



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL
INSTITUTO POLITÉCNICO NACIONAL

UNIDAD MONTERREY

Análisis espectral del Ruido Bioquímico en un oscilador genético sintético

Tesis que presenta:

M. en C. Criseida Gabriela Zamora Chimal

para obtener el grado de

Doctora en Ciencias

en

Ingeniería y Física Biomédicas

Asesor: Dr. Eduardo Santillan Zeron

Apodaca, N.L. - Febrero de 2015

Contents

1	Introducción	1
2	Antecedentes	3
2.1	Generalidades	3
2.2	Ruta de la información genética	4
2.3	Dogma central de la biología molecular	4
2.4	Operón	5
2.5	Circuito genético sintético	6
2.6	Ruido Bioquímico	6
2.6.1	Importancia del Ruido bioquímico en biología.	8
3	Justificación y objetivos.	10
3.1	Justificación	10
3.2	Objetivo general y particulares	11
4	Marco teórico	12
4.1	Cinética química estocástica	13
4.2	Ecuación Química Maestra	13
4.2.1	Deducción de la ecuación química maestra (EQM)	15
4.2.2	Algoritmo de Gillespie	16
4.3	Análisis armónico de Fourier	17
5	Modelo matemático del oscilador genético simple construido por Stricker et al.	21
5.1	Parámetros del modelo	30
5.2	Modelo determinista	32
5.3	Modelo estocástico	35
6	Métodos	35
7	Resultados	39
7.1	Resultados del modelo matemático para los promotores de Lutz y Bujard	39

7.2	Rangos de regulación medidos <i>in silico</i>	40
7.2.1	Promotor P_{lac}	41
7.2.2	Promotor P_{A1lac0_1}	41
7.2.3	Promotor P_{Llac0_1}	42
7.2.4	Resultados Generales	43
7.3	Resultados de las simulaciones del oscilador sintético más simple construido por Stricker et al.	44
7.3.1	Resultados del modelo determinista	44
7.3.2	Resultados del modelo estocástico	46
7.4	Efecto de la fuerza de promotor en el periodo de oscilación	52
7.4.1	P_{lac}	53
7.4.2	P_{A1lac0_1}	55
8	Discusión y conclusiones	57
9	Anexos	60
9.1	Anexo 1. Estabilidad de la Ecuación Química Maestra	60
9.2	Anexo 2. Parámetros para el modelo matemático	63
9.2.1	Parámetros medidos in vitro para el promotor P_{Llac0_1}	63
9.2.2	Parámetros medidos in vitro para el promotor $P_{A1lac0-1}$	64
9.2.3	Parámetros medidos in vitro para el promotor P_{lac}	65
9.3	Parámetros usados para el modelo del oscilador genético sintético construido por Stricker et al.	66
	Referencias	67

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología y al pueblo de México por la beca otorgada para llevar a cabo mis estudios de doctorado en el programa de Ingeniería y Física Biomédicas del CINVESTAV Unidad Monterrey del IPN.

Agradezco al Departamento de Matemáticas del CINVESTAV Zacatenco por las facilidades para realizar mi estancia de tesis doctoral en su departamento.

1 Introducción

En años recientes, investigaciones teóricas y experimentales han estudiado los efectos del ruido bioquímico en la expresión de los genes. Este fenómeno es inherente a las reacciones bioquímicas que involucran especies químicas con bajo conteo molecular, tales como el ADN y el ARNm. Por lo tanto, los procesos de transcripción y traducción son inherentemente ruidosos. En base a lo anterior, es de interés estudiar y analizar cuál es el efecto de éste en el comportamiento dinámico de un sistema genético (operón más sus lazos de retroalimentación).

Por otro lado, el comportamiento oscilatorio ocurre en diferentes niveles de la organización biológica. Ha sido estudiado y caracterizado ampliamente en organismos superiores, en fenómenos como: el ciclo cardíaco, ritmos hormonales, ciclos circadiano y algunas oscilaciones bioquímicas, entre otros. Aunque en organismos simples no ha sido caracterizado ampliamente es de nuestro particular interés estudiar el comportamiento oscilatorio en una bacteria que tiene un circuito genético no encontrado en la naturaleza que induce un comportamiento oscilatorio.

Es por ello que en este trabajo se estudia un oscilador genético sintético simple tomando en cuenta al ruido bioquímico; ésto es, se analiza un proceso estocástico debido a la aleatoriedad de las reacciones químicas en los procesos de transcripción y traducción a través de un análisis espectral, usando técnicas de la transformada de Fourier.

El oscilador que se estudia fue implementado introduciendo dos vectores pZ diferentes (comercializado por *EXPRESSYS company*) en una cepa de *E. coli*. Un vector pZ está compuesto por una unidad promotora P_{Llac0_1} , seguido de un gen estructural $lacI$, el cual codifica para el represor LacI y este represor a su vez regula negativamente a su gen estructural vía el promotor P_{Llac0_1} . El otro vector pZ contiene un gen reportero $yemYFP$ controlado por el mismo promotor P_{Llac0_1} , este vector reporta la actividad del represor $LacI$. Este sistema simple ha sido implementado experimentalmente por medio de biología sintética por Stricker et al.

Se eligieron un conjunto de reacciones químicas para construir modelos matemáticos determinista y estocástico que reprodujeran el comportamiento oscilatorio de este circuito genético, teniendo especial cuidado en que los parámetros cinéticos de las reacciones químicas elegidas

fuesen referenciados experimentalmente. Nuestro modelo es construido primeramente con un modelo mínimo que representa la represión de la transcripción y la transcripción. Posteriormente es completado con las reacciones de traducción, dimerización y tetramerización así como las degradaciones de las especies químicas involucradas. Finalmente se comparan nuestros resultados calculados teóricamente en base a nuestro modelo con resultados reportados de forma experimental.

2 Antecedentes

2.1 Generalidades

E.coli es el organismo biológico procariota más utilizado como modelo de estudio en biología molecular y celular. Es una bacteria fácilmente manejable para la experimentación genética, y debido a esto, se ha adoptado extensamente como un organismo modelo para estudios de laboratorio. Los trabajos de laboratorio aunados a los modelos teóricos, aportan un estudio integral entre teoría y experimentación para responder preguntas biológicas, o sobre algún sistema biológico en particular que no puede ser analizado sólo experimentalmente.

La biología de sistemas es un enfoque científico innovador que busca entender cómo todos los componentes individuales de un sistema biológico interactúan en tiempo y espacio para determinar el funcionamiento del sistema. Permite la apreciación de una gran cantidad de datos de biología molecular e investigación genómica, integrándose con la fisiología para modelar la función compleja de las células, órganos y organismos completos; y con ello tener la posibilidad de mejorar nuestro conocimiento de la salud y enfermedad. Otro enfoque importante es entender la biología a nivel de un sistema, así, debemos examinar la dinámica y la estructura de la función celular y los organismos, en lugar de solamente sus características a nivel de elementos aislados. La comprensión a nivel de sistemas es el enfoque que defiende la Biología de Sistemas, el cuál requiere de un cambio en nuestra percepción sobre qué es lo que buscamos en la biología. Sin duda, la comprensión de la identidad molecular de los genes y proteínas es importante y por lo mismo es importante centrar nuestra atención en comprender la estructura de un sistema y su dinámica. Debido a que nuestro sistema no sólo consiste de los elementos que lo conforman, sus propiedades no se pueden entender con solo eso, sino que es necesario conocer como es su interacción molecular. Debemos saber cómo estos elementos moleculares se juntan y forman una estructura funcional, tal como las redes de regulación genética y su interacción bioquímica. Conocer las partes que integran un sistema nos permite conocer como está constituido, pero no nos permite conocer como una parte del sistema afecta a otra parte, lo cual es importante para entender el funcionamiento [1].

2.2 Ruta de la información genética

La unidad fundamental de información en los seres vivos es el gen. Un gen puede definirse como el segmento de ADN que codifica la información necesaria para que se sintetice un producto biológico funcional, como por ejemplo una proteína o una de las diversas clases de ARN. En las células existen dos tipos de ácidos nucleicos: el ácido desoxirribonucleico (ADN) y el ácido ribonucleico (ARN). El ácido desoxirribonucleico es una macromolécula que forma parte de todas las células. La molécula de ADN porta la información necesaria para el desarrollo de las características biológicas de un individuo y contiene los mensajes e instrucciones para que las células realicen sus funciones. El ácido ribonucleico (ARN) se forma a partir de información contenida en el ADN. En el proceso llamado transcripción, un sistema enzimático llamado polimerasa convierte la información genética de un segmento de ADN en una cadena de ARN con una secuencia de bases complementarias a las cadenas de ADN, intercambiando la timina por uracilo. Existen principalmente tres clases de ARN, ARN mensajero (ARNm) el cuál es portador de secuencias que determinan el orden en que se unirán los aminoácidos de una proteína dada; el ARN de transferencia (ARNt), el cual es un adaptador que lee la información codificada en el ARNm y transfiere el aminoácido adecuado a la cadena polipeptídica en crecimiento durante la síntesis proteica; y las moléculas de ARN ribosomal (ARNr), el cual se asocia con las proteínas formando la maquinaria (el ribosoma) para la síntesis de proteínas. Además de estos tipos de ARN existen otros, con otras funciones más específicas.

2.3 Dogma central de la biología molecular

El dogma fundamental de la biología propuesto por Crick en 1970 es la vía por la cual fluye la información genética, en la que se distinguen tres pasos principales. El primero es el proceso de la *replicación del ADN*; el segundo es la *transcripción*, en el cual las moléculas de ARN mensajero se sintetizan a partir de un molde de ADN; y el tercero es la *traducción*, en el cuál los ribosomas traducen el ARN mensajero en proteínas caracterizadas por una secuencia de aminoácidos.

La expresión genética se define como el conjunto de reacciones químicas que controlan la abundancia de los productos génicos e influye en la mayoría de los aspectos del compor-

tamiento celular. La expresión de los genes es un proceso complejo, regulado en varias etapas, se controla en la regulación de la transcripción del ADN, la regulación de la traducción del ARN mensajero, la modificación postraduccional de las proteínas, hasta otros procesos como las rutas de biosíntesis.

La regulación de la expresión genética establece un aspecto fundamental en la regulación del metabolismo celular, así como en las diferencias estructurales y funcionales de las células [2].

2.4 Operón

Cuando Jacob, Monod y colaboradores analizaron el sistema genético de la lactosa en *Escherichia coli*, los resultados de sus estudios permitieron establecer el modelo genético del Operón, que hoy en día nos permite comprender como tiene lugar la regulación de la expresión génica en bacterias. Jacob, Monod et al. publicaron sus resultados en la revista "*Comptes rendus hebdomadaires des séances de l'Académie des sciences*" en 1960; y recibieron en 1965 el Premio Nobel por estas investigaciones. Un operón se define desde entonces como una unidad genética funcional, formada por un grupo de genes capaces de ejercer una regulación de su propia expresión por medio de los sustratos con los que interaccionan las proteínas codificadas por sus propios genes u otros. Este complejo está formado por genes estructurales que codifican para la síntesis de proteínas, cuya expresión generalmente está regulada por otros dos factores de control: regiones de promotor y de operador.

Promotor: es el lugar de enlace de la ARN polimerasa. Es la región del ADN que contiene el sitio de inicio de la transcripción del ARN mensajero y los héxameros -10 y -33, por los cuales la polimerasa tiene gran afinidad. El promotor se encuentra antes de los genes estructurales, su orientación y posición dictará cuál de las dos hélices del ADN servirá como molde para sintetizar el ARN.

Región operadora: es una región del ADN (dentro o adjacente al promotor) con una secuencia que es reconocida por una proteína reguladora y controla la expresión de los genes estructurales asociados.

Genes estructurales: secuencias de ADN que tienen la información necesaria para codificar proteínas, las cuales generalmente participan como enzimas en rutas de biosíntesis.

Proteína reguladora: ésta se une a la región del operador; controla la expresión de los genes estructurales y es codificada por un gen regulador diferente, aunque en algunos casos es codificado por el mismo gen.

Inductor: sustrato o compuesto cuya presencia o ausencia induce la expresión del resto de los genes que conforman el operón. Actúa uniéndose a la proteína reguladora y activando la expresión, denominándose activador; o bien reprimiendo el gen, llamándose entonces represor [3].

2.5 Circuito genético sintético

La biología sintética es una rama de la ciencia, la cual, en uno de sus enfoques persigue el diseño y la construcción de sistemas no encontrados en la naturaleza; y que lleven a cabo funciones específicas, como por ejemplo: los así llamados "switches", osciladores genéticos, entre otros.

En base a la definición de operón, un circuito genético sintético conserva los elementos básicos ya mencionados y sus interacciones, para que con ello se lleve a cabo la expresión de los genes y el control de los mismos; pero su diseño no es encontrado de forma natural en un organismo celular, sino que es construido sintéticamente.

2.6 Ruido Bioquímico

Las reacciones químicas dentro de las células son un proceso intrínsecamente estocástico, produciendo fuertes variaciones aleatorias en el número de moléculas de las especies químicas. Este fenómeno es conocido como ruido bioquímico, originado por la naturaleza aleatoria de las reacciones químicas. El estudio del ruido bioquímico se ha centrado en conocer cuáles son las fuentes, procesamiento y consecuencias biológicas de las fluctuaciones estocásticas inherentes en las poblaciones, concentraciones o estados moleculares que controlan el comportamiento celular. Tales fluctuaciones son especialmente relevantes en sistemas con bajo conteo molecular [11].

Debido a que en una célula procariota, en general hay un número muy pequeño de genes que codifican una proteína específica, la concentración de esta proteína tiene fluctuaciones estocásticas, esto es que su número de moléculas tiene una distribución de probabilidad y

no sólo una media en su conteo molecular. A estas fluctuaciones y sus efectos se les conoce como ruido bioquímico [12].

Todo individuo en una población de organismos o células es único, aun los gemelos idénticos tienen diferentes huellas digitales. La variabilidad en las poblaciones puede ser debida a las diferencias genéticas, pero también al ambiente de desarrollo, estos dos contribuyen a la variación fenotípica celular. Diversas investigaciones se han enfocado en demostrar que esta variabilidad es inevitable en los sistemas biológicos, debido a la naturaleza aleatoria de las reacciones químicas dentro de una célula [13]. Incluso células con los mismos genes, en el mismo ambiente y con la misma historia muestran variaciones tanto físicas como en comportamiento, las cuáles pueden ser mínimas o relevantes. Por ejemplo, los gemelos idénticos tienen huellas digitales diferentes, o un gato clonado de una célula madre muestra marcadas diferencias en la distribución de colores del pelaje con respecto a su madre genética [14].

Si una reacción química involucra grandes números de moléculas, en el orden de 6.023×10^{23} , podemos hablar de concentraciones y la dinámica temporal de estas concentraciones es predecible hasta cierto punto. En las bacterias existe a lo más una docena de secuencias de DNA asociadas a un mismo gen, las cuáles generan a través de la transcripción pocas moléculas de mRNA. Por ejemplo, algunas especies de mRNA tiene de 5-10 copias por bacteria; pero todas tienen a lo más 11,000 copias por bacteria [4]. Estos bajos conteos moleculares de DNA y mRNA no nos permiten hablar de concentraciones. Con base en lo anterior, si una reacción química involucra pocas moléculas de alguna especie química, se pueden llegar a observar fluctuaciones o cambios aleatorios en la dinámica temporal del número de moléculas, tal es el caso de las especies químicas que están involucradas en la expresión de los genes, como el ADN y el ARNm. Más aun, algunas proteínas tienen conteo molecular bajo, como el represor lactosa del cual se ha medido que existen entre diez y veinte moléculas por bacteria [43].

En resumen, uno de los ejemplos clásicos de un proceso estocástico en biología es el ruido bioquímico debido a que las reacciones químicas dentro de las células son un proceso intrínsecamente estocástico y producen variaciones aleatorias en el número de moléculas de las especies químicas. Estas variaciones son más grandes en tanto que el número de moléculas involucradas en las reacciones es más pequeño, éste fenómeno es conocido como ruido bioquímico.

2.6.1 Importancia del Ruido bioquímico en biología.

Desde un punto de vista general, el ruido biológico se encuentra en la naturaleza y en los seres vivos. Por ejemplo, en las fluctuaciones de fenómenos ecológicos, en la propagación de epidemias, etc. Más aún, en organismos compuestos por células, desde los seres humanos hasta organismos unicelulares, el ruido bioquímico es relevante por sí mismo, la estocasticidad es una característica de la naturaleza [19]. La expresión genética es fundamentalmente un proceso con aleatoriedad inherente en los procesos de transcripción y traducción, lo cual ocasiona la variación de los niveles de mRNA y proteínas célula a célula aún en una población isogénica.

Los primeros intentos para caracterizar la expresión genética estocástica fueron experimentos en biología sintética, en los cuales se encontró que el comportamiento ruidoso de la expresión genética interfiere en la operación de los circuitos genéticos, tal es el caso del repressilator, el cual es un circuito que se compone por tres genes que se regulan negativamente ocasionando un comportamiento oscilatorio en las concentraciones de sus proteínas (LacI, tetR, cI). Los autores (Elowitz y Liebler, 2000) encontraron que la estocasticidad en la expresión de los genes tiene influencia en las oscilaciones [7].

En el trabajo de Elowitz et al., en el cual se cultivaron cepas de bacterias clonadas y se controlaron factores como temperatura, medio de cultivo y el mismo promotor para todas las plásmidos introducidos en ellas. Efectivamente se demostró que existen diferencias fenotípicas debido a la expresión de proteínas, aun en organismos clones. Aunado a ello se diferenciaron experimentalmente dos tipos de ruido genéticos: el intrínseco y el extrínseco [6]. Si consideramos un gen de interés, éstos tipos de ruido son caracterizados debido a las fuentes que los producen. El ruido intrínseco se refiere a la estocasticidad inherente a los procesos bioquímicos de la expresión de un gen, como la transcripción y la traducción. El ruido extrínseco a la variación fenotípica debida a otros componentes celulares, por ejemplo la ARN polimerasa. Las fuentes de ruido extrínseco son independientes del gen de interés pero actúan en él. Algunos ejemplos de fuentes de ruido extrínseco son los ribosomas o la ARN polimerasa, la cantidad de proteínas, la maquinaria que degrada al ARNm y el ambiente.

Un ejemplo *in silico* de ruido intrínseco (McAdams y Arkin, 1997) se estudia en la dinámica del interruptor fago λ , el cual puede estar en dos estados independientes llamados

lysis y lisogenia. En el estado de lisogenia, el fago λ se encuentra latente y se reproduce conjuntamente con la bacteria cuando ésta se divide por mitosis. En contraste, cuando el fago λ detecta que la bacteria puede morir, éste entra al estado de lysis y como otros virus produce tantas copias de sí mismo que la célula muere. Ellos mostraron que el ruido en la expresión de las proteínas reguladoras de éste sistema tiene un efecto para que algunas células activen la vía lítica mientras otras sigan la vía lisogénica [20].

Los eventos estocásticos de una sola molécula tiene importantes consecuencias biológicas, tales como la determinación del fenotipo de una célula o la producción de proteínas en ráfagas "bursting", ésto fue demostrado por Ji Yu et al (2006) [22], quienes experimentalmente monitorearon la producción de proteínas individuales en tiempo real, demostrando los burst traduccionales de un promotor reprimido.

Estudios realizados por Xie y colaboradores describen de forma experimental y cuantitativa que la expresión de los genes en bacterias es estocástica [16]. Cai et al. demuestra que es posible la observación en tiempo real de la expresión de la enzima β -galactosidasa en una sola célula de *E. coli* con una resolución de una sola molécula. Implementan un ensayo microfluídico en tiempo real y observan que la expresión de las proteínas ocurre en estallidos (burst) y tiene un bajo conteo molecular (0-20 copias de β -galactosidasa por burst). El número de proteínas por burst tiene una distribución exponencial. Aplicaron éste ensayo en células madre de ratón y en levadura. Éste ensayo microfluídico abre posibilidades para la caracterización de todo el sistema de la expresión de estas proteínas con bajo conteo molecular [15].

3 Justificación y objetivos.

3.1 Justificación

El ruido bioquímico es un fenómeno inherente a los sistemas genéticos, debido a los procesos involucrados en la expresión genética y al bajo conteo molecular de las especies químicas ADN y ARN. Éste fenómeno se debe a la aleatoriedad intrínseca de las reacciones químicas que se llevan a cabo en procesos como la transcripción y la traducción.

