



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS  
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD MONTERREY

Pulsos transcripcionales en el operón triptófano de *Escherichia coli* y su efecto en la  
dinámica estocástica del sistema

**Tesis que presenta:**

Emanuel Salazar Cavazos

**para obtener el grado de Maestro en Ciencias**

**en:**

Ingeniería y Física Biomédicas

**Director de tesis:**

Dr. Moisés Santillán Zerón

Apodaca, Nuevo León

Agosto de 2012

*Dedicada a mis padres  
y en memoria a mi tía Rita*

# Agradecimientos

Al Cinvestav y al CONACyT por el apoyo brindado para la realización de esta investigación.

A mis padres por traerme a este mundo y por el incondicional apoyo e inmenso cariño que me han dado a lo largo de toda mi vida.

A mi asesor y gran amigo Moy por la confianza que me tuvo y la libertad que me dio durante el desarrollo de este proyecto. Por sus astutos consejos, por servir como modelo a seguir y más aún por esas frecuentes pláticas llenas de sabiduría y humor.

A la Dra. Blanca Galindo y al Dr. Arturo Chávez por sus atinadas sugerencias para la escritura de esta tesis.

A mis compañeros de cubículo Norma y Luisillo por esas incontables discusiones científicas y por hacer más amena mi estadía en el Cinvestav.

A mis amigos Pato, Juanillo, Mauricio, Diana, Abraham, Flor, Irene y Daniela por todos esos momentos de diversión que compartimos estando en el Cinvestav.

A mis hermanos y en especial a mi sobrinita Sofi por brindarme momentos de alegría y recordarme lo importante que es conservar esa gran capacidad de asombro característica de un niño.

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Dogma central de la biología molecular . . . . .	1
1.2. Transcripción: ¿de forma continua o en pulsos? . . . . .	3
1.3. Operón <i>trp</i> de <i>E. coli</i> . . . . .	4
<b>2. Hipótesis y Objetivos</b>	<b>8</b>
2.1. Hipótesis . . . . .	8
2.2. Objetivo general . . . . .	8
2.3. Objetivos particulares . . . . .	8
<b>3. Metodología</b>	<b>10</b>
3.1. Desarrollo del modelo . . . . .	10
3.2. Variantes del modelo . . . . .	11
3.3. Simulando la des-represión del operón <i>trp</i> . . . . .	13
3.4. Algoritmo para las simulaciones estocásticas . . . . .	14
3.5. Herramientas para el análisis de los resultados . . . . .	15
<b>4. Resultados</b>	<b>18</b>
4.1. Validación del modelo . . . . .	18

4.2. Pulsos transcripcionales . . . . .	19
4.3. Ruido bioquímico . . . . .	21
4.4. Tiempo de respuesta . . . . .	24
4.5. Relación Costo-Beneficio . . . . .	25
<b>5. Discusión y Conclusiones</b>	<b>27</b>
5.1. Discusión . . . . .	27
5.2. Conclusiones . . . . .	29
<b>6. Trabajo a futuro</b>	<b>31</b>
<b>A. Desarrollo del modelo</b>	<b>32</b>
<b>B. Estimación de parámetros</b>	<b>40</b>
<b>Bibliografía</b>	<b>49</b>

# Índice de tablas

3.1. Comparación de la cepa nativa con las cepas hipotéticas equivalentes (CHE)	13
A.1. Lista de reacciones . . . . .	39

# Índice de figuras

1.1. Estructura de un operón . . . . .	2
1.2. Operón triptófano de <i>E. coli</i> . . . . .	5
1.3. Atenuación de la transcripción . . . . .	6
1.4. Represión . . . . .	6
1.5. Inhibición enzimática . . . . .	7
3.1. Representación gráfica de las reacciones consideradas en el modelo matemático	10
4.1. Validación del modelo con los resultados experimentales de Yanofsky y Horn (1994) . . . . .	18
4.2. Presencia de pulsos transcripcionales en células simuladas . . . . .	19
4.3. Caracterización de los pulsos transcripcionales . . . . .	20
4.4. Caracterización del ruido bioquímico en los niveles de ARNm y enzima para las distintas cepas simuladas . . . . .	22
4.5. Caracterización del ruido bioquímico en los niveles de enzima y de triptófano para las distintas cepas simuladas . . . . .	23
4.6. Tiempo de respuesta en la cepa nativa y en las cepas hipotéticas equivalentes (CHE) . . . . .	24
4.7. Relación Costo-Beneficio para la cepa nativa y las cepas hipotéticas equiva- lentes (CHE) . . . . .	26

# Nomenclatura

**ADN**- Ácido desoxirribonucleico

**ARN**- Ácido ribonucleico

**ARNm**- ARN mensajero

**ARNr**- ARN ribosomal

**ARNt**- ARN de transferencia

**CHE**- Cepa(s) hipotética(s) equivalente(s)

**FF**- Factor de Fano

$k_c$ - Constante de cooperatividad

$R_{Tot}$ - Número total de represores

***trp***- Triptófano

***trpE-A***- Genes estructurales del operón *trp*

***trpL***- Región líder del operón *trp*

***trpP***- Región promotora del operón *trp*

$\mu$ - Media

$\sigma$ - Desviación estándar

$\eta$ - Ruido bioquímico

$\tau_R$ - Tiempo de respuesta ante la des-represión



# Resumen

$$\text{Ruido bioquímico Poissoniano } (\eta) = \frac{1}{\sqrt{\text{Media}}} \quad (1)$$

$$\text{Ruido bioquímico "Real" } (\eta_R) = \frac{\text{Desviación estándar}}{\text{Media}} \quad (2)$$

$$\frac{\eta_R}{\eta_P} \quad (3)$$

$$\eta_R \geq \eta_P \quad (4)$$

La manera en que se lleva a cabo la transcripción influye en la variabilidad del número de moléculas de ARN mensajero (ARNm). Recientemente se ha observado que para ciertos genes en diversos organismos la transcripción puede ocurrir en forma de pulsos. Para esta forma de transcripción, en donde varios ARN mensajeros son producidos en cada pulso, se incrementa la magnitud del ruido bioquímico en el número de ARNm. El operón triptófano (*trp*) de *Escherichia coli* contiene los genes que codifican las enzimas requeridas para la síntesis de triptófano, que es un aminoácido necesario para el correcto crecimiento y sobrevivencia de las células. A pesar de que el operón *trp* ha sido ampliamente estudiado, no se ha investigado la posibilidad de que la transcripción en este sistema sea pulsátil. Para ello se creó una representación matemática de los procesos involucrados en la expresión y regulación del operón *trp*, y se llevaron a cabo simulaciones estocásticas de este sistema. Uno de los mecanismos de regulación en el operón *trp* es la represión. En ella destaca la presencia de una interacción

entre represores unidos a sitios adyacentes en el promotor. Esta interacción provoca que los represores permanezcan un mayor tiempo unidos al promotor. Los resultados indican que esta interacción hace posible la presencia de pulsos transcripcionales. Además, observamos que a pesar de que los pulsos transcripcionales en el sistema aumentan de manera importante el ruido bioquímico en los niveles de ARNm, este incremento no se observa para los niveles de enzima y triptófano. Para estudiar las implicaciones biológicas de los pulsos transcripcionales se creó una serie de cepas virtuales con diferentes valores para la fuerza de la interacción entre represores, variando a la vez el número de represores con el fin de mantener el mismo nivel de represión que en la cepa nativa. Después de comparar los resultados del comportamiento dinámico en las distintas cepas llegamos a la conclusión de que la fuerza de esta interacción en la cepa nativa representa una solución óptima, en términos del costo asociado a la producción de represores, así como de las ventajas inherentes a la disminución de los tiempos de respuesta (después de una ausencia repentina de triptófano extracelular) y de la intensidad en el ruido bioquímico del triptófano.

# Abstract

The way transcription is carried out influences the variability in the number of messenger RNA molecules (mRNAs). Recently, it has been shown that transcription of some genes can occur in pulses, in a variety of organisms. For this kind of transcription, where several mRNAs are produced in each pulse, the magnitude in the fluctuations of mRNA number is increased. The tryptophan (*trp*) operon of *Escherichia coli* contains the genes that codify for the enzymes required for tryptophan synthesis, which is an amino acid necessary for proper cell growth and survival. Although the *trp* operon has been extensively studied, the possibility that transcription in this system is pulsatile has not been explored. For that purpose, we created a mathematical representation of the processes involved in the expression and regulation of the *trp* operon, and stochastic simulations of this system were executed. One of the regulatory mechanisms of the *trp* operon is repression. In this mechanism, the presence of an interaction between repressors bound to adjacent sites in the promoter stands out. This interaction makes repressors remain bound to the promoter for a longer time. Our results suggest that this interaction makes possible the presence of transcriptional pulses. Furthermore, we observed that even though transcriptional pulses importantly increase the biochemical noise in mRNA levels, this increase is not seen for enzyme and tryptophan levels. To study the biological implications of transcriptional pulses we created a set of virtual strains with different values for the strength of the interaction between repressors, and at the same time varying the number of repressors with the aim of maintaining the same repression level as in the wild type strain. After comparing the results of the system dynamical behavior in the different strains we came to the conclusion that the strength of this interaction in the wild type strain represents an optimal solution, in terms of the cost associated to the production of repressors, as well as of the advantages concomitant to decreasing the response times (after a sudden absence of extracellular tryptophan) and the tryptophan biochemical noise intensity.

# Capítulo 1

## Introducción

### 1.1. Dogma central de la biología molecular

La perpetuación y evolución de los seres vivos es posible gracias a la capacidad que poseen los organismos de transferir a sus descendientes la información genética que hizo posible que éstos sobrevivieran. Por lo tanto, esta información debe ser almacenada en una molécula que sea suficientemente estable como para mantener su integridad, esta molécula es el ácido desoxirribonucleico (ADN).

El ADN consiste de dos cadenas polinucleotídicas antiparalelas que forman una estructura de doble hélice. Cada nucleótido consiste de un grupo fosfato por el cual se une un nucleótido con otro, una molécula del azúcar desoxirribosa y una base nitrogenada orientada hacia el interior de la doble hélice. Gracias a que en el ADN se pueden encontrar presentes cuatro distintas bases nitrogenadas (adenina, timina, guanina y citosina) es posible codificar la información genética en la secuencia en la que se encuentran estas bases. Cuando las bases adenina y timina se encuentran una enfrente de la otra se forman dos puentes de hidrógeno, mientras que cuando guanina y citosina lo hacen se forman tres. A este hecho se le llama complementariedad de bases.

El elemento básico hereditario es llamado gen. Los genes son secuencias de ADN a partir de las cuales se producen moléculas que llevarán a cabo una función en el organismo, pudiendo ser éstas moléculas de ácido ribonucleico (ARN) o proteínas. El proceso de transcripción es la producción de una molécula de ARN a partir de una de ADN, transfiriendo la información genética con ayuda de la complementariedad existente entre las bases nitrogenadas. Existen

varios tipos de ARN, pero los más ampliamente estudiados son el ARN mensajero (ARNm), el ARN de transferencia (ARNt) y el ARN ribosomal (ARNr). De especial interés para nuestra investigación es el ARN mensajero, debido a que la información almacenada en este tipo de molécula es utilizada para la fabricación de proteínas. El proceso por el cual se sintetizan las proteínas recibe el nombre de traducción y es llevado a cabo por complejos proteína-ARNr (ribosomas) y por los ARNt. Los ARNt son los encargados de transportar los aminoácidos hacia los ribosomas. En procariontes el proceso de traducción generalmente se lleva a cabo al mismo tiempo que la transcripción.

Un gen cuenta con dos regiones principales: la región reguladora y la región codificante. Varios genes pueden compartir una misma región reguladora conformando los llamados operones. A partir de un operón se produce un ARNm que contiene la información de los diferentes genes. Dentro de la región reguladora se encuentra el promotor y si se trata de un gen regulado, este contendrá en su interior operadores. El promotor permite que la ARN polimerasa se una al ADN y comience la transcripción, mientras que los operadores son sitios de unión para represores. Si un represor se encuentra unido al operador, la ARN polimerasa no se podrá unir al promotor y por lo tanto la transcripción será impedida (Fig. 1.1). Generalmente las diferentes proteínas producidas por un operón se encuentran involucradas en una misma función celular, como es el caso del sistema del operón triptófano de *Escherichia coli*, el cual se explicará a detalle más adelante.



**Figura 1.1: Estructura de un operón.** La flecha delgada indica el sitio de inicio de la transcripción, los rectángulos corresponden a los operadores (O1, O2 y O3) y las flechas en diferentes tonos indican las regiones codificantes de los distintos genes que conforman al operón.

## 1.2. Transcripción: ¿de forma continua o en pulsos?

Anteriormente sólo era posible observar la actividad transcripcional promedio en una población celular. Por lo tanto se consideraba que la transcripción en todos los genes ocurría de manera continua con una tasa promedio de producción de mensajeros, tal como ocurre en un proceso de Poisson. Si la transcripción se lleva a cabo de esta manera el número de mensajeros de un gen en la población celular debería seguir una distribución de Poisson.

En el trabajo de Peccoud y Ycart (1995) se propuso un modelo que consideraba una transición lenta entre estados activo e inactivo del promotor, pero no fue hasta años después que Golding et al. (2005) pudieron observar de manera directa que la generación de ARN mensajeros de un gen bajo un promotor regulado en la bacteria *E. coli* se llevaba a cabo en forma de pulsos. Los pulsos transcripcionales son los periodos en donde se lleva a cabo la transcripción de varios mensajeros, y estos pulsos están separados entre sí por periodos de inactividad transcripcional. Éstos y otros descubrimientos fueron posibles gracias al avance científico en las técnicas de biología molecular y de microscopía, las cuales han permitido la cuantificación de números bajos de moléculas en células individuales.

De manera interesante, se ha inferido a través del análisis de distribuciones poblacionales del número de ARNm que la levadura *Saccharomyces cerevisiae* presenta dos tipos de transcripción: como un proceso de Poisson en genes constitutivos y como pulsos transcripcionales en genes regulados (Zenklusen et al., 2008). En eucariotas se ha propuesto que la transcripción pulsátil puede ser causada por el paso de la cromatina entre sus diferentes estados: estado compacto, que impide el acceso a la ARN polimerasa; y estado relajado, que permite que se lleve a cabo la transcripción. Además, se ha propuesto que estos pulsos pueden ser generados por la permanencia del complejo transcripcional sobre el promotor después del inicio de la síntesis de un ARNm (Blake et al., 2003).

