light of the *aprE* transcription regulation.

## 2. Materials and methods

### 2.1. Computer analysis directed to identify putative $\sigma^N$ promoters

Based on the statistical algorithm proposed by Mulligan et al. [10], we developed a program to search  $\sigma^N$  promoter sequences. In this algorithm, each base of the  $-35$  and  $-10$  boxes as well as the ones near to them are weighed according to a score matrix. This score matrix was built up from the frequency of occurrence of each base on a list of 122 *B. subtilis*  $\sigma^N$  promoters compiled by Helmann [7]. The space between the  $-35$  and  $-10$  regions contributes in geometrical terms to the final score. The maximum possible score of a putative perfect promoter is 100. A cut-off value was used to select only the best promoters within a DNA region.

### 2.2. Computer analysis of the DNA static curvature at the *rraprE*

To visualize the DNA three-dimensional trajectory, we used the DNASTAR computer program (DNASTAR, Madison, WI, USA), which is based on the nearest-neighbor wedge model [3]. The analysis was made with 536 bp of the *rraprE* sequence previously reported by Valle and Ferrari [20].

### 2.3. Site-directed mutagenesis of *aprE* promoter and plasmid construction

Plasmid pT7-*aprE* was constructed by subcloning a 509-bp *EcoRI*-*Bam*HI fragment derived from plasmid pSG35.1 [4] into pT7 Blue (Novagen, Madison, WI, USA) and contains the *aprE* promoter and its first eight codons. pT7-*aprE* was used as a template for oligonucleotide-directed PCR mutagenesis according to the protocol described by Merino et al. [9]. PCRs were carried out with *Taq* DNA polymerase (Promega, Madison, WI, USA). The nt substitutions were as follows:  $-7$  (T  $\rightarrow$  G),  $-31$  (A  $\rightarrow$  C),  $-37$  (T  $\rightarrow$  G) (see Fig. 1). The PCR products were double-digested with *EcoRI* and *Bam*HI, cloned into the same restriction sites of pSG-PLK, and verified by DNA sequencing [14] to yield plasmids p-7G, p-31C and p-37G, respectively. pSG-PLK is a pSG35.1 derivative that carries a promoterless *lacZ* gene. In this plasmid, the *aprE* promoter has been replaced with the pUC19 poly-linker region. This modification makes the selection of transformants easier, because colonies that carry the

mutated *aprE* promoter are blue when they are plated on X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside) agar.

### 2.4. Electrophoretic analysis of circular permuted DNA fragments

The sequences of the primers used for the PCR analysis to map the curved loci are shown in Table 1. We utilized a PCR-based method [21] to construct circular permutations of the *rraprE*. These fragments were analyzed by 8% polyacrylamide gel electrophoresis (PAGE) at 4°C at a constant voltage of 8 V  $\text{cm}^{-1}$  in 89 mM Tris-borate (pH 8.3), 2 mM EDTA buffer.

### 2.5. Bacterial strains and transformations

*Escherichia coli* JM101 [*lac-pro*] *supE* *thiE* *traD30* *pro* (*B*) *lacI* *q* *Z* (*M15*)] was used as the host for plasmid constructions. Plasmids were obtained as described in [2]. *B. subtilis* strains constructed in this study were derived from BSRI [5]. Plasmids p-7G, p-31C and p-37G were linearized with *Pvu*I and used to transform *B. subtilis* BSRI, to construct the strains listed in Table 2. The transformation protocol was followed according to Anagnostopoulos and Spizizen [1]; the selection was made for chloramphenicol-resistance (5  $\mu\text{g ml}^{-1}$ ). Integration of plasmid DNA was confirmed by testing the amylase production by iodine staining, as described previously [15].

### 2.6. $\beta$ -Galactosidase assays

*B. subtilis* cells carrying the mutagenized *rraprE* were grown at 37°C in Shaeffer's sporulation media [16]. At regular time intervals, samples of 1.5 ml were taken and assayed for  $\beta$ -galactosidase specific activity as described in [5].

## 3. Results and discussion

### 3.1. Computer analysis of the *aprE* promoter

To identify and evaluate putative  $\sigma^N$  promoters within the 5' control region of the *aprE* gene (*rraprE*), we developed a computer program based on the statistical algorithm derived by Mulligan et al. [10]. With this program, a maximum value of 100 is obtained when all bases match the promoter extended consensus  $\text{aaaattgtTTCACAT-11 bp-tatgaTATAATaaaa}$  (see Section 2). When a cut-off value of 60 was used, the program only found two putative promoters within the first 100 bp upstream the *aprE* TIS. One of these promoters had a score value of 63 and corresponded to that previously proposed by Park et al. [11], located 7 bp upstream to the *aprE* TIS with a 19-bp space region between the  $-10$  and  $-35$  boxes. A second pro-



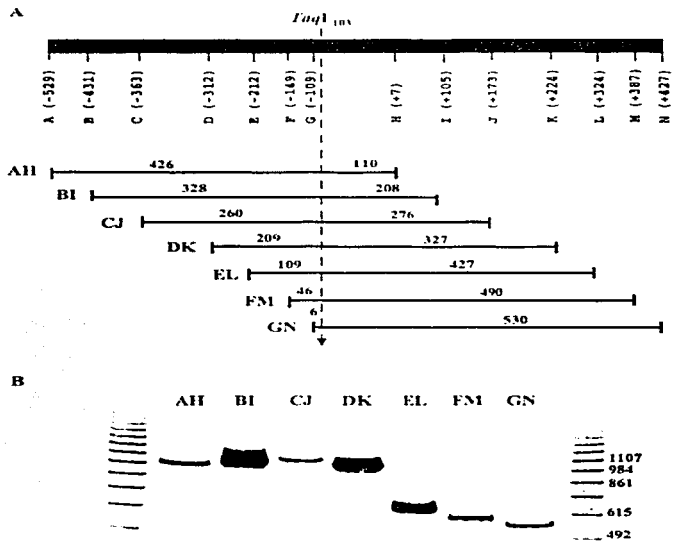


Fig. 2. Strategy to map the position of the bending loci on the *rrpE* gene. Several 536-bp DNA fragments were generated by PCR, using a set of primers complementary to different regions of the *aprE* gene and to its regulatory region (A), and analyzed by 8% PAGE (B). The positions of the primers used for PCR are relative to the TIS and their sequences are shown in Table 1. In each PCR-generated DNA fragment, the position of the *TaqI*-10x restriction site to both ends of each fragment is also indicated.

only affect one of the two analyzed  $-35$  boxes (see Fig. 1A). The BS-37G strain showed a 69% decrease in the *aprE'*-*lacZ* transcription rate. On the other hand, BS-31C did not yield a significant variation on its *aprE'*-*lacZ* transcription rate, as it can be inferred from their  $\beta$ -galactosidase activity, 81% with respect to the BSR6 strain (see Table 3). Therefore, our results indicate that, under the conditions tested, the *aprE* promoter has 17 bp between the  $-10$  and  $-35$  boxes and it is located 7 bp upstream of its TIS (Fig. 1A). Nevertheless, we could not discard the possibility that both promoter sequences may be functional and transcribe the *aprE* gene at particular conditions not assayed in our study. However, the conditions used in the present work are the same used by several other groups to study *aprE* regulation as well as many sporulation-associated processes.

### 3.3. Electrophoretic analysis of circular permuted DNA fragments of the *rrpE*

The presence of a static curvature DNA region on the *rrpE* has been reported [20]. However, the exact position of the sequence(s) responsible for such curvature was not identified. We used a PCR-based method to map curvature loci in this region. Using different sets of primers, we generated several DNA fragments of the same size, but with a specific DNA region permuted at different distances from the ends of the DNA fragments. This procedure allows the scanning for the contribution of different DNA regions to gel migration anomalies [21]. Because preliminary results suggested that the 160 bp preceding the  $-35$  region was involved in DNA bending [20], we focused our efforts on that region.

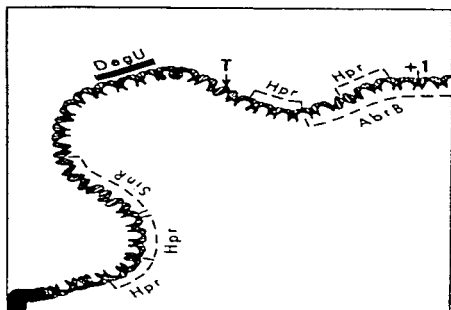


Fig. 3. Computer-generated structure of the *rraprE*. The sequence analyzed comprised the region from  $-415$  to  $+15$  of the published sequence by Valle and Ferrari [20]. The positions of the binding site for Hpr, SinR and AbrB proteins are indicated with broken lines [18]. The DNA region essential for activation by the *degC2* mutation [20] is also shown (black bar). Localization of the *TaqI*<sub>-103</sub> restriction site is indicated as T.

Diagrams of the probes used, as well as the gel migration of these PCR-generated DNA fragments, are presented in Fig. 2. As can be seen, all the DNA fragments migrated anomalously, different from the expected size of 536 bp (Fig. 2B). Furthermore, when the *TaqI*<sub>-103</sub> site was almost in the middle of the DNA segment (fragment CJ, Fig. 2A), the slowest migration was observed, suggesting that the center of bending was at, or very close to, this *TaqI*<sub>-103</sub> site. This interpretation was supported by the migration behavior of the BI and DK DNA segments, where the *TaqI*<sub>-103</sub> site is 208 or 209 bp from either end, and the fragments migrated almost identically.

In the case of the AH and EL DNA fragments, the *TaqI*<sub>-103</sub> site is located at the same distance from each end, and therefore we would expect to observe the same migration. Nevertheless, it was not the case. Fragment AH migrated more slowly and therefore we can assume it has an extra curved element that is not present in the fragment EL (Fig. 2B). Furthermore, if we consider the important

migration differences between fragments DK and EL, we can circumscribe this second curved locus to the DNA region located between 209 and 109 bp upstream the *TaqI*<sub>-103</sub> site. These data locate the major determinants of anomalous gel migration of the DNA fragments containing the *rraprE* around 100 and 200 bp upstream the TIS and more than five helix-turns away from the AbrB-binding site, consistent with a previous report by Strauch and Ayazifar [17].

### 3.4. Computer analysis of the static DNA curvature

Considering the complexity of the *rraprE*, revealed by the permutation assay described before, we decided to utilize a computer-based program to predict the three-dimensional trajectory of the *rraprE* DNA sequence. The predictions were made with the DNASTAR computer programs (see Section 2). This computer analysis showed that the DegU-binding site is surrounded by two curved loci, localized around 100 and 200 bp upstream the *aprE* TIS (Fig. 3). This result is in good agreement with our circular permutation assay.

There are numerous examples of curvature loci associated with regulatory elements/regions for transcription in both prokaryotes and eukaryotes [6,19]. Plaskon and Wartell [13] analyzed the distribution of sequence elements associated with DNA curvature for prokaryotic promoters and found a strong correlation between the presence of putative elements of curvature and the promoter strength for regions upstream of the associated genes. The localization of the intrinsically curved DNA upstream the RNAP-binding site might vary, ranging from  $-235$  bp in the distal activating region of streptococcal promoter *pelII* to  $-40$  bp in the early promoter of T5 and T7 phages (reviewed in [12]). In most of the known cases, only one curved upstream DNA region was present, but examples can be found in which two regions placed in phase contribute to promoter activation [8]. In our case, we postulate that the two curved regions presented in the *rraprE* could play an active role in the formation of the transcription open complex, by mediating the interactions between the RNAP and the DegU regulatory protein. DegU has been involved in the transcriptional control of several genes that are expressed during the transition state, such as: levansucrase, neutral protease,  $\alpha$ -amylase,  $\beta$ -glu-

Table 3  
 $\beta$ -Galactosidase specific activity

Strain	$\beta$ -Gal specific activity ( $U \cdot mg^{-1} \cdot protein^{-1}$ ) <sup>a</sup>	Relative activity (percentage) <sup>b</sup>	Time <sup>c</sup>
BSR6	1648.5	100	5
BS-7G	65.1	4	4
BS-31C	1333.3	81	5
BS-37G	50.8	31	3

<sup>a</sup>The values shown are the averages of three independent experiments and represent the maximum specific activity for each strain assayed.

<sup>b</sup>The relative activities were calculated with respect to BSR6 strain activity.

<sup>c</sup>Period (hours) when the maximum level of  $\beta$ -galactosidase activity is reached, after the transition state (T0).

canase, serine protease (reviewed in [20]). However, there is not a clear consensus sequence for its binding. We consider that the curved loci present in the *rrapE* might form a specific three-dimensional domain, that could modulate the interactions between the set of proteins that control *aprE* transcription. Currently, we are testing this hypothesis.

#### Acknowledgements

We thank Edmundo Calva, Guillermo Gosset and Beatriz Palmeros for critical reading of the manuscript and Mercedes Enzaldo for her excellent technical assistance. Janet Jan was supported by scholarship from Consejo Nacional de Ciencia y Tecnología (CONACyT) México and Dirección General de Estudios de Posgrado (DGEPE) UNAM-México.

#### References

- [1] Anagnostopoulos, C. and Spitzner, J. (1961) Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* 81, 741-746.
- [2] Birnboim, H.C. and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7, 1513-1523.
- [3] Bolshoy, A., McNamara, P., Harrington, R.E. and Trifonov, E.N. (1991) Curved DNA without A-A: Experimental estimation of all 16 DNA wedge angles. *Proc. Natl. Acad. Sci. USA* 88, 2312-2316.
- [4] Ferran, E., Henner, D.J., Perego, M. and Hoch, J.A. (1988) Transcription of *Bacillus subtilis* and expression of subtilisin in sporulation mutants. *J. Bacteriol.* 170, 289-295.
- [5] Ferran, E., Howard, S.A. and Hoch, J. (1986) Effect of stage 0 sporulation mutations on subtilisin expression. *J. Bacteriol.* 166, 173-179.
- [6] Gourse, R.L., de Boer, H.A. and Nomura, M. (1986) DNA determinants of rRNA synthesis in *E. coli*: Growth rate dependent regulation, feedback inhibition, upstream activation, antitermination. *Cell* 44, 197-205.
- [7] Helmmann, J.D. (1995) Compilation and analysis of *Bacillus subtilis*  $\sigma^A$ -dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res.* 23, 2351-2360.
- [8] McAllister, C.F. and Achberger, E.C. (1989) Rotational orientation of upstream curved DNA affects promoter function in *Bacillus subtilis*. *J. Biol. Chem.* 264, 10451-10455.
- [9] Merino, F., Osuna, J., Bolivar, F. and Suberón, X. (1992) A general PCR-based method for single or combinatorial oligonucleotide-directed mutagenesis on pUCM13 vectors. *Biotechniques* 12, 509-510.
- [10] Mulligan, M.E., Hawley, D.K., Entriken, R. and McClure, W.R. (1984) *Escherichia coli* promoter sequences predicted in vitro RNA polymerase selectivity. *Nucleic Acids Res.* 12, 789-800.
- [11] Park, S.S., Wong, S.L., Wang, L.F. and Doi, R.H. (1989) *Bacillus subtilis* subtilisin gene (*aprE*) is expressed from a  $\sigma^A$  ( $\sigma^{43}$ ) promoter in vitro and in vivo. *J. Bacteriol.* 171, 2657-2665.
- [12] Perez-Martin, J., Rivas, F. and Lorenzo, V. (1994) Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol. Rev.* 58, 208-290.
- [13] Plaskon, R.R. and Wartell, R.M. (1987) Sequence distributions associated with DNA curvature are found upstream of strong *E. coli* promoters. *Nucleic Acids Res.* 15, 785-796.
- [14] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- [15] Sekiguchi, J., Takada, N. and Okada, H. (1975) Genes affecting the productivity of alpha-amylase in *B. subtilis* Marburg. *J. Bacteriol.* 121, 688-694.
- [16] Shaeffer, P., Millet, J. and Aubert, J. (1965) Catabolic repression of bacterial sporulation. *Proc. Natl. Acad. Sci. USA* 54, 704-711.
- [17] Strauch, M.A. and Avazifar, M. (1995) Bent DNA is found in some but not all regions recognized by the *Bacillus subtilis* AbrB protein. *Mol. Gen. Evol.* 246, 756-760.
- [18] Strauch, M.A. and Hoch, J.A. (1992) Control of post-exponential gene expression by transition state regulators. In: *The Biology of Bacilli. Applications to Industry* (Doi, R.H. and McLaughlin, M., Eds.), pp. 105-121. Butterworth-Heinemann.
- [19] Struhl, K. (1985) Naturally occurring poly (dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. USA* 82, 8419-8423.
- [20] Valle, F. and Ferran, E. (1989) Subtilisin: a redundantly temporally regulated gene. In: *Regulation of Prokaryotic Development* (Smith, I., Slepceck, R.A. and Setlow, P., Eds.), pp. 131-146. American Society for Microbiology, Washington, DC.
- [21] Valle, F. (1994) PCR-based method to map the bending locus of DNA molecules. *PCR Methods Appl.* 4, 44-45.
- [22] Yang, M.Y., Ferran, E. and Henner, D.J. (1984) Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create an in vitro-derived deletion mutation. *J. Bacteriol.* 160, 15-21.