Por otra parte, el comportamiento oscilatorio en organismo superiores ha sido estudiado y caracterizado en diferentes procesos fisiológicos, tales como el ciclo cardiaco, los ciclos circadianos, oscilaciones del potencial de membrana, entre otros. Sin embargo, en organismos simples como las células procariotas o bacterias aún no ha sido tan ampliamente caracterizado, aunque se sabe que existen ciclos circadianos en cianobacterias.

Es nuestro interés analizar que sucede con el comportamiento oscilatorio en un organismo procariota, cuando el sistema de reacciones es visto como un proceso estocástico debido al ruido bioquímico. Así, al ser un proceso estocástico que tiene un comportamiento oscilatorio llevamos a cabo un análisis en el espacio de frecuencias usando técnicas de la transformada de Fourier. Primeramente se desarrolla un modelo determinista y se calcula su frecuencia de oscilación. Como nuestro sistema es un proceso estocástico, el periodo de oscilación es una variable aleatoria, por lo que implementamos una técnica para calcular el periodo de oscilación a través de la función de correlación y la transformada de Fourier. Uno de nuestros principales objetivos es analizar si el ruido bioquímico tiene algún efecto en el comportamiento oscilatorio. Por lo anterior, elegimos un oscilador genético sintético que se haya implementado *in vivo* y sea lo suficientemente simple para que sus parámetros de cinética química sean conocidos y finalmente comparar nuestros resultados teóricos con los reportes experimentales.

3.2 Objetivo general y particulares

Objetivo General Estudiar los efectos del ruido bioquímico en un oscilador genético sintético *in vivo* a través de su análisis espectral.

Objetivos Particulares

1. Desarrollar un modelo matemático de un oscilador genético sintético que se haya implementado *in vivo* y lo suficientemente simple para que la mayoría de sus parámetros de cinética química hayan sido determinados experimentalmente.
2. Validar el modelo matemático con base en resultados experimentales.
3. Implementar una técnica efectiva y práctica para calcular el espectro de potencia de un proceso estocástico y aplicarla al oscilador genético arriba indicado.
4. Comparar las frecuencias de resonancia medidas *in vivo* del oscilador genético sintético con las simulaciones del modelo determinista (ODE) y las del modelo estocástico (algoritmo de Gillespie).
5. Analizar si el ruido bioquímico tiene algún efecto en la frecuencia de resonancia.

4 Marco teórico

Un sistema biológico puede ser visualizado por un conjunto de reacciones químicas que representan las interacciones moleculares participantes en él. A partir de esto, existen dos caminos para representar a las reacciones químicas en ecuaciones, la modelación determinista y la estocástica. Estos dos tipos de modelado tienen diferentes alcances y se usan dependiendo de la pregunta que queramos contestar en relación a nuestro sistema biológico. Para un análisis general de un sistema biológico, estos métodos son complementarios. En la modelación determinista se usan ecuaciones diferenciales ordinarias, con retardo o parciales, mientras que en la modelación estocástica se usa la ecuación química maestra, algoritmos de simulación estocástica como por ejemplo el algoritmo de Gillespie, entre otros. Sin embargo, es necesario remarcar que la única manera para estudiar al ruido bioquímico es usando modelación estocástica.

En los modelos deterministas no se puede incluir el ruido bioquímico dentro de sus ecuaciones, pero sí en la Ecuación Química Maestra (EQM) y el algoritmo de Gillespie. Estos últimos proporcionan diferente información del mismo proceso. La EQM determina la evolución temporal de las distribuciones de probabilidad del conteo de moléculas de las especies químicas y el algoritmo de Gillespie la evolución de las trayectorias (realización del procesos estocástico). Una trayectoria muestra el cambio del número de moléculas de una especie química a través del tiempo, toman valores en los números naturales pues describen conteos moleculares [24].

Las trayectorias calculadas con el algoritmo de Gillespie, están asociadas a la EQM. A partir de una misma condición inicial, se generan diferentes trayectorias (alrededor de un millón), se registra el número de moléculas en un tiempo fijo t_0 y con esta información se genera una distribución de probabilidad del conteo molecular en el tiempo t_0 , esta distribución de probabilidad es la misma que se calcula como solución de la EQM.

De igual forma, las ecuaciones de Langevin y Fokker Planck proporcionan diferente información del mismo proceso. La ecuación de Langevin determina la evolución temporal de las trayectorias y la ecuación de Fokker Planck la de las distribuciones de probabilidad. Las trayectorias calculadas son continuas, tomando valores en los reales [25].

En éste trabajo se usará la ecuación química maestra (EQM) y el algoritmo de Gillespie,

ya que en un primer paso de nuestro análisis necesitamos calcular el conteo molecular de las especies químicas.

4.1 Cinética química estocástica

La cinética química estocástica es utilizada para describir la evolución en el tiempo de un reactor químico bien agitado y homogéneo, en el cual las reacciones químicas sucedan de manera discreta y aleatoria. En un sistema que contiene reacciones químicas aleatorias sus ecuaciones de cinética enzimática se pueden escribir en términos de propensiones, bajo el supuesto que cada propensión mide la probabilidad por unidad de tiempo de que se lleva a cabo una reacción química determinada.

4.2 Ecuación Química Maestra

El ruido bioquímico puede ser estudiado a través de la ecuación química maestra (EQM), la cuál es una ecuación diferencial lineal que rige la evolución temporal de la probabilidad $P(x, t)$ de que el sistema esté en el estado $X(t)$ al tiempo $t > 0$, donde $X(t)$ es un vector que contiene el conteo molecular de todas las especies moleculares del sistema en el tiempo t [24].

Considere un sistema químico bien agitado con n diferentes especies químicas, las cuáles indexamos con el entero $j = 1, 2, \dots, n$ y con m diferentes reacciones químicas involucradas e indexadas por el entero $k = 1, 2, \dots, m$.

El estado del sistema es representado por un vector $x(t) \in \mathbb{N}^n$, cuya entrada $x_j(t)$ denota al conteo molecular de la j ésima especie química al tiempo t , para toda $j = 1, 2, \dots, n$.

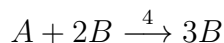
Cada una de las reacciones que se lleva a cabo en este reactor químico está indexada por un entero $k = 1, 2, \dots, m$; y entonces la k ésima reacción puede ser presentada como:



donde $a(k, x) \geq 0$ es la propensión de la reacción química y $v[k] \in \mathbb{Z}^n$ es el vector estequiométrico, el cuál indica cuantas moléculas de cada especie química se crearon o destruyeron en la reacción.

La función $a(k, x)$ es obviamente cero si cualquier entrada de los vectores x o $x + v[k]$ es negativa, pues no puede existir un conteo molecular negativo.

Por ejemplo, considere la siguiente reacción química:



Donde:

A y $2B$ son las especies químicas reactantes.

$3B$ es el producto de la reacción.

4 es la constante de reacción

Basándonos en (1) podemos reescribir la reacción anterior de la siguiente forma:

$$(x_A, x_B, \dots) \xrightarrow{2x_A x_B (x_B - 1)} (x_A - 1, x_B + 1, \dots). \quad (2)$$

Donde:

x_A, x_B representan el número de moléculas de la especie A y B , respectivamente. El lado derecho de la reacción es un vector $v[k]$ que contiene el número de moléculas que se sintetizan o se degradan. Mientras que $2x_A x_B (x_B - 1)$ es la propensión de que se lleve a cabo la reacción.

El movimiento de las moléculas en solución es de naturaleza aleatoria y por ende las reacciones químicas en solución son un proceso estocástico. Sin embargo, nosotros no somos capaces de observar en general esta aleatoriedad cuando el conteo molecular de cada especie química es del orden del número de Avogadro (6.023×10^{23}), ya que la mayoría de nuestros instrumentos no tienen todavía la resolución necesaria. Esta aleatoriedad sí es evidente cuando una de las especies químicas principales tiene un bajo conteo molecular, por ejemplo menos de 100 moléculas. La formulación estocástica de la cinética química es una forma de analizar el hecho que las colisiones en un sistema de moléculas en solución y en equilibrio térmico se producen de una manera esencialmente aleatoria. Por simplicidad se supone que todas las reacciones químicas son instantáneas. Debido a que el tiempo de espera entre dos reacciones químicas consecutivas es aleatorio, nuestro sistema es un proceso estocástico a tiempo continuo. Como el conteo molecular solo puede cambiar en cantidades discretas, nuestro sistema es un proceso estocástico de tiempo continuo con estados discretos.

Dado que el número de moléculas es finito, podemos suponer la existencia de un número natural $\sigma \gg 1$ tal que el conteo molecular $x_j(t)$ de la j ésima especie química sea un elemento

del conjunto

$$x_j(t) \in \Sigma := \{0, 1, 2, \dots, \sigma\}.$$

Por otro lado, al describir a nuestro sistema como un proceso estocástico, se puede calcular la probabilidad $P(x, t)$ que nuestro sistema esté en el estado $x \in \Sigma^n$ al tiempo $t > 0$.

La evolución temporal de la probabilidad $P(x, t)$ se rige por un sistema lineal de ecuaciones diferenciales llamado la ecuación química maestra (EQM):

$$\frac{\partial P(x, t)}{\partial t} = \sum_{k=1}^m [a(k, x - v[k])P(x - v[k], t) - a(k, x)P(x, t)] \quad (3)$$

4.2.1 Deducción de la ecuación química maestra (EQM)

Supongamos que $x(t)$ es el estado del sistema en el tiempo $t > t_0$, el valor de cada función de propensión $a(k, x(t)) > 0$, para un índice k fijo, se interpreta como sigue: para todo incremento temporal dt , el producto $a(k, x(t))dt$ es la probabilidad de que la k -reacción

$$x(t) \xrightarrow{a(k, x(t))} x(t) + v[k] \quad (4)$$

sucediera en el intervalo de tiempo infinitesimal $[t, t + dt)$. La ecuación química maestra (EQM) puede ser derivada tomando en consideración las leyes de adición y multiplicación de probabilidades. Así, calculamos $P(x; t + dt)$ como la suma de todas las probabilidades de las $m + 1$ diferentes formas posibles en las cuales el sistema puede llegar al estado $x(t + dt)$.

Las $m + 1$ formas de llegar al estado $x(t + dt)$ son las siguientes:

$$\begin{aligned} x(t) &\rightarrow x(t + dt) && \text{con probabilidad} && 1 - \sum_{k=1}^m a(k, x)dtP(x, t), \\ x(t) - v[1] &\rightarrow x(t + dt) && \text{con probabilidad} && a(1, x - v[1])dtP(x - v[1], t), \\ x(t) - v[2] &\rightarrow x(t + dt) && \text{con probabilidad} && a(2, x - v[2])dtP(x - v[2], t), \\ x(t) - v[m] &\rightarrow x(t + dt) && \text{con probabilidad} && a(m, x - v[m])dtP(x - v[m], t). \end{aligned}$$

Nótese que la suma de todas las probabilidades arriba indicadas es igual a uno. La pro/-ba/-bi/-li/-dad de estar en el estado $x(t + dt)$ al tiempo $t + dt$ se calcula entonces como:

$$\begin{aligned} P(x, t + dt) &= P(x, t) \left[1 - \sum_{k=1}^m a(k, x)dt \right] \\ &+ \sum_{k=1}^m a(k, x - v[k])dtP(x - v[k], t). \end{aligned}$$

El primer término, en la parte derecha de la ecuación anterior representa la probabilidad de que no ocurra ninguna reacción en el intervalo $[t, t + dt)$ y el segundo término representa la probabilidad de que ocurra la k ésima reacción en el intervalo $[t, t + dt)$, comenzando en el estado $x - v[k]$ y finalizando en el estado x [27].

A partir de la ecuación anterior podemos obtener la ecuación diferencial que determina la dinámica temporal de la probabilidad $P(x, t)$. Agrupando términos, dividiendo por dt y calculando el límite cuando dt converge a cero, obtenemos la ecuación química maestra (EQM):

$$\begin{aligned} \frac{\partial P(x, t)}{\partial t} &= \lim_{dt \rightarrow 0} \frac{P(x, t + dt) - P(x, t)}{dt} \\ &= \sum_{k=1}^m \left[a(k, x - v[k])P(x - v[k], t) - a(k, x)P(x, t) \right] \end{aligned} \quad (5)$$

Es fácil verificar que la ecuación química maestra (EQM) es una ecuación diferencial lineal y estable (Ver Anexo 1)[28].

4.2.2 Algoritmo de Gillespie

Por medio de éste algoritmo se puede llevar a cabo una simulación de naturaleza estocástica (a tiempo continuo y estados discretos) de un sistema con reactantes químicos. Cuando se hace una simulación de la evolución temporal de las especies químicas participantes, se obtiene una realización temporal de la posible evolución de los conteos moleculares de todas las especies químicas participantes.

El algoritmo de Gillespie consta de los siguientes pasos. Considerando que el sistema está en el estado $X(t) \in \mathbb{N}^n$ al tiempo t , las condiciones iniciales se definen en el primer paso, y de ahí en adelante se calcula la evolución del sistema.

- 1 Inicializar con el tiempo $t = t_0$ y el estado del sistema $X(t) = x_0(t)$, siendo $X(t)$ un vector que contiene número de moléculas inicial de cada especie química participante en el tiempo t .
- 2 Dado el estado $X(t)$ y el tiempo t , evaluar las propensiones del sistema $a(k, x(t))$ de acuerdo a (4) y su suma $a_0(x(t)) = \sum_{k=1}^m a(k, x(t))$. Generar el tiempo de espera $\tau = \frac{1}{a_0(X)} \ln \frac{1}{r_1}$ donde r_1 es un número aleatorio con distribución uniforme en el intervalo $[0, 1]$.

- 3 Generar un número aleatorio con distribución uniforme r_2 en el intervalo $[0, 1]$ y elegir el índice j , para conocer la reacción asociada que se producirá en el siguiente paso temporal. Para ello, el índice debe cumplir la siguiente condición: $\sum_{k=1}^{m-1} a_k < r_2 a_0 \leq \sum_{k=1}^m a_k$.
- 4 Actualizar el estado $X(t + \tau) = x(t) + v[k]$, donde $v[k]$ es el cambio en el número de moléculas causada por la k ésima reacción. Además actualizar el tiempo $t + \tau \mapsto t$.
- 5 Registrar los datos del sistema actualizado. Regresar al paso 2 o terminar la simulación si el tiempo actual t es mayor que el tiempo de simulación.

4.3 Análisis armónico de Fourier

Una variable aleatoria $v(\omega)$ puede ser vista como algún valor medido de una variable en un experimento ω y Ω es un conjunto universal que contiene a todos los experimentos posibles. Así $\omega \in \Omega$ representa un experimento en particular sin ambigüedad alguna.

Siguiendo el modelo anterior podemos ver un proceso estocástico $X^\omega(t)$ como una función que depende del tiempo $0 \leq t \leq T$ y que resulta de hacer una medición en algún experimento indexado por la variable universal ω . Así, $X^\omega(t)$ representa el valor medido en el tiempo de una variable en un experimento $\omega \in \Omega$ con respecto al tiempo.

Como estamos interesados en conocer el comportamiento oscilatorio de $X^\omega(t)$ uno puede calcular entonces la transformada de Fourier de $X^\omega(t)$ con respecto al tiempo $0 \leq t \leq T$; y así obtener un nuevo proceso estocástico $Z^\omega(s)$ que depende de la frecuencia s y de la misma variable universal ω . Por lo anterior, decimos que el espectro de un proceso estocástico en el tiempo $X^\omega(t)$ es otro nuevo proceso estocástico en la frecuencia $Z^\omega(s)$, pues $Z^\omega(s)$ representa el espectro de una medida hecha en el experimento en particular $\omega \in \Omega$. Para obtener información del espectro podemos calcular sus momentos probabilísticos principales como la esperanza, varianza y correlación.

Si consideramos a $X^\omega(t)$ como un proceso estocástico, los coeficientes $c^\omega(n)$ de su serie de Fourier también son un proceso estocástico. Pues para $T > 0$, $c^\omega(n)$ está dado por la conocida fórmula:

$$c^\omega(n) = \frac{1}{T} \int_{t=0}^T X^\omega(t) \exp \left[\frac{-2\pi i n t}{T} \right] dt \quad \forall n \in \mathbb{Z} \text{ entero.}$$

La esperanza de los coeficientes se calcula fácilmente. Como la esperanza es un operador lineal, ésta conmuta con la integral:

$$E(c^\omega(n)) = \int_{t=0}^T \frac{E(X^\omega(t))}{T} \exp \left[\frac{-2\pi i n t}{T} \right] dt$$

Recordemos que la ecuación química maestra es en particular asintóticamente estable (ver anexo A), así podemos suponer que existe un tiempo muy grande $t_0 \gg 1$ para el cual la probabilidad $P(x, t) = P^*(x)$ es independiente del tiempo $t > t_0$ e igual a la probabilidad en estado estacionario. Lo anterior implica que la esperanza $E(X^\omega(t))$ es constante cuando $t > t_0$, pues:

$$E(X^\omega(t)) = \int x P(x, t) dx = \int x P^*(x) dx.$$

y el último término no depende del tiempo t .

Más aun, la esperanza de los coeficientes $c^\omega(n)$ es entonces igual a:

$$\begin{aligned} E(c^\omega(n)) &= \int_{t=t_0}^{t_0+T} \frac{E(X^\omega(t))}{T} \exp \left[\frac{-2\pi i n t}{T} \right] dt \\ &= \begin{cases} E(X^\omega(t)) & \text{si } n = 0, \\ 0 & \text{si } n \in \mathbb{Z} \text{ y } n \neq 0. \end{cases} \end{aligned}$$

dado que $E(X^\omega(t)) = \int x P(x, t) dx = \int x P^*(x) dx$ cuando $t > t_0$ [29].

La esperanza no nos proporciona mucha información con respecto a los coeficientes estadísticos de $X^\omega(t)$, pero la varianza si pues recordemos que en el caso que nos ocupa hay colas pesadas o fluctuaciones fuertes.

Recordemos que:

$$\sigma^2(c^\omega(n)) = E(|c^\omega(n)|^2) - |E(c^\omega(n))|^2$$

Al término $E(c^\omega(n))$ ya lo conocemos, y por el otro lado (aceptando que \bar{c} es el complejo

conjugado de $c \in \mathbb{C}$)

$$\begin{aligned}
E(|c^\omega(n)|^2) &= E(c^\omega(n)\overline{c^\omega(n)}) = \\
&= E\left[\left(\int_{t=t_0}^{t_0+T} \frac{X^\omega(t)}{T} \exp\left[\frac{-2\pi int}{T}\right] dt\right) \times\right. \\
&\quad \left.\times \left(\int_{s=t_0}^{t_0+T} \frac{X^\omega(s)}{T} \exp\left[\frac{2\pi ins}{T}\right] ds\right)\right] = \\
&= E\left(\int_{t=t_0}^{t_0+T} \int_{s=t_0}^{t_0+T} \frac{X^\omega(t)X^\omega(s)}{T^2} \exp\left[\frac{2\pi in(s-t)}{T}\right] dsdt\right) \\
&= \int_{t=t_0}^{t_0+T} \int_{s=t_0}^{t_0+T} \frac{E(X^\omega(t)X^\omega(s))}{T^2} \exp\left[\frac{2\pi in(s-t)}{T}\right] dsdt
\end{aligned}$$

En conclusión, la esperanza $E(|c^\omega(n)|^2)$ del n -ésimo armónico frecuencial (coeficiente $c^\omega(n)$) es igual a la transformada de Fourier de la covarianza $E(X^\omega(t)X^\omega(s))$ con respecto a los tiempos s y t .

