



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

INSTITUTO DE BIOTECNOLOGIA

ESTUDIO DE LA UNIÓN DE TPP POR EL “RIBOSWITCH” DE
TIAMINA, MEDIANTE MUTACIONES PUNTUALES DE LAS
BASES CONSERVADAS EN LAS ASAS DEL “RIBOSWITCH”

TESIS

QUE PARA OBTENER EL GRADO DE
M A E S T R A E N C I E N C I A S
P R E S E N T A :
BIOLOGA NANCY GUADALUPE ONTIVEROS PALACIOS

DIRECTOR DE TESIS: DR. JUAN MIRANDA RÍOS

JURADO DE EXAMEN:

PRESIDENTE DR. ENRIQUE MERINO PÉREZ
SECRETARIOS DR. JUAN MIRANDA RÍOS
VOCAL: DR. ALFREDO TORRES LARIOS
SUPLENTE DR. CEI LEANDER GASTON ABREU GOODGER
SUPLENTE DR. ARTURO C. BECERRA BRACHO





Universidad Nacional
Autónoma de México



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.

ÍNDICE

1.	Prefacio	2
2.	Marco teórico	3
2.1.	¿Qué es un riboswitch?	3
2.2.	La regulación génica	4
2.3.	Mecanismo de Regulación de los riboswitches	6
2.4.	Biosíntesis de tiamina	8
3.	Antecedentes	9
3.1.	Estudios de regulación	9
3.2.	Estudios de genómica comparativa	12
3.3.	Estudios estructurales	13
4.	Propuesta	16
5.	Hipótesis	16
6.	Objetivos	16
6.1.	General	16
6.2.	Particulares	16
7.	Materiales	17
7.1.	Microorganismos	17
7.2.	Plásmidos	18
8.	Métodos	19
8.1.	Selección de la secuencia mínima de regulación (<i>rrthiM</i>) y de las bases conservadas en las asas del riboswitch	19
8.2.	Clonación de la <i>rrthiM</i>	19
8.3.	Mutagénesis sitio dirigida	20
8.4.	Fusiones traduccionales	22
8.5.	Ensayos de β -galactosidasa	23
9.	Resultados y discusión	25
9.1.	Selección de la región reguladora del gene <i>thiM</i>	25
9.2.	Selección de las bases conservadas en las asas del riboswitch.	26
9.3.	Condiciones para ver la represión por tiamina.	29
9.4.	Fenotipos en las mutantes	31
9.4.1.	Actividad de β -galactosidasa en las mutantes	31
9.4.2.	Estructura del riboswitch de tiamina y resumen de fenotipos de las mutantes	32
9.4.3.	Mutantes en la hélice sensora de HMP	33
9.4.4.	Mutantes en la hélice sensora del pirofosfato (PP)	34
10.	Conclusión	38
11.	Referencias	39
12.	Apéndice 1. Los distintos riboswitches	43
13.	Apéndice 2. “Conserved regulatory motifs in bacteria: Riboswitches and beyond”	47

Es nuestro interés señalar esta perspectiva para los estudiantes que desean realizar investigación en evolución; no tienen que decidir entre ser paleontólogos, taxónomos, genetistas o ecólogos, sino simplemente evolucionistas; solo deben tener muy claros los problemas que quieren estudiar, los principios de la biología evolutiva y la mente abierta para aprender técnicas y métodos nuevos
Juan Núñez-Farfán y Luis E. Eguiarte (1999)

1. Prefacio

La motivación de este trabajo ha sido estudiar al RNA y en particular estudiar a los ribo-interruptores o “riboswitches” que son moléculas de RNA a las cuales, de forma muy personal, encuentro fascinantes.

¿Y por qué estudiar a los riboswitches?. Primero que nada, el RNA es una molécula muy versátil. Esta molécula es capaz de autoreplicarse, funcionar como transcrito, catalizar distintos procesos (incluyendo procesos básicos como la síntesis peptídica) y regular la expresión génica a distintos niveles. Si ésto no fuera suficiente, yo creo que el RNA puede seguir sorprendiéndonos aún con otras funciones.

Antes del 2001, cuando formalmente se empieza a definir el mecanismo de regulación por riboswitches, ya se habían reportado varios casos de represión génica dependiente de un metabolito en genes relacionados a la biosíntesis del mismo. Durante mucho tiempo los estudios se dirigieron sin mucho éxito a encontrar proteínas reguladoras que estuvieran involucradas en el proceso. Nadie quería suponer que una sencilla molécula de RNA pudiera reconocer específicamente a un metabolito y sin embargo lo hace. En el caso de los riboswitches es interesante ver cómo un RNA, que es una molécula monótona, ya que básicamente está formada por la combinación de 4 tipos de bases (adenina, uracilo, guanina y citosina) es capaz de funcionar como sensor y regulador de manera eficiente. En este caso es quizá la solución más sencilla para la regulación. Aunque se han realizado varios estudios de la interacción entre riboswitches y sus ligandos, aún estamos lejos de definir en forma detallada cómo es que se dan estas interacciones.

Este trabajo es también una continuación a lo realizado en la licenciatura, donde se trabajó en un método de búsqueda de riboswitches, el cual fue reportado en la tesis “Identificación de secuencias de regulación “riboswitches” en la biosíntesis de vitaminas y cofactores” y en el propongo abordar el estudio de cómo se da el reconocimiento entre el RNA y su molécula efectora para el caso del riboswitch de tiamina desde una perspectiva molecular, estudiando las consecuencias que se dan en la regulación cuando se modifican bases conservadas de la secuencia del riboswitch. Este riboswitch es un caso muy interesante porque está ampliamente distribuido en eubacterias y a diferencia de otros riboswitches, está presente en arqueobacterias y algunos eucariotes.

1. Marco teórico

1.1. ¿Qué es un riboswitch?

Un riboswitch es un segmento de RNA¹ capaz de reconocer a un metabolito específico y regular la expresión de los genes a los que se encuentra asociado. Están formados por dos dominios: uno que reconoce al ligando, llamado “aptámero” y otro que funciona como plataforma de expresión (Fig. 1).



Fig. 1. Esquema de los dominios que componen a un riboswitch. La respuesta al ligando puede variar prendiendo o apagando al gen.

Los riboswitches tienen características muy particulares, algunas son:

1. El mRNA forma un sitio de unión altamente selectivo a un metabolito específico (moléculas efectoras como: vitaminas y cofactores), sin intervención de proteínas.
2. La unión del metabolito provoca un reconocimiento allostérico de la estructura del mRNA que lo estabiliza y provoca a su vez alteraciones en la expresión genética.
3. El mRNA toma estructuras alternativas dependiendo de la unión del metabolito.
4. La regulación se da en genes de la biosíntesis del metabolito.

Los primeros estudios que dieron indicio del mecanismo de acción de los riboswitches reportaron que algunos metabolitos como las vitaminas cianocobalamina, tiamina, riboflavina o aminoácidos como metionina y lisina, causan la represión de genes involucrados en la biosíntesis y transporte de los mismos [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.]. En estos genes se buscaron experimentalmente las proteínas que pudiesen estar regulando los genes, sin embargo no se encontraron. Uno de los primeros casos estudiados es el del gene *btuB* de *Escherichia coli* y del operón *cob* en *Salmonella typhimurium* [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.], se observó que a determinada concentración de vitamina B₁₂ se inhibe la expresión de estos genes. Ahora se sabe que estos transcritos reconocen a la vitamina B₁₂ y que la unión de ésta reprime su síntesis. Desde entonces

¹ RNA: ácido ribonucleico. Puede estar en forma de RNA mensajero o transcrito (mRNA), de RNA de transferencia (tRNA) o RNA ribosomal (rRNA).

se han reportado secuencias de RNA, capaces de reconocer metabolitos específicos y regular la expresión de los genes relacionados a su biosíntesis y transporte (Apéndice 1).

1.2. La regulación génica en los procariotes

Los procariotes coordinan la expresión de varios genes agrupándolos de forma contigua en un mismo transcrito. Cada transcrito está compuesto por una secuencia promotora², una secuencia operadora³ y por uno o varios genes⁴. Asimismo, la regulación puede darse a nivel de la transcripción o a nivel de la traducción. En la transcripción: activando al gene (control positivo) o apagándolo (control negativo) y mediante modificaciones al transcrito. En la Tabla 1 se presenta un panorama general de los mecanismos de regulación presentes en procariotes [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.].

Tabla 1
Mecanismos de regulación génica

1) A nivel transcripcional

- a) **Inicio de la transcripción**
 - i) Secuencia promotor y factores σ alternativos de la RNAPol
 - ii) Interacción de proteínas reguladoras (factores de transcripción) sobre sitios específicos de DNA
- b) **Terminación prematura de la transcripción o atenuación**
 - i) Terminación intrínseca
 - ii) Terminación por un factor
- c) **Estabilidad y modificaciones del mRNA**

2) A nivel traduccional.

- a) **Eficiencia de unión del ribosoma al mRNA**
- b) **Interferencia con el sitio "Shine Dalgarno"**
 - i) Por intervención de proteínas
 - ii) Por intervención de RNAs interferentes
 - iii) Por la estructura secundaria del transcrito

3) A nivel post-traducciona

Estabilidad y localización subcelular de la proteína; modificaciones a la proteína y función de la proteína

Inicio de la transcripción

Existen 2 tipos de promotores, los fuertes, con un alto índice de transcripción, y los débiles, con una transcripción más reducida, que dependen de su afinidad por la enzima RNAPol. La RNAPol es una enzima compuesta por seis subunidades (2 α , β , β' , ω y σ), de las cuales la subunidad σ es la encargada de reconocer a la secuencia promotora, esta subunidad es variable y reconoce secuencias específicas de promotor, de esta forma se coordina la expresión de un conjunto de genes que contienen una misma secuencia promotora [¡Error!No se encuentra el origen de la referencia.].

² Secuencia promotora: (o promotor) es la secuencia que reconoce la RNA polimerasa (RNAPol) a partir de la cual se inicia la transcripción.

³ Secuencia operadora: (o operador) región de DNA en bacterias que controla la transcripción de un gen adyacente, generalmente por la unión de una proteína reguladora.

⁴ Operón: unidad de expresión genética que consiste en uno o varios genes, además de la secuencia promotora y operadora que regulan su transcripción.

Muchos otros procesos de regulación dependen de proteínas como factores de transcripción, las cuales reconocen secuencias específicas de DNA ayudando a determinar los genes que se han de transcribir o bloqueando la transcripción en otros casos. Los factores que favorecen la transcripción, son llamados reguladores positivos o activadores y los que bloquean la transcripción, son reguladores negativos o represores. Para que la proteína sea capaz de interactuar con el DNA es necesario que esté en su estado activo para lo cual se necesita la unión de un cofactor a la misma, que generalmente son metabolitos relacionados al gen que regulan.

Terminación prematura de la transcripción o atenuación

Cuando los genes están regulados por atenuación, el líder del transcrito⁵ adopta una estructura de terminador transcripcional⁶ que pausa la transcripción y provoca que la RNA polimerasa (RNAPol)⁷ termine y libere tanto el transcrito como el templado de DNA. La atenuación transcripcional puede ser dirigida por otros factores, un ejemplo es la terminación dependiente de proteínas Rho, en la cual se requiere la unión del hexámero Rho a sitios en la región líder del transcrito y la interacción de éste con la RNAPol pausada para causar la terminación transcripcional [**Error!No se encuentra el origen de la referencia.**].

Estabilidad y modificaciones del mRNA

Los transcritos primarios de los rRNA y tRNA se modifican por la eliminación de secuencias y muchas bases y azúcares de los tRNA son modificados.

Eficiencia de unión del ribosoma al mRNA

En el caso de la regulación traduccional, el primer nivel de regulación se da por la eficiencia de unión del ribosoma a la secuencia “Shine-Dalgarno” (SD)⁸ del mRNA lo cual influye directamente los niveles de proteínas individuales en la célula.

Interferencia con el sitio “Shine-Dalgarno”

Por otro lado la regulación por interferencia con la secuencia SD se da por proteínas reguladoras, RNA antisentido o por la estructura secundaria del transcrito que pueden bloquear la secuencia SD e impiden la unión del ribosoma al mRNA [**Error!No se encuentra el origen de la referencia.**].

⁵ Líder del transcrito: región entre el inicio de la transcripción y el primer gene estructural del DNA.

⁶ Terminador transcripcional: es una estructura de tipo tallo-asa (horquilla) seguida por una serie de residuos de uracilos.

⁷ RNA polimerasa: Enzima que cataliza la formación del RNA a partir de ribonucleótidos 5'-trifosfatos utilizando una hebra de DNA o RNA como hebra.

⁸ Secuencia “Shine-Dalgarno” (SD): Secuencia de un mRNA que reconoce el ribosoma necesaria para que esté se asocie al mRNA e inicie la traducción.

1.3. Mecanismo de Regulación de los riboswitches

En las bacterias donde los procesos de transcripción y traducción están acoplados, la regulación por riboswitches se puede dar en ambos niveles. El mecanismo general de acción consiste en la unión del ligando a la secuencia aptámera del riboswitch, lo cual provoca un cambio conformacional en su estructura y la formación subsecuente de un terminador o un secuestrador [;Error!No se encuentra el origen de la referencia.].

Regulación por un terminador transcripcional. En ausencia del ligando parte de la secuencia que forma al terminador se asocia a una secuencia complementaria para dar lugar a un antiterminador, permitiendo la transcripción. En presencia del ligando se favorece la unión del antiterminador con una secuencia complementaria formando un anti-antiterminador, lo que permite la formación del terminador transcripcional y la disociación del complejo de transcripción. Este mecanismo generalmente está presente en bacterias gram-positivas (Fig. 2A).

Regulación por un inhibidor traduccional. En ausencia del ligando la secuencia del secuestrador de la “Shine Dalgarno” (SD) o anti-SD se encuentra asociada a una secuencia complementaria (anti-anti-SD), dejando libre la secuencia SD para que el ribosoma se una. En presencia del ligando, la secuencia SD se mantiene unida a la secuencia del secuestrador (anti-SD) impidiendo que el ribosoma se asocie al RNA y bloqueando la traducción. Este mecanismo generalmente está presente en bacterias gram-negativas (Fig. 2B).

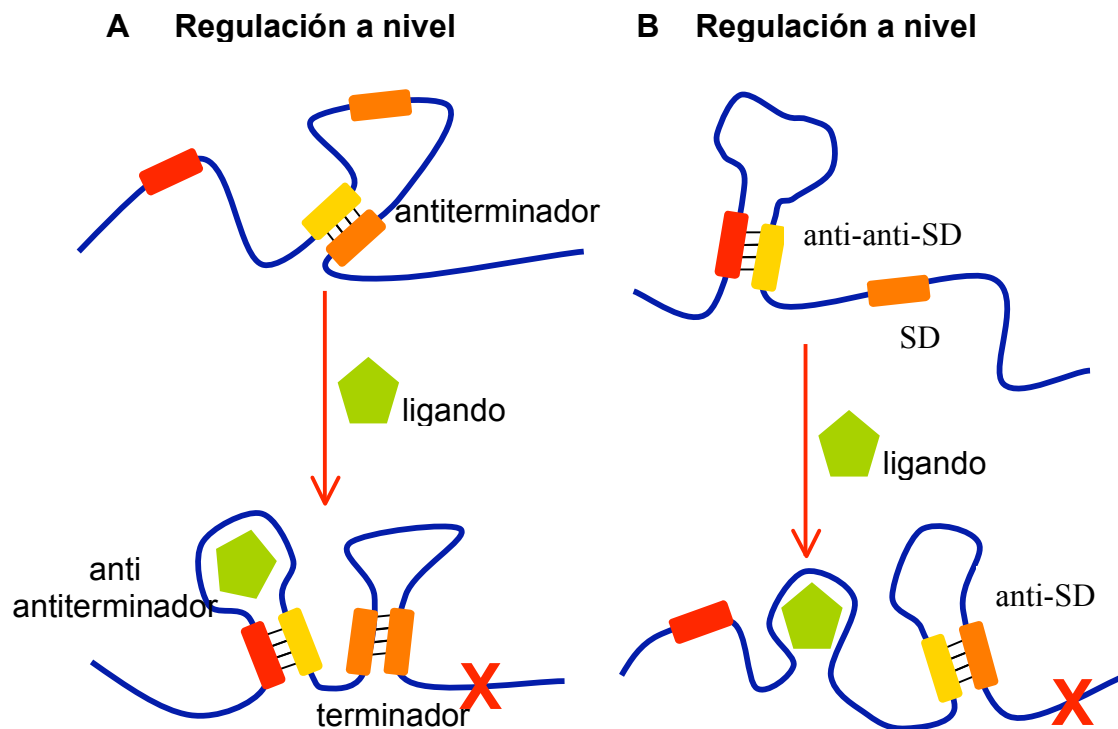


Fig. 2. Mecanismos de regulación en los riboswitches. A. La unión del ligando favorece la formación de un terminador. B. La unión del ligando secuestra a la secuencia de inicio de la traducción “Shine Dalgarno”.

Adicionalmente, se sabe que los riboswitches de glucosamina, acetina y glutina, tienen variaciones en el mecanismo de expresión que controlan.

En el caso del riboswitch de glucosamina, la unión de glucosamina-6-fosfato (GlcN₆P), provoca un cambio conformacional en el RNA que lo convierte en una ribozima capaz de catalizar el corte del transcrito al que está asociado el RNA [;Error!No se encuentra el origen de la referencia.]. En el de adenina se activa la transcripción de los genes a los que está asociado, mediante un rearrreglo en su estructura que evita la formación de un atenuador [;Error!No se encuentra el origen de la referencia.]. En el de glicina se han observado dos dominios de unión al ligando en una misma secuencia que funcionan de manera cooperativa, activando genes que están encargados de la degradación de glicina [;Error!No se encuentra el origen de la referencia.] (Apéndice 1).

También se han encontrado secuencias tipo riboswitch en eucariotes, como en el gene *thiC* de las plantas *Arabidopsis thaliana*, *Oriza sativa** (arroz) y *Poa secunda** (pasto) (*posible *thiC*), así mismo en genes de hongos implicados en biosíntesis de tiamina, como el gene *thiA* de *Aspergillus oryzae*, *nmt-1* de *Neurospora crassa* y *sti35* de *Fussrium oxysporum*. En plantas, estos se han localizado río arriba de las colas de poliA y se propone que interviene en la estabilidad del mRNA. En hongos están en intrones y se ha observado que interviene en el proceso de “splicing” [;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia.].

1.4. Biosíntesis de tiamina

La Tiamina Pirofosfato (TPP)⁹ es la forma activa de la vitamina B₁, el cual es un cofactor esencial para varias enzimas importantes del metabolismo de carbohidratos. No todos los organismos son capaces de sintetizar tiamina, las bacterias, varios protozoarios e incluso plantas y hongos sí la sintetizan.

La TPP es sintetizada a partir de dos principales precursores, la hidroximetilpirimidina (HMP) y el hidroxietiltiazol (HET). Como se observa en la ruta metabólica de la Fig. 2, *thiC* produce HMP, que es después fosforilada por una cinasa codificada por el gene *thiD*. El HET se forma a partir de tirosina o glicina (tirosina en *Escherichia coli* y glicina en *Bacillus subtilis*), de cisteína y de 1-deoxi-d-xilulosa fosfato, a través de una serie de reacciones que involucran los genes *thiF*, *thiS*, *thiG*, *thiI*, *thiH* en *Escherichia coli* o *thiO* en *Bacillus subtilis* y *Rhizobium etli*. Así mismo *thiM* codifica para HET, una cinasa de la vía de síntesis de tiazol. A partir de HMP-PP y HET-P por la acción del producto codificado por el gene *thiE* se forma el monofosfato de tiamina que es el precursor del pirofosfato de tiamina. La segunda fosforilación es llevada a cabo por la cinasa *ThiL*. La vía de síntesis de tiamina se puede revisar con más detalle en la base de datos de KEGG (Kyoto Encyclopedia of Genes and Genomes) [<http://www.genome.jp/kegg/pathway/map/map00730.html>] [¡Error!No se encuentra el origen de la referencia.].

En *Escherichia coli*, los genes que codifican para la biosíntesis de la tiamina, se encuentran organizados en dos unidades transcripcionales *thiMD*, *thiCEFSGH* y tiene también una unidad transcripcional *thiBPQ* que codifica para un transportador de tiamina. Todos estos genes están regulados por el riboswitch de tiamina (Fig. 3).

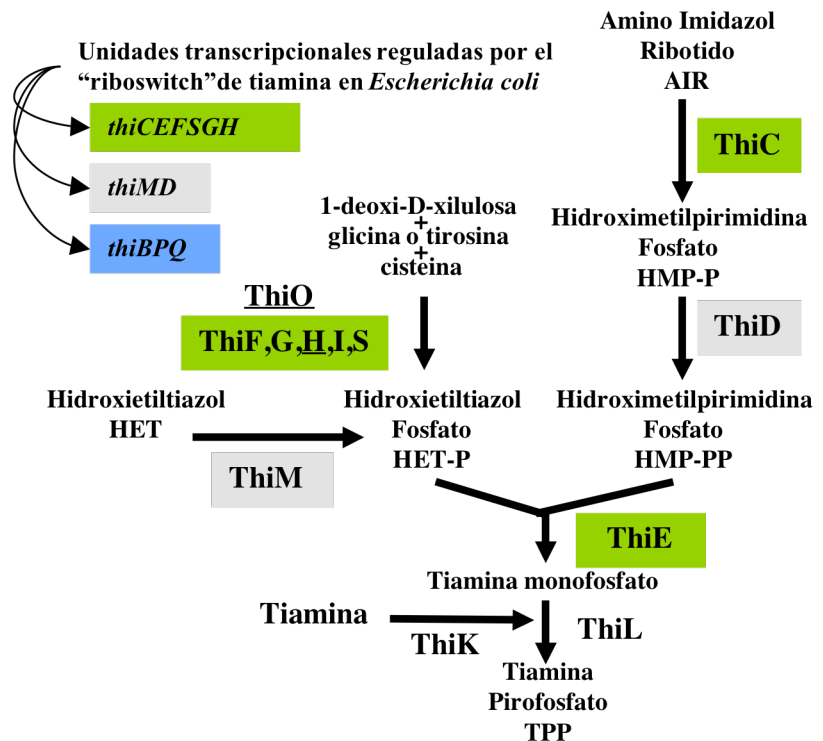


Fig. 3. Biosíntesis de tiamina en procariotes. En la vía se encuadran los genes regulados por el riboswitch de tiamina. *ThiBPQ* es un transportador tipo ABC de tiamina sus compuestos fosforilados (TP y TPP).