## **VI.2 MANUSCRITO 2:**

**“ Construction of protein overproducer strains in  
*Bacillus subtilis* by an integrative approach ”**

**Applied Microbiology and Biotechnology (Sometido)**

J. Jan F. Valle F. Bolivar E. Merino\*

**Construction of protein overproducer strains in *Bacillus subtilis* by an integrative approach.**

Molecular Microbiology Department. Instituto de Biotecnología, Universidad Nacional Autónoma de México. Apdo Postal 510-3. Cuernavaca, Morelos 62271. México.

\* Corresponding autor:

Tel.: (52-73) 29 16 01

Fax: (52-73) 17 23 88

e-mail: merino@ibt.unam.mx

## Abstract

We evaluated the effect of several genetic factors that had been reported with a role in the induction of the expression of important levels of recombinant protein in *Bacillus subtilis*. We have utilized the  $\beta$ -galactosidase reporter protein from *Escherichia coli* as our model for the overproduction of heterologous proteins in *Bacillus subtilis*. The *lacZ* gene was expressed in *B. subtilis* using the regulatory region of the subtilisin gene *aprE*, whose transcription is directed by the  $\sigma^A$  factor and its expression is highly controlled by negative and positive regulators. In this study, we considered factors modulating the transcription and translation initiation rates, genetic and mRNA stability and the effect of specific genetic backgrounds such as *degU32* and *hpr2*; that until now, had been studied independently in the construction of protein overproducer strains of *B. subtilis*. We obtained a very significant increase in the  $\beta$ -galactosidase activity, more than 100-fold, by changing the native -35 promoter box to the consensus TTGACA sequence. The effects of site-directed mutagenesis on this and other important regulatory *loci* are discussed in detail.

## Introduction

The overproduction of homologous and heterologous polypeptides by recombinant cells is widely used for research and commercial purposes. Overproducing strains from a broad range of prokaryotic and eukaryotic cells have been constructed through genetic engineering. Among them, members of the *Bacillus* group are particularly notorious since they produce many metabolites and enzymes with commercial value (Zukowski 1992). The most common *bacilli* used in industry are *B. licheniformis*, *B. amyloliquefaciens* and *B. subtilis*, being the former, the best studied from different points of view. Furthermore, because of its GRAS (generally recognized as safe) status, *B. subtilis* is a natural candidate for the production of metabolites and proteins utilized in the food and pharmaceutical industries.

Several genetic and process factors are involved in the performance of an overproducing strain and the production of each particular metabolite or protein present their own challenges. However, there are some common factors that through the years, have shown to be important. One of them, is obtaining the proper expression level of the relevant gene(s) during the process. In bacteria, this has been achieved through the use of a small set of well characterized regulatory regions that allow the controlled expression of the genes of interest.

In the present study, we utilized the regulatory region of the *aprE* gene of *B. subtilis* as our working model for the construction of protein overproducer strains. The *aprE* gene codes for the extracellular protease subtilisin, a valuable enzyme manufactured by the biotechnology industry (Debadov 1982). The development of recombinant protein production systems using *B. subtilis* as a host organism, especially those driven by the subtilisin promoter, provides an important tool for research and commercial production in this area (Oyama et al. 1989).

Although subtilisin synthesis is not required for sporulation (Stahl and Ferrari 1984), its production is triggered by mechanisms common to those events responsible for the sporulation initiation, and hence, it has served as a model for developmentally-associated gene expression (Sonenshein 1989). The *aprE* gene is transcribed by sigma A ( $\sigma^A$ ) and its expression is controlled by several regulators, such as: DegU/DegS, AbrB, Hpr and SinR (Valle and Ferrari 1989).

Here we report the results of experiments directed towards the optimization of the rate of transcription of the *aprE* promoter, as well as the translation frequency and stability of the mRNA produced.

## Materials and methods

### Bacterial strains and transformations procedures

The strains used in this work are listed in Table 1. *E. coli* JM101, transformed by standard procedures (Sambrook et al. 1989), was used for plasmid constructions and amplifications. *B. subtilis* strains were derived from BB80. The integrative plasmids listed in Table 1 were linearized with *Pst*I or *Xmn*I to transform *B. subtilis* BB80 and its derivatives were used to construct the strains detailed in Table 2. The transformation protocol was as described by Anagnostopoulos and Spizizen (1961); chloramphenicol was used as the selection resistance marker (5ug/ml). Integration of plasmid DNA was confirmed by testing the absence of amylase production by iodine staining as described by Sekiguchi et al. (1975). To measure the *aprE* gene expression, we used an *aprE*'-*lacZ* translational fusion integrated in the *amy* locus, as previously described by Ferrari et al. (1986).

### $\beta$ -galactosidase activity measurements

*B. subtilis* cells were grown at 37°C in Shaeffer's sporulation media (Shaeffer et al. 1965). At regular time intervals, samples of 1.5 ml were taken for measuring  $\beta$ -galactosidase specific activity as described by Ferrari et al. (1986).

Site-directed mutagenesis of the transcription and translation regulatory regions of the *aprE* gene

Plasmid pT7-*aprE* was constructed by the cloning of the 509 base pairs (bp) *EcoRI-BamHI* fragment derived from plasmid pSG35.1 (Ferrari et al. 1988), that contains the *aprE* promoter and the first eight codons of the structural gene, into the plasmid pT7 (Novagen). This new plasmid was used as a template for single or combinatorial oligonucleotide directed PCR mutagenesis accordingly with the protocol described by Merino et al. (1992). PCRs were carried out with *Taq* DNA polymerase (Promega Co.) in a Perkin Elmer PCR System. The nucleotide substitutions introduced into the 5' *aprE* regulatory region were as follows: A-34→T, C-33→G, T-32→A, A-31→C, A-12→G, G-1→A (see Fig. 1A). The final products of the PCR mutagenesis were verified by determining the nucleotide sequence using the dideoxy chain termination method described by Sanger et al. (1977). These DNAs were digested with *EcoRI* and *BamHI*, and cloned into the same restriction sites of the integrative plasmid pSG-PLK in order to yield the plasmids listed in the Fig. 1B. Plasmid pSG-PLK is a pSG35.1 derivative in which the *EcoRI-BamHI* region has been replaced by the polylinker derived from pUC19, leaving a promoterless *lacZ* gene. This change in pSG-PLK provides an easier selection of transformants because colonies plated on X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) agar are blue only if they carry the *aprE* promoter.

#### Transcriptional fusion of the *aprE'-lacZ* gene and the *crv* transcription terminator fragment

To introduce the transcription terminator of the *crvIII*A gene downstream the *lacZ* gene in our overproducer strains, we constructed the integrative plasmid pTerm. This plasmid is a p1515Tc derivative (Guérout-Fleury et al 1995) that contains a DNA fragment carrying 1154 bp of the 3' end from the *lacZ* gene, 621 bp of the *B. subtilis amyE* locus and 47 bp from the *crvIII*A transcription terminator of *B. thuriangiensis* (Wong and Chang 1986). These DNA fragments were cloned by the introduction of specific restriction sites at their ends using the PCR reaction. First, we synthesized oligonucleotides that corresponded to the *crvIII*A terminator sequence; the ends of them had *BamHI* and *XbaI* restriction sites, respectively. This DNA fragment was cloned downstream the *lacZ* gene in order to evaluate its effect on the specific mRNA stability. The nucleotide sequence of the hairpin was obtained from a previous report (Wong and Chang 1986). To facilitate the identification of clones containing the *crvIII*A

terminator, we generated a new *DraI* restriction site in it. This was done by a single base pair change G→C in the middle of the inverted and repeated sequence that corresponds to a loop in the mRNA. The secondary structure of the modified *crjIII.A* terminator was analyzed with the “Fold RNA program” of the GCG package (Wisconsin Package, Version 9.1-UNIX, Sep. 1997, Genetic Computer Group, Inc. Oxford Molecular Group, Inc. Madison, Wis.) and we confirmed that the secondary structures of the native and modified terminator were identical. The carboxy terminal end of *lacZ* was cloned into the vector by placing *SacI* and *XbaI* sites in the oligonucleotide ends utilized in the PCR reaction. Finally, the carboxy terminal end of *amyE* gene was cloned into the *HindIII-EcoRI* sites of the vector. These two DNA fragments carrying the *lacZ* and *amyE* genes were used as substrates for homologous recombination with a previously integrated *amyE::lacZ loci* derived from strains BSR6 and JJ1.

## Results

### Transcription initiation

Recently we have identified more precisely the *aprE* promoter using site-directed mutagenesis (Jan et al. 1999 submitted). Some characteristics of the wild type *aprE* regulatory region (*aprE*-WT) are as follow: a promoter with an almost perfect  $\sigma^A$  consensus sequence on its -10 box (TAcAAT); the -35 region (TactaA) has only two of the six nucleotides characteristic of the consensus sequence TTGACA. These two boxes are separated by 17 bp. Another characteristic of this regulatory region is the presence of an AT rich sequence upstream the -35 box, that is important for the recognition of the  $\alpha$ - subunit of the RNA polymerase (RNAP) (Ross et al. 1993). Taking all these data into account, and with the purpose of improving the interactions of the RNAP with this regulatory region, we decided to perform site-directed mutagenesis of the -35 region of *aprE* promoter, to generate a -35 consensus region. For this, we introduced four nucleotide changes (ACTA -34 to -31→ TGAC). We named this modified regulatory region as *rraprE*-TTGACA. For its analysis, *rraprE*-TTGACA was cloned upstream of the *lacZ* reporter gene and integrated into the *amy* locus of the *B. subtilis* chromosome,

generating strain JJ1. Identical construction with wild type *aprE* promoter has been reported previously (Ferrari et al., 1988). The  $\beta$ -galactosidase activity of strains carrying these *lacZ* fusions was assayed. It was found that the modifications to the -35 box, increased 106-fold  $\beta$ -galactosidase levels, with respect to the parental BSR6 strain that carries the *aprE* wild type promoter (See Table 2 and Fig. 2). It is important to mention that the timing of expression of the *rraprE*-WT and the *rraprE*-TTGACA was identical (see below), suggesting that the changes on the -35 box did not affect the regulatory mechanism(s) that block *aprE* expression during exponential growth.

#### Transcription termination and mRNA stability

It has been reported that the *cryIII*A terminator derived from *B. thuringiensis* is a very efficient transcription terminator that increase the half life of certain mRNA molecules (Wong and Chang 1986) and consequently, the amount of protein produced per mRNA produced. To explore if the *cryIII*A terminator could provide a benefit in the levels of  $\beta$ -galactosidase produced in our experimental system, we cloned the *cryIII*A terminator at the 3' end of the *aprE*'-*lacZ* gene fusion. The effect of this terminator on the  $\beta$ -galactosidase activity and presumably the mRNA stability was analyzed in strains JJ14 (BSR6 *lacZ*::pTerm::amyE) and JJ15 (JJ1 *lacZ*::pTerm::amyE). The insertion of the *cryIII*A terminator in these strains produced an increase in the  $\beta$ -galactosidase activity of 12% and 7%, respectively, with regard to their parental strains lacking the *cryIII*A terminator (See Table 2).

#### Translational initiation; changing the GTG start codon of *aprE*'-*lacZ* to ATG

Translation initiation requires a specific initiation codon and a properly spaced sequence upstream of the initiation codon that has potential base pairing with the 3' end of the 16S ribosomal RNA (the Shine and Dalgarno or SD sequence) (Gold et al. 1981). The recent sequencing of *B. subtilis* genome indicated that the initial codon for 78% of the *B. subtilis*



genes is ATG, while GTG and TTG are used 9 and 13% respectively (Rocha et al. 1999). The first codon of the *aprE* gene is GTG and we wanted to quantify the effect of changing GTG to ATG on *aprE* expression.

For such a purpose, we constructed plasmids pATG-WT and pATG-TTGACA that carry the mutated ATG start codon of the *aprE*'-lacZ translational fusion. This fusion comprises the first eight codons of *aprE* and the *lacZ* reporter gene. The only difference between these two plasmids is that pATG-WT carries the *rraprE*-WT while pATG-TTGACA carries the *rraprE*-TTGACA. The G→A mutation in the *aprE*'-lacZ start codon was verified by DNA sequencing. These translational fusions were integrated by double homologous recombination into the *amyE* locus of the *B. subtilis* BSR1 strain, generating strains JJ8 and JJ11, respectively (See material and methods). As shown in Table 2 and Fig. 3, these two strains produced similar β-galactosidase activities levels as their parental BSR6 and JJ1 strains, which carry the wild type GTG start codon.

#### Changing the *aprE*'-lacZ ribosome binding sequence

Another important factor that could affect the translational initiation rate is the interaction of the 3' end of the 16S rRNA with the mRNA. It has been reported that the *B. subtilis* SD consensus sequence AAGGAGG, increases the rate of translational initiation in other organisms (McLaughlin 1981; Ringquist et al. 1992; Lee et al. 1996). The native SD sequence of the *aprE* gene is AAGGAGA. Therefore, we decided to make it consensus by changing the first A to a G. This change was done in the *rraprE*-WT and *rraprE*-TTGACA regulatory regions. In this way, strains JJ9 and JJ12 were generated. In both cases, this change had a negative effect on the β-galactosidase levels. JJ9 strain produced only 24% of the β-galactosidase activity detected in its predecessor BSR6 strain; while JJ12 strain yielded 51% of the activity in relation to the parental JJ1 strain. The data for these β-galactosidase activities are also summarized in Table 2, exhibited in Fig. 3, and commented in detail in the Discussion section.

#### Combinatory effects on the translational level

To analyze the effect of the modifications to the SD sequence and to the initiation codon simultaneously, we combined these two modifications in a single *aprE*'-lacZ fusion. (see Fig. 1 and Table 1). Once again, we tested these modifications with the *rraprE*-WT and *rraprE*-TTGACA regulatory regions. The resulting strains were named as JJ10 and JJ13, respectively. In both strains, there was a decrease in the  $\beta$ -galactosidase levels. As shown in Table 2 and Fig. 3; strains JJ10 and JJ13 produced 14 and 63% respectively, of the  $\beta$ -galactosidase levels detected in their parental strains BSR6 and JJ1.

#### Effect of *hpr2* and *degU32* backgrounds on *aprE*'-lacZ expression

It has been reported that the *hpr2* and *degU32* mutations have a positive and accumulative effect on *aprE* expression. Both mutations increase the total amount of *aprE* mRNA produced (Ferrari et al. 1988). To explore the effect of these mutations in the expression levels promoted by the *rraprE*-WT and *rraprE*-TTGACA regulatory regions, the *hpr2* and *degU32* mutations were introduced in strains BSR6 and JJ1 (see Table 1). The effect of each mutation individually and their combination was analyzed in this way.

The  $\beta$ -galactosidase levels of these strains are shown in Table 2. The results obtained with the *rraprE*-WT regulatory region were consistent with the ones reported previously by Ferrari et al. (1988): the *hpr2*, *degU32* and *hpr2*-*degU32* backgrounds produced increases of 3.3, 36.3 or 66-fold of  $\beta$ -galactosidase activity with respect to the wild type strain (BSR6). Interestingly, strains carrying the *rraprE*-TTGACA regulatory region were also affected positively by some of these mutations. Strain JJ2 (*hpr2*) showed an increase of 1.8 fold over the JJ1 strain, while the JJ4 (*degU32*) strain had a marginal increase. However, the combination of these two mutations (*hpr2*-*degU32*) in strain JJ6, produced a 2.7 fold increase on  $\beta$ -galactosidase level.

#### Measurement of $\beta$ -galactosidase synthesis

To have a direct estimation of the amount the  $\beta$ -galactosidase protein produced by our best overproducing strain JJ6 (BSR1 *amyE::pTTGACA hpr2 degU32*), we analyzed its total protein profile by SDS-PAGE. For this, it was grown in Schaeffer medium (see material and methods) and samples were taken at regular intervals. The cell-free extracts obtained by sonication were analyzed by SDS-PAGE and stained with Coomassie brilliant blue (Fig. 4).  $\beta$ -galactosidase protein was observed as a band with a molecular mass of 116 kDa. The sporulation process started five hours after the inoculation. At this time, the expression of the recombinant protein started and reached its maximum two hours later, corresponding roughly to 10% of the total intracellular protein.

## Discussion

Several expression systems have been used to overproduce heterologous polypeptides in different species of *Bacillus*. This has been achieved by the combination of strong but regulable promoters, a variety of translation/secretion signals, as well transcriptional terminators. All these factors have been used in different genetic backgrounds, vectors and under a variety of growth conditions, making difficult to evaluate the impact of each one of these factors in the overall process. In this work we analyzed the effect of several modifications to the transcription/translation signals of the *aprE* gene, using an *aprE*- $\beta$ -galactosidase fusion to quantify the effects.