Reescribiendo las fórmulas anteriores tenemos que para $t_0 \gg 1$ suficientemente grande:

$$\begin{aligned}
\sigma^2(c^\omega(0)) &= E(|c^\omega(0)|^2) - |E(X^\omega(t))|^2 \\
&= \int_{t=t_0}^{t_0+T} \int_{s=t_0}^{t_0+T} \frac{E(X^\omega(t)X^\omega(s))}{T^2} dt ds - |E(X(t_0))|^2
\end{aligned}$$

y si $n \in \mathbb{Z}$ con $n \neq 0$:

$$\begin{aligned}
\sigma^2(c^\omega(n)) &= E(|c^\omega(n)|^2) \\
&= \int_{t=t_0}^{t_0+T} \int_{s=t_0}^{t_0+T} \frac{E(X^\omega(t)X^\omega(s))}{T^2} \exp\left[\frac{2\pi in(s-t)}{T}\right] dsdt
\end{aligned}$$

La varianza $\sigma^2(c^\omega(n))$ con respecto al n -ésimo armónico frecuencial con $n \neq 0$ es igual a la transformada de Fourier de la covarianza $E(X^\omega(t)X^\omega(s))$ con respecto a los tiempos s y t .

El problema se reduce entonces a calcular la covarianza $E(X^\omega(t)X^\omega(s))$ con respecto a los tiempos s y t mayores o iguales que t_0 , la cual puede ser calculada como el promedio:

$$E(X^\omega(t)X^\omega(s)) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N X_i(t)X_i(s)$$

dónde $X_i(t)$ es un valor medido de una variable dada del i -ésimo experimento, suponiendo que se han realizado $N \gg 1$ experimentos bajo las mismas condiciones.

P_{lac} TTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGG \tilde{A} ATTGTGAGC \dot{G} GATAACAATT
 $P_{LtetO-1}$ TCCTATCA \dot{G} TGATAGAGATTGACATCCTATCA \dot{G} TGATAGAGATACTGAGCAC \tilde{A} TC
 $P_{A1lacO-1}$ TTGACTTGTGAGC \dot{G} GATAACAATGATACTTAGATTC \tilde{A} ATTGTGAGC \dot{G} GATAACAATT
 $P_{LlacO-1}$ ATAAATGTGAGC \dot{G} GATAACATTGACATTGTGAGC \dot{G} GATAACAGATACTGAGCAC \tilde{A} T

Figure 1: Secuencias (de pares de bases) de los promotores diseñados por Lutz y Bujard y el promotor P_{lac} . Las unidades -10 and -33 son hexámeros que reconoce la polimerasa para iniciar la transcripción y estan marcados con rectángulos, la letra \tilde{A} con una tilde marca el sitio de inicio de la transcripción. Las secuencias correspondientes de los operadores lac y tet (sitios de unión para el represor $LacI$ y $TetR$) están subrayadas; tienen una longitud entre 18 y 21 pares de bases, y sus pares de bases centrales están indicados con un punto arriba de la letra \dot{G} . Los promotores P_{lac} , P_{A1lacO_1} y P_{LlacO_1} incluyen al operador lac , mientras que el P_{LtetO_1} incluye al promotor tet . Los pares de bases centrales en los promotores P_{A1lacO_1} , P_{LlacO_1} y P_{LtetO_1} están separados por 32, 22 y 24 pares de bases [30].

5 Modelo matemático del oscilador genético simple construido por Stricker et al.

En 1997 Lutz y Bujard desarrollaron los promotores controlables P_{lac} , P_{A1lacO_1} , P_{LlacO_1} y P_{LtetO_1} en *E. coli*, cuyas secuencias de pares de bases se muestran en la figura 1. Los promotores sintéticos P_{LlacO_1} y P_{LtetO_1} tiene dos secuencias operadoras lac y tet respectivamente, los cuales se ubican adyacentemente a los hexámeros -10 y -33 en el promotor fago lambda P_L (ver figura 1), a éstos promotores se les pueden unir los represores $LacI$ y $TetR$, según corresponda, y son a su vez inducidos por el *IPTG* y la anhidrotetraciclina [30].

EXPRESSYS es una compañía que comercializa los así llamados vectores pZ . Estos vectores contienen un modulo con uno de los cinco promotores sintéticos diseñados por Lutz y Bujard, seguidos por un sitio de unión a ribosoma y un sitio de clonado múltiple comúnmente ocupado por el gen $lacZ$ o el gen $luciferasa$, aunque también puede ser ocupado por un gen

de interés. Los vectores pZ tienen un segundo modulo que contiene los posibles orígenes de replicación y un tercer módulo con uno de los cinco marcadores de resistencia a antibiótico [31]. Los vectores pZ han sido utilizado en diferentes experimentos *in vivo*, como por ejemplo por Elowitz MB y Leibler S. (2000), Elowitz MB et al. (2002), Sara Hooshangi et al. (2005), Stricker et al. (2008), entre otros [6] [7] [8] [9]. En particular Stricker et al. usaron éstos vectores para implementar dos osciladores genéticos.

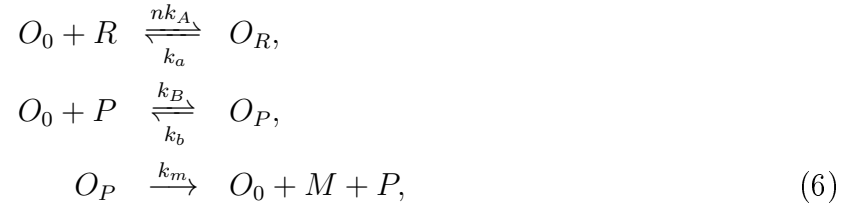
Stricker et al. (2008) construyeron *in vivo* dos osciladores genéticos sintéticos usando los vectores pZ . Uno de los osciladores tiene dos lazos de retroalimentación, uno positivo y otro negativo (lazos que compiten entre ellos), mientras que el segundo es más simple y tiene un sólo lazo de retroalimentación negativa. En éste trabajo analizamos el oscilador más simple debido a que su simplicidad significa que la frecuencia de resonancia y la existencia de oscilaciones son muy sensibles a las variaciones en los parámetros, construimos un modelo matemático del oscilador simple (lazo de retroalimentación negativa) construido por Stricker et al.

El modelo matemático se describe en dos partes. Primeramente se desarrolla un modelo matemático que involucra las interacciones del promotor con la polimerasa y el represor $LacI$, en este modelo se describen las transiciones de estados del operador y la dinámica de la represión en el promotor. Posteriormente este primer modelo se expande como un modelo que represente al oscilador genético sintético de Stricker et al. considerando un conjunto de reacciones químicas que involucren a la represión de la transcripción, la síntesis de ARNm, la traducción de proteínas, la dimerización del represor lactosa para la formación de dímeros $LacI$, la tetramerización para la formación del represor $LacI$ (tetrámero) y un lazo de autoregulación negativa del represor $LacI$; ver figura 4 para más detalles.

Las secuencias mostradas en la figura 1 describen la estructura de los diferentes promotores diseñados por Lutz y Bujard. Los operadores lac y tet están localizados junto al sitio de inicio de la transcripción (río abajo), junto al hexamero -33 o entre el hexamero -10 y -33. Los hexámeros -33 y -10 son el principal sitio de unión de la polimerasa, así cuando la polimerasa está unida al promotor se da inicio a la transcripción. Por otro lado, cuando un represor $LacI$ o $TetR$ está unido al sitio operador situado en el promotor, la polimerasa no se une al promotor y se reprime el inicio de la transcripción. La polimerasa y el represor se excluyen físicamente uno del otro por lo que existe una competencia por unirse a la secuencia

promotora.

Con el objetivo de modelar la competencia entre la polimerasa y los represores por el promotor, se definen los siguientes estados del operador: O_0 es el estado donde el operador se encuentra libre, no hay polimerasa ni represor; O_P es el estado en el cual la polimerasa P está unida al promotor; O_R cuando el operador está ocupado por un represor R , el cuál puede ser $LacI$ o $TetR$. Las reacciones químicas que describen las transiciones entre estos estados se muestran a continuación:



donde M denota un ARN mensajero naciente, k_m es la tasa de inicio de la transcripción y los parámetros k_A , k_a , k_B y k_b son las tasas de asociación y disociación de los complejos *represor LacI-operador* y *polimerasa-promotor*. La constante n denota el número de operadores activos localizados en la parte central del promotor, por lo que $n = 1$ para el promotor P_{lac} y $n = 2$ para los promotores P_{LlacO_1} y P_{A1lacO_1} . Si $n = 2$, existen dos operadores en el promotor y la constante cinética de asociación nk_A es el doble de la constante original k_A , dado que el operador tiene dos opciones en las cuales puede unirse. La transcripción se supone como una reacción instantánea debido a que la traducción del ARNm naciente comienza en cuanto los dos primeros segmentos de la región líder son transcritos, pues éstos son los sitios de unión al ribosoma.



En la reacción anterior existe una transición de estados, del estado O_P al estado libre de polimerasa y represor O_0 . Por lo cual, la reacción de la transcripción se toma en cuenta en el conjunto de reacciones que modelan las transiciones de estado del operador. La reacción comienza con una polimerasa P unida al operador (cuando aún no se abre la cadena de ADN), después la polimerasa abre la cadena de ADN, libera los sitios promotores (estado O_0), transcribe el ARNm naciente (M) en el proceso de elongación y finalmente se separa del ADN.

Debido a que los promotores P_{LlacO_1} y P_{A1lacO_1} contienen dos operadores, debemos considerar el caso cuando dos represores están unidos a operador (representado por el estado O_{2R}). La reacción química asociada es la siguiente:



Podemos suponer que la constante k_2 es mucho menor que la constante k_A , debido a que la separación entre los operadores es muy pequeña, sus pares de bases centrales (indicados con un punto en la figura 1) están separados por 32 pb o menos de acuerdo a la topografía presentada en la figura 1. Sin embargo, un represor unido al operador interfiere físicamente con la unión de un segundo represor a un operador adyacente. Por lo tanto, si consideramos el tamaño (1080 pares de bases) y la forma en V del represor *LacI* [54], podemos suponer que el estado O_{2R} no sucede en los promotores P_{LlacO_1} y P_{A1lacO_1} .

Se dice que los operadores *lac* presentan cooperatividad cuando un tetrámero *LacI* se puede unir simultáneamente a dos operadores formando un loop de represión [32], ya que de ésta forma se ayudan para producir una represión grande. Müller, Oeheler et al. han determinado que para separaciones entre 57 y 98 pares de bases, la cooperatividad entre dos operadores *lac* es máxima solo cuando los pares de bases centrales están separados por 57, 70, 81 y 92 pares de bases (ver figura 2). La regularidad en la serie de números anterior puede deberse a que dos operadores deben estar alineados en el mismo lado de la hélice de ADN; y la hélice de ADN presenta un giro completo aproximadamente cada 11 pares de bases [44]. Por consiguiente, podemos esperar que la cooperatividad también sea máxima cuando los pares de bases centrales están separadas por distancias en la secuencia 26, 37 ó 48 pb. Como los operadores en P_{LlacO_1} y P_{A1lacO_1} están separados por 32 y 22 pares de bases respectivamente podemos suponer que no hay cooperatividad entre los operadores de los promotores P_{LlacO_1} y P_{A1lacO_1} .

Con base en lo anterior se consideran los siguientes estados para el promotor: O_0 , O_P y O_R . Haciendo uso de notación, considere a $P(O_0)$, $P(O_P)$ y $P(O_R)$ como las probabilidades de que el operador esté libre, formando el complejo represor/operador y formando el complejo polimerasa/promotor, respectivamente. El estado de equilibrio para las reacciones en (6),

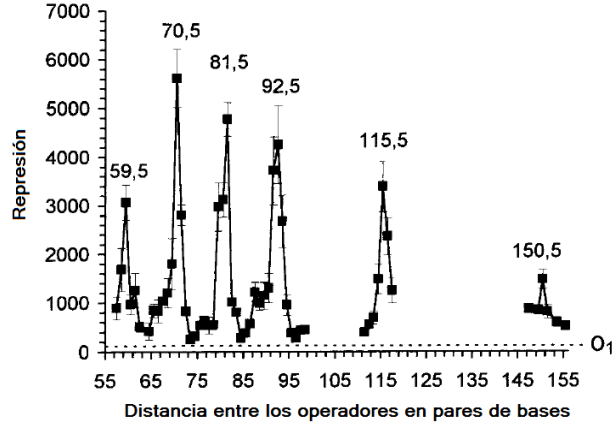


Figure 2: Represión de los operadores en función de la distancia entre ellos medida en pares de bases. Figura obtenida del trabajo de Müller et al. [54]

está determinado por las siguientes ecuaciones:

$$\begin{aligned} P(O_R) &= \frac{nk_A}{k_a} P(O_0), \\ P(O_P) &= \frac{nk_B[P]}{k_b + k_m} P(O_0). \end{aligned} \quad (9)$$

Las probabilidades $P(O_P) + P(O_R) + P(O_0) = 1$, podemos calcular fácilmente la siguiente ecuación después de dividir por $P(O_P)$:

$$\frac{1}{P(O_P)} = 1 + \frac{nk_A}{k_B} \left(\frac{k_b + k_m}{k_a} \right) \frac{[R]}{[P]} + \left(\frac{k_b + k_m}{k_B[P]} \right) \quad (10)$$

El último término $\frac{k_b + k_m}{k_B[P]}$ es muy pequeño con respecto al 1, pues es del orden de 10^{-3} ; así que lo podemos eliminar. Lo anterior podemos comprobarlo sustituyendo los valores de los parámetros k_b , k_m , k_B y $[P]$ indicados en el anexo 2. Por ejemplo, en el caso del promotor $P_{LlacO-1}$, los parámetros toman los valores $k_B = 7 \times 10^9 / M/s$, $k_b = 1.4 \times 10^7 / s$, $k_m = 2 / (60s)$ y $[P] = 8300 \text{ mpb} = 4.61 \times 10^{-6} M$, así que la fracción $\frac{k_b + k_m}{k_B[P]}$ es igual a 5.2×10^{-3} .

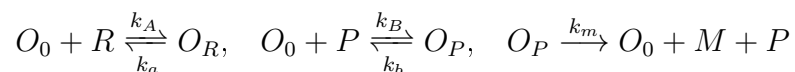
La probabilidad de que el promotor forme el complejo *polimerasa/promotor* es aproximadamente igual a uno cuando el número de represores $[R]$ es igual a cero, y así la ecuación (10) nos permite calcular el rango de regulación del promotor. Esta ecuación se usará con más detalle en la sección de resultados. El rango de regulación está definido como el cociente de la expresión del reportero (o fuerza del promotor) medida en condiciones de inducción dividido entre la expresión medida en condiciones de represión máxima (ver Tabla 1 en [30]).

El modelo arriba desarrollado para los promotores de Lutz y Bujard puede ser usado como un bloque para construir modelos de sistemas más complejos, por lo que en primera instancia se construye un modelo de un oscilador sintético genético implementado *in vivo* por Stricker et al. Este oscilador sintético está basado en el promotor P_{Llac0_1} y tiene un lazo de retroalimentación negativa (ver figura 3).

Stricker et al. implementaron *in vivo* dos osciladores genéticos sintéticos usando los vectores pZ comercializados por la compañía EXPRESSYS (esta compañía ofrece herramientas y reactivos para la expresión de genes en huéspedes procariontes basada en el sistema de expresión pZ). Uno de los osciladores tiene dos lazos de retroalimentación, uno negativo y uno positivo, mientras que el segundo tiene un sólo lazo de retroalimentación negativo [9]. En éste trabajo se analiza el oscilador simple debido a su propia simplicidad debido a que la frecuencia de resonancia y los periodos de oscilación son sensibles a las variaciones de parámetros, como en el caso del repressilator. Elowitz y Leibler implementaron *in vivo* una versión del repressilator (un oscilador genético sintético con un lazo de retroalimentación negativo), y observaron que una pequeña variación en los parámetros produce efectos considerables en la dinámica del oscilador [7].

Stricker et al. implementaron *in vivo* el oscilador simple, introduciendo dos vectores diferentes pZ dentro de una cepa *E.coli*. El primer vector pZ contiene una copia del gen estructural controlado por el promotor $P_{Llac0-1}$. Éste promotor es reprimido por un represor $LacI$, cuando este último se encuentra unido a la región operadora, pero también este promotor inicia la transcripción cuando la polimerasa está unida a él y por lo tanto no está reprimido por el represor $LacI$. Los genes $lacI$ son transcritos y traducidos para formar el represor $LacI$, por lo que la síntesis de $LacI$ forma un lazo de retroalimentación negativa simple que reprime su propia expresión. El segundo vector pZ contiene un gen $yemGFP$ controlado por el mismo promotor P_{Llac0_1} , este gen es un reportero de la actividad de $LacI$.

El modelo completo del oscilador se construye a partir de las reacciones de represión y transcripción, que describen la dinámica del promotor P_{Llac0_1} , y que fueron introducidas en (6).



Los procesos de transcripción y traducción han sido estudiados ampliamente, se sabe que

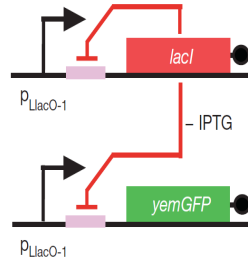


Figure 3: Esquema del lazo de retroalimentación negativa del gen *lacI*. Figura obtenida de Stricker et al. (2008). [9]

el proceso de *transcripción* es una secuencia de reacciones químicas que tiene como finalidad la síntesis del ARNm (deHaseth et al.) [17]; y de igual forma la *traducción* es un conjunto de reacciones que tiene como finalidad la síntesis de proteínas. Sin embargo, estos procesos se pueden representar en un sólo paso; siempre y cuando se haga la suposición de que no hay interrupciones y se introduzca un retardo temporal. Ésto es, la transcripción se representa como una única reacción química con retardo; y de igual forma otra reacción química se representa a la traducción; cada reacción tiene sus propias tasas de síntesis del ARNm y de proteínas respectivamente. Naturalmente el ARNm y las proteínas tienen su propia tasa de degradación. Para completar el modelo se añaden las reacciones de transcripción de ARNm (M) y traducción de proteínas (E), reacciones de dimerización para la formación de dímeros(D) y tetramerización para la formación de tetrámeros (R) del represor *LacI*, reacciones de interacción entre el inductor *IPTG* con el represor *lacI* (R) y degradación tanto del ARNm como de proteínas E , D y R . Ver figura 4. El represor *LacI* es una molécula tetrámerica, cuya función represora es inhibida por los inductores allolactosa y el 1-isopropil- β -D-tiogalactosida (*IPTG*). El represor *LacI* está formado por dos subunidades diméricas D , las cuales son capaces de unirse independientemente al operador *lac*. Una molécula de *IPTG* puede unirse a un dímero D disminuyendo su afinidad por el operador *lac* en un orden de 10^3 [32] [33].

Analicemos la reacción de traducción. Ésta tiene un retardo τ_E ; el cuál es debido al tiempo que el ribosoma tarda en traducir el ARNm a polipéptidos. El valor de éste retardo depende del número de codones que conforman al represor *LacI* (360 codones). La reacción

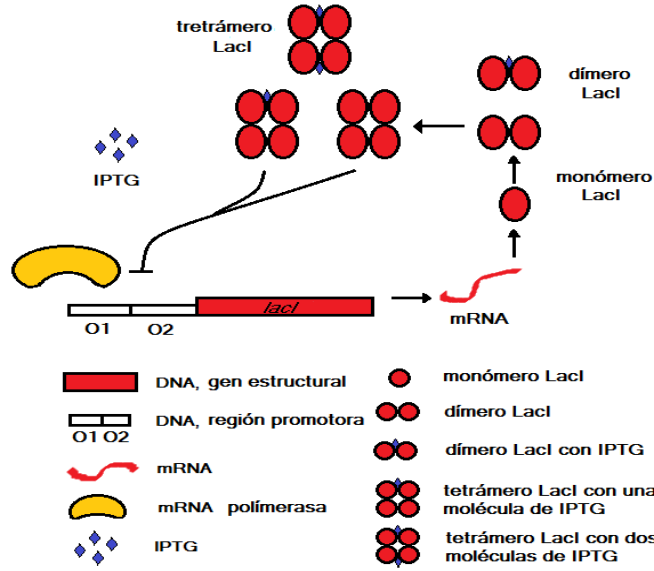


Figure 4: Representación esquemática del lazo de retroalimentación negativa del oscilador genético sintético

de traducción es:



donde E representa los monómeros de represor $LacI$ traducidos a partir del ARN mensajero M con una tasa de traducción k_E .

A continuación incluimos las reacciones de dimerización y tetramerización.