⁹ La tiamina es esencial como coenzima en el metabolismo de los α -ceto ácidos y α -ceto azúcares. Es cofactor de enzimas como el piruvato deshidrogenasa y α -cetogluturato deshidrogenasa que son vitales para el metabolismo de los carbohidratos, y la transcetolasa, enzima necesaria para el ciclo de las pentosas y síntesis de acetilcolina. [¡Error!No se encuentra el origen de la referencia.]

1. Antecedentes

1.1. Estudios de regulación

El riboswitch de tiamina ha sido ampliamente estudiado, ya que es uno de los primeros casos reportados. Así mismo la síntesis de tiamina es una vía importante, dada la importancia de la tiamina como cofactor. Para el 2001 se sabía que la tiamina tenía un efecto de autorepresión en genes de su biosíntesis, sin embargo poco se sabía del mecanismo de regulación [¡Error!No se encuentra el origen de la referencia.]. En ese año, Miranda-Ríos *et al.* reportaron la represión por tiamina de un grupo de genes organizados en un operón (operón *thiCOGE*) que codifican para la síntesis de tiamina en *Rhizobium etli* [¡Error!No se encuentra el origen de la referencia.]. Para estudiar la regulación en este operón se realizaron una serie de fusiones transcripcionales de la secuencia líder del gene *thiC* y de distintas deleciones de la misma al gene *lacZ* y se cuantificó la represión por tiamina. Se encontró que esta represión es un evento postranscripcional y que a su vez el elemento “thi-box” que corresponde a una secuencia identificada entre el nucleótido +24 y +147 de la región líder, es indispensable para la expresión de los genes ya que en la deleción de ésta, se pierde la expresión. La “thi-box” es una de las partes importantes en lo que hoy se conoce como la secuencia del riboswitch de tiamina.

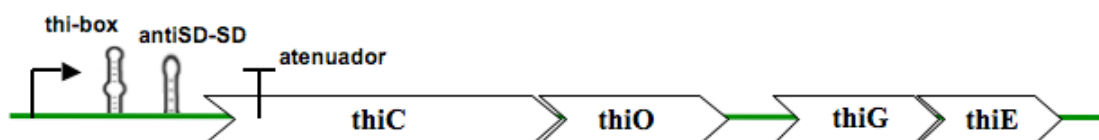


Fig. 4. Esquema de el operón *thiCOGE* de *Rhizobium etli*. En el se muestran las estructuras thi-box y el secuestrador de la secuencia SD en la región líder del operón y el atenuador identificado en la región codificante del gene *thiC*.

El operón *thiCOGE* de *Rhizobium etli* está regulado a nivel traduccional por la formación de un secuestrador de la secuencia SD y además presenta una secuencia conservada capaz de formar un terminador transcripcional, este atenuador se localiza en la región codificante del gene *thiC* (Fig. 4).

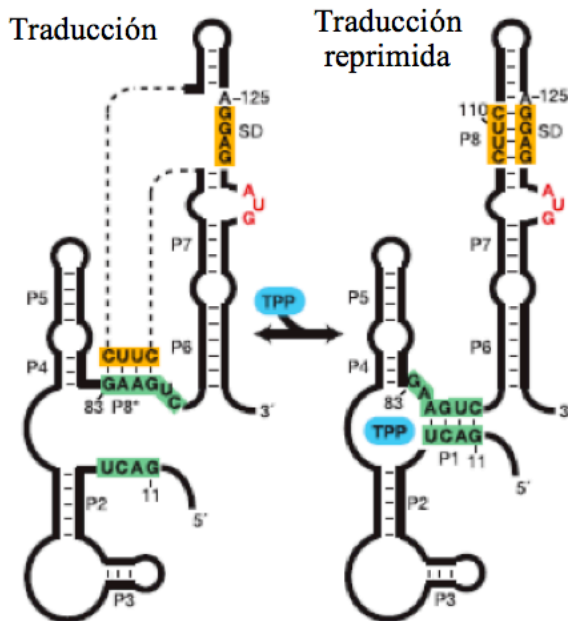
Por otro lado se hizo una comparación de secuencias de homólogos con la base de datos GenBank¹ y la región líder del gene *thiC* de *Rhizobium etli* y se encontró a la “thi-box” en la región 5’ no traducida de genes de biosíntesis de tiamina. En *Escherichia coli* la secuencia se encontró en las tres unidades transcripcionales de la biosíntesis de tiamina: *thiCOGE*, *thiMD* y *thiBPQ*, para las cuales ya se conocía que su expresión es regulada negativamente por TPP [¡Error!No se encuentra el origen de la referencia.]. A partir de estos estudios se identificó a la “thi-box” como el elemento de regulación involucrado en la represión de genes de biosíntesis de tiamina.

Haciendo un análisis filogenético se observó que la “thi-box” está presente y altamente conservada en bacterias gram-positivas y gram-negativas [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.]. En el 2002 Rodionov *et al.* realizaron un análisis comparativo más

¹ GenBank: es la base de datos del NIH, que contiene la anotación de todas las secuencias de DNA públicamente disponibles.

amplio, con 103 secuencias genómicas y reportaron 170 secuencias reguladoras en 78 genomas. En el 2003 Sudarsan *et al.* reportaron que el riboswitch de tiamina también estaba presente en eucariotes, en el gene *thiC* de *Arabidopsis thaliana* y en otras plantas como *Oriza sativa* y *Poa secunda*. También se encontró en hongos como *Neurospora crassa* y *Fussarium oxysporum* [¡Error!No se encuentra el origen de la referencia.]. Hasta este momento este riboswitch es el más ampliamente distribuido filogenéticamente [¡Error!No se encuentra el origen de la referencia.].

En el 2002, Winkler *et al.* estudiaron el pegado de la tiamina y su derivado tiamina pirofosfato (TPP) al riboswitch mediante ensayos de rompimiento espontáneo del RNA (“in-line probing”) ² [¡Error!No se encuentra el origen de la referencia.]. Los autores utilizaron las secuencias de los riboswitches del gene *thiC* y del gene *thiM* de *Escherichia coli* y observaron que el control genético era diferente, en *thiC* se da a nivel de la transcripción y en *thiM* a nivel de la traducción, observaron que la unión del ligando al riboswitch de *thiM* provoca un cambio conformacional que secuestra al sitio SD y como consecuencia hay una reducción en la expresión génica [¡Error!No se encuentra el origen de la referencia.].



espontáneo tanto en presencia como en ausencia de TPP, lo cual nos indica que siempre

Winkler *et al.* (2002)

Fig. 5. Representación del riboswitch de tiamina y del mecanismo de regulación para el gene *thiM* de *Escherichia coli*. En ausencia de TPP la secuencia SD está libre por lo que los genes se pueden traducir, sin embargo si la TPP se asocia al riboswitch, la secuencia SD queda secuestrada y se evita la síntesis de los genes.

En este trabajo se realizaron pruebas de “in-line” que miden el rompimiento espontáneo en una molécula de RNA, con lo cual se puede ver la modulación y los cambios conformacionales de la secuencia, cuando el ligando está asociado al riboswitch y cuando está ausente. [¡Error!No se encuentra el origen de la referencia.] En estas pruebas se observó que hay distintos patrones de rompimiento en las secuencias en presencia y ausencia de tiamina. En el gene *thiM* se observó que la secuencia complementaria a la secuencia SD, resiste el rompimiento espontáneo tanto en presencia como en ausencia de TPP, lo cual nos indica que siempre esta comprometida: ya sea asociándose al un segmento del riboswitch (P8) en ausencia de TPP o a la secuencia SD en presencia de TPP (Fig. 5).

² La prueba de “in-line” es una estrategia molecular que tiene como principio el rompimiento espontáneo de una secuencia de RNA por ataque nucleofílico de sus propias bases.

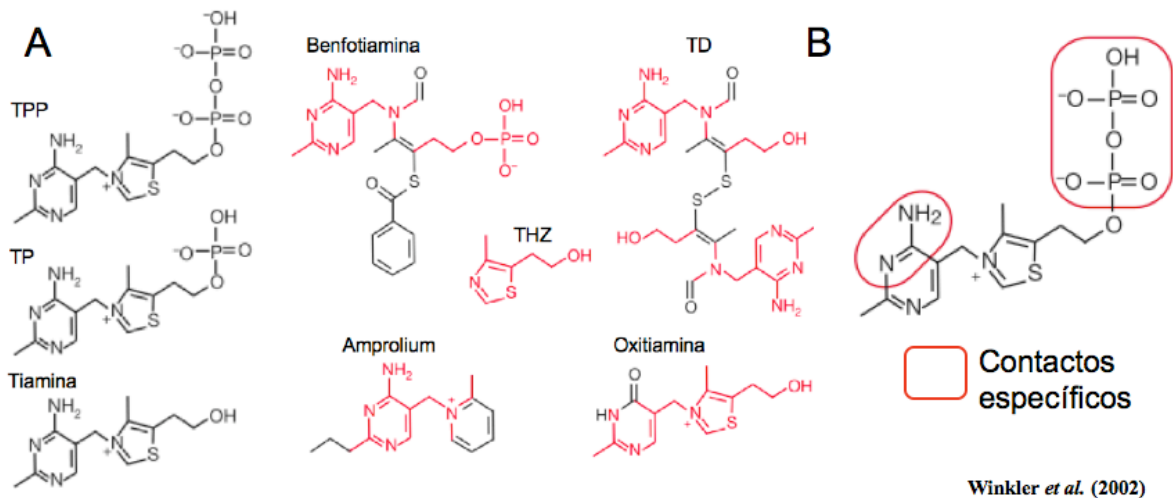


Fig. 6. Molécula de la tiamina pirofosfato. A. La molécula de TPP está formada por una pirimidina, un tiazol y dos fosfatos, en el esquema se muestran los análogos de TPP que se utilizaron en los experimentos de Winkler et al (2002). B. Grupos funcionales de TPP que se propone son importantes en el reconocimiento molecular.

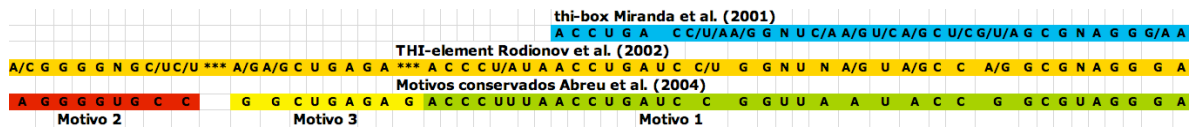
Para medir la afinidad del riboswitch por su ligando, se calculó la constante de disociación aparente³ para la TPP mediante diálisis al equilibrio⁴ y se obtuvo que para el gene *thiM* es de 600nM y para *thiC* es de 100nM en *Escherichia coli* según Winkler *et al.* [¡Error!No se encuentra el origen de la referencia.]. Para identificar los residuos importantes de TPP involucrados en su reconocimiento por el riboswitch, se hicieron ensayos de “in-line” utilizando análogos de TPP. A partir de éstos se observó que el riboswitch es capaz de discriminar en más de 1000 veces entre TPP, tiamina monofosfato (TP) y tiamina. El mayor reconocimiento se da por TPP y en orden decreciente, por: TP, tiamina, Benfotiamina, Amprolium, Oxitiamina, 4-metil-5-β-hidroxiethyltiazole (THZ) y tiamina dizulfido (TD). En el caso del gene *thiM*, se observó que TP y TPP sí inducen modulación estructural a diferencia de los otros análogos. A partir del análisis con análogos de TPP se propuso que tanto el grupo pirimidina, como los grupos fosfato, son elementos importantes en el reconocimiento molecular de TPP (Fig. 6).

1.2. Estudios de genómica comparativa

³ La constante de disociación (Kd), refleja la cantidad de molécula efectora (metabolito), necesaria para convertir la mitad de RNA en su estructura alternativa. Mientras menor sea el valor, mayor es la capacidad que tiene la secuencia de sensar a la molécula efectora.

⁴ La diálisis al equilibrio es una estrategia a través de la cual se pueden medir concentraciones de ligando asociado y no asociado a una molécula. El método se basa en que el ligando libre puede dializarse a través de una membrana que separa dos cámaras, hasta que la concentración de ligando libre está en equilibrio. Una de las cámaras contiene la molécula que reconoce al ligando y por tanto se puede medir una concentración de ligando libre y la de ligando secuestrado. Esta técnica usualmente necesita de un marcaje radiactivo para seguir los cambios de concentración del ligando.

A nivel secuencia se han realizado varios estudios de comparación genómica para definir la secuencia consenso que reconoce a la TPP y se ha visto que hay segmentos muy conservados y segmentos variables en las secuencias de los riboswitches en los distintos genes. En el 2002 Rodionov *et al.* hicieron una búsqueda más extensa de secuencias conservadas “thi-box” y reportaron una secuencia más ampliada a la cual denominaron “THI-element” (Fig. 7) [;Error!No se encuentra el



origen de la referencia.].

Para el 2004 en un estudio reportado por Abreu-Goodger *et al.*, se presentó un método para comparar secuencias intergénicas agrupadas por función, utilizando la base de datos COG's (Cluster of Orthologous Group of proteins) [;Error!No se encuentra el

Fig. 7. Secuencias conservadas que se han identificado del riboswitch de tiamina.

origen de la referencia.]. Este método busca motivos conservados, a partir de los cuales podemos identificar elementos de regulación como los riboswitches, entre otros. Los riboswitches tienen regiones muy conservadas y regiones variables y éste método permite identificar de forma independiente sus regiones más conservadas y evaluar cada residuo en función de qué tan probable es encontrar esa base en la secuencia, a su vez la combinación de los motivos conservados nos define secuencias de mayor tamaño (Apéndice 2). Los datos que se obtuvieron de la búsqueda de motivos conservados están disponibles en: (http://www.ibt.unam.mx/biocomputo/conserved_motifs.html, grupo 0012), y se puede consultar información como: en cuántas secuencias se encontró, de cuántos genomas y en cuántas phyla, a su vez se puede acceder mediante una liga a la lista total de secuencias que integran al grupo, también se mencionan elementos biológicos que coinciden con los motivos encontrados y se da tanto los motivos cómo el orden principal en el que estos se encuentran en las secuencias (Fig. 8).

Group0012	(Best:1.3e-10)	86_[2]_1_[3]_7_[1]_69_[4]_135 (2.5e-26)
KEGG 1e-188	MOTIF1:	[1]=CTGAGGCCAGACCCCTTGAACCTGATCCGGTTAATACCGGCTAGGGAA
218 106 12	THI-box_Thiamine	[2]=CACTAGGGGTGCC
Gene info	MOTIF4:	[3]=TGGGCTGAGAG
	FNR_nuoAp1	[4]=CTCITTGTGGATGTGGTTTTTTTTTT

Fig. 8. Motivos conservados del riboswitch de tiamina. El grupo 12 corresponde a metabolismo de tiamina v presenta 4 motivos.

A partir de este estudio se identificó una combinación de motivos muy fuerte presente en genes de biosíntesis de tiamina. La secuencia se encontró en 218 regiones reguladoras de 106 genomas, distribuidas en 12 órdenes filogenéticos [;Error!No se encuentra el origen de la referencia.].

Del análisis de los cuatro motivos que definen al riboswitch:

[1] el primer motivo contiene la “thi-box”

[2] y [3] el motivo 2 y 3 sobrelapan con parte de la estructura del “THI-element” (elemento de regulación ampliado de “thi-box”) [;Error!No se encuentra el origen de la referencia.].

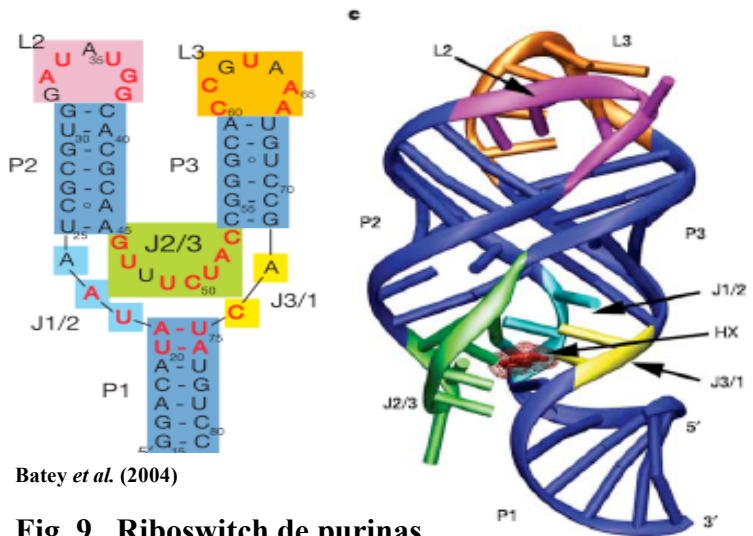
[4] el cuarto motivo es aparentemente un terminador transcripcional que no siempre está presente porque muchos genes son regulados a nivel traduccional.

1.3. Estudios estructurales

Después de identificar que los metabolitos se asocian de forma específica al RNA, se buscó identificar las interacciones que se llevan a cabo entre estas moléculas. Estas interacciones son importantes porque representan un mecanismo de reconocimiento RNA-ligando muy específico.

Riboswitch de purinas

El primer riboswitch natural que se cristalizó con su ligando fue el **Riboswitch de purinas**. Éste fue reportado por Batey *et al.* en el 2004 que cristalizaron el dominio de unión del riboswitch de guanina del operón *xpt-pbuX* de *Bacillus subtilis* acoplado a una molécula de hipoxantina [;Error!No se encuentra el origen de la referencia.].⁵ En ese mismo año, Serganov *et al.* reportaron la estructura del dominio de unión del mismo riboswitch pero esta vez unido a una molécula de guanina y la del dominio de unión del riboswitch de adenina del gene *add* de *Vibrio Vulnificus* unido a una molécula de adenina [;Error!No se encuentra el origen de la referencia.].⁵ De acuerdo a la estructura reportada se muestra que el mRNA envuelve al ligando, haciendo contactos con todos los grupos funcionales del mismo, lo cual explica la especificidad por la hipoxantina observada en estudios bioquímicos (Fig. 9).



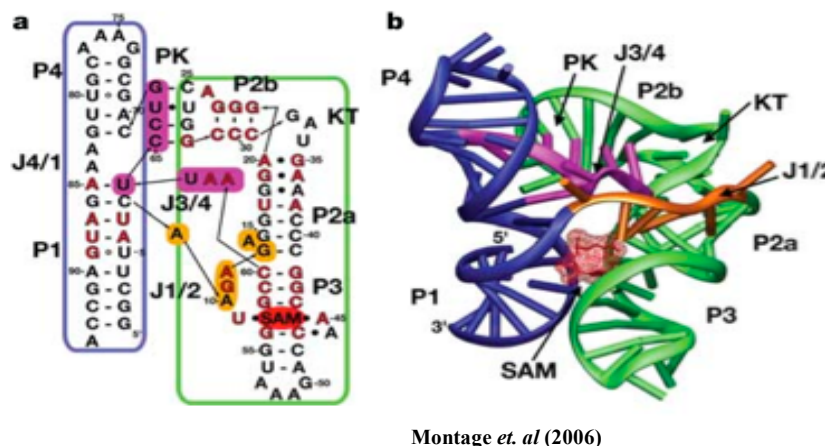
Batey *et al.* (2004)

Fig. 9. Riboswitch de purinas.

El mismo año, Serganov *et al.* reportaron la estructura del dominio de unión del mismo riboswitch pero esta vez unido a una molécula de guanina y la del dominio de unión del riboswitch de adenina del gene *add* de *Vibrio Vulnificus* unido a una molécula de adenina [;Error!No se encuentra el origen de la referencia.].⁵ De acuerdo a la estructura reportada se muestra que el mRNA envuelve al ligando, haciendo contactos con todos los grupos funcionales del mismo, lo cual explica la especificidad por la hipoxantina observada en estudios bioquímicos (Fig. 9).

Riboswitch de Metionina

La estructura del riboswitch de metionina fue reportada por Montagne y Batey en junio del 2006 ellos reportan la estructura del riboswitch del gene *metF* de *Thermoanaerobacter tengcongensis* asociado a su ligando S-adenosilmetionina [;Error!No se encuentra el origen de la referencia.].⁵ El riboswitch también está formado por dos hélices paralelas y el ligando



Montagne *et al.* (2006)

Fig. 10. Riboswitch de metionina.