De manera general, los pulsos de transcripción pueden ser causados por la unión y separación

lenta de los represores y activadores a las secuencias reguladoras del gen (Golding et al., 2005). Recientemente se han realizado experimentos cuyos resultados indican que la presencia de una especie de cooperatividad entre el represor y sus operadores permite la existencia de pulsos transcripcionales en el operón lactosa de *E. coli* (Choi et al., 2008). En este caso la llamada cooperatividad se debe a que un mismo represor se puede unir simultáneamente a dos operadores, provocando que la velocidad de disociación del represor y sus operadores sea mucho menor que la esperada si el represor sólo se pudiera unir a uno u a otro de los operadores al mismo tiempo.

La transcripción pulsátil sólo ha sido estudiada a fondo en muy pocos sistemas génicos. A pesar de que el operón triptófano de *E. coli* es uno de los sistemas más estudiados, la posibilidad de que su transcripción sea llevada a cabo en forma de pulsos no ha sido explorada. En este trabajo estudiamos la posibilidad de la existencia de la transcripción pulsátil en este operón. Por lo tanto, a continuación se describen las características de este sistema.

### **1.3. Operón *trp* de *E. coli***

El triptófano es un aminoácido presente en la mayoría de las proteínas producidas por la bacteria y es, por lo tanto, necesario para el correcto crecimiento y sobrevivencia de la bacteria. Cuando la cantidad de triptófano obtenida por la célula a partir del medio que la rodea no es suficiente, ésta lo sintetiza a partir de compuestos más simples. Los últimos pasos de la síntesis del triptófano son llevados a cabo por una serie de enzimas codificadas por el operón triptófano. La función del operón triptófano es producir las enzimas sintetizadoras de triptófano cuando dicho aminoácido no pueda ser obtenido en cantidades suficientes a partir del medio circundante.

El operón triptófano (Fig. 1.2) contiene cinco genes estructurales: *trpE*, *trpD*, *trpC*, *trpB* y *trpA*, los cuales codifican para polipéptidos que formarán las enzimas responsables de

la producción de triptófano a partir de corismato. La transcripción del operón es iniciada principalmente en el promotor *trpP*, que contiene tres operadores. Además, posee un promotor débil *trpP<sub>2</sub>* que provoca la expresión constitutiva de los últimos tres genes del operón (Jackson y Yanofsky, 1972).

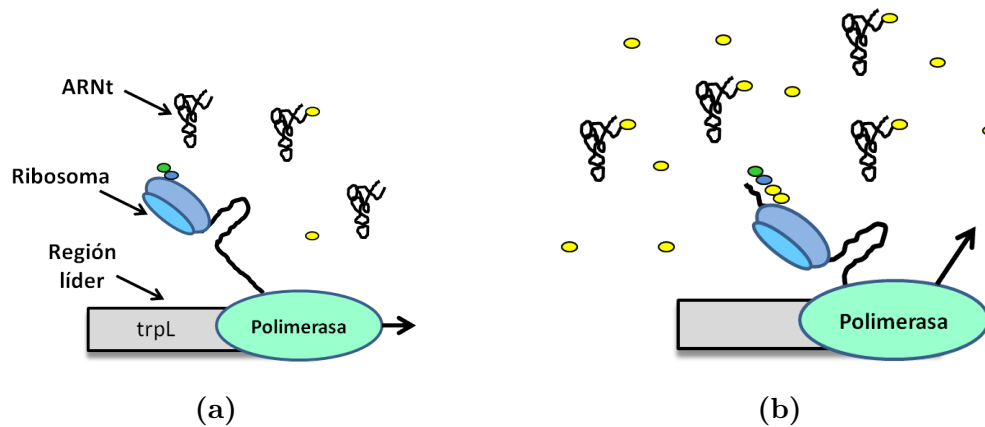


**Figura 1.2: Operón triptófico de *E. coli*.** El operón está compuesto por el promotor *trpP*, sus tres operadores  $O^{(1)}$ ,  $O^{(2)}$  y  $O^{(3)}$ , una región líder *trpL*, cinco genes estructurales *trpE-A* y el promotor interno *trpP<sub>2</sub>*.

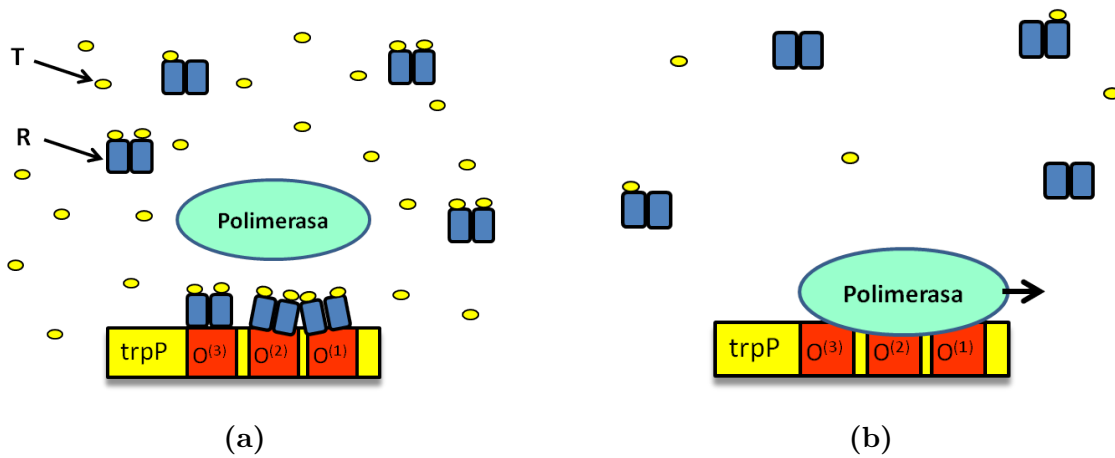
Río arriba a los genes estructurales se encuentra lo que se le conoce como la región líder (*trpL*), la cual codifica para un péptido que contiene dos aminoácidos triptófano consecutivos en su secuencia. Este péptido permite a la célula censar los niveles de triptófano en el citoplasma y regular la transcripción del operón. Este mecanismo de regulación es llamado atenuación y sucede cuando la concentración del ARNt del triptófano cargado (con una molécula de triptófano unida) es alta, lo que permite a los ribosomas avanzar rápidamente sobre el ARNm y formar una estructura en el ARNm que provoca la terminación prematura de la transcripción (Fig. 1.3b). Por el contrario, cuando los niveles intracelulares de triptófano son extremadamente bajos, los ribosomas avanzan lentamente y se facilita la formación de otra estructura en el ARNm (Yanofsky et al., 1984). Esta estructura permite que se continúe normalmente la transcripción del operón, a como es representado en la figura 1.3a.

Existen tres importantes mecanismos de regulación negativa en el operón triptófico de *E. coli*: la represión, la antes mencionada atenuación de la transcripción y la inhibición enzimática. La represión es mediada por la unión de represores a los tres operadores que se encuentran dentro del promotor *trpP*. El represor del operón *trp* es un homodímero y por sí solo se une muy débilmente a los operadores. Cuando a cada una de las dos subunidades se le une una molécula de triptófano este complejo es llamado represor activo y es capaz de unirse con mucha mayor fuerza a los operadores (Fig. 1.4).





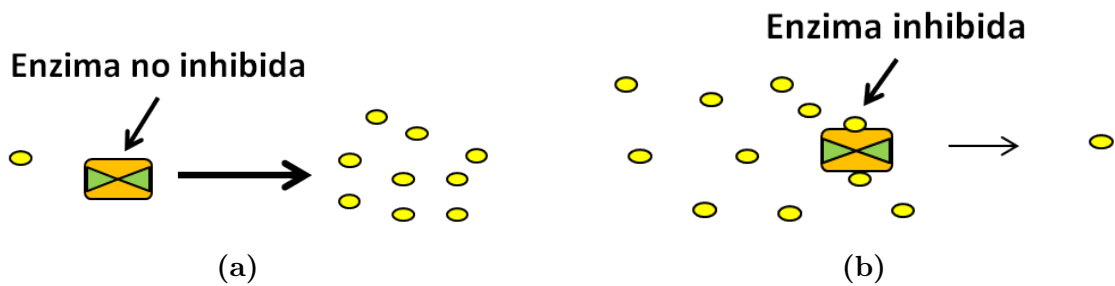
**Figura 1.3: Atenuación de la transcripción.** (a) Cuando las concentraciones de triptófano intracelular son muy bajas, la atenuación será disminuida, y por lo tanto la producción de ARN mensajeros será concluida exitosamente con mayor frecuencia. (b) Si los niveles de triptófano intracelular son altos la atenuación ocurrirá más a menudo.



**Figura 1.4: Represión.** (a) Al haber mucho triptófano intracelular la mayoría de los represores tendrán unidas dos moléculas de triptófano (represores activos) y serán capaces de asociarse a los operadores e impedir la unión de la ARN polimerasa. De especial importancia es la cooperatividad presente entre represores una vez que estos se han unido a los operadores  $O^{(1)}$  y  $O^{(2)}$ . (b) Cuando los niveles intracelulares de triptófano son bajos la transcripción iniciará con mayor frecuencia.

Cuando dos represores se encuentran unidos simultáneamente a los dos primeros operadores, estos interaccionan entre sí provocando que la velocidad de disociación de los represores sea más lenta, o en otras palabras, que la estabilidad del complejo operador-represor aumente. Este fenómeno es referido como cooperatividad, debido a que los represores cooperan entre ellos para aumentar el tiempo que permanecerán unidos a sus respectivos operadores (Fig. 1.4a).

Los genes *trpE* y *trpD* codifican para la enzima antranilato sintasa. Esta enzima tiene una estructura tetramérica formada por dos subunidades del polipéptido TrpE y dos del TrpD (Zalkin, 1973). Esta enzima cataliza la primera reacción de la ruta de biosíntesis del triptófano y es regulada negativamente por la unión de una molécula de triptófano a cada una de las subunidades TrpE de la enzima antranilato sintasa (Fig. 1.5).



**Figura 1.5: Inhibición enzimática.** (a) En ausencia o escasez de triptófano la enzima antranilato sintasa no se encuentra inhibida. (b) En abundancia de triptófano la enzima se encuentra inhibida.

# Capítulo 2

## Hipótesis y Objetivos

### 2.1. Hipótesis

Las interacciones entre los represores y sus sitios de unión en el operón triptófano de *Escherichia coli* son capaces de originar pulsos transcripcionales, afectando de manera importante la dinámica del sistema.

### 2.2. Objetivo general

Predecir la existencia o ausencia de pulsos transcripcionales en el operón triptófano de *Escherichia coli* y estudiar sus implicaciones biológicas.

### 2.3. Objetivos particulares

- 1.- Analizar si las interacciones entre represores y operadores pueden dar lugar a pulsos transcripcionales en el operón *trp* de *E. coli*.
- 2.- Investigar cuáles son los efectos dinámicos de la transcripción pulsátil si ésta existe en el sistema.
- 3.- Estudiar la manera en que los lazos de retroalimentación de represión, atenuación de la transcripción e inhibición enzimática afectan a los pulsos transcripcionales en el operón *trp*

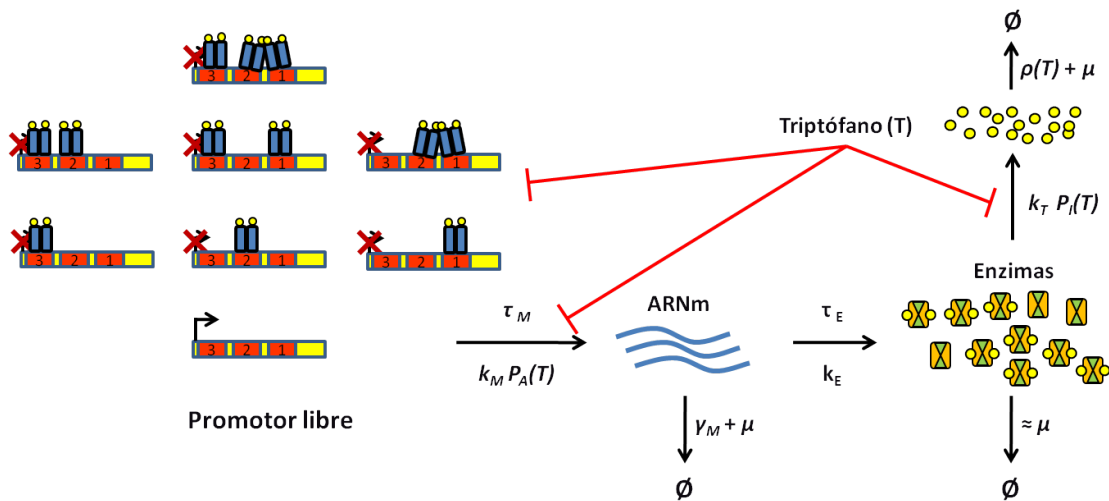
de *E. coli*, así como al comportamiento dinámico del sistema.

# Capítulo 3

## Metodología

### 3.1. Desarrollo del modelo

Se desarrolló un modelo matemático que nos permitió representar al operón *trp* de *E. coli* y sus tres lazos de retroalimentación negativa (Fig. 3.1).



**Figura 3.1:** Representación gráfica de las reacciones consideradas en el modelo matemático. Las líneas rojas indican a los tres lazos de retroalimentación negativo en los el triptófano actúa: la represión, la atenuación de la transcripción y la inhibición enzimática. *trpL*: región líder;  $\tau_M$ : retraso en la transcripción;  $k_M$ : tasa máxima de transcripción;  $P_A(T)$ : probabilidad de que no ocurra atenuación;  $\gamma_M$ : tasa de degradación del ARNm;  $\mu$ : tasa de dilución;  $\emptyset$ : desaparición de la molécula indicada;  $\tau_E$ : retraso en la traducción;  $k_E$ : tasa máxima de producción de enzimas;  $k_T$ : tasa máxima de producción de triptófano;  $P_I(T)$ : probabilidad de que no ocurra inhibición enzimática;  $\rho(T)$ : tasa de consumo de triptófano.

En la parte izquierda de la figura 3.1 podemos observar los 8 posibles estados en los que se puede encontrar el promotor. Existen un total de 24 reacciones que representan las posibles transiciones entre estos estados. Cuando el promotor se encuentra libre de represores la transcripción podrá ser iniciada. El número de ARN mensajeros producidos exitosamente dependerá de la tasa máxima de transcripción ( $k_M$ ) y de la probabilidad de que no ocurra la atenuación ( $P_A(T)$ ).