De las reacciones anteriores, a partir de la síntesis del monómero E de represor $LacI$, dos monómeros se unen formando un dímero D , el cual puede unirse al operador lac . Más aún, los dímeros pueden asociarse a tetrámeros, conformando al represor $LacI$ R . De las reacciones anteriores k_D y k_R son las constantes de dimerización y tetramerización respectivamente, así como k_d y k_r son las constantes de disociación de dímeros en monómeros y de tetrámeros en dímeros.

El represor $LacI$ es una molécula tetrámerica, cuya función represora es inhibida por los inductores allolactosa y el 1-isopropil- β -D-tiogalactosida ($IPTG$). El represor $LacI$ está

formado por dos subunidades diméricas D , las cuales son capaces de unirse independientemente al operador lac . Una molécula de $IPTG$ puede unirse a un dímero D disminuyendo su afinidad por el operador lac en un orden de 10^3 [32] [33]. Las siguientes reacciones representan las diferentes interacciones de $IPTG$ con dímeros D y tetrámeros R .

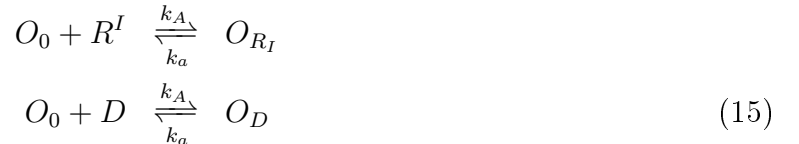


donde k_I y k_i son las constantes de asociación y disociación del IPTG con un dímero o una subunidad dimérica del tetrámero. El $2k_I$ aparecerá ya que el represor tetrámero tiene dos sitios a los cuales se pueden unir el IPTG. D^I es un dímero contaminado con $IPTG$, R^I y R^{2I} es un tetrámero contaminado con una y dos moléculas de IPTG respectivamente.

Los dímeros contaminados también pueden participar en el proceso de tetramerización, lo cual podemos ver en las siguientes reacciones:

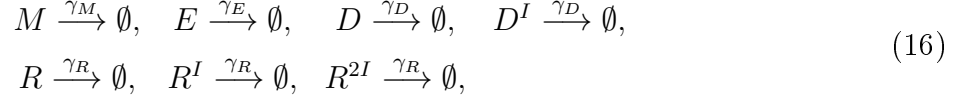


El represor tetrámero R^I contaminado con una molécula de $IPTG$ puede unirse al operador lac por medio de la subunidad dimérica libre de $IPTG$, si ésto ocurre la polimerasa no tiene disponibilidad de unirse al promotor y se impide el inicio de la transcripción. De igual forma, si un represor dimérico D se une al operador, frena la transcripción. Todo lo anterior lo podemos representar en las siguientes reacciones:



Es importante mencionar que la propensión k_A para la formación del complejo O_{R^I} y O_D es la mitad de la propensión del valor de la propensión para formar el complejo O_R presentado en las reacciones (6) de transición de estados del operador, ésto es debido a que sólo tenemos la posibilidad de ocupar un operador con la subunidad dimérica libre de IPTG, pues tenemos la otra subunidad dimérica contaminada por $IPTG$, en el caso R^I .

Finalmente se completa el modelo del oscilador con un lazo de retroalimentación negativa construido por Stricker et al. incluyendo las degradaciones de ARNm y de las diferentes proteínas (monómeros E , dímeros D , D^I y tetrámeros R , R^I y R^{2I}).



5.1 Parámetros del modelo

Los parámetros para las reacciones (6), (7), (11), (12), (13), (14), (15) y (16) se muestran en la siguiente tabla:

$$\begin{aligned}
 k_A &= 233.3/mpb/min, & IPTG &= [0, 1000] mpb \\
 k_a &= 0.0372/min, & k_r &= 0/min \\
 k_B &= 0.4667/mpb/min, & k_D &= 0.02/min \\
 k_b &= 0.0408/min, & k_R &= 0.02/min \\
 k_I &= 0.0015/mpb/min, & k_d &= 0/min \\
 k_M &= 2/min, & k_i &= 0/min \\
 k_E &= 10/min & \gamma_E &= 0.1/min \\
 Pl &= 2 mpb & \gamma_R &= 0.1/min \\
 P &= 8000 mpb & \gamma_D &= 0.1/min \\
 \gamma_M &= 0.639/min & \tau_E &= 0.375 min \\
 n &= 1, 2
 \end{aligned}$$

donde Pl es el número de vectores que contienen el gen estructural $lacI$; Stricker et al. suponen que hay dos vectores [36]. Ver referencias de parámetros en el anexo 2.

Posteriormente se analiza éste conjunto de reacciones con dos procedimientos de la modulación matemática: determinista y estocástica. Ésto es, se construyó un modelo representado por medio de ecuaciones diferenciales ordinarias con retardo (en la traducción) y también se llevaron a cabo simulaciones estocásticas a través del algoritmo de Gillespie con el objetivo de considerar al ruido bioquímico. Ambos modelos se desarrollan con el mismo conjunto de reacciones químicas y se analizan sus resultados. A continuación presentamos nuevamente las reacciones (6), (7), (11), (12), (13), (14), (15) y (16) para facilitar su análisis:

Tabla 1. Reacciones químicas que modelan el oscilador sintético con un lazo de retroalimentación negativa construido por Stricker et al.

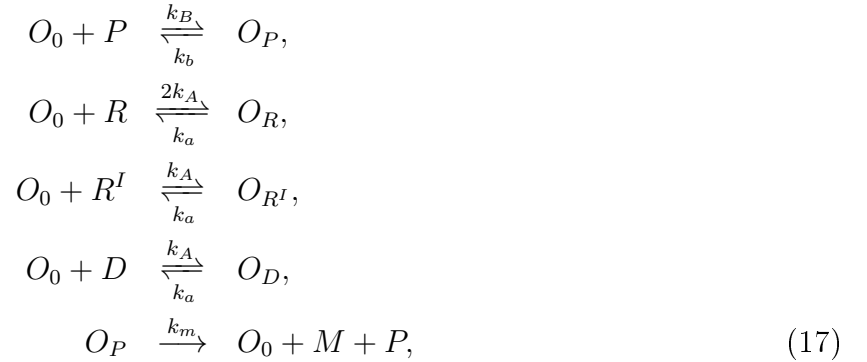
$O_0 + P \xrightleftharpoons[k_b]{k_E} O_P$	Unión de polimerasa al promotor.
$O_0 + R \xrightleftharpoons[k_a]{2k_A} O_R$	Unión de represor LacI R al sitio operador lac .
$O_0 + R^I \xrightleftharpoons[k_a]{k_A} O^{R^I}$	Unión de represor LacI R con una molécula de $IPTG$
$O_0 + D \xrightleftharpoons[k_a]{2k_A} O_D$	Unión de dímero LacI D al sitio operador lac
$O_P \xrightarrow{k_m} O_0 + M + P$	Transcripción del gen $lacI$,
$M \xrightarrow{\gamma_M} \emptyset$	Degradación de ARNm $lacI$,
$M \xrightarrow{k_E} E + M$	Traducción de ARNm a monómero E
$E \xrightarrow{\gamma_E} \emptyset$	Degradación de monómero E ,
$2E \xrightleftharpoons[k_d]{k_D} D$	Dimerización,
$D \xrightarrow{\gamma_d} \emptyset$	Degradación de dímero D ,
$2D \xrightleftharpoons[k_r]{k_R} R$	Tetramerización,
$R \xrightarrow{\gamma_r} \emptyset$	Degradación de tetrámero R ,
$D + I \xrightleftharpoons[k_i]{k_I} D^I$	Asociación de una molécula de $IPTG$ al dímero,
$D + D^I \xrightleftharpoons[k_r]{k_R} R^I$	Tetramerización,
$R + I \xrightleftharpoons[k_i]{2k_I} R^I$	Asociación de una molécula de $IPTG$ al tetrámero,
$2D^I \xrightleftharpoons[k_r]{k_R} R^{2I}$	Tetramerización,
$R^I + I \xrightleftharpoons[k_i]{k_I} R^{2I}$	Asociación de una molécula de $IPTG$ a R^I ,
$D^I \xrightarrow{\gamma_d} \emptyset$	Degradación de D^I ,
$R^I \xrightarrow{\gamma_r} \emptyset$	Degradación de R^I ,
$R^{2I} \xrightarrow{\gamma_r} \emptyset$	Degradación de R^{2I} .

5.2 Modelo determinista

Para el modelo completo basado en el oscilador simple con retroalimentación negativa construido experimentalmente por Stricker et al. se usa el promotor P_{Llac0_1} . En las siguientes subsecciones se utiliza éste promotor para desarrollar sus modelos matemáticos. El promotor P_{Llac0_1} controla al gen $lacI$, el cual codifica para el represor $LacI$ R que reprime al promotor P_{Llac0_1} . El modelo determinista se construyó a partir de las reacciones químicas propuestas anteriormente, descritas en (6)-(8) y (11)-(16).

Para construir el sistema de ecuaciones diferenciales ordinarias que describan la dinámica temporal del sistema de reacciones químicas arriba descrito, debemos empezar por calcular la dinámica temporal de probabilidades $P(O_0)$, $P(O_P)$, $P(O_D)$, $P(O_R)$ y $P(O_{R^I})$ de que el promotor se encuentre en alguno de los cinco estados O_0 , O_P , O_D , O_R y O_{R^I} . Éstas son las ecuaciones de la ecuación química maestra asociadas a las reacciones (6) y (15); a continuación se muestran las reacciones completas para el promotor P_{Llac0_1} :

Las transiciones entre los estados del operador O_0 , O_P , O_R , O_{R^I} y O_D se enuncian nuevamente en las siguientes reacciones químicas para facilitar la presentación de los resultados:



La correspondiente ecuación química maestra es la siguiente:

$$\begin{aligned}
 \frac{dP(O_P)}{dt} &= k_B P(O_0)[P] - k_b P(O_P) - k_m P(O_P), \\
 \frac{dP(O_D)}{dt} &= k_A P(O_0)[D] - k_a P(O_D), \\
 \frac{dP(O_R)}{dt} &= 2k_A P(O_0)[R] - k_a P(O_R), \\
 \frac{dP(O_{R^I})}{dt} &= k_A P(O_0)[R^I] - k_a P(O_{R^I}).
 \end{aligned} \tag{18}$$

El valor exacto de $P(O_0)$ se deduce fácilmente del hecho que la suma de todas las probabil-

idades es igual a uno. Así que:

$$P(O_0) = 1 - P(O_P) - P(O_D) - P(O_R) - P(O_{R^I}). \quad (19)$$

Seguido de lo anterior, ahora se toman en cuenta las reacciones de transcripción y traducción descritas en (7) y (11), las cuales representamos con ecuaciones diferenciales ordinarias, en dónde se encuentran términos que representan a la síntesis y degradación del ARN mensajero y del represor *LacI* en su forma monomérica. Con el objetivo de hacer más realista nuestro modelo, consideramos un retardo τ_E en la traducción de E , ya que aunque la transcripción y traducción en procariotas se da casi en forma simultánea, si existe un retardo en la síntesis de proteínas debido al número de codones que traduce el ribosoma.

Diferentes trabajos como Cao et al. (2012), Elowitz et al. (2000), etc. han utilizado sistemas de ecuaciones diferenciales ordinarias para representar la dinámica de síntesis y degradación de ARNm y proteínas, así como Hidde de Jong (2002), Hernández-Valdez et al. (2010), entre otros han usado ecuaciones diferenciales con retardo para el mismo fin [34] [7] [35] [57]. Las siguientes ecuaciones representan la dinámica temporal de ARNm (M) y de monómeros *LacI* (E) de nuestro modelo:

$$\begin{aligned} \frac{dM}{dt} &= k_m Pl - \gamma_m M, \\ \frac{dE}{dt} &= k_E M(t - \tau) - 2k_D E^2 + 2k_d D - \gamma_E E, \end{aligned} \quad (20)$$

dónde los términos $-2k_D E^2 + 2k_d D$ que aparecen en la ecuación inmediata anterior representan la relación entre las dinámicas de los monómeros (E) y dímeros D . Ésto es, el término $-2k_D E^2$ representa cuando un dímero D se forma y dos monómeros desaparecen. De igual forma, $2k_d D$ representa cuando un dímero desaparece tomando dos monómeros.

Sean D dímeros de *LacI*, D^I dímero con una molécula de *IPTG*, R represor *LacI*, R^I el represor *LacI* con una molécula de *IPTG* y R^{2I} el represor *LacI* con dos moléculas de *IPTG*. Para completar nuestro sistema de ecuaciones diferenciales consideramos la dinámica de los dímeros D y tetrameros R , así como los dímeros y tetrameros contaminados con *IPTG* D^I , R^I , R^{2I} . Éstas ecuaciones diferenciales se construyen a partir del principio de la Ley de acción de masas y se muestran a continuación:

$$\begin{aligned}
\frac{dD}{dt} &= k_D E^2 - k_d D - 2k_R D^2 + 2k_r R - k_I D I + k_i D^I + k_r R^I - k_R D D^I - \gamma_D D, \\
\frac{dD^I}{dt} &= k_I D I - k_R D D^I - 2k_R (D^I)^2 + 2k_r R^{2I} - k_i D^I + k_r R^I - \gamma_D D^I, \\
\frac{dR}{dt} &= k_R D^2 - 2k_I R I + k_i R^I - k_r R - \gamma_R R, \\
\frac{dR^I}{dt} &= k_R D D^I - k_r R^I + 2k_I R I - k_i R^I - k_I R^I I + k_i R^{2I} - \gamma_R R^I, \\
\frac{dR^{2I}}{dt} &= k_R (D^I)^2 - k_r R^{2I} + k_I R^I I - k_i R^{2I} - \gamma_R R^{2I}.
\end{aligned} \tag{21}$$

La solución de éste sistema de ecuaciones diferenciales con retardo, nos dará las concentraciones de las especies químicas a través del tiempo.

El modelo completo al cual llamaremos de transiciones lentas está conformado por las ecuaciones diferenciales descritas en (18), (20), (21) y (19). Otra manera de representar el modelo matemático es un modelo simplificado al cual llamaremos de transiciones rápidas. Para ello, se calcula la probabilidad $P(O_P)$ suponiendo que el sistema de reacciones descrito en (17) está en estado estacionario y luego se sustituye en el siguiente sistema.

$$\begin{aligned}
\frac{dM}{dt} &= k_m P I P(O_P) - \gamma_M M, \\
\frac{dE}{dt} &= k_E M(t - \tau) - 2k_D E^2 + 2k_d D - \gamma_E E, \\
\frac{dD}{dt} &= k_D E^2 - k_d D - 2k_R D^2 + 2k_r R - k_I D I + k_i D^I + k_r R^I - k_R D D^I - \gamma_D D, \\
\frac{dD^I}{dt} &= k_I D I - k_R D D^I - 2k_R (D^I)^2 + 2k_r R^{2I} - k_i D^I + k_r R^I - \gamma_D D^I, \\
\frac{dR}{dt} &= k_R D^2 - 2k_I R I + k_i R^I - k_r R - \gamma_R R, \\
\frac{dR^I}{dt} &= k_R D D^I - k_r R^I + 2k_I R I - k_i R^I - k_I R^I I + k_i R^{2I} - \gamma_R R^I, \\
\frac{dR^{2I}}{dt} &= k_R (D^I)^2 - k_r R^{2I} + k_I R^I I - k_i R^{2I} - \gamma_R R^{2I}.
\end{aligned} \tag{22}$$

donde $P(O_P)$ es la probabilidad de que una polimerasa esté unida al promotor y se pueda iniciar el proceso de transcripción.

Considerando que la suma de las probabilidades $P(O_P) + P(O_0) + P(O_D) + P(O_R) + P(O_{R^I})$ es igual a 1 y que las probabilidades se encuentran en estado estacionario, es fácil calcular $P(O_P)$ como sigue:

$$P(O_P) = \frac{\frac{k_B}{k_b + k_m} [P]}{1 + \frac{k_B}{k_b + k_m} [P] + 2 \frac{k_A}{k_a} [R] + \frac{k_A}{k_a} ([R^I] + [D])}. \tag{23}$$

Esta función se define para describir la regulación del inicio de la transcripción por medio de las interacciones del represor *LacI* y los operadores O_R . $P(O_P)$ es la probabilidad que una polimerasa esté unida al promotor $P_{Llac0-1}$ y por tanto éste promotor no esté reprimido por un represor *LacI*. Esta función toma en cuenta las interacciones del represor R , el inductor *IPTG* con el represor, la polimerasa P y los sitios de operación O_R .

Hasta ahora se han descrito dos modelos de ecuaciones diferenciales. Uno descrito en las ecuaciones (18), (20) y (21), el cual es el modelo completo y al que nosotros llamamos modelo de transiciones lentas y el otro es el modelo simplificado descrito en las ecuaciones (22), al que llamamos modelo de transiciones rápidas en el que calculamos a $P(O_P)$ en estado estacionario (ver ecuación 23).

5.3 Modelo estocástico

En el modelo estocástico se toman en cuenta el mismo conjunto de reacciones químicas usadas para la construcción de los modelos deterministas. El objetivo de este modelado estocástico es tomar en cuenta la aleatoriedad inherente de las reacciones químicas, lo cual no se logra al desarrollar el modelo con ecuaciones diferenciales. Así, con éste modelo se analiza el comportamiento oscilatorio del sistema de reacciones químicas incluyendo al ruido bioquímico. Aunado a lo anterior, al igual que en el modelado determinista también se realizan dos modelos estocásticos. Uno es el modelo completo al cual llamamos de transiciones lentas, en el cual se implementa un algoritmo de simulación estocástica con retardo que incluye a todas las reacciones descritas en el apartado anterior (ver tabla 1, pag. 33) y el otro modelo es el modelo simplificado o llamado modelo de transiciones rápidas en el cual se calcula a $P(O_P)$ en estado estacionario (ver ecuación 23) y se sustituye en la reacción de transcripción.

6 Métodos

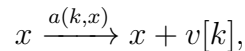
Se realizaron simulaciones deterministas y estocásticas para analizar el comportamiento oscilatorio del sistema. Para el sistema determinista se utilizó Mathematica para resolver el sistema de ecuaciones diferenciales con retardo, graficar las soluciones y así medir los periodos y frecuencias de oscilación de las concentraciones del represor *LacI*. Para el modelo

estocástico se implementó el algoritmo de Gillespie con el objetivo de tomar en cuenta la naturaleza aleatoria de las reacciones químicas participantes.

Las simulaciones con el algoritmo de Gillespie considera al sistema de reacciones (descritas en la tabla 1) como un proceso estocástico, por lo que cada simulación (trayectoria) es diferente a la anterior, lo que dificulta determinar cual es el periodo de oscilación de una trayectoria determinada. El periodo es una variable aleatoria.

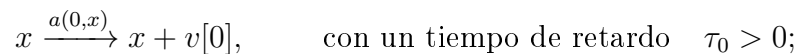
El algoritmo de Gillespie se usa para simular reacciones químicas instantáneas, sin embargo, en la reacción de traducción de nuestro sistema tenemos incluido el retardo $\tau = 0.375$. Por lo anterior implementamos un algoritmo de simulación estocástica (SSA) que incluya un retardo en la reacción de traducción, basado en el algoritmo de Gillespie y modificado para incluir una reacción con retardo en la traducción. Este algoritmo está referenciado en el trabajo de Cai et al. [15], en donde demuestra que el método de rechazo desarrollado por Barrio et al. es adecuado para incluir reacciones con retardo. Por simplicidad nosotros adaptamos el algoritmo para el caso cuando sólo hay una reacción no consumista con retardo (a la cual llamamos reacción 0).

Recordando, una reacción que se lleva a cabo en un reactor químico está indexada por un entero $k = 1, 2, \dots, m$; y entonces la k ésima reacción puede ser presentada como:



donde $a(k, x) \geq 0$ es la propensión de la reacción química y $v[k] \in \mathbb{Z}^n$ es el vector estequiométrico, el cuál indica cuantas moléculas de cada especie química se crearon o destruyeron en la reacción.