⁵ PDBs: riboswitch de *xpt-pbuX* acoplado a hipoxantina, 1U8D; riboswitch de *xpt-pbuX* acoplado a guanina, 1Y26; riboswitch de *ydhL* acoplado a adenina, 1Y27; riboswitch de *metF* acoplado a metionina, 2GIS.

interacciona con bases en cada una de las hélices (Fig. 10). En este caso el ligando adopta una conformación compacta en la cual el motivo de metionina queda sobre el anillo de adenina. La conformación compacta del ligando crea dos caras que interaccionan de forma diferente con el riboswitch. Una de éstas caras interacciona por puentes de hidrógeno con una de las hélices y el otro interacciona mediante fuerzas de “Van der Waals” con la otra.

Riboswitch de Tiamina

Para el 2005 ya era muy claro el mecanismo de acción del riboswitch de tiamina, sin embargo poco se sabía del reconocimiento estructural del RNA con su ligando y de los cambios en la estructura del mismo. Dado que la molécula de TPP contiene dos fosfatos y el esqueleto fosfodiéster del RNA confiere una carga alta de iones, el riboswitch debe contrarrestar los efectos repulsivos entre las cargas del pirofosfato de tiamina y las cargas del RNA. La solución a esto es la interacción específica con iones y en ese año un estudio realizado por Yamauchi *et al.* analizaron la estructura secundaria y los cambios estructurales que se dan en el riboswitch del gene *thiA* de *Aspergillus oryzae* en presencia y en ausencia de tiamina mediante dicroísmo circular⁶ y observaron que el ion Mg^{2+} es importante para el plegamiento del riboswitch [¡Error!No se encuentra el origen de la referencia.]. En este estudio se observó que la constante de asociación del riboswitch con la tiamina a 20°C, incrementa de 1.2×10^6 a $50 \times 10^6 M^{-1}$ cuando la concentración de Mg^{2+} se incrementa de 0 a 1 mM.

El riboswitch de tiamina es el segundo riboswitch natural cristalizado; la estructura fue reportada por tres grupos de forma independiente: El grupo de Thore *et al.* [¡Error!No se encuentra el origen de la referencia.] y el de Serganov *et al.* [¡Error!No se encuentra el origen de la referencia.] en mayo del 2006 y Edwards y Ferré-D’Amare [¡Error!No se encuentra el origen de la referencia.] en septiembre del mismo año. En el caso del grupo de Thore *et al.* se cristalizó el riboswitch eucariote del gene *thiC* de *Arabidopsis thaliana* y en el caso del grupo de Serganov *et al.* y de Edwards and Ferré-D’Amare se cristalizó el del gene *thiM* de *Escherichia coli* con su ligando TPP.

A diferencia de los trabajos reportados por Thore *et al.* y Serganov *et al.* en el trabajo de Edwards y Ferré-D’Amare, se cristalizó al riboswitch asociado a distintos cationes divalentes y asociado a distintos análogos de TPP, como se muestra en la Tabla 2, de este trabajo ellos concluyen que los análogos tienen una disminución en el reconocimiento óptimo con respecto a TPP, donde los análogos monofosforilados (TMP y BTP) tienen una unión subóptima pero estructuralmente similar a la de TPP y la piritiamina no fosforilada (PT) tiene una unión marginal. Así mismo establecen que a la TPP se asocian dos cationes divalentes y en el caso de ligandos monofosforilados (TMP, BTP) también se asocian dos cationes divalentes a la molécula.

⁶ El dicroísmo circular es una estrategia que se utiliza para explorar la conformación y cambios estructurales en moléculas, principalmente en proteínas. Se basa en utilizar rayos de luz circularmente polarizada. Los rayos de luz polarizada inicialmente están en fase y son de una misma amplitud, al pasar por un medio ópticamente activo, interaccionan de manera diferente con los centros quirales de las moléculas presentes, lo cual induce un desfase y un cambio en la magnitud de la luz polarizada. Entonces la rotación en el plano y la diferencia en absorción de la luz polarizada varía de acuerdo a una longitud de onda, lo cual permite ver espectros (gráficas de la rotación o elipicidad contra longitud de onda). Los espectros obtenidos son un reflejo global de la conformación de una molécula.

Tabla 2

Estructuras reportadas para el riboswitch de tiamina				
PDB	Fuente	Ligando	Catión divalente	Resolución
2CKY	gene <i>thiC</i> de <i>Arabidopsis thaliana</i>	TPP	Mg ²⁺	2.9 Å
2GDI	gene <i>thiM</i> de <i>Escherichia coli</i>	TPP	Mg ²⁺	2.05 Å
2HOJ	gene <i>thiM</i> de <i>Escherichia coli</i>	TPP	Mn ²⁺	2.5 Å
2HOK	gene <i>thiM</i> de <i>Escherichia coli</i>	TPP	Ca ²⁺	3.2 Å
2HOL	gene <i>thiM</i> de <i>Escherichia coli</i>	TPP	Ba ²⁺	2.9 Å
2HOM	gene <i>thiM</i> de <i>Escherichia coli</i>	TMP	Ca ²⁺	2.85 Å
2HOO	gene <i>thiM</i> de <i>Escherichia coli</i>	BTP	Ca ²⁺	3.0 Å
2HOP	gene <i>thiM</i> de <i>Escherichia coli</i>	PT	Ca ²⁺	2.5 Å

En la estructura del riboswitch de tiamina se observan dos dominios de interacción, uno que reconoce al motivo 4-amino-5-hidroximetil-2-metil-pirimidina (HMP) y otro que reconoce al motivo pirofosfato (PP) de TPP (Fig. 11). El dominio que interacciona con el PP está compuesto por los tallos: P4 y P5; y por las asas: J4/5 y L5, que corresponden a la “thi-box” y el dominio que interacciona con el HMP, está formado por los tallos: P2 y P3; y por las asas: J2/3 y L3. En la estructura se identifica por lo menos un Mg²⁺.

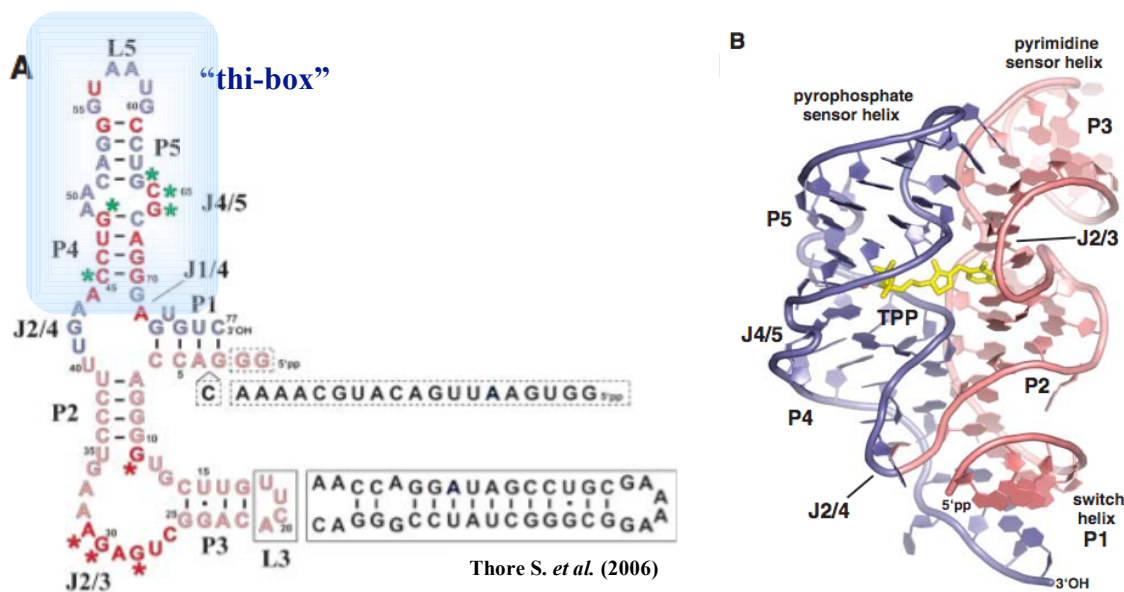


Fig. 11. Riboswitch de tiamina. Riboswitch de *thiC* de *Arabidopsis thaliana* reportada por Thore S. et al. (2006). A. Estructura secundaria. B. Estructura del riboswitch. La estructura está formado por dos dominios: una hélice sensora de la pirimidina (rosa) y una hélice sensora del pirofosfato de TPP (azul), que corresponde a

1. Propuesta

La especificidad, alta afinidad por el ligando y la conservación en la secuencia que presenta el riboswitch de tiamina, son características muy interesantes para estudiar la interacción RNA-ligando. En este trabajo se propone estudiar el reconocimiento entre el riboswitch de tiamina y TPP, mediante mutaciones puntuales en las bases conservadas en las asas de este riboswitch. Estudiar las mutantes es una forma de observar *in vivo* el efecto que tiene la modificación de estas bases en el reconocimiento de TPP y en la regulación.

1. Hipótesis

Dado que existen bases muy conservadas en las asas del riboswitch de tiamina y dado el reconocimiento específico de éste por TPP, proponemos que estas bases son importantes en el pegado de TPP y que cambios en ellas pueden modificar la interacción entre TPP y el riboswitch de tiamina, afectando la regulación.

1. Objetivos

1.1. General

Estudiar el pegado de TPP al riboswitch de tiamina mediante mutaciones puntuales.

1.2. Particulares

- En base a la secuencia conservada del riboswitch de tiamina, definir las bases que pueden estar involucradas en el pegado de TPP y la secuencia mínima de regulación.
- Realizar los cambios específicos en las bases seleccionadas mediante mutagénesis dirigida.
- Analizar el efecto de los cambios en la regulación por medio de fusiones traduccionales al gene reportero *lacZ*.
- Cuantificar el cambio de expresión en función de la actividad de β -galactosidasa en las fusiones traduccionales.
- Y en base a los resultados de expresión, identificar las mutaciones que afectan la regulación por tiamina.

1. Materiales

1.1. Microorganismos

Escherichia coli K-12 (MG1655)

MG1655 es una cepa derivada de la cepa W1485, la cual se obtuvo de un cultivo “stock” del laboratorio de Joshua Lederberg, de la cepa K-12 original. La cepa original *Escherichia coli* K-12 se obtuvo de una muestra de cultivo de un paciente con difteria en Palo Alto, CA en 1922. El genoma de la cepa MG1655 fue secuenciado por el laboratorio de Blattner por su aproximación a la *Escherichia coli* silvestre la cual se mantiene como cepa de laboratorio con mínima manipulación génica [**¡Error!No se encuentra el origen de la referencia.**]. Esta cepa tiene la ventaja de no tener modificaciones que alteren la regulación o el metabolismo básico de *Escherichia coli* por lo que se pueden hacer modificaciones independientes a otros cambios.

Genotipo

F- lambda- *ilvG- rfb-50 rph*

Escherichia coli (DH5 α)

La cepa DH5 α permite hacer ensayos fenotípicos: azul/blanco por α complementación con β -galactosidasa, con vectores pUC. Las células complementadas, sintetizan ambos fragmentos de la β -galactosidasa y forman colonias azules en medio con X-Gal. Cuando el DNA es insertado en el sitio múltiple de clonación del vector (MCS), localizado entre el codón 6-7 del gene *lacZ* se inactiva el fragmento N-terminal de la β -galactosidasa y bloquea la α -complementación. Por lo tanto las bacterias que contienen plásmidos recombinantes forman colonias blancas. No contiene *lac^q*, por lo que no es necesario inducir con IPTG. Codifica para X-Gal, lo cual permite una fácil selección del DNA recombinante en librerías génicas o subclonación de plásmidos recombinantes. La cepa tiene las mutaciones *recA1* y *endA1* que incrementan la estabilidad del inserto y mejoran la calidad del DNA plasmídico. Su eficiencia de transformación es de $1.0 \times 10^7 \sim 1.0 \times 10^8$ cfu/ μ g con pUC18 [**¡Error!No se encuentra el origen de la referencia.**].

Genotipo

F⁻, ϕ 80dlacZ Δ M15, Δ (*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsdR17*(rk⁻, mk⁺), *phoA*, *supE44*, λ , *thi-1*, *gyrA96*, *relA1*

1.2. Plásmidos

Vector pUC19

El plásmido pUC19 es un vector de alto número de copias, con un tamaño de 2686pb que contiene un sitio múltiple de corte (“polylinker”) donde se encuentran los sitios de restricción *EcoR*I y *Bam*H1, entre otros un gene de resistencia a ampicilina (*amp^r*) y un origen de replicación (pMB1ori) (Fig. 12). Contiene el pMB1 replicón *rep*, que es el responsable de la replicación del plásmido (fuente - plásmido pBR322). El alto número de copias de los plásmidos pUC es el resultado del gene *rop* y de una mutación puntual en *rep* del pMB1, también contiene el gene *bla*, que codifica para β -lactamasa, lo cual le confiere resistencia a ampicilina (*amp^r*) (fuente – plásmido pBR322). Difiere del pBR322 en dos mutaciones puntuales: (1) La región del operón *lac* de *Escherichia coli* contiene el sitio de unión de la proteína CAP, promotor P_{lac} , sitio de unión al represor *lac* y (2) la parte 5'-terminal del gene *lacZ* que codifica para la β -galactosidasa (fuente –M13mp). A través de este fragmento (inducido por IPTG), se lleva a cabo la α complementación (intra- alélica) con DNA que contienen una forma defectiva de la β -galactosidasa (ejem. cepa DH5 α) [;Error!No se encuentra el origen de la referencia.].

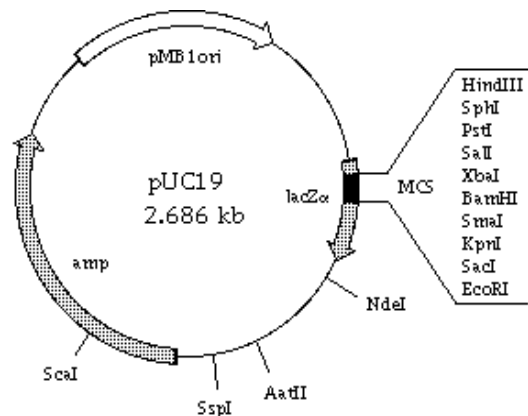


Fig. 12. Plásmido pUC19

Vector pRS552

El plásmido pRS552 es un vector multicopia para fusiones traduccionales a *lacZ*. Tiene asociado el gene truncado *lacZ* (- 8 codones, que eliminan a la secuencia promotora) (fuente – pMC1404). Mide 12.34kb y está formado por un sitio múltiple de corte que contiene a *EcoR*I y *Bam*H1, el gene de resistencia a ampicilina (*amp^r*) y un origen de replicación (pMB1ori). Tiene los genes silvestres *lacY*, *lacA* y el terminador transcripcional del operón *lac* (Fig. 13). Este es un vector derivado del plásmido pBR322 que contiene los genes *ori*, *bla* y porciones distales de *tet* (*tet'*) de éste. Entre los genes *bla* y *tet'* el plásmido tiene varios elementos: cuatro copias contiguas, de un terminador transcripcional fuerte, *T1* (fuente - operón *rnB* de *Escherichia coli*), el cual bloquea los promotores río arriba del plásmido. Tiene el gene *kan* de selección a kanamicina (fuente –Tn903), entre el gene *bla* y *T1* [;Error!No se encuentra el origen de la referencia.].

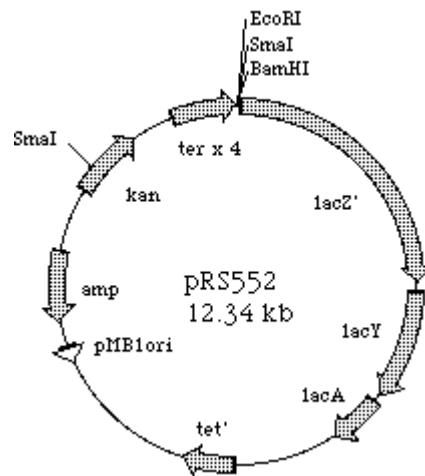


Fig. 13. Plásmido pRS552

1. Métodos

1.1. Selección de la secuencia mínima de regulación (*rrthiM*) y de las bases conservadas en las asas del riboswitch

Para seleccionar la región reguladora con la que se trabajó, se analizaron las regiones reguladoras de las unidades transcripcionales de genes de biosíntesis de tiamina en *Escherichia coli* (*thiMD*, *thiCEFSGH* y *thiBPQ*) y se eligió la secuencia más corta que incluye los motivos del riboswitch (*thiDM* = 75 bases). Se decidió utilizar esta secuencia tanto por su tamaño como por los trabajos previos donde se ha utilizado al riboswitch del operón de *thiM* (Winkler, W. *et al.* (2002)) [26]. El operón de *thiM* codifica para genes relacionados a la biosíntesis de tiamina: *thiM* para una hidroxietiltiazol cinasa y *thiD* para una hidroximetilpirimidin cinasa.

Para definir la secuencia mínima de regulación del gene *thiM* (*rrthiM*), se buscó su región promotora y la secuencia del riboswitch de tiamina. La secuencia promotora del gene se designó en base al análisis de predicción de operones de regulonDB (http://regulondb.ccg.unam.mx:80/html/Computational_Predictions.jsp) [**Error!No se encuentra el origen de la referencia.**] y la secuencia del riboswitch se obtuvo de los datos reportados en “Conserved regulatory motifs in bacteria: riboswitches and beyond” (Abreu-Goodger *et al.* (2004)) (Apéndice 2) [**Error!No se encuentra el origen de la referencia.**]. Esta secuencia contiene los tres primeros motivos identificados que se complementan para formar la estructura básica del riboswitch del riboswitch.

La selección de las bases conservadas en las asas del riboswitch se hizo de acuerdo a la estructura secundaria formada por los 3 motivos que componen al riboswitch y a los valores de probabilidad de cada base [**Error!No se encuentra el origen de la referencia.**]. Se seleccionaron 5 bases en el motivo 1 (G60, A61, U68, C77, G78), 1 base en el motivo 2 (G19) y 6 bases en el motivo 3 (C38, U39, G40, A41, G42, A43), estas bases son prácticamente invariables (>97% de los casos analizados presentan la misma base) Se seleccionó también la base U79A como control, ya que es una base menos conservada (56% conservada).

1.2. Clonación de la *rrthiM*

Se amplificó la *rrthiM* por una reacción en cadena de polimerasa (PCR) y después se clonó en el vector pUC19.

Diseño y preparación de oligos

Para la amplificación se diseñaron los oligos: *thiM*_up y *thiM*_do, que son complementarios a la *rrthiM* y tienen los sitios de restricción *EcoR1* y *BamH1* respectivamente. Para diseñarlos se utilizó el programa oligo 4.0 y se sintetizaron mediante el Método de Fosfito-Triéster en Fase Sólida, en la unidad de síntesis del Instituto de Biotecnología, UNAM.

*thiM*_up (5' → 3'): TTAG**GAATTC**ACTGCCGTTTTTCCTCGTTTAC *EcoR1*

*thiM*_do (3' → 5'): TAAG**GATCC**CAGCAGGTCGACTTGCATAGTTTG *BamH1*

Amplificación de la *rrthiM*

Con los oligos thiM_up y thiM_do se amplificó la *rrthiM*, utilizando como templado el DNA total de la cepa MG1655 de *Escherichia coli* y se realizó una reacción de amplificación con Ampli Taq DNA polimerasa en un volumen de 30µl con: 1µl del oligo thiM_up (20picomol/µl), 1µl del oligo thiM_do (20picomol/µl), 1µl de mezcla de deoxiribonucleósidos 5'-trifosfato (dNTPs) (0.25mM), 3µl de PCR Buffer II (500mM, 100mM Tris-HCl, pH 8.3), 1µl del DNA templado (200ng), 1.5µl de MgCl₂ (25mM), 19µl de H₂O y 0.5µl de Ampli Taq DNA polimerasa (5u/µl).

Las condiciones del PCR, fueron: una primera desnaturalización del DNA a 94°C durante 7min; luego 30 ciclos de: desnaturalización a 94°C durante 1.5min, anillamiento a 50°C durante 1.5min y una polimerización a 70°C durante 2min y un último ciclo de polimerización a 70°C durante 5min.

Digestión y ligación de la *rrthiM* y puc19

Se purificó el producto de PCR (*rrthiM*) mediante el método de purificación por columna y el pUC19 mediante extracción de DNA y purificación por columna con QIAquick PCR Purification Kit y QIAprep Spin Miniprep Kit” (QIAGEN).

Se digirió el fragmento de la *rrthiM* y el vector pUC19 con las enzimas EcoR1 y BamH1 en un volumen final de 50µl: 30µl del fragmento de la *rrthiM* y pUC19 respectivamente (120µM y 200ng/µl respectivamente), 8µl de H₂O, 5µl de Buffer EcoR1 (50mM NaCl, 100mM Tris-HCl, 10mM MgCl₂, 0.025% Triton X-100 (pH 7.5)), 5µl de BSA (150mM NaCl), 1µl de la enzima EcoR1 (20u/µl) y 1µl de la enzima BamH1 (20u/µl).

Se ligó la *rrthiM* y el vector pUC19 en una reacción con T4 DNA ligasa, en un volumen final de 120µl: 50µl de fragmento de la *rrthiM* y pUC19 respectivamente, 12µl de buffer ligasa (50mM Tris-HCl, pH 7.5, 10mM MgCl₂, 10mM dithiothreitol, 1mM ATP, 25µg/ml suero albúmina bovino), 6µl de H₂O, 2µl de T4 DNA ligasa (400u/µl).