Debido a que la ARN polimerasa avanza más lento que el ribosoma, debe pasar cierto tiempo antes que la traducción pueda ser iniciada ( $\tau_M$ ). La disminución en el número de moléculas de ARNm depende de la tasa de degradación del ARN mensajero ( $\gamma_M$ ) y de la tasa de dilución debida al crecimiento celular ( $\mu$ ).

La traducción se llevará a cabo con una tasa máxima  $k_E$  y con un retraso entre el inicio y conclusión de la fabricación de la enzima  $\tau_E$ . La pérdida de enzimas estará dada principalmente por la dilución ( $\mu$ ). La degradación de enzimas es despreciada en el modelo debido a que ésta es llevada a cabo en días y su consideración no modifica notoriamente la dinámica del sistema. La producción de triptófano por cada una de estas enzimas producidas dependerá de la tasa máxima con la que la enzima antranilato sintasa produce dicho aminoácido ( $k_T$ ) y de la probabilidad de que ésta no se encuentre inhibida por el triptófano ( $P_I(T)$ ). Por otro lado, la disminución en el número de moléculas de triptófano está determinado por el consumo de triptófano para la fabricación de proteínas ( $\rho(T)$ ) y por la dilución ( $\mu$ ).

El desarrollo del modelo es presentado a detalle en el Apéndice A (pág. 32).

## 3.2. Variantes del modelo

Para explorar el papel que juegan los diferentes lazos de retroalimentación en la dinámica del sistema, se realizaron diversas modificaciones en el modelo original para simular el comportamiento de 3 distintas cepas genéticamente modificadas de *E. coli*. Cada una de estas cepas

presenta una modificación que provoca la eliminación de uno de los lazos de retroalimentación con los que cuenta el sistema del operón *trp*.

Una de estas cepas no cuenta con el mecanismo de represión. Para lograr esto experimentalmente se puede mutar o remover el promotor que regula la fabricación del represor, impidiendo así la producción del mismo. Por lo tanto, para simular esta cepa se redujo el número total de represores ( $R_{Tot}$ ) de 400 a 0.

Para eliminar experimentalmente el segundo lazo de retroalimentación, la atenuación de la transcripción, uno podría eliminar del genoma de la bacteria el segmento de ADN que codifica para el péptido líder. Al eliminar este segmento, se evita la posibilidad de que al haber suficiente triptófano intracelular, se forme una estructura en el ARN que provoque la terminación prematura de la transcripción. Para simular la ausencia de atenuación en nuestro modelo, sólo es necesario darle un valor constante de 1 a la función que determina la probabilidad de que no ocurra atenuación ( $P_A$ ).

La inhibición enzimática puede ser disminuida casi por completo al mutar los aminoácidos de la enzima a los cuales se unen las moléculas de triptófano. El equivalente de esta cepa en el modelo puede ser creado asignándole a la función de la probabilidad de que no ocurra inhibición enzimática ( $P_I$ ) un valor constante de 1.

Posteriormente, con el fin de explorar el efecto que tiene la cooperatividad sobre la dinámica del sistema del operón *trp*, se creó una cepa hipotética en la cual no existe cooperatividad entre los represores contiguamente unidos al promotor. Para simular esta modificación basta con asignar a la constante de cooperatividad ( $k_c$ ) un valor de 1.

Con la intención de estudiar las razones evolutivas de esta cooperatividad, se crearon una serie de cepas hipotéticas con distintas combinaciones en los valores de  $k_c$  y del número total de represores (ver Tabla 3.1). Para poder comparar adecuadamente las características de estas cepas, se buscó que los valores de  $k_c$  y  $R_{Tot}$  permitieran mantener el mismo nivel de expresión que posee la cepa nativa, por lo que de ahora en adelante estas cepas serán

llamadas cepas hipotéticas equivalentes (CHE). El nivel de expresión es calculado a partir de la función que determina la probabilidad de que el promotor se encuentre libre,  $P(Pr_{0,0,0})$ , representada dentro de las ecuaciones B.3 en la página 42.

**Tabla 3.1:** Comparación de la cepa nativa con las cepas hipotéticas equivalentes (CHE)

Cepa	Constante de cooperatividad ( $k_c$ )	Número de represores ( $R_{Tot}$ )
CHE 1	792	100
CHE 2	185	200
Cepa nativa	40	400
CHE 3	6.8	800
CHE 4	1	1330

Al aumentar la cooperatividad existente entre los represores la estabilidad del complejo represor-operador aumenta y, por consiguiente, la fuerza de represión también aumenta. Por lo tanto, para mantener el mismo nivel de expresión al realizar este cambio es necesario disminuir  $R_{Tot}$ . De manera contraria, cuando se disminuye la cooperatividad es necesario aumentar  $R_{Tot}$ .

### 3.3. Simulando la des-represión del operón *trp*

Un experimento comúnmente utilizado para el estudio de la regulación del operón triptófano de *E. coli* es el de la des-represión. Este consiste en dos pasos generales: el primero es cultivar a las células en un medio que contenga triptófano y, el segundo, transferirlas a un medio sin triptófano. Cuando existe triptófano en el medio, el operón se encuentra reprimido casi por completo. Al transferir las células al medio sin triptófano, los niveles intracelulares de triptófano disminuyen drásticamente, provocando que los represores se disocien del promotor (des-represión) y comience la producción masiva de las enzimas codificadas por el operón.

El hecho de que el nombre del experimento no mencione una des-atenuación, se debe a que



la atenuación en la transcripción dejará de ocurrir sólo en los primeros minutos después de la des-represión, cuando el número intracelular de moléculas de triptófano sea cercano a cero. Una vez que la producción de las enzimas sintetizadoras de triptófano comienza, la atenuación vuelve a ocurrir normalmente.

Para simular el primer paso general de la des-represión, los valores iniciales del número de enzimas ( $E_0$ ) y triptófano ( $T_0$ ) de nuestro sistema toman los valores cuantificados experimentalmente en células cultivadas en un medio con triptófano. El segundo paso se logra simular al no incluir en el modelo una entrada de triptófano a partir del medio, lo cual es equivalente a tener un medio sin triptófano.

### 3.4. Algoritmo para las simulaciones estocásticas

Dado que en el modelo fueron considerados ciertos procesos en los cuales las tasas de reacción eran tan altas que provocarían un gran aumento en el costo computacional al realizar las simulaciones, los procesos fueron divididos en procesos rápidos y procesos lentos, basándonos en la estrategia propuesta por Haseltine y Rawlings (2002). Los procesos rápidos tienen velocidades de reacción suficientemente rápidas para permitirnos realizar una aproximación adiabática y suponer que la reacción alcanza instantáneamente un estado de equilibrio. Por otro lado, en los procesos lentos la velocidad de reacción no nos permite hacer esta aproximación y por lo tanto estos procesos son tomados en cuenta directamente por el algoritmo de simulaciones estocásticas.

Para realizar las simulaciones estocásticas se utilizó el algoritmo diseñado por Cai (2007), el cual es un método exacto que permite la incorporación de retrasos entre el inicio de la síntesis de una molécula y la terminación en su fabricación, como es el caso de la producción de ARNm y proteínas. Una descripción detallada de este algoritmo puede ser encontrada en el artículo de Cai (2007). Este algoritmo fue implementado utilizando el lenguaje de programación *Python*. Empleando este algoritmo se simuló el experimento de des-represión

para una población de mil células por cada una de las cepas.

### **3.5. Herramientas para el análisis de los resultados**

Los algoritmos que se describen de forma resumida en esta sección fueron implementados en el programa MATLAB, mientras que las representaciones gráficas de las dinámicas y de la comparación de datos fueron creadas con ayuda de los programas MATLAB y SigmaPlot.

Para permitir el análisis posterior de los datos se crearon dos algoritmos, uno para calcular el comportamiento promedio del sistema y el otro para obtener la distribución en los niveles de las especies moleculares en la población celular simulada. El primero de ellos toma los valores de los niveles moleculares de todas las simulaciones de una cepa, los suma y posteriormente los divide entre el número de simulaciones, obteniendo así el comportamiento promedio de cada cepa simulada. Este promedio fue utilizado posteriormente para la validación del modelo y para el cálculo de los tiempos de respuesta.

El segundo algoritmo también toma los niveles moleculares de todas las simulaciones, pero a diferencia del algoritmo anterior, sólo toma los valores a un determinado tiempo de simulación. Estos valores son almacenados en un archivo, logrando así que cada archivo contenga la distribución de cada una de las cepas simuladas. Utilizando las funciones predeterminadas de MATLAB podemos obtener los parámetros de estas distribuciones, como por ejemplo su media ( $\mu$ ) y su desviación estándar ( $\sigma$ ), valores que son utilizados para la caracterización de los pulsos transcripcionales y la variabilidad en el sistema.

Para la construcción de todas las distribuciones se tomaron los valores al minuto 500 de simulación, el cual es tiempo suficiente para que la dinámica promedio del sistema en todas las cepas simuladas se encuentre en estado estacionario.

Antes de poder confiar en los resultados obtenidos a partir del modelo construido, este último debe ser validado. Para lograr esto, el comportamiento dinámico del sistema simulado debe

ser similar al comportamiento observado experimentalmente. En nuestro caso, múltiples validaciones fallidas nos sirvieron para guiar la incorporación de aquellos aspectos del sistema que anteriormente no habían sido considerados en el modelo.

Para caracterizar los pulsos transcripcionales utilizamos el factor de Fano (FF), el cual es aproximadamente igual al número promedio de mensajeros producidos por cada pulso (Raj y van Oudenaarden, 2009) y está dado por:

$$FF = \frac{\sigma_M^2}{\mu_M} \quad (3.1)$$

en donde  $\sigma_M$  y  $\mu_M$  representan respectivamente la desviación estándar y la media de la distribución del número de ARN mensajeros que existe en la población celular simulada.

Es de particular interés conocer la magnitud de la variabilidad en los niveles de las distintas especies moleculares (ARNm, enzima y triptófano) e investigar cómo se transmite esta variabilidad desde la transcripción hasta la producción de triptófano. Además, es importante analizar cuáles son los efectos de los distintos lazos de retroalimentación y de la cooperatividad sobre estas variaciones. Esta variabilidad puede ser cuantificada por el ruido bioquímico ( $\eta$ ), el cual ha sido definido por Elowitz et al. (2002) como:

$$\eta = \frac{\sigma_X}{\mu_X} \quad (3.2)$$

en donde  $\sigma_X$  y  $\mu_X$  son respectivamente la desviación estándar y la media de la distribución del número de moléculas de la especie  $X$  en la población celular.

Dado que el objetivo final del operón *trp* es la producción de triptófano, una manera adecuada de saber qué tan rápido responde el sistema ante el estímulo de la des-represión es analizando el tiempo de respuesta ( $\tau_R$ ) en los niveles de triptófano. Este tiempo es considerado como el tiempo que transcurre desde el inicio del experimento de des-represión, hasta que el conteo promedio de moléculas de triptófano alcanza el 95 % del valor correspondiente en el estado estacionario. Para obtener este valor se diseñó un algoritmo que detecta el tiempo al cual el nivel promedio de triptófano sobrepasa el 95 % del valor en el estado estacionario.

Una forma comúnmente utilizada para realizar un análisis costo-beneficio de la implementación de un cambio en un proceso es utilizando una función del siguiente tipo:

$$\psi = B - C \quad (3.3)$$

en donde  $B$  representa los beneficios que brinda el realizar el cambio y  $C$  son los costos que conlleva su implementación. Entre mayor sea el valor que tome la función  $\psi$ , mayor será la ventaja que proporcionará la implementación del cambio.

Siguiendo este razonamiento para la existencia y fuerza de la cooperatividad en las cepas hipotéticas equivalentes (Ver Tabla 3.1), se formuló la siguiente función:

$$\psi = \frac{1}{\eta^*} + \frac{1}{\tau_R^*} - R_{Tot}^* \quad (3.4)$$

en donde  $\eta$  representa el ruido en los niveles de triptófano,  $\tau_R$  el tiempo de respuesta,  $R_{Tot}$  el número total de represores y \* indica que estos valores se encuentran normalizados debido a la diferencia en la magnitud de estas tres variables. La normalización se llevo a cabo con respecto a los valores de la cepa con el mayor número de represores.

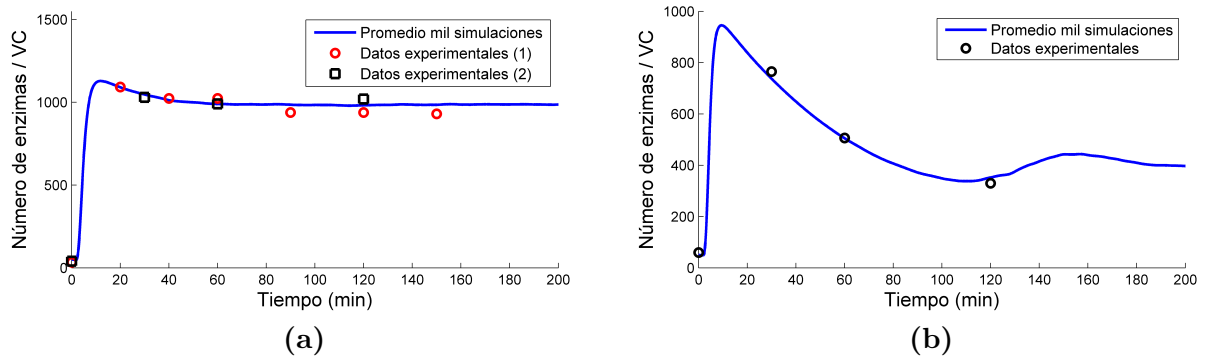
Podemos observar en la ecuación 3.4 que las primeras dos variables son positivas y se encuentran en el denominador. Esto se debe a que la reducción de ambos, el ruido y el tiempo de respuesta, representa un beneficio para la célula. Por el contrario, el tercer término tiene un signo negativo, el cual indica que el incremento en el número de represores provoca un aumento en el costo energético que representa la producción de los mismos.

# Capítulo 4

## Resultados

### 4.1. Validación del modelo

Para confirmar que el modelo desarrollado representa adecuadamente el comportamiento real del sistema, utilizamos los resultados obtenidos por Yanofsky y Horn (1994) al realizar el experimento de des-represión en cultivos celulares de las cepa nativa y de la cepa que carece de inhibición enzimática. En estos experimentos se cuantificó la actividad de la enzima antranilato sintasa mediante colorimetría. Asumiendo que la actividad enzimática es proporcional a los niveles de la enzima, se hizo una comparación entre estos resultados y la dinámica promedio de los niveles de enzima obtenida en las simulaciones de ambas cepas (Fig. 4.1).



