



UNIVERSIDAD DE SEVILLA
Dpto. de Ciencias de la Computación
e Inteligencia Artificial

043

442

Resumen de la tesis doctoral

**Sistemas P: Un marco computacional para la
Biología de Sistemas**

Francisco José Romero Campero

UNIVERSIDAD DE SEVILLA



600495227

TESIS DOCTORAL

UNIVERSIDAD DE SEVILLA
SECRETARÍA GENERAL

Queda registrada esta Tesis Doctoral
en folio 165 número 338 del libro
de correspondencia.
Sevilla.

26-11-07
El Jefe del Negociado de Teles

[Handwritten signature]

Índice general

I	Introducción	5
1.	Introducción	7
1.1.	Organización de la tesis	8
2.	Modelado Celular	13
2.1.	Realidad y modelos.	13
2.2.	El proceso de modelado	14
2.3.	Biología de sistemas	14
2.4.	Aproximaciones al modelado de sistemas celulares	15
2.5.	Modelado de sistemas celulares basado en ecuaciones diferenciales	15
2.6.	Modelado estocástico de sistemas celulares	16
2.7.	Modelado computacional de sistemas celulares	17
2.7.1.	Redes de Petri	17
2.7.2.	π -calculus	18
II	Un Marco de Modelado	19
3.	Sistemas P: Un Marco para el Modelado de Sistemas Celulares	21
3.1.	Sistemas P: Un Modelo de Computación Bio-inspirado	21
3.2.	Especificaciones y Modelos de Sistemas	22
3.3.	Teoría Cinética de Gillespie para la Evolución de los Modelos de Sistemas P	27
3.3.1.	Algoritmo Multicompartimental de Gillespie	27
3.3.2.	Algoritmo de Gillespie multicompartimental	27
3.4.	Model Checking Probabilístico sobre Modelos de Sistemas P usando PRISM	29
3.4.1.	Model Checking Probabilístico con PRISM	32
3.4.2.	Transformando Especificaciones de Sistemas P en PRISM	32
3.4.3.	Análisis de Modelos de Sistemas P usando PRISM	33
4.	Especificaciones de Sistemas P de Sistemas Celulares	35
4.1.	Especificaciones de Compartimentos	36
4.2.	Especificaciones en Sistemas P de las Interacciones Proteína-Proteína	37
4.3.	Especificaciones en Sistemas P de Redes de Transcripción	41
4.3.1.	Especificación de las Redes de Transcripción usando Objetos	42
4.3.2.	Especificación de las Redes de Transcripción usando Cadenas	44
4.4.	Modelos de Sistemas P y SBLM (Systems Biology Markup Language)	48

4.4.1.	Traduciendo Formatos SBML en Modelos de Sistemas P	49
5.	Modularización en P sistemas	53
5.1.	Módulos en biología celular y P sistemas	53
5.2.	Una librería de módulos de P sistemas básicos	53
5.2.1.	Módulos de interacciones proteínicas	53
5.2.2.	Módulos de regulación de genes	54
5.3.	Modelado de módulos celulares mediante la combinación de módulos de P sistemas	55
5.3.1.	Autoregulation in Transcription Networks	55
III	Casos de estudio	57
6.	Modelado de la Regulación de Genes en Procaiotas	59
6.1.	Control de Expresión de Genes en el Lac Operon	59
6.2.	Modelado de la Regulación del Lac Operon usando Sistemas P	60
6.2.1.	Una Especificación de Sistema P del Sistema Lac Operon	61
6.2.2.	Modelos de Sistemas P del Sistema de Regulación del Lac Operon	66
6.3.	Análisis de la Expresión de Genes en Lac Operon	67
7.	Modelado de la Traducción de Señales	69
7.1.	Traducción de Señales	69
7.2.	Algoritmo Determinista con Tiempo de Espera	69
7.3.	Modelado de una Cascada de Señales del EGFR	71
7.3.1.	Una Especificación de Sistema P para una Cascada de Señales del EGFR	71
7.3.2.	Modelos de Sistemas P para el análisis de una Cascada de Señales EGFR	76
7.4.	Modelado de la Apoptosis Inducida por FAS	77
7.4.1.	Una Especificación de Sistema P de la Apoptosis Inducida por FAS	79
7.4.2.	Análisis de la Apoptosis Inducida por FAS usando Modelos de Sistemas P	82
8.	Un Modelo del Quorum Sensing	85
8.1.	Quorum Sensing en Vibrio Fischeri	85
8.2.	Una Especificación de Sistema P del Quorum Sensing en Vibrio Fischeri	86
8.3.	Modelos de Sistemas P del Quorum Sensing en Vibrio Fischeri	89
IV	Conclusiones	93
9.	Conclusiones	95

Parte I
Introducción

Capítulo 1

Introducción

El campo *Biología de Sistemas*¹ tiene como objetivo profundizar en el conocimiento de la dinámica de los sistemas celulares a través de las interacciones entre los diferentes sistemas. A diferencia de la aproximación clásica en la que el estudio de la dinámica de los procesos celulares se realizaba a través de la identificación y caracterización de cada una de sus componentes moleculares. Esta aproximación no captura el hecho de que el funcionamiento de los sistemas celulares emerge de las interacciones entre sus diferentes componentes. Debido a la complejidad de los procesos celulares y a la gran cantidad de datos producido por los experimentos en laboratorios, se hace imprescindible el modelado matemático-computacional, la simulación y el análisis que son técnicas esenciales en este campo.

La Biología de Sistemas constituye un campo interdisciplinar que entremezcla la Biología clásica, las Matemáticas y la Informática. Podemos pensar que esta nueva disciplina producirá una nueva generación de científicos capaces de entender y aplicar conceptos, técnicas y de servirse de fuentes de inspiración provenientes de cada una de esas áreas.

La Biología de Sistemas es un campo relativamente próximo a otra joven disciplina, la Biología Sintética [15, 98]. El fin último del modelado de sistemas celulares es obtener un conocimiento necesario para poder ser capaces de controlar su funcionamiento, lo que nos puede permitir construir nuestro propio sistema celular que satisfaga unas propiedades fijadas a priori (que es el principal objetivo de la Biología Sintética). Esto abre un amplio espectro a posibles aplicaciones en farmacología y tratamientos novedosos en importantes enfermedades, como el cáncer.

Las ecuaciones diferenciales ordinarias (ODEs) constituyen el primer marco para el modelado en Biología de Sistemas, quizás debido a que se ha utilizado desde hace décadas para modelar procesos dinámicos. Sin embargo, la aproximación macroscópica, continua y determinista inherente a las ODEs es, cuando menos, cuestionable en sistemas celulares con un bajo número de moléculas, en reacciones lentas o en estructuras no homogéneas [8, 41]. Para estos casos se han propuesto aproximaciones mesoscópicas, estocásticas y discretas. Algunos de los marcos computacionales alternativos usados para modelar sistemas celulares han sido las redes de Petri [46], el álgebra de procesos (π -cálculo [101], bioambientes [100], brane cálculo [26], κ -cálculo [Danos2004, etc.]), sistemas basados en agentes [54], etc.

No obstante, en la mayoría de estas aproximaciones no se tiene en cuenta el papel crucial que suelen jugar las membranas biológicas en la compartimentalización y funcionamiento de

¹término acuñado por Ideker [59])

las células de los organismos vivos.

En esta memoria se presentan los sistemas P como un marco de modelado computacional que integra los aspectos estructurales y dinámicos de los sistemas celulares en una forma comprensiva y relevante, proporcionando la formalización necesaria para poder realizar análisis matemáticos y computacionales. Todas las aproximaciones citadas anteriormente fueron formalizadas a partir de diferentes fuentes de inspiración, y antes de ser aplicadas al modelado de sistemas celulares. Por el contrario, los sistemas P constituyen un modelo no convencional de computación inspirado en la estructura y el funcionamiento de las células [91]. Por tanto, los conceptos usados en los sistemas P están más próximos a los usados en biología celular y molecular que las abstracciones de los otros formalismos o aproximaciones.

Los tres ingredientes esenciales de un sistemas P son los siguientes: una estructura de membranas (a modo de célula), conteniendo un determinado número de membranas distribuidas jerárquicamente y delimitando regiones o compartimentos; multiconjuntos de objetos y cadenas colocadas en el interior de los compartimentos delimitados por las membranas (abstracciones de las sustancias químicas); y reglas de reescritura (abstracciones de las reacciones químicas) asociadas a cada compartimento que describen la evolución de los objetos y cadenas colocados en los compartimentos

Más que una alternativa a los marcos clásicos de modelado en los sistemas celulares, los sistemas P constituyen una aproximación complementaria para ser usada cuando las aproximaciones clásicas puedan plantear problemas. Los sistemas P son capaces de representar el carácter discreto de la cantidad de componentes de un sistema celular mediante el uso de reglas de reescritura sobre multiconjuntos de objetos y cadenas que representan moléculas, y cadenas que describen la organización de genes sobre el genoma. La estocasticidad inherente, el ruido externo y la incertidumbre de los sistemas celulares se captura en los sistemas P mediante el uso de estrategias estocásticas basadas en el algoritmo de Gillespie [40, 41, 42, 43, 44].

El principal hecho diferencial de los sistemas P es el uso de la llamada *Estructura de membranas* que representa la compartimentalización en la organización estructural de las células vivas. En todos los casos de estudio que se presentan en esta memoria, las membranas en los sistemas P de especificación describirán las regiones relevantes que corresponden a los sistemas celulares objetos de estudio.

1.1. Organización de la tesis

La memoria está organizada en cuatro partes que cubren diferentes aspectos del trabajo desarrollado por el autor usando sistemas P como marco computacional para la especificación y simulación de sistemas celulares.

■ Part I: Introducción

La primera parte del texto constituye la introducción de este trabajo y consta de dos capítulos. El primero de ellos es una especie de resumen extendido de la tesis, con una descripción de la organización del texto y un breve sumario de cada capítulo.

El segundo capítulo presenta una breve panorámica acerca de la problemática en el proceso de modelado de sistemas celulares en el marco de la Biología de Sistemas. Además, se

describen sucintamente algunas aproximaciones a dicho modelado: ecuaciones diferenciales (sección 2.5), modelos estocásticos (sección 2.6) y modelos computacionales (sección 2.7).

■ Part II: Marco de Modelado

La segunda parte constituye el núcleo central de la memoria. En ella se presenta un marco de modelado computacional basado en sistemas P, estructurado en tres capítulos.

En el primero de ellos, los sistemas P se presentan como un modelo de sistemas celulares que satisfacen las condiciones de relevancia, comprensibilidad, extensibilidad y tratabilidad matemático-computacional. Los principales conceptos son los de sistemas P de especificación y familias de modelos de sistemas P. En la sección 3.3 se introduce una extensión del conocido como algoritmo de Gillespie, a la estructura compartimentalizada de los sistemas P (el *algoritmo Multicompartimental de Gillespie*). Finalmente, en la sección 3.4 se discute una metodología sobre modelos de sistemas P usando PRISM, que permite realizar model checking probabilístico.

El capítulo 4 presenta los principios de la especificación en sistemas P. En la sección 4.2 se presenta una enumeración de las interacciones proteínas-proteínas que pueden ser especificadas usando reglas de reescritura en sistemas P. La descripción de redes de transcripción se discute en la siguiente sección ofreciendo dos posibilidades: los procesos que involucran proteínas, genes y mRNA pueden ser representados usando bien reglas de reescritura sobre multiconjuntos de objetos (sección 4.3.1) o bien usando reglas de reescritura sobre multiconjuntos de objetos y cadenas (sección 4.3.2). Finalmente, en la sección 4.4 se introduce un método para traducir a sistemas P, modelos de sistemas celulares especificados usando SBML (Systems Biology Markup Language), un Lenguaje derivado de XML para la representación de redes de reacciones bioquímicas.

En el capítulo 5 se analiza el concepto de módulos de un sistema P. En la sección 5.2 se describen una serie de módulos básicos que representan los subsistemas elementales más importantes de un sistema celular. Finalmente, en la sección 5.3 se usa la autorregulación en las redes de transcripción para ilustrar cómo los módulos de sistemas P pueden combinarse entre sí para producir módulos más complejos.

■ Parte III: Casos de Estudio

En la tercera parte de esta tesis se estudian tres casos de estudio para ilustrar los principios y técnicas de modelado discutidos en la parte II.

En el capítulo 6 se analizan los principios básicos para el modelado de la regulación de genes en células procariontas mediante el uso de sistemas P, estudiando el comportamiento del sistema del Lac Operon bajo diferentes condiciones del entorno.

En el capítulo 7 comienza con una breve discusión acerca de los principios generales que rigen la traducción de señales y se introduce una versión determinista del algoritmo multicompartimental de Gillespie, denominado *algoritmo determinista de tiempo de espera*. Se estudian dos sistemas de traducción de señales: una cascada de señales del factor de crecimiento epidérmico (EGFR), y una cascada de señales apoptóticas mediatizadas por la proteína FAS.

En el capítulo 8 se presenta el último caso de estudio relativo a los mecanismos de comunicación entre bacterias, conocido como *quorum sensing*. Para ello, se describe brevemente este fenómeno en la bacteria marina *Vibrio fischeri* y se analizan algunos principios para el desarrollo de especificaciones y modelos para colonias de bacterias en el marco de los sistemas P.

■ Parte IV: Conclusiones

La última parte de la memoria está dedicada a la presentación de algunas conclusiones, discutiéndose algunos logros y limitaciones del uso de los sistemas P como marco de modelado computacional en Biología de Sistemas. Se presentan algunos trabajos y líneas de investigación para el futuro.

Muchas de las contribuciones técnicas de esta memoria han sido publicadas en revistas internacionales. Las partes no publicadas han sido sometidas recientemente a revisión.

La mayor parte de las simulaciones realizadas en esta memoria usan una extensión del algoritmo clásico de Gillespie: el algoritmo Multicompartimental de Gillespie, cuyo uso en el marco de los sistemas P fue presentado por primera vez en el un artículo de la revista *Transactions on Computational Systems Biology*.

- Pérez-Jiménez, M.J., Romero-Campero, F.J. (2006) P Systems, a new computationl modelling tool for systems biology, *Transactions on Computational Systems Biology VI, LNBI, 4220*, 176 – 197.

Relativo a la adaptación de los sistemas P al modelado de sistemas celulares, se presentaron dos comunicaciones en sendos Congresos Internacionales que, posteriormente, fueron seleccionados para su publicación en un número especial del *Lecture Notes in Computer Science* dedicado al mismo. En ellos se describen el uso del paralelismo acotado en lugar de la estrategia clásica de los sistemas P basadas en la estrategia del paralelismo maximal.

- Bernardini, F., Romero-Campero, F.J., Gheorghe, M., Pérez-Jiménez M.J. (2006) A Modeling Approach Based on P Systems with Bounded Parallelism, *Lecture Notes in Computer Science, 4361*, 49 – 65.
- Bernardini, F., Gheorghe, M., Krasnogor, N., Muniyandi, R.C., Pérez-Jiménez, M.J., Romero-Campero, F.J. (2005) On P Systems as a Modelling Tool for Biological Systems, *Lecture Notes in Computer Science, 3850*, 114 – 133.

Una investigación dirigida hacia el uso de técnicas de model checking probabilístico y simbólico en sistemas P se presente en un Congreso Internacional.

- Romero-Campero, F.J., Gheorghe, M., Bianco, L., Pescini, D., Pérez-Jiménez, M.J., Ceterchi, R. (2006) Towards Probabilistic Model Checking on P Systems Using PRISM, *Lecture Notes in Computer Science, 4361*, 477 – 495.

Los cuatro artículos indicados cubren, básicamente, el contenido del capítulo 3, en donde se presenta el marco de modelado basado em sistemas P.

Una comparación de este modelado computacional con otros marcos en Biología de Sistemas (más específicamente, con el π -calculus) es analizada en el siguiente artículo publicado en la revista *Progress in Natural Science*.

- Romero-Campero, F.J., Gheorghe, M., Ciobanu, G., Auld, J., Pérez-Jiménez, M.J. (2007) Cellular modelling using P systems and process algebra, *Progress in Natural Science*, **17** (4), 375 – 383.

Los principios de especificación de las interacciones proteínas-proteínas y las redes de transcripción fueron presentadas por primera vez en el artículo siguiente de la revista *Biosystems*.

- Romero-Campero, F.J., Pérez-Jiménez, M.J. Modelling Gene Expression Control Using P Systems: The Lac Operon, A Case Study, *Biosystems* in press.

Las discusiones presentadas en el capítulo 6 son continuación de las ideas presentadas en el artículo anterior. En el capítulo 6 se describe la regulación de genes en el Lac Operon. Este sistema fue usado como un caso de estudio para ilustrar los principios de especificación en sistemas P.

Los dos casos de estudio descritos en el capítulo 7 y que ilustran cómo los sistemas de traducción de señales pueden ser especificados e ilustrados usando sistemas P, son presentados en los dos siguientes artículos, uno de ellos publicado en la revista *Progress in Natural Science* y el otro como comunicación en un Congreso Internacional. paper.

- Chereku, S., Paun, A., Romero-Campero, F.J., Pérez-Jiménez, M.J., Ibarra, O.H. (2007) Simulating Fas-induced apoptosis by using P systems, *Progress in Natural Science*, **17** (4), 424 – 431.
- Paun, A., Pérez-Jiménez, M.J., Romero-Campero, F.J. (2006) Modeling Signal Transduction Using P Systems, *Lecture Notes in Computer Science*, **4361**, 100 – 122.

Finalmente, el siguiente artículo, que será publicado en la revista *Artificial Life*, estudia el uso de los sistemas P para el modelado de las interacciones en colonias de bacterias; en particular, presenta el primer modelado computacional del fenómeno Quorum Sensing, con un caso de estudio dedicado a la bacteria marina *Vibrio Fisheri*

- Romero-Campero, F.J., Pérez-Jiménez, M.J. (2008) A Model of the Quorum Sensing System in *Vibrio Fischeri* Using P Systems, *Artificial Life*, **14**, in press.

Capítulo 2

Modelado Celular

2.1. Realidad y modelos.

El uso de modelos es intrínseco a cualquier actividad científica. Los científicos usan a menudo abstracciones de la realidad como diagramas, gráficos, etc. con la intención de describir y comprender la realidad que examinan.

Aunque los biólogos están familiarizados con el modelado, el modelado computacional matemático y cuantitativo ha permanecido fuera de la corriente principal de investigación debido a la falta de técnicas apropiadas tanto en la parte experimental como en la parte matemática/computacional. Sin embargo, los progresos obtenidos durante finales del siglo pasado tanto en biología celular como en ciencias de la computación han llevado a estas dos disciplinas a una situación en la que ambas se pueden beneficiar la una de la otra. Los principios de este siglo fueron testigos del nacimiento de una nueva disciplina que integra los campos clásicos de biología, matemática y ciencias de la computación, la biología de sistemas.

El desarrollo de modelos es central en la biología de sistemas. Un modelo es una abstracción del mundo real dentro de un marco computacional/matemático que destaca algunos aspectos claves del fenómeno modelado mientras que se ignoran ciertas características que se asumen ser irrelevantes.

Está ampliamente aceptado que un buen modelo celular debe satisfacer al menos las cuatro siguientes propiedades: relevancia, comprensibilidad, extensibilidad y tratabilidad matemática y computacional.

- Un modelo debe ser relevante capturando las características esenciales del sistema modelado presentando un marco unificador tanto de la estructura del sistema y como de su dinámica.
- Un modelo debe facilitar una mejor comprensibilidad del sistema celular real en lugar de producir un formalismo complicado y difícil de descifrar.
- Un modelo debe ser fácilmente extensible y modificable para incluir nuevo conocimiento o eliminar hipótesis falsas.
- Un modelo debe ser computable en el sentido de que debe ser posible su implementación en computadores tradicionales posibilitando su análisis computacional y matemático.

Existen principalmente tres usos de los modelos celulares:

- Los modelos celulares son usados como un marco formal para resumir y sintetizar grandes cantidades de conocimiento empírico disperso.
- La obtención de un modelo capaz de realizar predicciones sobre el comportamiento de un sistema celular es quizás el principal objetivo del modelado de procesos celulares.
- El conocimiento adquirido sobre un sistema celular a través de su modelado nos da la posibilidad de construir, restringir y/o manipular sistemas celulares reales de tal manera que produzca un comportamiento deseado.

2.2. El proceso de modelado

El proceso de modelado es un conjunto de reglas semiformales que nos guían en el desarrollo de un modelo, su formulación en un lenguaje formal, su implementación y su análisis.

El proceso de modelado consiste en las siguientes etapas:

- Identificación y delimitación del sistema.
- Definición y formulación formal
- Implementación en computadores
- Validación y calibración
- Análisis y testeo

2.3. Biología de sistemas

La vida es una de los fenómenos más interesantes del universo. La aproximación clásica al estudio de la vida, biología, se centró en el análisis de cada uno de los componentes celulares aislados. Sin embargo, se ha probado que esta dirección es errónea. Los procesos celulares surgen como un comportamiento emergente de las múltiples interacciones entre sus componentes individuales. La funcionalidad de los procesos celulares no es la suma de los comportamientos individuales de sus componentes.

Gracias a los grandes avances tanto en biología celular como en ciencias de la computación obtenidos durante finales del pasado siglo y principios de este la comunidad científica se encuentra posicionada para desarrollar un estudio sistemático de células, tejidos, órganos y organismos en general y, en particular, de los principales procesos celulares como comunicación celular, división, homeostasis, etc. Este nuevo campo se ha denominado biología de sistemas y se basa en la integración de biología, matemática y ciencias de la computación.

2.4. Aproximaciones al modelado de sistemas celulares

Existen múltiples formas de modelar sistemas celulares. Aquí nos es imposible presentar una clasificación exhaustiva de las diferentes aproximaciones al modelado celular. Por lo tanto, nos limitaremos a enumerar las principales aproximaciones al modelado celular.

De acuerdo al escalado del espacio, desde una visión general del sistema hasta la representación exhaustiva de cada componente, podemos distinguir entre modelos macroscópicos, mesoscópicos y microscópicos.

Según el tipo de análisis que se realice sobre los modelos se distingue entre modelos cualitativos y cuantitativos.

Basados en la representación de los distintos componentes de un modelo podemos diferenciar entre modelos discretos y continuos.

Los modelos se pueden clasificar en deterministas o estocásticos dependiendo de las dinámicas que sigan.

Dependiendo del punto de partida del desarrollo del modelo, bajo o alto nivel, y como se incorpora nueva información al modelo, podemos distinguir entre modelos top-down y bottom-up.

2.5. Modelado de sistemas celulares basado en ecuaciones diferenciales

Hoy en día las ecuaciones diferenciales ordinarias (EDO) son el marco de modelado más ampliamente usado en biología de sistemas. Sin embargo, el uso de este marco está basado en dos condiciones:

1. Se asume que las células son volúmenes homogéneos bien mezclados donde la concentración de las diferentes especies moleculares no cambia con respecto al espacio.
2. Se asume que la concentración de las diferentes especies moleculares cambia de forma continua y determinista en el tiempo.

Ambas de las previas condiciones no se cumplen en sistemas celulares con un número pequeño de moléculas, reacciones lentas y/o estructuras organizadas en diferentes compartimentos. Este tipo de sistemas celulares no es extraño. Sin embargo, el uso de EDO ha sido exitoso en el modelado de múltiples sistemas celulares.

En el marco de modelado de las EDO la concentración de cada especie molecular en un instante de tiempo t se representa mediante una variable, $X_i(t)$. Cada reacción química se representa utilizando una ley cinética en particular. Para cada especie molecular se escribe una ecuación diferencial que describe como su concentración química cambia a largo del tiempo. Las leyes cinéticas más utilizadas son la ley de acción de masas, la ley del decaimiento exponencial y dinámicas de Michaelis-Menten.

Para resolver un sistema de EDO se deben imponer además unas condiciones iniciales que representan las concentraciones iniciales de las diferentes especies moleculares. Un sistema de EDO junto con unas condiciones iniciales constituye un problema de valores iniciales bien

fundamentado. La existencia y unicidad de solución de este tipo de problema está garantizada bajo débiles condiciones de regularidad.

De acuerdo a lo mencionado arriba EDO no pueden utilizarse en sistemas celulares no homogéneos en los cuales la concentración de las especies moleculares varía en el espacio. En este caso se pueden utilizar Ecuaciones en Derivadas Parciales (EDP). En este trabajo sólo trataremos brevemente las EDP reacción-difusión. En este caso, la densidad de cada especie molecular en un instante de tiempo t y en un punto del espacio \mathbf{x} se representa utilizando una variable $X_i(t, \mathbf{x})$. También en este caso se asume que la concentración de cada especie molecular varía de forma continua y determinista en el tiempo de acuerdo con leyes cinéticas específicas. Además, en este caso se añade un término de difusión que representa el movimiento de las especies moleculares en el espacio.

2.6. Modelado estocástico de sistemas celulares

Existen sistemas celulares en los que el uso de ecuaciones diferenciales no es adecuado. En su lugar se utilizan procesos estocásticos.

En la representación de sistemas celulares utilizando procesos estocásticos también se parte de la asunción de que tenemos un volumen fijo homogéneo.

El estado del sistema en un instante de tiempo t se representa mediante un vector $\mathbf{X}(t) = (X_1(t), \dots, X_N(t))$ donde $X_i(t)$ representa el número de moléculas de la especie molecular i . A cada interacción molecular R_j se le asocia una constante c_j . Utilizando esta constante y el número de reactivos en el sistema en el instante t se calcula la propensidad de R_j en el instante t , $a_j(\mathbf{X}(t))$ de tal forma que $a_j(\mathbf{X}(t))dt$ sea la probabilidad de que tenga lugar una interacción molecular del tipo R_j en el intervalo de tiempo $[t, t + dt)$.

En una primera aproximación uno está interesado en calcular $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$, la probabilidad de que el sistema se encuentre en el estado $\mathbf{X}(t) = \mathbf{x}$, dado que el estado en el instante inicial era $\mathbf{X}(t_0) = \mathbf{x}_0$. Esto da lugar a la Ecuación Química Maestra (EQM):

$$\frac{dP(\mathbf{x}, t | \mathbf{x}_0, t_0)}{dt} = \sum_{j=1}^M [a_j(\mathbf{x} - \mathbf{v}_j)P(\mathbf{x} - \mathbf{v}_j, t | \mathbf{x}_0, t_0) - a_j(\mathbf{x})P(\mathbf{x}, t | \mathbf{x}_0, t_0)] \quad (2.1)$$

La EQM consiste en un sistema de ecuaciones diferenciales con tantas ecuaciones como posibles estados. Esto la hace intratable y se requiere un cambio en el punto de vista.

Se ha diseñado un algoritmo que permite explorar de forma exacta el espacio de estados asociado a la EQM denominado algoritmo de Gillespie. En este caso lo que se pretende calcular es $p(\tau, j | \mathbf{x}, t)$ la probabilidad de que cuando el sistema se encuentra en el estado \mathbf{x} en el instante t la próxima interacción molecular tendrá lugar en el intervalo infinitesimal $[t + \tau, t + \tau + d\tau)$ y será del tipo R_j . Esto da lugar al siguiente algoritmo:

1. Inicializar el tiempo de simulación $t = t_0$ y el estado del sistema $\mathbf{X}(t_0) = \mathbf{x}_0$.
2. Computar todas las propensidades asociadas con las interacciones moleculares $a_j(\mathbf{x})$ y su suma $a_0(\mathbf{x})$.
3. Generar dos números aleatorios según la distribución uniforme en el intervalo unidad, r_1 y r_2 , y calcular τ , el tiempo de espera, y j la interacción molecular a aplicar como sigue:

$$\tau = \frac{1}{a_0(\mathbf{x})} \ln \left(\frac{1}{r_1} \right) \quad (2.2)$$

$$j = \text{el menor entero satisfaciendo } \sum_{j'=1}^j a_{j'}(\mathbf{x}) > r_2 a_0(\mathbf{x}) \quad (2.3)$$

4. Aplicar una vez la interacción molecular R_j y actualizar el tiempo de simulación $t \leftarrow t + \tau$.
5. Almacenar (t, \mathbf{x}) y volver al paso 2 o parar la simulación.

2.7. Modelado computacional de sistemas celulares

La complejidad de los sistemas celulares hacen necesario el uso de computadores para develar sus mecanismos yacientes. Sin embargo, la mayoría de los modelos computacionales se implementan ad hoc careciendo de un marco definido, consistente y formal. Recientemente, se han propuesto varios formalismos bien fundamentados para el diseño de modelos computacionales. En este trabajo tan sólo presentaremos brevemente los dos marcos computacionales más extendidos, redes de Petri y álgebra de procesos, más específicamente π -calculus.

2.7.1. Redes de Petri

Las redes de Petri son una herramienta matemática/computacional para el modelado y análisis de sistemas discretos con comportamiento concurrente. Presentan un marco para representar la estructura de las interacciones de sistemas discretos, simular su comportamiento y probar propiedades del sistema.

Una red de Petri consiste en un gráfico formado por dos tipos de nodos llamados lugares L y transiciones T respectivamente. Los lugares se conectan a las transiciones y las transiciones a los lugares mediante arcos dirigidos con un peso asociado. Dada una transición se distingue entre lugares de entrada, aquellos con arcos que llegan hasta la transición y lugares de salida aquellos con arcos que salen de la transición y llegan a ellos. Normalmente una red de Petri se representa gráficamente dibujando los lugares como círculos y las transiciones como rectángulos.