El producto de la ligación: *rrthiM* + pUC19 se limpió de sales por el método de purificación de membrana y se transformó en células DH5α electrocompetentes. Posteriormente se purificó DNA de plásmido de colonias y se mandó secuenciar por el método de “Taq FS Dye Terminator Cycle Sequencing Fluorescence-Based Sequencing”, por la unidad de secuencia del Instituto de Biotecnología, UNAM.

1.3. Mutagénesis sitio dirigida

Los oligos que se utilizaron para los PCR mutagénicos se diseñaron de acuerdo al programa: “Quick Change ® multi Kit primer design”, <http://labtools.stratagene.com/QC>. Se sintetizaron los oligos de la cadena sentido (5' > 3') de cada mutante mediante el Método de Fosfito-Triéster en Fase Sólida, en la unidad de síntesis del Instituto de Biotecnología, UNAM. Los oligos utilizados se muestra en la Tabla 3.

Tabla 3
Oligos mutagénicos

Oligo	Secuencia del Oligo	T_m
G19C	5'-ACCAAACGACTCGGG G TGCCCTTCTGCGTGAA-3'	57°C
Antisentido G19C	5'-TTCACGCAGAAGGG C CCCGAGTCGTTTGGT-3'	57°C
C38G	5'-CCCTTCTGCGTGAAGG G TGAGAAATACCCGTATC-3'	80.4°C
Antisentido C38G	5'-GATACGGGTATTTCT C CCCTTACGCAGAAGGG-3'	80.4°C
U39A	5'-CCTTCTGCGTGAAGG C AGAGAAATACCCGTATCA-3'	79.2°C
Antisentido U39A	5'-TGATACGGGTATTTCT T GCCTTACGCAGAAGG-3'	79.2°C
G40C	5'-CTTCTGCGTGAAGG C TAGAAATACCCGTATCAC-3'	79.2°C
Antisentido A107C	5'-GTGATACGGGTATTTCT G AGCCTTACGCAGAAG-3'	79.2°C
A41U	5'-CTGCGTGAAGGCTG T GAAATACCCGTATCAC-3'	79.1°C
Antisentido A41U	5'-GGTGATACGGGTATTT C ACAGCCTTACGCAG-3'	79.1°C
G42C	5'-CTGCGTGAAGGCTG C AAATACCCGTATCAC-3'	79.1°C
Antisentido G42C	5'-GGTGATACGGGTATTT G TCAGCCTTACGCAG-3'	79.1°C
A43U	5'-CTGCGTGAAGGCTGAG T AATACCCGTATCACCTG-3'	80.4°C
Antisentido A43U	5'-CAGGTGATACGGGTATT A CTCAGCCTTACGCAG-3'	80.4°C
G60C	5'-AATACCCGTATCAC C TATCTGGATAATGCCAGC-3'	78°C
Antisentido G60C	5'-GCTGGCATTATCCAGAT G AGGTGATACGGGTATT-3'	78°C
A61U	5'-ATACCCGTATCACCTG T TCTGGATAATGCCAGCG-3'	79.2°C
Antisentido A61U	5'-CGCTGGCATTATCCAG A CAGGTGATACGGGTAT-3'	79.2°C
U68A	5'-TATCACCTGATCTGG A AATGCCAGCGTAGGGAA-3'	78°C
Antisentido U68A	5'-TTCCCTACGCTGGCATT T TCCAGATCAGGTGATA-3'	78°C
C77G	5'-CTGGATAATGCCAG G GTAGGGAAGTCACGG-3'	78.9°C
Antisentido C77G	5'-CCGTGACTTCCCT A CCTGGCATTATCCAG-3'	78.9°C
G78C	5'-CTGGATAATGCCAG C TAGGGAAGTCACGGAC-3'	80.3°C
Antisentido G78C	5'-GTCCGTGACTTCCCT A GCTGGCATTATCCAG-3'	80.3°C
U79A	5'-CTGGATAATGCCAG C AGGGAAGTCACGGAC-3'	80.3°C
Antisentido U79A	5'-GTCCGTGACTTCCCT T CGCTGGCATTATCCAG-3'	80.3°C

Para obtener las mutantes se utilizó “QuickChange ® Multi Site-Directed Mutagenesis Kit” (Stratagene). Este “Kit” utiliza sólo un oligo para amplificar DNA templado de doble cadena (dsDNA) y amplifica a partir de vectores pUC. El proceso de mutagénesis se divide en tres etapas: La etapa 1 que corresponde a varios ciclos de PCR, donde se utiliza la enzima DNA polimerasa *PfuTurbo* para amplificar el oligo que tiene el cambio, la mella en el DNA después es cerrada por la mezcla de la enzima “QuickChange ® Multi”. En la segunda etapa, el producto de PCR se pone a digerir con la enzima DpnI, que es una endonucleasa (secuencia blanco 5'Gm⁶ATC-3') específica para DNA metilado y hemimetilado, con la cual se elimina el DNA parental. En la tercera etapa, el producto digerido con DpnI es transformado en células DH5α electrocompetentes.

Reacciones de PCR

Se realizó una reacción de PCR para cada uno de los oligos en 25µl: 2.5µl de Buffer “QuickChange ® Multi Reaction”, 0.75µl de “QuickSolution”, 4µl de templado (DNA plasmídico de la construcción que contiene la *rrthiM*: *pthiM*) (50ng), de 1.3 a 1.6µl de oligo según la concentración de cada uno (100ng), 1µl de la mezcla de dNTPs, 1µl de la mezcla de la enzima “QuickChange ® Multi” (2.5u/µl) y H₂O para un volumen final de 25µl.

Los parámetros para los ciclos de PCR fueron: Un primer ciclo de 1min a 95°C, 30 ciclos de 1min a 95°C, 1min a 55°C y 6min a 65°C (2min por cada Kb) y un ultimo ciclo de 10min a 72°C. El producto del PCR mutagénico se digirió con 1ul de la enzima DpnI por reacción (10u/ul) y se transformó en células DH5α electrocompetentes. Por último se secuenciaron los DNAs de colonias de cada mutante y se seleccionaron las que tenían el cambio propuesto.

1.4. Fusiones traduccionales

Se hicieron fusiones traduccionales de la *rrthiM* y de cada fragmento mutante, al vector pRS552. Las construcciones realizadas son fusiones traduccionales de los 5 primeros codones y las dos primeras bases del sexto codón de la *rrthiM* en marco al noveno codón de *lacZ*.

Amplificación, digestión y ligación de la *rrthiM* y mutantes a pRS552

Se purificaron los plásmidos que contienen la *rrthiM* y cada una de las regiones mutantes, por el método de columna “QIAprep®”, “HiSpeed Plasmid Midi” (QIAGEN) y se digirieron con *EcoR*I y *Bam*H1: *pthiM*/mutante (10µg), 7µl de Buffer *EcoR*I (50mM NaCl, 100mM Tris-HCl, 10mM MgCl₂, 0.025% Triton X-100 (pH 7.5), 7µl de BSA (150mM NaCl), 1ul de la enzima *EcoR*I (20u/µl), 1µl de la enzima *Bam*H1 (20u/µl) y H₂O para un volumen final de 70µl.

Se amplificó la *rrthiM* (igual a la primera amplificación) y de cada una de las mutantes a partir del DNA plasmídico mediante PCR y se purificó el fragmento por el método de columna “QIAprep®”, “PCR purification Kit” (QIAGEN).

Se digirió el vector pRS552 (0.1µg/µl) con las enzimas EcoR1 y BamH1 en un volumen final de 30µl: 22µl de H₂O, 3µl de Buffer EcoR1 (50mM NaCl, 100mM Tris-HCl, 10mM MgCl₂, 0.025% Triton X-100 (pH 7.5)), 3µl de BSA (150mM NaCl), 1µl de la enzima EcoR1 (20u/µl) y 1µl de la enzima BamH1 (20u/µl).

Se digirió cada uno de los fragmentos amplificados por PCR con las enzimas EcoR1 y BamH1 en un volumen final de 60µl: 4µl H₂O, 2µl de Buffer EcoR1 (50mM NaCl, 100mM Tris-HCl, 10mM MgCl₂, 0.025% Triton X-100 (pH 7.5)), 1µl BSA 100x (150mM NaCl), 2µl de la enzima EcoR1 (20u/µl) y 2µl de la enzima BamH1 (20u/µl).

Se ligó la *rrthiM* y cada una de las mutantes con el vector pRS552 previamente digeridos y purificados, en una reacción con T4 DNA ligasa: *rrthiM*/mutante (1µg), 6µl de buffer ligasa (50mM Tris-HCl, pH 7.5, 10mM MgCl₂, 10mM dithiothreitol, 1mM ATP, 25µg/ml suero albúmina bovino), 1µl de T4 DNA ligasa (400u/µl) y H₂O para un volumen de 60µl.

El producto de la ligación (*rrthiM*/mutantes + pRS552) se limpió de sales por el método de purificación de columna y se transformó en células electrocompetentes MG1655. A partir de las colonias recuperadas se purificó DNA plasmídico y se pusieron reacciones de PCR para amplificar los fragmentos mutantes, con las mismas condiciones que se utilizaron para amplificar la *rrthiM* y se revisó por secuencia.

1.5. Ensayos de β-galactosidasa

Las clonas de *Escherichia coli* con las diferentes fusiones traduccionales se crecieron en 100ml de medio mínimo con antibiótico kanamicina al 10%¹ en presencia y en ausencia de tiamina (tiamina 1%) en el medio, a 37°C y a 2500rpm. A las 24hr se centrifugaron 30ml del medio por 5min a 6000rpm y se resuspendió la pastilla en 50ml de MM Km₁₀, se dejó crecer el cultivo por 12hr más en las mismas condiciones, hasta obtener una densidad de cultivo mínima de 1.2 densidades ópticas a 600nm (DO₆₀₀).

Para medir la actividad de la enzima β-galactosidasa a la cual están fusionadas las distintas regiones de regulación (silvestre y mutantes), se hicieron ensayos con el sustrato orto-nitrofenil-β-D-galactosa (ONPG). Los ensayos se hicieron por triplicado en cada ocasión. Se tomaron alícuotas de 1ml del cultivo de cada una de las fusiones y se permeabilizaron las células adicionándoles 30µl de cloroformo CHCl₃ y 15µl de SDS al 0.1%. Las células se mezclaron y se dejaron en reposo, a los 15min. se mezclaron nuevamente y se dejaron reposar por otros 15min. Para medir actividad se tomaron 100µl de células permeabilizadas y se pusieron en 0.80ml de buffer Z con β-mercaptoetanol, (buffer Z: 60mM Na₂HPO₄, 40mM NaH₂PO₄, 10mM KCl, 1mM MgSO₄.7H₂O; más 27µl β-mercaptoetanol por cada 10ml de buffer Z, 50mM final). Se adicionó 0.1ml de ONPG (26mM) y se registró el tiempo 0. La reacción se llevó a cabo a 37°C por 5min. y se detuvo con 500ml de Na₂CO₃ 1M. Por último se midió densidad óptica a 420nm (DO₄₂₀) que es la densidad a la que absorbe la molécula de orto-nitrofenil (ONP) que se debe liberar por la acción de la enzima β-galactosidasa.

¹ El Medio Mínimo (MM) contiene: para 100ml, 20ml de sales M95X (Na₂ HPO₄, KH₂PO₄, NaCl, NH₄Cl), 10µl CaCl₂ 1M, 100µl MgSO₄.7H₂O, 1ml de glucosa 20% y 78.89ml de H₂O, más 100µl de km₁₀.

Los resultados se normalizaron por concentración de proteína total, para lo cual se midió las concentraciones de proteína por espectrofotometría a 595nm mediante el método de Bradford (Bio-Rad Protein Assay Catalog 500-0006). Primero se hizo una curva estándar de albúmina (BSA) con concentraciones de 1 a 30 μg de proteína por ml (Fig. 14). Se midió la densidad óptica de las reacciones de “Bradford” a 595 (DO_{595}) y los μg contenidos en la muestra se calcularon a partir de la ecuación lineal de la curva de BSA: $y = 0.0447x + 0.066$, donde x son los μg de proteína ($x = (\text{DO}_{595} - \text{ordenada}) / \text{pendiente}$) y los $\mu\text{g}/\mu\text{l}$, corresponden a los μg entre los μl utilizados (100 y 200 μl).

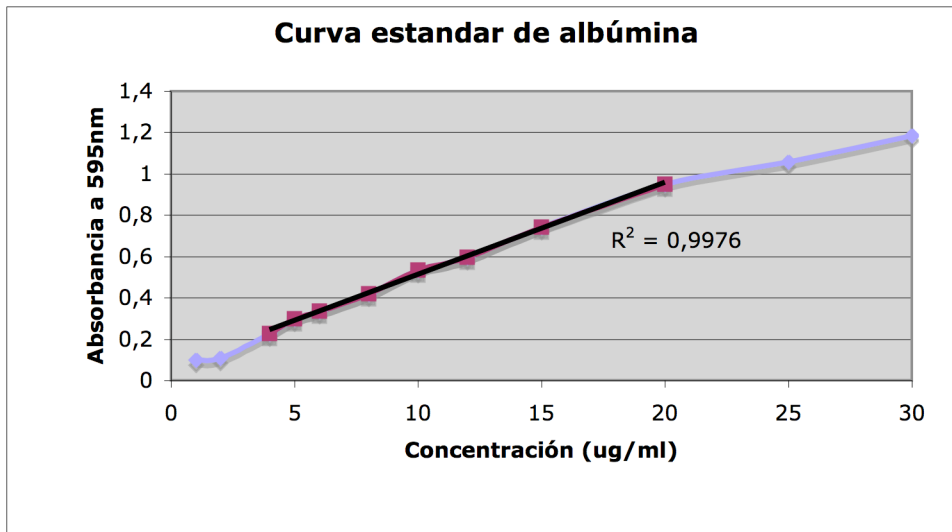


Fig. 14. Curva estándar de albúmina. Para la curva se midieron concentraciones desde 1 hasta 30 μg de proteína BSA. Para la regresión lineal se tomaron los valores de 4 a 20 μg que ajustan a un $r^2=0.9$

Por último se calculó la actividad específica (actividad específica de la enzima β -galactosidasa) en cada uno de los ensayos midiendo la densidad óptica de cada muestra a 420nm (DO_{420}) y la concentración de proteína por ensayos de “Bradford”. La actividad específica es igual a: (nmolas de ONPG consumidas/tiempo de reacción)*miligramos de proteína. Donde nmolas de ONPG consumidas es igual a la $\text{DO}_{420} * 1000$, tiempo de reacción: 5min y miligramos de proteína obtenidos por los ensayos de “Bradford” según la concentración de proteína en cada muestra.

1. Resultados y discusión

1.1. Selección de la región reguladora del gene *thiM*

La secuencia promotora específica para el gene *thiM* de *Escherichia coli* no se ha determinado experimentalmente, por lo que se realizó una búsqueda de la secuencia promotora utilizando la herramienta de predicción de operones de la base de datos de regulación transcripcional: RegulonDB.

(http://regulondb.ccg.unam.mx:80/html/Computational_Predictions.jsp) [**Error!No se encuentra el origen de la referencia.**, **Error!No se encuentra el origen de la referencia.**]

De ésta se obtuvieron 5 posibles secuencias promotoras (Fig. 15), de las cuales se seleccionó la secuencia: -35:TTTACA, -10:TACCAT, ya que ésta coincide con los primeros dos promotores predichos (que son las secuencias más fuertes identificadas) y presenta mayor similitud a la secuencia consenso -35, -10 (TTGACA - N17 - TATAAT). Para evitar no incluir la secuencia promotora, se seleccionó una secuencia de 17pb antes de la secuencia propuesta.

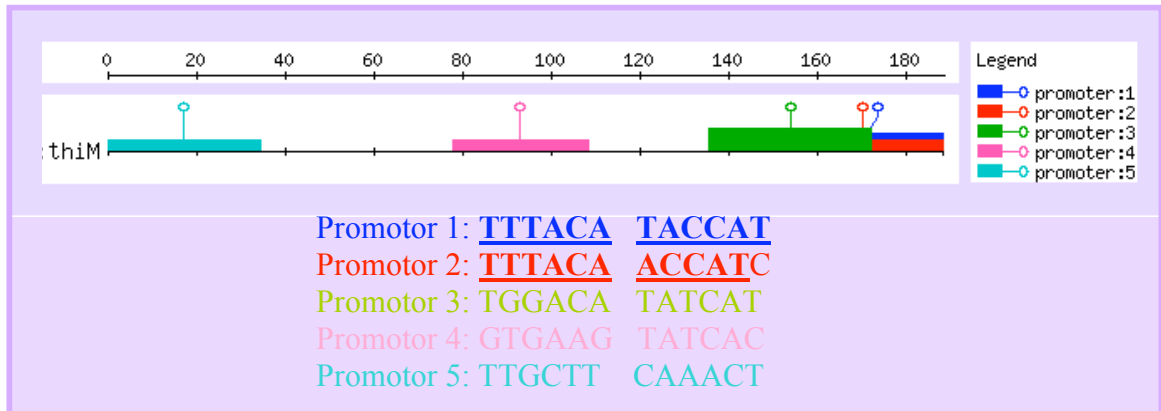


Fig. 15. Predicción de operones de regulonDB. En la figura se representa la secuencia líder del gene *thiM* y las posibles secuencias promotoras. El promotor 1 y 2, coinciden con la secuencia seleccionada, el promotor 4 sobrelapa con la secuencia del riboswitch y el promotor 5 con la secuencia SD. El extremo 5' está en el punto 0 al inicio el gene y el 3' en la posición 180.

La *rrthiM* (región reguladora de *thiM*) con la que se trabajó, está formada por 219 pares de bases (pb): 202pb río arriba del codón de inicio (ATG) y 17pb que incluyen a los primeros 5 codones del gene *thiM* (Fig. 16). Esta secuencia contiene al promotor y a los tres motivos conservados del riboswitch de tiamina. En la secuencia se seleccionaron también los primeros 5 codones de *thiM* y dos bases del sexto que en la fusión traduccional al vector pRS552 quedan en marco con el noveno codón del gene *lacZ*.

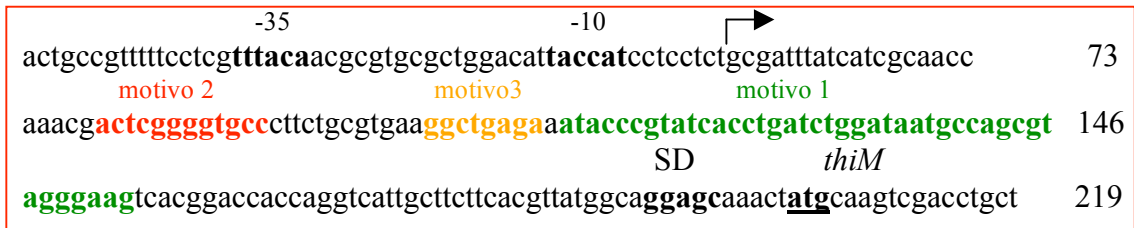


Fig. 16. Secuencia mínima de regulación. El la figura se muestra la región promotora (caja -35 y -10), el posible sitio de inicio de la transcripción (→), en orden los tres motivos que forman al riboswitch (motivo 2, 3 y 1), la secuencia “Shine Dalgarno” (SD) y el inicio de traducción (ATG).

1.2. Selección de las bases conservadas en las asas del riboswitch.

Para seleccionar las bases conservadas en las asas del riboswitch, se utilizaron los motivos conservados obtenidos en el trabajo de “Conserved regulatory motifs in bacteria: riboswitches and beyond” (Abreu-Goodger *et al.* 2004) (Apéndice 2) [¡Error!No se encuentra el origen de la referencia.]. Cada motivo fue evaluado de forma independiente, en ellos se identificó la parte de la secuencia más conservada y se eliminaron las no conservadas, quedando los motivos de la siguiente forma:

- Motivo 1 ACCCUUGAACCUGAUCCGGUAAUACCGGCGUAGGGAAG (40pb), se eliminaron las primeras 10 bases**
- Motivo 2 AGGGGUGCC (9pb), se eliminaron las primeras 4 bases**
- Motivo 3 GGCUGAGAG (9pb), se eliminaron las primeras 2 bases**

El método de selección de motivos del trabajo de Abreu-Goodger *et al.* (2004), tiene parámetros generales para buscar regiones conservadas en genomas totales. En casos específicos como lo son los riboswitches la búsqueda debe ser más dirigida, sin embargo los datos reportados nos da una buena aproximación (ver apéndice 2). También se reporta un cuarto motivo para el grupo 12 (CTTTTTTTGAATGTGGTTTTTTTTTTT) que parecería corresponder a un atenuador y que no se encuentra presente en la *rrthiM* [¡Error!No se encuentra el origen de la referencia.].