The *aprE* gene is an interesting model to develop systems for the overproduction of natural or recombinant proteins. *aprE* is induced in a natural way at the end of the exponential growth, when the maximum biomass is reached, minimizing the possibility of mutations and plasmid segregation. Natural induction avoids the necessity of artificial inducers such as those systems base on chemicals, like IPTG, oxygen (Walsh and Koshland 1985; Jensen et al. 1993; Makrides 1996) or physical induction by heat (Bujard et al. 1983; Mieschendahl and Muller-Hill 1985). This natural induction of the *aprE* gene represents an important economical advantage and also because of its simplicity in the fermentation processes, specially at a large industrial scale. The first modification studied was designed to increase the rate of transcription. Our approach consisted in the modification of the -35 region of the *aprE* promoter to make it identical to the consensus sequence identified for  $\sigma^A$  dependent promoters (Helmann 1995). We found that by

changing 4 bases on this region, more than 100-fold increase in the  $\beta$ -galactosidase activity was obtained. Interestingly, this new *aprE* promoter still responded to temporal regulation as the wild type *aprE* promoter. Furthermore, it also retained its ability to respond to certain mutations that are known to increase *aprE* transcription, i.e. the *hpr2* and *degU32* mutations. *hpr2* mutation eliminates the negative regulator Hpr while *degU32* makes DegU to remain phosphorylated for longer periods of time. This is particularly interesting because one of the 4 binding sites for the Hpr regulator that have been identified on the *aprE* regulatory region overlaps with the promoter (Kallio et al.1991). In this sense, the -35 region corresponds to the first 6 bases of the 22 protected from DNase I cleavage by the binding of Hpr. These results may indicate that the modified bases do not establish critical contacts with Hpr or that the presence of 3 other Hpr binding sites in the same region, diminish the effect of the mutations.

In this paper we demonstrate that the changes introduced into the *aprE* regulatory region could increase the  $\beta$ -Gal activity even more (1.6 folds See Table 2) than the increase obtained by the *hpr2 degU32* mutations, commonly present in strains used at industry. Furthermore, we believe that this is particularly relevant because the presence of the *hpr2 degU32* mutations have several pleiotropic effects, like the hyper-production of several other extracellular degradative enzymes; propensity to cellular lysis and the lack of motility and competence (Kunst et al. 1994). The absence of some these pleiotropic effects could explain why, while in both strains BSR6 and JJ1  $\beta$ -galactosidase activity was obtained 10 hours after  $T_0$ , the accumulated  $\beta$ -galactosidase started to decrease in the BSR6 strain, but remained constant for 20 hour in the JJ1 strain (See Fig. 2).

Another aspect considered in this work was the change of the initiation codon GUG of the *aprE* gene. As mentioned before, the inspection of the *Bacillus* genome indicated that the frequency of use of AUG, UUG and GUG as initiation codons was 78%, 13% and 9% respectively. Interestingly, the same GUG codon is used in the orthologous genes of *Bacillus stearothermophilus* and *Bacillus amyloliquifaciens*. After changing the initiation codon GUG to ATG, we expected a positive effect in the expression of our *aprE'-lacZ* fusion. However, it did not have any significant modification on the amount of  $\beta$ -galactosidase produced. It has been proposed that start codon frequency is partly a result of mRNA structure avoidance (Rocha et al., 1999). In this sense, on average the *in silico* mutation of an AUG implies

significantly more stable structures (energy 7% lower for GUG). The formation of stable structures at the beginning of the mRNA can affect translation by affecting ribosome binding and/or movement (Rocha et al., 1999). The fact that we did not measure any effect of the change of the initiation codon, suggest that the change had no significant impact the alternative structures formed at the beginning of the *aprE-lacZ* mRNA. Here it is important to stress that the formation of these alternative structures will depend on the initiation codon as well as on the downstream sequence. Because of this, the lack of effect of the GUG to AUG change should not be generalized.

Regarding the change of the ribosome binding site in the *aprE'-lacZ* gene fusion, our results suggest that the nucleotide A at the position +49 is relevant for the ribosome binding and has some influence on the yield of the protein synthesis. It has been reported that SD sequences closer to the *B. subtilis* AAGGAGG consensus, increased the translation efficiency due to the easier recognition of the ribosomes (Band and Henner 1984; Vellanoweth et al. 1992); however in our particular case, we did not find this effect. One possible explanation for the lack of a positive effect in the construction of the consensus SD sequence, is that altering the natural *aprE* SD sequence could also generate a secondary structure in the leader region that could block the translation initiation. Nevertheless, as far as our computer analysis could tell, this possibility was found not to be true. The translation initiation site was free of stable interactions that could mask the SD sequence or the initiation codon, as it happens in the wild type SD sequence. However, there are other factors that could influence the translation efficiency like the space and the sequence between the initiation codon and the SD sequence. One element to be considered is that in a consensus sequence, the more frequently utilized nucleotides are in general considered important, but are they the most feasible for a particular gene? Sometimes it is not possible to predict the effect of the context and a nonconsensus base at a specific position could be important for a particular mRNA, as it was the case for the *aprE* SD sequence. It is important to notice that our modified *aprE'-lacZ* reporter gene and the T7 mRNA late genes had the same AAGGAGG sequence. A low efficiency to translate the late mRNA of the *E. coli* phage T7 by *B. subtilis* ribosomes despite the presence of a strong SD sequence has been reported by Hager and Rabinowitz (1985). Therefore, a strong SD sequence

may not always be sufficient for efficient translation of a mRNA by *B. subtilis* ribosomes, but it is in fact, a necessary one.

The cloning of the *cryIII.A* transcription termination at the 3' end of the *lacZ* gene produced a minor increase of the  $\beta$ -galactosidase activity, in contrast to other cases where the same transcription termination has been used to stabilize the mRNA of the penicillinase gene of *Bacillus licheniformis* (*penP*) or the interleukin 2 cDNA from the human Jurkat cell line (Wong and Chang et al. 1986). In our case, it is important to consider the nature of the mRNA decay. In *E. coli* the main degradation of the *lacZ* mRNA happens at the 5' end (Bechhofer and Dubnae 1987; Agaisse and Lereclus 1996; Carrier and Keasling 1997). Therefore the addition of stable secondary structures at the 3' end could not be sufficient to stabilize its mRNA. We believe that the cloning of this transcription termination at the 3' end of any other gene could perhaps increase its stability; although in the particular case of *lacZ* gene, it did not work so efficiently. Therefore, it would be necessary to design experiments to protect the *lacZ* mRNA against the 5' end of the exonucleases action. Accordingly to this, Carrier and Keasling (1997) have developed a system that consists in the introduction of hairpins into the leader region between the start of the transcription and the translation of the gene of interest. The mRNA containing the hairpin structure exhibited a half-life three times longer than the half-life of the mRNA with no hairpin, resulting in an increase of both, mRNA stability and protein levels. Furthermore, Jurgen et al. (1998) have studied the importance of the 3' end of the *lacZ* mRNA. They concluded that the stability is not determined by specific sequences at the 3' end of the transcript but rather by sequences upstream the translational start codon and suggested that the *lacZ* mRNA degradation is probably the result of sequential cleaves of naked mRNA at vulnerable sites exposed by transit of the last translating ribosome (Subbarao and Kenell 1988). This work shows that a high level of a specific recombinant protein, around 10% of the total protein, can be produced in certain *B. subtilis* strains. It is important to consider that this increase was achieved only by a single copy of the gene, integrated in the chromosome. In this regard it has been reported that the integration of DNA into bacterial genomes is a feasible approach for the stabilization of heterologous genes in industrially important bacteria. The pressure exerted on autonomous replicons appears to be absent or reduced when genes are inserted into the bacterial chromosome (Peredelchuk and Bennet 1997). In contrast, the

expression systems based on plasmids present difficulties to maintain the genetic information due to their segregational instability properties (Balbás et al. 1993). This problem is commonly solved by the introduction of antibiotic resistance genes; nevertheless, the addition of high concentration of antibiotics to the media is required. Other important problem related to plasmids is the rearrangement and deletion events that are significantly selected, because of the loss of the genes to be expressed, often impart a cell growth advantages over the producer strain. In our study, these drawbacks were minimized by the integration of the genetic material into the bacterial chromosome. Furthermore, it is possible to increase the productivity by multiple integration or amplification of the DNA region of interest; there are several methods for this purpose (Mori et al. 1988; Kiel et al. 1995; Balbás et al. 1996; Le Borgne et al. 1998). This system can also be applied to others genes of commercial relevance, with the advantage that the *aprE* promoter avoids the use of expensive elements for large scale fermentation induction.

In this study, we only made modifications at the genetic level; however, there are some others possibilities to raise the level of gene expression related with the fermentation conditions of the culture. It has been reported that it is possible to increase the productivity of a system, if a higher cell density is attained by the use of O<sub>2</sub>-enriched air for sparging (Fass et al. 1989), removing growth inhibitory metabolites by ultrafiltration (Lee and Chang 1990), using computer control (Paalme et al. 1997) or combinations of these strategies (Martínez et al. 1997). Therefore, the 100-fold increase in the  $\beta$ -galactosidase expression obtained in our study (strain JJ1 ) could be considered as an advantageous starting point for an optimized scale-up process.

**Acknowledgements** We thank Mercedes Enzaldo for her technical assistance. Jan-Roblero J. was supported by scholarship from Consejo Nacional de Ciencia y Tecnología.(CONACyT) México and Dirección General de Estudios de Posgrado (DGEP). UNAM-México.

## References

- Agaisse H, Lereclus D (1996) STAB-SD: a Shine-Dalgarno sequence in the 5' untranslated region is a determinant of mRNA stability. *Mol. Microbiol.* 20: 633-643
- Anagnostopoulos C, Spizizen J (1961) Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol.* 81: 741-746
- Balbás P, Alexeyev M, Shokolenko I, Bolivar F, Valle F (1996) A pBRINT family of plasmid for integration of cloned DNA into the *Escherichia coli* chromosome. *Gene* 172: 65-69
- Balbás P, Alvarado X, Bolivar F, Valle F (1993) Plasmid pBRINT: a vector for chromosomal insertion of cloned DNA. *Gene* 136: 211-213
- Band L, Henner DJ (1984) *Bacillus subtilis* requires a "stringent" Shine-Dalgarno region for gene expression. *DNA* 3: 17-21
- Bechhofer DH, Dubnau D (1987) Induced mRNA stability in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 84: 498-502
- Bujard H., Baldarum C, Brunner M, Deuschle U, Gentz R, Hughes J, Kammerer W, Stuber D (1983) In: Papas M, Rosenberg J, Chirikjian G, (eds) *Gene Amplification and Analysis* Elsevier, Amsterdam, vol. 3, pp 65
- Carrier TA, Keasling JD (1997) Engineering mRNA stability in *E. coli* by the addition of synthetic hairpins using a 5' cassette system. *Biotechnol. Bioeng.* 55: 577-580
- Debadov VG (1982) The Industrial Use of *Bacilli*. In: Dubnau DA (ed) *The Molecular Biology of the Bacilli*. Academic Press: New York/London, vol 1, pp 331-370.
- Fass R, Clem TR, Shiloach J (1989) Use of a novel air separation system in a fed-batch fermentative culture of *Escherichia coli*. *Appl. Environ. Microbiol.* 55: 1305-1307
- Ferrari E, Henner DJ, Perego M, Hoch JA (1988) Transcription of *Bacillus subtilis* subtilisin and expression of subtilisin in sporulation mutants. *J. Bacteriol.* 170: 289-295.
- Ferrari E, Howard SA, Hoch J (1986) Effect of stage 0 sporulation mutations on subtilisin expression. *J. Bacteriol.* 166:173-9
- Gold L, Pribnow D, Schoneider T, Shinedlin S, Singer BS, Stormo G (1981) Translational initiation in prokaryotes. *Annu. Rev. Microbiol.* 35: 365-403
- Guérout-Fleury AM, Shazand K, Frandsen, Stragier P (1995) Antibiotic-resistance cassettes for *Bacillus subtilis*. *Gene* 167: 335-336



- Hager PW, Rabinowitz JC. (1985) Inefficient translation of T7 late mRNA by *Bacillus subtilis* ribosomes. Implications for species-specific translation. *J. Biol. Chem.* 260: 15163-7
- Helmann, J.D. (1995) Compilation and analysis of *Bacillus subtilis*  $\sigma^A$ -dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res.* 23: 2351-2360.
- Jensen PR, Westerhoff HV, Michelsen O (1993) The use of *lac*-type promoters in control analysis. *Eur. J. Biochem.* 211: 181-191
- Jurgen B, Schweder T, Hecker M (1998) The stability of mRNA from the *gsiB* gene of *Bacillus subtilis* is dependent on the presence of a strong ribosome binding site. *Mol. Gen. Genet.* 258: 538-545
- Kallio PT, Fagelson JE, Hoch JA, Strauch MA (1991) The transition state regulator Hpr of *Bacillus subtilis* is a DNA-binding protein. *J. Biol. Chem.* 266:13411-13417.
- Kiel JAKW, ten Berge AM, Borger P, Venema G (1995) A general method for the consecutive integration of single copies of a heterologous gene at multiples locations in the *Bacillus subtilis* chromosome by replacement recombination. *Appl. Environ. Microbiol.* 61: 4244-4250
- Kunst F, Msadek T, Rapoport G (1994) Signal transduction network controlling degradative enzyme synthesis and competence in *Bacillus subtilis* In: Piggot PJ, Moran CP, Youngman P (eds) Regulation of bacterial differentiation. American Society for Microbiology, Washington, DC pp. 1-20
- Le Borgne S, Palmeros B, Valle F, Bolivar F, Gosset G (1998) pBRINT-Ts: a plasmid family with a temperature-sensitive replicon designed for chromosomal integration into the *lacZ* gene of *Escherichia coli*. *Gene* 223: 213-219
- Lee K, Holland-Staley CA, Cunningham PR (1996) Genetic analysis of the Shine-Dalgarno interaction: Selection of alternative functional mRNA-rRNA combinations. *RNA* Cambridge University Press 2: 1270-1285.
- Lee YL, Chang HN. (1990) High cell density culture of a recombinant *Escherichia coli* producing penicillin acylase in a membrane cell recycle fermentor. *Biotechnol. Bioeng.* 36: 330-337

- Makrides SC (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* 60:512-38
- Martínez A, Ramírez OT, Valle F (1997) Improvement of culture conditions to overproduce  $\beta$ -galactosidase from *Escherichia coli* in *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.* 47: 40-45.
- McLaughlin JR, Murray CL, Rabinowitz JC (1981) Unique features in the ribosome binding site sequence of the gram-positive *Staphylococcus aureus* beta-lactamase gene. *J. Biol. Chem.* 256: 11283-11291
- Merino E, Osuna J, Bolívar F, Soberón X (1992) A general PCR-based method for single or combinatorial oligonucleotide-directed mutagenesis on pUC/M13 vectors. *Biotechniques* 12: 509-510
- Mieschendahl M, Muller-Hill B. (1985) F'-coded temperature-sensitive lambda cI857 repressor gene for easy construction and regulation of lambda promoter-dependent expression systems. *J. Bacteriol.* 164: 1366-9
- Mori M, Hashigushi KI, Yoda K, Yamasaki M (1988) Designed gene amplification on the *Bacillus subtilis* chromosome. *J. Gen. Microbiol.* 134: 85-95
- Oyama H, Yoshimoto T, Takeshita T, Tsuru D (1989) Secretion of *Escherichia coli* Aminopeptidase P in *Bacillus subtilis* Using the Prepro-Structure Coding Region of Subtilisin Amylosacchariticus. *J. Ferment. Bioeng.* 68: 289-292
- Paalme T, Elken R, Kahru A, Vanatalu K, Vilu R (1997) The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach. *Antonie Van Leeuwenhoek* 71: 217-30
- Peredelchuk MY, Bennett GN (1997) A method for construction of E. Coli with multiple DNA insertions in the chromosome. *Gene* 187: 231-238
- Ringquist S, Shinedling S, Barrick D, Green L, Binkley J, Stormo GD, Gold L (1992) Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. *Mol. Microbiol.* 6: 1219-1229.
- Rocha EP, Danchin A, Viari A (1999). Translation in *Bacillus subtilis*: roles and trends of initiation and termination, insights from a genome analysis. *Nucleic Acids Res.* 27: 3567-76.

- Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, Ishihama A, Severinov K, Gourse RL (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* 262: 1407-1413.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a laboratory manual* (2<sup>nd</sup> edn) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-7
- Sekiguchi J, Takada N, Okada H (1975) Genes affecting the productivity of  $\alpha$ -amylase in *Bacillus subtilis*. *J. Bacteriol.* 121: 688-694
- Shaeffer P, Millet J, Aubert J (1965) Catabolic repression of bacterial sporulation. *Proc Natl Acad Sci USA* 54: 704-11
- Sonenshein AL (1989) Metabolic Regulation of Sporulation and Other Stationary-Phase Phenomenon. In: Smith I, Slepecky RA, Setlow P (eds) *Regulation of Prokaryotic Development*. American Society for Microbiology, Washington, DC pp 109-130
- Stahl, ML, Ferrari, E. (1984) Replacement of the *Bacillus subtilis* Subtilisin Structural Gene With an In Vitro-Derived Deletion Mutation. *J Bacteriol.* 158: 411-418
- Subbarao MN, Kenell D (1988) Evidence for endonucleolytic cleavages in decay of *lacZ* and *lacI* mRNAs. *J. Bacteriol.* 170: 2860-5
- Valle F, Ferrari E (1989) In: Smith I, Slepecky RA, Setlow P (eds) *Regulation of Prokaryotic Development*. American Society for Microbiology, Washington, DC pp 131-146.
- Vellanoweth RL, Rabinowitz JC (1992) The influence of ribosome-binding-site elements on translational efficiency in *Bacillus subtilis* and *Escherichia coli* *in vivo*. *Mol. Microbiol.* 6: 1105-14
- Walsh K, Koshland DE Jr. (1985) Characterization of rate-controlling steps *in vivo* by use of an adjustable expression vector. *Proc. Natl. Acad. Sci. USA* 82: 3577-3581
- Wong HC, Chang S (1986) Identification of a positive retroregulator that stabilizes mRNAs in bacteria. *Proc. Natl. Acad. Sci. USA* 83: 3233-3237
- Zukowski MM (1992) Production of commercially valuable products. In: Doi RH, McGlouglin M (eds) *Biology of bacilli: applications to industry*. Butterworth-Heinemann, Stoneham, Mass pp 311-337

**Table 1** *B. subtilis* strains

Strain	Description	Reference
BSR1	<i>Δnpr hisA glyB</i>	Lab stock
BSR2	<i>Δnpr hisA hpr2</i>	Lab stock
BSR3	<i>Δnpr glyB degU32 cat</i>	Lab stock
BSR6	BSR1 <i>amyE::pSG35.1 cat</i>	This work
JJ3	BSR6 <i>hpr2</i>	This work
JJ5	BSR6 <i>degU32</i>	This work
JJ7	BSR6 <i>hpr2 degU32</i>	This work
JJ8	BSR1 <i>amyE::pATG</i>	This work
JJ9	BSR1 <i>amyE::pSD</i>	This work
JJ10	BSR1 <i>amyE::pATGSD</i>	This work
JJ14	BSR6 <i>lacZ::pTerm::amyE-front tc</i>	This work
JJ1	BSR1 <i>amyE::pTTGACA cat</i>	This work
JJ2	JJ1 <i>hpr2 cat</i>	This work
JJ4	JJ1 <i>degU32</i>	This work
JJ6	JJ1 <i>hpr2 degU32</i>	This work
JJ11	JJ1 <i>amyE::pATG</i>	This work
JJ12	JJ1 <i>amyE::pSD</i>	This work
JJ13	JJ1 <i>amyE::pATGSD</i>	This work
JJ15	JJ1 <i>lacZ::pTerm::amyE-front tc</i>	This work

Abbreviations: *cat*, chloramphenicol acetyltransferase gene; *tc* tetracycline resistance gene, *amyE-front*, the 5' end of the *amyE* gene.

**Table 2**  $\beta$ -galactosidase activity in different overproducer strains

Strain <sup>a</sup>	Maximum specific activity (U/mgprot <sup>b</sup> )	Relative activity with respect to BSR6	Relative activity with respect to JJ1	Time <sup>c</sup>
<b>Effect of <i>hpr2</i> and <i>degU32</i> backgrounds</b>				
BSR6	2 452	1X	.009X	10
JJ3	8 105	3.3X	.03X	14
JJ5	89 000	36.3X	0.3X	5
JJ7	162 330	66X	0.6X	8
JJ1	261 360	106X	1X	7
JJ2	472 090	192X	1.8X	11
JJ4	335 810	137X	1.3X	9
JJ6	712 300	290X	2.7X	23
<b>Effect of the <i>cryIIIA</i> terminator</b>				
JJ14	2747	1.12X	.01X	10
JJ15	278 800	114X	1.067X	7
<b>Effect of the ATG and SD mutations</b>				
JJ8	2 370	1X	0.009	8
JJ9	595	0.24X	0.002	6
JJ10	355	0.14X	5.3E-7	5
JJ11	268 060	109X	1X	7
JJ12	132 880	54X	0.51X	5
JJ13	163 720	67X	0.63X	5

<sup>a</sup> The relevant genotype or mutation from each strain is exhibited

<sup>b</sup> The values shown are the averages of three independent experiments. The activity assayed is  $\beta$ -gal expressed from the *aprE*<sup>-</sup>-*lacZ* fusion.

<sup>c</sup> Period (hours) when the maximum level of  $\beta$ -galactosidase activity is reached, after the transition state ( $T_d$ ).

## Legends

**Fig. 1-A** Representation of the mutations incorporated at the regulatory region of the *aprE'-lacZ* fusion gene. Numbers indicate the nucleotide positions relative to the transcription start site (+1). The bases that correspond to the promoter sequences, boxes -10 and -35, are exhibited by squares. The analyzed modifications are shown on the top of the specific nucleotides.

**Fig. 1-B** Plasmids constructed in this study. Nucleotides in bold face represent the mutations introduced at the *aprE* regulatory region.

**Fig. 2**  $\beta$ -galactosidase activity in different *B. subtilis* backgrounds. (A) Estimation of the biomass as units of optic density. (B) Specific  $\beta$ -galactosidase activity in strains: JJ1 (BSR1 *amyE*::pTTGACA) and BSR6 (BSR1 *amyE*::pSG35.1); this last strain was considered as the wild type.

**Fig. 3** Effect of mutations affecting the translation initiation rate. (A) Mutations incorporated at the *aprE'-lacZ* fusion carrying the -35 box consensus, TTGACA. (B) Mutations incorporated at the *aprE'-lacZ* carrying the wild type -35 box sequence.

**Fig. 4**  $\beta$ -galactosidase synthesis in strain JJ6 detected by SDS-PAGE. The number above each lane represents the samples taken at specific intervals, accordingly with the specific expression time (See text). LacZ protein (116kDa) was used as the molecular marker.

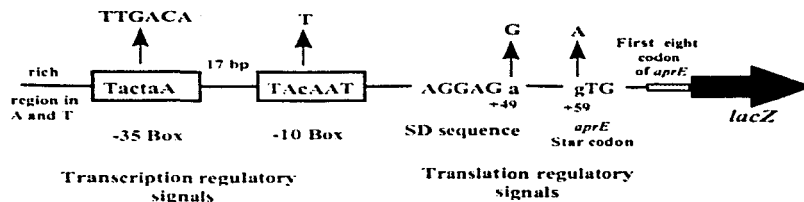


Fig. 1A

Plasmid	DNA sequence				Reference
	-35 Box	-10 Box	SD sequence	star codon	
pSG35.1	TACTAA	TACAAT	---AGGAGA	----GTG	Ferrari et al. 1988
pSG-TTGACA	TTGACA	TACAAT	---AGGAGA	----GTG	This work
PATG	TACTAA	TACAAT	---AGGAGA	----ATG	This work
pATG-TTGACA	TTGACA	TACAAT	---AGGAGA	----ATG	This work
PSD	TACTAA	TACAAT	---AGGAGG	----GTG	This work
pSD-TTGACA	TTGACA	TACAAT	---AGGAGG	----GTG	This work
PATGSD	TACTAA	TACAAT	---AGGAGG	----ATG	This work
pATG-TTGACA	TTGACA	TACAAT	---AGGAGG	----ATG	This work

Fig. 1B

Fig. 2

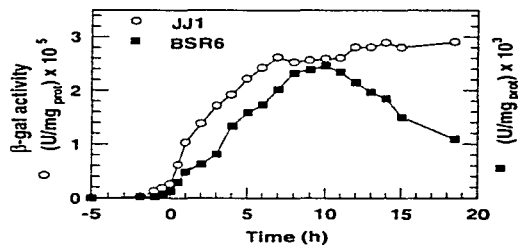
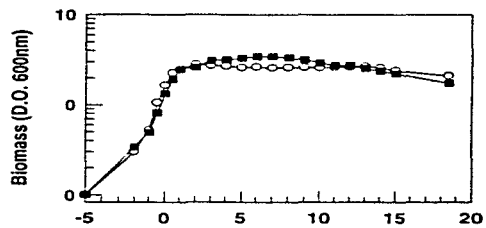




Fig. 3

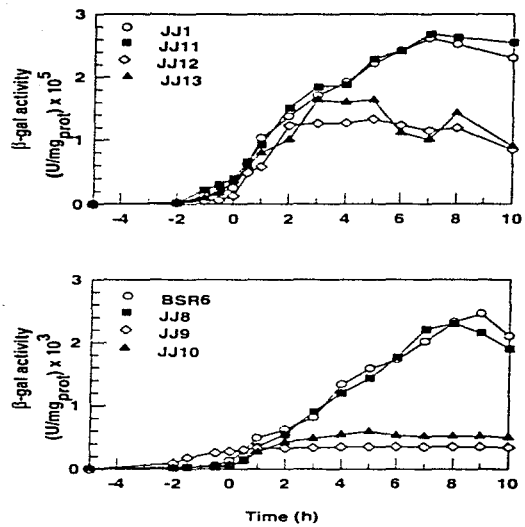
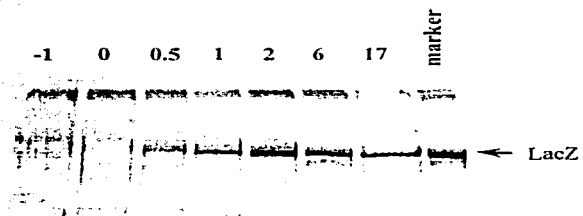


Fig. 4



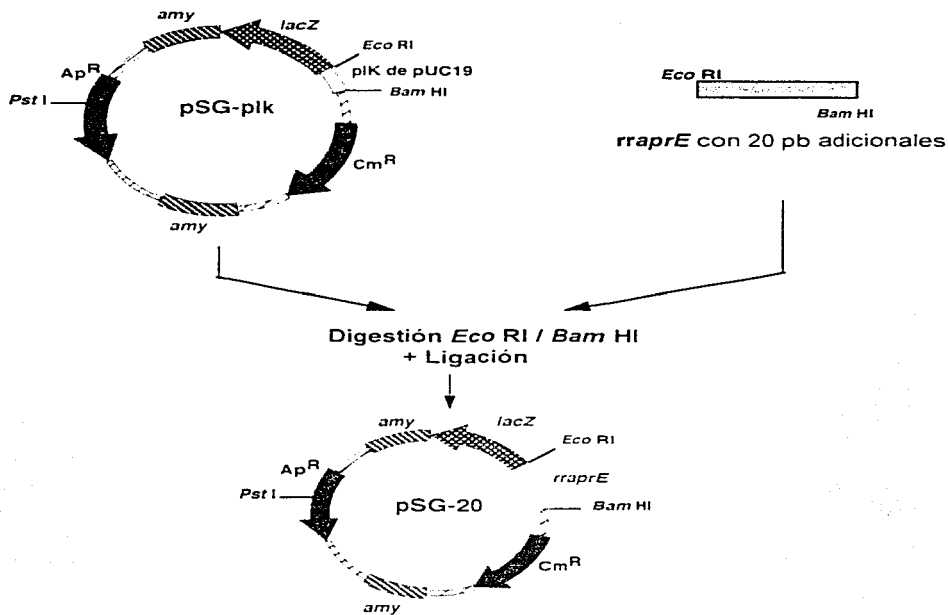
## **VII Experimentos no incluidos en los artículos**

## VII.1 Adición de 20 pb entre los sitios de unión 3 y 4 de la proteína Hpr en la *rraprE*.

Dentro de nuestro grupo de investigación se realizaron experimentos donde se insertaron múltiplos de 5 pb entre los sitios de pegado 3 y 4 de la proteína represora Hpr (Sampieri A., Tesis de Maestría. En preparación). Estos experimentos fueron realizados con la finalidad de conocer la topología del DNA en la *rraprE* y ver si la inserción de las bases producía giros que afectaran de alguna manera la unión de la proteína represora Hpr y sus posibles contactos con otras proteínas reguladoras. Se observó que la inserción de 20 pb sobre la región de regulación silvestre del gen *aprE* [*rraprE*(WT)] producía, al menos fenotípicamente, un efecto positivo sobre la expresión de  $\beta$ -Galactosidasa (Fig. 6). Razón por la cual insertamos esta secuencia sobre la región de regulación del gen *aprE* que lleva la caja -35 consenso, *rraprE*(TTGACA), y analizamos su efecto. Los plásmidos integrativos pSG-20WT y pSG-20TTGACA fueron construidos para este fin. (Ver Fig. 7).

pSG-20WT y pSG-20TTGACA al ser digeridos con *Pst*I y transformados en la cepa BSR1 se integran por doble recombinación homóloga en el *locus* de *amyE*, obteniéndose las cepas JJ17 y JJ16, respectivamente. Antes de proceder a la transformación se verificó mediante secuencia, que los plásmidos únicamente presentaran las mutaciones deseadas; además, la integración fue probada al sembrar las cepas en cajas de almidón (Sekiguchi *et al.*, 1975). Consideramos de interés construir durante el proceso las correspondientes cepas con genotipo *degU*<sup>32</sup>, mutación que hace que el regulador positivo DegU se mantenga mayor tiempo en estado fosforilado. Las mutantes resultantes las llamamos JJ18 y JJ19. La construcción de estas mutantes estuvo basada en la siguiente propuesta: la inserción de 20 pb afectaría la unión del represor Hpr y el efecto que produjera sería similar al obtenido en mutantes bajo el fondo *hpr2* (mutantes previamente caracterizadas, que no producen el represor Hpr). Por consiguiente, esperábamos alcanzar niveles de actividad similares a la cepa JJ3 (*rraprE*-wt, *hpr2*) y JJ2 (*rraprE*-TTGACA, *hpr2*), con las cuales se obtuvieron aumentos de 3.3 y 1.8 veces más actividad, respectivamente (considerando como silvestre la cepa BSR6 y JJ1, en cada caso) con la ventaja de no tener efectos pleiotrópicos del fondo *hpr2*.





**Fig. 7 Construcción de los plásmidos pSG-20WT y pSG-TTGACA**

El plásmido pSG-20WT lleva la *rraprE*(WT) y pSG-20TTGACA la *rraprE* (TTGACA). La secuencia AGCTTAATCCAGTAACAGTA fue insertada entre los sitios de pegado 3 y 4 de la proteína represora Hpr.

La descripción de las cepas, que llevan la inserción de 20 pb. se presenta a continuación:

**Cepas construidas:**

- JJ17** *amy::rraprE* (W.T.)  $\Omega$ 20pb *lacZ'*  $Cm^R$   $\Delta$ *nrp* *his<sup>-</sup> gly<sup>-</sup>*  
**JJ19** *amy::rraprE* (W.T.)  $\Omega$ 20pb *lacZ'*  $Cm^R$   $\Delta$ *nrp degU<sup>32</sup> his<sup>-</sup> gly<sup>-</sup>*  
**JJ16** *amy::rraprE* (TTGACA)  $\Omega$ 20pb *lacZ'*  $Cm^R$   $\Delta$ *nrp his<sup>-</sup> gly<sup>-</sup>*  
**JJ18** *amy::rraprE*(TTGACA)  $\Omega$ 20pb *lacZ'*  $Cm^R$   $\Delta$ *nrp degU<sup>32</sup> his<sup>-</sup> gly<sup>-</sup>*

Las cepas BSR6 y JJ1 sirvieron como control. Los resultados de actividad enzimática de  $\beta$ -galactosidasa se exhiben en la siguiente tabla:

**Comparación de cepas**

	Act. Específica de $\beta$ -galactosidasa (U/mg-prot) <sup>a</sup>	Act. Relativa (Porcentaje) <sup>b</sup>	Tiempo (h) <sup>c</sup>
BSR6 (wt)	2 452	1	10
JJ17 (wt,20pb)	6 051	2.5	12
JJ19 (wt,20pb,degUh32)	22 720	9.3	7
JJ1 (ttgaca)	261 360	1	7
JJ16 (ttgaca,20pb)	347 800	1.3	13
JJ18 (ttgaca,20pb,degUh32)	387 710	1.5	10

**NOTAS:**

<sup>a</sup> Los valores exhibidos son el promedio de 3 experimentos independientes y representan la actividad específica máxima de cada cepa.

<sup>b</sup> La actividad relativa fue calculada con respecto a la actividad de la cepa BSR6.

<sup>c</sup> Periodo (horas) cuando el nivel máximo de actividad de  $\beta$ -galactosidasa es alcanzado, después del estado de transición (T0).