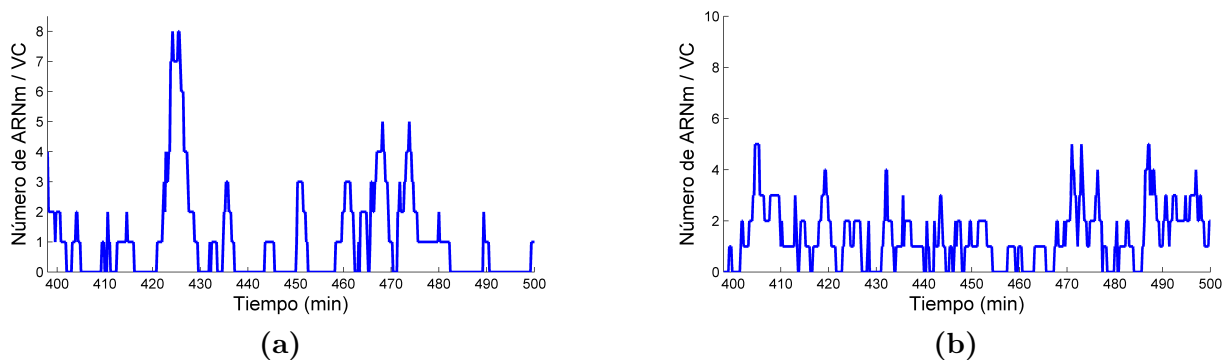
**Figura 4.1: Validación del modelo con los resultados experimentales de Yanofsky y Horn (1994).** Validación del modelo empleando los resultados experimentales obtenidos con (a) la cepa nativa y con (b) la cepa sin inhibición enzimática. VC: Volumen Celular; min: minutos.

En estas gráficas podemos observar que para ambas cepas las simulaciones reproducen correctamente el comportamiento observado experimentalmente. Para la dinámica experimental en la cepa sin inhibición enzimática no existen suficientes mediciones temporales como para realizar una comparación más detallada con los resultados del modelo. Es importante prestar atención al sobre-impulso que ocurre en la producción de enzima después de la des-represión. Este sobre-impulso permite a la célula un incremento más rápido en los niveles de triptófano

necesarios para un crecimiento óptimo (Bhartiya et al., 2006).

## 4.2. Pulsos transcripcionales

Al analizar la dinámica del número de ARNm en la simulación estocástica de una célula de la cepa nativa podemos observar que la transcripción se da en forma de pulsos, con sus característicos periodos de actividad transcripcional separados por periodos en donde no hay producción de ARN mensajeros (Fig. 4.2a). La actividad e inactividad transcripcional pueden ser identificadas en la dinámica por el aumento y disminución, respectivamente, del número de ARNm en la célula.



**Figura 4.2: Presencia de pulsos transcripcionales en células simuladas.** Dinámica en el número de ARNm de una célula simulada de (a) la cepa nativa y de (b) la cepa hipotética equivalente sin cooperatividad. VC: Volumen Celular; min: minutos; CHE: Cepa Hipotética Equivalente.

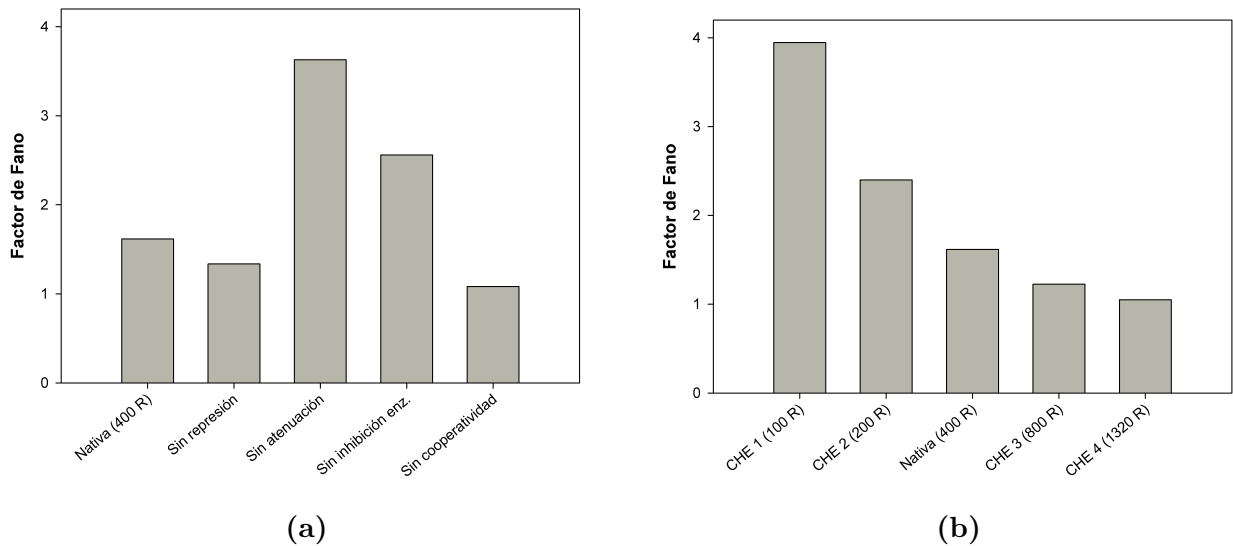
Al comparar estos resultados con los obtenidos a partir de la cepa hipotética equivalente que no cuenta con cooperatividad, observamos una diferencia cualitativa en la distribución de los eventos transcripcionales a lo largo del tiempo (Fig. 4.2). En esta cepa la transcripción se lleva a cabo de una manera más uniforme, a pesar de que sus niveles de expresión son semejantes a los de la cepa nativa.

Para estudiar de manera cuantitativa la forma en que se lleva a cabo la transcripción utilizamos como herramienta el factor de Fano (FF), el cual es aproximadamente igual al número de ARNm producidos por cada pulso transcripcional (Raj y van Oudenaarden, 2009). Por

consiguiente, si el FF tiene un valor cercano a 1 esto nos indica que la transcripción se está llevando a cabo básicamente como un proceso de Poisson, o dicho de otra manera, sin pulsos transcripcionales. El FF está dado por:

$$FF = \frac{\sigma_M^2}{\mu_M} \quad (4.1)$$

en donde  $\sigma_M$  es la desviación estándar y  $\mu_M$  la media de la distribución del número de ARN mensajeros en la población celular simulada. Esta distribución fue construida con los valores de las células cuando el sistema ha alcanzado un estado estacionario.



**Figura 4.3: Caracterización de los pulsos transcripcionales.** Comparación del número promedio aproximado de ARN mensajeros producidos por pulso para la cepa nativa, **(a)** las cepas con los lazos de retroalimentación modificados y **(b)** las cepas hipotéticas equivalentes. R: Represores.

El valor del FF para la cepa nativa es de alrededor 1.6. De las tres cepas que carecen de algún lazo de retroalimentación sólo en la cepa sin represión ocurre una disminución en el comportamiento de transcripción pulsátil, indicando la importancia de esta para la generación de los pulsos transcripcionales (Fig. 4.3a). Por otro lado, la cepa sin atenuación tuvo un aumento en el FF, explicado por el hecho de que al no haber terminación prematura en la transcripción, la producción de ARNm en cada pulso se incrementa. A la vez, al incrementar el número de ARNm producidos se refleja en un aumento en los niveles de triptófano, lo

cual provoca un aumento en la estabilidad de los represores y por lo tanto un incremento del comportamiento pulsátil. De la misma manera, la cepa que carece de inhibición enzimática tendrá niveles muy elevados de triptófano provocando el subsecuente incremento en el valor del FF.

Podemos observar que al eliminar la cooperatividad existente entre los represores el comportamiento pulsátil prácticamente desaparece, lo cual se infiere a partir del hecho que el Factor de Fano para esta cepa tiene un valor de 1 (Fig. 4.3a). Al no haber cooperatividad, la estabilidad del complejo represor-operador disminuye, provocando que su cinética sea tan rápida que no permite la existencia de periodos largos de inactividad transcripcional, necesarios para la existencia de los pulsos de producción de ARNm.

En concordancia con estos resultados, la cepa hipotética equivalente sin cooperatividad (CHE 4) también posee un FF con un valor de 1 (Fig. 4.3b). Además, se puede observar que en estas cepas existe un incremento en el comportamiento pulsátil (FF) conforme se aumenta la cooperatividad.

### 4.3. Ruido bioquímico

Para cuantificar las variaciones de las diversas especies moleculares en las poblaciones celulares simuladas se hizo uso del llamado ruido bioquímico, el cual está dado por:

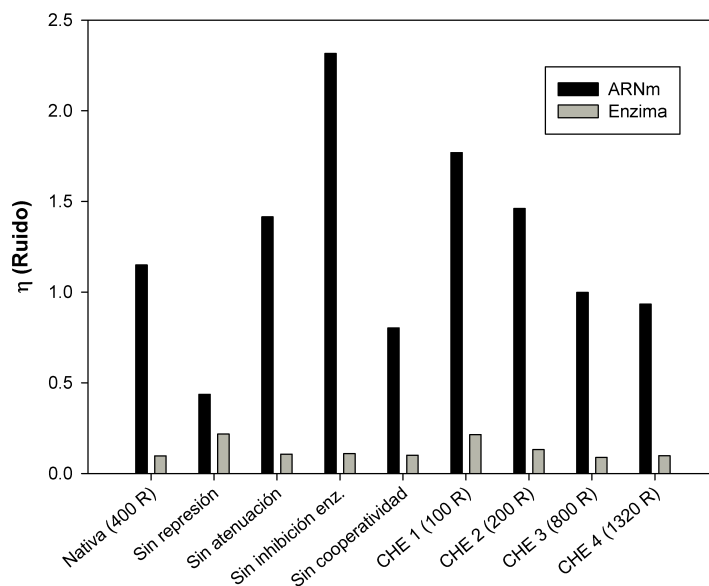
$$\eta = \frac{\sigma_X}{\mu_X} \quad (4.2)$$

en donde  $\sigma_X$  y  $\mu_X$  son la desviación estándar y la media, y  $X$  se refiere a la especie molecular de que se trata (ARNm, Enzima o Triptófano).

Al comparar los niveles de ruido del ARNm con los de la enzima, podemos observar una disminución generalizada de la variabilidad en el número de enzimas con respecto a la variabilidad en el ARNm (Fig. 4.4). Esta disminución se debe al hecho de que la cinética para la enzima, en particular la degradación, es mucho más lenta que para el ARNm. Por lo tanto, la



degradación lenta de las enzimas permite que las variaciones rápidas en el número de ARNm sufran una especie de promediación a lo largo del tiempo.

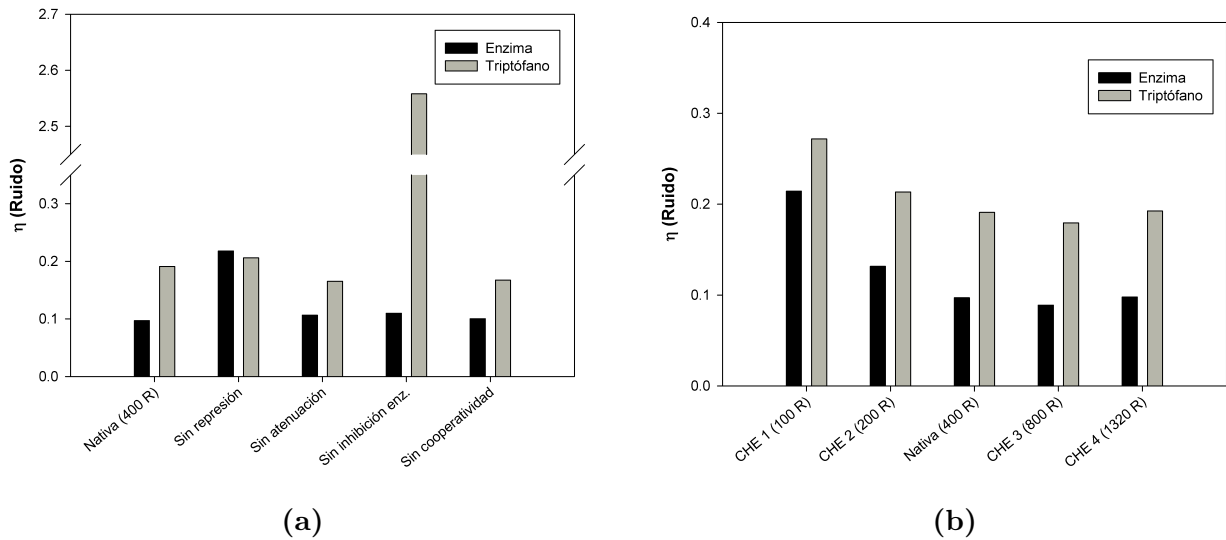


**Figura 4.4:** Caracterización del ruido bioquímico en los niveles de ARNm y enzima para las distintas cepas simuladas. R: Represores; enz.: enzimática; CHE: Cepa Hipotética Equivalente.

De manera opuesta, al comparar los niveles de ruido de las enzimas con los del triptófano, observamos que el ruido generalmente aumenta para el número de moléculas de triptófano, en especial cuando no está presente la inhibición enzimática (Fig. 4.5a). Esta propagación y amplificación del ruido es posible dado que la cinética de la síntesis y degradación de triptófano es más rápida que la de la enzima (Thattai y van Oudenaarden, 2002; Bruggeman et al., 2009).

Por otro lado, el lazo de retroalimentación negativa de la inhibición enzimática permite una reducción drástica en los niveles de ruido en el número de moléculas de triptófano. Esto se infiere a partir de que en la cepa sin inhibición enzimática el ruido bioquímico en el triptófano aumenta considerablemente (Fig. 4.5a).

Comparando los niveles de ruido en la enzima de la cepa nativa con los de las cepas sin un



**Figura 4.5: Caracterización del ruido bioquímico en los niveles de enzima y de triptófano para las distintas cepas simuladas.** Comparación de la intensidad del ruido bioquímico en los niveles de enzima y triptófano para la cepa nativa, (a) las cepas con los lazos de retroalimentación modificados y (b) las cepas hipotéticas equivalentes. R: Represores; enz.: enzimática; CHE: cepa hipotética equivalente.

lazo de retroalimentación o sin la cooperatividad, observamos que al no existir la represión, el ruido aumenta a más del doble (Fig. 4.5a).