De acuerdo a los datos reportados, se identificó el orden en que los motivos se presentan:

..._[2]_1_[3]_?_[1]_..._[4]_...
 Motivo 2 - Motivo 3 - Motivo 1

AGGGGUGCC-GGCUGAGAG-ACCCUUGAACCCUGAUCCGGUAAUACCGGCGUAGGGA

Y con esta firma se construyó una estructura secundaria consenso:

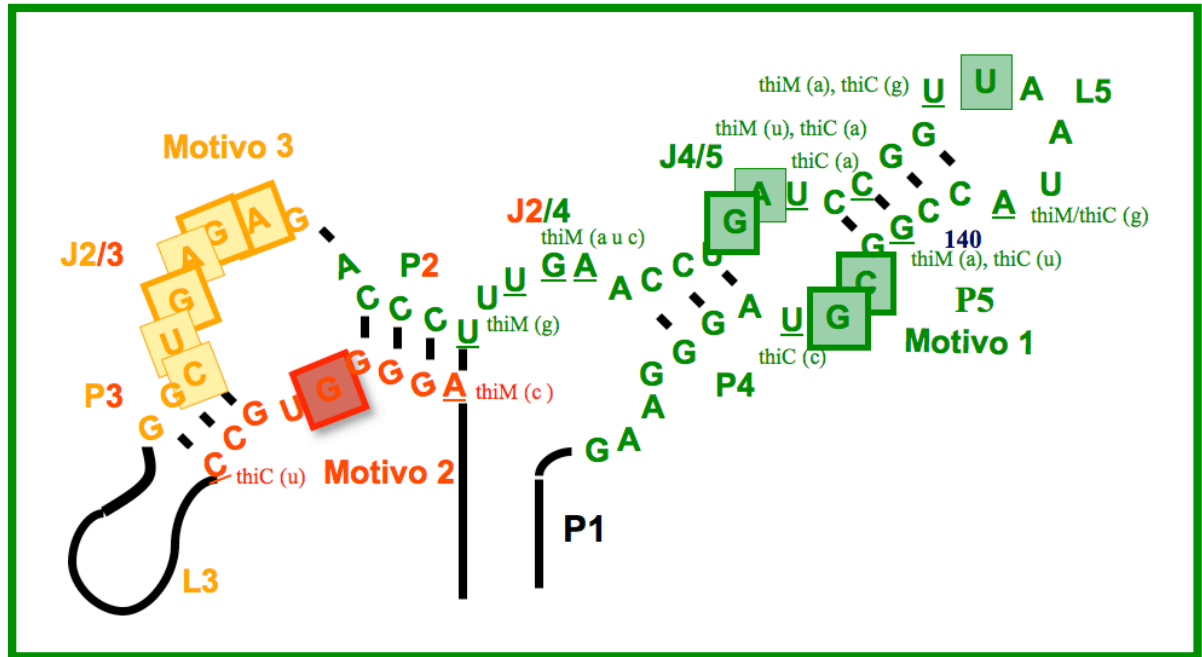


Fig. 17. Secuencia y estructura secundaria consenso del riboswitch de tiamina. La secuencia está compuesta por los tres motivos conservados. En el esquema se muestran las bases que difieren del consenso, para los genes *thiC* y *thiM* de *Escherichia coli*. En la estructura, P1, P2, P3, P4 y P5, corresponden a los tallos; L3, J2/3, J2/4, J4/5 y L5 corresponden a las asas. Las bases conservadas en las asas que se seleccionaron están marcadas con cuadros.

La estructura secundaria se generó de acuerdo a la secuencia consenso y a las estructuras secundarias propuestas anteriormente: Miranda-Ríos *et al.* (2001) y Winkler *et al.* (2002) [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.] (Fig. 17). A partir de ésta se identificaron las bases que forman las asas y se comparó los valores de frecuencia de aparición de cada base de los motivos. Se seleccionaron 12 bases conservadas con valores por arriba de 0.97 (97% de conservación) que después se identificaron en la *rrthiM* como: G60, A61, A35, C77, G78 en el motivo 1, G19 en el motivo 2 y C38, U39, G40, A41, G42 y A43 en el motivo 3 (la numeración va de acuerdo a la antes dada para el riboswitch del gene *thiM* de *Escherichia coli* (Winkler *et al.* (2002)) [¡Error!No se encuentra el origen de la referencia.]. También se seleccionó como control negativo a la base U79 que está menos conservada (50% de conservación). En la Fig. 18 se presentan las tablas de probabilidad de los motivos seleccionados.

MOTIVO 1				
A	C	G	U	
0,75	0,04	0,07	0,14	A47
0,18	0,69	0,01	0,12	C48
0,04	0,89	0,01	0,07	C49
0,00	0,93	0,02	0,04	C50
0,10	0,18	0,27	0,46	U51
0,07	0,17	0,07	0,69	U52
0,25	0,17	0,13	0,45	U53
0,09	0,01	0,73	0,17	G54
0,74	0,07	0,11	0,07	A55
1,00	0,00	0,00	0,00	A56
0,00	0,96	0,02	0,15	C57
0,00	0,97	0,00	0,03	C58
0,00	0,04	0,01	0,96	U59
0,00	0,00	1,00	0,00	G60
0,97	0,03	0,00	0,00	A61
0,14	0,03	0,01	0,82	U62
0,08	0,77	0,06	0,09	C63
0,11	0,47	0,04	0,38	C64
0,31	0,01	0,67	0,15	G65
0,03	0,00	0,97	0,00	G66
0,26	0,09	0,20	0,45	U67
0,01	0,01	0,00	0,99	U68
0,69	0,25	0,01	0,06	A69
0,80	0,00	0,20	0,00	A70
0,04	0,12	0,06	0,79	U71
0,64	0,01	0,34	0,02	A72
0,00	0,97	0,00	0,03	C73
0,01	0,66	0,01	0,31	C74
0,38	0,03	0,47	0,13	G75
0,13	0,04	0,82	0,01	G76
0,00	1,00	0,00	0,00	C77
0,00	0,00	1,00	0,00	G78
0,26	0,05	0,12	0,56	U79
0,90	0,01	0,07	0,03	A80
0,05	0,00	0,95	0,00	G81
0,01	0,02	0,96	0,00	G82
0,20	0,01	0,79	0,00	G83
0,98	0,00	0,02	0,00	A84
0,66	0,05	0,13	0,15	A85
0,09	0,19	0,53	0,19	G19

MOTIVO 2				
A	C	G	U	
0,52	0,29	0,14	0,05	A15
0,00	0,03	0,97	0,00	G16
0,02	0,02	0,97	0,00	G17
0,11	0,00	0,89	0,00	G18
0,00	0,00	1,00	0,00	G19
0,23	0,02	0,05	0,71	U21
0,06	0,23	0,92	0,00	G22
0,00	0,86	0,00	0,14	C23
0,05	0,59	0,04	0,32	C24

MOTIVO 3				
A	C	G	U	
0,13	0,03	0,78	0,06	G36
0,08	0,00	0,90	0,02	G37
0,00	1,00	0,00	0,00	C38
0,00	0,00	0,00	1,00	U39
0,00	0,00	1,00	0,00	G40
0,99	0,00	0,01	0,00	A41
0,00	0,01	0,99	0,00	G42
1,00	0,00	0,00	0,00	A43

Fig. 18. Matrices de probabilidad de los motivos. Cada columna corresponde a una base, al lado derecho esta la secuencia que se va seleccionando, en gris están las bases que están comprometidas en tallos y en amarillo las bases seleccionadas con un valor de frecuencia de aparición de arriba del 97% y de 56% en el control. La búsqueda se realizó sobre un total de 218 secuencias relacionadas a la biosíntesis de tiamina. La numeración corresponde a la antes dada por Winkler *et al.* (2002) para el riboswitch [;Error!No se encuentra el

1.3. Condiciones para ver la represión por tiamina.

La *rrthiM*¹ se clonó al vector pUC19 utilizando los sitios *EcoR*I y *Bam*H1 y se secuenció para revisar su secuencia. Después se realizó una subclonación al vector pRS552 y se obtuvo el vector que contiene la fusión traduccional de la *rrthiM* al gene reportero de β -galactosidasa, a esta construcción se le llamo *pthiM*. Después a partir de un PCR mutagénico se obtuvieron las distintas mutantes y se hicieron las fusiones traduccionales. La construcción *pthiM* y las fusiones traduccionales de cada mutante se clonaron en la cepa MG1655 de *Escherichia coli* que es una cepa silvestre y es protótrofa para tiamina y a partir de estas se hicieron ensayos de actividad de β -galactosidasa.

Primero se caracterizó *pthiM* y se establecieron las condiciones de crecimiento para ver un fenotipo de represión por tiamina. El medio LB (Luria-Bertani) es un medio rico que se utiliza comúnmente para cultivos de bacterias, éste tiene una concentración alta de tiamina (suficiente para mantener reprimida la expresión de la construcción). Para evitar la alta concentración de tiamina los cultivos se hicieron en medio mínimo (MM) que sólo contiene sales, NH_4Cl como fuente de nitrógeno y glucosa como fuente de carbono. A pesar de que los cultivos carecen de tiamina, la tiamina que traen las bacterias desde su inoculación por su crecimiento previo en LB es suficiente para reprimir y evitar que se vea un marcado efecto en la actividad de β -galactosidasa entre los cultivos con tiamina y sin tiamina. Para evitar esto, se llevaron a cabo dos cultivos consecutivos, un cultivo inicial en 10ml en MM durante 24 hr del cual se centrifugaron 30ml y se resuspendieron en un segundo cultivo de 50ml de MM durante 12 hr. Los cultivos se pusieron en condiciones estándar de crecimiento para *Escherichia coli* (37°C a 250rpm).

Para registrar el crecimiento y la actividad de β -galactosidasa en los cultivos se tomaron alícuotas cada hora durante las 12 hrs del segundo crecimiento en MM. Se registraron las densidades ópticas a 600 nm (DO_{600}) para cada muestra y se llevaron a cabo ensayos de actividad de cada alícuota. Se graficó la curva de crecimiento del cultivo y se cuantificó la actividad de *pthiM* en presencia y ausencia de tiamina (Fig. 20). De estos registros se observó una represión promedio de 8.5 veces menos actividad en presencia de tiamina a partir de las 6 y hasta las 12 hrs de crecimiento. En contraste, en los ensayos de actividad de β -galactosidasa de Winkler *et al.* (2002), se reportaron valores de 10 veces menos actividad en presencia de tiamina². Si se compara la curva de crecimiento y los valores de actividad, se observa que la represión empieza al inicio de la fase logarítmica de crecimiento (aprox. a las 6 hr) lo cual corresponde a una concentración de cultivo mayor de $\text{DO}_{600} = 0.5$. Estos ensayos nos ayudaron a definir a partir de qué densidad de cultivo y a qué tiempo es recomendable tomar la muestra; en el trabajo las muestras se tomaron a las 12 hrs, a más de 0.5 DO_{600} .

Para definir las condiciones de los ensayos de β -galactosidasa en las que se obtienen los valores más altos de represión, se utilizó una concentración de sustrato de

¹ La *rrthiM* contiene al promotor y a los tres motivos conservados del riboswitch de tiamina. Contiene también los primeros 5 codones de *thiM* y dos bases del sexto que en la fusión traduccional al vector pRS552 quedan en marco con el noveno codón del gene *lacZ*.

² Los valores de β -galactosidasa de los ensayos reportados por Winkler *et al.* (2002) están dados en función de unidades Miller (unidad Miller = $1000 * (\text{DO}_{420} * 1.75 \text{DO}_{550}) / \text{tiempo}(\text{min}) * V(\text{ml}) * \text{DO}_{600}$), para comparar estos datos, los primeros ensayos de actividad de *pthiM* también están en unidades Miller, sin embargo los ensayos de actividad con las mutantes se reportan como actividad específica.

26 mM (8mg/ml) de ONPG a la cual se sabe que la actividad de la enzima β -galactosidasa es óptima [¡Error!No se encuentra el origen de la referencia.] y se hicieron ensayos con distintas cantidades de células permeabilizadas (de 50 a 250 μ l) y ensayos parando la reacción a distintos tiempos (de 3 a 9 min.)³ en los cuales se midió la DO₄₂₀ en la que se cuantifica el ONP producido en la reacción (Fig. 19).

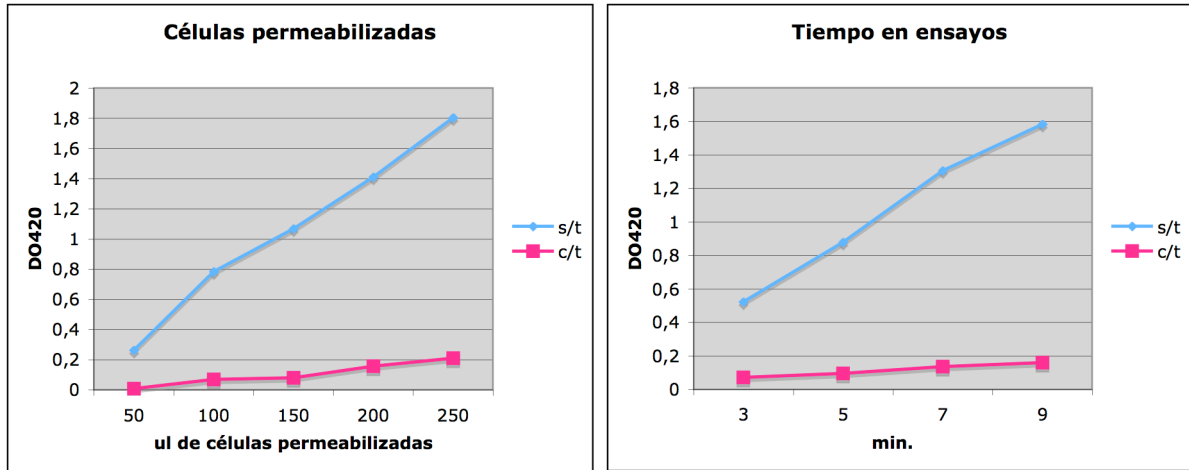


Fig. 19. Condiciones de linealidad en los ensayos.

Se decidió tomar 100 μ l de células permeabilizadas para realizar los ensayos y parar la reacción a los 5 min para trabajar en condiciones de linealidad. Con estas condiciones se ven diferencias de hasta 11 veces menos expresión de β -galactosidasa en el cultivo con tiamina (ex. 0.78 DO₄₂₀ en el cultivo sin tiamina y 0.07 DO₄₂₀ en el cultivo con tiamina).

Para cuantificar la expresión intrínseca del gene de la β -galactosidasa (*lacZ*) en la cepa silvestre MG1655 y en la que tiene al vector sin inserto (pRS552) se hicieron ensayos de actividad específica y se vió que la cepa silvestre no presenta expresión y que la cepa con el vector pRS552 tiene una expresión similar a la que presenta pthiM reprimido (pthiM con tiamina), es decir en condiciones de represión para pthiM la actividad de β -galactosidasa baja al nivel de la del vector. (Fig. 17).

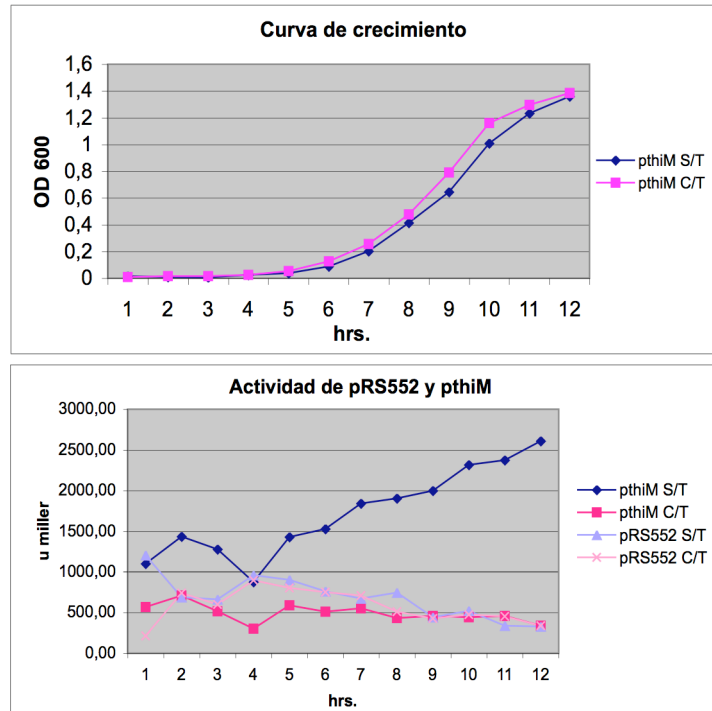


Fig. 20. Curva de crecimiento y valores de β -galactosidasa para pthiM y pRS552. Los valores de represión por tiamina son más marcados después de las 6hr de crecimiento, cuando el cultivo entra en la fase logarítmica

³ La enzima β -galactosidasa es activa a pH 10-11, para inactivarla se adiciono Na₂CO₃ que baja el pH, lo hace más alcalino.

1.4. Fenotipos en las mutantes

9.4.1. Actividad de β -galactosidasa en las mutantes

A partir de pthiM se realizaron PCRs mutagénicos, se secuenciaron las candidatas y se seleccionaron las que tienen el cambio. Se trató de que el cambio fuera lo más drástico posible, por lo que las bases pirimidinas se cambiaron por bases purinas y viceversa. Después se hizo una subclonación al vector pRS552, y se obtuvieron las fusiones traduccionales de cada mutante al gene *lacZ*, se transformaron los plásmidos en la cepa MG1655 de *Escherichia coli* y a partir de éstas, se realizaron los ensayos de actividad (medida como actividad específica⁴).

En la tabla 4 se reportan los valores de actividad de β -galactosidasa de las fusiones traduccionales y el nivel de represión obtenida. Los datos reflejan los fenotipos obtenidos en todos los experimentos. El total de experimentos se reportan en el Apéndice 3.

Tabla 4
Actividades de β -galactosidasa y represión promedio

Clona	Sin tiamina	Con tiamina	Represión
pRS552	2 ± 0.0	1 ± 0.2	-
pthiM	18 ± 0.6	3 ± 0.3	6
G19C	29 ± 0.8	29 ± 0.7	-
C38G	26 ± 4	21 ± 2	-
U39A	33 ± 1	26 ± 0.5	-
G40C	25 ± 0.8	26 ± 0.9	-
A41U	29 ± 0.5	23 ± 0.3	-
G42C	25 ± 4.9	26 ± 0.5	-
A43U	28 ± 1.6	24 ± 0.5	-
G60C	22 ± 1	20 ± 0.4	-
A61U	3 ± 1.4	2 ± 0.4	s/actividad
U68A	27 ± 0.3	28 ± 1.1	-
C77G	9 ± 0.1	7 ± 1.1	s/actividad
G78C	1 ± 0.5	2 ± 0.4	s/actividad
U79A	9 ± 0.8	4 ± 0.4	2

Mutantes en la hélice sensora del HMP
Mutantes en la hélice sensora del PP

Los ensayos de actividad se hicieron con los mismos parámetros establecidos para pthiM y se realizaron por triplicado, se realizó un total de tres experimentos en los que se reproduce en todos los casos el fenotipo de la mutante. Los resultados obtenidos para cada experimento se reportan como el promedio de los tres ensayos. En cada experimento se incluyeron al vector sin inserto (pRS552) y a la fusión pthiM como controles positivos y a la mutante U79A como control negativo. La mutante U79A (56% conservada) presenta el mismo fenotipo que pthiM pero con un nivel de represión

⁴ Actividad específica: (nmolas de ONPG consumidas/tiempo de reacción)*miligramos de proteína. Donde nmolas de ONPG consumidas es igual a la DO₄₂₀*1000. La concentración de proteína se obtuvo por ensayos de "Bradford".

menor. Mientras que *pthiM* tiene valores de represión promedio de 6 veces, la mutante U79A tiene valores de 3 veces.

9.4.2. Estructura del riboswitch de tiamina y resumen de fenotipos de las mutantes

De acuerdo a las estructuras reportadas [;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia.] el riboswitch está compuesto por dos hélices paralelas que reconocen de forma independiente al motivo pirimidina (4-amino-5-hidroximetil-2-metil-pirimidina (HMP)) y al motivo pirofosfato (PP) de TPP. De acuerdo a la estructura secundaria del riboswitch (Fig. 17), éste está formado por 5 tallos nombrados de P1 a P5 y 5 asas nombradas J2/3, J2/4, J4/5, L3 y L5. La hélice sensora de la pirimidina está formado por: (P2-J2/3-P3-L3) y la hélice sensora del pirofosfato está formada por: (J2/4-P4-J4/5-P5-L5) que

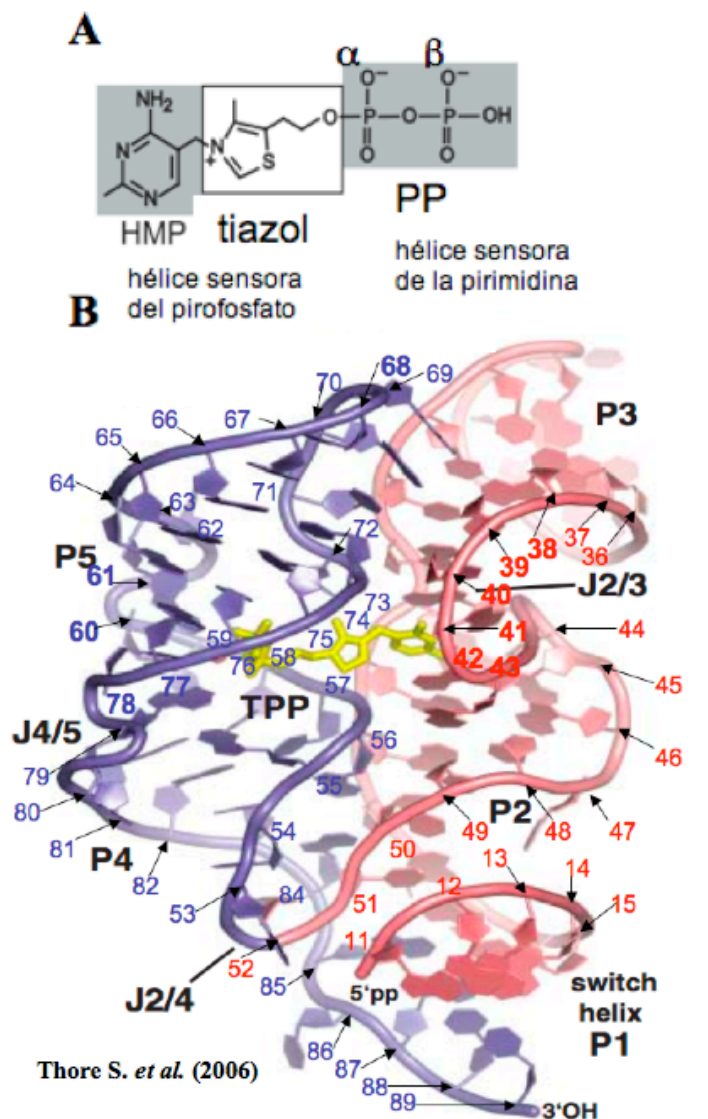


Fig. 21. Estructura del riboswitch de tiamina. A. Muestra la molécula de la tiamina pirofosfato (TPP), compuesta por un motivo hidroxietilpirimidina (HMP), un motivo tiazol y un motivo pirofosfato (PP), el fosfato ligado al tiazol corresponde al fosfato α ($P\alpha$) y el segundo fosfato, corresponde al fosfato β ($P\beta$) B. Riboswitch de *thiC* de *Arabidopsis thaliana* reportado por Thore *et al.* (2006) formado por la hélice sensora del pirofosfato (azul) y por la hélice sensora de la pirimidina (rosa). La numeración esta adaptada a la dada a la secuencia del riboswitch de *thiM*.

corresponde a la “thi-box” (Fig. 21). En el trabajo se analizaron mutantes en ambas hélices sensoras y el resumen de los fenotipos se muestra en el esquema de la Fig. 22.