Aunque en nuestras construcciones no obtuvimos incrementos de similar magnitud, nuestros resultados sugieren que la adición de 20 pb entre los sitios de unión 3-4 hacen que el represor Hpr no pueda unirse tan eficientemente y que por ende su efecto represor es inferior respecto al de la cepa silvestre.

Es importante mencionar que dentro de las inquietudes del grupo se encuentra la de obtener mutantes sobreproductoras y que no presenten efectos pleiotrópicos. Antes de la obtención de la cepa JJ1, únicamente contábamos con cepas bajo los fondos *hpr2* y *degU<sup>32</sup>*, con las cuales se alcanzaron niveles considerables de actividad. Sin embargo estas cepas presentan efectos pleiotrópicos; entre ellos, lisarse fácilmente, inconveniente trascendental cuando se trabaja a nivel fermentador. Consideramos de importancia la obtención de las cepas que llevan la adición de 20 pb porque permiten tener efectos positivos en *cis* que no originen efectos pleiotrópicos. En este sentido, las cepas JJ16 y JJ17 no presentan efectos pleiotrópicos. Además con la cepa JJ16 se obtiene un incremento adicional, del 30%, en la actividad de  $\beta$ -Gal en relación con JJ1. Otra característica sobresaliente es que JJ16 presenta regulación similar a la cepa que lleva el promotor silvestre, BSR6. Esto es, la expresión de la proteína recombinante se inicia después de  $T_0$  (Ver. Fig. 8) y en consecuencia el crecimiento del cultivo en su parte vegetativa no se ve afectado, lo que permite niveles de biomasa final elevados.

Se ha reportado que Hpr se une a 4 regiones separadas sobre el promotor *aprE* y han sugerido un modelo de su interacción sobre el DNA (Strauch *et al.*, 1991). El modelo se basa en la formación de una horquilla de represión, tal como la reportada en otros sistemas; operón arabinosa (Huo *et al.*, 1988), operón lactosa (Oehler *et al.*, 1990). Es posible que la inserción de 20 pb afecte de tal manera que el tentativo efecto cooperativo de la proteína represora Hpr disminuya; lo que no es el caso en cepas *hpr2*, donde existe la completa remoción de Hpr. Es indudable que ésta es solamente una hipótesis y que se requieren de más experimentos que nos den la pauta sobre la interacción de Hpr. Sin embargo debemos recordar que en el promotor del gen existe la unión de otras proteínas; como lo es la proteína Sin que tiene efecto represor y la proteína DegU, a la cual se le ha adjudicado efecto activador sobre la expresión de *aprE*.



Por tanto, los experimentos deben planearse de tal manera que no se enmascare el efecto que pudieran tener estas proteínas.

En relación a las mutantes JJ19 y JJ18, fueron construídas esperando un aumento similar al obtenido con las dobles mutantes *degU<sup>32</sup>-hpr2*. Con ellas, tampoco se obtuvo el incremento esperado. Sabemos que la falta de incremento no son efectos de carga metabólica debido a que tenemos una cepa, JJ7, con la cual hemos alcanzado niveles por arriba de los obtenidos con las cepas presentadas en este apartado.

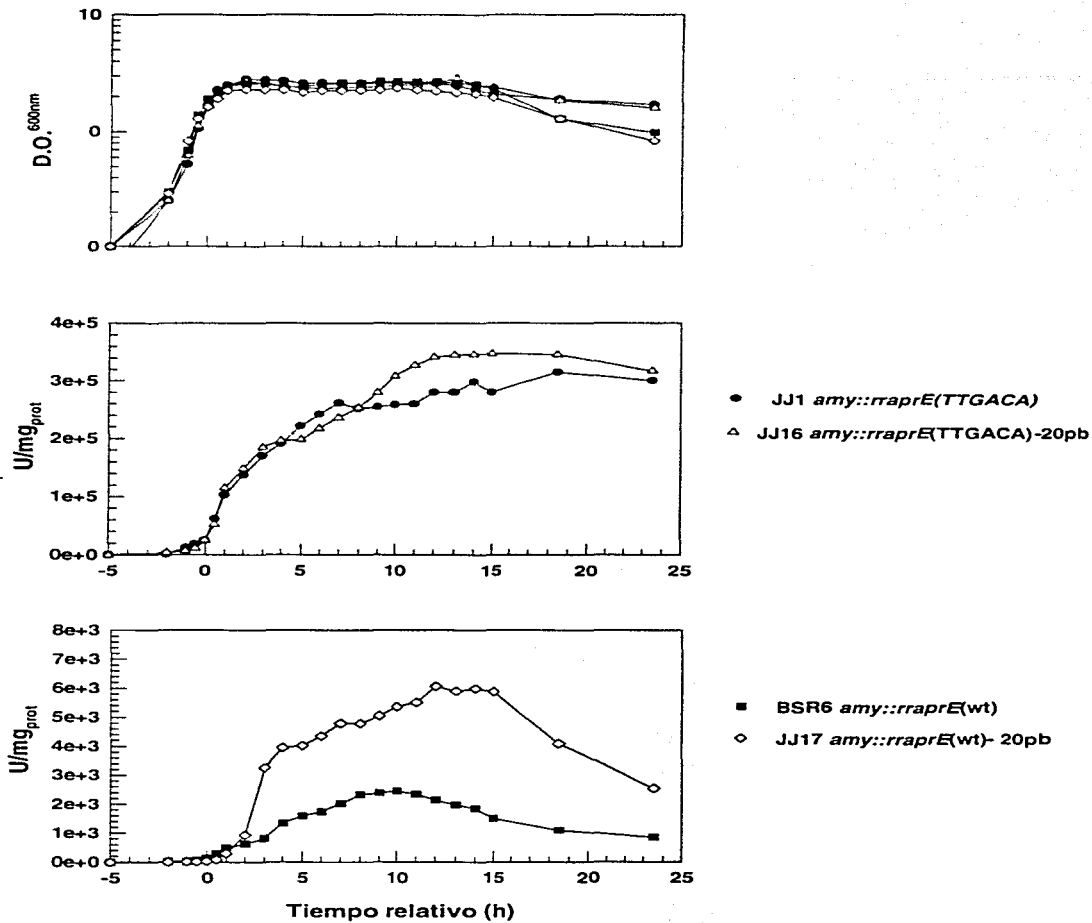


Fig. 8 Curva de crecimiento y actividad específica de  $\beta$ -Gal de las cepas JJ1, JJ16, BSR6 y JJ17

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

## VII DISCUSION GENERAL

## Caracterización del promotor del gen *aprE* de *B. subtilis*

La expresión genética en procariotes comienza con la unión específica de la RNAP al promotor con la RNAP para iniciar la síntesis del RNA. La cantidad de transcrito sintetizado se encuentra íntimamente relacionada con la velocidad de iniciación efectiva; a su vez la cantidad y vida media del mRNA producido repercute en la cantidad de proteína recombinante (McClure, 1985).

El inicio de la transcripción es un proceso que involucra múltiples etapas, en el que intervienen parámetros termodinámicos y cinéticos. Las etapas involucradas en el proceso se ilustran a continuación:

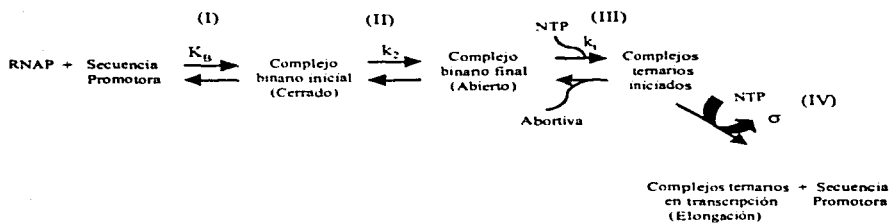


Fig. 9 Etapas en el proceso de la transcripción.

**Etapa I:** Reconocimiento específico. El complejo RNAP- $\sigma$  se une a la doble hélice de DNA en el promotor formando un complejo binario "cerrado". Este complejo es reversible y es descrito por una constante de equilibrio ( $K_b$ ).

**Etapa II:** Formación del complejo abierto. El complejo cerrado es convertido en un complejo abierto por la "apertura" reversible de una región de 10- a 15-pb de DNA en el sitio de inicio

de la transcripción. La apertura de las cadenas comienza arriba del sitio de inicio de la transcripción (+1) y se extiende hasta este mismo, produciéndose el complejo abierto binario final. Para promotores fuertes, la conversión hacia un complejo abierto es irreversible, por lo que esta reacción ocurre a una velocidad constante  $k_2$ . Esta reacción es rápida.

**Etapa III:** Unión de nucleótidos. La siguiente etapa es la incorporación de los primeros nucleótidos (NTP), que son complementarios a la cadena templado. Esto genera un complejo ternario que contiene RNA así como también DNA y RNAP, el cual es descrito por una velocidad constante  $k_3$ , que es más rápida que  $k_2$ . Sin embargo es posible que la enzima libere la cadena, comprometiendo a la iniciación de novo del proceso "iniciación abortiva".

**Etapa IV:** Elongación (Liberación del promotor). Una vez que 7-12 nucleótidos del transcrito se sintetiza, la subunidad  $\sigma$  es liberada y la elongación continua.

Por consiguiente la expresión eficiente requiere que cada una de estas etapas se lleve a cabo a una alta velocidad; por consiguiente, es necesario que las velocidades de las etapas reversas (dirección de disociación) sean pequeñas en comparación con la siguiente etapa, así la dirección hacia delante es esencialmente irreversible.

Se han realizado varios esfuerzos con la finalidad de especificar las características estructurales comunes de los promotores, principalmente los reconocidos por el factor  $\sigma^{70}$  de la RNAP de *E. coli* (Lisser y Margalit, 1994) y de su homólogo  $\sigma^A$  en *B. subtilis* (Helmann, 1995). La mayoría de los promotores conocidos se han determinado a través de experimentos basados principalmente en identificar el sitio de inicio de la transcripción, a través de mapeo por nucleasa S1 o por extensión de primero (Brosius *et al.*, 1982; Izui *et al.*, 1985; Lee y Roeder, 1981).

En años posteriores el análisis de cómputo ha sido empleado como una herramienta para la búsqueda de posibles promotores dentro de una secuencia de DNA (Mulligan *et al.*, 1984; Chen *et al.*, 1997). El trabajo descrito aquí, ratifica que una secuencia de DNA que contenga un promotor, puede ser analizada semi-cuantitativamente para su localización y estimación de su posible fuerza *in vivo*. El algoritmo que utilizamos fue el descrito por Mulligan *et al.* (1984) y está basado en adjudicar valores a cada nucleótido de acuerdo a una matriz de peso para promotores  $\sigma^{70}$  de *E. coli*. La calificación obtenida relaciona el grado con la cual una secuencia de DNA se acerca al consenso de promotores  $\sigma^{70}$  de *E. coli*. Con el objeto de analizar los posibles promotores en la *trappE*, hemos construido una matriz de frecuencia relativa en base a una lista de 122 promotores  $\sigma^7$  de *B. subtilis* compilado por Helmann (1995); al igual que la matriz de Mulligan *et al.* (1984) para promotores de *E. coli*, nuestra matriz para *B. subtilis* contempla 16 pares de bases de la región -35, incluyendo 9 nucleótidos río arriba de la caja canónica -35 y una sola base hacia abajo; además de 14 pares de bases de la región -10, 5 nucleótidos hacia arriba y 3 hacia abajo de la caja -10 (Fig. 10). Por tanto, sería interesante incluir bases adicionales arriba de la caja -35, elemento "up", las cuales mejoran la actividad del promotor por la unión de la subunidad  $\alpha$  de la RNAP o por facilitar la curvatura del DNA. En este sentido, se ha reportado que el elemento "up" en *E. coli* tiene el consenso: -59 nnAAA(A/T)(A/T)T(A/T)TTTTnnAAAAnn -38 (Ross *et al.*, 1998; Estrem *et al.*, 1998). Los residuos de la subunidad  $\alpha$  involucrados con la unión a DNA son altamente conservados (Gaal *et al.*, 1996; Murakami *et al.*, 1996); por lo tanto, las secuencias de DNA reconocidas por la subunidad  $\alpha$  tienen alta probabilidad de ser conservadas. Además se ha reportado la presencia del elemento "up" en eubacterias. En promotores de *B. subtilis* ha sido identificado el elemento "up" en los promotores de: flagelina, *hag* (Fredrick *et al.*, 1995), del estado de esporulación, *spoVG* (Banner *et al.*, 1983), del fago SP82, *AluI56* (McAllister y Achberger, 1988). Sin embargo, pese a la similitud de las secuencias que reconoce  $\sigma^{70}$  de *E. coli* y  $\sigma^7$  de *B. subtilis*, no puede asegurarse que el promotor extendido sea 100% funcional en todos los genes transcritos en *B. subtilis*. Existen casos tal como, el promotor *lacUV5* de *E. coli* que es transcrito pobremente en *B. subtilis* a pesar de su acercamiento al consenso (Henkin y Sonenshein, 1987). Además se ha reportado que el dinucleótido TG (en la posición -15 y -14) y una secuencia rica en As cercana al -43 son

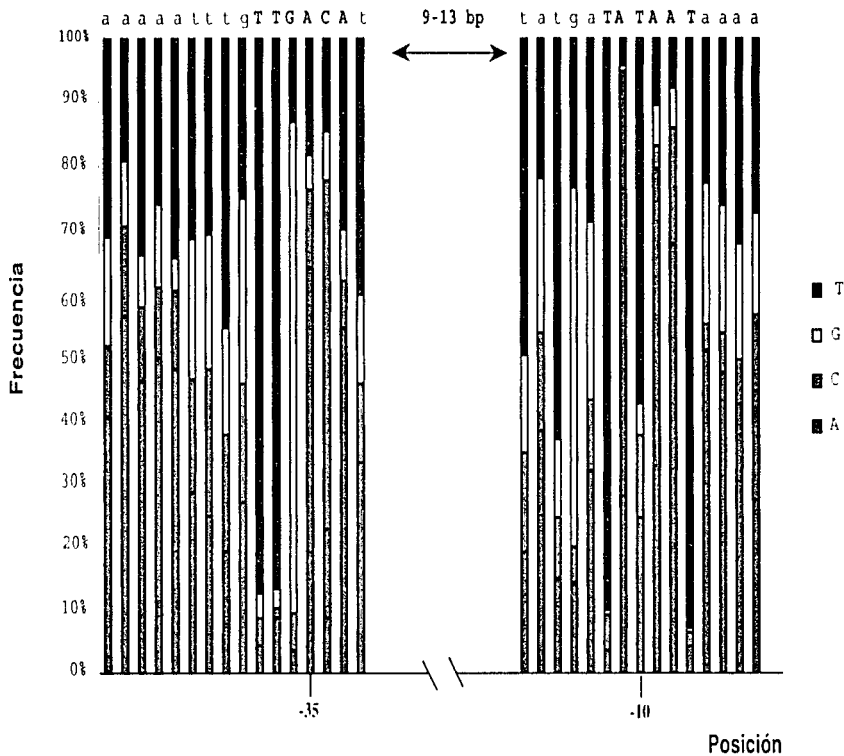


Fig. 10 Frecuencia relativa de nucleótidos en promotores  $\sigma^A$  de *Bacillus subtilis* compilados por Helmann (1995). 122 promotores fueron alineados y la frecuencia relativa de cada base en diferentes posiciones fue evaluada y normalizada.

conservados en promotores de bacterias gram-positivas pero no en *E. coli* (Graves y Rabinowitz, 1986; Moran *et al.*, 1982; Voskuil *et al.*, 1995; Voskuil y Chambliss, 1998). De acuerdo a lo anterior, consideramos de importancia la construcción de una matriz de frecuencia relativa ampliada, que incluya tanto las regiones "up" como aquellas relevantes en promotores de bacterias gram-positivas.

Este tipo de análisis únicamente sugiere posibles promotores; la disposición de tal evidencia debe ser obtenida en una manera más informada y eficiente a través de la evaluación experimental. En el artículo "Characterization of the 5' subtilisin (*aprE*) regulatory region from *Bacillus subtilis*" abordamos ambos tipos de análisis. El valor máximo que se obtiene al correr el programa es 100, cuando todas las bases coinciden con la secuencia promotora consenso aaaaatttgTTGACAt — 11pb — tatgaTATAATaaaa. Al realizar la evaluación de la *rraprE*, para obtener sus promotores putativos, encontramos dos secuencias con calificaciones tan cercanas una de otra (63 y 61) que no fue posible mediante este análisis excluirse una de la otra; por tal razón fue necesario realizar la mutagénesis sitio-específico. El criterio que utilizamos fue la mutación de residuos altamente conservados por los nucleótidos menos frecuente en esa posición [Fig. 1A Artículo 1. "Characterization of the 5' subtilisin (*aprE*) regulatory region from *Bacillus subtilis*". pag. 41]. Además nuestro programa sugiere que el promotor reportado previamente por Strauch y Hoch (1992a) es poco probable de ser funcional ya que obtuvo una calificación de 37; existen al menos 71 otras secuencias con igual o mayor calificación. Este artículo ilustra como la localización de promotores hipotéticos es una tarea relativamente sencilla, el problema es comprobar experimentalmente cual de los sitios es el funcional; además discute como dos posibles promotores superpuestos pueden ser funcionales *in vivo* bajo diferentes condiciones.