Una serie de fichas se colocan en los distintos lugares de una red de Petri. El estado de una red de Petri se determina por el número de fichas colocadas en sus lugares. Una transición está habilitada cuando el número de fichas en sus lugares de entradas es mayor o igual al peso asociado a los arcos que los unen con la transición. Cuando una transición se dispara consume de los lugares de entrada un número de fichas igual al peso de los arcos correspondientes y produce en los lugares de salida un número de fichas igual al peso de los arcos correspondientes.

Un sistema molecular puede representarse como una red de Petri como sigue. Cada especie molecular se representa mediante un lugar, el número de moléculas se representa mediante el número de fichas asociadas con los lugares y cada interacción molecular mediante una transición cuyos lugares de entradas son los reactivos y cuyos lugares de salida son los productos de la interacción molecular.

Con este marco sólo se puede realizar un estudio cualitativo de los sistemas moleculares. Con la finalidad de obtener un marco cuantitativo se desarrollaron las redes de Petri estocásticas

en las que el tiempo de espera para el disparo de las transiciones se determina según una distribución exponencial cuyo parámetro se calcula de acuerdo con una constante asociada a cada transición y el número de fichas en los lugares de entrada.

2.7.2. π -calculus

El π -calculus fue introducido por Milner, Parrow y Walker como un lenguaje formal para describir procesos móviles que se comunican a través de canales de comunicación. En π -calculus el nombre de nuevos canales de comunicación se pueden transmitir a través de canales de comunicación lo que aporta una movilidad de canales de comunicación.

El π -calculus tiene una semántica simple y una teoría algebraica tratable. A partir de procesos simples y acciones atómicas se pueden construir procesos más complejos. Los procesos se definen a través de la composición paralela $P|Q$, elección no determinista $P + Q$, replicación $!P$ y la creación de un canal de comunicación local μxP .

Un sistema molecular se representa en el π -calculus como un sistema de procesos que interactúan a través de canales de comunicación específicos. Cada especie molecular o dominio de una especie molecular se representa mediante procesos. El número de moléculas de una determinada especie molecular se representa mediante el número de procesos. Finalmente, cada interacción molecular se representa mediante un canal de comunicación.

De la misma forma que para la redes de Petri este marco es sólo cualitativo para introducir información cuantitativa se asocia una constante a cada canal de comunicación. Entonces, el tiempo de espera para realizar una comunicación se determina mediante una distribución exponencial cuyo parámetro se calcula utilizando la anterior constante y el número de procesos que intentan comunicarse a través de dicho canal de comunicación.

Parte II

Un Marco de Modelado

Capítulo 3

Sistemas P: Un Marco para el Modelado de Sistemas Celulares

En el capítulo 1 se han discutido diferentes aproximaciones para el modelado de sistemas celulares. A pesar de que cada una de esas aproximaciones captura alguna información relevante relativa a esos sistemas y a sus componentes, ninguna de ellas captura íntegramente la dinámica y los detalles estructurales de las distintas componentes. Además, como ya hemos indicado, esto debería ser realizado manteniendo los requisitos de relevancia, comprensibilidad, extensibilidad y tratabilidad matemático-computacional. En este capítulo se presentan los sistemas P como una nueva aproximación al modelado de sistemas celulares verificando los requisitos antes mencionados.

3.1. Sistemas P: Un Modelo de Computación Bio-inspirado

La *Computación Celular con Membranas* (*Membrane Computing*) es una rama emergente de la Computación Natural introducida por Gheorghe G. Păun en [89], que está inspirada en la estructura y funcionamiento de las células de los organismos vivos. Desde entonces ha recibido una atención especial por parte de la comunidad científica y ha sido designada por el *Institute of Scientific Information*, ISI, de USA como *Fast Emerging Research Front in Computer Science* (octubre de 2003) y el artículo fundacional *Computing with Membranes* ([89]) fue nominado por el ISI como *Fast Breaking Paper* en *Computer Science* (febrero de 2003).

Este nuevo modelo parte del supuesto de que los procesos que tienen lugar en la estructura compartimentalizada de una célula viva pueden ser interpretados como procedimientos de cálculo; es decir, como computaciones. La Computación Celular con Membranas parte de la observación de que las membranas biológicas juegan un papel relevante en el funcionamiento de las células. La membrana celular separa y, por tanto, protege el espacio interno celular del entorno externo. Las membranas interiores definen la estructura de la célula identificando un número de compartimentos internos que permiten proteger al núcleo que contiene toda la información genética. Las membranas están involucradas en muchas reacciones químicas que tienen lugar dentro de los compartimentos y, además, actúan como canales selectivos de comunicación entre las células y su entorno. Todos estos hechos son formalizados mediante la introducción de

los *sistemas P*, también denominados *sistemas de membranas*. Los principales ingredientes de un sistema *P* son los siguientes:

- Una *estructura de membranas* que contiene una estructura jerarquizada de membranas que delimitan una serie de regiones o compartimentos. Formalmente, una estructura de membranas es un árbol enraizado cuya raíz se denomina *piel* y cuyas hojas se denominan *membranas elementales*.
- Unos multiconjuntos de objetos o cadenas sobre un alfabeto colocados en los compartimentos delimitados por las membranas, que son abstracciones de las sustancias químicas contenidas en las distintas regiones.
- Unas reglas de reescritura asociadas a cada compartimento que describen las reglas de evolución de los objetos y las cadenas situadas en su interior. Estas reglas son abstracciones de las reacciones químicas que se producen en las regiones.

En la aproximación original, una *computación* en un sistema de membranas se obtiene por la aplicación reiterada de las reglas de evolución de manera maximal y no determinista; es decir, en cada paso, en cada compartimento, todos los objetos y cadenas que puedan evolucionar por medio de cualquier regla, deben evolucionar en paralelo al mismo tiempo. Desde su introducción, han aparecido multitud de variantes como los sistemas *P* con reglas del tipo symport/antiport rules [90], con promotores y/o inhibidores [60], con membranas activas, sistemas *P* de tejidos, sistemas *P* con spiking Neuronales [61], etc.

Es importante resaltar el hecho de que, de acuerdo con la motivación original, los sistemas *P* no trataban de proporcionar un modelo comprensivo y completo de las células vivas sino más bien explorar la naturaleza computacional de una serie de hechos relativos a las membranas biológicas. Por ello, se ha estudiado exhaustivamente la potencia computacional de distintas variantes de sistemas *P*, así como su capacidad para resolver problemas computacionalmente duros en tiempo polinomial (por supuesto, intercambiando recursos de tiempo por espacio). Recientemente, los sistemas *P* están siendo utilizados para modelar fenómenos biológicos dentro del marco de la Biología de Sistemas, presentando modelos de sistemas oscilatorios [35], de traducción de señales [92], de control de la regulación de genes [107], del quorum sensing [108], y de las metapoblaciones [93].

3.2. Especificaciones y Modelos de Sistemas

En esta sección se introducen los conceptos de *especificaciones de sistemas P* que constituyen la principal estructura usada para el análisis de sistemas celulares. A cada especificación particular se le asocia unos conjuntos de *parametros* que permiten, por instanciación, producir los *modelos de sistemas P*.

Definición 3.2.1 (Especificación de Sistema P) *Una especificación de sistema P es una tupla*

$$\Pi = (O, L, \mu, M_1, M_2, \dots, M_n, R_{l_1}, \dots, R_{l_m})$$

en donde:

- O es un alfabeto finito de símbolos que representa objetos;
- $L = \{l_1, \dots, l_m\}$ es un alfabeto finito de símbolos que representa etiquetas para los distintos compartimentos;
- μ es una estructura de membranas que contiene $n \geq 1$ membranas identificadas inyectivamente con valores en $\{1, \dots, n\}$ y etiquetadas por elementos de L ;
- $M_i = (l_i, w_i, s_i)$, para cada $1 \leq i \leq n$, es la configuración inicial de la membrana i , con $l_i \in L$, la etiqueta de esta membrana, $w_i \in O^*$ un multiconjunto finito de objetos, y s_i un conjunto finito de cadenas sobre O ;
- $R_{l_t} = \{r_1^{l_t}, \dots, r_{k_{l_t}}^{l_t}\}$, para cada $1 \leq t \leq m$, es un conjunto finito de reglas de reescritura asociadas con cada compartimento del tipo representado por la etiqueta $l_t \in L$, de una de las siguientes formas:
 - Reglas de reescritura de multiconjuntos:

$$r_j^{l_t} : \text{obj}_1 [\text{obj}_2]_l \xrightarrow{c_j^{l_t}} \text{obj}'_1 [\text{obj}'_2]_l \quad (3.1)$$

en donde $\text{obj}_1, \text{obj}_2, \text{obj}'_1, \text{obj}'_2 \in O^*$ son multiconjuntos finitos de objetos y l una etiqueta de L . Un multiconjunto de objetos, obj se representa por $\text{obj} = o_1 + o_2 + \dots + o_m$, con $o_1, \dots, o_m \in O$. La cadena vacía, que representa el multiconjunto vacío, se denotará por λ . Por simplicidad, escribiremos o^n en lugar de $\overbrace{o + \dots + o}^n$.

- Reglas de reescritura de cadenas:

$$r_j^{l_t} : [\text{obj}_1 + \text{str}_1; \dots; \text{obj}_p + \text{str}_p]_l \xrightarrow{c_j^{l_t}} [\text{obj}'_1 + \text{str}'_{1,1} + \dots + \text{str}'_{1,i_1}; \dots; \text{obj}'_p + \text{str}'_{p,1} + \dots + \text{str}'_{p,i_p}]_l \quad (3.2)$$

Una cadena str se representa como sigue: $\text{str} = \langle s_1.s_2 \dots s_i \rangle$ where $s_1, \dots, s_i \in O$. En este caso, cada multiconjunto de objetos obj_j y cadena str_j , $1 \leq j \leq p$, son reemplazados por un multiconjunto de objetos obj'_j y cadenas $\text{str}'_{j,1} \dots \text{str}'_{j,i_j}$.

Obsérvese que a cada regla se le asocia una constante $c_j^{l_t}$, denominada constante estocástica y que será usada para calcular la propensidad de la regla, de acuerdo con el contexto actual del compartimento en donde la regla está localizada.

Definición 3.2.2 (Parámetros) Dado una especificación de sistema $P \Pi = (O, L, \mu, M_1, \dots, M_n, R_{l_1}, \dots, R_{l_m})$, el conjunto de parámetros de Π , $\mathcal{P}(\Pi) = (\mathcal{M}_0(\Pi), \mathcal{C}(\Pi))$, consiste en lo siguiente:

1. Los multiconjuntos iniciales $\mathcal{M}_0(\Pi) = (M_1, \dots, M_n)$ asociados con cada compartimento.
2. Las constantes $\mathcal{C}(\Pi) = (r_j^{l_t})_{\substack{1 \leq j \leq k_{l_t} \\ 1 \leq t \leq m}}$, asociadas a cada regla de reescritura en R_{l_1}, \dots, R_{l_m} .

Definición 3.2.3 (Modelos de sistemas P) Sea Π una especificación de sistema P con parámetros $\mathcal{P}(\Pi) = (\mathcal{M}_0(\Pi), \mathcal{C}(\Pi))$, \mathbb{M}_0 , y \mathbb{C} una familia de posibles valores para los multiconjuntos iniciales $\mathcal{M}_0(\Pi)$ y para las constantes $\mathcal{C}(\Pi)$. Una familia de modelos de sistemas P , $\mathcal{F}(\Pi, \mathbb{M}_0, \mathbb{C})$, se obtiene de Π , \mathbb{M}_0 y \mathbb{C} mediante la instanciación de los parámetros $\mathcal{P}(\Pi)$ usando valores de \mathbb{M}_0 y \mathbb{C} .

A continuación, se ilustran las definiciones previas usando el sistema celular descrito en la figura 3.1. Nuestro ejemplo consiste en una señal S que es transportada al citoplasma mediante un receptor R localizado en la superficie celular. Una vez que ha llegado al citoplasma, la señal S interactúa con una proteína Tf formando el complejo $S-Tf$ que permite la activación de los genes $gene-Tf$ y $gene-R$, que codifican la proteína Tf y el receptor R , respectivamente.

La siguiente especificación de sistema P describe el sistema:

$$\Pi_0 = (O, L, \mu, M_1, M_2, M_3, R_{env}, R_{surf}, R_{cyto})$$

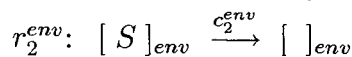
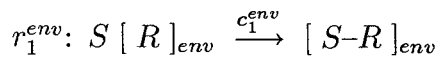
en donde:

- En el alfabeto O se representan las entidades moleculares presentes en el sistema.

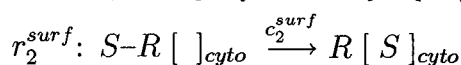
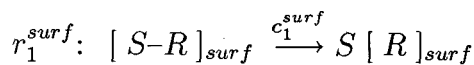
$$O = \{S, R, S-R, Tf, S-Tf, gene-Tf, S-Tf-gene-Tf, gene-R, S-Tf-gene-R\}$$

- Las etiquetas $L = \{env, surf, cyto\}$ especifican las regiones del sistema: el entorno, la superficie celular y el citoplasma.
- La estructura de membranas μ consta de tres membranas que definen tres compartimentos o regiones involucradas en el sistema: el entorno, la superficie celular y el citoplasma que están identificados con los números 1, 2 y 3, y etiquetadas por env , $surf$ y $cyto$, respectivamente.
- Los multiconjuntos iniciales M_1 , M_2 y M_3 representan las condiciones iniciales de los distintos compartimentos.
- El conjunto de reglas de reescritura R_{env} , R_{surf} y R_{cyto} describen las interacciones moleculares que tienen lugar en cada uno de los compartimentos.

- El conjunto de reglas $R_{env} = \{r_1^{env}, r_2^{env}\}$ representa el enlace de la señal S al receptor R , r_1^{env} ; y la degradación de la señal S (r_2^{env}).



- El conjunto de reglas $R_{surf} = \{r_1^{surf}, r_2^{surf}, r_3^{surf}\}$ especifica la rotura del enlace de la señal S y el receptor R (r_1^{surf}), el desplazamiento de la señal S al citoplasma (r_2^{surf}), y la internalización o difusión del receptor R en el citoplasma (r_3^{surf}).



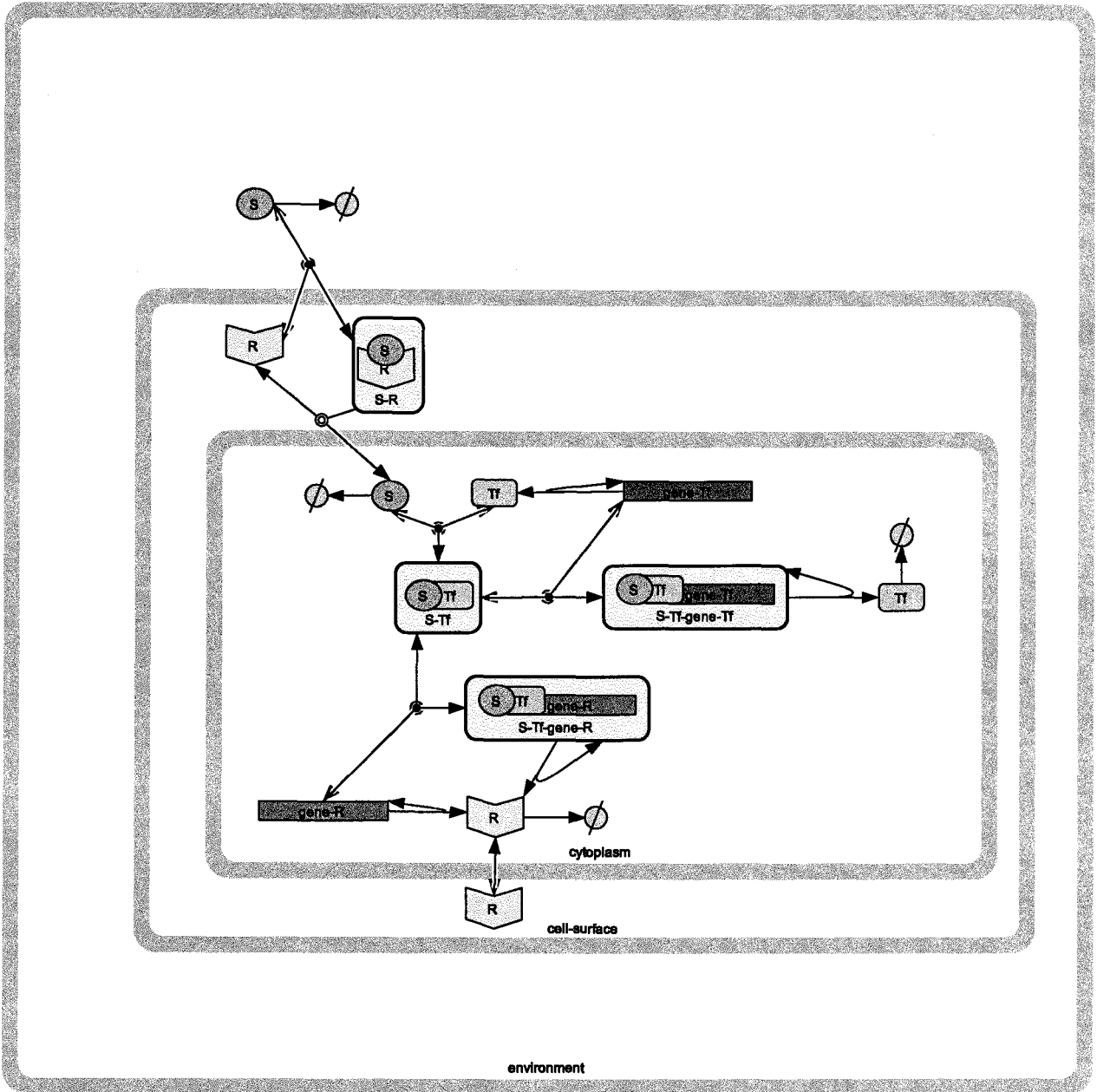


Figura 3.1: Un ejemplo de sistema celular.

$$r_3^{surf}: R []_{cyto} \xrightarrow{c_2^{cyto}} [R]_{cyto}$$

- El conjunto de reglas R_{cyto} especifica las interacciones moleculares que tienen lugar en el citoplasma:

$$r_1^{cyto}: [gene-Tf]_{cyto} \xrightarrow{c_1^{cyto}} [gene-Tf + Tf]_{cyto}$$

$$r_2^{cyto}: [gene-R]_{cyto} \xrightarrow{c_2^{cyto}} [gene-R + R]_{cyto}$$

Estas dos reglas representan la producción a nivel básico de la proteína Tf (r_1^{cyto}) y del receptor R (r_2^{cyto}).

$$r_3^{cyto}: [S + Tf]_{cyto} \xrightarrow{r_3^{cyto}} [S-Tf]_{cyto}$$

$$r_4^{cyto}: [S-Tf]_l \xrightarrow{c_4^{cyto}} [S + Tf]_l$$

Estas reglas describen las interacciones entre la señal S y la proteína Tf permitiendo la formación y disociación del complejo the complex $S-Tf$.

$$r_5^{cyto}: [S-Tf + gene-Tf]_{cyto} \xrightarrow{c_5^{cyto}} [S-Tf-gene-Tf]_{cyto}$$

$$r_6^{cyto}: [S-Tf-gene-Tf]_{cyto} \xrightarrow{c_6^{cyto}} [S-Tf + gene-Tf]_{cyto}$$

$$r_7^{cyto}: [S-Tf-gene-Tf]_{cyto} \xrightarrow{c_7^{cyto}} [S-Tf-gene-Tf + Tf]_{cyto}$$

$$r_8^{cyto}: [S-Tf + gene-R]_{cyto} \xrightarrow{c_8^{cyto}} [S-Tf-gene-R]_{cyto}$$

$$r_9^{cyto}: [S-Tf-gene-R]_{cyto} \xrightarrow{c_9^{cyto}} [S-Tf + gene-R]_{cyto}$$

$$r_{10}^{cyto}: [S-Tf-gene-R]_{cyto} \xrightarrow{c_{10}^{cyto}} [S-Tf-gene-R + R]_{cyto}$$

Las reglas anteriores especifican la activación de los genes $gene-Tf$ y $gene-R$ por interacción directa con $S-Tf$ (r_5^{cyto} , r_6^{cyto} , r_8^{cyto} y r_9^{cyto}); y la producción activa de Tf y R (r_7^{cyto} y r_{10}^{cyto}).

$$r_{11}^{cyto}: [R]_{cyto} \xrightarrow{c_{11}^{cyto}} R []_{cyto}$$

$$r_{12}^{cyto}: [Tf]_{cyto} \xrightarrow{c_{12}^{cyto}} []_{cyto}$$

$$r_{13}^{cyto}: [R]_{cyto} \xrightarrow{c_{13}^{cyto}} []_{cyto}$$

$$r_{14}^{cyto}: [S]_{cyto} \xrightarrow{c_{14}^{cyto}} []_{cyto}$$

La regla r_{11}^{cyto} representa la difusión fuera del citoplasma o la localización del receptor en la superficie celular. Las reglas r_{12}^{cyto} , r_{13}^{cyto} y r_{14}^{cyto} especifican la degradación en el citoplasma de la proteína Tf , el receptor R y la señal S , respectivamente.

Una vez que la correspondiente especificación, Π_0 , ha sido formulada hay que identificar los parámetros asociados, $\mathcal{P}(\Pi_0)$. Estos parámetros consisten en:

- Los multiconjuntos iniciales, $\mathcal{M}_0(\Pi_0) = (M_1, M_2, M_3)$, en donde $M_1 = (l_1, w_1, s_1)$ se asocia al entorno, $M_2 = (l_2, w_2, s_2)$ se asocia a la superficie celular, y $M_3 = (l_3, w_3, s_3)$ se asocia al citoplasma.
- Las constantes de las reglas $\mathcal{C}(\Pi_0) = (c_1^{env}, c_2^{env}, c_1^{surf}, c_2^{surf}, c_3^{surf}, c_1^{cyto}, \dots, c_{14}^{cyto})$.

Entonces ya estamos en condiciones de definir una familia de modelos de sistemas P mediante la especificación de los valores asignados a los parámetros $\mathcal{M}_0(\Pi_0)$ and $\mathcal{C}(\Pi_0)$.

Los posibles valores asociados con los multiconjuntos iniciales, $\mathbb{M}_0 = (\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3)$, nos permitirá estudiar el comportamiento del sistema bajo diferentes condiciones iniciales. En nuestro caso, estudiaremos la evolución del mismo para diferentes número de señales S en el entorno.

$$\mathbf{M}_1 = \{(env, \lambda, \lambda), (env, S^{50}, \lambda)\}$$

$$\mathbf{M}_2 = \{(surf, \lambda, \lambda), (surf, R^3, \lambda)\}$$

$$\mathbf{M}_3 = \{(cyto, gene-Tf + gene-R, \lambda), (cyto, R^3 + gene-Tf + gene-R, \lambda)\}$$

3.3. Teoría Cinética de Gillespie para la Evolución de los Modelos de Sistemas P

En el campo de la Computación Celular con Membranas, juegan un papel esencial el aspecto discreto de las diferentes componentes así como el carácter compartimentalizado y distribuido de la estructura en donde las computaciones tienen lugar. Este no es el caso cuando se usa la semántica no-determinista y paralela maximal que ha sido estudiada en muchas variantes [28, 36]. En esta sección, la aproximación original es reemplazada por varias estrategias que asocian diferentes tasas a las reglas en función de la configuración actual del sistema. Nuestra estrategia se basará en la teoría cinética y estocástica de Gillespie [40, 41, 42, 43, 44].

Para ello, se le asociará a cada regla una constante c que dependa de las propiedades físicas de las moléculas involucradas en la reacción descrita por la regla, así como de otros parámetros físicos como la temperatura. Dicha constante representará la probabilidad por unidad de tiempo para que una reacción particular pueda ser ejecutada y es usada para hallar la *propensidad* de cada regla que determinará la probabilidad y el tiempo necesario para su ejecución.

A nivel microscópico del funcionamiento de procesos celulares, las interacciones entre las moléculas siguen las leyes de la física, en concreto la famosa ley \sqrt{n} , que establece que los niveles de aleatoriedad o fluctuación del sistema son inversamente proporcionales a la raíz cuadrada del número de partículas. En consecuencia, los sistemas con un bajo número de moléculas muestran altas fluctuaciones y la aplicación de las aproximaciones clásicas deterministas y/o continuas para el modelado de sistemas celulares es, cuando menos, cuestionable.

A continuación se presentan diferentes estrategias para la evolución de los modelos de sistemas P basadas en el algoritmo de Gillespie.

3.3.1. Algoritmo Multicompartimental de Gillespie

En esta sección se presentan dos algoritmos que proporcionan sendas semánticas de sistemas celulares con membranas especialmente adecuadas para el modelado y la simulación de procesos celulares.

3.3.2. Algoritmo de Gillespie multicompartimental

El algoritmo (clásico) de Gillespie es un algoritmo de Monte Carlo para la simulación estocástica de interacciones moleculares que tienen lugar en un determinado volumen. Ha sido utilizado con éxito para la simulación estocástica de procesos biológicos que tienen lugar en un

determinado volúmen fijo. Ahora bien, los sistemas P constan de una estructura que delimitan regiones (volúmenes), cada una con su propio conjunto de reglas y multiconjunto de objetos. Por ello, definimos una extensión de dicho algoritmo (*Algoritmo de Gillespie Multicompartimental*) a fin de que cada célula (compartimento u orgánulo) evolucione de acuerdo con el algoritmo de Gillespie.

Dado el estado de un compartimento $M_i = (l_i, w_i, s_i)$, la siguiente regla a aplicar se calcula como sigue:

1. Para cada regla $r_j \in R_{l_i}$ se calcula su propensidad, $a_j(M_i)$, multiplicando la constante estocástica $c_j^{l_i}$ asociada a la regla $r_j \in R_{l_i}$ por el número de posibles combinaciones de los distintos objetos y subcadenas presentes en la parte izquierda de la regla en relación con el contenido actual de las membranas involucradas en la regla.
2. Calcular la suma de todas las propensidades:

$$a_0(M_m) = \sum_{r_j \in R_{l_i}} a_j(M_i)$$

3. Generar dos números aleatorios r_1 y r_2 a partir de la distribución uniforme en el intervalo unidad, y seleccionar τ_i y j_i de acuerdo con:

$$\tau_i = \frac{1}{a_0(M_i)} \ln \left(\frac{1}{r_1} \right) \quad (3.3)$$

$$j_m = \text{el menor entero tal que } \sum_{j=1}^{j_i} a_j(\mathbf{x}) > r_2 a_0(M_i) \quad (3.4)$$

Entonces, los compartimentos definidos por las membranas se ordenan en una cola de prioridad. La primera regla a aplicar aparece en la cabeza de la cola. Dependiendo del tipo de regla que se ha seleccionado, el estado de uno o dos compartimentos deberá ser actualizado. Por tanto, se recalcularán los tiempos de espera y la regla a ser aplicada en estos compartimentos. El algoritmo termina cuando el tiempo de simulación alcanza o excede un valor prefijado.

El *algoritmo multi-compartimental de Gillespie* podemos describirlo con más precisión como sigue:

- **Inicialización**

- Poner el tiempo de simulación a $t = 0$.
- Para cada membrana m en la estructura μ hacer:
 1. Ejecutar el algoritmo de Gillespie clásico en m , devolviendo (τ_m, j, m) .
- Ordenar la lista de ternas (τ_m, j, m) con relación a τ_m (en orden ascendente), formando una lista de reglas a aplicar.

- **Iteración**

- Elegir la primera terna de la lista, (τ_m, j, m) .

- Poner el tiempo de simulación a $t = t + \tau_m$.
- Actualizar el tiempo de espera para el resto de las ternas, restando τ_m .
- Aplicar una vez la regla r_j modificando las membranas afectadas.
- Para cada membrana m' afectada por esa aplicación, eliminar la terna $(\tau'_{m'}, j', m')$ y ejecutar el algoritmo de Gillespie clásico en el nuevo contexto, obteniendo una nueva terna $(\tau''_{m'}, j'', m')$.
- Añadir la terna $(\tau''_{m'}, j'', m')$ a la lista y ordenarlas de nuevo de acuerdo con los tiempos de espera (en orden ascendente).
- Iterar el proceso.

• Finalización

- Si el tiempo de la simulación t alcanza o excede a un tiempo maximal prefijado, finalizar el proceso.

Como ejemplo del uso de esta estrategia hemos elegido dos modelos de la familia $\mathcal{F}(\Pi_0, (\mathbb{M}_0, \mathbb{C}))$ definida en la sección previa:

$(\Pi_0; (\mathbb{M}_1(1), \mathbb{M}_2(1), \mathbb{M}_3(1)), C)$ y $(\Pi_0; (\mathbb{M}_1(2), \mathbb{M}_2(2), \mathbb{M}_3(2)), C)$

que representan las diferentes condiciones iniciales (ausencia o presencia de señales en el entorno, respectivamente), y ambos comparten las mismas constantes

$C = (0,16, 10^{-3}, 2, 4, 10^{-2}, 10^{-2}, 10^{-3}, 0,1, 8, 5, 10, 0,2, 3, 10, 8 \times 10^{-2}, 5 \times 10^{-3}, 5 \times 10^{-3}, 10^{-3}, 3 \times 10^{-3})$

En las figuras 3.2 y 3.3, se presentan dos realizaciones de estos modelos y una estimación de la probabilidad de tener diferentes números de moléculas al cabo de 500 *sec* usando 1000 diferentes ejecuciones.