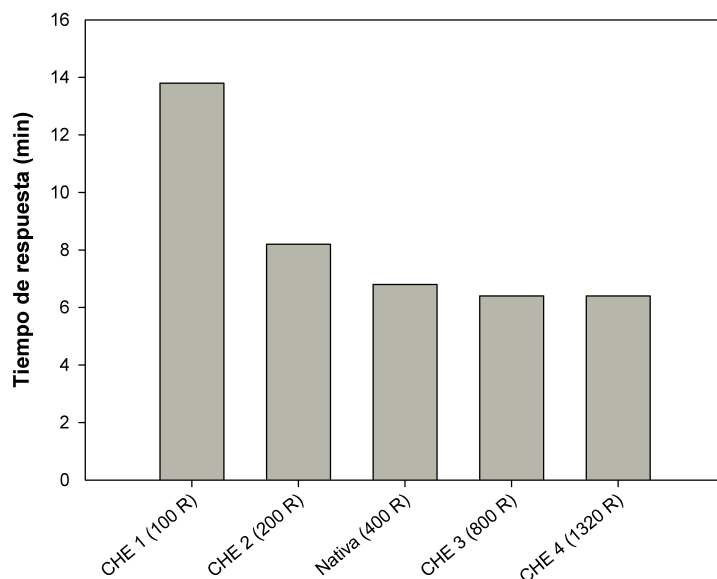
Si ahora analizamos la variaciones en los niveles de enzima para las cepas hipotéticas equivalentes (CHE) y la cepa nativa, nos damos cuenta que conforme aumentamos la fuerza de la cooperatividad y disminuimos el número de represores, el ruido aumenta (Fig. 4.5b). La principal razón de este aumento en la variabilidad es debida a que el incremento de la cooperatividad permite periodos cada vez más largos de inactividad transcripcional, provocando que la caída en los niveles de enzima entre dos periodos de actividad transcripcional sea cada vez más grande.

Es importante notar que aunque *E. coli* utilizara 1320 represores en lugar de 400, se predice que la magnitud del ruido no se vería afectada (Fig. 4.5b). Esto representaría para la célula invertir energía para producir más de 900 represores adicionales, sin lograr un efecto notorio sobre la reducción del ruido. En cambio, si *E. coli* utilizara 200 o 300 represores menos que la

cepa nativa, se predice que el ruido aumentaría de manera importante. Muy probablemente un aumento de tal magnitud en la variabilidad del triptófano repercutiría en el desempeño de aquellas células cuyos valores de triptófano estén por debajo de los necesarios para un crecimiento óptimo, bajo las condiciones en las que se encuentren las células.

## 4.4. Tiempo de respuesta

Centrándonos en el estudio del efecto que tendrían las modificaciones en las CHE sobre el tiempo con el que las células responden ante la des-represión, analizamos la dinámica promedio de los niveles de triptófano y determinamos el tiempo en el que cada cepa alcanza el 95 % de su nivel de triptófano en el estado estacionario.



**Figura 4.6: Tiempo de respuesta en la cepa nativa y en las cepas hipotéticas equivalentes (CHE).** Comparación del tiempo al cual la dinámica promedio del triptófano en cada cepa alcanza el 95 % del valor en estado estacionario. min:minutos; R: Represores; CHE: Cepa(s) Hipotética(s) Equivalente(s).

Podemos observar en la figura 4.6 que de manera similar al ruido en la enzima, el tiempo de respuesta aumenta conforme aumentamos la cooperatividad en las cepas. Este incremen-

to en el tiempo de respuesta lo atribuimos a que al aumentar la cooperatividad, también aumentará el tiempo que tardan los represores en disociarse del promotor después de la des-represión. Además, es importante notar que al modificar la cooperatividad de tal manera que aumentamos el número de represores por arriba del de la cepa nativa, el tiempo de respuesta disminuye muy levemente. De manera opuesta, cuando el número de represores es menor que el de la cepa nativa, el tiempo de respuesta aumenta considerablemente. Esto último sería perjudicial para la sobrevivencia de la población bacteriana ante un agotamiento repentino de las fuentes externas de triptófano.

## 4.5. Relación Costo-Beneficio

Para investigar más a fondo las posibles implicaciones que tiene la existencia de la cooperatividad en el operón *trp* de *E. coli* decidimos formular una función que relacionara los costos y beneficios que conllevan las modificaciones presentes en las CHE (ver Tabla 3.1) y la cepa nativa. La función Costo-Beneficio ( $\psi$ ) está dada por:

$$\psi = \frac{1}{\eta^*} + \frac{1}{\tau_R^*} - R_{Tot}^* \quad (4.3)$$

en donde  $\eta$  representa el ruido en los niveles de triptófano,  $\tau_R$  el tiempo de respuesta,  $R_{Tot}$  el número total de represores y \* indica la normalización de estos valores con los de la CHE sin cooperatividad.

Al analizar los valores que toma esta función para las distintas CHE y para la cepa nativa, observamos que hasta cierto punto el empleo de la cooperatividad permite la disminución del costo energético por los represores y a la vez un aumento en la relación Costo-Beneficio (Fig 4.7). Después de este punto de inflexión, el valor que toma  $\psi$  empieza a disminuir. Más importante aún es el hecho de que este punto en donde la relación Costo-Beneficio tiene el valor máximo corresponde a la cepa nativa. Esto nos da un indicio de que el sistema ha sufrido una optimización en la fuerza de cooperatividad a lo largo del tiempo, y de que entre los factores que moldearon esta optimización muy probablemente se encuentran el costo

energético de los represores, las variaciones en el sistema y el tiempo de respuesta.

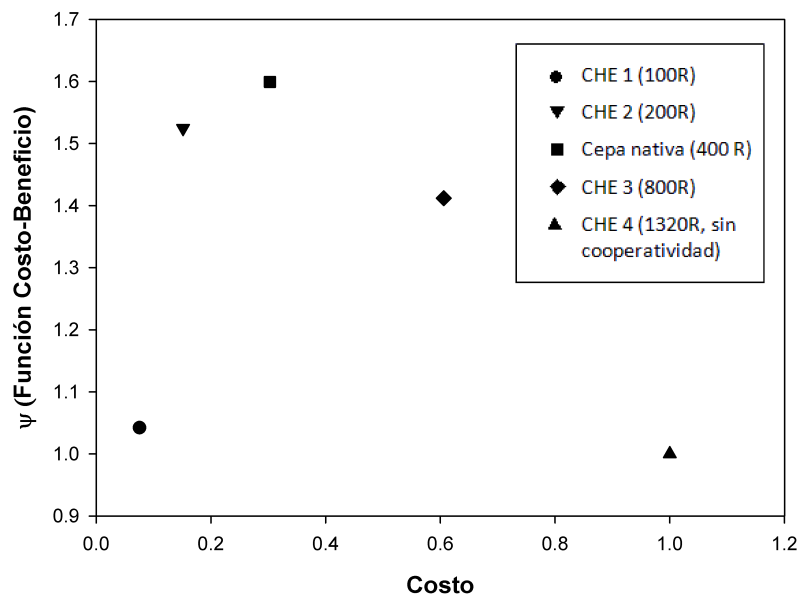


Figura 4.7: Relación Costo-Beneficio para la cepa nativa y las cepas hipotéticas equivalentes (CHE). Costo: Costo energético debido a represores (normalizado); R: Represores.

# Capítulo 5

## Discusión y Conclusiones

### 5.1. Discusión

En este trabajo se desarrolló un modelo matemático del operón *trp* de *E. coli* para investigar la existencia de los pulsos transcripcionales en el sistema, así como estudiar el papel que juegan los lazos de retroalimentación y en especial la cooperatividad, en la generación de estos pulsos y en la dinámica estocástica del sistema.

Gracias al gran esfuerzo invertido para incorporar los aspectos relevantes del sistema y para calcular los parámetros de la manera más adecuada, fue posible la creación de un modelo matemático que reprodujera adecuadamente experimentos de des-represión realizados por Yanofsky y Horn (1994). El modelo no sólo reproduce los resultados experimentales observados en la cepa nativa, sino que además los resultados experimentales de la cepa sin inhibición son reproducidos por su contraparte virtual.

Investigando la importancia que tienen los diferentes lazos de retroalimentación para la generación de la transcripción pulsátil, identificamos a la represión como un componente clave para que se dé el comportamiento de la producción pulsátil de ARNm. Para estudiar más a fondo que característica de la represión hace posible este comportamiento, se analizó una cepa hipotética sin cooperatividad. Interesantemente, los resultados indican que la cooperatividad es totalmente necesaria para que sea posible la generación de pulsos transcripcionales en el operón *trp* de *E. coli*.

Al analizar el papel de los tres lazos de retroalimentación negativa con los que cuenta el sistema sobre la variabilidad en la expresión del operón, observamos que la presencia de la represión permite disminuir los niveles de ruido en el número de enzimas. Este resultado concuerda con las observaciones experimentales de Dublanche et al. (2006), en donde se

obtuvo que la intensidad del ruido en los niveles de proteína eran menores en un sistema de regulación génica con represión, que en su versión con la ausencia de este mecanismo de regulación.

Adicionalmente, se estudió la propagación del ruido que se da desde los niveles de ARNm hasta los niveles de triptófano. Se encontró que al pasar de ARNm a proteína, la magnitud del ruido bioquímico disminuye drásticamente. Esto representa una ventaja para las célula considerando que variaciones excesivas en los niveles de enzima y, por lo tanto, de triptófano tendrían un efecto perjudicial en el desempeño de la población celular. Al disminuir estas variaciones aseguramos que la mayoría de las células cuenten con los niveles necesarios de triptófano intracelular. La reducción del ruido bioquímico que se da entre el paso de ARNm a proteína sucede gracias a que la cinética de la síntesis y degradación del ARNm es mucho más rápida que la cinética de degradación de la enzima (Bruggeman et al., 2009). Por otro lado, el hecho de que la cinética en los niveles de enzima sea mucho más lenta que la cinética de los niveles de triptófano provoca una amplificación importante del ruido en el triptófano cuando no está presente la inhibición enzimática. En estos resultados se ve reflejada la importancia que tiene la inhibición enzimática en la reducción de las fluctuaciones en los niveles de triptófano.

Se ha propuesto que las variaciones en los niveles de uno de los intermediarios metabólicos, provocadas por ejemplo por variaciones en los niveles de enzima, no son propagadas al resto de la ruta metabólica (Levine y Hwa, 2007). Considerando que todas las enzimas encargadas de la transformación de corismato a triptófano en *E. coli* son producidas a partir del mismo ARNm, podemos deducir que para este caso las variaciones en la enzima antranilato sintasa serán propagadas en buena medida a los niveles intracelulares de triptófano.

Con el fin de investigar los aspectos evolutivos detrás de la presencia de la cooperatividad en el operón *trp*, se crearon una serie de cepas hipotéticas equivalentes (CHE), las cuales tienen modificaciones en la constante de cooperatividad y el número de represores, pero a la vez poseen niveles de expresión muy similares a los de la cepa nativa. Comparando las características de la dinámica del sistema de las CHE con las de la cepa nativa, observamos

que el ruido en las distintas especies moleculares y el tiempo de respuesta aumentan conforme se aumenta la cooperatividad. El aumento en el ruido y tiempo de respuesta sólo es considerable cuando la cooperatividad posee un valor superior al de la cepa nativa. Esto nos indica la importancia que tiene la cooperatividad para permitir producir una cantidad menor de represores sin modificar notoriamente la fuerza de represión, el ruido en el sistema y el tiempo de respuesta.

Para estudiar la relación que hay entre el gasto energético en la producción de represores, el ruido y el tiempo de respuesta, utilizamos una función Costo-Beneficio ( $\psi$ ) que permitió la integración de estas tres características. Al calcular y comparar los valores de  $\psi$  para las distintas CHE y para la cepa nativa, encontramos que el valor más alto que toma esta función corresponde al de la cepa nativa. Este resultado sugiere que muy probablemente la fuerza de cooperatividad presente en el sistema del operón *trp* de *E. coli* ha estado sujeta a una presión evolutiva, en la cual han jugado un papel importante ciertas características que afectan al desempeño de la célula, como lo son el costo energético, el ruido en el sistema y el tiempo de respuesta.

## 5.2. Conclusiones

A partir de la discusión presentada anteriormente se pueden recopilar las siguientes conclusiones:

1. La cooperatividad entre represores es responsable de la existencia de pulsos transcripcionales en el operón triptófano de *Escherichia coli*.
2. Los pulsos transcripcionales presentes en el sistema aumentan las fluctuaciones en los niveles de ARN mensajero, pero no en los niveles de enzima y triptófano.
3. La atenuación de la transcripción y la inhibición enzimática no son determinantes en la existencia de pulsos transcripcionales en el operón *trp* de *E. coli*. Los efectos de estos



mecanismos de regulación sobre el ruido bioquímico son los esperados.

4. El nivel de cooperatividad en la cepa nativa representa una solución óptima, en terminos del costo asociado a la producción de represores, así como de las ventajas inherentes a la disminución de los tiempos de respuesta y de los niveles de ruido bioquímico.

# Capítulo 6

## Trabajo a futuro

A partir del modelo se generaron una serie de datos que pueden ser comparados con resultados experimentales, como es el caso de las dinámicas y distribuciones para las distintas especies moleculares en las diferentes cepas.

Actualmente en el Cinvestav Monterrey se están llevando a cabo experimentos que se prevé permitirán la obtención de la distribución de proteínas cuya expresión se encuentra bajo el control del promotor *trp*. Para ello se están utilizando cultivos de la bacteria *E. coli* transformados con un plásmido que contiene al gen de la Proteína Verde Fluorescente (GFP, por sus siglas en inglés) bajo el control del promotor *trp*. El uso de microscopía permitirá la cuantificación de los niveles de GFP en células individuales, obteniendo por consiguiente su distribución en la población celular. Para estos experimentos se están empleando la cepa nativa y la cepa sin inhibición enzimática, por lo que podríamos hacer la comparación de los resultados experimentales con los resultados obtenidos a partir de sus contrapartes virtuales.

Adicionalmente, se podrían hacer comparaciones con resultados obtenidos con la técnica utilizada por Golding et al. (2005) para la cuantificación en tiempo real de la producción de ARNm, o la técnica utilizada por Zenklusen et al. (2008) para la obtención de la distribución del número de ARNm en la población celular.