Por lo que, la reacción $k = 0$ está representada por:



En una reacción no consumista los reactantes de una reacción que aún no finaliza pueden participar en una nueva reacción y los conteos moleculares sólo cambian cuando la reacción ha finalizado. El pseudocódigo para el método de rechazo de Barrio et al. se describe en los siguientes pasos:

- 1 Inicializar con el tiempo $t = t_0$ y el estado del sistema $X(t) = x_0(t)$, siendo $X(t)$ un vector que contiene número de moléculas inicial de cada especie química participante en el tiempo t .
- 2 Dado el estado $X(t)$ y el tiempo t , evaluar las propensiones del sistema $a(k, x(t))$ de acuerdo a (4) y su suma $a_0(x(t)) = \sum_{k=1}^m a(k, x(t))$. Generar el tiempo de espera $\tau = \frac{1}{a_0(X)} \ln \frac{1}{r_1}$ donde r_1 es un número aleatorio con distribución uniforme en el intervalo $[0, 1]$. Si existe una reacción con retardo que finalice en el intervalo $(t, t + \tau)$, encontrar el tiempo $t_d \in [t, t + \tau]$ en el cual finaliza la primera reacción con retardo y descartar τ . Actualizar el vector $X(t_d) = X(t) + v[0]$, actualizar el tiempo $t = t_d$ y repetir el paso 2. Si no existe una reacción con retardo que finalice en el intervalo $(t, t + \tau)$, continuar con el paso 3.
- 3 Generar un número aleatorio r_2 en el intervalo $[0, 1]$ y elegir el índice j , para conocer la reacción asociada que se producirá en el siguiente paso temporal. Para ello, el índice debe cumplir la siguiente condición: $\sum_{k=0}^{m-1} a_k < r_2 a_0 \leq \sum_{k=0}^m a_k$. Si la reacción elegida es la reacción con retardo, guardar el tiempo de finalización $t + \tau_0$ a la lista de reacciones que están en proceso de finalizar. Si la reacción elegida es una reacción instantánea entonces seguir al paso 4.
- 4 Actualizar el estado $X(t + \tau) = x(t) + v[k]$, donde $v[k]$ es el cambio en el número de moléculas causada por la k ésima reacción. Además actualizar el tiempo $t + \tau \mapsto t$.
- 5 Regresar al paso 2 o terminar la simulación si el tiempo actual t es mayor que el tiempo de simulación.

Por lo anterior nosotros calculamos un millón de trayectorias $X_{k(t)}$ con el algoritmo de Gillespie. Cada trayectoria $x_j(k)$ es simulada de 0 a 500 minutos y se mide el conteo molecular de un represor tetramérico R a través del tiempo t . Se borraron los primeros 200 minutos de todas las trayectorias con el objetivo de eliminar el estado estacionario; después muestreamos el resto de la trayectoria (300 minutos) guardando el número de moléculas cada minuto (teniendo 300 muestras por cada trayectoria). Finalmente calculamos la transformada rápida de Fourier de cada trayectoria muestreada y trabajamos con el espectro de potencias promedio.

Así, si $C_k(n)$ son los coeficientes de la transformada rápida de Fourier de una trayectoria muestreada $x_k(t)$, entonces el espectro de potencias promedio está dado por:

$$C(n) = \frac{1}{N} \sum_{k=1}^N |C_k(n)|^2 \quad \text{para } n = 0, 1, 2, \dots, 299$$

donde $N = 10^6$ es el número de trayectorias calculadas (número de simulaciones).

7 Resultados

En este capítulo se presentan los resultados obtenidos al simular tanto el modelo determinista como el modelo estocástico. Las variables que mediremos en todas las simulaciones son la frecuencia de resonancia y el periodo de oscilación, para comparar nuestros resultados contra las mediciones experimentales realizadas por Stricker et al [9].

Primeramente se presenta el modelo para los promotores de Lutz y Bujard que incluye sólo las reacciones de los estados del operador (ecuaciones (10) y (17)), para luego compararlos contra las mediciones experimentales del rango de regulación. Posteriormente se presentan los resultados del modelo determinista en dos casos: modelo completo (llamado de transiciones lentas) de acuerdo a las ecuaciones (18), (20) y (21); y modelo simplificado (llamado de transiciones rápidas) de acuerdo a (22), enfatizando en el efecto que tienen las transiciones de los estados del operador en el periodo de oscilación. Finalmente se hace el mismo análisis en el modelo estocástico: transiciones lentas y transiciones rápidas, sin embargo, en el modelo estocástico es necesario mencionar que el periodo no se mide en una forma trivial como en el modelo determinista. En los resultados de los modelos estocásticos se calcularon los espectros de potencia de los conteos moleculares del represor R en su forma tetramérica y monomérica, los espectros de potencia nos permiten encontrar las periodicidades que no se pueden identificar a simple vista, además de que el periodo es una variable aleatoria al considerar como un proceso estocástico al sistema de reacciones químicas debido a la influencia del ruido bioquímico.

7.1 Resultados del modelo matemático para los promotores de Lutz y Bujard

Como se mencionó anteriormente Lutz y Bujard diseñaron diferentes promotores sintéticos, tales como P_{Llac0_1} , P_{A1lac0_1} y P_{lac} . De los cuales P_{lac} es el promotor obtenido directamente de la cepa *E. coli* silvestre y P_{Llac0_1} fue usado por Stricker et al. en sus experimentos.

7.2 Rangos de regulación medidos *in silico*

Lutz y Bujard realizaron experimentos en los cuales midieron la fuerza del promotor y con ello calcularon el rango de regulación de los promotores P_{Llac0_1} , P_{A1lac0_1} , que hemos considerado hasta ahora. El rango de regulación está definido como el cociente de la expresión del reportero (o fuerza del promotor) medida en condiciones de inducción dividido entre la expresión medida en condiciones de represión máxima (ver Tabla 1 en [30]).

La fuerza del promotor se mide experimentalmente en función de la actividad de la luciferasa (ver figura 4 en [30]). A continuación se describe una sencilla expresión para el cálculo de la fuerza del promotor la cual se usa para calcular el rango de regulación de los promotores P_{Llac0_1} , P_{A1lac0_1} y P_{lac} .

$$\text{Fuerza del promotor} = \alpha[\text{Luciferasa}] = \alpha\beta[\text{ARNm Luciferasa}] = \alpha\beta\gamma P(O_P)$$

donde: $P(O_P)$ es la probabilidad de que la polimerasa se una al promotor y se inicie el proceso de transcripción, γ es la constante de transcripción, β es la constante de traducción y α es una constante de proporcionalidad de la concentración de la enzima luciferasa.

Lutz y Bujard reportan experimentalmente una fuerza de promotor para el promotor P_{Llac0_1} , en dos casos distintos, uno con presencia del inductor *IPTG* (igual a 21630) y el otro en ausencia del *IPTG* (igual a 35). Lo anterior podemos expresarlo como sigue:

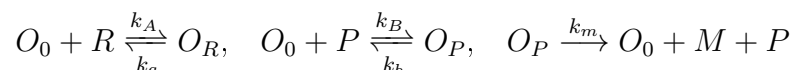
$$\text{Fuerza del promotor} = \alpha\beta\gamma P(O_P \text{ con } IPTG) = 21630$$

$$\text{Fuerza del promotor} = \alpha\beta\gamma P(O_P \text{ sin } IPTG) = 35$$

Por lo anterior el rango de regulación, al cual abreviaremos como RR , se calcula de la siguiente forma:

$$RR = \frac{P(O_P \text{ con } IPTG)}{P(O_P \text{ sin } IPTG)} = 620$$

La $P(O_P \text{ con } IPTG)$ es aproximadamente 1. El conteo molecular de R está alrededor de 10 – 20 represores por célula [43], si existe *IPTG* induciendo a la célula, los represores se contaminan y podemos suponer que el número de represores libres R es cero. Con base a las ideas anteriores, calculamos el rango de regulación $RR = \frac{1}{P(O_P)}$, donde $P(O_P)$ está calculada en 24 y es la probabilidad de que una polimerasa esté unida al promotor y se inicie la transcripción. En esta sección se valida el modelo para los promotores de Lutz y Bujard, por lo que se valida el sistema de reacciones (17) que describen los estados del operador.



Lo anterior se lleva a cabo substituyendo los valores correspondientes a las constantes cinéticas químicas medidas *in vitro* por diferentes autores y se comparan los cálculos contra rangos de regulación medidos *in vivo*. Comparar por ejemplo la ecuación (10) contra la siguiente:

$$\frac{1}{P(O_P)} = 1 + \frac{nk_A}{k_B} \left(\frac{k_b + k_m}{k_a} \right) \frac{[R]}{[P]}. \quad (24)$$

La ecuación anterior se usa para calcular los rangos de regulación para cada uno de los promotores, como se explica posteriormente.

7.2.1 Promotor P_{lac}

Oehler, Amoyual, et al. (1994) midieron *in vivo* un rango de regulación de 200 para el promotor P_{lac} (ver [48], p. 3350), el cual tiene un sólo operador ($n = 1$) ubicado después del sitio de transcripción, este experimento se realiza en una cepa que contiene 50 tetrámeros por *E.coli*. Con el modelo para los promotores de Lutz y Bujard nosotros calculamos un rango de regulación de 215.2 substituyendo los siguientes parámetros en (24):

$$\begin{aligned} k_A &= 7 \times 10^9 / M/s & [R] &= 50 \text{ mpb}, \\ k_a &= 6.2 \times 10^{-4} / s, & k_m &\approx 0.0033 / s, \\ k_B &= 2 \times 10^6 / M/s, & [P] &= 8300 \text{ mpb}, \\ k_b &= 3.3 \times 10^{-3} / s, & n &= 1. \end{aligned}$$

7.2.2 Promotor P_{A1lac0_1}

Lutz y Bujard (1997) midieron *in vivo* un rango de regulación de 350 para el promotor P_{A1lac0_1} [30], cuando éste se encuentra controlando a un gen reportero luciferasa en un vector pZ . Este promotor está construido con el promotor bacteriófago P_{A1} , al cual se le insertan dos secuencias operadoras ($n = 2$), una de ellas se encuentra entre los hexameros -10 y -33 y la otra secuencia operadora se encuentra río abajo del sitio de inicio de transcripción (ver figura 1). Este experimento fué medido en una cepa creada por los mismo autores, en la cual se induce una sobreproducción constitutiva del represor $LacI$, ésto se logra poniendo una copia del gen $laci$ bajo el control de un promotor constitutivo mutante $laci^q$, con el cual se producen 10 veces la cantidad de represores $LacI$ (ver Tabla 1 en [44]). Al substituir los valores de los parámetros cinéticos medidos *in vitro* mostrados en la siguiente lista obtenemos un valor de 372 de la ecuación (24).

$$\begin{aligned}
k_A &= 7 \times 10^9 / M/s & [R] &= 90 \text{ mpb}, \\
k_a &= 6.2 \times 10^{-4} / s, & k_m &\approx 0.09 / s, \\
k_B &= 6 \times 10^7 / M/s, & [P] &= 8300 \text{ mpb}, \\
k_b &= 9.2 \times 10^{-4} / s, & n &= 2.
\end{aligned}$$

7.2.3 Promotor P_{Llac0_1}

Lutz y Bujard midieron *in vivo* un rango de regulación de 620 para el promotor P_{Llac0_1} cuando éste controla al gen reportero luciferasa en un vector pZ (ver Tabla 1 en [30], p.1205). Este promotor está construido a partir de un promotor colifago lambda P_L al cual se le añaden dos operadores lac ($n = 2$), uno de ellos se encuentra entre los hexámeros -10 y -33 (ver figura 1). En este experimento también se induce una sobreproducción constitutiva del represor $LacI$ poniendo una copia del gen $laci$ bajo el control del promotor $laci^q$, el cual produce diez veces más la cantidad de represores que en una bacteria *E.coli* silvestre. Nosotros obtenemos el valor de 595.9 de la ecuación (24) sustituyendo los parámetros medidos *in vitro* enlistados a continuación:

$$\begin{aligned}
k_A &= 7 \times 10^9 / M/s & [R] &= 90 \text{ mpb}, \\
k_a &= 6.2 \times 10^{-4} / s, & k_m &\approx 0.0333 / s, \\
k_B &= 1.4 \times 10^7 / M/s, & [P] &= 8300 \text{ mpb}, \\
k_b &= 6.8 \times 10^{-4} / s, & n &= 2.
\end{aligned}$$

7.2.4 Resultados Generales

En la siguiente tabla se resumen los cálculos para cada promotor con sus parámetros correspondientes y los rangos de regulación obtenidos a partir de la ecuación (10).

Tabla 2. Rangos de regulación medidos *in silico* al calcular $\frac{1}{P(O_P)}$.

Experimento	Promotor y RR	Parámetros exp.	RR calculado
Lutz y Bujard	$P_{LlacO-1}$	$\frac{k_A}{k_B} = \frac{7 \times 10^9 / s / M}{1.4 \times 10^7 / s / M}$	$\frac{1}{P(O_P)} = 595.9$
$DH5\alpha Z1$	dos O_1 a 22pb RR=620	$\frac{k_b+k_m}{k_a} = \frac{3.398 \times 10^{-2}}{6.2 \times 10^{-4}}$ $\frac{n[R]}{[P]} = \frac{2 \times 90}{8300}$	
Lutz y Bujard	$P_{AlacO-1}$	$\frac{k_A}{k_B} = \frac{7 \times 10^9}{6 \times 10^7}$	$\frac{1}{P(O_P)} = 372$
$DH5\alpha Z1$	dos O_1 a 33pb RR=350	$\frac{k_b+k_m}{k_a} = \frac{9.092 \times 10^{-2}}{6.2 \times 10^{-4}}$ $\frac{n[R]}{[P]} = \frac{2 \times 90}{8300}$	
Oehler y Amouyal	P_{lac} , un O_1	$\frac{k_A}{k_B} = \frac{7 \times 10^9}{2 \times 10^6}$	$\frac{1}{P(O_P)} = 215.2$
$BMH8117$	RR=200	$\frac{k_b+k_m}{k_a} = \frac{6.3 \times 10^{-3}}{6.2 \times 10^{-4}}$ $\frac{n[R]}{[P]} = \frac{50}{8300}$	

7.3 Resultados de las simulaciones del oscilador sintético más simple construido por Stricker et al.

A continuación presentamos los resultados de las simulaciones, tanto deterministas como estocásticas, que hemos llevado a cabo con los modelos construidos en la sección 5.

7.3.1 Resultados del modelo determinista

Stricker et al. construyeron dos osciladores sintéticos en su trabajo titulado: *A fast, robust and tunable synthetic gene oscillator* publicado en el año 2008. Uno de ellos, el más simple, está construido por medio de dos vectores pZ introducidos en una *E.coli*. Cada vector pZ tiene el promotor P_{LacO_1} (diseñado por Lutz y Bujard) el cual controla respectivamente al gen $lacI$ y al gen reportero $yemGFP$. El gen $lacI$ codifica para el represor $LacI$, el cual reprime la transcripción de su mismo gen, formando así un lazo de retroalimentación negativo. Stricker et al., reporta en sus resultados que éste sistema oscilatorio tiene un periodo de oscilación de aproximadamente 30 min. En ésta sección verificamos si nuestro modelo reproduce los resultados experimentales medidos por Stricker et al.

Modelo de transiciones lentas Se llevaron a cabo simulaciones deterministas y estocásticas a partir del sistema de reacciones químicas descritas en la tabla 1 de la sección 5. Para el modelo determinista se uso un método de integración numérica para resolver a las ecuaciones diferenciales con retardo y se graficó la dinámica temporal del conteo molecular de las especies químicas. Llamamos transiciones lentas cuando usamos el modelo completo de ecuaciones diferenciales, descritas en (18), (20) y (21). Los resultados se muestran en la siguiente tabla:

Tabla 3. Periodos de oscilación en diferentes condiciones de *IPTG*, usando el modelo determinista con transiciones lentas.

Caso	<i>IPTG</i> (mpb)	τ_{crit} (min)	Periodo
1	0	34	105 min
2	50	25	76 min
3	100	21	63 min
4	150	19	56 min
5	200	19	53 min
6	250	19	53 min
7	300	19	53 min
8	400	19	53 min
9	500	21	55 min

Donde τ_{crit} , es el valor mínimo del retardo necesario para observar oscilaciones.

Se observa que el periodo disminuye conforme aumenta el número de moléculas de *IPTG* y se reporta un retardo crítico τ_{crit} para cada concentración de *IPTG*. Aunado a lo anterior, se llevaron a cabo simulaciones fijando las concentraciones de *IPTG* y aumentando el retardo a partir del valor τ_{crit} y se observa que el periodo aumenta si el retardo aumenta para una concentración fija de inductor *IPTG*.

En este modelo se observa que no hay congruencia con el periodo medido experimentalmente por Stricker et al., pues los periodos comienzan a crecer en concentraciones mayores de *IPTG*, alrededor de 500 mpb. El periodo mínimo que se observa es de 53 min.

Modelo de transiciones rápidas Para el modelo de transiciones rápidas (22), se calcula la $P(O_P)$ (probabilidad de que una polimerasa se una al promotor) en estado estacionario y se sustituye en la ecuación diferencial que describe la dinámica de la síntesis del represor *LacI* en su forma monomérica (E). De igual manera que en el caso anterior se resuelve numéricamente el sistema de ecuaciones diferenciales descrito en (22) y se gráfica la dinámica temporal del conteo molecular de las especies químicas. Los resultados se encuentran en la siguiente tabla:

Tabla 4. Periodos de oscilación en diferentes condiciones de *IPTG*, usando el modelo determinista con transiciones rápidas.

Caso	<i>IPTG</i> (mpb)	τ_{crit} (min)	Periodo
1	0	28	88 min
2	50	17	55 min
3	100	14	45 min
4	150	12	38 min
5	200	11	35 min
6	250	10	32 min
7	300	10	30 min
8	400	10	29 min
9	500	9	27 min

Siendo τ_{crit} el valor mínimo del retardo para observar oscilaciones.

Se observa que el periodo disminuye conforme aumenta el número de moléculas de *IPTG* y se reporta un retardo crítico τ_{crit} para cada concentración de *IPTG*. Aunado a lo anterior, se llevaron a cabo simulaciones fijando la concentraciones de *IPTG* y aumentando el retardo a partir del τ_{crit} y se observa que el periodo aumenta si el retardo aumenta para una concentración fija de inductor.

En este modelo se observa que si hay congruencia con el periodo medido experimentalmente por Stricker et al. para concentraciones de *IPTG* mayores a 300 moléculas. Sin embargo, ésto se logra si el retardo en la traducción es de 10 min aproximadamente, lo cual es mucho más grande que el retardo establecido en los parámetros $\tau_E = 0.375 \text{ min}$.

En el modelo determinista el modelo de transiciones lentas incrementa los periodos de oscilación en comparación con las transiciones rápidas, lo cual muestra que al considerar las transiciones lentas el oscilador es más lento que cuando se usan transiciones rápidas para los cambios de estados del operador (ver figura 5).

7.3.2 Resultados del modelo estocástico

El modelo determinista con transiciones rápidas reproduce el comportamiento oscilatorio del oscilador sintético construido experimentalmente, pero hay un problema al ajustar el retardo a más de 25 veces del retardo calculado en los parámetros considerando la velocidad de la

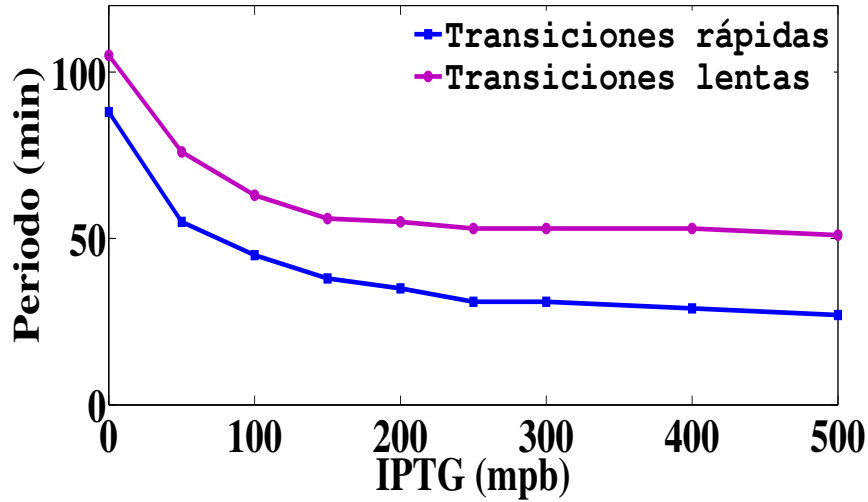


Figure 5: Resultados del modelo determinista. En el eje horizontal se muestra el número de moléculas de *IPTG* y en el eje vertical el periodo de oscilación medido en una concentración de *IPTG* específica.

elongación de la cadena de péptidos que sucede en la traducción de monómeros *E*. Por lo anterior, se implementa un algoritmo de simulación estocástica, basado en el algoritmo de Gillespie (ver métodos), con el objetivo de añadir el ruido intrínseco de las reacciones químicas (ruido bioquímico).