3.4. Model Checking Probabilístico sobre Modelos de Sistemas P usando PRISM

Muchas investigaciones en Biología de Sistemas se centran en el desarrollo de modelos suficientemente fiables que sean capaces de revelar nuevas propiedades que sean muy difíciles o imposible de descubrir directamente mediante experimentos de laboratorio. Una cuestión fundamental es qué se puede hacer con ese tipo de modelos, además de realizar simulaciones ¿Es suficiente realizar simulaciones para obtener realmente nuevo conocimiento? Esta cuestión ha sido abordada con detalle para modelos deterministas en los que se ha desarrollado una teoría que ha sido utilizada en modelos basados en ODEs. No obstante, la situación es bien distinta en el caso de modelos estocásticos y este campo está abierto a avances teóricos que nos ayuden a razonar sobre los sistemas con más detalle y mejor precisión.

Existen varios intentos en esta dirección que consiste en aplicar herramientas de model checking a modelos computacionales de sistemas celulares con el fin de analizar de manera automática propiedades de tales modelos. Los métodos así desarrollados no tratan de reemplazar la simulación clásica y las aproximaciones basadas en ODEs, sino que esos métodos deberían

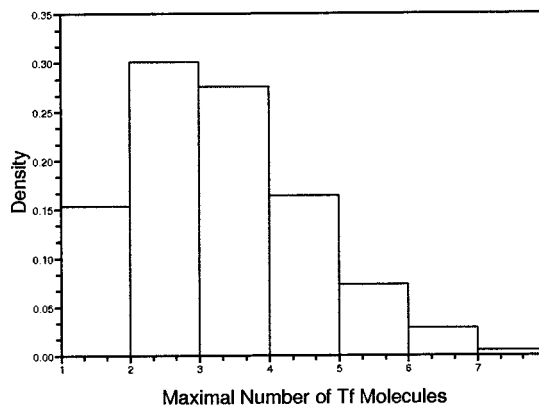
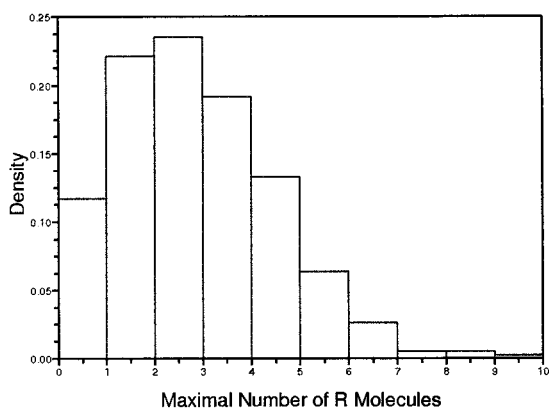
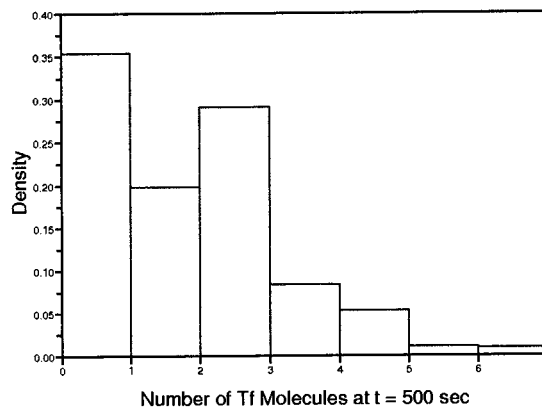
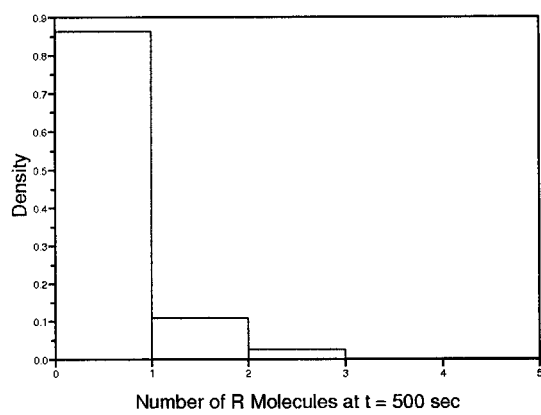
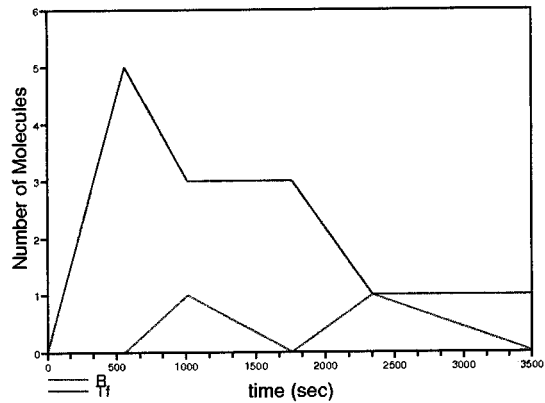
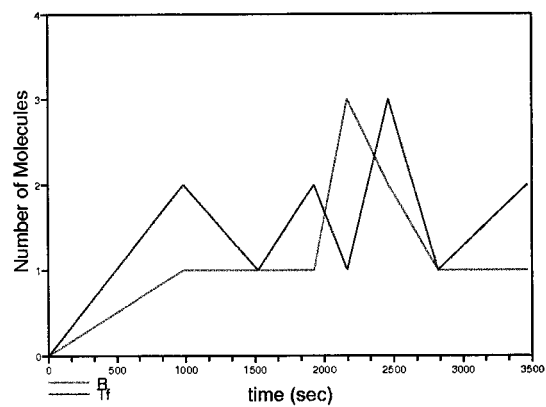


Figura 3.2:

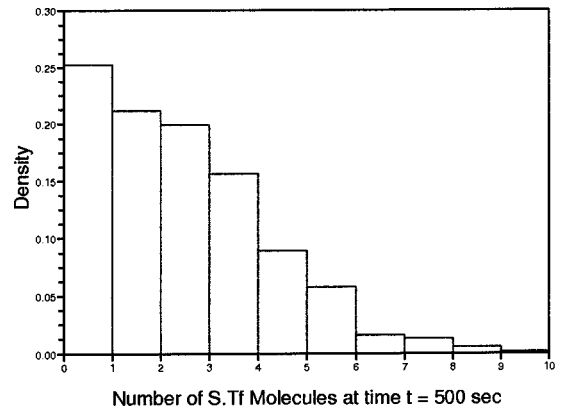
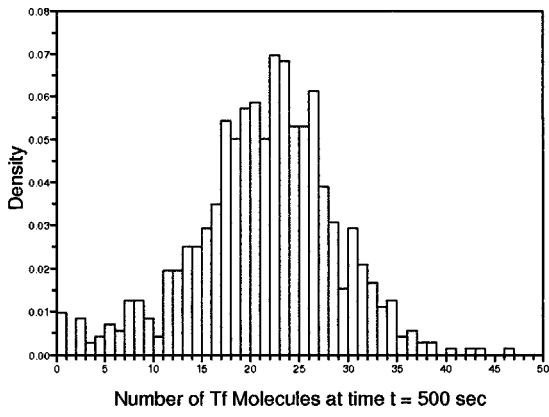
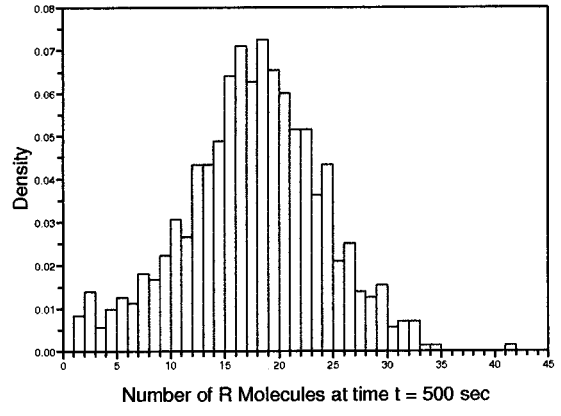
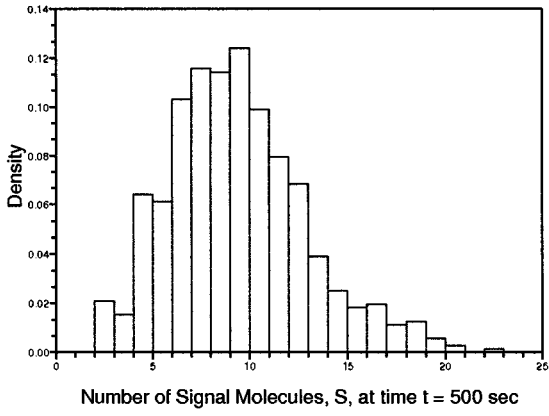
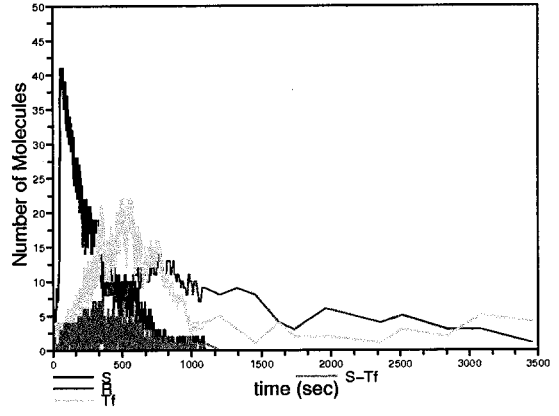
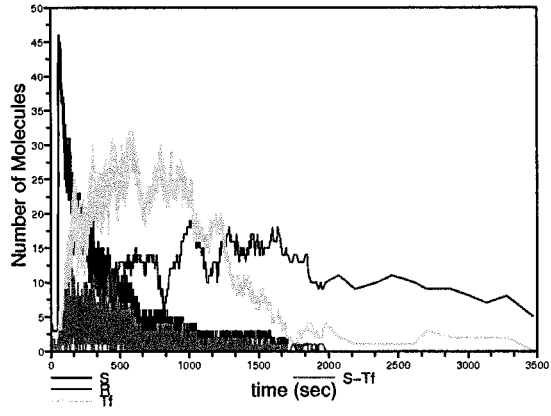


Figura 3.3:

ser usados conjuntamente con otras aproximaciones a fin de conseguir ampliar el conocimiento acerca de interacciones complejas de los sistemas celulares.

A este respecto, existen algunos estudios previos que analizan el uso de técnicas de model checking para modelos de sistemas P [7, 29]. En esta sección, se propone el uso de un modelo simbólico y probabilístico de model checking basado en PRISM (Probabilistic and Symbolic Model Checker) [133].

3.4.1. Model Checking Probabilístico con PRISM

El *Model checking* es un método formal bien establecido y ampliamente usado para verificar la corrección de sistemas reales. El *Model checking probabilístico* es una variante probabilística del model checking clásico al que se le ha añadido información cuantitativa acerca de las transiciones que se realizan y de los instantes en que suceden [111]. Usualmente, dicho modelo trabaja con *cadena de Markov* o *procesos de decisión de Markov*. En esta memoria, basta considerar *cadena de Markov con tiempo-continuo* (CTMC).

El *Model checking probabilístico* ha sido utilizado con éxito, como técnica de verificación formal, al análisis de sistemas complejos de amplio espectro, que incluyen protocolos de comunicación y seguridad, algoritmos distribuidos, etc. [73].

Los métodos analíticos del model checking probabilístico consta de tres pasos diferentes:

1. En primer lugar se debe diseñar un modelo matemático preciso del sistema que va a ser analizado. En esta memoria, las especificaciones y modelos de sistemas P serán usados como la descripción formal requerida en este paso.
2. A continuación se debe traducir al lenguaje específico de model checking probabilístico, PRISM en nuestro caso, el modelo formal que se desea analizar.
3. Finalmente, algunas propiedades del modelo deben ser identificadas y expresadas formalmente usando una lógica temporal. Ello permite al chequeador analizar dichas propiedades de manera automática.

Una de las mayores ventajas del model checking probabilístico consiste en que se trata de una aproximación exhaustiva; es decir, que todos los posibles comportamientos del sistema son susceptibles de ser analizados.

En esta sección usaremos la herramienta software PRISM (Probabilistic and Symbolic Model Checker) para realizar model checking probabilístico sobre modelos de sistemas P. Para un estudio detallado de PRISM véase [133].

3.4.2. Transformando Especificaciones de Sistemas P en PRISM

Consideremos una especificación de un sistema P $\Pi = (O, L, \mu, M_1, M_2, \dots, M_n, R_{l_1}, \dots, R_{l_m})$, las respectivas componentes serán descritas en PRISM como sigue:

- **Estructura de membranas:** Cada membrana en μ está unívocamente determinada por un identificador i , $1 \leq i \leq n$. Usaremos módulos para describir el comportamiento de las membranas. A este respecto, se introduce un módulo denominado `compartment_i` para cada membrana i .

- **Alfabeto y multiconjuntos iniciales:** Para cada objeto $obj \in O$ presente en el compartimento definido por una membrana i , se declara una variable local obj_i en el módulo `compartment_i`.

El valor inicial de esta variable viene determinado por el multiconjunto inicial $M_i = (l_i, w_i, s_i)$ asociado a la membrana i . Teniendo presente que los multiconjuntos iniciales son parte de los parámetros de una especificación de sistema P, se define una constante, `ini_obj_i`, sin valor asignado, que permitirá considerarla como un parámetro del modelo PRISM.

El rango de los valores de las variables que representan objetos será determinado experimentalmente o extraído de la literatura. Con el fin de especificar estos rangos, se declaran dos constantes `upb_obj_i` and `lob_obj_i`.

- **Reglas de reescritura:** Los comandos se usan en PRISM para describir las reglas de reescritura de una especificación de sistema P PRISM, y no admiten el uso directo de cadenas; por ello, nos restringiremos al uso de reglas que afectan a objetos (son las que describen interacciones proteína-proteína).

En general, dichas reglas necesitan dos membranas para interactuar e intercambiar objetos de una manera sincronizada. Supongamos que la membrana i está contenida en la membrana k descrita por el módulo `compartment_k`. En el caso en que la regla afecte a dos compartimentos diferentes, los módulos que los representa sincronizarán la aplicación de dos comandos diferentes, usando la etiqueta `r_j_l_i`. En el caso en que la regla sólo involucre a un compartimento, la etiqueta `r_j_l_i` también será usada.

Por tanto, dada una regla

$$r_j^i : obj_1 [obj_2]_l \xrightarrow{c_j^i} obj'_1 [obj'_2]_l$$

siendo $obj_1 = o_1^1 + \dots + o_{n_1}^1, obj_2 = o_1^2 + \dots + o_{n_2}^2, obj'_1 = oo_1^1 + \dots + oo_{m_1}^1, obj'_2 = oo_1^2 + \dots + oo_{m_2}^2$ multiconjuntos finitos, y c_j^i la constante estocástica asociada a la regla, las variables `o_1_1_k, ..., o_n_1_1_k, oo_1_1_k, ..., oo_m_1_1_k` especifican los objetos de obj_1 y obj'_1 en el módulo `compartment_k`, y los objetos `o_1_2_i, ..., o_n_2_2_i` and `oo_1_2_i, oo_m_2_2_i` representan los objetos de obj_2 y obj'_2 en el módulo `compartment_i`.

Las constantes estocásticas asociadas a las reglas son parte de los parámetros de la especificación del sistema P y serán especificadas en los módulos de PRISM usando constantes sin valor asignado.

3.4.3. Análisis de Modelos de Sistemas P usando PRISM

En esta sección se ilustra cómo usar PRISM para analizar modelos de sistemas P, considerando el ejemplo de la figura 3.1. Aquí nos centraremos en el modelo $(\Pi_0; (M_1(1), M_2(1), M_3(1)), C)$ presentado en la sección 3.3 y que representa las condiciones iniciales sin señales.

A la hora de analizar un modelo en PRISM, lo primero que hay que hacer es asociar una apropiada recompensa a los correspondientes estados y transiciones:

- A cada estado se asocia una recompensa que viene a ser el número de una molécula particular. Se usa una constante `molec` para identificar la molécula seleccionada en ese instante.

```

const int molec;
rewards "molecules"
molec = 1 : s_1;
molec = 2 : s_3;
molec = 3 : r_2;
:
molec = 11 : s_Tf_gene_r_3;
endrewards

```

- De manera similar se asocia una recompensa igual a 1 a cada transición que represente que la regla se ha aplicado una sólo vez. Se usa una constante `rule` para identificar qué regla se analiza.

```

const int rule;
rewards "rules"
[ r_1_env ] rule = 1 : 1;
[ r_2_env ] rule = 2 : 1;
[ r_1_surf ] rule = 3 : 1;
:
[ r_14_cyto ] rule = 19 : 1;
endrewards

```

Una vez que las correspondientes recompensas han sido asignadas, se puede usar PRISM para chequear algunas propiedades del modelo.

Capítulo 4

Especificaciones de Sistemas P de Sistemas Celulares

En el capítulo 2 se ha discutido brevemente distintas aproximaciones para el modelado de sistemas celulares. Los formalismos de esos modelos fueron establecidos con anterioridad a su uso en dichos sistemas. Por contra, los sistemas P están inspirados directamente de la estructura y el funcionamiento de las células de los organismos vivos.

En este capítulo se presentan los sistemas P como un marco formal para la especificación y simulación de sistemas celulares que integran aspectos estructurales y dinámicos, proporcionando los formalismos necesarios para realizar análisis matemáticos y computacionales.

Los sistemas P representan una aproximación complementaria más que una alternativa a otros modelos clásicos, como por ejemplo el basado en ODEs. En algunas situaciones, los modelos con aproximaciones macroscópica, continua y determinista son cuestionables: por ejemplo:

- En los sistemas celulares con una estructura compartimentalizada y heterogénea, con bajo número de moléculas y reacciones lentas.
- En procesos de regulación de genes, como la transcripción o la traducción (que son procesos inherentemente concurrentes y discretos) y el enlace de factores de transcripción a genes promotores, que tienen básicamente un comportamiento booleano.
- En procesos en los que importa el orden en el que están situados diferentes genes sobre el genoma.

Los sistemas P constituyen un modelo no convencional de computación que representa de manera explícita el carácter discreto de la cantidad de componentes de un sistema celular mediante el uso de reglas de reescritura sobre multiconjuntos de objetos que representan moléculas, y cadenas que describen la organización de los genes en el genoma. La estocasticidad inherente a dichos sistemas se captura mediante el uso de estrategias estocásticas como el *algoritmo multicompartimental de Gillespie's Algorithm* introducido en la sección 3.3.

En lo que sigue se presenta con todo detalle cómo son especificadas las principales componentes de los sistemas celulares y cómo son simuladas usando sistemas P.

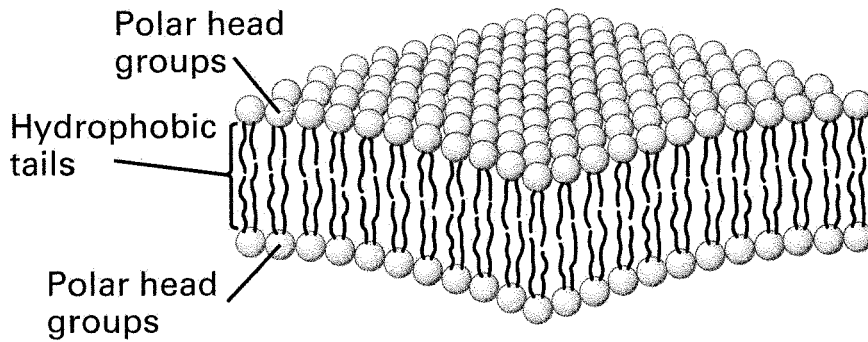


Figura 4.1: Estructura bi-capa de una membrana biológica [76]

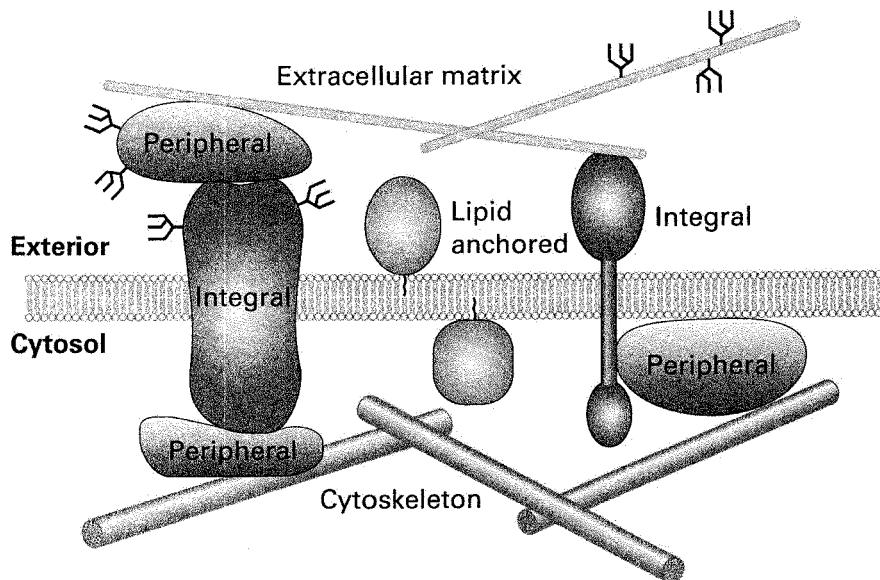


Figura 4.2: Proteínas asociadas a una bi-capa lípida [76]

4.1. Especificaciones de Compartimentos

Las membranas biológicas juegan un papel crucial en la organización estructural y el funcionamiento de células tanto procariontas como eucariotas. La unidad estructural básica de cualquier membrana es un bi-capa fosfolípida que posee una parte hidrófoba y otra hidrófila (ver figura 4.1).

Las células procariontas, que son las más pequeñas ($\sim 1 - 2\mu m$), presentan una estructura simple que consta de un único compartimento rodeado por una membrana plasmática.

La función de una membrana biológica no se reduce a una serie de compartimentos debidamente jerarquizados en el interior de los cuales se almacenan moléculas específicas que reaccionan entre sí, sino que, además, controlan el movimiento de moléculas entre un compartimento y otro. Estos procesos son ejecutados por determinadas proteínas de transporte que están asociadas a las membranas, ver figura 4.2. La importancia de este tipo de proteínas radica en el hecho de que, aproximadamente, una tercera de los genes codifican este tipo de proteínas.

En el marco de los sistemas P, las membranas se usan para definir regiones relevantes en

los sistemas celulares. En esta memoria se usarán dos membranas diferentes para especificar las dos regiones relevantes asociadas al compartimento celular:

1. Una primera membrana representará la superficie celular y definirá una región en la que se situarán las proteínas periféricas. Por ejemplo, los procesos que involucran moléculas de transporte y señalización se representarán a través de reglas asociadas a esta región.
2. Otra membrana describirá el interior acuoso del compartimento y, por tanto, estará dentro de la anterior membrana. El conjunto de objetos y cadenas que especifican las proteínas y otras moléculas localizadas en el lumen, serán colocadas en la región definida por esta membrana.

En aquellos casos en los que no intervienen proteínas de transporte ni hay procesos de señalización, se suprimirá la membrana que representa la superficie celular.

4.2. Especificaciones en Sistemas P de las Interacciones Proteína-Proteína

Las interacciones proteínicas son, posiblemente, uno de los procesos más importantes que tienen lugar en las células vivas. La descripción teórica y experimental de dichas interacciones pertenecen al campo de la química cinética, que tiene como objetivo fundamental el cálculo de la probabilidad o propensidad de una interacción, con el fin de determinar la tasa de reactantes que son transformados en productos.

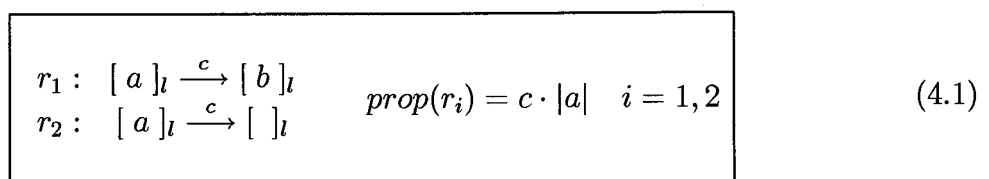
En esta sección se presenta un esquema de modelado en sistemas P para cada una de dichas interacciones, y las respectivas propensidades se calculan de acuerdo con la teoría cinética estocástica de Gillespie. [40, 41, 42, 43, 44].

Para cada uno de dichos esquemas, se presenta una gráfica obtenida mediante CellDesigner, un editor de diagrama estructurado para redes bioquímicas y reguladoras de genes. [37].

■ Transformación y degradación:

Una molécula a puede reaccionar para producir otra molécula b , o bien puede ser degradada por la propia maquinaria celular (figura 4.3). La dinámica de estas reacciones se modela usando la *ley de decaimiento exponencial*. Esta ley establece que la tasa de la reacción o su propensidad es proporcional al número de moléculas del reactante a .

Los procesos de transformación y degradación pueden especificarse en sistemas P usando reglas de reescritura:



Como ya se discutió en la sección 2.6, para este tipo de reacciones, la constante mesoscópica y estocástica, c , usada en aproximaciones similares a la que hacemos aquí, es igual a la constante macroscópica y determinista, k , usada en ODEs.

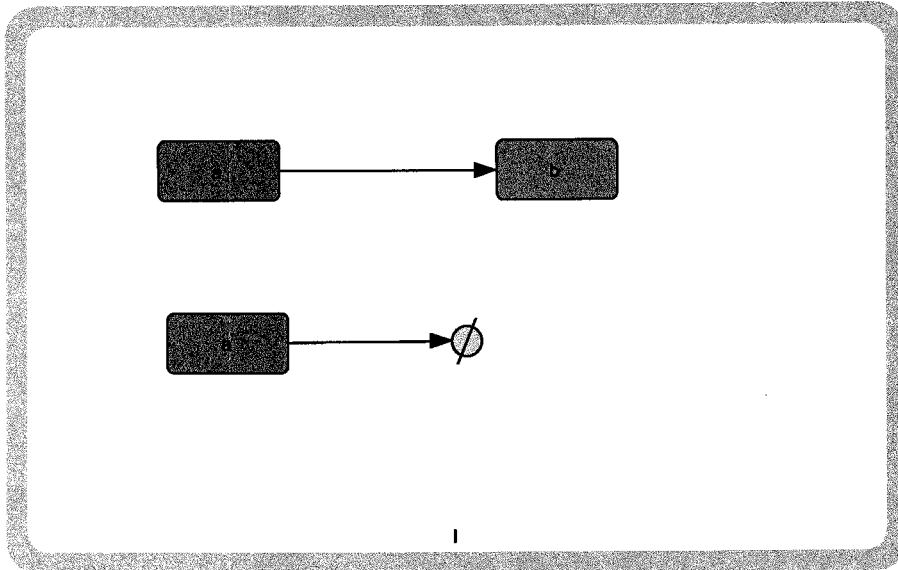


Figura 4.3: Representación gráfica de la transformación y degradación.

■ **Formación de complejos y disociación:**

Dos moléculas, a y b , pueden colisionar y producir un complejo c . A su vez, dicho complejo puede disociarse en dos componentes, d y e , que podrían haberse modificado como producto de la interacción.

En bioquímica, estas reacciones se denominan de formación (heterodímera si $a \neq b$ y homodímera si $a = b$) o disociación de complejos, ver figura 4.2.

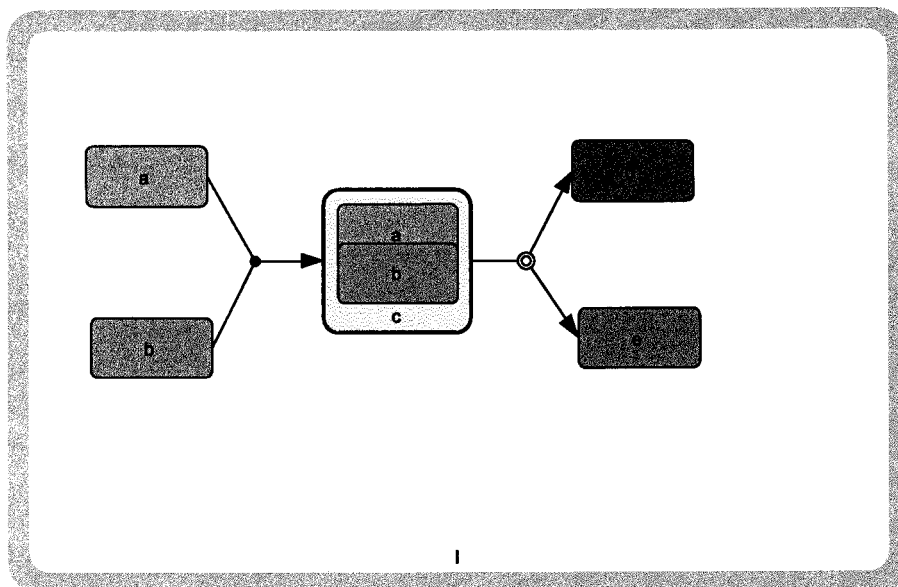


Figura 4.4: Representación gráfica de la formación y disociación de complejos.