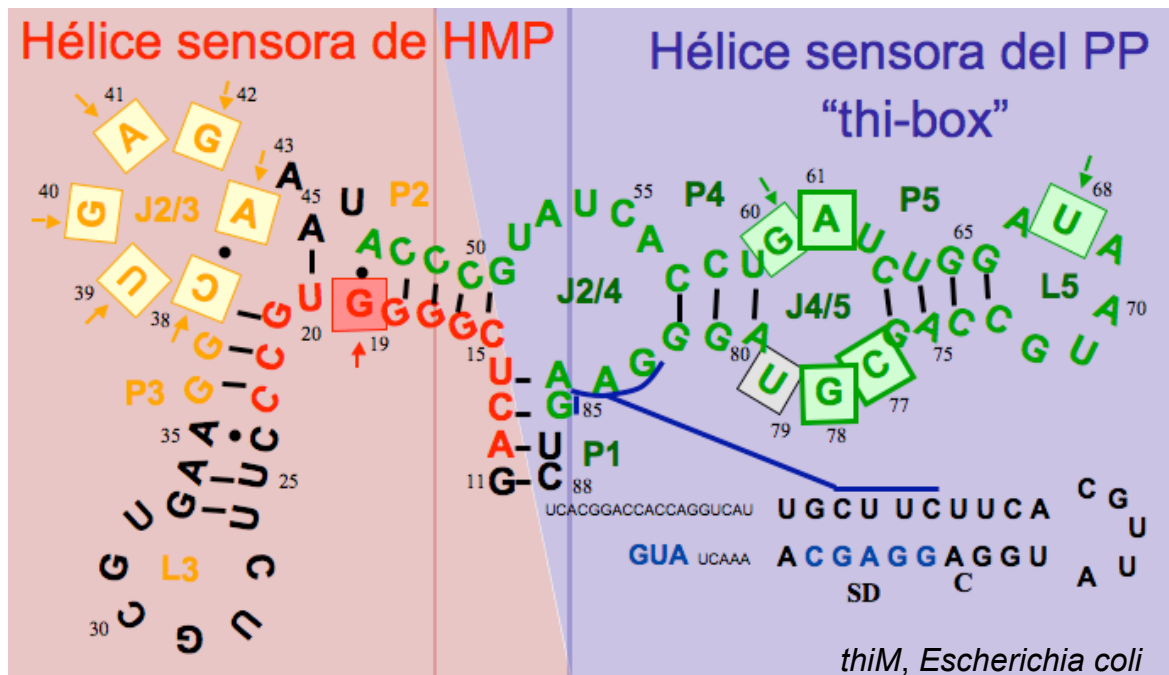


Fig. 22. Resumen de los fenotipos de las mutantes. Las bases mutadas están marcadas en cuadros, las flechas señalan todos los cambios que tuvieron pérdida de la represión por tiamina y las bases con cuadro remarcado son las que perdieron la expresión. En la posición 79 se encuentra el control.

9.4.3. Mutantes en la hélice sensora del HMP

En las mutantes del asa J2/3 (6 bases del motivo 3: C38, U39, G40, A41, G42, A43 y la base G19 del motivo 2), se observó un fenotipo de pérdida de la represión por tiamina (Tabla 4 y Fig. 22); para explicar estos fenotipos, se buscaron las interacciones de estas bases en los datos reportados de estructura [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.]. De acuerdo a las estructuras, el motivo HMP se acomoda entre la base G42 y A43 mediante interacciones de apilamiento (“stacking”) (Fig. 23) y tiene contacto directo con la base G40 y G19 por enlaces de hidrógeno (Fig. 21). En la base G19 el oxígeno 2’ forma un enlace de hidrógeno con la posición N1’ del HMP. En el caso de la base G40, ésta adopta una conformación hacia fuera de la hélice (“flipped out”) que le permite aparearse por enlaces de hidrógeno con el anillo pirimidina del HMP mediante un apareamiento tipo Watson-Crick/Sugar-Edge.

Serganov A. *et al.* (2006)

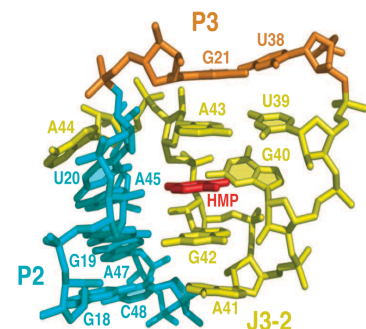


Fig. 23. Interacciones de “stacking”. Las interacciones de “stacking” estabilizan el asa J2/3

La secuencia conservada que va de la base U39 a la A43, se estructura formando una asa tipo T (“T-loop”) [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.], donde la base U39 y A43 tienen un enlace trans-Watson-Crick/Hoogsteen que sirve como un broche para la estructura “T-loop”. La base C38 que se localiza después de las bases que forman el “T-loop” y

está comprometida formando el extremo del tallo P3 con la base G21 (Fig. 24). Las estructuras “T-loop” se han identificado en varios RNAs incluyendo: tRNAs, tmRNAs, ribozimas, el extremo 3’ de genoma viral de plantas (“turnip yellow mosaic virus” – TYMV) y en RNA ribosomal [¡Error!No se encuentra el origen de la referencia.]; este motivo estructural se caracteriza por el enlace UA *trans* Watson-Crick/Hoogsteen apilado a un par de bases con enlaces Watson-Crick en la base de la estructura. En el riboswitch de tiamina esto corresponde al enlace U39•A43 y al inicio del tallo P3 (C38-G21). Así mismo la base A43 adopta una conformación C2’-endo que permite el apilamiento continuo de las bases A41, G42, HMP, A43 y G21, a su vez la conformación se estabiliza por las interacciones de apilamiento y por la formación de triples contactos entre las bases (G19•A47)•G42 y (G18-C48)•A41. La base G40 girada permite el apareamiento con el HMP.

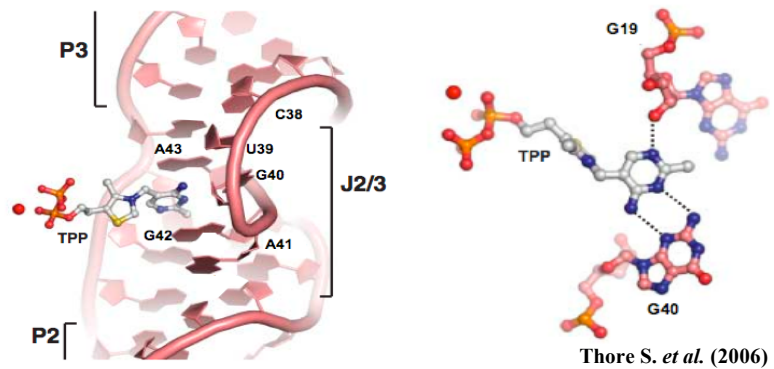


Fig. 24. Motivo HMP. La hélice sensora del HMP se estructura formando una asa tipo T, en ella el HMP se intercala entre la base G42 y A43 y forma enlaces directos con G40 y G19

Después de identificar las interacciones de las bases mutadas, se concluye que la pérdida de la represión es consecuencia directa de afectar al pegado del motivo HMP de la TPP, ya sea por la modificación de las bases que hacen contactos directos (G42 y A43, que tienen interacciones hidrofóbicas (“stacking”) con HMP; G19 y G40, que tienen contactos directos por enlaces de hidrógeno) o por la modificación de las bases que forman el “T-loop” (U39 a A43) y en el caso de la mutante C38G, porque se evita la formación del enlace (C38-G21) y afecta por tanto a la formación del “T-loop”.

9.4.4. Mutantes en la hélice sensora del pirofosfato (PP)

Como se planteó en la hipótesis del proyecto, al mutar las bases conservadas de las asas del riboswitch de tiamina obtenemos fenotipos donde se ve afectada la regulación. Lo interesante es que no sólo se perdió la represión en muchas de las mutantes, sino que también se perdió la expresión en las mutantes: A61U, C77G y G78C que corresponden al asa J4/5 (Tabla 4 y Fig. 22). De acuerdo a los datos de estructura [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.], estas bases son importantes para que se reconozca el PP ya sea por contacto directo con la molécula o por interacciones mediadas por iones bivalentes (Mg^{2+}).

Según Serganov *et al.* (2006) y Edwards y Ferré-D’Amare (2006) en el riboswitch de *thiM* de *Escherichia coli* [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.], presenta dos

apareamientos no canónicos⁵ entre las bases G60•G78 y A61•C77⁶ y a su vez estas bases se asocian mediante iones metal (Mg^{2+}) a la TPP. Uno de ellos se encuentra asociado al fosfato β de la TPP (Mg1) y otro se encuentra coordinando ambos fosfatos (Mg2) de ésta (Fig. 25).

Mutante G60C La mutante G60C presenta un fenotipo de pérdida de la represión (Tabla 4 y Fig. 22). En el caso de la mutante G60C el cambio podría conservar o incluso reforzar la interacción con la base G78 por un enlace guanina-citosina (G-C), por lo que el cambio no modifica la complementariedad del enlace no canónico G60•G78 (Fig. 26). Por lo tanto la pérdida de la represión parece ser la consecuencia de que la mutante afecte la interacción con los iones que se asocian a la TPP. De acuerdo a las estructuras la base G60 está coordinada por su grupo O^6 carbonil con el Mg1 que a su vez interacciona con el fosfato β del motivo PP (Fig. 25). En este caso puede ser que el cambio de una purina (guanina) por una pirimidina (citosina) que es una molécula más pequeña tenga como consecuencia afectar las interacciones con el Mg^{2+} y esto evite la unión de la TPP y por tanto la represión, una alternativa para explorar la hipótesis de que el reconocimiento está directamente asociado al tamaño de la molécula (cambio de purina a pirimidina) sería mutar la base G60 por adenina, para conservar la condición de purina y probar en ésta la represión. En conclusión la base guanina es importante para hacer contactos con el Mg1 y para el reconocimiento del motivo PP.

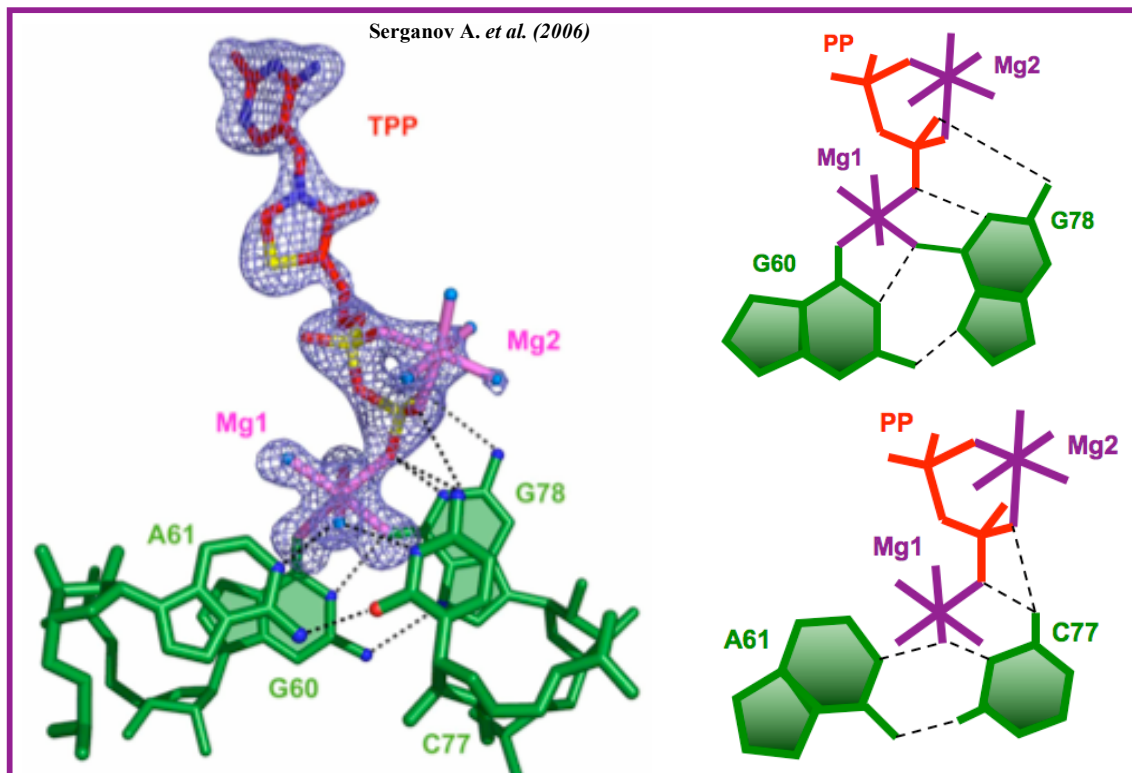
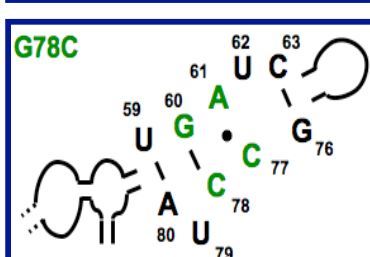
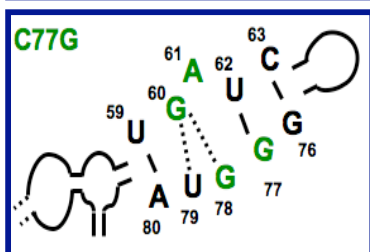
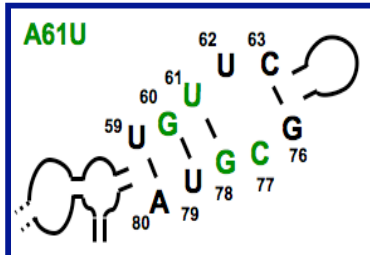
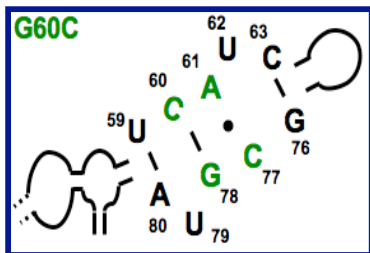
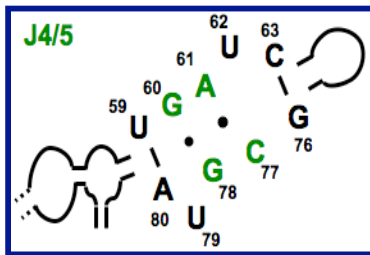
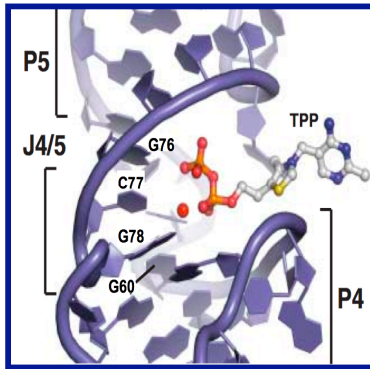


Fig. 25. Interacciones entre el PP de la TPP, el Mg1 y Mg2 y las bases G60, A61, C77 Y G78. Las aguas que coordinan a los iones Mg se muestran como esferas azules en los iones Mg.

⁵ Los apareamientos no canónicos son enlaces distintos a las interacciones “Watson-Crick”



Mutante A61C Cuando se analizó la mutante A61U, se observó que se pierde la actividad tanto en presencia como en ausencia de tiamina, esto nos sugiere que el fenotipo es independiente de la tiamina y más bien es consecuencia de la modificación en la estructura. En el RNA los apareamientos entre guanina y uracilo (G-U “wobble”) son comunes como los apareamientos canónicas A-U, G-C [¡Error!No se encuentra el origen de la referencia.] y en el caso de la mutante A61U el cambio evita el enlace con la C77, pero puede permitir un apareamiento entre A61U y G78 (U-G), lo cual a su vez puede evitar el apareamiento de G60•G78 y favorecer un apareamiento entre la G60 y U79 (G-U), éstos cambios implican dos enlaces “Watson-Crick” en las bases del asa J4/5, que refuerzan el tallo de la “thi-box” (Fig. 26), lo que puede provocar que se establezca la conformación represora. Por lo tanto la hipótesis es que la mutante estabiliza la estructura de represión y por tanto los genes se reprimen independientemente de si hay tiamina o no. Por otro lado la base A61 está coordinada a través de una molécula de agua con el Mg1 que está interaccionando con el fosfato β del PP, por lo que la mutante afecta a esta interacción.

Mutante C77G La mutante C77G también tiene un fenotipo de pérdida de actividad. En este caso se afecta el enlace A61•C77 [¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.] y el cambio puede reforzar el tallo de la “thi-box” favoreciendo un enlace entre la mutante C77G y la U62, así mismo la base G60 puede aparearse con la G78 como está reportado para la estructura o bien con la base U79 (Fig. 26). Si el tallo de la “thi-box” está reforzado entonces la estructura de represión puede ser más estable y como en el caso de la mutante A61G los genes se reprimen independientemente de si hay tiamina o no. Por otro lado la mutante C77G afecta las interacciones que tiene la base C77 con los iones Mg y con la TPP (Fig. 26). Según los datos reportados por Serganov *et al.* (2006) [¡Error!No se encuentra el origen de la referencia.], la base C77 tiene un contacto

Fig. 26. Posibles cambios en las mutantes del asa J4/5. En la imagen se muestra el asa J4/5 y se esquematizan los enlaces no canónicos G60•G78 y A61•C77 del riboswitch y las mutantes que evitan éstos enlaces, pero que pueden favorecer otros enlaces entre las bases que pueden reforzar la estructura y por consecuencia el tallo de la “thi-box”.

⁶ En la estructura reportada por Thore *et al.* (2006) en el riboswitch de *trpE* de *Escherichia coli* la base A61 no se aparea con la C77.

coordinado por H₂O con el Mg1 y tiene contactos mediante enlaces de hidrógeno con los oxígenos del fosfato β de la TPP. Por consiguiente la mutante afecta el contacto con el Mg1 y con la TPP.

Mutante G78C En el caso de la mutante G78C que también pierde la actividad en presencia y ausencia de tiamina, el cambio evita el enlace G60•G78 y favorecer un enlace Watson-Crick G60-C78 que es más fuerte (Fig. 26) por lo que se propone que como en el caso de las bases A61 y C77 la mutante estabiliza la conformación represora. Por otro lado la mutante está afectando las interacciones de la base con el Mg1 y con la TPP, ya que de acuerdo a la estructura reportada por Serganov *et al.* (2006) [**¡Error!No se encuentra el origen de la referencia.**] la base G78 está coordinada al Mg1 por la interacción Mg-O⁶ carbonil de la base G78 y está interaccionando con la TPP por enlaces de hidrógeno con los átomos de oxígeno del fosfato β de esta (Fig. 25).

Además de los contactos que se dan por el reconocimiento de la TPP, hay otras dos áreas de contacto que ayudan a conectar las dos hélices sensoras del riboswitch. Una que corresponde a la región L5-P3 y otra que corresponde a la región J2/4. En la primera, algunas de las bases del asa L5 hacen contacto por interacciones Watson-Crick con el tallo P3 y en la segunda según Thore *et al.* (2006) [**¡Error!No se encuentra el origen de la referencia.**] la interacción de la base A84 con las bases G16 y C50 de la hélice sensora de HMP y con la base U54 de la hélice sensora de PP mediante enlaces de hidrógeno, forma una plataforma de interacción, que estabiliza la región J2/4 donde se juntan las dos hélices sensoras y el tallo P1. La base A84 es muy importante dentro de toda la estructura, dado que hace triples contactos justo en la interfase de los dominios y a través de ellos junta las cuatro diferentes cadenas del RNA, funcionando como una bisagra que estabiliza a toda la estructura. Por otro lado según Serganov *et al.* (2006) [**¡Error!No se encuentra el origen de la referencia.**], la región J2/4 se estabiliza por enlaces de hidrogeno y apilamiento entre tétradas de nucleótidos, donde los pares de bases G17-C49 y G16-C50 del P2 forman enlaces con los pares no canónicos A56•G83 y A53•A84 del asa J2/4 respectivamente.