Las simulaciones con el algoritmo de Gillespie considera al sistema de reacciones (descritas en la tabla 1) como un proceso estocástico, por lo que cada simulación (trayectoria) es diferente a la anterior, lo que dificulta determinar cual es el periodo de oscilación de una trayectoria determinada. Por lo anterior nosotros calculamos un millón de trayectorias $X_{k(t)}$ con el algoritmo de Gillespie. Cada trayectoria $x_j(k)$ es simulada de 0 a 500 minutos y se mide el conteo molecular de un represor tetramérico *R* a través del tiempo *t*. Se borraron los primeros 200 minutos de todas las trayectorias con el objetivo de eliminar el estado estacionario; después muestreamos el resto de la trayectoria (300 minutos) guardando el número de moléculas cada minuto (teniendo 300 muestras por cada trayectoria). Finalmente calculamos la transformada rápida de Fourier de cada trayectoria muestreada y trabajamos con el espectro de potencias promedio. Así, si $C_k(n)$ son los coeficientes de la transformada rápida de Fourier de una trayectoria muestreada $x_k(t)$, entonces el espectro de potencias

promedio está dado por:

$$C(n) = \frac{1}{N} \sum_{k=1}^N |C_k(n)|^2 \quad \text{para } n = 0, 1, 2, \dots, 299 \quad (25)$$

donde $N = 10^6$ es el número de trayectorias calculadas (número de simulaciones).

Los resultados son presentados en la figura 6; y puede verse claramente que la frecuencia de resonancia es aproximadamente igual a $\frac{1}{30min}$ cuando la concentración de *IPTG* es $[I] = 100 \text{ mpb}$. Esta frecuencia de resonancia corresponde a un periodo similar medido *in vivo* (de 30 minutos) por Stricker et al. Es importante observar en la figura 6 que la magnitud de las oscilaciones disminuye en tanto que la concentración del *IPTG* se incrementa. Este efecto sucede debido a que el *IPTG* es capturado por los represores tetraméricos *R* inactivándolos. Sin embargo, cuando calculamos el espectro de potencias promedio del conteo molecular del represor monomérico *E*, también observamos una frecuencia de resonancia aproximada de $\frac{1}{30min}$ para una concentración de *IPTG* de $[I] = 100 \text{ mpb}$; y además podemos ver de la figura 7 que la magnitud de las oscilaciones no disminuye.

En la figura 6 y 7 se muestran los espectros de potencias de los conteos moleculares para diferentes concentraciones de *IPTG* usando el promotor P_{LlacO_1} .

De lo observado en los espectros de potencia, en la siguiente tabla se reportan los periodos de oscilación para diferentes concentraciones de *IPTG*, cuando usamos las transiciones lentas.

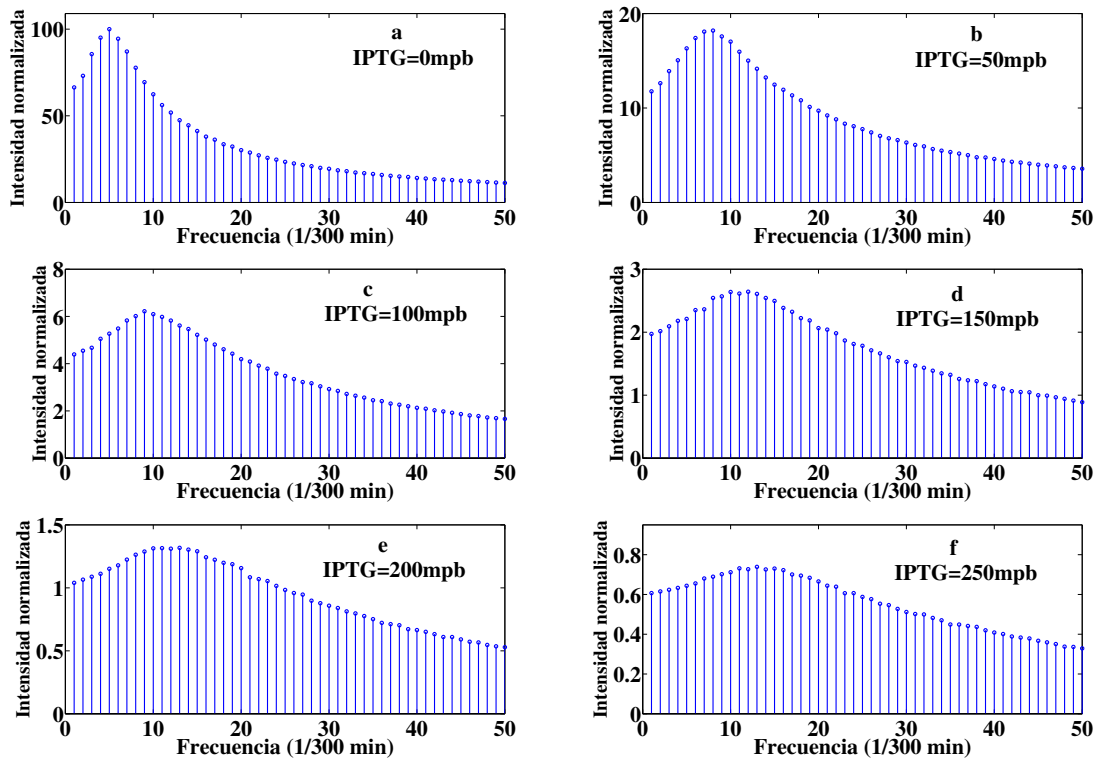


Figure 6: Espectro de potencias del conteo molecular del represor *lac* tetramérico *R* con diferentes concentraciones de IPTG: 0mpb, **a**; 50mpb, **b**; 100mpb, **c**; 150mpb, **d**; 200mpb, **e**; 250mpb, **f**. En el eje horizontal se indica la frecuencia en unidades de $\frac{1}{300 \text{ min}}$, y en el eje vertical se muestra la intensidad normalizada en unidades arbitrarias.

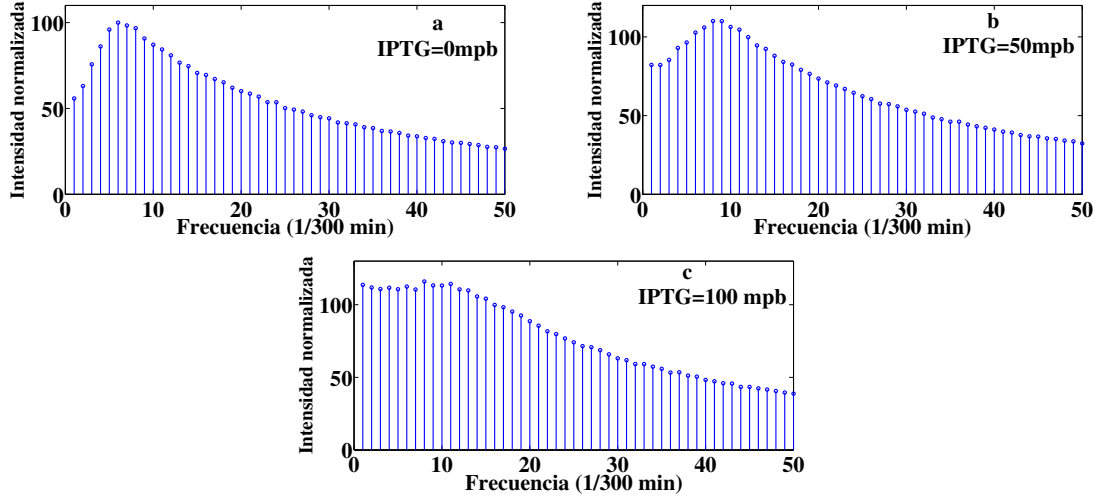


Figure 7: Espectro de potencias del conteo molecular del represor *lac* monomérico E con diferentes concentraciones de IPTG: 0mpb , **a**; 50mpb , **b**; 100mpb , **c**. En el eje horizontal se indica la frecuencia en unidades de $\frac{1}{300\text{min}}$, y en el eje vertical se muestra la intensidad normalizada en unidades arbitrarias.

Modelo de transiciones lentas.

Tabla 5. Periodos de oscilación en diferentes condiciones de *IPTG*, usando el modelo estocástico con transiciones lentas.

Caso	<i>IPTG</i> (mpb)	Periodo
1	0	60 min
2	50	42.5 min
3	100	30 min
4	150	25 min
5	200	25 min
6	250	25 min

El modelo estocástico usando las transiciones lentas (modelo completo) tiene el mismo periodo que el medido experimentalmente (≈ 30 min), cuando se usa una concentración de *IPTG* de 100 moléculas y sin tener que ajustar el valor del retardo. Es importante hacer énfasis en este resultado puesto que Stricker et al. midieron experimentalmente un periodo de oscilación de aproximadamente 30 minutos en su oscilador sintético con un lazo de

retroalimentación negativa. Nosotros al desarrollar un modelo de reacciones químicas que reproduce el comportamiento oscilatorio, observamos que sólo incluyendo el ruido bioquímico y con una concentración de *IPTG* de 100 *mpb* nuestros resultados tienen congruencia con los resultados experimentales sin tener que ajustar parámetros. Por lo que decimos que el modelo estocástico que incluye todas las reacciones químicas propuestas en la tabla 1 de la sección 5, es el más cercano a la realidad biológica.

Modelo de transiciones rápidas.

Tabla 6. Periodos de oscilación en diferentes condiciones de *IPTG*, usando el modelo estocástico con transiciones rápidas descritas en las reacciones 22.

Caso	<i>IPTG</i> (mpb)	Periodo
1	0	27.27 min
2	50	21.42 min
3	100	18.75 min
4	150	13.63 min
5	200	13.63 min
6	250	11.53 min
7	300	10.71 min
8	350	10.34 min
9	400	10 min
10	500	10 min

En este modelo los periodos de oscilación están por debajo del valor medido experimentalmente.

En la figura 8 se concentra la información de los resultados de los modelos de transiciones rápidas y lentas (los dos casos considerados en el modelo estocástico). Ahí se puede observar que al usar transiciones lentas (modelo completo) en el modelo estocástico aumentan los periodos de oscilación en comparación con las transiciones rápidas (modelo simplificado); al aumentar la concentración de *IPTG* permanece ésta tendencia. Por lo que las transiciones en los estados del operador son un bloque que disminuye la velocidad o la frecuencia de las oscilaciones.

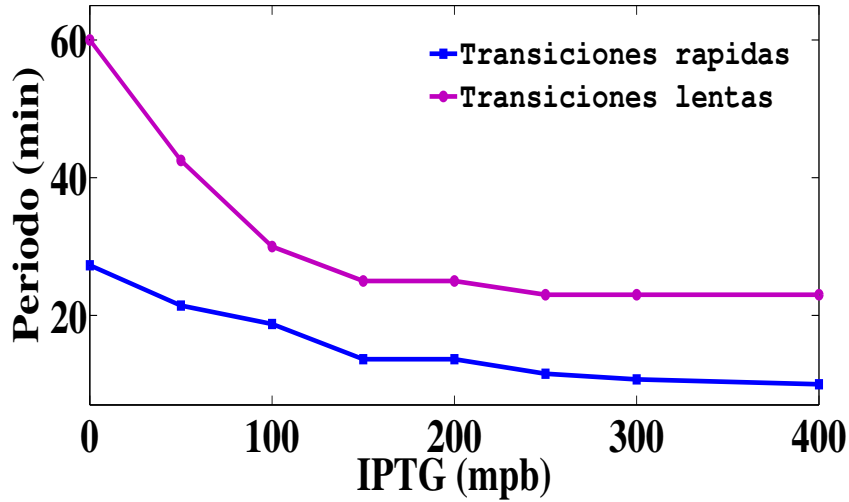


Figure 8: Resultados del modelo estocástico. En el eje horizontal se muestra el número de moléculas de *IPTG* y en el eje vertical el periodo de oscilación medido en una concentración de *IPTG* específica.

De lo anterior se concluye que el modelo estocástico es más realista y que el modelo que incluye a las reacciones que describen los cambios de estado del operador (modelo completo o de transiciones lentas) disminuye la velocidad de las oscilaciones y aumenta los periodos de oscilación.

7.4 Efecto de la fuerza de promotor en el periodo de oscilación

El modelo estocástico tiene congruencia con las mediciones experimentales medidas por Stricker et al. Al tener validado el modelo de reacciones químicas que se proponen, tanto con los rangos de regulación como con la dinámica del oscilador sintético construido por Stricker et al. nosotros queremos conocer si al usar un promotor sintético diferente a $P_{Llac0-1}$, existe algún efecto en la frecuencia de oscilación en el oscilador sintético con un lazo de retroalimentación negativa. Para lo cual usamos los promotores estudiados anteriormente: P_{lac} y P_{Allac0_1} .

Para simular el uso de un promotor distinto cambiamos el valor de las constantes de asociación y disociación de la polimerasa con el promotor, nombradas como k_B y k_b respectivamente. Estas constantes tienen un valor distinto para los promotores P_{lac} y P_{Allac0_1} (ver

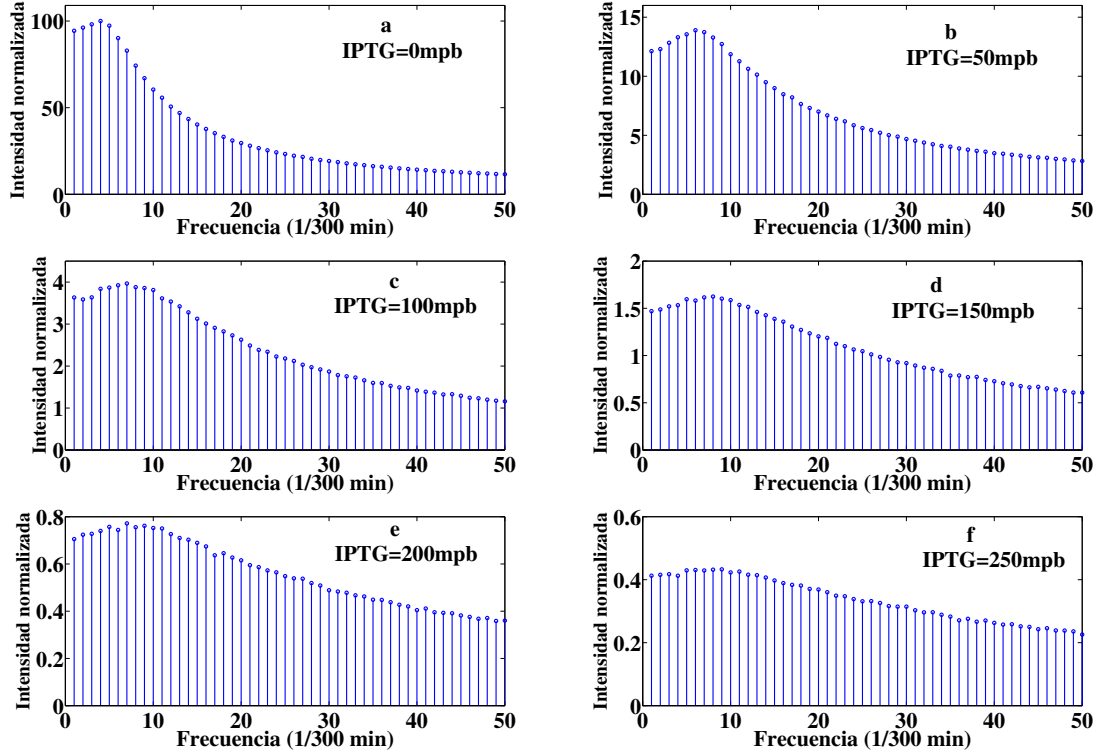


Figure 9: Espectro de potencias del conteo molecular del represor *lac* tetramérico R usando el promotor P_{lac} con diferentes concentraciones de IPTG: $0mpb$, **a**; $50mpb$, **b**; $100mpb$, **c**; $150mpb$, **d**; $200mpb$, **e**; $250mpb$, **f**. En el eje horizontal muestra la frecuencia en unidades de $\frac{1}{300 \text{ min}}$, y en el eje vertical se muestra la intensidad normalizada en unidades arbitrarias.

anexo 2).

7.4.1 P_{lac}

El promotor P_{lac} tiene un rango de regulación de 200, éste rango de regulación fué medido por Lutz y Bujard [30]. Se calcularon los espectros de potencia del conteo molecular de tetrámeros y monómeros de represor $LacI$, con el objetivo de observar si existe algún efecto en la frecuencia de oscilación en comparación con el promotor P_{Llac0_1} , (promotor que usó Stricker et al.).

En la figura 9 se muestran los espectros de potencias del conteo molecular del represor $LacI$ en su forma tetrámerica usando el promotor P_{lac} , el cual controla al gen $lacI$. Se observa que en comparación de los espectros calculados cuando se usa el promotor $P_{Llac0-1}$, la existencia del máximo local no es tan evidente como en los espectros calculados usando el

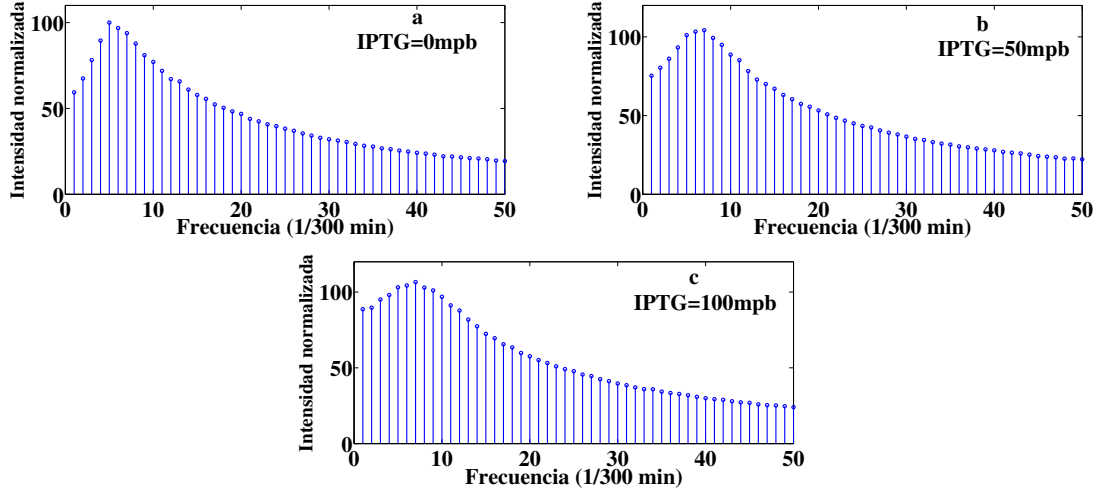


Figure 10: Espectro de potencias del conteo molecular del represor *lac* monomérico E usando el promotor P_{lac} con diferentes concentraciones de IPTG: $0mpb$, **a**; $50mpb$, **b**; $100mpb$, **c**. En el eje horizontal se indica la frecuencia en unidades de $\frac{1}{300 \text{ min}}$, y en el eje vertical se muestra la intensidad normalizada en unidades arbitrarias.

promotor $P_{LlacO-1}$. Por lo cual, podemos afirmar que si existen oscilaciones, pero serán muy difíciles de identificar en un experimento real, ya que van a ser enmascaradas por el comportamiento estocástico del ruido bioquímico. Lo anterior tiene congruencia con la figura 11, en la cual se muestran los espectros de potencia del conteo molecular de los monómeros usando este promotor. En la figura 10 si se pueden diferenciar los máximos locales, situando en una frecuencia de oscilación de $\frac{7}{300 \text{ min}}$ para una concentración de $IPTG$, $[I] = 100 \text{ mpb}$, lo que implica un periodo de oscilación de aproximadamente 42.8 min; pero para poder observar éstas oscilaciones en la concentración del represor monomérico en un experimento real, debemos tener un instrumento capaz de medir concentraciones de represores monoméricos dentro de una célula, ya que un gen reportero sería inútil en este caso. Con base a lo anterior hacemos notar que si nuestro modelo de reacciones es correcto, entonces no se deben de observar oscilaciones al usar el promotor P_{lac} y un gen reportero simple, y en el mejor de los casos estas oscilaciones deben estar fuertemente enmascaradas por el ruido bioquímico.