La dinámica de estas reacciones siguen la *ley de acción de masas*, que establece que la tasa o propensidad de la reacción es directamente proporcional al producto del número

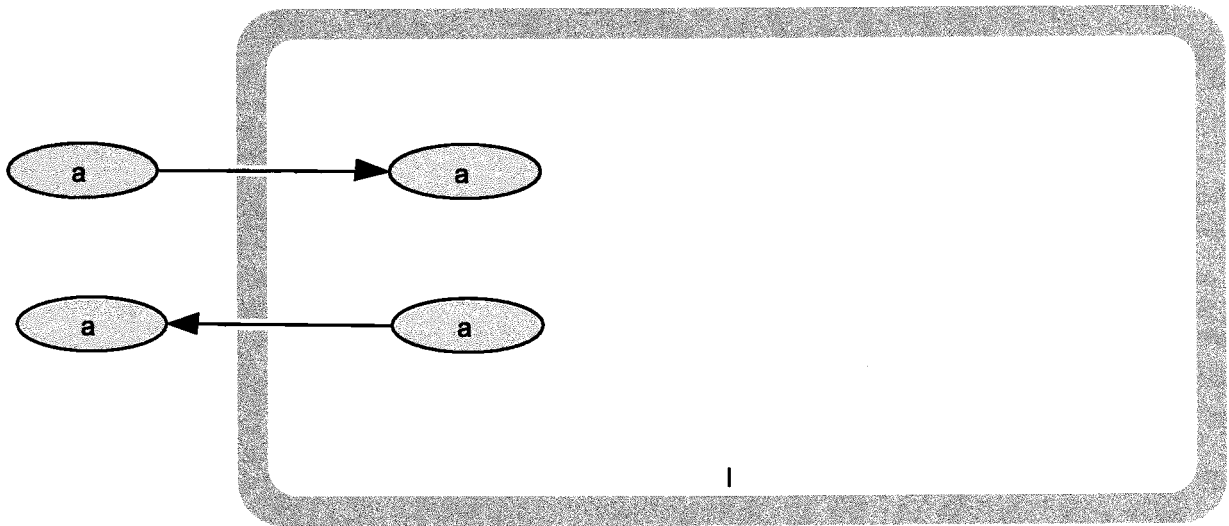


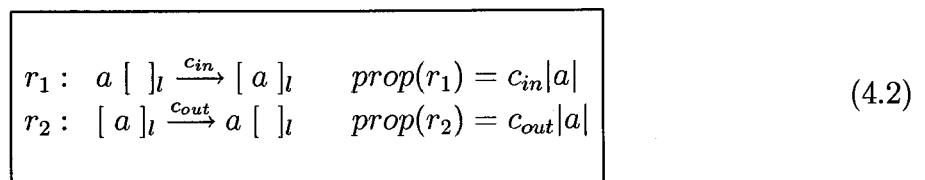
Figura 4.5: Difusión a través de la membrana celular.

de moléculas reactantes. Por tanto, dos constantes c_{cf} y c_{cd} se asociarán a las reacciones de formación y disociación de complejos, que permitirán calcular sus respectivas propensidades.

■ **Difusión hacia dentro y hacia fuera:**

Pequeñas moléculas pueden moverse lentamente por difusión pasiva a través de las membranas sin necesidad de ayudarse de proteínas de transporte y sin consumir energía metabólica alguna. El movimiento suele tener lugar desde regiones de alta concentración a regiones de baja concentración (4.5).

Las reglas de reescritura (4.2) especifican en sistemas P la difusión de un objeto a hacia fuera o hacia dentro de un compartimento identificado por la etiqueta l . Dos constantes c_{in} y c_{out} se asocian a dichas reglas a fin de calcular sus respectivas propensidades.



■ **Enlace y rotura de enlace:**

Uno de los pasos más importantes en el proceso de convertir señales en respuesta celular, *traducción de señales*, es el enlace de moléculas señalizadoras a lugares complementarios extracelulares (receptores), que permiten la activación del receptor (4.6).

Obsérvese que en el proceso de traducción de señales, la membrana celular juega un papel esencial ya que es el lugar en donde están colocados los receptores. Por ello, en nuestra

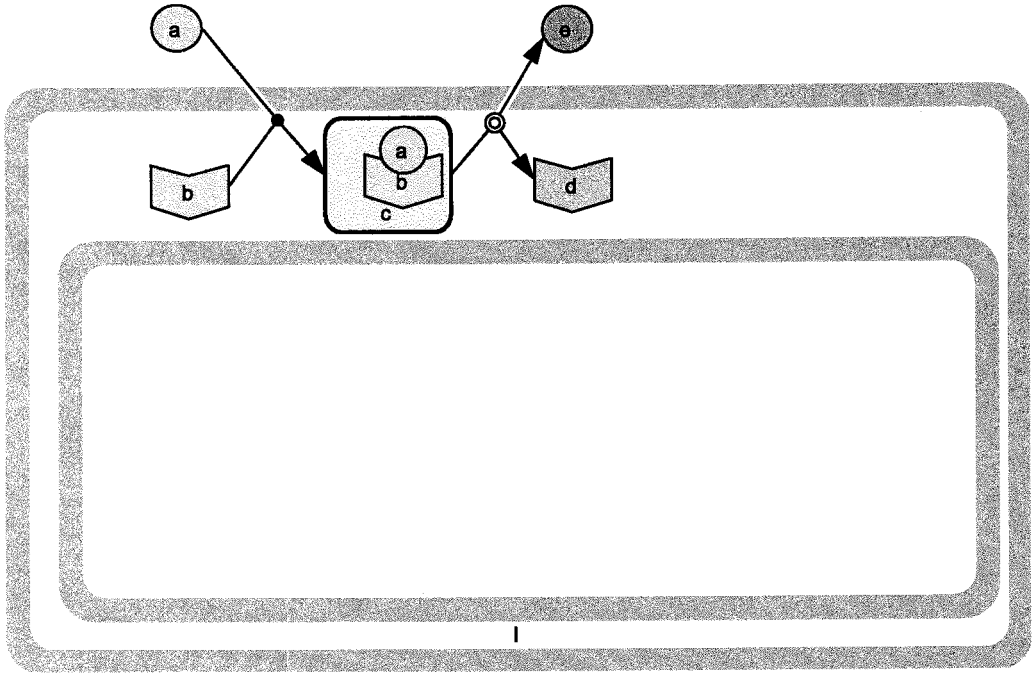
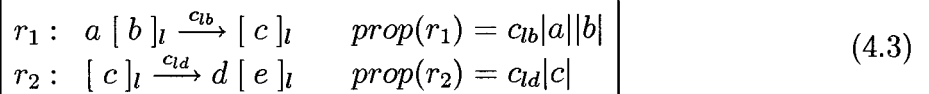


Figura 4.6: Enlace de un ligando a un receptor transmembrana.

especificación, la membrana plasmática será representada por la región determinada por dos membranas: una de ellas representa la superficie celular y la otra el citoplasma.

En sistemas P, el enlace de un ligando a su receptor (y la rotura de ese enlace) son especificadas mediante las reglas de reescritura (4.3). Para la regla de enlace, el objeto a representa el ligando fuera de la superficie celular (identificada por la etiqueta l) y el objeto b representa el receptor. Estos dos objetos son reemplazados por el objeto c (el complejo receptor-ligando). La reacción de rotura de enlace (4.6) se especifica reemplazando el objeto c por el objeto d , que representa el ligando, y el objeto e , que representa el receptor liberado.



Dichas reacciones tienen asociadas dos constantes, c_{cf} y c_{cd} , que permiten calcular sus respectivas propensidades de acuerdo con la ley de acción de masas.

■ **Reclutamiento y liberación:**

El enlace de un ligando a un receptor produce un cambio en la conformación del dominio citosólico que dispara el reclutamiento de algunas proteínas citoplasmática. Estas proteínas son, posteriormente, transformadas o liberadas en el citoplasma, produciendo respuestas celulares (4.7).

Las reglas (4.4) especifican en los sistemas P los procesos de reclutamiento y liberación de proteínas. En la regla de reclutamiento, r_{rt} , el receptor activo se representa por el

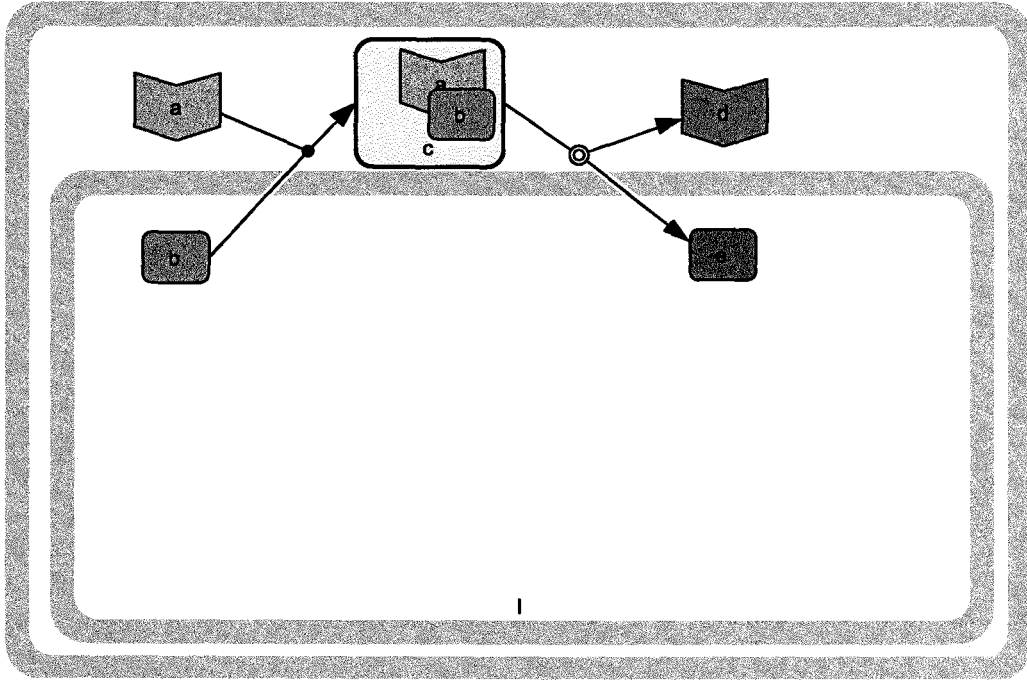


Figura 4.7: Reclutamiento de una proteína del citoplasma por un receptor transmembrana.

objeto a , colocado en el compartimento identificado por la etiqueta l , en donde el objeto b representa la proteína que es reclutada. Estos objetos son reemplazados por el objeto c fuera del compartimento l .

Recíprocamente, en las reglas de liberación, r_{rl} , el objeto c fuera del compartimento l se reemplaza por los objetos d fuera del compartimento y el objeto e dentro del compartimento.

Dos constantes, c_{rt} y r_{rl} , se asocian a dichas reglas y permiten el cálculo de sus respectivas propensidades de acuerdo con la ley de acción de masas.

$$\begin{array}{l}
 r_{rt} : a [b]_l \xrightarrow{c_{rt}} c []_l \quad \text{prop}(r_{rt}) = c_{rt}|a||b| \\
 r_{rl} : c []_l \xrightarrow{c_{rl}} d [e]_l \quad \text{prop}(r_{rl}) = c_{rl}|c|
 \end{array} \tag{4.4}$$

4.3. Especificaciones en Sistemas P de Redes de Transcripción

Las células vivas pueden percibir señales complejas del entorno a través de interacciones moleculares descritas en la sección anterior, produciendo una serie de respuestas mediante la producción de adecuadas proteínas que están codificadas en genes específicos. La tasa de producción de esas proteínas están reguladas por proteínas específicas, denominadas *factores de transcripción*, que se enlazan a genes incrementando o decrementando la tasa de transcripción que tienen grabado en el mRNA. En este sentido, las células usan factores de transcripción como

una representación interna del entorno y del estado interno de la célula. La interacción entre factores de transcripción y genes que permiten la alteración de la tasa de producción de ciertas proteínas son descritas por medio de *redes de transcripción*. En esta sección se presenta unos esquemas de especificación de sistemas P para redes de transcripción en células procariotas (que las consideramos por simplicidad).

4.3.1. Especificación de las Redes de Transcripción usando Objetos

En una aproximación simplista, podemos considerar que los procesos de transcripción y traducción son reacciones individuales. En tal situación, los genes y operons serán especificados mediante objetos individuales que producen en un único paso sus complementarios mRNA, que también se representarán por objetos individuales. La producción de una proteína a partir del mRNA se describirá en un sólo paso. Finalmente, los procesos involucrados en el control de la expresión del gen, se especifican a través de reglas de reescritura sobre multiconjuntos de objetos.

En lo que sigue, se presenta unos esquemas de especificación de sistemas P usando objetos y reglas de reescritura sobre multiconjuntos de objetos con el fin de describir los procesos que tienen lugar en las redes de transcripción.

- **Transcripción y traducción:**

El dogma de la biología molecular establece que la información genética está codificada en el DNA. Esta información se transcribe en el mRNA que se traduce en proteínas (4.8).

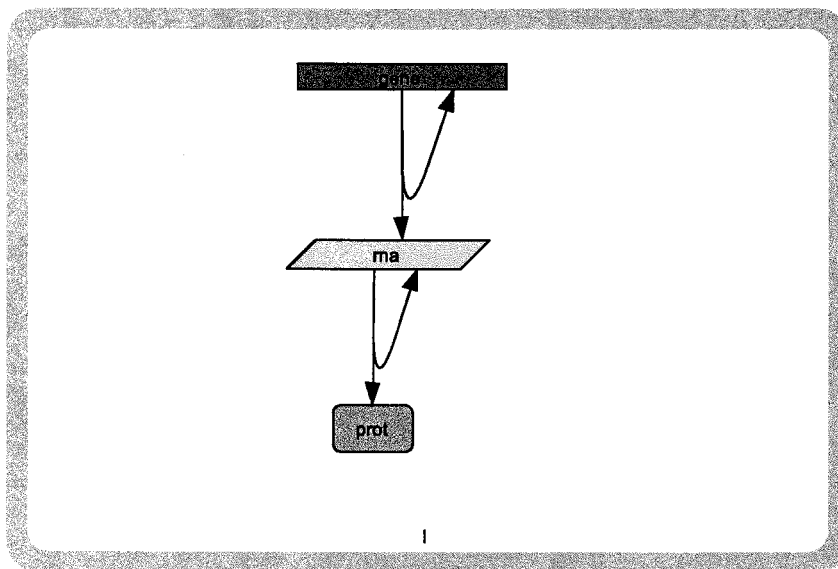


Figura 4.8: Transcripción y traducción de un gen que codifica la proteína *prot*.

La información lineal contenida en genes y mRNAs se representan por objetos individuales. Los objetos *gene*, *rna* y *prot* especifican los trozos de DNA que consisten en el gen, su complementario mRNA, y su gen producto o proteína, respectivamente.

La transcripción de un gen en su complementario mRNA se describe por la regla de reescritura r_{tc} . De forma similar se describe la traducción mediante una regla r_{tl} . Las

propensidades de dichas reglas se calculan de la misma forma que las indicadas en la sección anterior.

$$\begin{array}{l}
 r_{tc} : [gene]_l \xrightarrow{c_{tc}} [gene + rna]_l \quad prop(r_{tc}) = c_{tc}|gene| \\
 r_{tl} : [rna]_l \xrightarrow{c_{tl}} [rna + prot]_l \quad prop(r_{tl}) = c_{tl}|rna|
 \end{array}
 \tag{4.5}$$

En una especificación más simplista, se puede ignorar el paso intermedio de producción del mRNA y admitir que la proteína es producida directamente a partir del gen en un simple proceso descrito en (4.6).

$$r_{prod} : [gene]_l \xrightarrow{c_{prod}} [gene + prot]_l \quad prop(r_{prod}) = c_{prod}|gene|
 \tag{4.6}$$

■ **Enlace y rotura de enlace de factores de transcripción en genes:**

Las reglas de reescritura (4.5) sólo permiten especificar genes cuya expresión no dependa de los factores de transcripción y, por tanto, de ninguna señal interna o externa. En esos casos hay que introducir nuevas especificaciones.

Existen, básicamente, dos tipos de factores de transcripción: *activadores* y *represores*. Si bien ambos tipos enlaza a genes (4.9) tienen efectos opuestos ya que mientras los activadores incrementan la tasa de transcripción de genes, los represores la disminuye.

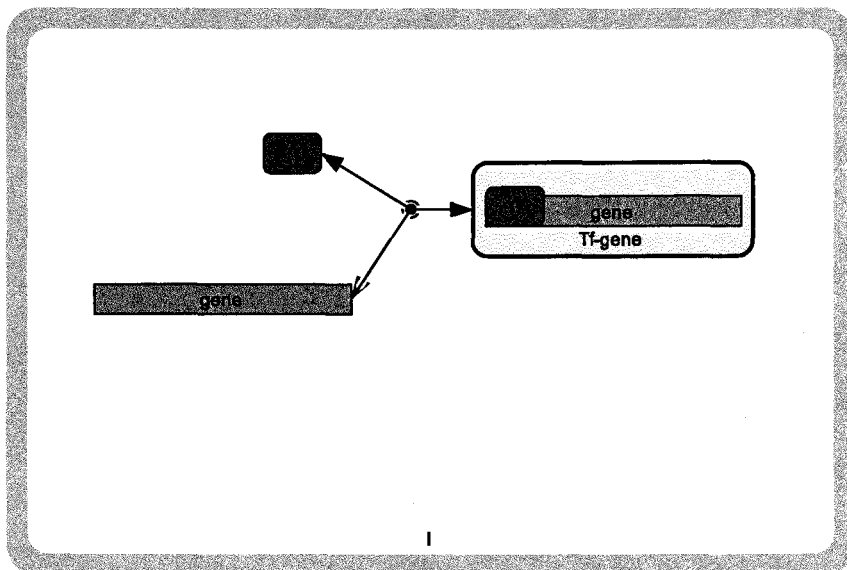


Figura 4.9: Enlace y rotura del enlace de un factor de transcripción Tf a un gen.

Los procesos de enlace y rotura de enlace de un factor de transcripción a un gen puede ser descrito mediante reglas similares a las de formación y disociación de complejos. La

especificación de sistema P (4.7) constituye la descripción de estos procesos mediante reglas de reescritura sobre multiconjuntos de objetos.

La regla r_{gon} representa el enlace de un factor de transcripción, representado por el objeto Tf , a un gen, especificado por un objeto $gene$. El proceso inverso se especifica mediante la regla r_{goff} .

$$\begin{array}{l}
 r_{gon} : [Tf + gene]_l \xrightarrow{c_{gon}} [Tf-gene]_l \quad prop(r_{gon}) = c_{gon}|Tf||gene| \\
 r_{goff} : [Tf-gene]_l \xrightarrow{c_{goff}} [Tf + gene]_l \quad prop(r_{goff}) = c_{goff}|Tf-gene|
 \end{array} \quad (4.7)$$

4.3.2. Especificación de las Redes de Transcripción usando Cadenas

Con el fin de modelar con mayor precisión algunos procesos involucrados en las redes de transcripción, incorporaremos algunos hechos como retrasos, que emerge de la organización lineal de genes en el DNA y la diferencia de tiempo entre la transcripción y la traducción. Además, el orden en que los genes estén colocados en los operons puede ser relevantes. Por ello, usaremos estructuras lineales, como cadenas de un alfabeto, a fin de producir modelos relevantes de redes de transcripción de genes.

En esta sección se propone el uso de cadenas para representar la estructura lineal de las hebras de DNA y RNA, así como el uso de reglas de reescritura sobre cadenas para describir el enlace y la rotura de enlaces de los factores de transcripción a genes, así como los procesos de transcripción y traducción como procesos paralelos y concurrentes.

- **Enlace y rotura de enlaces de factores de transcripción a lugares específicos del DNA:**

El esquema de especificación de sistema P (4.8) describe el enlace y la rotura de enlaces de un factor de transcripción representado por el objeto Tf a un operador especificado por la subcadena $\langle site \rangle$. La regla r_{tfb} describe el enlace del factor de transcripción. El proceso inverso se describe mediante la regla r_{tfd} .

$$\begin{array}{l}
 r_{tfb} : [Tf + \langle site \rangle]_l \xrightarrow{c_{tfb}} [\langle site' \rangle]_l \quad prop(r_{tfb}) = c_{on}|Tf||\langle site \rangle| \\
 r_{tfd} : [\langle site' \rangle]_l \xrightarrow{c_{tfd}} [Tf + \langle site \rangle]_l \quad prop(r_{tfd}) = c_{off}|\langle site' \rangle|
 \end{array} \quad (4.8)$$

La constante c_{tfb} asociada a la regla r_{tfb} representa la afinidad entre el factor de transcripción y el operador.

- **Transcripción:**

La transcripción es el proceso por el cual una proteína específica, denominada **RNA polimerasa** sintetiza, a partir de una sucesión de DNA, la correspondiente sucesión

complementaria de RNA. Durante la síntesis de RNA, el lenguaje del DNA que consta de sucesiones de A (adenina), G (guanina), C (citosina) y T (timina), se copia o transcribe en el lenguaje del RNA, en donde la T (timina) se reemplaza por U (uracilo).

Para realizar la transcripción, la RNA polimerasa realiza distintas funciones, como aparecen en la figura 4.10. A continuación, detallamos los esquemas de sistemas P que usaremos para especificar esas interacciones.

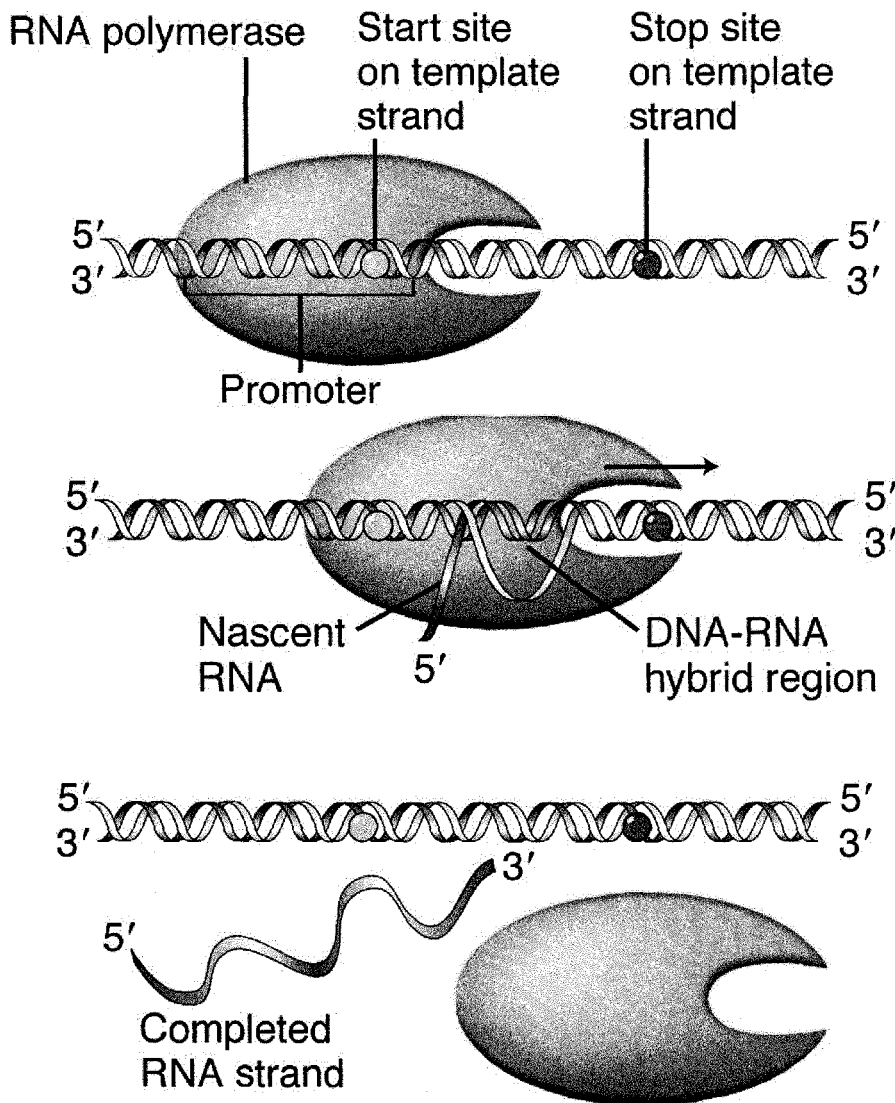


Figura 4.10: Etapas de la transcripción [76]

- En primer lugar, la RNA polimerasa reconoce y enlaza reversiblemente a un lugar específico al comienzo del gen, llamado *promotor*, representado por la subcadena $\langle prom \rangle$.

Las reglas de reescritura (4.9) describen los procesos de reconocimiento del promotor por la RNAP, el enlace de la RNAP al promotor y la rotura del enlace del RNAP del promotor.

$$\begin{aligned}
r_{rb} : [\text{RNAP} + \langle \text{prom} \rangle]_l &\xrightarrow{c_{rb}} [\langle \text{prom.RNAP} \rangle]_l \\
\text{prop}(r_{rb}) &= c_{on} | \text{RNAP} | | \langle \text{prom} \rangle | \\
r_{rd} : [\langle \text{prom.RNAP} \rangle]_l &\xrightarrow{c_{rd}} r [\text{RNAP} + \langle \text{prom} \rangle]_l \\
\text{prop}(r_{rd}) &= c_{rd} | \langle \text{prom.RNAP} \rangle |
\end{aligned} \tag{4.9}$$

- Durante la **iniciación de la transcripción**, la RNA polimerasa disuelve las hebras de DNA a fin de disponer de las bases de la cadena a copiar. Después de que varios ribonucleótidos han sido polimerizados, la RNA polimerasa se disocia del promotor que vuelve a estar disponible para otras polimerasas.

Este proceso se describe mediante reglas de reescritura, r_{ti} , sobre multiconjuntos de cadenas (4.10).

$$\begin{aligned}
r_{ti} : [\langle \text{RNAP.site}_{ini} \rangle]_l &\xrightarrow{c_{ti}} [\langle \overline{\text{site}_{ini}}.\overline{\text{site}_{ini}}.\text{RNAP} \rangle]_l \\
\text{prop}(r_{ti}) &= c_{ini} | \langle \text{RNAP.site}_{ini} \rangle |
\end{aligned} \tag{4.10}$$

- Durante la etapa de **elongación de la hebra**, la RNA polimerasa se mueve a lo largo de la hebra a copiar añadiendo nucleótidos a la cadena de RNA que se va creando. La regla de reescritura r_{el} (4.11) describe el proceso de elongación del mRNA.

$$\begin{aligned}
r_{el} : [\langle \overline{\text{site}_{ini}}.w.\text{RNAP.site}_{mid} \rangle]_l &\xrightarrow{c_{el}} [\langle \text{site}_{mid}.\overline{\text{site}_{ini}}.w.\overline{\text{site}_{mid}}\text{RNAP} \rangle]_l \\
\text{prop}(r_{el}) &= c_{el} | \langle \overline{\text{site}_{ini}}.w.\text{RNAP.site}_{mid} \rangle |
\end{aligned} \tag{4.11}$$

La constante c_{el} se usa para calcular la propensidad de la regla r_{el} . Esa constante describe el número medio de veces que la RNA polimerasa transcribe un lugar de longitud site_{mid} .

- El último paso en la síntesis del RNA es la **finalización de la transcripción**, cuando una molécula completa de RNA es liberada de la RNA polimerasa y la polimerasa se disocia de la molécula de DNA que se estaba copiando. La regla r_{ter} (4.12) describe dicho paso y la constante c_{ter} se usa para calcular su propensidad.

$$\begin{aligned}
r_{ter} : [\langle \overline{\text{site}_{ini}}.w.\text{RNAP.site}_{ter} \rangle]_l &\xrightarrow{c_{ter}} [\text{RNAP} + \langle \text{site}_{ter} \rangle; \langle \overline{\text{site}_{ini}}.w.\overline{\text{site}_{ter}} \rangle]_l \\
\text{prop}(r_{ter}) &= c_{ter} | \langle \overline{\text{site}_{ini}}.w.\text{RNAP.site}_{ter} \rangle |
\end{aligned} \tag{4.12}$$

■ **Traducción:**

Es el proceso por el cual la secuencia de nucleótidos de un mRNA se usa para ordenar y unir aminoácidos en una cadena polipéptida. Esta sucesión de aminoácidos determina la estructura tridimensional de la cadena polipéptida y su actividad. Por tanto, la traducción es fundamental en la producción de proteínas funcionales.

Análogamente a lo que sucedía con la transcripción, los procesos de traducción se pueden estructurar en tres etapas: inicio, elongación y finalización (4.11).