Por la parte teórica, sería interesante la creación de un código que permita simular el transporte de triptófano entre células vecinas a través del medio extracelular. Es posible que con esta incorporación, el modelo sea capaz de reproducir las oscilaciones sincronizadas que fueron observadas en el trabajo de Bliss et al. (1982) en la cepa sin inhibición enzimática después de la des-represión.

# Apéndice A

## Desarrollo del modelo

Siguiendo las consideraciones mencionadas en la Sección Algoritmo para las simulaciones estocásticas (3.4), los procesos fueron divididos en dos tipos: rápidos y lentos. Para los procesos rápidos se obtienen las ecuaciones en el estado de equilibrio, y con ellas se calcula el número de moléculas de la especie de la que se esté tratando. Por otro lado, las reacciones que describen a los procesos lentos serán tomadas directamente para realizar las simulaciones estocásticas.

Para la construcción del modelo se considera un volumen celular constante ( $V_c$ ). Por lo tanto, a lo largo de esta sección nos referiremos a la concentración de cierta especie molecular (número de moléculas/ $V_c$ ) simplemente como el número de moléculas de la especie en cuestión.

Empecemos describiendo la ecuación que representa la interacción entre el triptófano y los represores. El represor ( $R$ ) está conformado por dos subunidades idénticas, cada una con un sitio capaz de unir una molécula de triptófano ( $T$ ). Las reacciones que representan estas interacciones son:



en donde  $K_{RT}$  representa la constante de disociación entre el represor y el triptófano.  $K_{RT}$  está dividida entre dos en la primera reacción debido a que el triptófano tiene la posibilidad de asociarse a cualquiera de los dos sitios en el represor. Por el contrario, en la segunda reacción el triptófano se puede disociar de cualquiera de las dos posiciones del represor, y por lo tanto  $K_{RT}$  se multiplica por dos. La rapidez con que se llevan a cabo estas reacciones nos permiten considerar este proceso en estado de cuasi-equilibrio, lo cual nos servirá para obtener el número promedio de represores activos, aquellos asociados con dos moléculas de triptófano. Las ecuaciones que representan el estado de equilibrio de las reacciones A.1 son:

$$(R)(T) = \frac{1}{2}(K_{RT})(RT) \quad y \quad (RT)(T) = 2(K_{RT})(RT_2) \quad (\text{A.2})$$

Si consideramos que el número total de represores ( $R_{Tot}$ ) es constante, entonces:

$$R_{Tot} = R + RT + RT_2 \quad (\text{A.3})$$

A partir de las ecuaciones A.2 y A.3 podemos obtener el número promedio de represores activos:

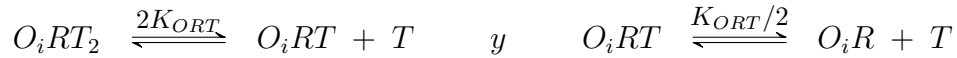
$$RT_2 = R_{Tot} \left( \frac{T}{T + K_{RT}} \right)^2 \quad (\text{A.4})$$

Estos represores se podrán unir a los tres diferentes operadores de la siguiente manera:



en donde  $O_i$  representa al operador  $i$ , con  $i$  tomando valores de 1, 2 y 3; y  $k_{i,RT_2}^+$  representa la tasa de asociación del represor activo al operador  $i$ , con el signo  $+$  indicando que se trata de la tasa de asociación.

Al igual que el represor libre, el represor que se encuentra unido al operador, alcanza rápidamente un equilibrio entre sus diferentes estados con ninguna, una o dos moléculas de triptófano: para  $i=1,2$



y para  $i=3$



en donde  $K_{ORT}$  es la constante de disociación entre el represor unido al operador y el triptófano. La razón de utilizar constantes de disociación distintas para el operador 3 es explicada a detalle en el artículo de Tabaka et al. (2008). En resumidas cuentas, si esta consideración no se hiciera, la afinidad entre el represor sin triptófano unido y el operador resultaría incluso menor que la afinidad del represor por una secuencia no específica en el ADN, lo cual es inconsistente con la realidad.

Al disociarse las moléculas de triptófano del represor, la estabilidad del complejo operador-represor disminuye, provocando un incremento en la tasa de disociación de este complejo. Para calcular la tasa promedio con la que el represor se disociará del operador, primero debemos conocer cuál es la distribución de probabilidad entre los tres estados del represor unido, la cual dependerá de los niveles intracelulares de triptófano. Para ello, necesitamos primero obtener las ecuaciones de equilibrio de las reacciones A.6:

para  $i=1,2$

$$2(K_{ORT})(O_iRT_2) = (O_iRT)(T) \quad y \quad \frac{1}{2}(K_{ORT})(O_iRT) = (O_iR)(T)$$

y para  $i=3$

$$(K_{RT})(O_iRT_2) = (O_iRT)(T) \quad y \quad (K_{ORT})(O_iRT) = (O_iR)(T) \quad (\text{A.7})$$

Luego, considerando que la suma de las probabilidades de los tres estados debe ser 1, tenemos que:

$$1 = P(O_iRT_2) + P(O_iRT) + P(O_iR) \quad (\text{A.8})$$

Por último, para calcular las probabilidades de los distintos estados, resolvemos las ecuaciones A.7 y A.8 obteniendo que:

para  $i=1,2$

$$P(O_iR) = \left( \frac{K_{ORT}}{K_{ORT} + T} \right)^2$$

$$P(O_iRT) = \frac{2(K_{ORT})(T)}{(K_{ORT} + T)^2} \quad (\text{A.9})$$

$$P(O_iRT_2) = \left( \frac{T}{K_{ORT} + T} \right)^2$$

y para  $i=3$

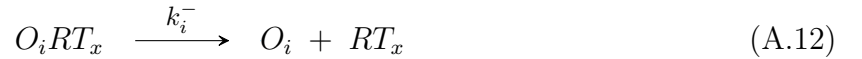
$$\begin{aligned}
P(O_i R) &= \frac{K_{ORT}}{K_{ORT} + T + \frac{T^2}{K_{RT}}} \\
P(O_i RT) &= \frac{T}{K_{ORT} + T + \frac{T^2}{K_{RT}}} \\
P(O_i RT_2) &= \frac{T^2}{K_{RT}(K_{ORT} + T + \frac{T^2}{K_{RT}})}
\end{aligned} \tag{A.10}$$

Ponderando las contribuciones que tiene cada estado sobre la tasa de disociación promedio del represor ( $k_i^-$ ), obtenemos que:

$$k_i^- = k_{i,R}^- P(O_i R) + k_{i,RT}^- P(O_i RT) + k_{i,RT_2}^- P(O_i RT_2) \tag{A.11}$$

en donde  $k_{i,R}^-$ ,  $k_{i,RT}^-$  y  $k_{i,RT_2}^-$  corresponden a las tasas con las que los distintos estados del represor se disocian del operador  $i$ . Por lo tanto, la reacción de disociación entre los represores y los operadores individuales queda de la siguiente manera:

para  $i=1,2,3$



en donde  $O_i RT_x$  representa al complejo entre el operador ( $O_i$ ) y el represor en cualquiera de sus tres estados ( $RT_x$ ).

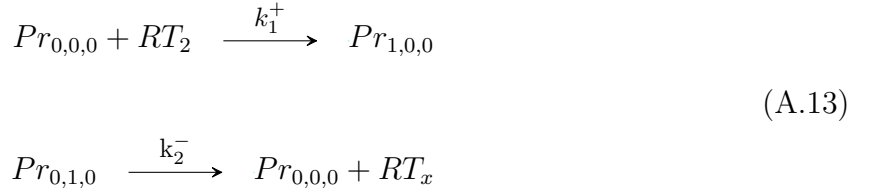
Hasta ahora hemos tratado a los operadores individualmente. Para integrar la información de los tres operadores en una sola entidad, el promotor, utilizaremos la siguiente representación:

$$Pr_{x,y,z}$$

en donde  $x$ ,  $y$  y  $z$  hacen alusión respectivamente al primer, segundo y tercer operador del promotor. Estas tres variables tomarán un valor de 0 cuando el operador al que hacen referencia se encuentra libre y 1 cuando este se encuentre ocupado por un represor. Por lo tanto, tendremos 8 estados del promotor, incluyendo aquel donde el promotor se encuentra

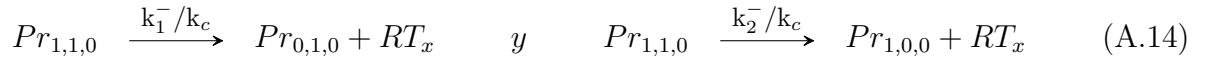
totalmente libre ( $Pr_{0,0,0}$ ), aquellos donde un operador se encuentra ocupado por un represor (e.g.  $Pr_{1,0,0}$ ), dos de los tres se encuentren ocupados (e.g.  $Pr_{1,1,0}$ ), o todos los operadores se encuentren ocupados ( $Pr_{1,1,1}$ ).

Si representamos todas las posibles reacciones entre estos estados, encontraremos que en 12 de ellas un represor se asocia a uno de los operadores, mientras que en otras 12 se disocia. Para ejemplificar, mostramos una reacción de asociación y otra de disociación:



en donde podemos apreciar que se considera que sólo el represor activo ( $RT_2$ ) se puede asociar al operador, y que el represor se puede disociar estando en cualquiera de sus tres estados ( $RT_x$ ).

Ahora que consideramos a los operadores conjuntamente, debemos tomar en cuenta la cooperatividad que existe entre los represores unidos a los operadores 1 y 2. Por lo tanto, las reacciones donde la cooperatividad estará involucrada serán:



en donde  $k_c$  representa la constante de cooperatividad.

Cuando el promotor se encuentre libre, la ARN polimerasa será capaz de unírsele y comenzar la transcripción. Por lo tanto, la reacción de producción de ARNm estará dada por:



en donde  $k_M$  es la tasa máxima promedio con que se lleva a cabo la transcripción,  $\tau_M$  es el retraso entre el inicio de la transcripción y el instante en el que un ribosoma se le puede unir al mensajero y llevar a cabo la traducción sin ser frenado por la polimerasa, y  $P(A)$  representa la probabilidad de que no ocurra atenuación durante la producción de ARNm. Santillan y

Zeron (2004) formularon una ecuación que representa adecuadamente esta probabilidad:

$$P_A(T) = \frac{1 + 2\alpha \frac{T}{K_G+T}}{(1 + \alpha \frac{T}{K_G+T})^2} \quad (\text{A.16})$$

en donde  $\alpha$  y  $K_G$  son parámetros a estimar.

La disminución en el número de ARNm va a estar determinada por dos factores: la degradación, y la dilución por crecimiento celular. Estos dos procesos pueden ser representados conjuntamente por la siguiente reacción:



en donde  $M$  representa al ARNm,  $\gamma_M$  es la tasa con que este se degrada,  $\mu$  es la tasa de dilución, y el símbolo de conjunto vacío  $\emptyset$  indica la desaparición del ARNm del sistema.

A partir del ARNm los ribosomas llevarán a cabo la producción de la enzima (E) antranilato sintasa. Esta producción está representada por la siguiente reacción:



en donde  $k_E$ , la tasa de síntesis de enzimas por cada ARNm, está dividida entre 2 para tomar en cuenta que la enzima antranilato sintasa está compuesta por dos copias de cada una de sus subunidades E y D; y  $\tau_E$  representa el retardo en la traducción, que estaría dado por el tiempo que pasa desde que el ribosoma se une al mensajero hasta que termina la producción de la enzima.

Al igual que el ARNm, las enzimas podrán ser degradadas o diluidas, pero el hecho de que la tasa de dilución ( $\mu$ ) sea mayor que la tasa de degradación ( $\gamma_E$ ) por cerca de 2 órdenes de magnitud, nos permite despreciar esta última. Por lo tanto, la reacción que determinará la disminución en el número de enzimas será:



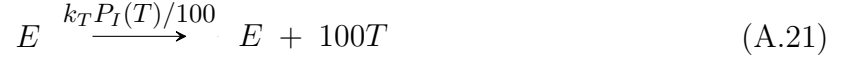
La enzima antranilato sintasa puede producir triptófano con una tasa máxima  $k_T$ . La inhibición enzimática provoca que esta tasa disminuya conforme la concentración de triptófano



aumenta. Para tomar en cuenta este fenómeno, utilizaremos una función que represente la probabilidad de que la enzima no se encuentre inhibida por el triptófano ( $P_I(T)$ ). Esta función fue propuesta por Caligiuri y Bauerle (1991) a partir de datos experimentales:

$$P_I(T) = \frac{K_I^n}{K_I^n + T^n} \quad (\text{A.20})$$

en donde  $K_I$  es la constante de disociación del complejo triptófano-enzima, y  $n$  representa el coeficiente de interacción, mejor conocido como coeficiente de Hill. Dado que los eventos de producción de triptófano son muy frecuentes y representan una carga computacional elevada, podemos suponer que la producción de triptófano se lleva a cabo en grupos de 100 moléculas. Al hacer simulaciones de prueba con y sin esta suposición, se observó que el hacer esta suposición no afecta el comportamiento del sistema. Por lo tanto, la reacción de producción de triptófano (T) queda representada de la siguiente manera:



Siguiendo el mismo razonamiento, la reacción que determinará la disminución en los niveles de triptófano será:



en donde  $\rho(T)$  es la tasa de utilización de triptófano libre para la fabricación de proteínas, y  $\mu$  es la antes mencionada tasa de dilución. Para considerar que la fabricación de proteínas disminuye cuando la disponibilidad de triptófano es demasiado baja, proponemos que  $\rho(T)$  estará dada por la siguiente función

$$\rho(T) = \rho_{max} \frac{T}{T + K_\rho} \quad (\text{A.23})$$

en donde  $\rho_{max}$  es la tasa máxima de consumo de triptófano y  $K_\rho$  es una constante que permita representar el comportamiento antes mencionado.

A manera de resumen, en la tabla A.1 se muestra una lista de las reacciones que son utilizadas para llevar a cabo las simulaciones estocásticas, así como sus correspondientes propensiones efectivas.