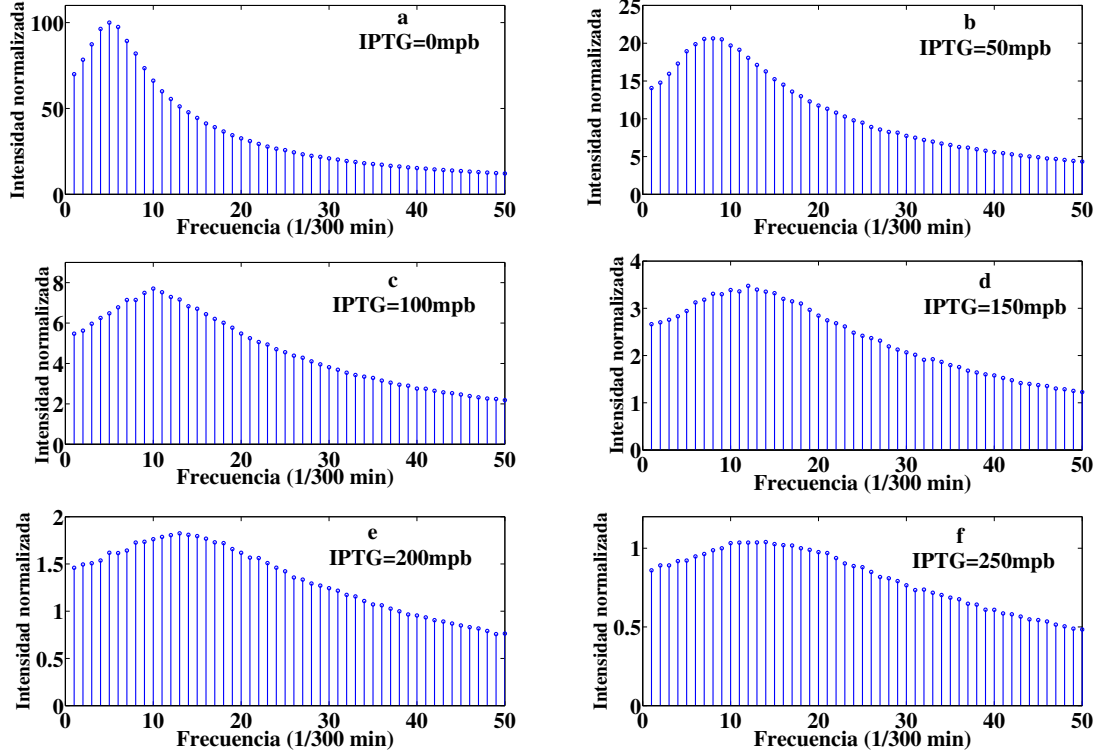


Figure 11: Espectro de potencias del conteo molecular del represor *lac* tetramérico *R* usando el promotor P_{A1lac0_1} con diferentes concentraciones de *IPTG*: $0mpb$, **a**; $50mpb$, **b**; $100mpb$, **c**; $150mpb$, **d**; $200mpb$, **e**; $250mpb$, **f**. En el eje horizontal se indica la frecuencia en unidades de $\frac{1}{300 \text{ min}}$, y en el eje vertical se muestra la intensidad normalizada en unidades arbitrarias.

7.4.2 P_{A1lac0_1}

El promotor P_{A1lac0_1} tiene un rango de regulación de 350, de igual manera se calculan los espectros de potencia del conteo molecular de tetraméros y monómeros del represor *LacI*. Se usa éste promotor para estudiar si existe algún efecto en los periodos de oscilación del oscilador sintético al usar el operador P_{A1} .

En comparación con el promotor P_{Llac0_1} , el rango de regulación es menor. Al cambiar el promotor por el P_{A1lac0_1} las oscilaciones se mantienen y aproximadamente en la concentración de *IPTG*, $[I] = 100mpb$ la frecuencia de resonancia es de $\frac{1}{30}$ (ver figura 11).

Lo anterior nos sugiere que si en nuestro modelo de reacciones propuesto se usa el promotor P_{A1lac0_1} se esperaría que el periodo de oscilación fuera casi igual al periodo medido por Stricker et al. cuando usan el promotor P_{Llac0_1} .

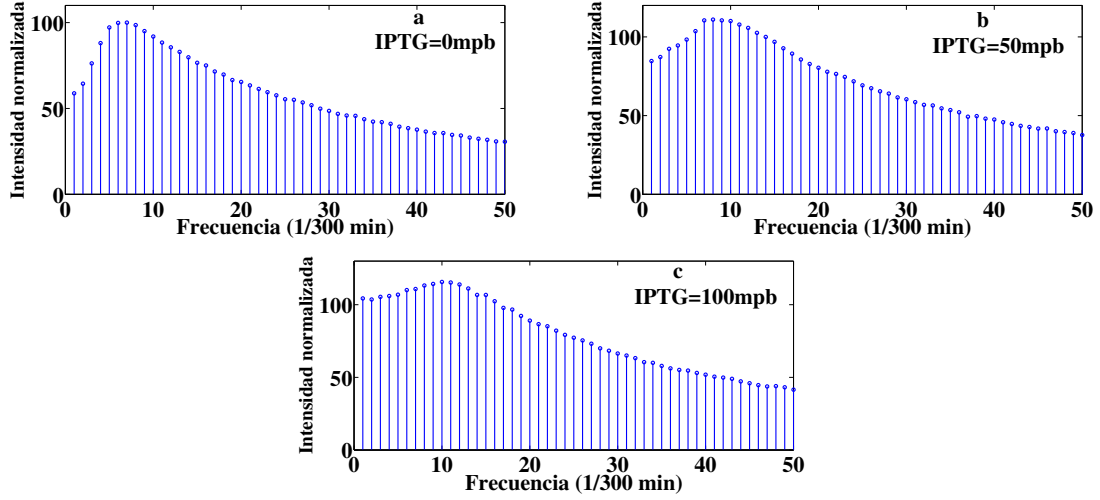


Figure 12: Espectro de potencias del conteo molecular del represor *lac* monomérico E usando el promotor $P_{A11lac0_1}$ con diferentes concentraciones de $IPTG$: $0mbp$, **a**; $50mbp$, **b**; $100mbp$, **c**. En el eje horizontal se indica la frecuencia en unidades de $\frac{1}{300 \text{ min}}$, y en el eje vertical se muestra la intensidad normalizada en unidades arbitrarias.

Al calcular los espectros de frecuencias del conteo molecular de los monómeros usando el promotor $P_{A11lac0_1}$, la amplitud las oscilaciones tampoco disminuye al variar la concentración el $IPTG$ (ver figura 12), como lo observado usando el promotor P_{Llac0_1} .

Por lo anterior, se observa que al usar diferentes promotores sintéticos con diferentes rangos de regulación, el comportamiento oscilatorio permanece al usar los promotores P_{lac} y $P_{A11lac0_1}$. Más aún, el periodo de oscilación medido con los experimentos propuestos se acerca al periodo medido experimentalmente por Stricker et al. al usar el promotor P_{Llac0_1} (de aproximadamente 30 min). También podemos decir al usar el promotor P_{lac} , promotor que tiene un rango de regulación menor en comparación de los otros (de 200) las oscilaciones son más ruidosas.

8 Discusión y conclusiones

Es sabido que los modelos deterministas no son el método adecuado para modelar sistemas con bajo conteo molecular pues es necesario tomar en cuenta la aleatoriedad intrínseca de las reacciones químicas. En éste estudio se comprobó que el modelo estocástico donde se toma en cuenta al ruido bioquímico se acerca más a la realidad biológica, pues se encuentra que los periodos de oscilación medidos a través del cálculo de los espectros de potencia de los conteos moleculares tienen congruencia con los periodos medidos experimentalmente por Stricker et al [9] de aproximadamente 30 min.

Del modelo matemático de este sistema (oscilador genético sintético simple), se concluye que es importante tomar en cuenta al ruido bioquímico para modelar éste sistema y que el ruido bioquímico es necesario para obtener un comportamiento oscilatorio sin tener que ajustar el retardo (τ). Ésto es congruente con resultados anteriores, donde se afirma que el modelado determinista no es apropiado para modelar sistemas con bajo conteo molecular [56], como lo es nuestro sistema, el cual como lo vimos en los diferentes casos tiene un número de represores en orden de centenas. En comparación del modelo que publicó Stricker et al. con el nuestro, hacemos una notable diferencia en los parámetros utilizados para el sistema de reacciones propuesto en éste trabajo, pues no necesitamos ajustar los parámetros de cinética química y son obtenidos de la literatura con referencia a experimentos *in vitro*.

El modelo que es congruente con la realidad biológica es cuando se consideran las transiciones en los estados del operador (llamado de transiciones lentas) y al ruido bioquímico. Al realizar las simulaciones considerando las transiciones en los estados del operador, se observa un efecto en la velocidad de las oscilaciones. Es fácil ver que al tomar en consideración los cambios en el estado del operador las frecuencias de oscilación disminuyen, por lo que este bloque de reacciones hace lentas las oscilaciones del sistemas y aumenta los periodos de oscilación.

Aunado a lo anterior es sabido que el cálculo del espectro de potencias es ampliamente usado para el análisis de señales eléctricas. Nuestro modelo estocástico, simulado por medio de un algoritmo de simulación estocástica nos proporciona el cambio de moléculas a través del tiempo (trayectoria), tomando en cuenta la naturaleza discreta del cambio en los conteos moleculares a través del tiempo. Sin embargo, las periodicidades en una trayectoria se en-

cuentran enmascaradas por el ruido bioquímico, por lo que el espectro de potencias de estas trayectorias hace posible encontrar las periodicidades escondidas y con ello determinar la frecuencia y periodo de oscilación cuando el sistema de reacciones es visto como un proceso estocástico (debido a la influencia del ruido bioquímico). Aunque, el método del cálculo de la potencia espectral ya ha sido estudiado y aplicado en otros ámbitos, en este trabajo lo aplicamos a un sistema oscilatorio genético encontrando las periodicidades.

Por otro lado, diferentes sistemas genéticos construidos *sintéticamente* se basan en los promotores de Lutz y Bujard, éstos promotores se encuentran en los vectores *pZ* comercializados por la compañía EXPRESSYS. El modelo matemático desarrollado para los promotores P_{Llac0_1} , P_{A1lac0_1} y P_{lac} es exacto, flexible y simple. Es exacto, porque al sustituir los valores de los parámetros cinéticos medidos experimentalmente, encontramos que los rangos de regulación predichos por nuestro modelo son casi idénticos a los rangos medidos experimentalmente para tres promotores distintos. El hecho que el mismo modelo funcione y haya predicho correctamente los rangos de regulación de tres promotores distintos nos permite afirmar que nuestro modelo también es flexible. Finalmente, el modelo es simple puesto que en su construcción solamente se usan tres reacciones (reversibles) para describir la competencia que existe entre la polimerasa y el represor *LacI* por ocupar su sitio de unión en el promotor. Este modelo puede ser utilizado como un bloque que describe a la transcripción y represión de la transcripción en modelos de sistemas genéticos más complejos que usen éstos promotores sintéticos construidos por Lutz y Bujard.

Más aún, el hecho que el modelo es flexible y exacto nos permite afirmar que las suposiciones hechas al construir el modelo son válidas. En particular, podemos deducir que la cooperatividad es virtualmente nula entre dos operadores *lac* cuando sus pares de bases centrales están separados por 22 ó 32 pb, pues ésta es una de las suposiciones hechas al construir nuestro modelo. Recordemos que dos operadores *lac* cooperan para producir una represión mayor cuando un represor *LacI* en su forma tetramérica se une a los dos formando un loop de represión; así que la cooperatividad es la probabilidad de que un mismo represor se una a dos operadores. Müller et al. han encontrado experimentalmente que, para dos operadores *lac* cuyos pares de bases centrales tienen una separación de entre 57 a 92 pb, los máximos y mínimos de la cooperatividad están distribuidos periódicamente [54]. Ésto es, la cooperatividad presenta un máximo cuando la separación entre los pares de bases centrales

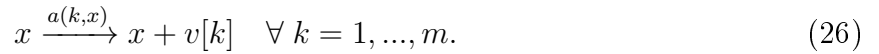
de los operadores es de 59, 70, 81 y 92 pb; mientras que la cooperatividad presenta un mínimo cuando la separación es de 63, 74, 84 y 96 pb. Nuestros resultados extienden los resultados por Müller et al. para separaciones de 22 y 32 pb. El hecho de que nuestro modelo es flexible y exacto también nos permite afirmar que es imposible que dos represores *LacI* se unan simultáneamente a dos operadores *lac* cuando sus pares de bases centrales están separadas por 32 pb o menos.

Finalmente, dado que nuestro modelo es predictivo, podemos hacer predicciones sobre el periodo de oscilación que se obtendría si se usaran los promotores P_{lac} y $P_{A1lacO-1}$ en la topología del oscilador sintético construido por Stricker et al. Con lo cual, se verificaría que efecto tienen los rangos de regulación en el comportamiento oscilatorio del sistema genético sintético.

9 Anexos

9.1 Anexo 1. Estabilidad de la Ecuación Química Maestra

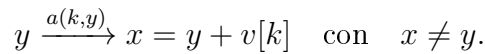
Para demostrar que la ecuación química maestra (EQM) es estable necesitamos calcular los eigenvalores de la matriz B asociada a la EQM, la cual es una ecuación diferencial lineal $\dot{P} = BP$. El primer paso entonces es calcular la matriz B . Recordemos que las m reacciones químicas se representan por:



La matriz B se construye a partir de las propensiones $a(k, x)$

$$B[x, y] = \begin{cases} -\sum_{k=1}^m a(k, x) & \text{si } x = y, \\ a(1, y) & \text{si } x = y + v[1], \\ \vdots & \vdots \\ a(m, y) & \text{si } x = y + v[m], \\ 0 & \text{en otro caso.} \end{cases}$$

Notese que la entrada $B[x, y]$ es igual a la propensión $a(k, y) > 0$ cuando existe una reacción química,



Es importante remarcar que la entrada $B[x, y]$ representan una reacción química que va del estado y al estado x . Por otro lado, si $x = y$ la entrada $B[x, x]$ es estrictamente negativa e igual a $-\sum_{k=1}^m a(k, x)$. Las entradas de la diagonal principal son negativas ya que la suma de los elementos en una columna es igual a cero, esto es:

$$\begin{aligned} \sum_{x \in \Sigma^n} B[x, y] &= a(1, y) + a(2, y) + \dots + a(n, y) - \sum_{k=1}^m a(k, y) \\ &= 0 \end{aligned}$$

Nosotros sabemos que B es la matriz asociada a la EQM ya que:

$$\begin{aligned}
 \frac{dP(x, t)}{dt} &= \sum_{y \in \Sigma^n} B[x, y]P(y, t) \\
 &= B[x, x]P[x, t] + \sum_{y \neq x} B[x, y]P(y, t) \\
 &= - \sum_{k=1}^m a(k, x)P[x, t] + \sum_{k=1}^m a(k, x - v[k])P(x - v[k], t)
 \end{aligned}$$

Como podemos ver, al considerar todos los casos de la matriz B obtenemos la ecuación química maestra (EQM).

Procedamos ahora a calcular los eigenvalores λ de la matriz B . Recordemos que un eigenvalor $\lambda \in \mathbb{C}$ y su eigenvector asociado $V \in \mathbb{C}^{(\sigma+1)^n}$ son un par que satisface $BV = \lambda V$. Un resultado que vamos a usar todo el tiempo es el hecho que la matriz B y su transpuesta B' tienen las dos exactamente los mismos eigenvalores (aunque no los mismos eigenvectores).

Empecemos por demostrar que $\lambda = 0$ es un eigenvalor de B , lo cual es equivalente a demostrar que $\lambda = 0$ es un eigenvalor de la transpuesta B' , pero también es equivalente a demostrar que existe un eigenvector $V \neq 0$ tal que $B'V = 0$.

El eigenvector buscado es

$$V = (1, 1, \dots, 1)'$$

ya que

$$\begin{aligned}
 (B'V)[x] &= \sum_{y \in \Sigma^n} B'[x, y]V[y] = \sum_{y \in \Sigma^n} B[y, x] \\
 &= a(1, x) + a(2, x) + \dots + a(n, x) - \sum_{k=1}^m a(k, x) = 0
 \end{aligned}$$

Por lo tanto 0 es un eigenvalor.

Recordemos que las entradas de la transpuesta $B'[x, y] = B[y, x]$ son:

$$B'[x, y] = \begin{cases} -\sum_{k=1}^m a(k, y) & \text{si } y = x, \\ a(1, x) & \text{si } y = x + v[1], \\ a(2, x) & \text{si } y = x + v[2], \\ \vdots & \vdots \\ a(m, x) & \text{si } y = x + v[m], \\ 0 & \text{en otro caso.} \end{cases}$$

De lo anterior podemos ver que $B'V = 0$, donde V es el eigenvector de B' asociado a $\lambda = 0$. Esto significa que $\lambda = 0$ también es un eigenvalor de B . Suponiendo que W es un eigenvector de B asociado a $\lambda = 0$, entonces $BW = 0$ y más aun W es un múltiplo de las distribuciones en estado estacionario de la ecuación química maestra $\dot{P} = BP$.

Calculemos los eigenvalores $\lambda \neq 0$ de la matriz B . Como B y su transpuesta B' tienen los mismos eigenvalores, pero tal vez no los mismos eigenvectores, podemos tomar a $\lambda \neq 0$ como un eigenvalor de B' . Sea $V \in \mathbb{C}^{(\sigma+1)^n}$ un eigenvector no cero asociado al eigenvalor $\lambda = \alpha + ib$. Así que $B'V = (\alpha + ib)V$. De otra forma, para todo $x \in \Sigma^n$:

$$\begin{aligned} (\alpha + ib)V[x] &= (B'V)[x] = \sum_{y \in \Sigma^n} B'[x, y]V[y] \\ &= \sum_{k=1}^m \left(-a(k, x)V[x] + a(k, x)V[x + v[k]] \right) \end{aligned}$$

Reagrupando algunos términos podemos escribir:

$$\left[\alpha + ib + \sum_{k=1}^m a(k, x) \right] V[x] = \sum_{k=1}^m a(k, x)V[x + v[k]]$$

Existe un elemento particular $V[z]$ tal que $0 \neq |V[z]| \geq |V[x]|$ para todo $x \in \Sigma^n$ (el vector V es finito, así que entre sus elementos siempre hay un máximo). Aplicando el valor absoluto a la ecuación anterior (y usando las propiedades del valor absoluto $|xy| = |x| |y|$, $|x+y| \leq |x| + |y|$) se obtiene la siguiente fórmula:

$$\begin{aligned}
\left| \alpha + ib + \sum_{k=1}^m a(k, z) \right| |V[z]| &= \left| \left[\alpha + ib + \sum_{k=1}^m a(k, z) \right] V[z] \right| \\
&= \left| \sum_{k=1}^m a(k, z) V[z + v[k]] \right| \\
&\leq \sum_{k=1}^m |a(k, z) V[z + v[k]]| \\
&= \sum_{k=1}^m |a(k, z)| |V[z + v[k]]| \\
&\leq \sum_{k=1}^m a(k, z) |V[z]|
\end{aligned}$$

Factorizando el término $|V[z]| \neq 0$, obtenemos:

$$\left| \alpha + ib + \sum_{k=1}^m a(k, z) \right| \leq \sum_{k=1}^m a(k, z)$$

Definamos:

$$r := \sum_{k=1}^m a(k, z) > 0.$$

Por todo lo anterior podemos deducir que $|\alpha + r + ib| \leq r$. Elevando al cuadrado:

$$r^2 \geq |\alpha + r + ib|^2 = (\alpha + r)^2 + b^2 = \alpha^2 + 2\alpha r + r^2 + b^2.$$

Entonces:

$$0 \geq \alpha^2 + b^2 + 2\alpha r.$$

Esta igualdad solamente se satisface si $\alpha < 0$, o si $\alpha = b = 0$.