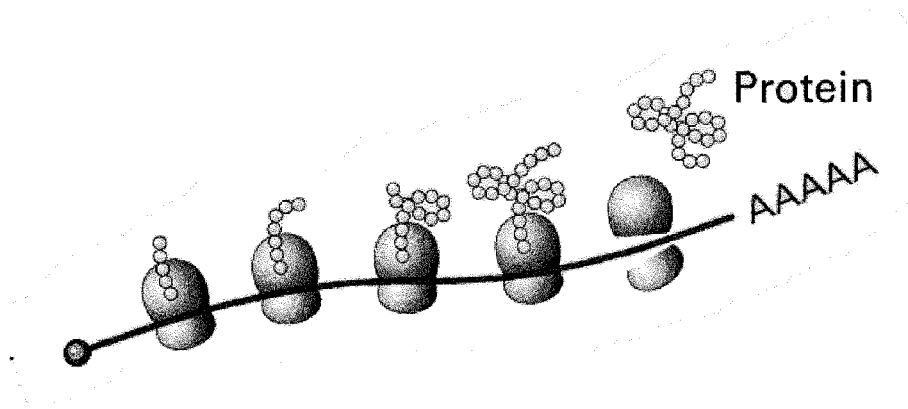


Figura 4.11: Etapas de la traducción [76]

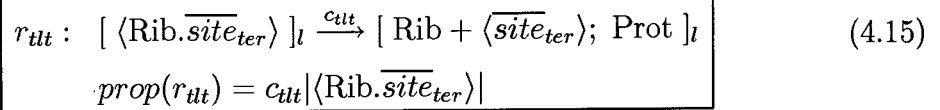
- En nuestra especificación de sistema P para la etapa de iniciación, los codones de iniciación se describen usando la subcadena $\langle \overline{site}_{ini} \rangle$ en la cadena que describe el mRNA y los ribosomas son representados por el objeto Rib. La regla r_{tli} describe el enlace de un ribosoma, Rib, al codón de iniciación, $\langle \overline{site}_{ini} \rangle$.

$$\begin{aligned}
 r_{tli} : [\text{Rib} + \langle \overline{site}_{ini} \rangle]_l &\xrightarrow{c_{tli}} [\langle \overline{site}_{ini} \cdot \text{Rib} \rangle]_l \\
 prop(r_{tli}) &= c_{tli} | \text{Rib} | | \langle \overline{site}_{ini} \rangle |
 \end{aligned}
 \tag{4.13}$$

- Los ribosomas dirigen la elongación de un polipéptido moviéndose a lo largo de una cadena de mRNA e interactuando con varias proteínas. La regla (4.14) describe un paso de la elongación. La constante c_{tle} asociada a esa regla representa el número medio de lugares representados por \overline{site}_{mid} que son traducidos por unidad de tiempo.

$$\begin{aligned}
 r_{tle} : [\langle \text{Rib} \cdot \overline{site}_{mid} \rangle]_l &\xrightarrow{c_{tle}} [\langle \overline{site}_{mid} \cdot \text{Rib} \rangle]_l \\
 prop(r_{tle}) &= c_{tle} | \langle \text{Rib} \cdot \overline{site}_{mid} \rangle |
 \end{aligned}
 \tag{4.14}$$

- La etapa final en la producción de una cadena de polipéptido consiste en la finalización. Los ribosomas se disocian de un mRNA y liberan una cadena completa cuando alcanzan determinados codones que hacen de punteros finales. Este proceso se describe mediante la regla r_{tlt} (4.15).



4.4. Modelos de Sistemas P y SBLM (Systems Biology Markup Language)

La diversidad de aproximaciones y métodos en el modelado celular demanda un formato común, una especie de *lingua franca*, que permita la comunicación entre los aspectos más importantes de un modelo. El lenguaje SBLM (Systems Biology Markup Language) fue desarrollado como consecuencia de esta demanda. Se trata de un formato que permite representar redes bioquímicas y es, básicamente, una variante de un *eXtensible Markup Language* (XML) que codifica las principales componentes de las redes bioquímicas. Un modelo SBML consta de una lista de funciones, unidades, compartimentos, especies, parámetros, reglas, reacciones y eventos. Esta es la estructura básica de una especificación SBML Level 2 (versión 1):

```
<?xml version="3.0" encoding="UTF-8"?>
<sbml xmlns="http://www.sbml.org/sbml/level2" level="2" version="1">
  <model id=" ">
    <listOfFunctionDefinitions>
      ...
    </listOfFunctionDefinitions>
    <listOfUnitDefinitions>
      ...
    </listOfUnitDefinitions>
    <listOfCompartments>
      ...
    </listOfCompartments>
    <listOfSpecies>
      ...
    </listOfSpecies>
    <listOfParameters>
      ...
    </listOfParameters>
    <listOfRules>
      ...
    </listOfRules>
    <listOfReactions>
      ...
    </listOfReactions>
    <listOfEvents>
      ...
    </listOfEvents>
  </model>
</sbml>
```

Aquí sólo se describen las listas de unidades, compartimentos, especies parámetros y reacciones. Para una descripción completa de la versión 1 del SBML Level 2, ver [135].

- La *lista de unidades* permite la definición y redefinición de las unidades usadas en este modelo. Los modelos mesoscópicos usan normalmente el número de moléculas como unidades mientras que los modelos macroscópicos trabajan con unidades de concentración.
- La *lista de compartimentos* enumera los compartimentos de un modelo., el cual debe tener, al menos un compartimento y cada uno de ellos tendrá un identificador, *id*.
- La *lista de especies* establece cuáles son las especies moleculares del modelo, cada una de las cuales deben estar convenientemente identificadas mediante un *id*, pudiendo especificarse también un compartimento en donde las especies están situadas, la cantidad inicial de concentración, o bien otros atributos.
- La *lista de parámetros* contiene la definición de las constantes que son usadas en el modelo. También es posible declarar parámetros locales en la ley cinética de reacción.
- La *lista de reacciones* consta de una enumeración de las reacciones del sistema. Una reacción consiste de una lista de reactantes, una lista de productos y una ley cinética. La lista de reactantes y la lista de productos contienen los identificadores de las especies reactantes y productos de la reacción.

El lenguaje SBML no debería ser mirado como una alternativa a otras representaciones sino, más bien, como un formato electrónico que se puede usar en conjunción con las especificaciones de cualquier otro modelo.

4.4.1. Traduciendo Formatos SBML en Modelos de Sistemas P

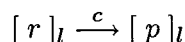
Todas las componentes de un fichero SBML Level 2, versión 1, no pueden ser traducidas a un modelo de sistema P. En esta sección sólo analizaremos cómo se traducen las *unidades*, *especies*, *compartimentos* y *reacciones*.

- Las **unidades** constituyen la primera sección que debe ser analizada cuando se traduce un fichero SBML en un modelo de sistema P. Para ello, es necesario chequear las unidades de las cantidades iniciales de las especies y parámetros.
- Los **compartimentos** se consideran en el formato SBML como regiones acotadas en las que las especies están localizadas. En los modelos de sistemas P, las membranas no se corresponden necesariamente con las membranas de células reales delimitadas por compartimentos.
- Las **especies** son tratadas en el formato SBML como entidades bioquímicas simples e indivisibles, lo que permite su especificación como objetos en el alfabeto del modelo de sistema P.
- Las **reacciones** en el formato SBML se describen en términos de reactantes, productos y una ley cinética opcional. SBML no impone ninguna restricción adicional al número de reactantes y productos. La ley cinética asociada a una reacción puede ser una expresión matemática compleja especificada en MathML, sin detallar si las tasas son deterministas o estocásticas.

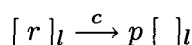
Los siguientes pasos muestran cómo traducir las reacciones de un fichero SBML en reglas de sistemas P, distinguiendo entre reacciones de primer o segundo orden.

1. **Reacciones de primer orden:** Las reacciones con un único reactante, r , pueden ser descritas usando diferentes tipos de reglas, en función del número y localización de los productos.

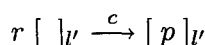
- Reacciones con un único producto p localizado en una membrana l (*regla de transformación*):



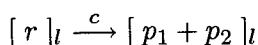
- Reacciones con un único producto p localizado en una membrana l' fuera de la membrana l (*regla de difusión hacia fuera*):



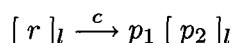
- Reacciones con un único producto p localizado en una membrana situada dentro de una membrana l , identificada por la etiqueta l' (*regla de difusión hacia dentro*):



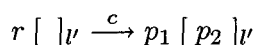
- Reacciones con dos productos, p_1 y p_2 , localizados en la membrana l (*regla de disociación de complejos*):



- Reacciones con dos productos, p_1 y p_2 , con p_1 localizado en la membrana l' situada fuera de la membrana l y p_2 localizado en la membrana l (*regla de rotura de enlace*):

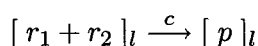


- Reacciones con dos productos, p_1 y p_2 , con p_1 localizado en una membrana l y p_2 localizado en una membrana situada dentro de otra etiquetada por l' . (*regla de liberación*):

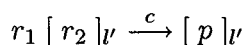


2. **Reacciones de segundo orden:** Las reacciones con dos reactantes, r_1 y r_2 , y un único producto p pueden ser descritas usando diferentes tipos de reglas que dependen de la localización de r_1 , r_2 y de p .

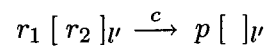
- Reacciones con los reactantes r_1 y r_2 localizados en la misma membrana l y con un único producto p localizado, también, en la misma membrana l (*regla de formación de complejos*):



- Reacciones con el reactante r_1 localizado en una membrana l y tanto el otro reactante r_2 como el producto p están localizados en la membrana l' inmediatamente exterior a ella (*regla de enlace*):



- Reacciones con el reactante r_1 y el producto p localizados en una membrana l que es inmediatamente exterior a la membrana l' donde está situado el otro reactante r_2 (*regla de reclutamiento*):



Capítulo 5

Modularización en P sistemas

5.1. Módulos en biología celular y P sistemas

Las funciones celulares se llevan a cabo rara vez por moléculas individuales, la mayoría de las funciones celulares surgen como un comportamiento emergente de las interacciones entre módulos formados por muchas especies moleculares. Aquí nos referiremos a módulo para designar una entidad discreta que realiza una función específica separable de las funciones del resto de los módulos.

Esta separación surge del aislamiento físico de especies moleculares en distintos compartimentos y de la especificidad en las interacciones moleculares. Ambos casos son fácilmente representables en P sistemas utilizando membranas para el aislamiento físico y reglas de reescritura para la especificidad molecular.

Definición 5.1.1 *Un módulo de P sistema es un conjunto de reglas de P sistemas que representan las interacciones moleculares que ocurren repetidamente en muchos sistemas celulares. Un módulo de P sistema se identifica con un nombre y tres conjuntos de variables V , C y Lab . V representa las variables que se pueden instanciar usando los objetos o cadenas que describen las entidades moleculares involucradas en las interacciones que forman el módulo. C representa las constantes estocásticas asociadas a cada regla. Lab especifica las etiquetas de los compartimentos involucrados en las reglas. Formalmente, un módulo, mod , con variables V , constantes C and etiquetas L se escribe como $mod(V, C, L)$. Módulos más complejos se pueden construir a partir de módulos más simples mediante simple unión de conjuntos.*

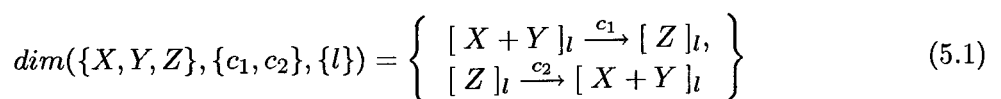
5.2. Una librería de módulos de P sistemas básicos

A continuación presentamos una serie de módulos de P sistemas que describen los módulos moleculares más básicos:

5.2.1. Módulos de interacciones proteínicas

Primero nos centraremos en módulos formados por interacciones proteínicas.

- **Módulo de dimerización:**



- **Módulo de interacción enzimática :**

$$enz(\{X, Y, Z, W\}, \{c_1, c_2, c_3\}, \{l\}) = \left\{ \begin{array}{l} [X + Y]_l \xrightarrow{c_1} [Z]_l, \\ [Z]_l \xrightarrow{c_2} [X + Y]_l, \\ [Z]_l \xrightarrow{c_2} [X + W]_l \end{array} \right\} \quad (5.2)$$

Una simplificación del módulo anterior la constituye el modulo de producción:

$$prod(\{X, Z\}, \{c_1\}, \{l\}) = \{ [X]_l \xrightarrow{c_1} [X + Z]_l \} \quad (5.3)$$

- **Módulo de degradación:**

$$deg(\{X\}, \{c_1\}, \{l\}) = \{ [X]_l \xrightarrow{c_1} []_l \}$$

- **Módulo de difusión:**

$$\begin{aligned} diff_{in}(\{X\}, \{c\}, \{l\}) &= \{ X []_l \xrightarrow{c} [X]_l \} \\ diff_{out}(\{X\}, \{c\}, \{l\}) &= \{ [X]_l \xrightarrow{c} X []_l \} \end{aligned}$$

5.2.2. Módulos de regulación de genes

Aquí presentamos tres módulos de P sistemas representando los tres mecanismos básicos en la regulación de genes. Expresión constitutiva o independiente de factores de transcripción, regulación positiva y regulación negativa.

- **Módulo de expresión constitutiva:**

$$const(\{G, R, P\}, \{c_1, \dots, c_4\}, \{l\}) = \left\{ \begin{array}{l} prod(\{G, R\}, \{c_1\}, \{l\}) \quad \cup \\ prod(\{R, P\}, \{c_2\}, \{l\}) \quad \cup \\ deg(\{R\}, \{c_3\}, \{l\}) \quad \cup \\ deg(\{P\}, \{c_4\}, \{l\}) \end{array} \right\} \quad (5.4)$$

- **Módulo de regulación positiva :**

$$pos(\{A, G, G_{on}, R, P\}, \{c_1, \dots, c_6\}, \{l\}) = \left\{ \begin{array}{l} dim(\{A, G, G_{on}\}, \{c_1, c_2\}, \{l\}) \quad \cup \\ prod(\{G_{on}, R\}, \{c_3\}, \{l\}) \quad \cup \\ prod(\{R, P\}, \{c_4\}, \{l\}) \quad \cup \\ deg(\{R\}, \{c_5\}, \{l\}) \quad \cup \\ deg(\{P\}, \{c_6\}, \{l\}) \end{array} \right\} \quad (5.5)$$

- **Módulo de regulación negativa:**

$$\begin{aligned} lneg(\{Rep, G, G_{off}, R, P\}, \{c_1, \dots, c_6\}, \{l\}) &= \\ &= \left\{ \begin{array}{l} dim(\{Rep, G, G_{off}\}, \{c_1, c_2\}, \{l\}) \quad \cup \\ prod(\{G, R\}, \{c_3\}, \{l\}) \quad \cup \\ prod(\{R, P\}, \{c_4\}, \{l\}) \quad \cup \\ deg(\{R\}, \{c_5\}, \{l\}) \quad \cup \\ deg(\{P\}, \{c_6\}, \{l\}) \end{array} \right\} \end{aligned} \quad (5.6)$$

5.3. Modelado de módulos celulares mediante la combinación de módulos de P sistemas

En esta sección ilustramos como módulos celulares complejos se pueden construir a partir de los módulos de P sistemas básicos introducidos anteriormente. Utilizaremos los módulos celulares de autorregulación para ejemplificar nuestro método.

5.3.1. Autoregulation in Transcription Networks

- **Autorregulación negativa:**

La autorregulación negativa consiste en la regulación negativa de un gen por la propia proteína que codifica. Se puede obtener un módulo de P sistema para la autorregulación negativa utilizando el módulo *neg*:

$$NAR(\{G, G_{off}, R, P\}, \{c_1, \dots, c_6\}, \{l\}) = neg(\{P, G, G_{off}, R, P\}, \{c_1, \dots, c_6\}, \{l\}) \quad (5.7)$$

La propiedad principal de la autorregulación es que acelera el tiempo de respuesta de un gen, es decir, es tiempo necesario para alcanzar la mitad de su máxima expresión.

- **Autorregulación positiva:**

La autorregulación positiva se da cuando un gen está regulado positivamente por la proteína que él mismo codifica. De la misma manera que el caso anterior podemos diseñar un módulo de P sistema para este módulo celular utilizando el módulo previamente definido *pos*.

$$PAR(\{G, G_{on}, R, P\}, \{c_1, \dots, c_6\}, \{l\}) = pos(\{P, G, G_{on}, R, P\}, \{c_1, \dots, c_6\}, \{l\}) \quad (5.8)$$

Opuestamente al caso de autorregulación negativa, la autorregulación positiva retrasa el tiempo de respuesta.

Parte III

Casos de estudio

Capítulo 6

Modelado de la Regulación de Genes en Procaiotas

La estructura y funcionalidad de las células vivas están determinadas por un conjunto específico de proteínas que ellas mismas contienen, de manera codificada, en su genoma. El control de la expresión de genes determina qué genes serán expresados y, por tanto, pueden generar las proteínas que codifican. Por ello, el proceso de control de la expresión de genes es un aspecto fundamental en Biología molecular y es objeto de intensivas investigaciones en el marco de la Biología de Sistemas.

En muchas bacterias, o células procariotas, la expresión de genes está altamente regulada con el fin de producir las proteínas necesarias para responder a las alteraciones en el entorno. Por tanto, en un instante determinado una bacteria sintetiza únicamente aquellas proteínas que son necesarias para la supervivencia en ese momento.

En este capítulo se describen los principios básicos para la regulación de genes en células procariotas usando el marco de los modelos de sistemas P. Como un caso de estudio, analizaremos el sistema de regulación del Lac Operon que nos permitirá ilustrar los principios generales presentados en el capítulo 3.

6.1. Control de Expresión de Genes en el Lac Operon

La forma más directa de controlar la expresión de un gen es mediante el cálculo de su tasa de transcripción; es decir, la tasa en la que la RNA polimerasa transcribe el gen en moléculas de RNA mensajero (mRNA).

Añadiendo un nuevo sustrato al medio, es posible inducir la formación de nuevas enzimas capaces de metabolizar ese sustrato. Un ejemplo de este fenómeno sucede cuando se considera un cultivo de *Escherichia coli* (*E. coli*) que se alimenta de glucosa y transfiere al medio algunas sustancias que contienen lactosa.

Al principio las células están en estado quiescente; es decir, no metabolizan lactosa, decae su actividad metabólica y cesa la división celular. Sin embargo, el cultivo comienza a crecer repentinamente consumiendo lactosa. Durante el intervalo quiescente, las células llegaron a producir tres enzimas que no habían sido elaboradas con anterioridad: una permeasa, LacY, que transporta lactosa desde el medio al interior de la célula a través de la membrana plasmática; β -galactosidasa que hidroliza la lactosa en glucosa y galactosa; y una transacetilasa, LacA, cuya función aún no se conoce con exactitud.

Los genes que codifican estas proteínas involucradas en la recepción y consumo de lactosa, están localizados en una región del genoma de la *E. coli* genome llamado **Lac Operon**.

El gen *lacI* codifica una proteína denominada LacI que actúa como un represor, compuesto de

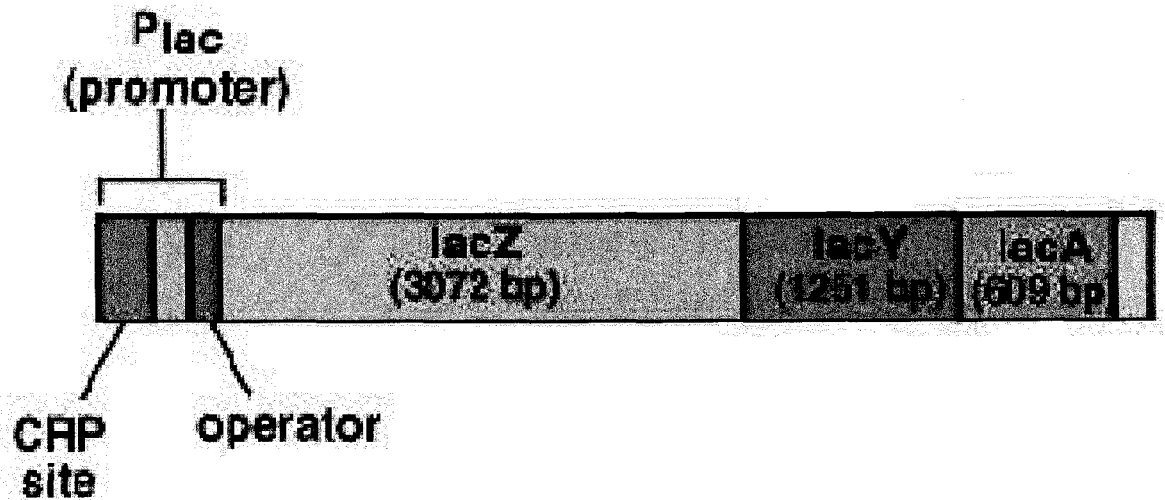


Figura 6.1: Lac Operon

cuatro péptidos idénticos. Una parte de esta molécula tiene un lugar que es capaz de reconocer y enlazarse a 24 pares de bases del Lac promotor, llamado Lac Operator, impidiendo a la RNA polimerasa la transcripción de los genes *lacZ*, *lacY* and *lacA* que codifican la β -galactosidasa, la permeasa y la transacetilasa, respectivamente.

El operon lactosa permite a la vez un control positivo y negativo en el sentido siguiente: el *lac* represor, *LacI*, regula negativamente la expresión mientras que el activador, *cAMP-CRP*₂, activa positivamente la expresión.

Existen cuatro estados básicos de expresión del Lac Operon:operon:

- NO Glucosa y NO Lactosa.

En estas condiciones no se producirá la transcripción del Lac Operon

- Presente la Glucosa pero NO la Lactosa.

En estas condiciones no se producirá la transcripción del Lac Operon

- Presente la Glucosa y la Lactosa.

En estas condiciones se producirá la transcripción del Lac Operon pero a bajo nivel.

- NO Glucosa pero abundante Lactosa.

En estas condiciones se producirá la transcripción del Lac Operon a máximo nivel.

La presencia de dos sistemas separados de control permite a la célula responder a las necesidades impuestas por los cambios en las condiciones de crecimiento.

6.2. Modelado de la Regulación del Lac Operon usando Sistemas P

En esta sección se desarrolla una especificación y una familia de modelos de sistemas P para el sistema de regulación del Lac Operon. Nuestros modelos ilustrarán los principios de especificación discutidos en el capítulo 3.

En los modelos propuestos, estudiaremos el comportamiento del sistema para diferentes condiciones iniciales en el cultivo (sin o con glucosa y sin o con lactosa).

6.2.1. Una Especificación de Sistema P del Sistema Lac Operon

A continuación se presenta una descripción detallada de una especificación de sistema P para el sistema de regulación del Lac Operon.

■ Especificación de las regiones relevantes en el sistema Lac Operon:

En una bacteria *E. coli* existen dos regiones relevantes:

- La *superficie celular de la bacteria* que contiene un conjunto de proteínas que controlan el movimiento de moléculas y detectan señales.
- El *lumen de la bacteria* o interior acuoso de la bacteria, en donde existen una serie de proteínas complementarias que están involucradas en procesos celulares específicos.

De acuerdo con los principios de especificación presentados en el capítulo 3, se considerarán tres membranas diferentes para especificar una bacteria *E. coli*.

1. La primera, identificada por la etiqueta s , representa la superficie celular de la bacteria que tendrá asociada un multiconjunto de objetos que describen las entidades moleculares asociadas con la membrana plasmática, y unas reglas de reescritura que especificarán procesos moleculares.
2. La segunda membrana, identificada por la etiqueta c , describirá el medio acuoso del interior de la bacteria y, por tanto, está situada dentro de la anterior membrana. Esta membrana tendrá asociada un multiconjunto de objetos y cadenas que describen las entidades moleculares situadas en el lumen de la bacteria, y unas reglas de reescritura que especificarán las interacciones moleculares que tienen lugar en la región determinada por esa membrana.
3. La tercera membrana, identificada por la etiqueta e , es el cultivo que provocará un comportamiento distinto de la bacteria, según las fuentes de azúcar (glucosa y/o lactosa).

■ Especificación de las entidades moleculares del sistema Lac Operon:

Entre las entidades moleculares involucradas en el sistema Lac Operon, existen dos conjuntos significativamente diferente de moléculas:

- Por una parte, tenemos proteínas y enzimas involucradas en la selección de azúcar del cultivo (por ejemplo, EIICB y la enzima β -galactosidasa). Estas moléculas pueden ser consideradas como entidades individuales sin estructura interna.

Los objetos usados en esta memoria para representar las proteínas y enzimas involucradas en el Lac Operon determinan el alfabeto O_{prot} .

$$O_{prot} = \{ EIIA, EIIA \sim P, EIICB, EIICB \sim P, Gluc, Gluc \sim P, EIICB \sim P - -Gluc, Lact, LacY, Lact - LacY, AC, AC - -EIIA, AC - -EIIA \sim P, ATP, AC - -EIIA \sim P - -ATP, LacY - EIIA, \beta\text{-Galac}, Lact, \beta - -Galac - -Lact, Allolact, LacI, LacI - -Allact, CRP, CRP - -cAMP, CRP - -cAMP_2, RNAP \}$$

- Por otra parte, tenemos los genes que codifican la información necesaria para la síntesis de algunas proteínas y enzimas descritas en el alfabeto O_{prot} . La estructura lineal de estas entidades moleculares, genes y mRNAs, será especificadas usando cadenas. Los lugares relevantes en el Lac Operon y en el correspondiente mRNA serán representados usando símbolos de los alfabetos O_{dna} y O_{rna} , respectivamente. Una descripción detallada de los lugares representados por los símbolos de O_{dna} y O_{rna} son presentados en la figura 6.2.

$$O_{dna} = \{ \text{cap}, \text{cap}^{CRP-cAMP_2}, \text{op}, \text{op}^{LacI}, \text{lacZ}_s, \text{lacZ}_m, \text{lacZ}_e, \text{lacY}_s, \text{lacY}_m, \text{lacY}_e, \text{lacA}_s, \text{lacA}_m, \text{lacA}_e \}$$

$$O_{rna} = \{ \text{Rib}, \overline{\text{op}}, \overline{\text{lacZ}}_s, \overline{\text{lacZ}}_m, \overline{\text{lacZ}}_e, \overline{\text{lacY}}_s, \overline{\text{lacY}}_m, \overline{\text{lacY}}_e, \overline{\text{lacA}}_s, \overline{\text{lacA}}_m, \overline{\text{lacA}}_e \}$$

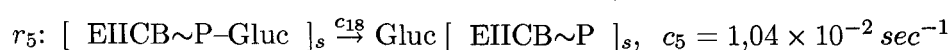
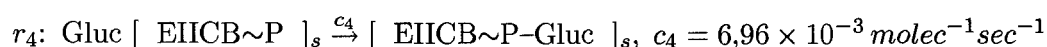
Symbol	Site
cap	Free CAP site where the activator CRP-cAMP ₂ binds
cap ^{CRP-cAMP₂}	CAP site occupied by the activator CRP-cAMP ₂ binds
op	Free operator site where the repressor LacI binds
op ^{LacI}	Operator site occupied by the repressor LacI binds
lacZ _s lacY _s lacA _s	Sites marking the start point of the lacZ, lacY and lacA gene respectively
lacZ _m lacY _m lacA _m	Sites located in the middle of the lacZ, lacY and lacA gene respectively
lacZ _e lacY _e	Sites located in the final point of the lacZ and lacY gene respectively.
lacA _e	Site marking the end of gene lacA which coincides with the transcription termination site of the lac operon
$\overline{\text{op}}$	Site marking the starting point of the mRNA transcript
$\overline{\text{lacZ}}_s$ $\overline{\text{lacZ}}_e$ $\overline{\text{lacY}}_s$ $\overline{\text{lacY}}_e$ $\overline{\text{lacA}}_s$ $\overline{\text{lacA}}_e$	Sites marking the beginning and end of the reading frames in the mRNA for the genes lacZ, lacY and lacA respectively
$\overline{\text{lacZ}}_m$ $\overline{\text{lacY}}_m$ $\overline{\text{lacA}}_m$	Sites located in the middle of reading frames of lacZ, lacY and lacA gene respectively

Figura 6.2:

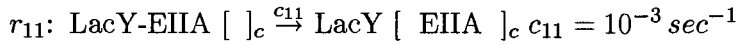
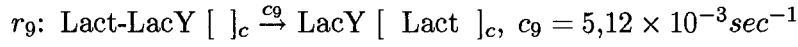
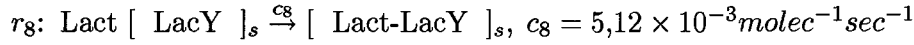
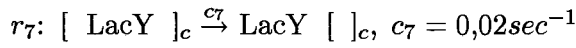
■ Especificación de las interacciones moleculares en el sistema Lac Operon:

A continuación se presenta una descripción detallada de las reglas de reescritura que especifican los procesos moleculares en el sistema Lac Operon.

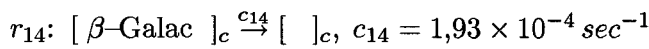
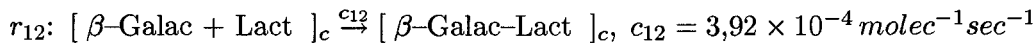
– Reglas que describen el sistema de transporte de la glucosa.



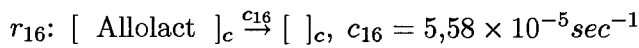
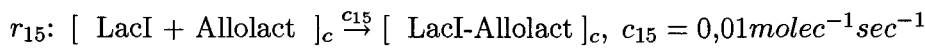
– Reglas que describen el sistema de transporte de la lactosa.