**Tabla A.1:** Lista de reacciones

Reacción	Propensión efectiva
$Pr_{0,0,0} + RT_2 \rightarrow Pr_{1,0,0}$	$k_1^+ Pr_{0,0,0} RT_2$
$Pr_{1,0,0} \rightarrow Pr_{0,0,0} + RT_x$	$k_1^- Pr_{1,0,0}$
$Pr_{0,0,0} + RT_2 \rightarrow Pr_{0,1,0}$	$k_2^+ Pr_{0,0,0} RT_2$
$Pr_{0,1,0} \rightarrow Pr_{0,0,0} + RT_x$	$k_2^- Pr_{0,1,0}$
$Pr_{0,0,0} + RT_2 \rightarrow Pr_{0,0,1}$	$k_3^+ Pr_{0,0,0} RT_2$
$Pr_{0,0,1} \rightarrow Pr_{0,0,0} + RT_x$	$k_3^- Pr_{0,0,1}$
$Pr_{1,0,0} + RT_2 \rightarrow Pr_{1,1,0}$	$k_2^+ Pr_{1,0,0} RT_2$
$Pr_{1,1,0} \rightarrow Pr_{1,0,0} + RT_x$	$k_2^- Pr_{1,1,0}$
$Pr_{1,0,0} + RT_2 \rightarrow Pr_{1,0,1}$	$k_3^+ Pr_{1,0,0} RT_2$
$Pr_{1,0,1} \rightarrow Pr_{1,0,0} + RT_x$	$k_3^- Pr_{1,0,1}$
$Pr_{0,1,0} + RT_2 \rightarrow Pr_{1,1,0}$	$k_1^+ Pr_{0,1,0} RT_2$
$Pr_{1,1,0} \rightarrow Pr_{0,1,0} + RT_x$	$k_1^- Pr_{1,1,0}$
$Pr_{0,1,0} + RT_2 \rightarrow Pr_{0,1,1}$	$k_3^+ Pr_{0,1,0} RT_2$
$Pr_{0,1,1} \rightarrow Pr_{0,1,0} + RT_x$	$k_3^- Pr_{0,1,1}$
$Pr_{0,0,1} + RT_2 \rightarrow Pr_{1,0,1}$	$k_1^+ Pr_{0,0,1} RT_2$
$Pr_{1,0,1} \rightarrow Pr_{0,0,1} + RT_x$	$k_1^- Pr_{1,0,1}$
$Pr_{0,0,1} + RT_2 \rightarrow Pr_{0,1,1}$	$k_2^+ Pr_{0,0,1} RT_2$
$Pr_{0,1,1} \rightarrow Pr_{0,0,1} + RT_x$	$k_2^- Pr_{0,1,1}$
$Pr_{1,1,0} + RT_2 \rightarrow Pr_{1,1,1}$	$k_3^+ Pr_{1,1,0} RT_2$
$Pr_{1,1,1} \rightarrow Pr_{1,1,0} + RT_x$	$k_3^- Pr_{1,1,1}$
$Pr_{1,0,1} + RT_2 \rightarrow Pr_{1,1,1}$	$k_2^+ Pr_{1,0,1} RT_2$
$Pr_{1,1,1} \rightarrow Pr_{1,0,1} + RT_x$	$k_2^- Pr_{1,1,1}$
$Pr_{0,1,1} + RT_2 \rightarrow Pr_{1,1,1}$	$k_1^+ Pr_{0,1,1} RT_2$
$Pr_{1,1,1} \rightarrow Pr_{0,1,1} + RT_x$	$k_1^- Pr_{1,1,1}$
$Pr_{0,0,0} \rightarrow Pr_{0,0,0} + M$	$k_M P_A(T) Pr_{0,0,0}$
$M \rightarrow \emptyset$	$(\gamma_M + \mu) M$
$M \rightarrow M + E$	$k_E M/2$
$E \rightarrow \emptyset$	$\mu M$
$E \rightarrow E + 100T$	$k_T P_I(A) E/100$
$100T \rightarrow \emptyset$	$(\rho(T) + \mu) T/100$

# Apéndice B

## Estimación de parámetros

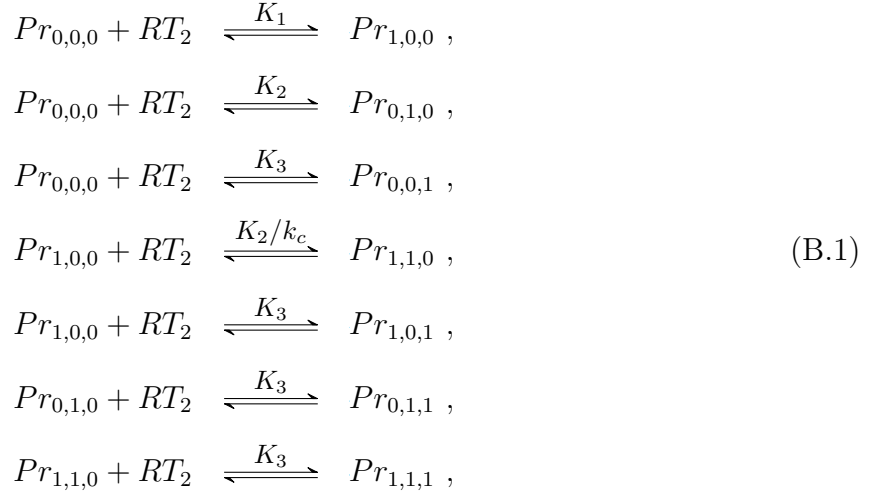
Usando microscopía electrónica Javelle et al. (2005) estimaron que el volumen celular promedio ( $V_c$ ) de *Escherichia coli* es de alrededor  $1\mu\text{m}^3$ . Como se mencionó en el apéndice de Desarrollo del modelo este volumen será tomado como constante y por lo tanto, el resto de los parámetros que indiquen concentración sólo serán representados como el número de moléculas, sobreentendiendo que nos estaremos refiriendo al número de moléculas/ $V_c$ .

Para determinar cuál es el estado inicial del promotor, se diseñó un algoritmo que permitiera escoger aleatoriamente los estados iniciales del promotor, considerando la combinación de dos probabilidades: la probabilidad que tiene una bacteria de contar con uno o con dos promotores; y la probabilidad de que cierto estado del promotor se encuentre presente en las condiciones iniciales de la bacteria.

Baker y Yanofsky (1972) calcularon que durante el crecimiento exponencial la bacteria *E. coli* contiene en promedio 1.8 copias del operón triptófano. Por lo tanto, la probabilidad de que una célula posea uno o dos operones *trp* será de 0.2 y 0.8, respectivamente. Dicho de otra manera, 1 de cada 5 bacterias contará con 1 operón, y las 4 bacterias restantes, con 2.

Por otro lado, para obtener la probabilidad de cada uno de los estados del promotor, es necesario primero seleccionar de la tabla de reacciones (Tabla A.1) un conjunto de reacciones de transición que interrelacione a los 8 estados del promotor. Para facilitar el manejo de este conjunto, una reacción de asociación y su correspondiente reacción de disociación son

mostradas en una sola reacción bidireccional. A continuación se presentan estas reacciones:



en donde  $K_i = k_i^-/k_i^+$  es la constante de disociación entre el represor y el operador  $i = 1, 2, 3$  a los niveles iniciales de triptófano ( $T_0$ ). A partir de ellas obtenemos sus respectivas ecuaciones en estado de equilibrio:

$$\begin{aligned}
Pr_{0,0,0} \frac{RT_2}{K_1} &= Pr_{1,0,0} , \\
Pr_{0,0,0} \frac{RT_2}{K_2} &= Pr_{0,1,0} , \\
Pr_{0,0,0} \frac{RT_2}{K_3} &= Pr_{0,0,1} , \\
Pr_{1,0,0} \frac{k_c RT_2}{K_2} &= Pr_{1,1,0} , \\
Pr_{1,0,0} \frac{RT_3}{K_3} &= Pr_{1,0,1} , \\
Pr_{0,1,0} \frac{RT_2}{K_3} &= Pr_{0,1,1} , \\
Pr_{1,1,0} \frac{RT_2}{K_3} &= Pr_{1,1,1} ,
\end{aligned} \tag{B.2}$$

Utilizando estas ecuaciones de equilibrio y sabiendo que la probabilidad de que el promotor se encuentre en cualquiera de sus 8 estados es igual a 1 ( $Pr_{0,0,0} + Pr_{1,0,0} + Pr_{0,1,0} + Pr_{0,0,1} +$

$Pr_{1,1,0} + Pr_{1,0,1} + Pr_{0,1,1} + Pr_{1,1,1} = 1$ ), se obtuvieron las ecuaciones que representan la probabilidad de cada uno de los estados del promotor:

$$\begin{aligned}
P(Pr_{0,0,0}) &= \frac{1}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)}, \\
P(Pr_{1,0,0}) &= \frac{\frac{RT_2}{K_1}}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)}, \\
P(Pr_{0,1,0}) &= \frac{\frac{RT_2}{K_2}}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)}, \\
P(Pr_{0,0,1}) &= \frac{\frac{RT_2}{K_3}}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)}, \\
P(Pr_{1,1,0}) &= \frac{\frac{k_c RT_2^2}{K_1 K_2}}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)}, \\
P(Pr_{1,0,1}) &= \frac{\frac{RT_2^2}{K_1 K_3}}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)}, \\
P(Pr_{0,1,1}) &= \frac{\frac{RT_2^2}{K_2 K_3}}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)}, \\
P(Pr_{1,1,1}) &= \frac{\frac{k_c RT_2^3}{K_1 K_2 K_3}}{\left(1 + \frac{RT_2}{K_3}\right) \left(1 + \frac{RT_2}{K_1} + \frac{RT_2}{K_2} + \frac{k_c RT_2^2}{K_1 K_2}\right)},
\end{aligned} \tag{B.3}$$

en donde  $P(Pr_{x,y,z})$  representa la probabilidad de que el promotor se encuentre en el estado indicado,  $K_i = k_i^-/k_i^+$  ( $i = 1, 2, 3$ ), y  $RT_2$  es el número de represores activos, el cual dependerá de la concentración inicial de triptófano intracelular.

Para aquellas bacterias que cuentan sólo con un promotor, estas probabilidades simplemente

son multiplicadas por la probabilidad de que la célula tenga un promotor, la cual es de 0.2.

En las bacterias que poseen dos promotores, el número de posibilidades se incrementa de 8 a 36, debido a las diferentes combinaciones que se pueden dar entre los dos promotores. Por lo tanto, se calcula la probabilidad de cada una de las posibles combinaciones, y se multiplican por la probabilidad de que la bacteria posea 2 operones *trp*, la cual es de 0.8.

Una vez obtenidas las probabilidades de todas las posibles opciones, se toman en cuenta sólo aquellas que tengan un valor mayor a 0.001. Estas probabilidades serán utilizadas por el algoritmo, que aleatoriamente seleccionará el estado inicial de el(los) promotor(es).

Del trabajo de Morse et al. (1968) obtenemos que el número de enzimas AS antes de la des-represión ( $E_0$ ) es de

$$E_0 \approx 50 \text{ moléculas}$$

Por otro lado, a partir de la página web *E. coli Statistics* ([http://ccdb.wishartlab.com/CCDB/cgi-bin/STAT\\_NEW.cgi](http://ccdb.wishartlab.com/CCDB/cgi-bin/STAT_NEW.cgi)), obtenemos que el número inicial de moléculas de triptófano es de

$$T_0 \approx 80,000 \text{ moléculas}$$

La constante de disociación entre el represor y el triptófano ( $K_{RT}$ ), fue obtenida por Arvidson et al. (1986) en condiciones experimentales de pH 7.4 y temperatura de 37°C, y tiene un valor de

$$K_{RT} \approx 44,160 \text{ moléculas}$$

A partir del trabajo de Gunsalus et al. (1986), en donde se calculó el número de represores en *E. coli* cultivada en un medio sin triptófano, obtuvimos que:

$$R_{Tot} \approx 400 \text{ moléculas}$$

Para las tasas de asociación ( $k_i^+$ ) y disociación ( $k_i^-$ ) entre los distintos estados del represor (R, RT y RT2) y cada uno de los operadores, se utilizaron los valores empleados por Tabaka et al. (2008). Estos datos, los cuales fueron calculados a partir de los trabajos experimentales de Grillo et al. (1999), Hurlburt y Yanofsky (1992), Zhang et al. (1994), y Jardetzky y Finucane (2007), son presentados a continuación:

$$\begin{array}{ll}
 k_{1,RT2}^+ \approx 8.1 \text{ moléculas}^{-2} \text{ min}^{-1}, & k_{1,RT2}^- \approx 6.0 \text{ moléculas}^{-1} \text{ min}^{-1}, \\
 k_{1,RT}^- \approx 19.2 \text{ moléculas}^{-1} \text{ min}^{-1}, & k_{1,R}^- \approx 60.0 \text{ moléculas}^{-1} \text{ min}^{-1}, \\
 k_{2,RT2}^+ \approx 0.312 \text{ moléculas}^{-2} \text{ min}^{-1}, & k_{2,RT2}^- \approx 0.198 \text{ moléculas}^{-1} \text{ min}^{-1}, \\
 k_{2,RT}^- \approx 6.6 \text{ moléculas}^{-1} \text{ min}^{-1}, & k_{2,R}^- \approx 66.0 \text{ moléculas}^{-1} \text{ min}^{-1}, \\
 k_{3,RT2}^+ \approx 0.3 \text{ moléculas}^{-2} \text{ min}^{-1}, & k_{3,RT2}^- \approx 36.0 \text{ moléculas}^{-1} \text{ min}^{-1}, \\
 k_{3,RT}^- \approx 72.0 \text{ moléculas}^{-1} \text{ min}^{-1}, & k_{3,R}^- \approx 810.0 \text{ moléculas}^{-1} \text{ min}^{-1},
 \end{array}$$

La constante de disociación entre el triptófano y el represor que se encuentra unido al operador ( $K_{ORT}$ ), se calculó asumiendo una condición de balance detallado, al igual que en el trabajo de Tabaka et al. (2008), obteniendo un valor de:

$$K_{ORT} \approx 1,980 \text{ moléculas}$$

La constante de cooperatividad entre los represores unidos al primer y segundo operador ( $k_c$ ) fue estimada experimentalmente por Yang et al. (1996), la cual tiene un valor de

$$k_c \approx 40$$

La tasa máxima promedio de transcripción de una copia del operón *trp* ( $k_M$ ) se infirió a partir de la tasa de transcripción del operón lactosa (*lac*), la cual es 18.2 moléculas/min (Kennell y Riezman, 1977); y a partir de comparaciones entre las fuerzas (tasas de transcripción) de los promotores *trp*, *lacUV5* (variante del promotor *lac*) y *lac*. El promotor *trp* es aproximadamente 3 veces más fuerte que el promotor *lacUV5* (De Boer et al., 1983), y este último es