Recordemos que el eigenvalor $\lambda = \alpha + ib$, entonces tenemos que $Re(\lambda) = \alpha < 0$ ó $\lambda = 0$.

De lo anterior concluimos que la EQM es estable si el eigenvalor $\lambda = 0$ es único.

En general los sistemas en la naturaleza son estables. Debido a esto es importante demostrar que la ecuación química maestra es estable, ya que en nuestro caso representa a un conjunto de reacciones de un sistema

9.2 Anexo 2. Parámetros para el modelo matemático

9.2.1 Parámetros medidos in vitro para el promotor P_{Llac01}

1. *Tasa de asociación del complejo represor LacI/operador, k_A .* Riggs et al. midieron in vitro $7(\pm 0.9) \times 10^9/M/s$. Tomamos el valor de $k_A = 7 \times 10^9/M/s$ [37].

2. *Tasa de disociación del complejo represor LacI/operador, k_a .* Riggs et al. midieron in vitro $6.2(\pm 1.3) \times 10^{-4}/s$. Tomamos el valor de $k_a = 6.2 \times 10^{-4}/s$ [37].
3. *Tasa de asociación del complejo polimerasa/promotor P_{Llac0_1} , k_B .* Lanzer y Bujard et al. (1988) midieron $1.4(\pm 0.2) \times 10^7/M/s$ para el promotor P_L con un operador *lac* entre -10 y -33 hexámeros (ver [46], p.8975). Tomamos el valor de $k_B = 1.4 \times 10^7/M/s$.
4. *Tasa de disociación del complejo polimerasa/promotor P_{Llac0_1} , k_b .* Brunner y Bujard (1987) midieron $k_b = 6.8 \times 10^{-4}/s$ para el promotor P_L (ver [47], p. 3141).
5. *Tasa de inicio de la transcripción k_m .* McClure (1980) observaron un retardo de 60 s en la apertura del complejo polimerasa/promotor P_L , por lo que tomamos el valor $k_m = 2/60s$ [41].
6. *Número de polimerasas por célula $[P]$.* Bremer y Dennis (1996) midieron entre 1500-11400 polimerasas por célula. Tomamos el valor de $[P] = 8 \times 10^3 mpb$ [45].
7. *Número de represores por célula $[R]$.* Müller et al. (1968) determinaron que un promotor mutante *laci^q* induce 80 y 100 moléculas de represor *LacI* por célula [44]. Nosotros tomamos $[R]=90$ represores por célula.

9.2.2 Parámetros medidos in vitro para el promotor $P_{A1lac0-1}$

1. *Tasa de asociación del complejo represor LacI/operador, k_A .* Riggs et al. midieron in vitro $7(\pm 0.9) \times 10^9/M/s$. Tomamos el valor de $k_A = 7 \times 10^9/M/s$ [37].
2. *Tasa de disociación del complejo represor LacI/operador, k_a .* Riggs et al. midieron in vitro $6.2(\pm 1.3) \times 10^{-4}/s$. Tomamos el valor de $k_a = 6.2 \times 10^{-4}/s$ [37].
3. *Tasa de asociación del complejo polimerasa/promotor P_{A1lac0_1} , k_B .* Lanzer y Bujard et al. midieron $6(\pm 0.5) \times 10^7/M/s$ para el promotor $P_{A1/O4}$ con un operador *lac* entre -10 y -33 hexámeros (ver [46], p.8975). Tomamos el valor de $k_B = 6 \times 10^7/M/s$.
4. *Tasa de disociación del complejo polimerasa/promotor P_{A1lac0_1} , k_b .* Brunner y Bujard (1987) midieron $k_b = 9.2 \pm 10^{-4}/s$ para el promotor P_{A1} (ver [47], p. 3141).

5. *Tasa de inicio de la transcripción k_m* . Dayton et al. (1984) midieron la tasa de apertura del ADN de 0.09/s para el promotor bacteriófago P_{A1} , por lo que tomamos el valor $k_m = 0.09/s$ [40].
6. *Número de polimerasas por célula $[P]$* . Bremer y Dennis (1996) midieron entre 1500-11400 polimerasas por célula. Tomamos el valor de $[P] = 9 \times 10^3$ [45].
7. *Número de represores por célula $[R]$* . Müller et al. (1968) determinaron que un promotor mutante $laci^q$ induce 80 y 100 moléculas de represor $LacI$ por célula [44]. Nosotros tomamos $[R]=90$ represores por célula.

9.2.3 Parámetros medidos in vitro para el promotor P_{lac}

1. *Tasa de asociación del complejo represor $LacI$ /operador, k_A* . Riggs et al. midieron in vitro $7(\pm 0.9) \times 10^9/M/s$. Tomamos el valor de $k_A = 7 \times 10^9/M/s$ [37].
2. *Tasa de disociación del complejo represor $LacI$ /operador, k_a* . Riggs et al. midieron in vitro $6.2(\pm 1.3) \times 10^{-4}/s$. Tomamos el valor de $k_a = 6.2 \times 10^{-4}/s$ [37].
3. *Tasa de asociación del complejo polimerasa/promotor P_{lac} , k_B* . Lanzer y Bujard (1988) midieron *in vitro* $2(\pm 1) \times 10^6/M/s$ para el promotor P_{lac} [3, p.8975]. Tomamos el valor de $k_B = 2 \times 10^6/M/s$ [46].
4. *Tasa de disociación del complejo polimerasa/promotor P_{lac} , k_b* . Meiklejohn and Gralla (1985) midieron un tiempo de vida media de 3.5 minutos del complejo en una cepa wild type en presencia de CRP-cAMP (ver [21], p.8975). Tomamos $k_b = \frac{\ln 2}{210s} = 0.0033/s$.
5. *Tasa de inicio de la transcripción k_m* . Malan et al. (1984) midieron la tasa de apertura del ADN de 0.003/s para el promotor lac silvestre con CRP-cAMP, por lo que tomamos el valor $k_m = 0.09/s$ [42].
6. *Número de polimerasas por célula $[P]$* . Bremer y Dennis (1996) midieron entre 1500-11400 polimerasas por célula. Tomamos el valor de $[P] = 8300mpb$ [45].
7. *Número de represores por célula $[R]$* . Oehler et al. (1994) inducen 200 subunidades monoméricas por célula, lo cual corresponde a 50 tetrámeros de represor por célula [48].

9.3 Parámetros usados para el modelo del oscilador genético sintético construido por Stricker et al.

Para el oscilador genético construido por Stricker et al. usan el promotor P_{LacO_1} , algunos parámetros fueron determinados en el apartado 9.2.1 de este anexo y se completan con la siguiente lista.

1. *Volumen de E. coli.* Volkmer y Heinemann (2011) midieron el volumen de una *E. coli* en diferentes condiciones de crecimiento, ellos reportaron una variación en el volumen de 1.6 a 4.4 *fl* [55]. Nosotros tomamos un volumen medio de 3 *fl*, por lo que hacemos la siguiente conversión de unidades:

$$1M = \frac{6.02 \times 10^{23} \text{ molecules}}{1 \text{ l}} \times \frac{3 \times 10^{-15} \text{ l}}{1 \text{ bacteria}} = 1.8 \times 10^9 \text{ mpb}$$

2. *Número de genes estructurales LacI por célula.* Existe un sólo gen estructural *lacI* en cada vector *pZ*. Stricker et al. determina que hay entre 1 a 3 vectores *pZ* por célula, por lo que tomamos 2 genes estructurales *lacI* [9].
3. *Tasa de inicio de la traducción del monomero LacI, k_E .* Chen et al. (1991) determinaron el valor de $k_E = 10/\text{min}$ [50].
4. *Tasas de dimerización y tetramerización, k_D y k_R .* Usamos los valores $k_D = k_R = 0.018/\text{mpb}/\text{min}$ obtenidos por Stricker et al. (material suplementario) [9].
5. *Tasa de disociación de dímeros y tetrámeros, k_d y k_r .* Éstos valores son muy pequeños en comparación de $k_D = k_R$, por lo que $k_d = k_r = 0/\text{mpb}/\text{min}$.
6. *Tasa de asociación del complejo inductor IPTG/represor LacI, k_I .* Dunaway et al. (1980) midieron *in vitro* el valor de $k_I = 4.6 \times 10^4/M/s$ [51]
7. *Tasa de disociación del complejo inductor IPTG/represor LacI, k_i .* Este valor es muy pequeño en comparación de k_I , por lo que $k_i = 0/\text{min}$.
8. *Tasa de degradación del ARN mensajero lacI, γ_M .* Varmus et al. midieron el tiempo de vida media de aproximadamente 1 minuto para el ARNm de *lacI*, por lo que $\gamma_M \approx \frac{\ln(2)}{1\text{min}}$ [52].

9. *Tasa de degradación del represor LacI*, γ_E , γ_D y γ_R . Asumimos que el monómero E , dímero D y tetrámero R del represor $LacI$ tienen la misma tasa de degradación. Chen et al. (1991) midió ésta tasa de degradación, $\gamma_E = \gamma_D = \gamma_R = 0.1/min$ [50].
10. *Tiempo de retardo en la traducción del monómero de LacI*, τ_E . El represor $LacI$ en su forma monomérica es una proteína de 360 aminoácidos. Bremmer and Dennis (1996) determinaron que la elongación de la cadena polipeptídica se encuentra entre 12 y 21 aminoácidos por segundo, por que el tiempo de retardo en la traducción $\tau_E \approx \frac{360 \text{ res.}}{16 \text{ res./s}} = 22.5s$ [45].

References

- [1] Choi. *Introduction to Systems biology*, 2007.
- [2] Lehninger. *Principles of biochemistry*, 2005.
- [3] Karp Gerald. *Biología celular y molecular: Conceptos y experimentos*, 2006.
- [4] Alberts et al., *Molecular biology of the cell*, 3rd ed., Garland Publishing, New York, 1994
- [5] Tan Chee Meng, Sandeep Somani and Pawan Dh. *Modelling and analysis of gene regulatory networks*, 2008.
- [6] Michael B. Elowitz et al. Stochastic gene expression in a single cell. *Science* 297, 1183, 2002.
- [7] A synthetic oscillatory network of transcriptional regulators Michael B. Elowitz and Stanislas Leibler *Nature* 403, 335-338, January 2000.
- [8] Hooshangi S, Stephan Thiberge, Ron Weiss. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci USA*, 102, 10, 3581-3586, 2005.
- [9] A fast, robust and tunable synthetic gene oscillator. Jesse Stricker, Scott Cookson, Matthew R. Bennett, William H. Mather, Lev S. Tsimring and Jeff Hasty. *Nature* 456, 516-519, November 2008.
- [10] Gilbert Walter and Benno Müller-Hill. *PNAS, Biochemistry*. vOL. 56, 1966.
- [11] Michael L. Simpson, Chris D. Cox, Michael S. Allen, James M. McCollum, Roy D. Dar, David K. Karig, John F. Cooke *Noise in biological circuits*. Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology. Vol. 1, pp 214–225, April 2009
- [12] Mads Kaern, Timothy C. Elston, William J. Blake and James J. Collins. *Stochasticity in gene expression: from theories to phenotypes*. *Nature Reviews Genetics*. Vol. 6, pp 451-464, April 2009
- [13] Narendra Maheshri and Erin K. O Shea *Living with Noisy Genes: How Cells Function Reliably with Inherent Variability in Gene Expression*. *Annual Review of Biophysics and Biomolecular Structure*. Vol. 36, pp.413-434, 2007.

- [14] Jonathan M. Raser and Erin K. O'Shea. *Noise in gene expression: origins, consequences and control*. Science. Vol. 45, pp. 41-56, 2005.
- [15] Long Cai, Nir Friedman, X. Sunney Xie. *Stochastic protein expression in individual cells at the single molecule level*. Nature. Vol 440, March 2006
- [16] Xie, Choi, Gene-Wei Li, Nam Ki Lee and Giuseppe Lia. *Single-molecule approach to molecular biology in living bacterial cells*. Annu. Rev. Biophysics. Vol. 37, pp. 417-444, 2008.
- [17] *RNA Polymerase-Promoter Interactions: the Comings and Goings of RNA Polymerase*. Journal of Bacteriology. June 1998, p. 3019–3025.
- [18] Ramon Grima and Santiago Schnell *Modelling reaction kinetics inside cells*. Essays Biochem. Vol. 106, No. 52, pp. 22275-22280, 2009.
- [19] Naoma Barkai and Stanislas Leibler. *Biological Rhythms: Circadian clocks limited by noise*. Nature. Vol. 403, pp. 267-268, 2000.
- [20] Adam Arkina, John Rossb, and Harley H. McAdamsa. *Stochastic Kinetic Analysis of Developmental Pathway Bifurcation in Phage λ Infected Escherichia coli Cells*. Genetics. Vol. 149, pp. 1633-1648, 1998
- [21] A. L. Meiklejohn, J. D. Gralla. Entry of rna polymerase at the lac promoter, Cell 43 (3), 769-776, 1985.
- [22] Ji Yu, Jie Xiao, Xiaojia Ren, Kaiqin Lao, X. Sunney Xie. *Probing gene expression in live cells, one protein molecule at a time*. Science. Vol. 311 no. 5767 pp. 1600-1603, 2006
- [23] Jhon G. Proakis and Dimitris G. Manolakis. *Digital Signal Processing. Principles, Algorithms and Applications*. Third Edition. Prentice-Hall. pp. 230-231, 1998.
- [24] Daniel T. Gillespie. *Exact stochastic simulation of coupled chemical reactions*. J. Phys. Chem. Vol. 81, pp. 2340–2361, 1977.
- [25] Daniel T. Gillespie. *The chemical Langevin and Fokker Planck Equation for the reversible isomerization reaction*. J. Phys. Chem. Vol. 106, pp. 5063-5071, 2002.

- [26] Daniel T. Gillespie. *The chemical Langevin equation*. J. Phys. Chem. Vol. 113, pp. 297-306, 2000.
- [27] DT Gillespie. *A rigorous derivation of the chemical master equation*. Physica A: Statistical Mechanics and its Applications. Vol. 188, pp. 404-425, 1992
- [28] E. S. Zeron and M. Santillán. *Numerical Solution of the Chemical Master Equation: Uniqueness and Stability of the Stationary Distribution for Chemical Networks, and mRNA Bursting in a Gene Network with Negative Feedback Regulation*. Methods in Enzymology, Volume 487, 2011.
- [29] Athanasios Papoulis *Probability, Random Variables and Stochastic Processes*. 3rd Edition, 1991.
- [30] Lutz and Bujard. *Independent and Tight Regulation of Transcriptional Units in Escherichia Coli Via the LacR/O, the TetR/O and AraC/I1-I2 Regulatory Elements*. Nucleic Acids Research. 1997.
- [31] Expressys: Tools for gene expression, <http://http://www.expressys.com>.
- [32] Mitchell Lewis. *The lac repressor*. pp. 521–548, 2005.
- [33] A.D. Riggs, S.Bourgeois, M.Cohn. *The lac repressor-operator interaction: iii. Kinetic studies*. J. Mol. Biol. 53 (3), pp. 401-417, 1970.
- [34] Jiguo Cao, Xin Qi, and Hongyu Zhao. *Modeling Gene Regulation Networks Using Ordinary Differential Equations*. Methods in Molecular Biology, vol. 802, 2012.
- [35] Hidde de Jong. *Modeling and Simulation of Genetic Regulatory Systems: A Literature Review*. Journal of Computational Biology. Volume 9, Number 1, Pp. 67–103, 2002.
- [36] Jesse Stricker, Scott Cookson, Matthew R. Bennett, William H. Mather, Lev S. Tsimring and Jeff Hasty. *Material suplementario de: A fast, robust and tunable synthetic gene oscillator*. Nature 456, 516-519, November 2008.
- [37] AD Riggs, S Bourgeois, M Cohn. *The lac repressor-operator interaction: III. Kinetic studies*. Journal of molecular biology, Vol. 53, pp 401-417, 1970.

- [38] R Knaus, H Bujard. *PL of coliphage lambda: an alternative solution for an efficient promoter*. The EMBO journal, Vol. 7, pp 2919, 1988.
- [39] F Ceroni, S Furini, S Cavalcanti. *A computational model of gene expression in an inducible synthetic circuit*. World Scientific, Vol. 15, pp 409-420, 2010.
- [40] C. J. Dayton, D. E. Prosen, K. L. Parker, C. L. Cech. *Kinetic measurements of Escherichia coli rna polymerase association with bacteriophage t7 early promoters*. J. Biol. Chem. 259 (3), 1616-1621, 1984.
- [41] W. R. McClure. *Rate-limiting steps in rna chain initiation*. Proc. Natl. Acad. Sci. U.S.A. 70 (10), p. 5634-5638, 1980.
- [42] T. Malan, A. Kolb, H. Buc, W. R. McClure. *Mechanism of crp-camp activation of lac operon transcription initiation activation of the p1 promoter*. J. Mol. Biol. 180 (4), 881-909, 1984 .
- [43] Gilbert and Müller-Hill. *Isolation of the lac repressor*. Proceedings of the National Academy of Sciences of the United States of America, Vol. 56, pp 1891-1898, 1966.
- [44] Benno Müller-Hill, Lawrence Crapo and Walter Gilbert. *Mutants that make more Lac repressor*. Proceedings of the National Academy of Sciences of the United States of America, Vol. 59, pp 1259-1264, 1968.
- [45] H Bremer and PP Dennis. *Modulation of chemical composition and other parameters of the cell by growth rate*. Escherichia coli and Salmonella: cellular and molecular biology, Vol. 2, pp 1553-1569, 1996.
- [46] M Lanzer and H Bujard . *Promoters largely determine the efficiency of repressor action*. Proceedings of the National Academy of Sciences, Vol.85, pp 8973-8977, 1988.
- [47] M Brunner and H Bujard . *Promoter recognition and promoter strength in the Escherichia coli system*. The EMBO journal, Vol.6, pp 3139, 1987.
- [48] S Oehler, M Amouyal, P Kolkhof, B von Wilcken-Bergmann, and B Müller-Hill. *Quality and position of the three lac operators of Ecoli define efficiency of repression*. The EMBO journal, Vol. 13, pp. 3348-3355, 1994.

- [49] Wolfgang Hillen, Christiane Gatz, Lothar Altschmied, Klaus Schollmeier, Iris Meier. *Control of expression of the Tn10 encoded tetracycline resistance genes: Equilibrium and kinetic investigation of the regulatory reactions*. J. Mol. Biol., Vol. 169, pp 707-721, 1983.
- [50] W.Chen, J.E. Bailey, S.B. Lee. *Molecular design of expression systems: Comparison of different control configurations using mechanism models*. Biotechnol. Bioeng. 38 (7) (1991) pp 679-687.
- [51] M. Dunaway, J. Olson, J. Rosenberg, O. Kallai, R. Dickerson, K. Matthews. *Kinetic studies of inducer binding to lac repressor-operator complex*. J. Biol. Chem. 255 (21) (1980) pp 10115-10119.
- [52] H.E. Varmus, R.L. Perlman, I. Pastan. *Regulation of lac messenger ribonucleic acid synthesis by cyclic adenosine 3', 5'- monophosphate and glucose*. J. Biol. Chem. 245 (9) (1970) 2259-2267.
- [53] T.Philip Malan, Annie Kolb, Henri Buc, William R. McClure. *Mechanism of CRP-cAMP activation of lac operon transcription initiation activation of the P1 promoter*. Journal of molecular biology, Vol. 180, pp 881-909, 1984.
- [54] Johannes Müller, Stefan Oehler and Benno Müller Hill. *Repression of lac Promoter as a function of distance, phase and quality of an auxiliary lac operator*. J.Mol.Biol., 1996.
- [55] Volkmer y Heinemann. *Condition-dependent cell volume and concentration of Escherichia coli to facilitate data conversion for systems biology modeling*. PlosONE, 2011.
- [56] Manuel Barrio, Kevin Burrage, André Leier, Tianhai Tian. *Oscillatory regulations of Hes1: Discrete Stochastic Delay Modelling and Simulation*. PLOS Computational Biology. 2006.
- [57] A. Hernández-Valdez, M. Santillán, and E. S. Zeron. *Cycling expression and cooperative operator interaction in the trp operon of Escherichia coli*. J. Theor. Biol. 263:340-352, 2010.