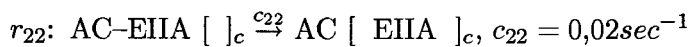
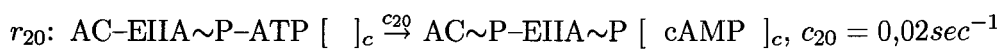
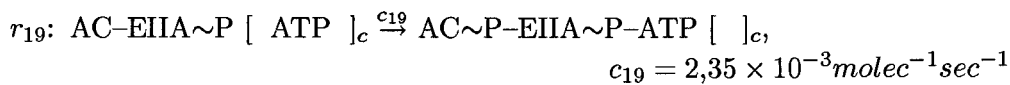
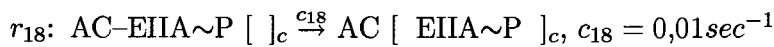
– Reglas que describen la actividad de la β -Galactosidasa.



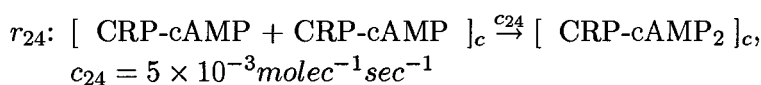
– Reglas que describen la actividad de la alolactosa.



– Reglas que describen la actividad de la Adenilata Ciclasa.



– Reglas que describen la formación del activador CRP-cAMP₂



– Reglas que describen la activación y represión del Lac Operon.

$$r_{25}: [\text{LacI} + \langle \text{op} \rangle]_c \xrightarrow{c_{25}} [\langle \text{op}^{\text{LacI}} \rangle]_c, \quad c_{25} = 0,2 \text{molec}^{-1} \text{sec}^{-1}$$

$$r_{26}: [\langle \text{op}^{\text{LacI}} \rangle]_c \xrightarrow{c_{26}} [\text{LacI} + \langle \text{op} \rangle]_c, \quad c_{26} = 5 \text{sec}^{-1}$$

$$r_{27}: [\text{CRP-cAMP}_2 + \langle \text{cap} \rangle]_c \xrightarrow{c_{27}} [\langle \text{cap}^{\text{CRP-cAMP}_2} \rangle]_c, \\ c_{27} = 0,01 \text{molec}^{-1} \text{sec}^{-1}$$

$$r_{28}: [\langle \text{cap}^{\text{CRP-cAMP}_2} \rangle]_c \xrightarrow{c_{28}} [\text{CRP-cAMP}_2 + \langle \text{cap} \rangle]_c, \quad c_{28} = 5 \text{sec}^{-1}$$

– Reglas que describen la iniciación de la transcripción en el Lac Operon.

$$r_{29}: [\text{RNAP} + \langle \text{cap} \rangle]_c \xrightarrow{c_{29}} [\langle \text{cap.RNAP} \rangle]_c, \quad c_{29} = 5 \times 10^{-4} \text{molec}^{-1} \text{sec}^{-1}$$

$$r_{30}: [\text{RNAP} + \langle \text{cap}^{\text{CRP-cAMP}_2} \rangle]_c \xrightarrow{c_{30}} [\langle \text{cap}^{\text{CRP-cAMP}_2} . \text{RNAP} \rangle]_c, \\ c_{30} = 0,02 \text{molec}^{-1} \text{sec}^{-1}$$

$$r_{31}: [\langle \text{RNAP.op} \rangle]_c \xrightarrow{c_{31}} [\langle \text{op. } \overline{\text{op}} . \text{RNAP} \rangle]_c, \quad c_{31} = 2 \text{sec}^{-1}$$

– Reglas que describen la elongación del mRNA en el Lac Operon.

$$r_{32}: [\langle \overline{\text{op}} . \text{RNAP} . \text{lacZ}_s \rangle]_c \xrightarrow{c_{32}} [\langle \text{lacZ}_s . \overline{\text{op}} . \overline{\text{lacZ}}_s . \text{RNAP} \rangle]_c, \quad c_{32} = 2 \text{sec}^{-1}$$

$$r_{33}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacZ}_m \rangle]_c \xrightarrow{c_{33}} [\langle \text{lacZ}_m . \overline{\text{op}} . w . \overline{\text{lacZ}}_m . \text{RNAP} \rangle]_c, \\ c_{33} = 2 \text{sec}^{-1}$$

$$r_{34}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacZ}_e \rangle]_c \xrightarrow{c_{34}} [\langle \text{lacZ}_e . \overline{\text{op}} . w . \overline{\text{lacZ}}_e . \text{RNAP} \rangle]_c, \\ c_{34} = 2 \text{sec}^{-1}$$

$$r_{35}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacY}_s \rangle]_c \xrightarrow{c_{35}} [\langle \text{lacY}_s . \overline{\text{op}} . w . \overline{\text{lacY}}_s . \text{RNAP} \rangle]_c, \\ c_{35} = 2 \text{sec}^{-1}$$

$$r_{36}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacY}_m \rangle]_c \xrightarrow{c_{36}} [\langle \text{lacY}_m . \overline{\text{op}} . w . \overline{\text{lacY}}_m . \text{RNAP} \rangle]_c, \\ c_{36} = 2 \text{sec}^{-1}$$

$$r_{37}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacY}_e \rangle]_c \xrightarrow{c_{37}} [\langle \text{lacY}_e . \overline{\text{op}} . w . \overline{\text{lacY}}_e . \text{RNAP} \rangle]_c, \\ c_{37} = 2 \text{sec}^{-1}$$

$$r_{38}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacA}_s \rangle]_c \xrightarrow{c_{38}} [\langle \text{lacA}_s . \overline{\text{op}} . w . \overline{\text{lacA}}_s . \text{RNAP} \rangle]_c, \\ c_{38} = 2 \text{sec}^{-1}$$

$$r_{39}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacA}_m \rangle]_c \xrightarrow{c_{39}} [\langle \text{lacA}_m . \overline{\text{op}} . w . \overline{\text{lacA}}_m . \text{RNAP} \rangle]_c, \\ c_{39} = 2 \text{sec}^{-1}$$

– Regla que describe el final de la transcripción en el Lac Operon

$$r_{40}: [\langle \overline{\text{op}} . w . \text{RNAP} . \text{lacA}_e \rangle]_c \xrightarrow{c_{40}} [\text{RNAP} + \langle \text{lacA}_e \rangle ; \langle \overline{\text{op}} . w . \overline{\text{lacA}}_e \rangle]_c, \\ c_{40} = 2 \text{sec}^{-1}$$

– Reglas que describen el inicio de la traducción.

$$r_{41}: [\text{Rib} + \langle \overline{\text{lacZ}}_s \rangle]_c \xrightarrow{c_{41}} [\langle \text{Rib} . \overline{\text{lacZ}}_s \rangle]_c, \quad c_{41} = 0,16 \text{molec}^{-1} \text{sec}^{-1}$$

$$r_{42}: [\text{Rib} + \langle \overline{\text{lacY}}_s \rangle]_c \xrightarrow{c_{42}} [\langle \text{Rib} . \overline{\text{lacY}}_s \rangle]_c, \quad c_{42} = 0,16 \text{molec}^{-1} \text{sec}^{-1}$$

$$r_{43}: [\text{Rib} + \langle \overline{\text{lacA}}_s \rangle]_c \xrightarrow{c_{43}} [\langle \text{Rib} . \overline{\text{lacA}}_s \rangle]_c, \quad c_{43} = 0,16 \text{molec}^{-1} \text{sec}^{-1}$$

– Reglas que describen la traducción y reglas de disociación del ribosoma.

$$\begin{aligned}
r_{44}: [\langle \text{Rib}.\overline{\text{lacZ}}_s \rangle]_c &\xrightarrow{c_{44}} [\langle \overline{\text{lacZ}}_s.\text{Rib} \rangle]_c, \quad c_{44} = 0,3\text{sec}^{-1} \\
r_{45}: [\langle \text{Rib}.\overline{\text{lacZ}}_m \rangle]_c &\xrightarrow{c_{45}} [\langle \overline{\text{lacZ}}_m.\text{Rib} \rangle]_c, \quad c_{45} = 0,3\text{sec}^{-1} \\
r_{46}: [\langle \text{Rib}.\overline{\text{lacZ}}_e \rangle]_c &\xrightarrow{c_{46}} [\beta\text{-Galac} + \text{Rib} + \langle \overline{\text{lacZ}}_e \rangle]_c, \quad c_{46} = 0,3\text{sec}^{-1} \\
r_{47}: [\langle \text{Rib}.\overline{\text{lacY}}_s \rangle]_c &\xrightarrow{c_{47}} [\langle \overline{\text{lacY}}_s.\text{Rib} \rangle]_c, \quad c_{47} = 0,3\text{sec}^{-1} \\
r_{48}: [\langle \text{Rib}.\overline{\text{lacY}}_m \rangle]_c &\xrightarrow{c_{48}} [\langle \overline{\text{lacY}}_m.\text{Rib} \rangle]_c, \quad c_{48} = 0,3\text{sec}^{-1} \\
r_{49}: [\langle \text{Rib}.\overline{\text{lacY}}_e \rangle]_c &\xrightarrow{c_{49}} [\text{LacY} + \text{Rib} + \langle \overline{\text{lacY}}_e \rangle]_c, \quad c_{49} = 0,3\text{sec}^{-1} \\
r_{50}: [\langle \text{Rib}.\overline{\text{lacA}}_s \rangle]_c &\xrightarrow{c_{50}} [\langle \overline{\text{lacA}}_s.\text{Rib} \rangle]_c, \quad c_{50} = 0,3\text{sec}^{-1} \\
r_{51}: [\langle \text{Rib}.\overline{\text{lacA}}_m \rangle]_c &\xrightarrow{c_{51}} [\langle \overline{\text{lacA}}_m.\text{Rib} \rangle]_c, \quad c_{51} = 0,3\text{sec}^{-1} \\
r_{52}: [\langle \text{Rib}.\overline{\text{lacA}}_e \rangle]_c &\xrightarrow{c_{52}} [\text{LacA} + \text{Rib} + \langle \overline{\text{lacA}}_e \rangle]_c, \quad c_{52} = 0,3\text{sec}^{-1}
\end{aligned}$$

– Reglas que describen la degradación del mRNA.

$$\begin{aligned}
r_{53}: [\langle \overline{\text{lacZ}}_s \rangle]_c &\xrightarrow{c_{53}} []_c, \quad c_{53} = 0,2\text{sec}^{-1} \\
r_{54}: [\langle \overline{\text{lacY}}_s \rangle]_c &\xrightarrow{c_{54}} []_c, \quad c_{54} = 0,2\text{sec}^{-1} \\
r_{55}: [\langle \overline{\text{lacA}}_s \rangle]_c &\xrightarrow{c_{55}} []_c, \quad c_{55} = 0,2\text{sec}^{-1}
\end{aligned}$$

En resumen, nuestra especificación de sistema P del sistema de regulación de genes del Lac Operon es la siguiente tupla:

$$\Pi_{lac} = (O, \{e, s, c\}, [[[]_3]_2]_1, M_1, M_2, M_3, R_e, R_s, R_c) \quad (6.1)$$

en donde::

- $O = O_{prot} \cup O_{dna} \cup O_{rna}$.
- Las etiquetas $\{e, s, c\}$ identifican los tipos de compartimentos definidos por las membranas de Π_{lac} .
- La estructura de membranas consta de tres membranas que definen las regiones relevantes del sistema.
- Los multiconjuntos iniciales M_1, M_2 and M_3 forman parte de los parámetros de nuestra especificación Π_{lac} .
- El conjunto de reglas de reescritura sobre multiconjuntos de objetos y cadenas R_e, R_s y R_c están asociados a los distintos compartimentos que integran el sistema.

$$R_e = \{r_4, r_8\}$$

$$R_s = \{r_2, r_3, r_5, r_6, r_9, r_{11}, r_{18}, r_{20}, r_{22}\}$$

$$R_c = \{r_1, r_7, r_{10}, r_{12}, r_{13}, r_{14}, r_{15}, r_{16}, r_{17}, r_{19}, r_{21}, r_{23}, \dots, r_{55}\}$$

6.2.2. Modelos de Sistemas P del Sistema de Regulación del Lac Operon

En esta sección se introduce una familia de modelos de sistemas P que nos va a permitir estudiar el comportamiento del sistema de regulación del Lac Operon bajo diferentes condiciones iniciales. Para ello, a partir de la especificación descrita, Π_{lac} , es necesario identificar los parámetros asociados a ella, $\mathcal{P}(\Pi_{lac})$, y especificar un conjunto de posibles valores a fin de generar una familia de modelos que procedan a realizar las simulaciones.

De acuerdo con la definición 3.2.2, los parámetros de $\mathcal{P}(\Pi_{lac})$ constan de los multiconjuntos iniciales y las constantes estocásticas asociadas a las reglas de reescritura:

$$\mathcal{P}(\Pi_{lac}) = (\mathcal{M}_0(\Pi_{lac}), \mathcal{C}(\Pi_{lac}))$$

- En nuestro caso, las constantes estocásticas asociadas a las reglas de reescritura $\mathcal{C}(\Pi_{lac})$ tendrán un valor fijo recogido en C y que fueron representadas durante la enumeración de las reglas. Estos valores han sido extraídos de la literatura o fueron calculados de acuerdo con la discusión realizada en la sección ??, a partir de algunas constantes deterministas usadas en modelos basados en ODEs, [58, 66, 69, 77, 106].
- Los posibles valores asociados a los multiconjuntos iniciales $\mathbf{M}_0 = (\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3)$ describirán cuatro diferentes condiciones iniciales en el entorno, $\mathbf{M}_1 = \{M_1^1, M_1^2, M_1^3, M_1^4\}$, y una única condición inicial para la superficie celular $\mathbf{M}_2 = \{M_2\}$ y para el citoplasma $\mathbf{M}_3 = \{M_3\}$. Concretamente son:

- Para el entorno:

- $M_1^1 = (e, \lambda, \lambda)$.
- $M_1^2 = (e, \text{Gluc}^{300000}, \lambda)$.
- $M_1^3 = (e, \text{Lact}^{300000}, \lambda)$.
- $M_1^4 = (e, \text{Gluc}^{300000} + \text{Lact}^{300000}, \lambda)$.

- Para la superficie celular:

- $M_2 = (s, \text{EIICB}^{2500} + \text{EIICB} \sim \text{P}^{15000} + \text{AC}^{10000} + \text{LacY}^{3000}, \lambda)$.

- Para el citoplasma:

- $M_3 = (c, w_3, s_3)$.

en donde w_3 representa los objetos iniciales

$$w_3 = \text{RNAP}^{300} + \text{Rib}^{3000} + \text{EIIA}^{2000} + \text{EIIA} \sim \text{P}^{13000} + \text{ATP}^{1000000} + \beta - \text{Galac}^{3000} + \text{LacI}^{1500} + \text{CRP}^{10000}$$

y s_3 representa la cadena inicial del Lac Operon:

$$\langle \text{cap. op.} \cdot \underbrace{\text{lacZ}_s \cdot \text{lacZ}_m \cdot \dots \cdot \text{lacZ}_m \cdot \text{lacZ}_e}_{30} \cdot \underbrace{\text{lacY}_s \cdot \text{lacY}_m \cdot \dots \cdot \text{lacY}_m \cdot \text{lacY}_e}_{12} \cdot \underbrace{\text{lacA}_s \cdot \text{lacA}_m \cdot \dots \cdot \text{lacA}_m \cdot \text{lacA}_e}_6 \rangle$$

Estos parámetros dan lugar a una familia, $\mathcal{F}_{lac}(\Pi_{lac}; (\mathbf{M}_0, C))$, que consta de cuatro modelos diferentes de sistemas P para la especificación del sistema Lac Operon Π_{lac} :

1. $PSM_1 = (\Pi_{lac}; (M_1^1, M_2, M_3), C)$
2. $PSM_2 = (\Pi_{lac}; (M_1^2, M_2, M_3), C)$

$$3. PSM_3 = (\Pi_{lac}; (M_1^3, M_2, M_3), C)$$

$$4. PSM_4 = (\Pi_{lac}; (M_1^4, M_2, M_3), C)$$

6.3. Análisis de la Expresión de Genes en Lac Operon

En esta sección se analiza el comportamiento del Lac Operon bajo diferentes condiciones iniciales mediante la ejecución de simulaciones de los cuatro modelos antes descritos, PSM_1 , PSM_2 , PSM_3 y PSM_4 . Para ello, usaremos el algoritmo multicompartimental de Gillespie descrito en la sección 3.3 y un simulador que implementa dicho algoritmo y que está disponible en [134].

Las condiciones iniciales que se estudian son: (a) ausencia de lactosa y glucosa; (b) abundancia de lactosa y ausencia de glucosa; (c) ausencia de lactosa y abundancia glucosa; (d) abundancia de lactosa y glucosa. Estas cuatro condiciones son descritas en los cuatro modelos de sistemas P, PSM_1 , PSM_2 , PSM_3 y PSM_4 descritos en la sección previa. Se analizará el comportamiento de cada modelo de sistema P, presentando la evolución en el tiempo del número de moléculas de algunas proteínas y enzimas principales, así como mediante la descripción de la configuración del Lac Operon que determina el nivel de expresión de los genes.

- Comportamiento del sistema **sin glucosa y sin lactosa** en el entorno: modelo PSM_1 .

En estas condiciones, un estado característico del Lac Operon obtenido en nuestro simulador es el siguiente:

$$\begin{aligned}
 & cap\text{-}CRP\text{-}cAMP_2 \text{ RNAP } op\text{-}LacI \text{ lacZ_s lacZ_m } \text{-op Rib} \\
 & \text{-lacZ_s } \text{-lacZ_m RNAP lacZ_m lacZ_m lacZ_m lacZ_m} \\
 & \text{lacZ_m lacZ_m lacZ_m } \text{-op Rib } \text{-lacZ_s } \text{-lacZ_m } \text{-lacZ_m} \\
 & \text{-lacZ_m } \text{-lacZ_m } \text{-lacZ_m } \text{-lacZ_m } \text{-lacZ_m } \text{-lacZ_m RNAP} \\
 & \text{lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m} \\
 & \text{lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m} \\
 & \text{lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_m lacZ_e} \\
 & \text{lacY_s lacY_m lacY_m lacY_m lacY_m lacY_m lacY_m lacY_m} \\
 & \text{lacY_m lacY_m lacY_m lacY_m lacY_e lacA_s lacA_m lacA_m} \\
 & \text{lacA_m lacA_m lacA_e}
 \end{aligned} \tag{6.2}$$

La configuración del Lac Operon en el modelo PSM_1 , $\langle cap^{CRP-cAMP_2}.RNAP.op^{LacI} \rangle$ produce un leve incremento en la expresión de los genes codificados.

- Comportamiento del sistema **con abundante glucosa y sin lactosa** en el entorno: modelo PSM_2 .

En estas condiciones ningún activador será enlazado en el sitio CAP del Lac Operon. Puesto que no existe lactosa en el entorno, no aparecerá aloactosa en el citoplasma y el represor se activará. Por tanto, en estas condiciones la configuración del switch será $\langle cap.op^{LacI} \rangle$, que sólo permite transcripción del Lac Operon a tasas muy bajas.

- Comportamiento del sistema **sin glucosa pero con abundante lactosa** en el entorno: modelo PSM_3 .

En estas condiciones, por una parte la ausencia de glucosa en el entorno permite a la enzima EIHA~P interactuar con AC en la síntesis de un alto número de cAMP. Por otra parte, puesto

que la lactosa es abundante en el entorno, será transportada por la permeasa LacY hacia el citoplasma. Luego, en este caso el Lac Operon será, a la vez, inducido (ningún represor será enlazado al operador) y activado (el activador será enlazado al sitio CAP) Por tanto, la configuración del switch será $\langle \text{cap}^{\text{CRP-cAMP}_2}.\text{op} \rangle$ y los genes codificados en el Lac Operon serán transcritos masivamente.

- Comportamiento del sistema **con abundante glucosa** y con **abundante lactosa** en el entorno: modelo PSM_4 .

En estas condiciones, la configuración del switch será $\langle \text{cap.op} \rangle$; es decir, el Lac Operon será inducido pero ningún activador CRP-cAMP₂ se enlazará al sitio CAP. Ello produce una tasa relativamente baja de transcripción de genes codificados en el Lac Operon.

Capítulo 7

Modelado de la Traducción de Señales

Uno de los procesos celulares más importantes consiste en la percepción de señales externas a través de receptores situados en la superficie celular. Las células observan el estado de cambio en el entorno mediante estos procesos y producen la respuesta correcta adaptando su funcionamiento de acuerdo con las señales que reciben. La importancia de las membranas en el funcionamiento y la estructura de los sistemas de traducción es clara y, por ello, los sistemas P parecen un marco adecuado para el desarrollo de modelos de tales sistemas.

7.1. Traducción de Señales

La traducción de señales es una colección de procesos celulares por el cual las células traducen señales extracelulares en respuestas específicas. Entre esos procesos caben destacar enlaces, roturas de enlaces de la señal al dominio extracelular del receptor, reclutamiento y liberación de proteínas desde y hacia el citoplasma, degradación molecular y transformación.

El modelado del comportamiento dinámico de las cascadas de señales no es nada fácil. El conocimiento que se tiene actualmente de los caminos y de sus interacciones es bastante incompleto. Además, el efecto de una señal puede cambiar el estado de toda la célula.

En la figura 7.1 se ilustra una sucesión ordinaria de eventos en las cascadas de señales que constituyen un sistema de traducción de señales: la señal (una sustancia química que actúa como ligando) se aproxima a la superficie celular y las células tratan de importarla bien a través de una serie de estímulos que permiten el enlace con su respectivo receptor (como sucede en el caso del Lac Operon), o bien a través de un receptor que sólo transmite al interior de la célula la *información*.

7.2. Algoritmo Determinista con Tiempo de Espera

En la sección 3.3.1 se presentó una generalización del algoritmo clásico de Gillespie a estructuras compartimentalizadas de sistemas P (*algoritmo multicompartimental de Gillespie*) y su uso estuvo motivado por el hecho de que sistemas con un bajo número de moléculas (entiéndase como tal, por debajo de mil) no pueden ser modelados usando ODEs.

Ahora bien, para sistemas con un alto número de moléculas son válidas las aproximaciones deterministas. En esta sección se presenta una versión determinista del algoritmo multicompartimental de Gillespie, que denominaremos *algoritmo determinista con tiempo de espera*.

En un modelo de sistema P, dado el estado de un compartimento i , $M_i = (l_i, w_i, s_i)$, la siguiente regla a aplicar y el tiempo de espera correspondiente es calculado como sigue:

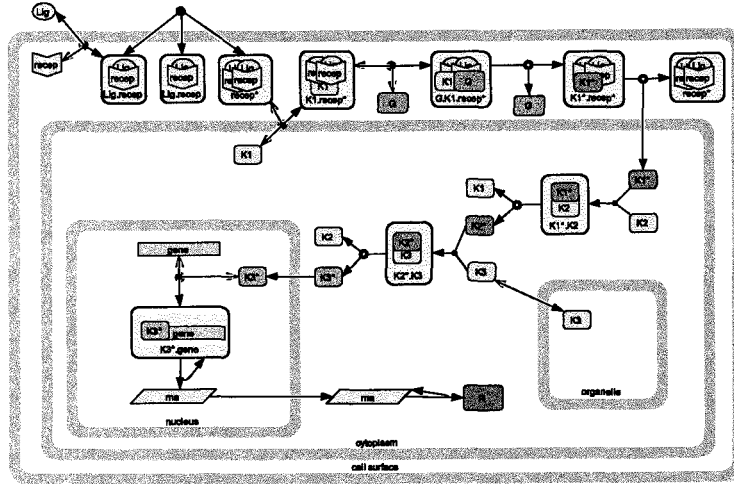


Figura 7.1: Traducción de señales.

1. Para cada regla en $r_j \in R_{l_i}$, calcular su velocidad, $v_j(M_i)$, multiplicando la constante estocástica c_j^i asociada a $r_j \in R_{l_i}$ por el número de posibles combinaciones de los objetos y subcadenas presentes en la parte izquierda de la regla.
2. Calcular el tiempo de espera asociado a $r_j \in R_{l_i}$ así:

$$\tau_j = \frac{1}{v_j(M_i)} \quad (7.1)$$

Entonces, las reglas asociadas a las membranas del sistema son ordenadas en una cola de prioridad, en orden ascendente respecto del tiempo de espera. Tras aplicar la primera regla y dependiendo del tipo que se trate, habrá que actualizar los compartimentos afectados así como los tiempos de espera de las reglas asociadas a esos compartimentos. El algoritmo *para* cuando se alcanza un tiempo de computación prefijado.

Una especificación del algoritmo es la siguiente:

- **Inicialización**

- Poner el tiempo de simulación a $t = 0$;
- Para cada regla r_j asociada a cada membrana i en μ , calcular la terna (τ_j, j, i) usando el procedimiento descrito anteriormente; construir una lista conteniendo todas esas ternas;
- Ordenar la lista de ternas (τ_j, j, i) de acuerdo con τ_j , en orden ascendente;

- **Iteración**

- Extraer la primera terna, $(\tau_{i_0}, j_{i_0}, i_0)$ de la lista;
- Poner el tiempo de simulación a $t = t + \tau_{i_0}$;
- Actualizar el tiempo de espera para las restantes ternas de la lista, restando τ_{i_0} ;
- Aplicar una sólo vez la regla $r_{j_{i_0}}$ en la membrana i_0 ;

- Para cada membrana i' afectada por la aplicación de la regla, quitar de la lista todas las ternas correspondientes a las reglas asociadas a i' , $(\tau_{i'}, j_{i'}, i')$;
- Para cada regla asociada a cada membrana i' afectada por la aplicación de la regla $r_{j_{i_0}}$, calcular su correspondiente terna;
- Añadir las nuevas ternas a la lista y ordenarlas de acuerdo con el tiempo de espera, reiterando el proceso.

- **Parada**

- La simulación termina cuando el tiempo t alcanza o excede un tiempo máximo prefijado.

7.3. Modelado de una Cascada de Señales del EGFR

El receptor del factor de crecimiento epidérmico (EGFR) pertenece a la familia de receptores de la quinasa tirosina y su enlace con el del factor de crecimiento epidérmico (EGF) induce un proceso de dimerización en el receptor y un proceso de autofosforilización en el dominio intracelular. Entonces se reclutan una serie de proteínas iniciándose una compleja cascada de señales. Dos caminos principales permiten la activación de la proteína Ras-GTP por hidrolización de Ras-GDP. El primero depende de la proteína Shc y el segundo no. La proteína Ras-GTP actúa como un *switch* que estimula la cascada de la Mitogen Activated Protein (MAP) que produce la fosforilización de las proteínas Raf, MEK and ERK, que juegan un papel crucial en la regulación de varias proteínas celulares y en factores de transcripción del núcleo celular. Se ha constatado que una disfunción de la expresión del EGFR, o de la producción del ligando está asociada fuertemente con procesos de génesis de tumores.

En esta sección se presenta una especificación de sistemas P y una familia de modelos para una cascada de señales del EGFR. Nuestra especificación y nuestros modelos sólo tratarán interacciones proteínicas; por tanto, sólo usaremos objetos para representar las entidades moleculares y reglas de reescrituras sobre multiconjuntos para describir las interacciones moleculares del sistema.

7.3.1. Una Especificación de Sistema P para una Cascada de Señales del EGFR

Las membranas de nuestro sistema P de especificación serán tres: el entorno (donde estarán localizadas las señales), etiquetado por e , la superficie celular (donde estarán el receptor y los complejos involucrados en el proceso), etiquetado por s , y el citoplasma (en donde la cascada se desarrolla propiamente), etiquetada por c . La raíz de la estructura es el entorno, la superficie celular es hija del entorno y el citoplasma es hijo de la superficie celular.