2.1 veces más débil que el promotor *lac* (Deuschle et al., 1986). Por lo tanto, el promotor *trp* es cerca de 1.43 veces más fuerte que el promotor *lac*, obteniendo entonces que

$$k_M \approx 26 \text{ moléculas/min}$$

La probabilidad de que la transcripción no sea interrumpida por el proceso de atenuación ( $P_A$ ) está representada en la ecuación A.16, la cual contiene dos parámetros,  $K_G$  y  $\alpha$ . Estos parámetros fueron ajustados para que la función reprodujera adecuadamente los resultados experimentales de Yanofsky et al. (1984). Estos resultados muestran que la probabilidad de que ocurra atenuación de la transcripción en bacterias cultivadas en un medio que simula ausencia de triptófano, es 6 veces menor que en bacterias cultivadas en medio mínimo, con o sin triptófano. Los valores que permiten satisfacer este comportamiento son

$$K_G \approx 1200 \text{ moléculas} \quad y \quad \alpha \approx 18.8$$

Para calcular el tiempo invertido entre el inicio de la transcripción y el inicio efectivo de la traducción ( $\tau_M$ ), debemos tomar en cuenta que la enzima antranilato sintasa (AS) es sintetizada a partir de los dos primeros genes del operón *trp* (*trpE* y *trpD*). Por lo tanto, debemos considerar el tiempo que la ARN polimerasa tarda en avanzar desde el inicio de la transcripción, hasta el codón de inicio del gen *trpD*. Sabiendo que la longitud de este segmento es de 1724 nucleótidos (nt) (Yanofsky et al., 1981), y que la ARN polimerasa avanza con una velocidad de 2700 nt/min en bacterias con tiempos de duplicación de 60 min (Bremer y Dennis, 1996), obtenemos que el tiempo que tomaría este proceso sería de 0.64 min. Por otro lado, debemos considerar que el ribosoma avanza más rápido (2880 nt/min) que la ARN polimerasa (Bremer y Dennis, 1996), por lo que debe pasar cierto tiempo para que la polimerasa no frene el avance del ribosoma sobre el ARNm. Sabiendo que la longitud de la región codificante del gen *trpD* es de 1596 nt, podemos calcular que este tiempo es de 0.04 min. Por consiguiente, si hacemos la suma entre estos dos tiempos, obtenemos que

$$\tau_M \approx 0.68 \text{ min}$$



Forchhammer et al. (1972) determinaron experimentalmente que las regiones del ARNm correspondiente a los genes *trpE* y *trpD* tienen vidas medias de 1 y 1.25 min, respectivamente. Por lo tanto, consideramos la degradación de la región con menor vida media, y el hecho de que la degradación ocurre como un proceso de decaimiento exponencial ( $M(t) = M_0 e^{-\gamma_M t}$ ). A partir de esto podemos calcular la tasa de degradación,  $\gamma_M$ , obteniendo

$$\gamma_M \approx 0.69 \text{ min}^{-1}$$

Mediciones experimentales realizadas por Bennett et al. (2009) indican que el tiempo de duplicación ( $t_2$ ) de la bacteria *E. coli* en medio mínimo con glucosa, durante crecimiento exponencial, es de alrededor 77 min. Con este dato calculamos la tasa de dilución ( $\mu = 1/t_2$ ), la cual tiene un valor de

$$\mu \approx 0.013 \text{ min}^{-1}$$

Para calcular la tasa con la que se inicia la producción de la enzima AS ( $k_E$ ), utilizamos dos datos. El primero es el del número de ribosomas que traducen cada ARNm del operón *trp*, el cual es de aproximadamente 30 ribosomas (Baker y Yanofsky, 1972). El segundo, es la vida promedio del ARNm, la cual a partir de  $\gamma_M$  obtenemos que es de 1.45 minutos. Por lo tanto,

$$k_E \approx 20.7 \text{ moléculas/min}$$

Los dos genes necesarios para sintetizar la enzima AS son traducidos independientemente. Por esta razón, el tiempo invertido en la traducción,  $\tau_E$ , estará dado por el tiempo que tarde el ribosoma en traducir el gen con la región codificante de mayor longitud. Este gen es *trpD*, cuya región codificante es de 1596 nt. A partir de esta longitud y la antes mencionada velocidad de traducción (2880 nt/min), podemos calcular que

$$\tau_E \approx 0.55 \text{ min}$$

La tasa máxima con que cada complejo enzimático produce triptófano ( $k_T$ ) fue calculada a partir del trabajo de Ito et al. (1969), obteniendo que

$$k_T \approx 300 \text{ moléculas/min}$$

Los valores de la constante de disociación entre el triptófano y la enzima ( $K_I$ ), y del coeficiente de Hill ( $n$ ) de la función de inhibición enzimática ( $P_I(T)$ ), fueron tomados del trabajo de Caligiuri y Bauerle (1991), y estos son:

$$K_I \approx 2500 \text{ moléculas} \quad y \quad n \approx 1.2$$

Para calcular la tasa máxima por minuto con la que el triptófano es utilizado para la fabricación de proteínas ( $\rho_{max}$ ), tomamos el número de moléculas de triptófano incorporadas en todas las proteínas de la célula, y lo dividimos entre el tiempo en que la célula, y por lo tanto también las proteínas, son duplicadas. Del sitio web *E. coli Statistics* obtenemos que una célula de *E. coli* contiene alrededor 2.6 millones de proteínas, cada una con una longitud promedio de 360 aminoácidos. Con este dato, y sabiendo que la abundancia de triptófano es de 1.1% del total de aminoácidos (Neidhardt, 1990), podemos calcular que habrá alrededor de 10.3 millones de moléculas de triptófano presentes en proteínas. Si el tiempo de duplicación es de 77 minutos, entonces

$$\rho_{max} \approx 134,000 \text{ moléculas/min}$$

Para la función de consumo de triptófano ( $\rho(T)$ ), un valor de

$$K_\rho \approx 1,000 \text{ moléculas}$$

permite representar que la producción de proteínas, y por lo tanto el consumo de triptófano, disminuye sólo cuando hay escasez de triptófano intracelular.

En el estado estacionario nuestro sistema alcanza niveles de enzima y triptófano cercanos a las 1000 y 4100 moléculas por célula, respectivamente. Estos valores se encuentran en el

mismo orden de magnitud que los datos obtenidos experimentalmente por Bliss et al. (1982) y Bennett et al. (2009).

# Bibliografía

- D. N. Arvidson, C. Bruce, y R. P. Gunsalus. Interaction of the *Escherichia coli* trp aporepressor with its ligand, L-tryptophan. *J Biol Chem*, 261(1):238–243, 1986.
- R. Baker y C. Yanofsky. Transcription initiation frequency and translational yield for the tryptophan operon of *Escherichia coli*. *J Mol Biol*, 69(1):89–102, 1972.
- B. D. Bennett, E. H. Kimball, M. Gao, R. Osterhout, S. J. Van Dien, y J. D. Rabinowitz. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol*, 5(8):593–599, 2009.
- S. Bhartiya, N. Chaudhary, K. V. Venkatesh, y F. J. Doyle. Multiple feedback loop design in the tryptophan regulatory network of *Escherichia coli* suggests a paradigm for robust regulation of processes in series. *Journal of the Royal Society, Interface / the Royal Society*, 3(8):383–91, June 2006.
- W. J. Blake, M. Kaern, C. R. Cantor, y J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, 2003.
- R. D. Bliss, P. R. Painter, y A. G. Marr. Role of feedback inhibition in stabilizing the classical operon. *J Theor Biol*, 97(2):177–193, 1982.
- H. Bremer y P. P. Dennis. Modulation of chemical composition and other parameters of the cell by growth rate. *Escherichia coli and Salmonella: cellular and molecular biology*, 2: 1553–1569, 1996.
- F. J. Bruggeman, N. Blüthgen, y H. V. Westerhoff. Noise management by molecular networks. *PLoS computational biology*, 5(9):e1000506, Sept. 2009.
- X. Cai. Exact stochastic simulation of coupled chemical reactions with delays. *Journal of Chemical Physics*, 2007.

- M. Caligiuri y R. Bauerle. Identification of amino acid residues involved in feedback regulation of the anthranilate synthase complex from *Salmonella typhimurium*. Evidence for an amino-terminal regulatory site. *J. Biol. Chem.*, 266(13):8328–8335, May 1991.
- P. J. Choi, L. Cai, K. Frieda, y X. S. Xie. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science*, 322(5900):442–446, 2008.
- H. A. De Boer, L. J. Comstock, y M. Vasser. The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proceedings of the National Academy of Sciences*, 80(1):21, 1983.
- U. Deuschle, W. Kammerer, R. Gentz, y H. Bujard. Promoters of *Escherichia coli*: a hierarchy of in vivo strength indicates alternate structures. *The EMBO journal*, 5(11):2987, 1986.
- Y. Dublanche, K. Michalodimitrakis, N. Kümmerer, M. Foglierini, y L. Serrano. Noise in transcription negative feedback loops: simulation and experimental analysis. *Molecular systems biology*, 2(1):41, Jan. 2006.
- M. B. Elowitz, A. J. Levine, E. D. Siggia, y P. S. Swain. Stochastic gene expression in a single cell. *Science (New York, N.Y.)*, 297(5584):1183–6, Aug. 2002.
- J. Forchhammer, E. N. Jackson, y C. Yanofsky. Different half-lives of messenger RNA corresponding to different segments of the tryptophan operon of *Escherichia coli*. *J Mol Biol*, 71(3):687–699, 1972.
- I. Golding, J. Paulsson, S. M. Zawilski, y E. C. Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, 2005.
- A. O. Grillo, M. P. Brown, y C. A. Royer. Probing the physical basis for trp repressor-operator recognition. *J Mol Biol*, 287(3):539–554, 1999.
- R. P. Gunsalus, A. G. Miguel, y G. L. Gunsalus. Intracellular Trp repressor levels in *Escherichia coli*. *J Bacteriol*, 167(1):272–278, 1986.

- E. L. Haseltine y J. B. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *The Journal of Chemical Physics*, 117(15):6959–6969, 2002.
- B. K. Hurlburt y C. Yanofsky. trp repressor/trp operator interaction. Equilibrium and kinetic analysis of complex formation and stability. *J Biol Chem*, 267(24):16783–16789, 1992.
- J. Ito, E. C. Cox, y C. Yanofsky. Anthranilate synthetase, an enzyme specified by the tryptophan operon of *Escherichia coli*: purification and characterization of component I. *J Bacteriol*, 97(2):725–733, 1969.
- E. N. Jackson y C. Yanofsky. Internal promoter of the tryptophan operon of *Escherichia coli* is located in a structural gene. *Journal of molecular biology*, 69(2):307–313, 1972.
- O. Jardetzky y M. Finucane. Tandem Interactions in the trp Repressor System may Regulate Binding to Operator DNA. *Structure and Biophysics. New Technologies for Current Challenges in Biology and Beyond*, pages 49–64, 2007.
- A. Javelle, G. Thomas, A. M. Marini, R. Krämer, y M. Merrick. In vivo functional characterization of the *Escherichia coli* ammonium channel AmtB: evidence for metabolic coupling of AmtB to glutamine synthetase. *Biochemical Journal*, 390(Pt 1):215, 2005.
- D. Kennell y H. Riezman. Transcription and translation initiation frequencies of the *Escherichia coli* lac operon. *J Mol Biol*, 114(1):1–21, 1977.
- E. Levine y T. Hwa. Stochastic fluctuations in metabolic pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 104(22):9224–9, May 2007.
- D. E. Morse, R. F. Baker, y C. Yanofsky. Translation of the tryptophan messenger RNA of *Escherichia coli*. *Proc Natl Acad Sci U S A*, 60(4):1428–1435, 1968.
- F. Neidhardt. *Physiology of the bacterial cell: a molecular approach*. 1990.

- J. Peccoud y B. Ycart. Markovian modeling of gene-product synthesis. *Theoretical population biology*, 48(2):222–234, 1995.
- A. Raj y A. van Oudenaarden. Single-molecule approaches to stochastic gene expression. *Annual review of biophysics*, 38:255–70, Jan. 2009.
- M. Santillan y E. S. Zeron. Dynamic influence of feedback enzyme inhibition and transcription attenuation on the tryptophan operon response to nutritional shifts. *J Theor Biol*, 231(2):287–298, 2004.
- M. Tabaka, O. Cybulski, y R. Holyst. Accurate genetic switch in Escherichia coli: novel mechanism of regulation by co-repressor. *J Mol Biol*, 377(4):1002–1014, 2008.
- M. Thattai y A. van Oudenaarden. Attenuation of noise in ultrasensitive signaling cascades. *Biophysical journal*, 82(6):2943–50, June 2002.
- J. Yang, A. Gunasekera, T. A. Lavoie, L. Jin, D. E. A. Lewis, y J. Carey. In vivo and in vitro studies of TrpR-DNA interactions. *Journal of molecular biology*, 258(1):37–52, 1996.
- C. Yanofsky y V. Horn. Role of regulatory features of the trp operon of Escherichia coli in mediating a response to a nutritional shift. *J Bacteriol*, 176(20):6245–6254, 1994.
- C. Yanofsky, T. Platt, I. P. Crawford, B. P. Nichols, G. E. Christie, H. Horowitz, M. Van-Cleemput, y A. M. Wu. The complete nucleotide sequence of the tryptophan operon of Escherichia coli. *Nucleic acids research*, 9(24):6647–6668, 1981.
- C. Yanofsky, R. L. Kelley, y V. Horn. Repression is relieved before attenuation in the trp operon of Escherichia coli as tryptophan starvation becomes increasingly severe. *J Bacteriol*, 158(3):1018–1024, 1984.
- H. Zalkin. Anthranilate synthetase. *Advances in enzymology and related areas of molecular biology*, pages 1–39, 1973.

D. Zenklusen, D. R. Larson, y R. H. Singer. Single-RNA counting reveals alternative modes of gene expression in yeast. *2-30 Nat Struct Mol Biol*, 15(12):1263–1271, 2008.

H. Zhang, D. Zhao, M. Revington, W. Lee, X. Jia, C. Arrowsmith, y O. Jardetzky. The solution structures of the trp repressor-operator DNA complex. *Journal of molecular biology*, 238(4):592–614, May 1994.