Mutante U68A La mutante U68A tiene un fenotipo de pérdida de la represión, está base pertenece al loop 5 (L5) y está altamente conservada a diferencia de las otras 5 bases que forman el loop, por lo que se propone que tiene una función importante en la estructura del riboswitch, lo contrastante es que en las estructuras reportadas [**¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.**], esta base no esta participando en enlaces con otras bases, a diferencia de las bases A69, A70 y U71 que forman enlaces de hidrógeno con bases en el tallo 3 (P3), mediante estos enlaces se une a la hélice sensora del PP y a la hélice sensora del HMP y se estructura al riboswitch en una conformación que engloba a la TPP.

Del análisis de todas las mutantes es interesante ver que una misma base cumple con más de una función importante en la regulación: por un lado en el reconocimiento molecular de TPP a través de iones bivalentes y por otro lado en el control estructural para evitar que el riboswitch se establezca en una posible estructura de represión, como sucede en las bases del asa J4/5, esto explica la conservación de éstas (en el 99% de los casos están conservadas), es decir que estas bases son prácticamente invariables e indispensables en el mecanismo de regulación del riboswitch de tiamina.

1. Conclusión

Las mutantes analizadas afectan la regulación por TPP en el riboswitch de tiamina. Nosotros proponemos que se afecta a la regulación porque las bases mutadas son importantes en la estructura del riboswitch y en las interacciones con TPP.

De forma específica, se puede concluir:

- ◆ Las bases mutadas del asa J2/3 pierden la represión por tiamina porque afectan el pegado del motivo HMP de la TPP:
 - Las bases G19 y G40 hacen contactos directos con el motivo HMP de la TPP
 - el motivo HMP de la TPP se acomoda entre las bases G42 y A43 mediante interacciones de apilamiento
 - La secuencia conservada que va de la U39 a la A43 forma una estructura llamada “T-loop”, donde las bases U39 y A43 tienen un enlace que cierra la estructura.
 - La mutante C38G, pierde la represión por tiamina, porque se evita la formación del enlace (C38-G21) y esto afecta a la formación del “T-loop”
- ◆ En las 4 bases mutadas del asa J4/5, G60C pierde la represión por tiamina y A61U, C77G y G78C pierden la actividad, estas bases participan en el reconocimiento del motivo PP de la TPP.
- ◆ La base G60 interacciona con el Mg1, por lo que se propone que la mutante pierde esté contacto y se pierde la interacción con TPP
- ◆ Las mutantes A61U, C77G y G78C pierden la actividad en presencia y ausencia de tiamina, lo cual nos sugiere que el fenotipo es independiente de la tiamina y más bien es consecuencia de la modificación en la estructura. Se propone que las mutantes modifican los enlaces en el asa J4/5 reforzando el tallo de la “thi-box” y esto estabiliza la conformación represora del riboswitch.
- ◆ La mutante U68A del L5 pierde la represión y aún no está muy clara su función, pero esta altamente conservada y la mutación tiene efecto sobre la regulación.

Una de las propuestas en el trabajo es utilizar las construcciones de las distintas mutantes para analizar los cambios en la estructura que afectan a la represión, en particular estudiar las mutantes en el asa J4/5 (A61U, C77g y G78c) que no presentan actividad, mediante pruebas estructurales con endonucleasas de restricción para probar o negar la hipótesis de que los cambios refuerzan el tallo de la “thi-box”.

Es interesante ver que no sólo podemos explorara las consecuencias más predecibles de mutar las bases de acuerdo a la estructura, nos estamos dando cuenta de que la plasticidad del RNA es mucha, no sólo lleva a cabo enlaces convencionales (A-U, G-C o G-U) sino que tiene todo un menú más de opciones y a partir de estudios moleculares como éste se puede explorar el mecanismo de regulación.

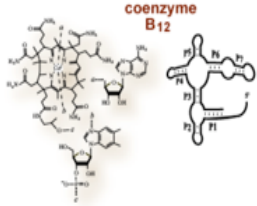
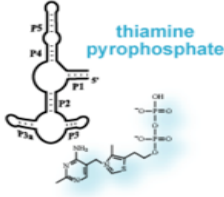
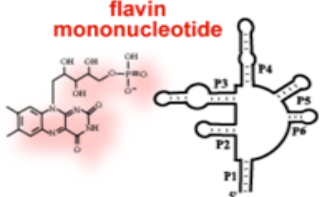
1. Referencias

1. Abreu-Goodger C. *et al.* (2004) Conserved regulatory motifs in bacteria: riboswitches and beyond. *Trends Genet.* 20(10):475-479.
2. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D. (1996) *Biología Molecular de la célula*. 3a. edición. Editorial Omega.
3. Bachean, B. and Neidhardt *et al.* (1996) *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. ASM press pp. 2460-2488.
4. Bailey, T.L. *et al.* (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers Proceedings of the 2nd International Conference on ISMB, AAAI Press, pp. 28–36.
5. Bailey, T.L. *et al.* (1998) Combining evidence using p-values: application to sequence homology searches. *Bioinformatics* 14, 48–54.
6. Barrick, J.E. *et al.* (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. U.S.A.* 101(17):6421-6426.
7. Batey, R.T. *et al.* (2004) Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature.* 432:411-415.
8. Begley, T. P. *et al.* (1999) Thiamin biosynthesis in prokaryotes. *Arch. Microbiol.* 171:293-300.
9. Bresler, *et al.* (1973) Mutations as possible replication errors in bacteria growing under conditions of thiamine deficiency. *Mutant Res.* 19(3):281-93.
10. Corbino, K.E. *et al.* (2005) Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biol.* 6:R70.
11. Cromie, M. *et al.* (2006) An RNA Sensor for Intracellular Mg²⁺. *Cell* 125:71-84.
12. Draper D.E. (2004) A guide to ions and RNA structure. *RNA.* 10:335-343.
13. Edwards T.E. and Ferré-D'Amaré A.R. (2006) Cristal Structures of the Thi-box Riboswitch Bound to Thiamine Pyrophosphate Analogs Reveal Adaptive RNA-Small Molecule Recognition. *Structure.* 14:1459-1468.
14. Ehresmann Ch. *et al.* (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res.* 15(22):9109-9128.
15. Epshtein, V. *et al.* (2003) The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 100(9):5052-5056.
16. Franklind C. V. and Kadner R. J. (1997) Multiple Transcribed Elements Control Expresión of the *Escherichia coli* *btuB* Gene. *J. of Bacterio.* 179:12.
17. Gelfand, M.S. *et al.* (1999) A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Genome Analysis.* 15:439-442.
18. Gilbert, W. and B. Müller-Hill. (1966) Insolation of the Lac Represor. *Proc. Nat. Acad. Sci.* 56:1891-1898.
19. Grundy, F.J. *et al.* (2003) The L box regulon: Lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc. Natl. Acad. Sci. U.S.A.* 100(21):12057-12062.
20. Grundy, F.J. and Henkin T.M. (2003) The T box and S box transcription termination control system. *Frontiers Biosci.* 8:D20-D31.
21. Grundy, F.J. and Henkin, T.M. (2004) Regulation of gene expression by effectors that bind to RNA. *Curr. Opin. Microbiol.* 7:1-6.
22. Hanson, S. *et al.* (2005) Molecular analysis of a synthetic tetracycline-binding riboswitch. *RNA.* 11:503-511.
23. Henkin, T. M. and Yanofski Ch. (2002) Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *BioEssays.* 24:700-707.
24. Huerta A. M. *et al.* (1997) RegulonDB: a database on transcriptional regulation in *Escherichia coli*. *Nucleic Acids Res.* 26(1):55-59.
25. Johansson, J. *et al.* (2002) An RNA thermosensor control expression of virulence genes in *Lysteria Monocystogenes*. *Cell.* 110:551-561.
26. Kadner, R.J. (1978) Repression of síntesis of the vitamin B₁₂ receptor in *Escherichia coli*. *J. Bacterial.* 136:1050-1057.
27. Kanehisa, M. and Goto, S. (2000) KEGG: Kyoto Enciclopedia Genes and Genomes. *Nucleic Acids Res.* 28:27-30.
28. Kubodera, T. *et al.* (2003) Thiamine-regulated gene expresión of *Aspergillus oryzae* *thiA* requires splicing of the intrón containing a riboswith-like domain in the 5'-UTR. *FEBS.* 555:516-520.

29. Lai, E. C. (2003) RNA Sensors and Riboswitches: Self-Regulating Messages. *Current Biology*. 13:R285-R291.
30. Leontis N.B. and Westhof E. (2006) Geometric nomenclature and classification of RNA base pairs. *RNA*. 7:499-512
31. Lim, J. *et al.* (2005) Molecular-Recognition Characteristics of SAM-Binding Riboswitches. *RNA*. 44:1-6.
32. Mandal, M. *et al.* (2003) Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell*. 113:577-586.
33. Mandal, M. and Breaker, R.R. (2004) Adenine Riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. & Mol. Biol.* 11:29-35.
34. Mandal, M. *et al.* (2004) A Glycine-Dependent Riboswitch That Uses Cooperative Binding to Control Gene Expression. *Science*. 306:275-279.
35. Mandal M. and Breaker R.R. (2004) Gene regulation by riboswitches. *Nature Rev.* 5:451-463.
36. Miller, J.H. Experiments in Molecular Genetics (Assay of β -Galactosidase). Cold Spring Harbor Laboratory. 1972. Pp. 352-355.
37. Miranda-Ríos, J. *et al.* (1997) Expression of thiamin biosynthetic genes (*thi*COGE) and production of symbiotic terminal oxidase *cbb* (3) in *Rhizobium etli*. *J. Bacteriol.* 179:6887-6893.
38. Miranda-Ríos, J. *et al.* (2001) A conserved RNA structure (*thi* box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc. Natl. Acad. Sci. USA*. 98: 9736-9741.
39. Mironov, A.S. *et al.* (2002) Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria. *Cell* 111:747-756.
40. Montagne, R.K. and Batey, R.T. (2006) Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature*. 441:1172-1175.
41. Moreno-Hagelsieb, G. and Collado-Vides, J. (2002) A powerful non homology method for the prediction of operons in prokaryotes. *Bioinformatics* 18:329-336.
42. Mulligan, J.T. *et al.* (1982) Mutations affecting regulation of methionine biosynthetic genes isolated by use of met-lac fusions. *J. Bacteriol.* 151:609-619.
43. Murphy, M.B. *et al.* (2003) Transcription termination control of the S box system: Direct measurement of S-adenosylmethionine by the leader RNA. *Proc. Natl. Acad. Sci. U.S.A.* 100(6):3083-3088.
44. Nagaswamy U. and Fox G.E. (2002) Frequent occurrence of the T-loop RNA holding motif in ribosomal RNAs. *RNA* 8:1112-1119.
45. Nahvi, A. *et al.* (2002) Genetic Control by a Metabolite Binding mRNA. *Chem. Biol.* 9:1043-1049.
46. Nahvi, A. *et al.* (2004) Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res.* 1(32):143-150.
47. Nelson, D.L. and Cox, M.M. (2001) *Lehninger Principios de bioquímica*. 3a. Edición. Editorial Omega.
48. Noeske, J. *et al.* (2006) Phosphate-Group Recognition by the Aptamer Domain of the Thiamine Pyrophosphate Sensing Riboswitch. *ChemBioChem*. 00:1-7.
49. Nou, X. and Kadner, R. J. (1998) Coupled Changes in Translation and Transcription during Cobalamin-Dependent Regulation of *btuB* Expression in *Escherichia coli*. *J. Bacteriol.* 180(24): 6719-6728.
50. Nou, X. and Kadner, R. J. (2000) Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proc. Natl. Acad. Sci. U.S.A.* 97:7190-7195.
51. Nudler, E. and Mironov A.S. (2004) The Riboswitch control of bacterial metabolism. *Trends Biochem Sci.* 29(1):11-16.
52. Patte, J.C. *et al.* (1998) The leader of the *Escherichia coli lycC* gene is involved in the regulation of LysC synthesis. *FEMS Microb. Lett.* 169:165-170.
53. Ravnun, S. and Andersson, D.I. (2001) An adenosyl-cobalamin (coenzyme-B12)-repressed translational enhancer in the cob mRNA of *Salmonella typhimurium*. *Mol. Microbiol.* 39(6):1585-1594.
54. Rodionov, D.A. *et al.* (2002) Comparative Genomics of Thiamin biosynthesis in Prokaryotes. *J. Biol. Chem.* 277: 48949-48959.
55. Rodionov, D.A. *et al.* (2003) Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Res.* 31(23):6748-6757.
56. Rosner, A. (1975) Control of lysine biosynthesis in *Bacillus subtilis*: inhibition of diaminopimelate decarboxylase by lysine. *J. Bacteriol.* 12(1):20-8.
57. Salgado H. *et al.* (2006) RegulonDB (version 5.0): *Escherichia coli* K-12 transcriptional regulatory network, operon organization, and growth conditions. *Nucleic Acids Res.* 34:D394-7.

58. Serganov, A. *et al.* (2004) Structural Basis for Discriminative Regulation of Gene Expression by Adenine- and Guanine-Sensing mRNAs. *Chem. & Biol.* 11:1729-1741.
59. Serganov, A. *et al.* (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature*.441:1167-1171.
60. Simons, R.W. *et al.* (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions *Gene* 53:85-96.
61. Soukup, G.A. and Breaker R.R. (1999) Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* 5:1308-1325.
62. Soukup, J. and Soukup G. (2004) Riboswitches exert genetic control through metabolite-induced conformational change. *Curr. Op. in Struc. Biol.* 14:344-349.
63. Sudarsan, N. *et al.* (2003) An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev.* 17(21): 2688-2697.
64. Sudarsan, N. *et al.* (2003) Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA.* 9:644-647.
65. Sudarsan, N. *et al.* (2005) Thiamine Pyrophosphate Riboswitches Are Targets for the Antimicrobial Compound Pyrithiamine. *Chem. & Biol.* 12:1325-1335.
66. Suess, B. *et al.* (2004) A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo. *Nucleic Acids Res.* 32(4):1610-1614.
67. Sussman, D. *et al.* (2000) The structural basis for molecular recognition by the vitamin B₁₂ RNA aptamer. *Nature Struct. Biol.* 7(1):53-57.
68. Tatusov, R.L. *et al.* (1997) A genomic perspective on protein families. *Science* 278, 631–637
69. Thore, S. *et al.* (2006) Structure of the Eukaryotic Thiamine Pyrophosphate Riboswitch with Its Regulatory Ligand. *Science.* 312(5777):1208-1211.
70. Tucker, B.J. and Breaker, R.R. (2005) Riboswitches as versatile gene control elements. *Curr. Op. in Struc. Biol.* 15:342-348.
71. Vellanoweth, R.L. (1993) Translation and its regulation. *Bacillus Subtilis and Other Gram-Positive Bacteria.* (Sonenshein A. L., ed. In chief), pp. 699-711.
72. Vitreschak, A.G. *et al.* (2002) Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* 30:3141-3151.
73. Vitreschak, A.G. *et al.* (2003) Regulation of the vitamin B₁₂ metabolism and transport in bacteria by a conserved RNA structural element. *RNA.* 9:1084-1097.
74. Vitrschak, A.G. *et al.* (2004) Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Gent.* 20(1):44-50.
75. Webb, E. *et al.* (1996) Thiamine pyrophosphate (TPP) negatively regulates transcription of some *thi* genes of *Salmonella typhimurium*. *J. Bacterial.* 178:2533-2538.
76. Westhof, E. and Fritsh, V. (2000) RNA holding: beyond Watson-Crick pairs. *Structure* 8(3): R55-R64.
77. Winkler, W. *et al.* (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature.* 419:952-956.
78. Winkler, W.C. *et al.* (2002) An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. U.S.A.* 99(25):15908-15913.
79. Winkler, W.C. and Breaker R.R. (2003) Genetic Control by Metabolite-Binding Riboswitches. *Chembiochem.* 4:1024-1032.
80. Winkler *et al.* (2004) Control of gene expression by a natural metabolite-responsive ribozyme *Nature.* 428:281-286.
81. Winkler, W.C. (2005) Metabolic monitoring by bacterial mRNAs. *Arch Microb.* 183:151-159
82. Wickiser, J.K. *et al.* (2005) The kinetics of ligand Binding by an Adenine-Sensing Riboswitch. *Biochem.* 44:13404-13414.
83. Woodcock, D.M. *et al.* (1989) *Nucl. Acids Res.* 17:3469-3478
84. Woodson, S. (2005) Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Op. in Chem. Biol.* 9:104-109.
85. Yamauchi, T. *et al.* (2005) Roles of Mg²⁺ in TPP-dependent riboswitch. *FEBS Letters* 579:2583-2588.
86. Yanisch-Perron, *et al.* (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene.* 33:103-119.

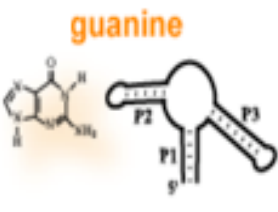
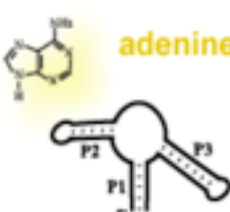
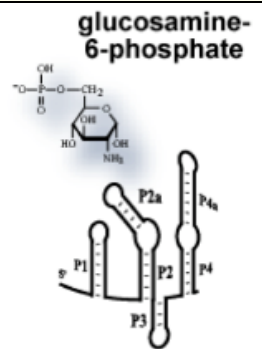
1. Apéndice 1. Los distintos riboswitches

Características de los riboswitches			
Riboswitch	Cobalamina	Tiamina	Riboflavina
Estructura propuesta			
Precursor	Cobalamina (Vitamina B ₁₂)	Tiamina (vitamina B ₁)	Riboflavina (vitamina B ₂)
Molécula efectora	5'- deoxi - 5' - adenosil Cobalamina (Ado-Cbl)	Tiamina pirofosfato (TPP)	Mononucleotido de Flavina (FMN) Flavin Adenin Dinucleotido (FAD)
Genes regulados^I	<i>btuB</i> (transportador de cobalamina) en <i>Escherichia. coli</i> y <i>Salmonella typhimurium</i> operón <i>cob</i> (síntesis de cobalamina) en <i>Salmonella typhimurium</i>	operón <i>thiCOGE</i> (síntesis de tiamina) en <i>Rizobium etli</i> operón <i>tenA</i> (síntesis de tiamina) en <i>Bacillus subtilis</i> <i>thiM</i> (kinasa de tiazol) en <i>Escherichia. coli</i>	operón <i>ribGBAH</i> (síntesis de riboflavina) en <i>Bacillus subtilis</i> <i>ypaA</i> (transportador de riboflavina) en <i>Bacillus subtilis</i>
Distribución taxonómica y regulación	Eubacterias Gram-negativas, Regulación a nivel de la traducción Regulación a nivel traduccional y transcripcional (<i>btuB E. coli</i> y <i>Salmonella</i>) Gram-positivas Regulación a nivel de la transcripción (<i>Bacillus/Clostridium group</i>)	Eubacterias Gram-negativas Regulación a nivel de la traducción Gram-positivas Regulación a nivel de la transcripción Arqueobacterias Eucariotes Plantas y Hongos^{II} Regulación a nivel de "splicing"- procesamiento y estabilidad del RNA	Eubacterias Gram-negativas Regulación a nivel de la traducción Gram-positivas Regulación a nivel de la transcripción Regulación a nivel transcripcional y traduccional (<i>ypaA</i>)

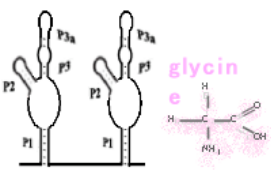
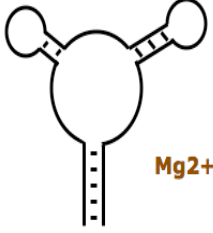
^I Se incluyen solo los primeros genes identificados y sobre los cuales se a realizado más investigación del su riboswitch asociado

^{II} Plantas: *Arabidopsis thaliana* (*thiC*), *Oriza sativa* (*thiC*), *Poa secunda* (posible *thiC*). Hongos: *Aspergillus oryzae* (*thiA*/síntesis de tiazol), *Neurospora crassa* (*nmt-1*), *Fusarium oxyssporum* (*sti35*), *Fusarium solana* (*sti35*)

	(Met)	(Met)	(Lys)
Molécula efectora	S-adenosil-L-Metionina (SAM)	S-adenosil-L-Metionina (SAM)	Lisina (Lys)
Genes regulados	<i>yiiJ</i> (síntesis de metionina) en <i>Bacillus subtilis</i> operón <i>metIC</i>	<i>metA</i> (<i>metX</i> , <i>metE</i> , <i>metK</i>) (síntesis de metionina) en <i>Agrobacterium tumefaciens</i>	<i>lysC</i> (monofuncional aspartocinasa) <i>lysA</i> (deamino descarboxilasa)
Distribución taxonómica y regulación	Eubacterias Gram-negativas Regulación a nivel traduccional Gram-positivas Regulación a nivel transcripcional	Eubacteria Gram-negativas (α -proteobacteria) ---	Eubacteria Gram-negativas Regulación a nivel traduccional Gram-positivas Regulación a nivel transcripcional
K_D Constante de disociación	~4 nM en <i>yiiJ</i> de <i>Bacillus subtilis</i>	~1 μ M en <i>metA</i> de <i>Agrobacterium tumefaciens</i>	-
Referencias	[;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia.]	[;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia.]	[;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia., ;Error!No se encuentra el origen de la referencia.]