El alfabeto de trabajo describe las entidades moleculares que consideramos relevantes:

Finalmente, las reglas asociadas a las distintas regiones describen las interacciones que se pueden producir entre las distintas entidades moleculares que conforman el sistema. Para cada regla se tiene una constante macroscópica determinista, obtenida de los modelos basados en ODEs que aparecen en [32, 68, 81, 115, 117, 119, 130]

- **Activación del Receptor**

Object	Molecular Entity
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
EGFR-EGF	EGFR and EGF complex
EGFR-EGF2, EGFR-EGF2*	Dimerised and Phosphorylated Receptor
TP1, TP2, TP3, TP4	Tyrosine Phosphatase 1,2,3,4
EGFR-EGF2*-TP1	Phosphorylated Receptor and TP1 complex
EGFR-EGF2-TP1	Receptor, TP1 complex
PLC, PCL*	Phospholipase C- γ and its phosphorylated state
EGFR-EGF2*-PLC	Phosphorylated Receptor and PLC complex
EGFR-EGF2*-PLC*	Phosphorylated Receptor and PLC* complex
PLC*-TP2	PLC* and TP2 complex
PLC-TP2	PLC and TP2 complex
PI3K, PI3K*	Phosphatidylinositol 3-kinase and its phosphorylated state
EGFR-EGF2*-PI3K	Phosphorylated receptor and PI3K complex
EGFR-EGF2*-PI3K*	Phosphorylated receptor and PI3K* complex
PI3K*-TP4	PI3K* and TP4 complex
PI3K-TP4	PI3K and TP4 complex
Grb2	Growth factor receptor binding protein 2
EGFR-EGF2*-Grb2	Phosphorylated receptor and Grb2 complex
SOS	Son of sevenless homologue protein
EGFR-EGF2*-Grb2-SOS	Phosphorylated receptor, Grb2, SOS complex
Grb2-SOS	Grb2 and SOS complex
Shc, Shc*	src homology 2 domain and its phosphorylated state
EGFR-EGF2*-Shc	Phosphorylated receptor and Shc complex
EGFR-EGF2*-Shc*	Phosphorylated receptor and Shc* complex
Shc*-TP3	Shc* and TP3 complex
Shc-TP3	Shc and TP3 complex
EGFR-EGF2*-Shc*-Grb2	Phosphorylated receptor, Shc* and Grb2 complex
EGFR-EGF2*-Shc*-Grb2-SOS	Phosphorylated receptor, Shc*, Grb2 and SOS complex
Shc*-Grb2-SOS	Shc*, Grb2 and SOS complex
Shc*-Grb2-SOS	Shc* and Grb2complex
Ras-GDP, Ras-GTP, Ras-GTP*	Ras protein and its hydrolised state
Raf, Raf*	Raf protein and its phosphorylated state
Ras-GTP-Raf	Ras-GTP and Raf complex
MEK, MEK-P, MEK-PP	Mitogen-activated Protein Kinase and its phosphorylated states
ERK, ERK-P, ERK-PP	External Regulated Kinase and its phosphorylated states
Raf*-MEK	Raf* and MEK complex
Raf*-MEK-P	Raf* and MEK-P complex
MEK-PP-ERK	MEK-PP and ERK complex
MEK-PP-ERK-P	MEK-PP and ERK-P complex
P2, P3	Phosphatases 2,3
MEK-P-P2	MEK-P and P2 complex
MEK-PP-P2	MEK-PP and P2 complex
ERK-P-P3	ERK-P and P3 complex
ERK-PP-P3	ERK-PP and P3 complex

Figura 7.2: Objetos en una especificación de sistema P en EGFR

Regla	Constante cinética
$r_1 : \text{EGF} [\text{EGFR}]_s \xrightarrow{c_1} [\text{EGFR-EGF}]_s$	$k_1 = 3 \times 10^{-3}$
$r_2 : [\text{EGFR-EGF}]_s \xrightarrow{c_2} \text{EGF} [\text{EGFR}]_s$	$k_2 = 6 \times 10^{-2}$
$r_3 : [\text{EGFR-EGF} + \text{EGFR-EGF}]_s \xrightarrow{c_3} [\text{EGFR-EGF2}]_s$	$k_3 = 1,1 \times 10^{-2}$
$r_4 : [\text{EGFR-EGF2}]_s \xrightarrow{c_4} [\text{EGFR-EGF} + \text{EGFR-EGF}]_s$	$k_4 = 0,814$
$r_5 : [\text{EGFR-EGF2}]_s \xrightarrow{c_5} [\text{EGFR-EGF2}^*]_s$	$k_5 = 2,71$
$r_6 : [\text{EGFR-EGF2}^*]_s \xrightarrow{c_6} [\text{EGFR-EGF2}]_s$	$k_6 = 0,0271$
$r_7 : \text{EGFR-EGF2}^* [\text{TP1}]_c \xrightarrow{c_7} \text{EGFR-EGF2}^* [\text{TP1}]_c$	$k_7 = 0,03$
$r_8 : \text{EGFR-EGF2}^* [\text{TP1}]_c \xrightarrow{c_8} \text{EGFR-EGF2}^* [\text{TP1}]_c$	$k_8 = 5,91$
$r_9 : [\text{EGFR-EGF2}^* [\text{TP1}]_s \xrightarrow{c_9} [\text{EGFR-EGF2} [\text{TP1}]_s$	$k_9 = 7,44$
$r_{10} : \text{EGFR-EGF2} [\text{TP1}]_c \xrightarrow{c_{10}} \text{EGFR-EGF2} [\text{TP1}]_c$	$k_{10} = 5,1 \times 10^{-3}$
$r_{11} : \text{EGFR-EGF2} [\text{TP1}]_c \xrightarrow{c_{11}} \text{EGFR-EGF2} [\text{TP1}]_c$	$k_{11} = 2 \times 10^{-4}$

■ Internalización y degradación del receptor

Regla	Constante cinética
$r_{12} : \text{EGFR} []_c \xrightarrow{c_{12}} [\text{EGFR}]_c$	$k_{12} = 5 \times 10^{-5}$
$r_{13} : [\text{EGFR}]_c \xrightarrow{c_{13}} \text{EGFR} []_c$	$k_{13} = 5 \times 10^{-3}$
$r_{14} : \text{EGFR-EGF2*} []_c \xrightarrow{c_{14}} [\text{EGFR-EGF2*}]_c$	$k_{14} = 5 \times 10^{-5}$
$r_{15} : [\text{EGFR}]_c \xrightarrow{c_{15}} []_c$	$k_{15} = 6,64 \times 10^{-4}$
$r_{16} : [\text{EGFR-EGF2*}]_c \xrightarrow{c_{16}} []_c$	$6,64 \times 10^{-4}$
$r_{17} : [\text{EGF}]_c \xrightarrow{c_{17}} []_c$	$k_{17} = 1,67 \times 10^{-4}$
$r_{18} : [\text{EGFR-EGF2*}-\text{Grb2-CPP}]_s \xrightarrow{c_{18}} [\text{EGFR-EGF2*}-\text{Grb2-CPP}]_s$	$k_{18} = 1,73 \times 10^{-7}$
$r_{19} : [\text{EGFR-EGF2*}-\text{Grb2-CPP}]_s \xrightarrow{c_{19}} [\text{EGFR-EGF2*}-\text{Grb2-CPP}]_s$	$k_{19} = 1,66 \times 10^{-3}$
$r_{20} : \text{EGFR-EGF2*}-\text{Grb2-CPP} []_c \xrightarrow{c_{20}} [\text{EGFR-EGF2*}-\text{Grb2-CPP}]_c$	$k_{20} = 0,003$
$r_{21} : [\text{CPP}]_c \xrightarrow{c_{21}} \text{CPP} []_c$	$k_{21} = 10^{-5}$

■ Camino PLC γ

Regla	Constante cinética
$r_{22} : \text{EGFR-EGF2*} [\text{PLC}\gamma]_c \xrightarrow{c_{22}} \text{EGFR-EGF2*}-\text{PLC}\gamma []_c$	$k_{22} = 0,1$
$r_{23} : \text{EGFR-EGF2*}-\text{PLC}\gamma []_c \xrightarrow{c_{23}} \text{EGFR-EGF2*} [\text{PLC}\gamma]_c$	$k_{23} = 1$
$r_{24} : [\text{EGFR-EGF2*}-\text{PLC}\gamma]_s \xrightarrow{c_{24}} [\text{EGFR-EGF2*}-\text{PLC}\gamma^*]_s$	$k_{24} = 10$
$r_{25} : [\text{EGFR-EGF2*}-\text{PLC}\gamma^*]_s \xrightarrow{c_{25}} [\text{EGFR-EGF2*}-\text{PLC}\gamma]_s$	$0,1$
$r_{26} : \text{EGFR-EGF2*}-\text{PLC}\gamma^* []_c \xrightarrow{c_{26}} \text{EGFR-EGF2*} [\text{PLC}\gamma^*]_c$	$k_{26} = 1,5 \times 10^{-4}$
$r_{27} : \text{EGFR-EGF2*} [\text{PLC}\gamma^*]_c \xrightarrow{c_{27}} \text{EGFR-EGF2*}-\text{PLC}\gamma^* []_c$	$k_{27} = 7,5 \times 10^{-8}$
$r_{28} : \text{PLC}\gamma^* []_s \xrightarrow{c_{28}} [\text{PLC}\gamma^*]_s$	$k_{28} = 1$
$r_{29} : [\text{PLC}\gamma^*]_s \xrightarrow{c_{29}} \text{PLC}\gamma^* []_s$	$k_{29} = 0,03$
$r_{30} : [\text{PLC}\gamma^* + \text{TP}_2]_c \xrightarrow{c_{30}} [\text{PLC}\gamma^*-\text{TP}_2]_c$	$k_{30} = 1,5 \times 10^{-4}$
$r_{31} : [\text{PLC}\gamma^*-\text{TP}_2]_c \xrightarrow{c_{31}} [\text{PLC}\gamma^* + \text{TP}_2]_c$	$k_{31} = 1,5 \times 10^{-3}$
$r_{32} : [\text{PLC}\gamma^*-\text{TP}_2]_c \xrightarrow{c_{32}} [\text{PLC}\gamma-\text{TP}_2]_c$	$k_{32} = 0,1$
$r_{33} : [\text{PLC}\gamma-\text{TP}_2]_c \xrightarrow{c_{33}} [\text{PLC}\gamma + \text{TP}_2]_c$	$k_{33} = 10^{-5}$
$r_{34} : [\text{PLC}\gamma \text{ TP}_2]_c \xrightarrow{c_{34}} [\text{PLC}\gamma-\text{TP}_2]_c$	$k_{34} = 1,8 \times 10^{-7}$

■ Camino Grb2

Regla	Constante cinética
$r_{35} : \text{EGFR-EGF2*} [\text{Grb2}]_c \xrightarrow{c_{35}} \text{EGFR-EGF2*}-\text{Grb2} []_c$	$k_{35} = 1,5 \times 10^{-3}$
$r_{36} : \text{EGFR-EGF2*}-\text{Grb2} []_c \xrightarrow{c_{36}} \text{EGFR-EGF2*} [\text{Grb2}]_c$	$k_{36} = 0,2$
$r_{37} : \text{EGFR-EGF2*}-\text{Grb2} [\text{SOS}]_c \xrightarrow{c_{37}} \text{EGFR-EGF2*}-\text{Grb2}-\text{SOS} []_c$	$k_{37} = 0,01$
$r_{38} : \text{EGFR-EGF2*}-\text{Grb2}-\text{SOS} []_c \xrightarrow{c_{38}} \text{EGFR-EGF2*}-\text{Grb2} [\text{SOS}]_c$	$k_{38} = 0,06$
$r_{39} : \text{EGFR-EGF2*}-\text{Grb2}-\text{SOS} []_c \xrightarrow{c_{39}} \text{EGFR-EGF2*} [\text{Grb2}-\text{SOS}]_c$	$k_{39} = 2,8 \times 10^{-3}$
$r_{40} : \text{EGFR-EGF2*} [\text{Grb2}-\text{SOS}]_c \xrightarrow{c_{40}} \text{EGFR-EGF2*}-\text{Grb2}-\text{SOS} []_c$	$k_{40} = 5,3 \times 10^{-5}$
$r_{41} : [\text{Grb2}-\text{SOS}]_c \xrightarrow{c_{41}} [\text{Grb2} + \text{SOS}]_c$	$k_{41} = 10^{-4}$
$r_{42} : [\text{Grb2} + \text{SOS}]_c \xrightarrow{c_{42}} [\text{Grb2}-\text{SOS}]_c$	$k_{42} = 6,7 \times 10^{-6}$

■ Camino Shc

Regla	Constante cinética
$r_{43} : \text{EGFR-EGF2}^* [\text{Shc}]_c \xrightarrow{c_{43}} \text{EGFR-EGF2}^*\text{-Shc} []_c$	$k_{43} = 0,1$
$r_{44} : \text{EGFR-EGF2}^*\text{-Shc} []_c \xrightarrow{c_{44}} \text{EGFR-EGF2}^* [\text{Shc}]_c$	$k_{44} = 1$
$r_{45} : [\text{EGFR-EGF2}^*\text{-Shc}]_s \xrightarrow{c_{45}} [\text{EGFR-EGF2}^*\text{-Shc}^*]_s$	$k_{45} = 20$
$r_{46} : [\text{EGFR-EGF2}^*\text{-Shc}^*]_s \xrightarrow{c_{46}} [\text{EGFR-EGF2}^*\text{-Shc}]_s$	$k_{46} = 0,2$
$r_{47} : \text{EGFR-EGF2}^*\text{-Shc}^* []_c \xrightarrow{c_{47}} \text{EGFR-EGF2}^* [\text{Shc}^*]_c$	$k_{47} = 5 \times 10^{-4}$
$r_{48} : \text{EGFR-EGF2}^* [\text{Shc}^*]_c \xrightarrow{c_{48}} \text{EGFR-EGF2}^*\text{-Shc}^* []_c$	$k_{48} = 3,56 \times 10^{-7}$
$r_{49} : [\text{Shc}^* + \text{TP}_3]_c \xrightarrow{c_{49}} [\text{Shc}^*\text{-TP}_3]_c$	$k_{49} = 5 \times 10^{-3}$
$r_{50} : [\text{Shc}^*\text{-TP}_3]_c \xrightarrow{c_{50}} [\text{Shc}^* + \text{TP}_3]_c$	$k_{50} = 9,5 \times 10^{-2}$
$r_{51} : [\text{Shc}^*\text{-TP}_3]_c \xrightarrow{c_{51}} [\text{Shc-TP}_3]_c$	$k_{51} = 0,15$
$r_{52} : [\text{Shc-TP}_3]_c \xrightarrow{c_{52}} [\text{Shc} + \text{TP}_3]_c$	$k_{52} = 2 \times 10^{-4}$
$r_{53} : [\text{Shc} + \text{TP}_3]_c \xrightarrow{c_{53}} [\text{Shc-TP}_3]_c$	$k_{53} = 4 \times 10^{-6}$
$r_{54} : \text{EGFR2}^*\text{-Shc}^* [\text{Grb2}]_c \xrightarrow{c_{54}} \text{EGFR2}^*\text{-Shc}^*\text{-Grb2} []_c$	$k_{54} = 1,5 \times 10^{-3}$
$r_{55} : \text{EGFR2}^*\text{-Shc}^*\text{-Grb2} []_c \xrightarrow{c_{55}} \text{EGFR2}^*\text{-Shc}^* [\text{Grb2}]_c$	$k_{55} = 4,95 \times 10^{-2}$
$r_{56} : \text{EGFR2}^*\text{-Shc}^*\text{-Grb2} []_c \xrightarrow{c_{56}} \text{EGFR2}^* [\text{Shc}^*\text{-Grb2}]_c$	$k_{56} = 6,5 \times 10^{-4}$
$r_{57} : \text{EGFR2}^* [\text{Shc}^*\text{-Grb2}]_c \xrightarrow{c_{57}} \text{EGFR2}^*\text{-Shc}^*\text{-Grb2} []_c$	$k_{57} = 1,4 \times 10^{-6}$
$r_{58} : [\text{Shc}^*\text{-Grb2}]_c \xrightarrow{c_{58}} [\text{Shc}^* + \text{Grb2}]_c$	$k_{58} = 1 \times 10^{-3}$
$r_{59} : [\text{Shc}^* + \text{Grb2}]_c \xrightarrow{c_{59}} [\text{Shc}^*\text{-Grb2}]_c$	$k_{59} = 1 \times 10^{-5}$
$r_{60} : \text{EGFR2}^*\text{-Shc}^*\text{-Grb2} [\text{SOS}]_c \xrightarrow{c_{60}} \text{EGFR2}^*\text{-Shc}^*\text{-Grb2-SOS} []_c$	$k_{60} = 0,015$
$r_{61} : \text{EGFR2}^*\text{-Shc}^*\text{-Grb2-SOS} []_c \xrightarrow{c_{61}} \text{EGFR2}^*\text{-Shc}^*\text{-Grb2} [\text{SOS}]_c$	$k_{61} = 0,03$
$r_{62} : \text{EGFR2}^*\text{-Shc}^*\text{-Grb2-SOS} []_c \xrightarrow{c_{62}} \text{EGFR2}^* [\text{Shc}^*\text{-Grb2-SOS}]_c$	$k_{62} = 1,1 \times 10^{-3}$
$r_{63} : \text{EGFR2}^* [\text{Shc}^*\text{-Grb2-SOS}]_c \xrightarrow{c_{63}} \text{EGFR2}^*\text{-Shc}^*\text{-Grb2-SOS} []_c$	$k_{63} = 2,37 \times 10^{-6}$
$r_{64} : [\text{Shc}^*\text{-Grb2-SOS}]_c \xrightarrow{c_{64}} [\text{Shc}^*\text{-Grb2-SOS}]_c$	$k_{64} = 0,06$
$r_{65} : [\text{Shc}^*\text{-Grb2-SOS}]_c \xrightarrow{c_{65}} [\text{Shc}^*\text{-Grb2-SOS}]_c$	$k_{65} = 0,03$
$r_{66} : [\text{Shc}^*\text{-Grb2-SOS}]_c \xrightarrow{c_{66}} [\text{Shc}^* \text{Grb2-SOS}]_c$	$k_{66} = 0,035$
$r_{67} : [\text{Shc}^* \text{Grb2-SOS}]_c \xrightarrow{c_{67}} [\text{Shc}^*\text{-Grb2-SOS}]_c$	$k_{67} = 2,5 \times 10^{-3}$
$r_{68} : \text{EGFR2}^*\text{-Shc}^* [\text{Grb2-SOS}]_c \xrightarrow{c_{68}} \text{EGFR2}^*\text{-Shc}^*\text{-Grb2-SOS} []_c$	$k_{68} = 0,5$
$r_{69} : \text{EGFR2}^*\text{-Shc}^*\text{-Grb2-SOS} []_c \xrightarrow{c_{69}} \text{EGFR2}^*\text{-Shc}^* [\text{Grb2-SOS}]_c$	$k_{69} = 0,1$

■ Actividad interna del receptor

Regla	Constante cinética
$r_{70} : [\text{EGF} + \text{EGFR}]_c \xrightarrow{c_{70}} [\text{EGFR-EGF}]_c$	$k_{70} = 3 \times 10^{-3}$
$r_{71} : [\text{EGFR-EGF}]_c \xrightarrow{c_{71}} [\text{EGF} + \text{EGFR}]_c$	$k_{71} = 0,06$
$r_{72} : [\text{EGFR-EGF} + \text{EGFR-EGF}]_c \xrightarrow{c_{72}} [\text{EGFR-EGF2}]_c$	$k_{72} = 1,1 \times 10^{-2}$
$r_{73} : [\text{EGFR-EGF2}]_c \xrightarrow{c_{73}} [\text{EGFR-EGF} + \text{EGFR-EGF}]_c$	$k_{73} = 0,814$
$r_{74} : [\text{EGFR-EGF2}]_c \xrightarrow{c_{74}} [\text{EGFR-EGF2}^*]_c$	$k_{74} = 2,71$
$r_{75} : [\text{EGFR-EGF2}^*]_c \xrightarrow{c_{75}} [\text{EGFR-EGF2}]_c$	$k_{75} = 0,0271$
$r_{76} : [\text{EGFR-EGF2}^* + \text{TP}_1]_c \xrightarrow{c_{76}} [\text{EGFR-EGF2}^*\text{-TP}_1]_c$	$k_{76} = 5,1 \times 10^{-3}$
$r_{77} : [\text{EGFR-EGF2}^*\text{-TP}_1]_c \xrightarrow{c_{77}} [\text{EGFR-EGF2}^* + \text{TP}_1]_c$	$k_{77} = 5,91$
$r_{78} : [\text{EGFR-EGF2}^*\text{-TP}_1]_c \xrightarrow{c_{78}} [\text{EGFR-EGF2-TP}_1]_c$	$k_{78} = 7,44$
$r_{79} : [\text{EGFR-EGF2-TP}_1]_c \xrightarrow{c_{79}} [\text{EGFR-EGF2} + \text{TP}_1]_c$	$k_{79} = 5,1 \times 10^{-3}$
$r_{80} : [\text{EGFR-EGF2} + \text{TP}_1]_c \xrightarrow{c_{80}} [\text{EGFR-EGF2-TP}_1]_c$	$k_{80} = 2 \times 10^{-4}$
$r_{81} : [\text{EGFR-EGF2}^* + \text{Grb2}]_c \xrightarrow{c_{81}} [\text{EGFR-EGF2}^*\text{-Grb2}]_c$	$k_{81} = 1,5 \times 10^{-3}$
$r_{82} : [\text{EGFR-EGF2}^*\text{-Grb2}]_c \xrightarrow{c_{82}} [\text{EGFR-EGF2}^* + \text{Grb2}]_c$	$k_{82} = 0,2$
$r_{83} : [\text{EGFR-EGF2}^*\text{-Grb2} + \text{SOS}]_c \xrightarrow{c_{83}} [\text{EGFR-EGF2}^*\text{-Grb2-SOS}]_c$	$k_{83} = 0,01$
$r_{84} : [\text{EGFR-EGF2}^*\text{-Grb2-SOS}]_c \xrightarrow{c_{84}} [\text{EGFR-EGF2}^*\text{-Grb2} + \text{SOS}]_c$	$k_{84} = 0,06$
$r_{85} : [\text{EGFR-EGF2}^*\text{-Grb2-SOS}]_c \xrightarrow{c_{85}} [\text{EGFR-EGF2}^* + \text{Grb2-SOS}]_c$	$k_{85} = 2,8 \times 10^{-3}$
$r_{86} : [\text{EGFR-EGF2}^* + \text{Grb2-SOS}]_c \xrightarrow{c_{86}} [\text{EGFR-EGF2}^*\text{-Grb2-SOS}]_c$	$k_{86} = 5,3 \times 10^{-5}$
$r_{87} : [\text{EGFR-EGF2}^* + \text{Shc}]_c \xrightarrow{c_{87}} [\text{EGFR-EGF2}^*\text{-Shc}]_c$	$k_{87} = 0,1$
$r_{88} : [\text{EGFR-EGF2}^*\text{-Shc}]_c \xrightarrow{c_{88}} [\text{EGFR-EGF2}^* + \text{Shc}]_c$	$k_{88} = 1$
$r_{89} : [\text{EGFR-EGF2}^*\text{-Shc}]_c \xrightarrow{c_{89}} [\text{EGFR-EGF2}^*\text{-Shc}^*]_c$	$k_{89} = 20$
$r_{90} : [\text{EGFR-EGF2}^*\text{-Shc}^*]_c \xrightarrow{c_{90}} [\text{EGFR-EGF2}^*\text{-Shc}]_c$	$k_{90} = 0,2$
$r_{91} : [\text{EGFR-EGF2}^*\text{-Shc}^*]_c \xrightarrow{c_{91}} [\text{EGFR-EGF2}^* + \text{Shc}^*]_c$	5×10^{-4}
$r_{92} : [\text{EGFR-EGF2}^* + \text{Shc}^*]_c \xrightarrow{c_{92}} [\text{EGFR-EGF2}^*\text{-Shc}^*]_c$	$3,56 \times 10^{-7}$

Regla	Constante cinética
$r_{93} : [\text{EGFR2}^* \text{-Shc}^* + \text{Grb2}]_c \xrightarrow{c^{93}} [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2}]_c$	$1,5 \times 10^{-3}$
$r_{94} : [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2}]_c \xrightarrow{c^{94}} [\text{EGFR2}^* \text{-Shc}^* + \text{Grb2}]_c$	$4,95 \times 10^{-2}$
$r_{95} : [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2}]_c \xrightarrow{c^{95}} [\text{EGFR2}^* + \text{Shc}^* \text{-Grb2}]_c$	$6,5 \times 10^{-4}$
$r_{96} : [\text{EGFR2}^* + \text{Shc}^* \text{-Grb2}]_c \xrightarrow{c^{96}} [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2}]_c$	$1,4 \times 10^{-6}$
$r_{97} : [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2} + \text{SOS}]_c \xrightarrow{c^{97}} [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2-SOS}]_c$	$k_{97} = 0,015$
$r_{98} : [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2-SOS}]_c \xrightarrow{c^{98}} [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2} + \text{SOS}]_c$	$k_{98} = 0,03$
$r_{99} : [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2-SOS}]_c \xrightarrow{c^{99}} [\text{EGFR2}^* + \text{Shc}^* \text{-Grb2-SOS}]_c$	$k_{99} = 1,1 \times 10^{-3}$
$r_{100} : [\text{EGFR2}^* + \text{Shc}^* \text{-Grb2-SOS}]_c \xrightarrow{c^{100}} [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2-SOS}]_c$	$k_{100} = 2,37 \times 10^{-6}$
$r_{101} : [\text{EGFR2}^* \text{-Shc}^* + \text{Grb2-SOS}]_c \xrightarrow{c^{101}} [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2-SOS}]_c$	$k_{101} = 0,5$
$r_{102} : [\text{EGFR2}^* \text{-Shc}^* \text{-Grb2-SOS}]_c \xrightarrow{c^{102}} [\text{EGFR2}^* \text{-Shc}^* + \text{Grb2-SOS}]_c$	$k_{102} = 0,1$

■ Camino PI3K

Regla	Constante cinética
$r_{103} : \text{EGFR-EGF2}^* [\text{PI3K}]_c \xrightarrow{c^{103}} \text{EGFR-EGF2}^* \text{-PI3K} []_c$	$k_{103} = 0,1$
$r_{104} : \text{EGFR-EGF2}^* \text{-PI3K} []_c \xrightarrow{c^{104}} \text{EGFR-EGF2}^* [\text{PI3K}]_c$	$k_{104} = 2$
$r_{105} : [\text{EGFR-EGF2}^* \text{-PI3K}]_s \xrightarrow{c^{105}} [\text{EGFR-EGF2}^* \text{-PI3K}^*]_s$	$k_{105} = 9,85$
$r_{106} : [\text{EGFR-EGF2}^* \text{-PI3K}^*]_s \xrightarrow{c^{106}} [\text{EGFR-EGF2}^* \text{-PI3K}]_s$	$k_{106} = 0,985$
$r_{107} : \text{EGFR-EGF2}^* \text{-PI3K}^* []_c \xrightarrow{c^{107}} \text{EGFR-EGF2}^* [\text{PI3K}^*]_c$	$k_{107} = 0,047$
$r_{108} : \text{EGFR-EGF2}^* [\text{PI3K}^*]_c \xrightarrow{c^{108}} \text{EGFR-EGF2}^* \text{-PI3K}^* []_c$	$k_{108} = 4,82 \times 10^{-5}$
$r_{109} : [\text{PI3K}^* \text{TP}_4]_c \xrightarrow{c^{109}} [\text{PI3K}^* \text{-TP}_4]_c$	$k_{109} = 7 \times 10^{-4}$
$r_{110} : [\text{PI3K}^* \text{-TP}_4]_c \xrightarrow{c^{110}} [\text{PI3K}^* \text{TP}_4]_c$	$k_{110} = 7 \times 10^{-3}$
$r_{111} : [\text{PI3K}^* \text{-TP}_4]_c \xrightarrow{c^{111}} [\text{PI3K} \text{-TP}_4]_c$	$k_{111} = 0,03$
$r_{112} : [\text{PI3K} \text{-TP}_4]_c \xrightarrow{c^{112}} [\text{PI3K} \text{TP}_4]_c$	$k_{112} = 8 \times 10^{-5}$
$r_{113} : [\text{PI3K} \text{TP}_4]_c \xrightarrow{c^{113}} [\text{PI3K} \text{-TP}_4]_c$	$k_{113} = 6,7 \times 10^{-6}$

■ Activación de las proteínas Ras y Raf

Regla	Constante cinética
$r_{114} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s \xrightarrow{c^{114}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GDP}]_s$	$k_{114} = 0,015$
$r_{115} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GDP}]_s \xrightarrow{c^{115}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s$	$k_{115} = 1,3$
$r_{116} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GDP}]_s \xrightarrow{c^{116}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GTP}]_s$	$k_{116} = 0,5$
$r_{117} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GTP}]_s \xrightarrow{c^{117}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GDP}]_s$	$k_{117} = 1 \times 10^{-4}$
$r_{118} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GTP}^*]_s \xrightarrow{c^{118}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GTP}]_s$	$k_{118} = 2,1 \times 10^{-3}$
$r_{119} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GTP}]_s \xrightarrow{c^{119}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GTP}^*]_s$	$k_{119} = 0,4$
$r_{120} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GTP}]_s \xrightarrow{c^{120}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s$	$k_{120} = 0,023$
$r_{121} : [\text{EGFR-EGF2}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s \xrightarrow{c^{121}} [\text{EGFR-EGF2}^* \text{-Grb2-SOS-Ras-GTP}]_s$	$k_{121} = 2,2 \times 10^{-4}$
$r_{122} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s \xrightarrow{c^{122}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GDP}]_s$	$k_{122} = 0,015$
$r_{123} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GDP}]_s \xrightarrow{c^{123}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s$	$k_{123} = 1,3$
$r_{124} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GDP}]_s \xrightarrow{c^{124}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GTP}]_s$	$k_{124} = 0,5$
$r_{125} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GTP}]_s \xrightarrow{c^{125}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GDP}]_s$	$k_{125} = 1 \times 10^{-4}$
$r_{126} : [\text{Ras-GTP} + \text{Raf}]_s \xrightarrow{c^{126}} [\text{Ras-GTP-Raf}]_s$	$k_{126} = 0,001$
$r_{127} : [\text{Ras-GTP-Raf}]_s \xrightarrow{c^{127}} [\text{Ras-GTP} + \text{Raf}]_s$	$k_{127} = 0,0053$
$r_{128} : [\text{Ras-GTP-Raf}]_s \xrightarrow{c^{128}} [\text{Ras-GTP}^* \text{Raf}^*]_s$	$k_{128} = 1$
$r_{129} : [\text{Ras-GTP}^* + \text{Raf}^*]_s \xrightarrow{c^{129}} [\text{Ras-GTP-Raf}]_s$	$k_{129} = 7 \times 10^{-4}$
$r_{130} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GTP}^*]_s \xrightarrow{c^{130}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GTP}]_s$	$k_{130} = 7,9 \times 10^{-3}$
$r_{131} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GTP}]_s \xrightarrow{c^{131}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GTP}^*]_s$	$k_{131} = 0,3$
$r_{132} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GTP}]_s \xrightarrow{c^{132}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s$	$k_{132} = 0,023$
$r_{133} : [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS} + \text{Ras-GDP}]_s \xrightarrow{c^{133}} [\text{EGFR-EGF2}^* \text{-Shc}^* \text{-Grb2-SOS-Ras-GTP}]_s$	$k_{133} = 2,2 \times 10^{-4}$
$r_{134} : \text{Raf}^* [\text{Phosp}_1]_c \xrightarrow{c^{134}} \text{Raf}^* \text{-Phosp}_1 []_c$	$k_{134} = 0,0717$
$r_{135} : \text{Raf}^* \text{-Phosp}_1 []_c \xrightarrow{c^{135}} \text{Raf}^* [\text{Phosp}_1]_c$	$k_{135} = 0,2$
$r_{136} : \text{Raf}^* \text{-Phosp}_1 []_c \xrightarrow{c^{136}} \text{Raf} [\text{Phosp}_1]_c$	$k_{136} = 1$