Actualmente se conoce la secuencia de genomas completos de diversos organismos y otros más están en este proceso, lo que genera un gran número de fragmentos de DNA secuenciados con función desconocida. El desarrollo de métodos confiables para identificar funciones en regiones de DNA, incluyendo promotores, juega un papel importante en el análisis de estas secuencias. Sin embargo, un dilema en desarrollar programas para el reconocimiento de promotores es descartar una gran porción de promotores falsos-positivos,



lo cual indica que este tipo de análisis preliminar es útil, pero como se mencionó con anterioridad, debe aún verificarse experimentalmente.

### **Curvatura de la región de regulación del gen *aprE***

La regulación del gen *aprE* es altamente controlada, interviniendo en su expresión diversas proteínas con efectos activadores o represores sobre su expresión, así mismo se ha sugerido la presencia de curvatura en esta región. El análisis presentado en el artículo "Characterization of the 5' subtilisin (*aprE*) regulatory region from *Bacillus subtilis*", es un estudio de la curvatura en la *rraprE*. Los *loci* curvos fueron determinados a través de fragmentos de DNA permutados circularmente de acuerdo al método previamente descrito por Valle (1994), ubicándose a 100 y 200 pb hacia arriba del inicio de la transcripción. Nuestros resultados coincidieron con los previamente reportados Strauch y Ayazifar (1995) y sugieren que la mayor curvatura de la *rraprE* se encuentra alrededor del sitio de unión de la proteína DegU; estas regiones curvas juegan un papel fundamental en la formación del complejo abierto de la transcripción, por mediar la interacción entre la RNAP y la proteína reguladora DegU. En este sentido se han reportado numerosos ejemplos de *loci* curvos asociados con regiones reguladoras de genes de origen procarionte y eucariote (Gourse et al., 1986; Lloubes et al., 1988; Struhl, 1985). Existe una correlación fuerte entre los elementos de curvatura y la fuerza del promotor, indicando que el DNA curvo situado hacia arriba de la caja -35 mejora la actividad transcripcional (Plaskon y Wartell, 1987). La localización de DNA intrínsecamente curvo puede variar; en el promotor *pectIII* de *Streptococcus* el DNA curvo se encuentra a -235 pb respecto al inicio de la transcripción y en los promotores tempranos de los fagos T5 y T7 a -40 pb (revisado por Pérez-Martín et al., 1994).

### **Construcción de cepas sobreproductoras**

Una de las estrategias más socorridas para la generación de cepas sobreproductoras en diferentes microorganismos es a través de rondas de mutagénesis química; sin embargo

conocer el genotipo exacto de estas cepas y la posibilidad de reproducirlo es difícil. Por otra parte, la mayoría de las cepas sobreproductoras han sido patentadas, lo que imposibilita su acceso.

En años recientes *B. subtilis* se ha utilizado como hospedero para el desarrollo de sistemas de expresión-secreción de proteínas de interés comercial, entre ellas proteasas, amilasas, celulasas y otras enzimas degradativas (Hoch, 1991). Dentro de sus características sobresalientes se encuentran: microorganismo GRAS por la FDA de los Estados Unidos de América, fácil de manipular a nivel genético, secreta con facilidad proteínas y metabolitos al medio de cultivo, sencillez para cultivarlo, su amplio conocimiento en el área de fermentaciones. Como consecuencia, en *B. subtilis* se han construido cepas para producir diversas enzimas y proteínas de interés; sin embargo estas cepas en la mayoría de los casos presentan efectos pleiotrópicos, interfiriendo en la expresión de genes involucrados en el crecimiento, lisis celular, motilidad y competencia (Kunst *et al.*, 1994).

El gen *aprE* de *B. subtilis* es uno de los modelos clásicos para entender los mecanismos reguladores del estado de transición; el grupo de investigación al cual pertenezco al igual que otros grupos de investigación, lo han adoptado para realizar diversos estudios. Una de las características más atractivas de este gen es que se expresa al final del crecimiento exponencial sin necesidad de inducción, lo cual representa una ventaja potencial si se considera la producción a nivel industrial. En este momento se alcanza la biomasa máxima, minimizando la posibilidad de mutaciones y la segregación de plásmidos.

Gran parte de los experimentos realizados en la presente tesis doctoral fueron hechos con la finalidad de construir cepas sobreproductoras de *B. subtilis*, tomando en cuenta parámetros involucrados en modular los eventos en el inicio de la transcripción y traducción, la estabilidad genética y del mRNA, así como también los efectos de mutaciones específicas in dos *loci* regulatorios del gen *aprE*: Hpr y DegU. En el artículo "Construction of protein overproducer strains in *Bacillus subtilis* by an integrative approach" se hace la descripción de las mutaciones efectuadas y de los efectos que estas produjeron de manera individual; así como también algunas combinaciones de las mismas.

Después de localizar el promotor funcional del gen *aprE*, a 7 pb del inicio de la transcripción, con 17 pb de separación entre las cajas -35 y -10 (Fig. 1A art. 1 pag. 41) fué posible proponer las modificaciones puntuales que se llevaron a cabo y que tienen efecto sobre las señales a nivel transcripcional y/o traduccional. El resultado más sorprendente fué obtenido con la cepa JJ1, la cual lleva la caja -35 consenso. Esta modificación produjo un incremento de 106 veces con respecto a la cepa considerada como silvestre, BSR6 (Fig. 1A y Tabla No. 2, artículo 2 "Construction of protein overproducer strains in *Bacillus subtilis* by an integrative approach", pag. 65 y 67). pensamos que esta mutación favorece el reconocimiento de la región promotora por la RNAP, facilitando la formación del complejo abierto e incrementando la velocidad de su formación. Además, JJ1 tiene gran potencial a nivel industrial, ya que con ella se alcanzaron valores de actividad de  $\beta$ -Gal por arriba de los obtenidos con la cepa de producción JJ7, que alcanza 66 veces más actividad que la cepa BSR6. La cepa JJ7 tiene una mutación en el regulador positivo DegU manteniéndose en estado fosforilado por periodos más largos, así como también la eliminación del regulador negativo Hpr. Es de importancia destacar que JJ1, a diferencia de JJ7, no presenta efectos pleiotrópicos (Kunst *et al.*, 1994) y su nivel de actividad se mantiene 20 h después de T0, lo que no sucede con las cepas BSR6 y JJ7. Aunado a ello, la cepa JJ1 presenta regulación genética similar a la cepa silvestre, BSR6; esto es, ambas cepas expresan la proteína recombinante después de alcanzar el estado de transición (T<sub>0</sub>).

La transferencia del genotipo *hpr2* a la cepa JJ1 (cepa JJ2) produjo un incremento adicional, del 80%. Sin embargo, la transferencia del genotipo *degUh32* (cepa JJ4) no causó un aumento significativo, lo que nos hace suponer que la mutación *degUh32* y la mutación que forma la caja -35 consenso favorecen el reconocimiento de la secuencia promotora por la RNAP de manera similar, razón por la cual no observamos un incremento significativo.

La activación de DegU parece ser similar a la estimulación de la proteína activadora catabólica, CAP. El papel común de CAP, en la activación de la transcripción, es estimular eventos que dirijan la formación de un complejo abierto en un conjunto de promotores con una secuencia consenso pobre en las cajas -35 o -10, como es el caso de la caja -35 del promotor del gen *aprE*. Diversos operones, incluyendo los promotores de lactosa, galactosa y arabinosa en *E. coli* involucran la regulación a través de AMPc-CAP (Adhya 1987; Gang *et al.*, 1997; Tagami y Aibam 1998). En este sentido, se ha reportado que los sitios de unión de la proteína DegU en la región de regulación de *aprE* cubren los nucleótidos de -188 a -169 hacia arriba del sitio de inicio de la transcripción (Dartois, V. *et al.* 1998). Es posible que la interacción entre la forma fosforilada de DegU (DegU-PO<sub>4</sub><sup>3-</sup>) y la RNAP se necesite solo temporalmente para activar la transcripción de la fusión silvestre *aprE-lacZ*, cuya caja -35 se acerca muy poco a la secuencia consenso. Podrían ocurrir dos vías en las cuales DegU active la transcripción: 1) por su interacción directa con la RNAP, ó 2) por actuar sobre el DNA, cambiando su estructura de alguna manera que asista la unión de la RNAP.

Consideramos que los cambios en la caja -35 de las cepas JJ1, JJ2, JJ4 y JJ6 produjeron una región -35 funcional, que podría interaccionar independientemente con la RNAP. Por lo que la formación del complejo abierto es más efectivo y probablemente no se requiera más activación por parte de DegU-PO<sub>4</sub><sup>3-</sup>. Creemos que las modificaciones de la caja -35 hacia el consenso TTGACA (cepa JJ1) y la mutación que lleva el gen *degU32* tienen efectos "redundantes" sobre el inicio de la transcripción de *aprE-lacZ*; por tanto, solo un incremento minoritario en la actividad de β-Gal fue obtenido cuando ambas mutaciones actuaron juntas (cepa JJ4).

Cuando la doble mutante *hpr2* y *degU32* fue analizada, especulamos que tendría efectos aditivos e incrementos proporcionales a las mutaciones *hpr2* y *degU32*; sin embargo, esto no ocurrió. Nuestros resultados sugieren que aunque ambas mutaciones llevan a cabo sus funciones de manera independiente, el efecto de estas mutaciones no es completamente aditivo, como puede ser visto en la cepa doble mutante JJ7 (BSR6 *hpr2degU32*). Sin embargo, conforme los niveles de β-Gal fueron más elevados por la presencia de estas

mutaciones *cis*, el efecto del fondo fue menos significativo. Una posible explicación es que la carga metabólica sea limitante en las cepas sobreproductoras. Casos similares se han reportado en la sobreproducción de  $\beta$ -lactamasa (Togna *et al.*, 1993; Gu *et al.*, 1994) y en la producción de la insulina humana recombinante en *E. coli* (Schmidt M *et al.*, 1999). Además, se ha reportado que la concentración de ribosomas libres podría repercutir en la traducción de un mRNA: por consiguiente, la cantidad de mRNA de *lacZ* de las cepas JJ6 y JJ7 podría no reflejar la actividad final de  $\beta$ -Gal (Carrier *et al.*, 1998; Vind *et al.*, 1993) y tendríamos que medir el RNAm de estas cepas.

Una de las modificaciones realizadas a nivel traduccional fue la mutación que correspondiente al codón de inicio de la traducción GTG  $\rightarrow$  ATG. Con ello esperábamos un efecto positivo sobre la síntesis de la proteína recombinante AprE- $\beta$ Gal; sin embargo, no obtuvimos modificación significativa. Uno de los resultados más inesperados, fue el obtenido a través de la mutación que forma la secuencia Shine-Dalgarno consenso AAGGAGG. Lejos de tener efecto positivo sobre la actividad de  $\beta$ -Gal, observamos una marcada disminución. Con la finalidad de analizar si la mutación puntual en la secuencia SD altera la estructura secundaria en la región líder y por ende el bloqueo del inicio de la traducción, evaluamos la secuencia mutagenizada a través de un análisis de cómputo. Dicho análisis descartó por completo esta posibilidad.

La secuencia SD consenso para genes de *B. subtilis* es idéntica a la de los genes tardíos del fago T7 de *E. coli*. Paradójicamente se reportó la pérdida de eficiencia para traducir el mensajero del fago T7 por los ribosomas de *B. subtilis* a pesar de tener una secuencia SD fuerte (Hager y Raminowitz, 1985). Es posible que los ribosomas de *B. subtilis* reconozcan eficientemente los sitios de inicio en el mensajero, pero exista alguna influencia aún no descrita de los residuos vecinos a ser traducidos.

Si bien en algunos casos la clonación de un terminador de la transcripción en el extremo 3' de un gen ha sido empleado como una herramienta para aumentar la vida media del mensajero (Wong y Chang *et al.*, 1986), en el caso particular del terminador

transcripcional del gen *cryIIIA* de *B. thuringiensis* clonado en el extremo 3' del gen *lacZ*, no produjo incremento considerable en la estabilidad del mRNA, evaluado a través de la actividad de  $\beta$ -Gal. Tal parece que la protección en el extremo 3' del mRNA de *lacZ* no es suficiente para contender con la degradación por nucleasas. En posteriores estudios deberá considerarse la clonación de estructuras que protejan la degradación en el extremo 5', tal como se ha reportado en otros casos (Bechhofer y Dubnae, 1987; Agaisse y Lereclus, 1996; Carrier y Keasling, 1997). Una discusión más amplia de este evento se describe en el apartado discuss (pag. 55) del artículo "Construction of protein overproducer strain by an integrative approach".

Cabe mencionar que todas las mutaciones fueron analizadas a través de la fusión traduccional *aprE'-lacZ* integrada en el cromosoma de *B. subtilis* por un solo evento de recombinación homóloga en el locus de *amyE*. En este sentido, ha sido ampliamente reportado que la integración a cromosoma estabiliza los genes heterólogos producidos industrialmente (Peredelchuk y Bennet, 1997). Además, es posible incrementar la productividad mediante integraciones múltiples o amplificaciones de la región de interés. Existen varios métodos reportados para ello (Mori *et al.*, 1988; Kiel *et al.*, 1995). Así mismo, el nivel de expresión de un gen puede incrementarse si las condiciones de fermentación del cultivo son manejadas de manera adecuada. En este sentido hay numerosos reportes en los cuales ha sido posible incrementar la productividad de un sistema; por ejemplo, cuando se esparce aire enriquecido con oxígeno para alcanzar una densidad celular mayor (Fass *et al.*, 1989), se remueven los metabolitos que inhiben el crecimiento (Lee y Chang 1990), se usa un control automatizado (Paalme *et al.*, 1997) o combinaciones de las estrategias anteriores (Martínez *et al.*, 1997). Por lo tanto, el incremento en la expresión de  $\beta$ -Gal obtenido con la cepa JJ1 en este estudio, es considerado como un punto de inicio en la optimización de un proceso a gran escala.

## **IX Conclusiones**

De los resultados experimentales obtenidos en el desarrollo de este trabajo se pueden obtener las siguientes conclusiones relevantes:

- El análisis teórico utilizado en nuestro estudio relevó dos secuencias promotoras hipotéticas del gen *aprE*. Estos promotores están localizados a 7pb del sitio de inicio de la transcripción, tienen la misma caja -10, pero diferente caja -35. Determinamos por mutagénesis sitio-específico, que solamente una de ellas (TACTAA— 17 pb —TACAAT) es funcional *in vivo* bajo las condiciones ensayadas en este estudio. Sin embargo, no se descarta la posibilidad que el segundo promotor pudiera transcribir bajo condiciones no analizadas en este trabajo.
- Hemos establecido mediante ensayos de permutaciones circulares, que hay al menos dos *loci* curvos en la *rraprE*. Estos fragmentos están localizados a los lados de la secuencia de unión de la proteína DegU (alrededor de -100 a -200 pb en relación al sitio de inicio de la transcripción) que podrían jugar un papel importante en la regulación de la transcripción del gen *aprE*.
- De los factores analizados en la construcción de cepas sobreproductoras de *B. subtilis*, el cambio más importante correspondió a la formación de la caja -35 consenso de los promotores dependientes de  $\sigma^A$  en la *rraprE*, formando la cepa JJ1. Este cambio produjo un incremento de 100 veces la actividad con respecto a la cepa considerada como silvestre (BSR6). Es de importancia mencionar que la adición de 20 pb entre los sitios de pegado 3 y 4 de la proteína Hpr, en la cepa JJ16, produjo un incremento adicional, del 30%. La regulación de los promotores modificados en las cepas JJ1 y JJ16 presentan patrones similares al del promotor *aprE* silvestre; es decir, su expresión se presenta al inicio de la fase de transición.



- El efecto del fondo genético *hpr2* y *degU32* no es tan notorio en la cepa JJI (2.7 veces) como lo es en la cepa BSR6 (66 veces); sin embargo deberá evaluarse si este beneficio compensa los efectos pleiotrópicos causados por estas mutaciones.
- El cambio efectuado en el codón de inicio de la traducción, GTG→ ATG, no tuvo impacto en la expresión de la proteína recombinante; a diferencia de la modificación para la formación de la secuencia SD consenso, AAGGAGG, que causó efecto negativo sobre la expresión de  $\beta$ -Gal.
- La clonación del terminador transcripcional del gen *cryIIIA* de *Bacillus thuringiensis* en el extremo 3' de la fusión *aprE-lacZ*, aumentó la estabilidad del mensajero en un 10%.