-continuación- Características de los Riboswitches			
Riboswitch	Purinas (Guanina)	Purinas (Adenina)	Glucosamina
Estructura propuesta			
Precursor	Purinas (Adenina, Guanina)	Purinas (Adenina, Guanina)	Glucosamina 6- fosfato (GlcN6P)
Molécula efectora	Hypoxantina y Guanina o Adenina ⁶	Hypoxantina y Guanina o Adenina ⁶	Glucosamina-6-fosfato (GlcN6P)
Genes	<i>xpt-pbuX</i>	<i>xpt-pbuX</i>	<i>glmS</i>

regulados	(xanthin fosforibosiltransferasa y transportador de Xanthin) en <i>Bacillus subtilis</i> <i>purEKBCSQLFMNHD</i> (síntesis de Inosine Monofosfato IMP)	(xanthin fosforibosiltransferasa y transportador de Xanthin) en <i>Bacillus subtilis</i> <i>purEKBCSQLFMNHD</i> (síntesis de Inosine Monofosfato IMP)	(glutamina-fructosa-fosfato aminotransferasa) en <i>Bacillus subtilis</i>
Distribución taxonómica y regulación	Eubacterias Gram-positivas Regulación a nivel transcripcional	Eubacterias Gram-positivas Regulación a nivel transcripcional	Eubacterias Gram-positivas Regulación a nivel transcripcional por el rompimiento del mensajero (ribozima)
K_D Constante de disociación	~5 nM en <i>xpt</i> de <i>Bacillus subtilis</i>	~5 nM en <i>xpt</i> de <i>Bacillus subtilis</i>	K _{max} = 3 min ⁻¹ K _{min} =0.003 min ⁻¹ Actividad media = ~200 μM
Referencias	[¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.]	[¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia., ¡Error!No se encuentra el origen de la referencia.]	[¡Error!No se encuentra el origen de la referencia.]

-continuación-		
Características de los Riboswitches		
Riboswitch	Glicina	Magnesio
Estructura propuesta		
Precursor	Glicina	Magnesio
Molécula efectora	Glicina	Mg ²⁺
Genes regulados	operón: <i>gcvT-gctvPA-gctvPB</i> (complejo de rompimiento de glicina) en <i>Bacillus subtilis</i> VC1422 (<i>gcvT</i>) (propuesto transportador de sodio y alanina) de <i>Vibrio cholerae</i>	<i>mgtA</i> (transportador de magnesio) en <i>Sallmonela enterica</i> serovar Typhimurium
Distribución	Eubacterias	Eubacterias

taxonómica y regulación	<p>Gram-positivas Regulación a nivel transcripcional</p> <p>*presenta dos secuencias aptaméricas (I y II) que ejercen un reconocimiento doble y un aumento en la modulación por la molécula efectora</p>	<p>Gram-negativas Regulación a nivel transcripcional</p> <p>*está regulado por dos pasos: el inicio de la transcripción está regulado por la proteína PhoP, controlado por el sensor PhoQ en respuesta a Mg²⁺ extracitoplasmico y por el paro transcripcional controlado por el riboswitch</p>
K_D Constante de disociación	<p>aptámero I = ~30μM aptámero II = ~10μM en VC1422 de <i>Vibrio cholerae</i></p>	<p>--</p>
Referencias	<p>[¡Error!No se encuentra el origen de la referencia.]</p>	<p>[¡Error!No se encuentra el origen de la referencia.]</p>

2. Apéndice 2. Método de selección de motivos

En el trabajo se utilizaron los datos obtenidos en el artículo: “Conserved regulatory motifs in bacteria: Riboswitches and beyond”¹ reportados por: Ceí Abreu-Goodger, Nancy Ontiveros-Palacios, Ricardo Ciria y Enrique Merino. A continuación se hace una breve descripción del método y los resultados obtenidos y un desarrollo más amplio de los programas que se utilizaron (MEME y MAST).

En “Conserved regulatory motifs in bacteria: Riboswitches and beyond” se utilizaron herramientas computacionales para analizar secuencias intergénicas en búsqueda de posibles elementos de regulación conservados. Lo primero que se hizo fue identificar genes relacionados en base a familias de genes ortólogos, utilizando la base de datos de proteínas ortólogas COG (Cluster of Orthologous Groups) (<http://www.ncbi.nlm.nih.gov/COG/>) [Tatusov, R.L. et al. (1997) A genomic perspective on protein families. *Science*. 24:631-637], estos genes se agruparon en operones mediante el método de predicción de operones en base a distancias intergénicas [Salgado, H. et al. (2000) A powerful non-homology method for the prediction of operons in procaryotes. *Bioinformatics*. 18:329-336]. Después a cada operón se le asignó su región intergénica iMUR (intergenical minimal upstream region) que corresponde a la región entre el primer gene del operón y el gene más próximo del siguiente operón. Para evitar la sobre representación de secuencias redundantes pertenecientes a organismos filogenéticamente cercanos, se hizo una comparación de todas las secuencias intergénicas entre si mediante el programa BLAST (Basic Local Alignment Search Tool) [Altschul, S.F. et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res.* 25, 3389-3402] y se eliminó a las secuencias redundantes. Para la búsqueda se utilizaron 164 genomas completamente secuenciados y se obtuvieron ~4000 grupos de regiones reguladoras ortólogas, cada uno perteneciente a un diferente COG.

Después de forma iterativa se utilizó el programa MEME (Multiple EM for Motif Elicitation) [Bailey, T.L. and Elkan, Ch. (1994) "Fitting a mixture model by expectation maximization to discover motifs in biopolymers", *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, AAAI Press, Menlo Park, California, pp. 28-36] para identificar motivos sobre representados en cada grupo de regiones reguladoras ortólogas y los motivos reconocidos se usaron para buscar en el total de secuencias intergénicas utilizando el programa MAST (Motif Alignment and Search Tool) [Bailey, T.L. and Gribskov, M. (1998) "Combining evidence using p-values: application to sequence homology searches". *Bioinformatics*, 14:48-54.] con este proceso se identifican nuevos miembros en cada grupo inicial y se remueven otros, el nuevo grupo se evalúa de la misma forma hasta que ningún miembro nuevo es encontrado. El resultado final es un grupo de secuencias intergénicas que comparten los mismos motivos conservados.

¹ Abreu-Goodger C. *et al.* (2004) Conserved regulatory motifs in bacteria: riboswitches and beyond. *Trends Gent.* 20(10):475-479

-MEME Y MAST-

MEME es un programa que obtiene motivos (secuencias altamente conservadas), de acuerdo a matrices específicas de posición. La matriz calcula la probabilidad que tiene cada posible letra de una ventana de secuencia, en cada posible posición (match), en este caso el valor de un match en un motivo comparado con una secuencia al azar es una variable discreta de la cual se puede calcular su distribución y a partir de esta distribución se puede calcular su valor de probabilidad (p-value). MEME utiliza un valor de proporción de probabilidad logarítmica (log likelihood ratio o LLR) para la ocurrencia de cada motivo y a partir del LLR se estima el valor de expectancia (e-value) del motivo, en un estimado de motivos con el mismo largo y con el mismo número de secuencias.

MAST es una herramienta de búsqueda de secuencias que contienen uno o más motivos en una base de datos dada. MAST está diseñado para buscar “matches” a partir de matrices específicas de posición, de esta forma MEME y MAST son programas complementarios

El programa MEME tiene distintas opciones para dirigir la búsqueda:

- El tipo de alfabeto que va a comparar: secuencias de aminoácidos [-protein] o secuencias nucleotídicas [-dna]
- La distribución de motivos: que el motivo esté por lo menos una vez en cada secuencia [-mod oops], que el motivo este cero o una vez por secuencia [-mod zoops] o que el motivo este cualquier número de veces en una secuencia de forma no superpuesta [-mod tcm]
- El máximo número de motivos a encontrar [-nmotifs]
- Para ajustar los resultados a un máximo valor de expectancia (e-value) [-evt]
- Para ajustar la cantidad de secuencias que presenten el motivo: [nsites], mínimo número de secuencias por motivo [-minsites], máximo número de secuencias por motivo [-maxsites]
- Para ajustar el largo del motivo (ventana):[-w], mínimo largo [-minw], máximo largo [-maxw], entre otros.

El archivo de salida de MEME está compuesto por varias tablas, la primer tabla contiene un resumen de los datos de entrada con que se corrió el programa y del archivo de entrada. En este caso considera:

[mod=zoops], que el motivo seleccionado pueda no estar o esté por lo menos una vez en cada secuencia

[nmotifs=4], busca máximo 4 motivos

[-evt=0.000001], que tengan máximo un valor de expectancia de $1e^{-06}$

[minw=11], [maxw=50], que tenga una longitud mínima de 11pb y máxima de 50pb

[minsites=2], [maxsites=137], que mínimo esté en dos secuencias y máximo en el total de las secuencias, entre otros parámetros. Además nos indica que el archivo de entrada consta de 137 secuencias [N=137] y da los valores de frecuencia a cada base de acuerdo al archivo [frecuencias in data set/background setter frequencies].

Ejemplo de los datos obtenidos para un archivo de salida de MEME Grupo 12 Metabolismo de tiamina

```

*****
COMMAND LINE SUMMARY
*****
This information can also be useful in the event you wish to report a
problem with the MEME software.

command: meme /home/cei/data/Results_iMURs_reduced/COG0351_1-COG0352_1-COG0422_1-
COG0476_1-COG0665_1-COG1992_1-COG2022_1-COG2104_1-COG0352_3-COG0477_5-.faa -dna -
mod zoops -nmotifs 4 -evt 0.000001 -minw 11 -maxw 50 -maxsize 2000000

model: mod=      zoops  nmotifs=    4  evt=      1e-06
object function= E-value of product of p-values
width: minw=    11  maxw=    50  minic=    0.00
width: wg=      11  ws=      1  endgaps=  yes
nsites: minsites= 2  maxsites= 137  wnsites=  0.8
theta: prob=    1  spmap=   uni  spfuzz=  0.5
em:   prior=   dirichlet  b=      0.01  maxiter=  50
       distance= 1e-05
data:  n=      41696  N=      137
strands: +
sample: seed=    0  seqfrac=  1
Letter frequencies in dataset:
A 0.280 C 0.204 G 0.227 T 0.290
Background letter frequencies (from dataset with add-one prior applied):
A 0.280 C 0.204 G 0.227 T 0.289
*****

```

Las siguientes tablas se presentan para cada motivo, en el grupo de secuencias 12:metabolismo de tiamina, se encontraron cuatro motivos conservados. El archivo de salida de MEME, reporta como primer motivo al que tiene mayor LLR y en orden decreciente de LLR el resto de los motivos. En la primer tabla por motivo se da una descripción, que incluye:

```

-----
MOTIF 1      width = 50  sites = 137  llr = 4635  E-value = 7.0e-1526
-----
MOTIF 2      width = 13  sites = 132  llr = 1260  E-value = 1.2e-175
-----
MOTIF 3      width = 11  sites = 90  llr = 929  E-value = 1.8e-111
-----
Motif 3 Description
-----
Simplified      A 3411:::a:1
pos.-specific  C 21::a:::2
probability     G 2489::a:a:4
matrix          T 321::a:::3

bits  2.3  *
      2.1  * * *
      1.8  * * * *
      1.6  * * * * *
Information  1.4  * * * * *
content     1.1  * * * * *
(14.9 bits) 0.9  * * * * *
           0.7  * * * * *
           0.5  * * * * *
           0.2  * * * * *
           0.0  -----

Multilevel
consensus  TAGGCTGAGAG
sequence   AG      T
           GT

```

[width]=longitud del motivo (en número de bases)

[sites]=cantidad de secuencias en las que se reporta

[llr]=proporción de probabilidad logarítmica (log likelihood ratio)

[E-value]=esperanza (en función del LLR)

una matriz de probabilidad simplificada y la secuencia consenso del motivo asociada a un gráfico que refleja el grado de conservación en cada base. En la secuencia consenso también se presentan las otras bases que pueden encontrarse en cada posición.

Por ejemplo el motivo 3 de este grupo está formado por una secuencia de 11pb que se encontró en 98 secuencias con LLR de 929 y un e-value de $-1.8e^{-111}$, su secuencia consenso es: TAGGCTGAGAG, sin embargo en la primer posición también se encontró A o G, en la segunda, G o T, y en la última T. Enseguida de esta tabla se tiene una lista de todas las secuencias que contienen el motivo ordenadas por su valor de probabilidad (p-

value), en ella se menciona en qué posición empieza el motivo y se da la secuencia del motivo y las bases contiguas.

En la siguiente tabla se da un diagrama por posiciones de dónde se encuentra el motivo en la secuencia, por ejemplo, en el gene 23097931 de *O. iheyensis* el motivo está entre las bases 125 y 154 de la secuencia.

Las últimas dos tablas corresponden a la matriz de valor por posición y a la matriz de probabilidad, con las que fue evaluado el motivo.

```
-----
-----
Motif 3 sites sorted by position p-value
-----
-----
Sequence name      Start  P-value      Site
-----
23097931|O_iheyensis      126  1.82e-07  GTACTATGTG TGGGCTGAGAG GAAAGCACCT
46198623|T_thermophilus_  87   5.02e-07  CCCGAATGGA AGGGCTGAGAG CTGGGTTTCT
15827072|M_leprae        68   5.02e-07  CCGAATCTGC GGGGCTGAGAG TGGGCACAGC
15643550|T_maritima      130  5.02e-07  CTCTATTCC  GGGGCTGAGAG GAGGACGGAA
16081564|T_acidophilum   318  5.02e-07  AGTCCATAA  GGGGCTGAGAG GATCCGGATG
16082177|T_acidophilum   220  5.02e-07  GCTCCATAAG GGGGCTGAGAG GATCCGGATG
15896266|C_acetobutylicu 233  7.27e-07  TAAAGACTT  TAGGCTGAGAG GAGAAATCCA
-----
-----
```

```
-----
-----
Motif 3 block diagrams
-----
-----
SEQUENCE NAME      POSITION P-VALUE  MOTIF DIAGRAM
-----
23097931|O_iheyensis      1.8e-07  125_[3]_154
46198623|T_thermophilus_  5e-07    86_[3]_142
15827072|M_leprae        5e-07    67_[3]_79
15643550|T_maritima      5e-07    129_[3]_117
16081564|T_acidophilum   5e-07    317_[3]_72
16082177|T_acidophilum   5e-07    219_[3]_58
15896266|C_acetobutylicu 7.3e-07  232_[3]_157
-----
-----
```

```
-----
-----
Motif 3 position-specific scoring matrix
-----
-----
log-odds matrix: a=length= 4 w= 11 n= 40326 bayes= 8.38342 E= 1.8e-111
  -1   -29   -3   20
  35  -188   65  -31
 -107 -261  178 -238
 -185 -1314  199 -370
 -1314  229 -1314 -1314
 -1314 -1314 -1314  179
 -1314 -1314  214 -1314
  182 -1314 -435 -1314
 -1314 -420  212 -1314
  184 -1314 -1314 -1314
 -165  -29   97   5
-----
-----
```

```
-----
-----
Motif 3 position-specific probability matrix
-----
-----
letter-probability matrix: a=length= 4 w= 11 n= 40326 E= 1.8e-111
0.277778 0.166671 0.222223 0.333328
0.355547 0.055572 0.355541 0.233340
0.133350 0.033352 0.777717 0.055582
0.077800 0.000023 0.899925 0.022252
0.000031 0.999912 0.000025 0.000032
0.000031 0.000023 0.000025 0.999921
0.000031 0.000023 0.999914 0.000032
0.988810 0.000023 0.011135 0.000032
0.000031 0.011133 0.988804 0.000032
0.999920 0.000023 0.000025 0.000032
0.088910 0.166671 0.444420 0.299999
-----
-----
```

La matriz de valores por posición (position-specific scoring matrix) es una matriz log-odds, que se calcula tomando 100 veces el logaritmo (base 2) del radio p/f en cada posición en el motivo donde p es la probabilidad de que una letra particular esté en esa posición en el motivo, y f es la frecuencia de cada base (background frequency). MAST toma como archivo de entrada a la información contenida en esta matriz, para hacer su búsqueda sobre el archivo inicial de secuencias. La matriz de probabilidad (position-specific probability) corresponde a la frecuencia observada de cada posible

letra, en ambas matrices las columnas corresponden a las bases (A, C, G, T) y las filas corresponden a cada posición del motivo.

 Combined block diagrams: non-overlapping sites with p-value < 0.0001

SEQUENCE NAME	COMBINED P-VALUE	MOTIF DIAGRAM
31544422 M_gallisepticum	4.44e-10	210_[4(9.84e-05)]_14_[4(8.01e-05)]_11_[1(1.05e-13)]_4_[4(8.88e-05)]_27
27367795 V_vulnificus_CM	1.38e-17	153_[2(1.59e-06)]_3_[3(1.11e-05)]_1_[1(1.49e-18)]_103
28872090 P_syringae	3.84e-17	145_[2(1.39e-06)]_7_[3(4.26e-06)]_3_[1(7.36e-18)]_171
17936589 A_tumefaciens_C	4.13e-17	51_[2(6.46e-05)]_27_[1(7.36e-18)]_52
42784386 B_cereus_ATCC_1	5.13e-25	77_[2(4.35e-08)]_11_[3(7.27e-07)]_10_[1(3.06e-19)]_45_[4(1.79e-07)]_55
16126272 C_crescentus	2.19e-18	58_[2(5.48e-05)]_6_[3(1.17e-05)]_5_[1(1.69e-20)]_88
15900613 S_pneumoniae_TI	1.86e-18	103_[2(3.87e-05)]_5_[1(3.67e-17)]_61_[4(9.11e-07)]_140
33241220 P_marinus_CCMP1	2.70e-19	57_[2(6.83e-09)]_19_[1(5.22e-17)]_61
19705080 F_nucleatum	2.12e-12	9_[4(5.45e-06)]_125_[3(7.17e-05)]_5_[1(4.96e-14)]_43

Por ultimo se da un resumen de motivos con todas las secuencias que se agruparon, las secuencias están ordenadas de acuerdo a su valor de probabilidad combinado para todos los motivos en la secuencia. Y se da un diagrama de la posición y el orden de los motivos que presenta cada secuencia. Por ejemplo, la secuencia 27367795 de *V. vulnificus* presenta al motivo dos en la posición 210, después a 14 bases presenta al motivo 3 y a una base al motivo 1.

3. Apéndice 3. Tablas de actividad de β -galactosidasa

Para cada experimento se realizaron ensayos por triplicado de cada mutante, los valores que se reportan son el promedio de ensayos y el error estándar indica que los valores obtenidos son semejantes en los ensayos. A su vez (-) representa el fenotipo de pérdida de represión y (s/actividad) el fenotipo de pérdida de actividad. Cómo se observa el fenotipo de represión es reproducible en los experimentos.

Clona	Experimento 1		
	Sin tiamina	Con tiamina	Represión
pRS552	2 ± 0.2	2 ± 0.2	-
pthiM	35 ± 2.4	4 ± 0.4	9
G19C	11 ± 0.2	15 ± 0.5	-
C38G	49 ± 0.9	48 ± 0.2	-
U39A	30 ± 0.7	29 ± 1.0	-
G40C	27 ± 1.0	28 ± 1.1	-
A41U	28 ± 0.3	25 ± 0.3	-
G42C	32 ± 0.7	31 ± 0.8	-
A43U	27 ± 0.3	23 ± 0.3	-
G60C	20 ± 0.4	18 ± 0.3	-
A61U	7 ± 2.2	5 ± 1.1	s/actividad
U68A	46 ± 1.5	47 ± 1.9	-
C77G	4 ± 0.5	3 ± 0.1	s/actividad
G78C	1 ± 0.8	2 ± 2.6	s/actividad
U79A	16 ± 1.0	4 ± 0.8	4

Clona	Experimento 2		
	Sin tiamina	Con tiamina	Represión
pRS552	1 ± 0.4	2 ± 0.3	-
pthiM	15 ± 1.3	2 ± 0.2	8
G19C	26 ± 0.2	25 ± 1.0	-
C38G	36 ± 0.9	33 ± 2.0	-
U39A	28 ± 0.4	23 ± 0.3	-
G40C	25 ± 0.7	21 ± 0.3	-
A41U	22 ± 0.3	19 ± 0.3	-
G42C	24 ± 1.0	21 ± 0.4	-
A43U	28 ± 1.0	22 ± 1.7	-
G60C	20 ± 0.7	17 ± 1.1	-
A61U	2 ± 0.5	1 ± 0.5	s/actividad
U68A	32 ± 0.8	32 ± 1.0	-
C77G	4 ± 0.3	3 ± 0.6	s/actividad
G78C	0 ± 0.2	0 ± 0.2	s/actividad
U79A	9 ± 1.7	3 ± 0.3	3

Clona	Experimento 3		
	Sin tiamina	Con tiamina	Represión
pRS552	2 ± 0.0	1 ± 0.2	-
pthiM	18 ± 0.5	3 ± 0.3	6
G19C	29 ± 0.8	29 ± 0.7	-
C38G	26 ± 4.0	21 ± 2.0	-
U39A	33 ± 1.0	26 ± 0.5	-
G40C	25 ± 0.8	26 ± 0.9	-
A41U	29 ± 0.5	23 ± 0.3	-
G42C	25 ± 4.9	26 ± 0.5	-
A43U	28 ± 1.6	24 ± 0.5	-
G60C	22 ± 1.0	20 ± 0.4	-
A61U	3 ± 1.4	2 ± 0.4	s/actividad
U68A	27 ± 0.3	28 ± 1.1	-
C77G	9 ± 0.1	7 ± 1.1	s/actividad
G78C	1 ± 0.5	2 ± 0.4	s/actividad
U79A	9 ± 0.8	4 ± 0.4	2