## **X. PERSPECTIVAS**

En base a los resultados obtenidos, para trabajos futuros se recomienda:

- Mejorar el método predictivo de reconocimiento de promotores mediante la construcción de una matriz de frecuencia relativa ampliada, que incluya tanto la región de reconocimiento de la subunidad alfa de la RNAP como aquellos nucleótidos relevantes en promotores de bacterias Gram-positivas.
- Con el fin de reducir la degradación del mRNA de *lacZ*, clonar estructuras secundarias en su extremo 5' que lo protejan de la acción por nucleasas.
- Con el propósito de incrementar la producción de la proteína recombinante, construir cepas no productoras de las proteasas mayoritarias (*aprE*, *isp1*).
- Evaluar la productividad de las cepas JJ1 y JJ16 a nivel fermentador; para que mediante posteriores estudios se puedan manipular las condiciones de crecimiento y del medio de cultivo con la finalidad de aumentar la cantidad de proteína recombinante.
- Fusionar la *rraprE* a otras proteínas de interés comercial.
- Incluir las mutaciones sobre los nucleótidos reportados importantes en el reconocimiento de la subunidad alfa de la RNAP, cuyo consenso es: -59 nnAAA(A/T)(A/T)T(A/T)TTTTnnAAAAnnn -38.

## **XI. REFERENCIAS**

- Adhya, S. (1987). The galactose operon. In: Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M., Umberger, H.E. (eds) *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology. American Society for Microbiology, Washington, D.C. pp. 1503-1512.
- Agaisse, H. and Lereclus, D. (1996). STAB-SD: a Shine-Dalgarno sequence in the 5' untranslated region is a determinant of mRNA stability. *Mol. Microbiol.* 20: 633-643.
- Arbige, M.V. and Picher, W.H. (1989). Industrial enzymology: a look towards the future. *Trends in Biotechnol.* 7: 330-335.
- Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A. 3<sup>rd</sup>, Steever, A.B., Wach, A., Philippsen, P. and Pringle, J.F. (1998). Heterologous modules for efficient and versatile PCR-based gene targetin in *Schizosaccharomyces pombe*. *Yeast* 14: 943-951.
- Bai, U., Mulec-Mandic, I. and Smith, I. (1993). SinI modulates the activity of SinR, a developmental swith protein of *Bacillus subtilis*, by protein-protein interaction. *Genes and Development* 7: 139-148.
- Banner, C.D., Moran, C.P. Jr. and Losick, R. (1983). Deletion analysis of a complex promoter for a developmentally regulated gene from *Bacillus subtilis*. *J. Mol. Biol.* 168: 351-365.
- Bechhofer, D.H. and Dubnau, D (1987). Induced mRNA stability in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 84: 498-502.
- Bolaños, V.M. (1994). Sobreproducción de la enzima  $\beta$ -galactosidase de *Escherichia coli* en *Bacillus subtilis*. Tesis de Maestría en Biotecnología. UACPyP-CCH, Universidad Nacional Autónoma de México.
- Bron, S., Bosma, P., Van Belkum, M. and Luxen, E. (1987). Stability function in the *Bacillus subtilis* plasmid pTA 1060. *Plasmid.* 18: 8-15.
- Brosius, J., Cate, R.L. and Perlmutter, A.P. (1982). Precise location of two promoters for the beta-lactamase gene of pBR322. S1 mapping of ribonucleic acid isolated from *Escherichia coli* or synthesized *in vitro*. *J Biol Chem.* 257: 9205-9210.
- Burbulys, D., Trach, K.A. and Hoch, J.A. (1991). Initiation of sporulation in *Bacillus subtilis* is controlled by a multicomponent phosphorelay. *Cell.* 64: 545-552.
- Busby, S. and Ebright, R.H. (1994). Promoter structure, promoter recognition and transcription activation in prokaryotes. *Cell* 79: 743-746.
- Carrier, T., Jones, K.L. and Keasling, J.D. (1998). mRNA stability and plasmid copy number effects on gene expression from an inducible promoter system. *Biotechnol. Bioeng.* 55: 666-672.

Carrier, T.A. and Keasling, J.D. (1997). Engineering mRNA stability in *E. coli* by the addition of synthetic hairpins using a 5' cassette system. *Biotechnol. Bioeng.* 55: 577-580.

Carter, H.L. 3d, Wang, L.F., Doi, R.H. and Moran, C.P. Jr. (1988). *rpoD* operon promoter used by sigma H-RNA polymerase in *Bacillus subtilis*. *J. Bacteriol.* 170: 1617-1621.

Chen, Q.K., Hertz, G.Z. and Stormo, G.D.(1997). PromFD 1.0: a computer program that predicts eukaryotic pol II promoters using strings and IMD matrices. *Comput Appl Biosci.* 13: 29-35.

Chibazakura, T., Kawamura, F. and Takahashi, H. (1991). Differential regulation of *spoOA* transcription in *Bacillus subtilis*: Glucose represses promoter swithing at the initiation of sporulation. *J. Bacteriol.* 173: 2625-2632.

Crutz, A.M., Steinmetz, M., Aymerich, S., Richter, R. and Le Coq, D. (1990). Induction of levansucrase in *Bacillus subtilis*: an antitermination mechanism negatively controlled by the phosphotransferase system. *J. Bacteriol.* 172: 1043-1050.

Dartois, V., Débaarbouillé, M., Kunst, F. and Rapoport, G. (1998). Characterization of a novel member of the DegS-DegU regulon affected by salt stress in *Bacillus subtilis*. *J. Bacteriol.* 180: 1855-1861.

Debarbouille, M., Martin-Verstraete, I., Kunst, F. and Rapoport, G. (1991). The *Bacillus subtilis sigL* gene encodes an equivalent of  $\sigma_{54}$  from gram-negative bacteria. *Proc. Natl. Acad. Sci. USA.* 88: 9092-9096.

DeHaseth, P.L., Zupancic, M.L. and Record, M.T. Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.* 180: 3019-3025.

Dod, B.J. and Balassa, G. (1973). The kinetics of extracellular protease production in an abnormal sporulation mutant of *B. subtilis*. *Biochimie.* 55: 1005-1006.

Doi, R.H. (1984). Genetic engineering in *Bacillus subtilis*. *Biotechnol. Genet. Eng. Rev.* 2:121-155.

Doi, R.H. and Wang, L.F. (1986). Multiple procaryotic ribonucleic acid polymerase sigma factors. *Microbiol. Rev.* 50:227-243.

Ehrlinch, S.D., Noirot, P., Petit, M.A., et al. (1986). In: *Genetic Engineering*, vol.8 (Setlow, J.K. y Hollaender, A., eds). Plenum Press. New York. pp. 71-83.

Estrem, S.T., Gaal, T., Ross, W. and Gourse, R.L. (1998). Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. USA.* 95: 9761-9766.

- Fahnestock, S.R. and Fisher, K.E. (1986). Expression of the staphylococcal protein A gene in *Bacillus subtilis* by gene fusions utilizing the promoter from a *Bacillus amyloliquefaciens* alpha-amylase gene. *J. Bacteriol.* 165: 796-804.
- Fahnestock, S.R., Saunders, C.W., Guyer, M.S., Lofdahl, S., Guss, B., Uhlen, M. and Lindberg, M. (1986). Expression of the staphylococcal protein A gene in *Bacillus subtilis* by integration of the intact gene into the *B. subtilis* chromosome. *J. Bacteriol.* 165: 1011-1014.
- Fass, R., Clem, T.R. and Shiloach, J. (1989). Use of a novel air separation system in a fed-batch fermentative culture of *Escherichia coli*. *Appl. Environ. Microbiol.* 55: 1305-1307.
- Ferrari, E., Henner, D.J., Perego, M. and Hoch, J.A. (1988). Transcription of *Bacillus subtilis* and expression of subtilisin in sporulation mutants. *J. Bacteriol.* 170: 289-295.
- Ferrari, E., Jarnagin, A.S. and Schmidt, B.F. (1993). Commercial production of extracellular enzymes. In: Sonenshein, A. L., Hoch, J.A., Losick, R. (eds) *Bacillus subtilis* and other Gram-positive bacteria: Biochemistry, physiology and molecular genetics. American Society for Microbiology, Washington, D.C., pp. 917-937.
- Foulger, D. and Errington, J. (1991). Sequential activation of dual promoters by different sigma factors maintains *spoIVJ* expression during successive developmental stages of *Bacillus subtilis*. *Mol. Microbiol.* 5: 1363-1373.
- Fredrick, K., Caramori, T., Chen, Y.F., Galizzi, A. and Helmann, J.D. (1995). Promoter architecture in the flagellar regulon of *Bacillus subtilis*: high-level expression of by the sigma D RNA polymerase requires an upstream promoter element. *Proc. Natl. Acad. Sci. USA.* 92: 2582-2586.
- Freese, E. (1991). Initiation of bacterial sporulation. In: Levinspn, H.S., Sonenshein, A.L., Tipper, D.J. (eds) *In sporulation and germination*. American Society for Microbiology, Washington, D.C., pp. 1-12.
- FürbaB, R. and Marahel, M.A. (1991). Mutant analysis of interaction of the *Bacillus subtilis* transcription regulator AbrB with the antibiotic biosynthesis gene *trcA*. *FEBS* 1: 153-156.
- Gaal, T., Ross, W., Blatter, E.E., Tang, H., Jia, X., Krishnan, V.V., Assa-Munt, N., Ebright, R.H., and Gourse, R.L. (1996). DNA-binding determinants of the alpha subunit of RNA polymerase: novel DNA-binding domain architecture. *Genes Dev.* 10: 16-26.
- Gang, J.B. and Harman, J.G. (1997). CRP: cAMP complex binding to the *lac* operator region induces a structural change in *lac* DNA. *Mol. Cell.* 7: 444-447
- Gartner, D., Geissendorfer, M. and Hillen, W. (1988). Expression of the *Bacillus subtilis* *xyl* operon is repressed at the level of transcription and is induced by xylose. *J. Bacteriol.* 170: 3102-9

Gaur, N.K., Dubnau, E. and Smith, I. (1986). Characterization of a cloned *Bacillus subtilis* gene that inhibits sporulation in multiple copies. *J. Bacteriol.* 168: 860-869.

Gaur, N.K., Oppenheim, J. and Smith, I. (1991). The *Bacillus subtilis sin* gene, a regulator of alternate developmental processes, codes for a DNA-binding protein. *J. Bacteriol.* 173: 678-686.

Gilman, M.Z., Wings, J.L. and Chamberlin, M.J. (1981). Nucleotide sequence of two *Bacillus subtilis* promoters used by *Bacillus subtilis* sigma-28 RNA polymerase. *Nucleic Acids Res.* 9: 5991-6000.

Glick, B.R. and Pasternak, J.J. (1994). The molecular biotechnology revolution. In: *Molecular Biotechnology principles and applications of recombinant DNA*. American Society for Microbiology, Washington, D.C. pp. 5-15.

Gold, L. and Stormo, G.D. (1990). High-level translation initiation. *Methods Enzymol.* 185: 89-93.

Gourse, R.L. de Boer, H.A. and Nomura, M. (1986). DNA determinants of rRNA synthesis in *E. coli*: Growth rate dependent regulation, feedback inhibition, upstream activation, antitermination. *Cell* 44: 197-205.

Graves, M.C. and Rabinowitz, J.C. (1986). In vivo and in vitro transcription of the *Clostridium pasteurianum* ferredoxin gene. Evidence for "extended" promoter elements in gram-positive organisms. *J. Biol. Chem.* 261: 11409-1115.

Grossman, A.D. (1991). Integration of developmental signal and the initiation of sporulation in *Bacillus subtilis*. *Cell.* 65, 5-8.

Grossman, A.D. and Losick, R. (1988). Extracellular control of spore formation in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 85: 4369-4373.

Gu, M.B., Todd, P. and Kompala, D.S. (1994) Analysis of foreign protein overproduction in recombinant CHO cells. Effect of growth kinetics and cell cycle traverse. *Ann. N. Y. Acad. Sci.* 721: 194-207

Hall, M.N., Gabay, J., Débarbouillé, M. and Schwartz, M. (1982). A role for mRNA secondary structure in the control of translation initiation. *Nature.* 295: 616-618

Harwood, C.F. and Archibald, A.R. (1990). Growth, maintenance, and general techniques. In: *Molecular biological methods for Bacillus*. Harwood, C.F. and Cutting, S.M (eds). John Wiley & Sons, Chichester, England, pp. 1-26.



- Harwood, C.R. (1992). *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. TIBTECH 10: 247-256.
- He, X.S., Bruckner, R. and Doi, R.H. (1991). The protease genes of *Bacillus subtilis*. Res. Microbiol. 142:797-803.
- Helmann, J.D. (1995). Compilation and analysis of *Bacillus subtilis* sigma A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic. Acids Res. 23: 2351-60.
- Henkin, T.M. and Sonenshein, A.L. (1987). Mutations of the *Escherichia coli lacUV5* promoter resulting in increased expression in *Bacillus subtilis*. Mol. Gen. Genet. 209: 467-474.
- Henner, D.J., Ferrari, E., Perego, M. and Hoch, J.A. (1988a). Location of the targets of the *hpr-97*, *sacU32* (Hy) and *sacQ36* (Hy) mutations in upstream regions of the subtilisin promoter. J. Bacteriol. 170: 296-300.
- Henner, D.J., Yang, M. and Ferrari, E. (1988b). Localization of *Bacillus subtilis sacU* (Hy) mutations to two linked genes with similarities to the conserved prokaryotic family of two-component signalling systems. J. Bacteriol. 170: 5102-5109.
- Higerd, T.B., Hoch, J.A. and Spizizen, J. (1972). Hyperprotease-producing mutants of *Bacillus subtilis*. J. Bacteriol. 112: 1026-1028.
- Hoch, J. A. (1991) Genetic analysis in *Bacillus subtilis*. Methods Enzymol. 204: 305-320.
- Honjo, M., Nakayama, A. and Fukazawa, K. (1990). A novel *Bacillus subtilis* gene involved in negative control of sporulation and degradative enzyme production. J. Bacteriol. 172: 1783-1790.
- Huo, L., Martin, K.J. and Schleif, R (1988). Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. Proc. Natl. Acad. Sci U.S.A. 85: 5444-8.
- Imanaka, T., Himeno, T. and Aiba, S. (1987). Cloning and nucleotide sequence of the penicillinase antirepressor gene *penJ* of *Bacillus licheniformis*. J Bacteriol. 169: 3867-72.
- Ito, J. and Spizizen, J. (1971). Increased rate of asporogenous mutations following treatment of *Bacillus subtilis* spores with ethyl methansulfonate. Mutant Res. 13: 93-96
- Izui, K., Miwa, T., Kajitani, M., Fujita, N., Sabe, H., Ishihama, A. and Katsuki, H. (1985). Promoter analysis of the phosphoenolpyruvate carboxylase gene of *Escherichia coli*. Nucleic Acids Res. 13: 59-71.
- Jacobs, M.F. (1995). Expression of the subtilisin Carlsberg-encoding gene in *Bacillus licheniformis* and *Bacillus subtilis*. Gene 152: 69-74.

Jan J., Valle F., Bolivar, F. and Merino, E. (2000). Characterization of the 5' subtilisin (*aprE*) regulatory region from *Bacillus subtilis*. FEMS Microbiology Letter. 183: 9-14.

Johnson, W.C., Moran, C.P. Jr. and Losick, R. (1983). Two RNA polymerase from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. Nature. 302: 800-804.

Kallio, P., Simonen M., Palva, I. and Sarvas, M. (1986). Synthesis of OmpA protein of *Escherichia coli* K12 in *Bacillus subtilis*. J. Gen. Microbiol. 132: 677-687.

Kiel, J.A., ten Berge, A.M., Borger, P. and Venema, G. (1995). A general method for the consecutive integration of single copies of a heterologous gene at multiples locations in the *Bacillus subtilis* chromosome by replacement recombination. Appl. Environ. Microbiol. 61: 4244-4250.

Kunisawa, T. (1995) Alteration in synonymous arginine codon preferences of *Bacillus subtilis* during sporulation. J. Theor. Biol. 172: 387-90.

Kunkel, B. (1991). Compartmentalized gene expression during sporulation in *Bacillus subtilis* in oxygen-controlled, glucose fed-batch fermentations. Appl. Microbiol. Biotechnol. 28: 404-408.

Kunst, F., Msadek, T. and Rapoport, G. (1994) Signal transduction network controlling degradative enzyme synthesis and competence in *Bacillus subtilis*. In: Piggot P.J., Moran, C.P., Youngman, P. (eds) Regulation of bacterial differentiation. American Society for Microbiology, Washington, D.C. pp. 1-20.

Kupsch, J. (1986). Analysis of structural and biological parameters affecting plasmid deletion in *Bacillus subtilis*. In 8<sup>th</sup> European Meeting on Genetic Transformation (Abstract). Uppasala, Sweden. pp. 65

LeDeaux, J.R., Yu, N. and Grossman, A.D. (1995). Different roles for KinA, KinB, and KinC in the initiation of sporulation in *Bacillus subtilis*. J. Bacteriol. 177: 861-863.

Lee, D.C. and Roeder, R.G. (1981). Transcription of adenovirus type 2 genes in a cell-free system: apparent heterogeneity of initiation at some promoters. Mol. Cell Biol. 7: 635-651.

Lee, S.Y. (1996). High cell-density culture of *Escherichia coli*. Trends in Biotechnol. 14: 98-105.

Lee, Y.L. and Chang, H.N. (1990). High cell density culture of a recombinant *Escherichia coli* producing penicillin acylase in a membrane cell recycle fermentor. Biotechnol. Bioeng. 36: 330-337.

- Lindsey, K., Wei, W., Clarke, M.C., McArdle, H.F., Rooke, L.M. and Topping, J.F. (1993). Tagging genomic sequences that direct transgene expression by activation of a promoter trap in plants. *Transgenic Res.* 2: 33-47.
- Lisser and Margalit (1994). Determination of common structural features in *Escherichia coli* promoters by computer analysis. *Eur J. Biochem.* 223: 823-830
- López, J.M. Dromerick, A. and Freese, G. (1981) Response of guanosine 5-triphosphate concentration to nutritional changes and its significance for *Bacillus subtilis*. *J. Bacteriol.* 146: 605-613.
- López, J.M., Marks, L.L. and Freese, E. (1979). The decrease of guanine nucleotides initiates sporulation of *Bacillus subtilis*. *Biochem. Biophys. Acta* 587: 238-252.
- López, J.M., Uratani-Wong B and Freese, E. (1980). Catabolite repression of enzyme synthesis does not prevent sporulation. *J. Bacteriol.* 141: 1447-1449.
- Lloubes, R., Granger-Schnarr, M., Lazdunski, C. and Schnarr, M. (1988). LexA repressor induces operator-dependent DNA bending. *J. Mol. Biol.* 204: 1049-1054.
- Ludin, B., Doll, T., Meilli, R., Kaech, S. and Matus, A. (1996). Application of novel vectors for GFP-tagging of protein to study microtubule-associated proteins. *Gene.* 173: 107-111.
- Marino, M.H. (1989). Expression systems for heterologous protein production. *BioPharm.* 2: 18-33.
- Martínez, A., Ramírez, O.T. and Valle, F. (1997). Improvement of culture conditions to overproduce  $\beta$ -galactosidase from *Escherichia coli* in *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.* 47: 40-45.
- McAllister, C.F. and Achberger, E.C. (1988) Effect of polyadenine-containing curved DNA on promoter utilization in *Bacillus subtilis*. *J. Biol. Chem.* 263: 11743-11749.
- McClure, W.R. (1985). Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* 54: 171-204.
- Moran, C.P., Jr., Lang, N., LeGrice, S.F.J., Lee, G., Stephens, M., Sonnenshein, A.L., Pero, J. and Losick, R. (1982). Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. *Mol. Gen. Genet.* 186: 339-346.
- Moran, C.P., Jr. (1989). Sigma factors and the regulation of transcription. In: Smith, I., Slepecky, R. and Setlow, P. (ed.). *Regulation of Prokaryotic Development*. American Society for Microbiology. Washington, D.C. pp. 167-184.

- Mori, M., Hashigushi, K.I., Yoda, K. and Yamasaki, M. (1988) Designed gene amplification on the *Bacillus subtilis* chromosome. *J. Gen. Microbiol.* 134: 85-95
- Mountain, A. (1989). *Bacillus*. In: Biotechnology Handbooks. (Harwood, CR., ed.) Plenum Press, New York, pp. 73-114.
- Msadek, T., Kunst, F., Henner, D., Klier, A., Rapoport, G. and Dedonder, R. (1990). Signal transduction pathway controlling synthesis of a class of degradative enzymes in *Bacillus subtilis*: Expression of the regulatory genes and analysis of mutations in *degS* and *degU*. *J. Bacteriol.* 172: 824-834.
- Mulligan, M.E., Hawley, D.K., Entirken, R. and McClure, W.R. (1984) *Escherichia coli* promoter sequences predicted in vitro RNA polymerase selectivity. *Nucleic Acids Res.* 12: 789-800.
- Murakami, K., Fujita, N., Ishihama, A. (1996). Transcription factor recognition surface on the RNA polymerase alpha subunit is involved in contact with the DNA enhancer element. *EMBO J.* 15: 4358-4367.
- Nagarajan, V., Albertson, H., Che, M. and Ribbe, J. (1992). Modular expression and secretion vectors for *Bacillus subtilis*. *Gene* 114: 121-126.
- Nagarajan, V., Ramaley, R. Alberrson, H. and Chen, M. (1993). Secretion of streptavidin from *Bacillus subtilis*. *Applied Environmental Microbiol.* 59: 3894-3898.
- Nicholson, W.L., Sun, D., Setlow, B. and Setlow, P. (1989). Promoter specificity of  $\sigma^G$ -containing RNA polymerase from sporulating cells of *Bacillus subtilis*: identification of a group of forespore-specific promoters. *J. Bacteriol.* 171: 2708-2718.
- Oehler, S., Eismann, E.R., Kramer, H. and Muller-Hill, B. (1990). The three operators of the *lac* operon cooperate in repression. *EMBO J.* 9:973-979.
- Paalme, T., Elken, R., Kahru, A., Vanatalu, K. and Vilu, R. (1997). The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach. *Antonie Van Leeuwenhoek* 71: 217-30
- Park, S.S., Wong, S.L., Wang, L.F. and Doi, R.H. (1989). *Bacillus subtilis* subtilisin gene (*aprE*) is expressed from a  $\sigma^A$  ( $\sigma^{43}$ ) promoter in vitro and in vivo. *J. Bacteriol.* 171: 2657-2665.
- Peredelchuk, M.Y. and Bennett, G.N. (1997). A method for construction of *E. Coli* with multiple DNA insertions in the chromosome. *Gene* 187: 231-238
- Perego, M., Spiegelman, G.B. and Hoch, J.A. (1988). Structure of the gene for the transition state regulator, *abrB*: regulator synthesis is controlled by the *spoOA* sporulation gene in *Bacillus subtilis*. *Molec. Microbiol.* 2: 689-699.

Perego, M., Wu, J., Spiegelman, G.B. and Hoch, J.A. (1991). Mutational dissociation of the positive and negative regulatory properties of the SpoOA sporulation transcription factor of *Bacillus subtilis*. *Gene* 100: 207-212.

Pérez-Martín, J., Rojo, F. and Lorenzo, V. (1994) Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol. Rev.* 58: 268-290.

Peschke, U., Beuck, V., Bujard, H., Gentz, R. and Le Grice, S. (1985). Efficient utilization of *Escherichia coli* transcriptional signals in *Bacillus subtilis*. *J. Mol. Biol.* 186: 547-555.

Petit, M.A., Joliff, G., Mesas, J.M., Klier, A., Rapoport, G. and Ehrlich, S.D. (1990). Hypersecretion of a cellulase from *Clostridium thermocellum* in *Bacillus subtilis* by induction of chromosomal DNA amplification. *Biotechnology*. 8: 559-63.

Pierce, A.J. (1992). Engineered *Bacillus subtilis* based subtilisin production systems: Physiological, genetic and morphological consideration. Ph.D. Thesis pp. 10-11.

Plaskon, R.R. and Wartell, R.M. (1987) Sequence distributions associated with DNA curvature are found upstream of strong *E. coli* promoters. *Nucleic Acids Res.* 15: 785-796.

Priest, G.F. (1989). Products and applications. In: Harwood, C.F. (ed). *Bacillus*, biotechnology handbooks, vol. 2. Plenum Press, New York. pp. 293-322.

Priest, G.F. (1993). Systematics and ecology of *Bacillus*. In: Sonenshein, A.L. Hoch, J.A. and Losick, R. (eds). *Bacillus subtilis* and other Gram-positive bacteria: Biochemistry, physiology and molecular genetics. American Society for Microbiology, Washington, D.C. pp. 3-16.

Robertson, J.B., Gocht, M., Marahiel, M.A. and Zuber, P. (1989). AbrB, a regulator of gene expression in *Bacillus*, interacts with the transcription initiation regions of a sporulation gene and an antibiotic biosynthesis gene. *Proc. Natl. Acad. Sci. USA.* 86: 8457-61.

Roels, S.A., Driks, A. and Losick, R. (1992). Characterization of spo1VA, a sporulation gene involved in coat morphogenesis in *Bacillus subtilis*. *J. Bacteriol.* 174: 575-585.

Ross, W., Aiyar, S.E., Salomon, J. and Gourse, R.L. (1998). *Escherichia coli* promoters with UP elements of different strengths: modular structure of bacterial promoters. *J. Bacteriol.* 180: 5375-83.

Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R.L. (1993). A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* 262: 1407-13.

Satola, S.W., Baldus, J.M. and Moran, C.P. (1992). Binding of SpoOA stimulates *spoIIIG* promoter activity in *Bacillus subtilis*. *J. Bacteriol.* 174: 1448-1453.

- Satola, S.W., Kirchman, P.A. and Moran, C.P. (1991). SpoOA binds to a promoter used by sigma A RNA polymerase during sporulation in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA. 88: 4533-4537.
- Schaeffer, P., Millet, J. and Aubertm J.P. (1965). Catabolic repression of bacterial sporulation. Proc. Natl. Acad. Sci. USA. 54: 704-711.
- Schmidt, M., Babu, K.R., Khanna, N., Marten, S. and Rinas, U. (1999). Temperature-induced production of recombinant human insulin in high-cell density cultures of recombinant *Escherichia coli*. J. Biotechnol. 68: 71-83
- Sekiguchi, J., Takada, N. and Okada, H. (1975) Genes affecting the productivity of alfa-amylase in *B. subtilis* Marburg. J. Bacteriol. 121: 688-694.
- Setlow, P (1994). DNA structure, spore formation and spore properties. In: Piggot, J.P., Moran, P.C. and Youngman, P. (ed). Regulation of bacterial differentiation. American Society for Microbiology, Washington, D.C. pp. 181-194.
- Shimotsu, H. and Henner, D.J. (1986). Construction of a single-copy integration vector and its use in analysis of regulation of the *trp* operon of *Bacillus subtilis*. Gene 43: 85-94.
- Slepecky, A.R. (1992). What is a *Bacillus*. In: Doi, H.R. and McGloughlin, M. (eds). Biology of *Bacilli*: Applications to industry. Butterworth-Heinemann, Boston. pp. 1-21.
- Sonenshein, A. L. (1989). Metabolic regulation of sporulation and other stationary-phase phenomena. In: Smith, I., Slepecky, R.A. and Setlow, P. (eds) Regulation of prokaryotic development: Structural and functional analysis of sporulation and germination. American Society for Microbiology, Washington, D.C. pp. 109-130.
- Stock, J. B., Ninfa, A.J. and Stock, A.M. (1989). Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. 53: 450-490.
- Strauch, M.A. and Ayazifar, M. (1995). Bent DNA is found in some but not all, regions recognized by the *Bacillus subtilis* AbrB protein. Mol. Gen. Genet. 246: 756-760.
- Strauch, M and Hoch, J.A. (1992a). Control of post-exponential gene expression by transition state regulators. In: Doi, H.R. and McGloughlin, M. (eds) Biology of *Bacilli*: Applications to industry. Butterworth-Heinemann, Boston. pp. 105-121.
- Strauch, M. and Hoch, J.A. (1992b). Sporulation in prokaryotes and lower eukaryotes. Current Opinion in Genetics and Development 2: 799-804.

- Strauch, M.A., Spiegelman, G.B., Perego, M., Johnson, W.C., Burbulys, D. and Hoch, J.A. (1989). The transition state transcription regulator AbrB of *Bacillus subtilis* is a DNA binding protein. *EMBO J.* 8: 1616-1621.
- Strauch, M., Webb, V., Spiegelman, G. and Hoch, J.A. (1990). The SpoOA protein of *Bacillus subtilis* is a repressor of the *aprB* gene. *Proc. Natl. Acad. Sci. USA.* 87: 1801-1805.
- Struhl, K. (1985) Naturally occurring poly (dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. USA* 82: 8419-8423.
- Sun, D., Fajardo-Cavazos, P., Susman, D., Tovar-Rojo, F., Cabrera-Martínez and Setlow, P. (1991). Effect of chromosome location of *Bacillus subtilis* forespore genes in their *spo* gene dependence and transcription by  $E\sigma^F$ : Identification of features of good  $E\sigma^F$ -dependent promoters. *J. Bacteriol.* 173: 7867-7874.
- Tagami, H. and Aiba, H. (1998). A common role of CRP in transcription activation: CRP acts transiently to stimulate events leading to open complex formation at a diverse set of promoters. *EMBO J.* 17: 1759-67
- Tatti, K.M., Carter III, H.L., Moir, A. and Moran, C.P. Jr. (1989). Sigma H-directed transcription of *citG* in *Bacillus subtilis*. *J. Bacteriol.* 171: 5928-5932.
- Tatti, K.M. and Moran, C.P. Jr. (1984). Promoter recognition by sigma-37 RNA polymerase from *Bacillus subtilis*. *J. Mol. Biol.* 175: 285-297.
- Togna, A.P., Shuler, M.L. and Wilson, D.B. (1993) Effects of plasmid copy number and runaway plasmid replication on overproduction and excretion of beta-lactamase from *Escherichia coli*. *Biotechnol. Prog.* 9: 31-9
- Valle F. (1994). PCR-based method to map the bending locus of DNA molecules. *PCR Methods Appl.* 4: 44-45.
- Valle, F. and Ferrari, E. (1989). Subtilisin: a redundantly temporally regulated gene. In : Smith, I., Slepecky, A.R. and Setlow, R. (eds). Regulation of prokaryotic development. American Society for Microbiology, Washington, D.C. pp. 131-146.
- Vellanowell, R.L. (1993). Translation and regulation. In: Hoch, J.A. and Losick, R. (eds.). *Bacillus subtilis* and other Gram-positive bacteria. American Society for Microbiology. Washington, D.C. pp. 699-711.
- Vind, J., Sorensen, M.A., Rasmussen, M.D. and Pedersen, S. (1993). Synthesis of proteins in *Escherichia coli* is limited by the concentration of free ribosomes. *J. Mol. Biol.* 231: 678-688.
- Voskuil, M.I. and Chambliss, G.H. (1998). The -16 region of *Bacillus subtilis* and other gram-positive bacterial promoters. *Nucleic Acids Res.* 26: 3584-3590.

- Voskuil, M.I., Voepel, K. and Chambliss, G.H. (1995). The -16 region, a vital sequence for the utilization of a promoter in *Bacillus subtilis* and *Escherichia coli*. *Mol. Microbiol.* 17: 271-279.
- Waldburger, L., González, D. and Chambliss, G.H. (1993). Characterization of a new sporulation factor in *Bacillus subtilis*. *J. Bacteriol.* 175: 6321-6327.
- Wang, L.F. and Doi, R.H. (1990). Complex character of *senS*, a novel gene regulating expression of extracellular-protein genes of *Bacillus subtilis*. *J. Bacteriol.* 172: 1939-1947.
- Wang, L.F., Wong, S.L., Lee, S.G., Kalyan, N., Hung, P.P., Hilliker, S. and Doi, R.H. (1988a). Expression and secretion of human atrial natriuretic  $\alpha$ -factor in *Bacillus subtilis* using the subtilisin signal peptide. *Gene* 69: 39-47.
- Wang, L.F., Wong, S.L., Park, S.S. and Doi, R.H. (1988b). Isolation and characterization of a novel *Bacillus subtilis* and *Bacillus natto* gene that enhances protease production. In: Ganesan, A. and Hoch, J.A. (ed). *Bacillus* Molecular Genetics and Biotechnology Applications. Academic Press Inc. New York. Vol. 3 pp. 45-60.
- Wong, S.L. (1995). Advances in the use of *Bacillus subtilis* for the expression and secretion of heterologous proteins. *Current Opinion in Biotechnol.* 6: 517-522.
- Wong, H.C and Chang, S. (1986). Identification of a positive retroregulator that stabilizes mRNAs in bacteria. *Proc. Natl. Acad. Sci. USA.* 83: 3233-3237.
- Yamashita, S., Kawamura, F. and Yoshikawa, H. (1989). Dissection of the expression signals of the *spoOA* gene of *Bacillus subtilis*: glucose sporulation specific expression. *J. Gen. Microbiol.* 135: 1335-1345.
- Yansura, D.G. and Henner, D.J. (1984). Use of the *Escherichia coli lac* repressor and operator to control gene expression in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA.* 81: 439-443.
- York, K., Kenney, T.J. and Satola, S. (1992). SpoOA controls the sigA-dependent activation of *Bacillus subtilis* sporulation-specific transcription unit *spoIIR*. *J. Bacteriol.* 174: 2648-2658.
- Zukowski, M.M. (1992). Production of commercially valuable products. In: Doi, H.R., and McGloughlin M. (eds). *Biology of Bacilli: Applications to industry*. Butterworth-Heinemann, Boston. pp. 311-337.