■ Activación de las proteínas MEK y ERK

Regla	Constante cinética
$\tau_{137} : \text{Raf}^* [\text{MEK}]_c \xrightarrow{c^{137}} \text{Raf}^* \text{-MEK} []_c$	$k_{137} = 0,0111$
$\tau_{138} : \text{Raf}^* \text{-MEK} []_c \xrightarrow{c^{138}} \text{Raf}^* [\text{MEK}]_c$	$k_{138} = 0,01833$
$\tau_{139} : \text{Raf}^* \text{-MEK} []_c \xrightarrow{c^{139}} \text{Raf}^* [\text{MEK-P}]_c$	$k_{139} = 3,5$
$\tau_{140} : \text{Raf}^* [\text{MEK-P}]_c \xrightarrow{c^{140}} \text{Raf}^* \text{-MEK-P} []_c$	$k_{140} = 0,0111$
$\tau_{141} : \text{Raf}^* \text{-MEK-P} []_c \xrightarrow{c^{141}} \text{Raf}^* [\text{MEK-P}]_c$	$k_{141} = 0,01833$
$\tau_{142} : \text{Raf}^* \text{-MEK-P} []_c \xrightarrow{c^{142}} \text{Raf}^* [\text{MEK-PP}]_c$	$k_{142} = 2,9$
$\tau_{143} : [\text{MEK-PP} + \text{Phosp}_2]_c \xrightarrow{c^{143}} [\text{MEK-PP-Phosp}_2]_c$	$k_{143} = 1,43 \times 10^{-2}$
$\tau_{144} : [\text{MEK-PP-Phosp}_2]_c \xrightarrow{c^{144}} [\text{MEK-PP} + \text{Phosp}_2]_c$	$k_{144} = 0,8$
$\tau_{145} : [\text{MEK-PP-Phosp}_2]_c \xrightarrow{c^{145}} [\text{MEK-P} + \text{Phosp}_2]_c$	$k_{145} = 0,058$
$\tau_{146} : [\text{MEK-P} + \text{Phosp}_2]_c \xrightarrow{c^{146}} [\text{MEK-P-Phosp}_2]_c$	$k_{146} = 2,5 \times 10^{-4}$
$\tau_{147} : [\text{MEK-P-Phosp}_2]_c \xrightarrow{c^{147}} [\text{MEK-P} + \text{Phosp}_2]_c$	$k_{147} = 0,5$
$\tau_{148} : [\text{MEK-P-Phosp}_2]_c \xrightarrow{c^{148}} [\text{MEK} + \text{Phosp}_2]_c$	$k_{148} = 0,058$
$\tau_{149} : [\text{MEK-PP} + \text{ERK}]_c \xrightarrow{c^{149}} [\text{MEK-PP-ERK}]_c$	$k_{149} = 1,1 \times 10^{-4}$
$\tau_{150} : [\text{MEK-PP-ERK}]_c \xrightarrow{c^{150}} [\text{MEK-PP} + \text{ERK}]_c$	$k_{150} = 0,033$
$\tau_{151} : [\text{MEK-PP-ERK}]_c \xrightarrow{c^{151}} [\text{MEK-PP} + \text{ERK-P}]_c$	$k_{151} = 16$
$\tau_{152} : [\text{MEK-PP} + \text{ERK-P}]_c \xrightarrow{c^{152}} [\text{MEK-PP-ERK-P}]_c$	$k_{152} = 1,1 \times 10^{-4}$
$\tau_{153} : [\text{MEK-PP-ERK-P}]_c \xrightarrow{c^{153}} [\text{MEK-PP} + \text{ERK-P}]_c$	$k_{153} = 0,033$
$\tau_{154} : [\text{MEK-PP-ERK-P}]_c \xrightarrow{c^{154}} [\text{MEK-PP} + \text{ERK-PP}]_c$	$k_{154} = 5,7$
$\tau_{155} : [\text{ERK-PP} + \text{Phosp}_3]_c \xrightarrow{c^{155}} [\text{ERK-PP-Phosp}_3]_c$	$k_{155} = 0,0145$
$\tau_{156} : [\text{ERK-PP-Phosp}_3]_c \xrightarrow{c^{156}} [\text{ERK-PP} + \text{Phosp}_3]_c$	$k_{156} = 0,6$
$\tau_{157} : [\text{ERK-PP-Phosp}_3]_c \xrightarrow{c^{157}} [\text{ERK-P} + \text{Phosp}_3]_c$	$k_{157} = 0,27$
$\tau_{158} : [\text{ERK-P} + \text{Phosp}_3]_c \xrightarrow{c^{158}} [\text{ERK-P-Phosp}_3]_c$	$k_{158} = 0,005$
$\tau_{159} : [\text{ERK-P-Phosp}_3]_c \xrightarrow{c^{159}} [\text{ERK-P} + \text{Phosp}_3]_c$	$k_{159} = 0,5$
$\tau_{160} : [\text{ERK-P-Phosp}_3]_c \xrightarrow{c^{160}} [\text{ERK} + \text{Phosp}_3]_c$	$k_{160} = 0,3$

7.3.2. Modelos de Sistemas P para el análisis de una Cascada de Señales EGFR

En esta sección se analiza el comportamiento del sistema y se estudia la robustez de la cascada de señales con relación al número de señales EGF en el entorno y al número de receptores EGFR en la superficie celular.

Una vez que se ha desarrollado un sistema P de especificación de una cascada de señales del EGFR, Π_{EGFR} , en esta sección se estudia el comportamiento del sistema bajo diferentes condiciones concretando valores a los parámetros asociados a la especificación, y se demuestra la robustez de la cascada de señales con relación al número de señales EGF en el entorno y al número de receptores EGFR en la superficie celular.

- Los multiconjuntos iniciales asociados al entorno $M_1 = \{M_1^{100}, M_1^{200}, M_1^{300}, M_1^{400}, M_1^{1000}, M_1^{2000}\}$ representarán las condiciones iniciales que corresponden a concentraciones de 100nM, 200nM, 300nM, 400nM, 1000nM y 2000nM de EGF en el entorno.
- Los multiconjuntos iniciales asociados a la superficie celular $M_2 = \{M_2^{100}, M_2^{1000}\}$ representarán las condiciones iniciales que corresponden a concentraciones de 100nM y 1000nM de receptores EGFR en dicha superficie. La misma concentración inicial de 200 nM de Ras-GDP se colocará en los multiconjuntos iniciales M_2^{100} and M_2^{1000} .
- Finalmente, existe un único multiconjunto inicial asociado al citoplasma, M_3 . que representará las condiciones iniciales que corresponden a concentraciones 250nM de Shc, 150nM de PLC $_{\gamma}$, 50nM de PI3K, 40nM de SOS, 80nM de Grb2, 100nM de TP $_1$, 450nM de TP $_2$, 450nM de TP $_3$, 125nM de TP $_4$, 80nM de Raf, 400nM de MEK, 400nM de ERK, 80nM de P $_1$, 80nM de P $_2$ y 300nM de P $_3$.

Mediante la combinación de estos parámetros, se obtiene una familia de 12 modelos de sistemas P , $\mathcal{F}_{EGFR}(\Pi_{EGFR}; ((M_1, M_2, M_3), C))$, que denotaremos por PSM_{EGFR}^{ij} , en donde $PSM_{EGFR}^{ij} = (\Pi_{EGFR}; (M_1^i, M_2^j, M_3), C)$.

- **Robustez del sistema con relación a las señales extracelulares EGF:**

Se obtiene analizando la evolución de los modelos

$$PSM_{EGFR}^{100,100}, PSM_{EGFR}^{200,100}, PSM_{EGFR}^{300,100}, PSM_{EGFR}^{400,100}, PSM_{EGFR}^{1000,100} \text{ and } PSM_{EGFR}^{2000,100}.$$

- **Efecto del número de receptores en la superficie celular**

Se analiza el efecto sobre la dinámica de la cascada de señales de una varización en el número de receptores en la superficie celular. Para ello se consideran los modelos $PSM_{EGFR}^{400,100}$ y $PSM_{EGFR}^{400,1000}$.

7.4. Modelado de la Apoptosis Inducida por FAS

Existen, básicamente, dos mecanismos de muerte celular: *necrosis* y *apoptosis*. La necrosis es una forma de muerte celular que suele ocurrir cuando las células son dañadas por heridas o lesiones produciéndose una rotura de la membrana plasmática que provoca una dispersión de los materiales intracelulares por el entorno. En cambio, la apoptosis es una forma de muerte celular que se produce como una secuencia *ordenada* de sucesos que culmina con el suicidio de la célula sin que el contenido intracelular de la célula llegue al entorno de la misma, evitando las inflamaciones típicas de la necrosis.

El término *apoptosis* (también denominado *muerte celular programada*) fue introducido por Kerr, Wyllie y Currie [67] para distinguir ambos tipos de muerte celular.

La apoptosis aparece de manera natural durante el desarrollo orgánico y en la eliminación de células que representen una amenaza para el organismo. Así por ejemplo, juega un papel importante en la homeostasis celular (en la médula ósea una ingente cantidad de leucocitos, aproximadamente $2 \cdot 20^6$ cada 24 horas). La apoptosis suele provocar una cascada de señales, a la vez intra y extracelular, que convergen en la activación de un grupo específico de proteínas denominadas *caspases*.

El mecanismo de la apoptosis está compuesto de condensación del contenido celular, fragmentación del ADN, rotura de la membrana nuclear y la formación de cuerpos apoptóticos que facilitan la fagocitación de la célula, proceso en el que pueden intervenir células vecinas. La apoptosis protege al resto del organismo de potenciales agentes dañinos, y su disfunción puede contribuir al desarrollo de enfermedades del sistema inmunológico y de cánceres. La apoptosis puede ser inducida a través de terapias anticancerígenas y por irradiación.

La *decisión* del suicidio de una célula puede provenir de la propia célula, del tejido circundante o bien del sistema inmunológico.

Existe diversas patologías asociadas a la apoptosis:

- Por inhibición: Cánceres, enfermedades autoinmunes, infecciones virales.
- Por sobreexpresión: SIDA, enfermedades neurodegenerativas, daño isquémico (infartos de miocardio, apoplejía, etc)

El Premio Nobel de Fisiología y Medicina del año 2002 fue otorgado a S. Brenner, H.R. Horvitz y J.E. Sulston por sus descubrimientos relativos a la regulación genética del desarrollo de órganos y de la apoptosis.

La familia de proteasas que regulan la apoptosis está clasificada en dos subgrupos. El primero consta de la caspasa 8, la caspasa 9 y la caspasa 10, y tienen una función iniciadora del proceso apoptótico. El segundo grupo tienen una función de efectores y contiene las caspasas 3, 6 y 7. Existen

otras moléculas con funciones efectoras, entre las que destaca la Apaf-1, que junto con el citocromo c estimula la producción de pro-caspasa 9.

Los otros reguladores de la apoptosis son la familia de los Bcl2 que, de acuerdo con su estructura molecular, la integran tres subgrupos: el primero tiene una función anti-apoptótica y la integran el Bcl2 y el Bcl-xL; los dos restantes subgrupos tienen funciones pro-apoptóticas y lo integran Bax, Bak, Bid y Bad.

La apoptosis puede ser disparada por una amplia variedad de estímulos, entre los que destacan el producido por un daño en el núcleo (*controlado* por la proteína p53), y las rutas de señales reguladas por la proteína FAS que parece ser uno de los procesos más relevantes para entender y combatir el cáncer, el SIDA, la apoplejía isquémica u otras enfermedades neurodegenerativas (Parkinson, Alzheimer, etc.).

La proteína FAS (también denominada CD95 o APO-1) es un receptor de la superficie celular que posee una región extracelular, un dominio sobre la propia membrana plasmática y una región intracelular. La apoptosis mediatizada por FAS comienza con el enlace del receptor FAS a su ligando FASL y finaliza con la fragmentación del ADN.

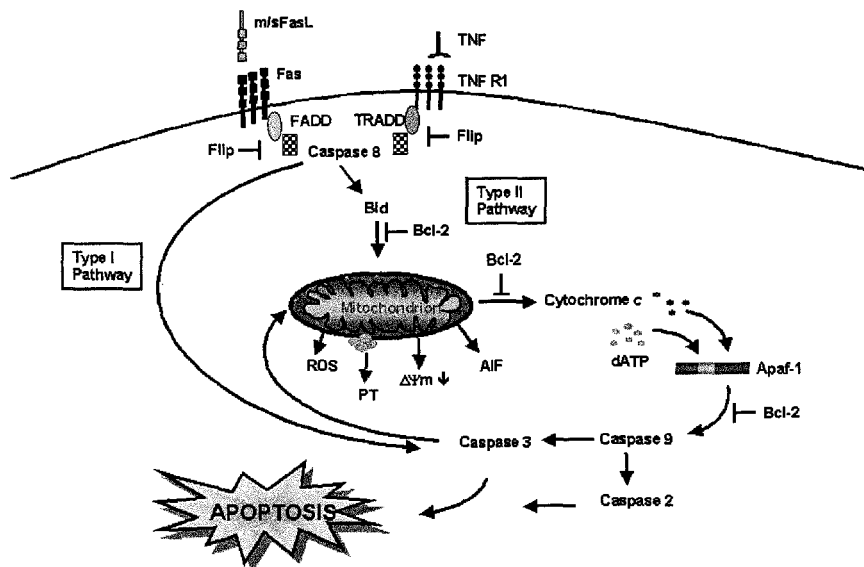


Figura 7.3: Caminos de la apoptosis regulada por FAS, de [2]

El ligando FASL es un trímero que al enlazar con el receptor FAS crea un complejo (clúster) que produce la activación del receptor (por medio de su trimerización). En el dominio intracelular, el receptor FAS contiene una secuencia denominada *dominio de muerte* (DD). El receptor activado recluta unas moléculas adaptadoras (como FADD, Fas-associating protein with death domain) que interactúan con el dominio de muerte del receptor FAS reclutando procaspasa 8 al complejo, la cual permite la liberación al citoplasma de la caspasa 8.

Finalmente, la caspasa 8 actúa como detonante de la activación de la caspasa 3 y lo puede hacer mediante dos caminos diferentes que han sido identificados por Scaffidi y otros [114]. Dichos caminos se denominan como tipo I (*camino de receptor de muerte*) y tipo II (*camino mitocondrial*). Ambos caminos divergen tras la activación de las caspasas iniciadoras y convergen en la activación de las caspasas ejecutoras de la apoptosis. Mediante el camino de tipo I, la caspasa 8 iniciadora se adhiere directamente a la procaspasa 3 activando la caspasa 3 ejecutora. El camino mitocondrial produce una cascada de señales bastante más complicada y está regulada por proteínas de la familia Bcl2: la caspasa 8 provoca una escisión en la proteína Bid (de la familia Bcl2), la cual produce unos poros en la mitocondria que permiten la liberación de numerosas proteínas, en particular de *citocromo c*, el

cual enlaza con la proteína Apaf-1 (apoptotic protease activating factor) activando la caspasa 9 que al adherirse a la procaspasa 3 activa la caspasa 3 que desencadena las últimas fases de la apoptosis.

Usando ODEs, en [55] se modelan los dos caminos activados por FAS a partir de la estimulación del FASL (ligando FAS) hasta la activación de la caspasa 3. En esta sección se presenta un modelo alternativo basado en sistemas P que permite mejorar una aproximación dada en [55], así como corroborar unas conjeturas formuladas en [114] y que no pudieron ser modeladas mediante ODEs.

7.4.1. Una Especificación de Sistema P de la Apoptosis Inducida por FAS

En esta sección se presenta una especificación de sistema P para la apoptosis mediada por FAS que se ha descrito anteriormente. Las regiones relevantes en la especificación son: el entorno etiquetado por e , en donde aparece las señales FASL; la superficie celular, etiquetada por s , en donde está situado el receptor FAS y tiene lugar la formación de complejos involucrados en el proceso; el citoplasma, etiquetado por c , en donde las cascadas tienen lugar, propiamente; y una región etiquetada por m que representará la mitocondria. La raíz de la estructura es el entorno, la superficie celular es hija del entorno, el citoplasma es hijo de la superficie celular y la mitocondria es hija del citoplasma.

En nuestro modelo hay 53 proteínas y complejos proteínicos que constituyen el alfabeto de trabajo O .

$$O = \{ \text{FASL, FAS, FASC, FADD, FASC-FADD, FASC-FADD}_2, \text{FASC-FADD}_3, \text{FASC-FADD}_2\text{-CASP8, FASC-FADD}_3\text{-CASP8, FASC-FADD}_2\text{-FLIP, FASC-FADD}_3\text{-FLIP, FASC-FADD}_2\text{-CASP8}_2, \text{FASC-FADD}_3\text{-CASP8}_2, \text{FASC-FADD}_2\text{-CASP8-FLIP, FASC-FADD}_3\text{-CASP8-FLIP, FASC-FADD}_2\text{-FLIP}_2, \text{FASC-FADD}_3\text{-FLIP}_2, \text{FASC-FADD-CASP8, FASC-FADD-FLIP, CASP8, FLIP, FASC-FADD}_3\text{-CASP8}_3, \text{FASC-FADD}_3\text{-CASP8}_2\text{-FLIP, FASC-FADD}_3\text{-CASP8-FLIP}_2, \text{FASC-FADD}_3\text{-FLIP}_3, \text{CASP8}_2^{P41}, \text{CASP8}_2^*, \text{CASP3, CASP8}_2^*\text{-CASP3, CASP3}^*, \text{CASP8}_2^*\text{-Bid, tBid, Bid, Bax, tBid-Bax, tBid-Bax}_2, \text{Smac, Smac}^*, \text{Cyto.c, Cyto.c}^*, \text{XIAP, Smac}^*\text{-XIAP, Apaf, Cyto.c}^*\text{-Apaf-ATP, CASP9, Cyto.c}^*\text{-Apaf-ATP-CASP9, Cyto.c}^*\text{-Apaf-ATP-CASP9}_2, \text{CASP9}^*, \text{CASP9}^*\text{-CASP3, CASP9-XIAP, CASP3}^*\text{-XIAP, Bcl2, Bcl2-Bax} \}$$

Las interacciones moleculares del sistema están especificadas a través de las siguientes reglas de reescritura:

Regla

Constante

$r_1 : FASL [FAS]_s \xrightarrow{c^1} [FASC]_s$	k_{1f}
$r_2 : [FASC]_s \xrightarrow{c^2} FASL [FASC]_s$	k_{1r}
$r_3 : FASC [FADD]_c \xrightarrow{c^3} FASC : FADD []_c$	k_{2f}
$r_4 : FASC-FADD []_c \xrightarrow{c^4} FASC [FADD]_c$	k_{2r}
$r_5 : FASC-FADD [FADD]_c \xrightarrow{c^5} FASC-FADD_2 []_c$	k_{2f}
$r_6 : FASC-FADD_2 []_c \xrightarrow{c^6} FASC-FADD [FADD]_c$	k_{2r}
$r_7 : FASC-FADD_2 [FADD]_c \xrightarrow{c^7} FASC-FADD_3 []_c$	k_{2f}
$r_8 : FASC-FADD_3 []_c \xrightarrow{c^8} FASC-FADD_2 [FADD]_c$	k_{2r}
$r_9 : FASC-FADD_2-CASP8 [FADD]_c \longrightarrow FASC-FADD_3-CASP8 []_c$	k_{2f}
$r_{10} : FASC-FADD_3-CASP8 []_c \xrightarrow{c^{10}} FASC-FADD_2-CASP8 [FADD]_c$	k_{2r}
$r_{11} : FASC-FADD_2-FLIP [FADD]_c \xrightarrow{c^{11}} FASC-FADD_3-FLIP []_c$	k_{2f}
$r_{12} : FASC-FADD_3-FLIP []_c \xrightarrow{c^{12}} FASC-FADD_2-FLIP [FADD]_c$	k_{2r}
$r_{13} : FASC-FADD_2-CASP8_2 [FADD]_c \xrightarrow{c^{13}} FASC-FADD_3-CASP8_2 []_c$	k_{2f}
$r_{14} : FASC-FADD_3-CASP8_2 []_c \xrightarrow{c^{14}} FASC-FADD_2-CASP8_2 [FADD]_c$	k_{2r}
$r_{15} : FASC-FADD_2-CASP8-FLIP [FADD]_c \xrightarrow{c^{15}} FASC-FADD_3-CASP8-FLIP []_c$	k_{2f}
$r_{16} : FASC-FADD_3-CASP8-FLIP []_c \xrightarrow{c^{16}} FASC-FADD_2-CASP8-FLIP [FADD]_c$	k_{2r}
$r_{17} : FASC-FADD_2-FLIP_2 [FADD]_c \xrightarrow{c^{17}} FASC-FADD_3-FLIP_2 []_c$	k_{2f}
$r_{18} : FASC-FADD_3-FLIP_2 []_c \xrightarrow{c^{18}} FASC-FADD_2-FLIP_2 [FADD]_c$	k_{2r}
$r_{19} : FASC-FADD-CASP8 [FADD]_c \xrightarrow{c^{19}} FASC-FADD_2-CASP8 []_c$	k_{2f}
$r_{20} : FASC-FADD_2-CASP8 []_c \xrightarrow{c^{20}} FASC-FADD-CASP8 [FADD]_c$	k_{2r}
$r_{21} : FASC-FADD-FLIP [FADD]_c \xrightarrow{c^{21}} FASC-FADD_2-FLIP []_c$	k_{2f}
$r_{22} : FASC-FADD_2-FLIP []_c \xrightarrow{c^{22}} FASC-FADD-FLIP [FADD]_c$	k_{2r}
$r_{23} : FASC-FADD_3 [CASP8]_c \xrightarrow{c^{23}} FASC-FADD_3-CASP8 []_c$	k_{2f}
$r_{24} : FASC-FADD_3-CASP8 []_c \xrightarrow{c^{24}} FASC-FADD_3 [CASP8]_c$	k_{2r}
$r_{25} : FASC-FADD_3 [FLIP]_c \xrightarrow{c^{25}} FASC-FADD_3-FLIP []_c$	k_{3f}
$r_{26} : FASC-FADD_3-FLIP []_c \xrightarrow{c^{26}} FASC-FADD_3 [FLIP]_c$	k_{3r}
$r_{27} : FASC-FADD_3-CASP8 [CASP8]_c \xrightarrow{c^{27}} FASC-FADD_3-CASP8_2 []_c$	k_{3f}
$r_{28} : FASC-FADD_3-CASP8_2 []_c \xrightarrow{c^{28}} FASC-FADD_3-CASP8 [CASP8]_c$	k_{3r}
$r_{29} : FASC-FADD_3-CASP8 [FLIP]_c \xrightarrow{c^{29}} FASC-FADD_3-CASP8-FLIP []_c$	k_{3f}
$r_{30} : FASC-FADD_3-CASP8-FLIP []_c \xrightarrow{c^{30}} FASC-FADD_3-CASP8 [FLIP]_c$	k_{3r}
$r_{31} : FASC-FADD_3-FLIP [CASP8]_c \xrightarrow{c^{31}} FASC-FADD_3-CASP8-FLIP []_c$	k_{3f}
$r_{32} : FASC-FADD_3-CASP8-FLIP []_c \xrightarrow{c^{32}} FASC-FADD_3-FLIP [CASP8]_c$	k_{3r}
$r_{33} : FASC-FADD_3-FLIP [FLIP]_c \xrightarrow{c^{33}} FASC-FADD_3-FLIP_2 []_c$	k_{3f}
$r_{34} : FASC-FADD_3-FLIP_2 []_c \xrightarrow{c^{34}} FASC-FADD_3-FLIP [FLIP]_c$	k_{3r}
$r_{35} : FASC-FADD_3-CASP8_2 [CASP8]_c \xrightarrow{c^{35}} FASC-FADD_3-CASP8_3 []_c$	k_{3f}
$r_{36} : FASC-FADD_3-CASP8_3 []_c \xrightarrow{c^{36}} FASC-FADD_3-CASP8_2 [CASP8]_c$	k_{3r}
$r_{37} : FASC-FADD_3-CASP8_2 [FLIP]_c \xrightarrow{c^{37}} FASC-FADD_3-CASP8_2-FLIP []_c$	k_{3f}
$r_{38} : FASC-FADD_3-CASP8_2-FLIP []_c \xrightarrow{c^{38}} FASC-FADD_3-CASP8_2 [FLIP]_c$	k_{3r}
$r_{39} : FASC-FADD_3-CASP8-FLIP [CASP8]_c \xrightarrow{c^{39}} FASC-FADD_3-CASP8_2-FLIP []_c$	k_{3f}
$r_{40} : FASC-FADD_3-CASP8_2-FLIP []_c \xrightarrow{c^{40}} FASC-FADD_3-CASP8-FLIP [CASP8]_c$	k_{3r}
$r_{41} : FASC-FADD_3-CASP8-FLIP [FLIP]_c \xrightarrow{c^{41}} FASC-FADD_3-CASP8-FLIP_2 []_c$	k_{3f}
$r_{42} : FASC-FADD_3-CASP8-FLIP_2 []_c \xrightarrow{c^{42}} FASC-FADD_3-CASP8-FLIP [FLIP]_c$	k_{3r}
$r_{43} : FASC-FADD_3-FLIP_2 [CASP8]_c \xrightarrow{c^{43}} FASC-FADD_3-CASP8-FLIP_2 []_c$	k_{3f}
$r_{44} : FASC-FADD_3-CASP8-FLIP_2 []_c \xrightarrow{c^{44}} FASC-FADD_3-FLIP_2 [CASP8]_c$	k_{3r}
$r_{45} : FASC-FADD_3-FLIP_2 [FLIP]_c \xrightarrow{c^{45}} FASC-FADD_3-FLIP_3 []_c$	k_{3f}

Regla	Constante
$r_{46} : FASC-FADD_3-FLIP_3 []_c \xrightarrow{c_{46}} FASC-FADD_3-FLIP_2 [FLIP]_c$	k_{3r}
$r_{47} : FASC-FADD_2 [CASP8]_c \xrightarrow{c_{47}} FASC-FADD_2-CASP8 []_c$	k_{3f}
$r_{48} : FASC-FADD_2-CASP8 []_c \xrightarrow{c_{48}} FASC-FADD_2 [CASP8]_c$	k_{3r}
$r_{49} : FASC-FADD_2 [FLIP]_c \xrightarrow{c_{49}} FASC-FADD_2-FLIP []_c$	k_{3f}
$r_{50} : FASC-FADD_2-FLIP []_c \xrightarrow{c_{50}} FASC-FADD_2 [FLIP]_c$	k_{3r}
$r_{51} : FASC-FADD_2-CASP8 [CASP8]_c \xrightarrow{c_{51}} FASC-FADD_2-CASP8_2 []_c$	k_{3f}
$r_{52} : FASC-FADD_2-CASP8_2 []_c \xrightarrow{c_{52}} FASC-FADD_2-CASP8 [CASP8]_c$	k_{3r}
$r_{53} : FASC-FADD_2-CASP8 [FLIP]_c \xrightarrow{c_{53}} FASC-FADD_2-CASP8 : FLIP []_c$	k_{3f}
$r_{54} : FASC-FADD_2-CASP8-FLIP []_c \xrightarrow{c_{54}} FASC-FADD_2-CASP8 [FLIP]_c$	k_{3r}
$r_{55} : FASC-FADD_2-FLIP [CASP8]_c \xrightarrow{c_{55}} FASC-FADD_2-CASP8-FLIP []_c$	k_{3f}
$r_{56} : FASC-FADD_2-CASP8-FLIP []_c \xrightarrow{c_{56}} FASC-FADD_2-FLIP [CASP8]_c$	k_{3r}
$r_{57} : FASC-FADD_2-FLIP [FLIP]_c \xrightarrow{c_{57}} FASC-FADD_2-FLIP_2 []_c$	k_{3f}
$r_{58} : FASC-FADD_2-FLIP_2 []_c \xrightarrow{c_{58}} FASC-FADD_2-FLIP [FLIP]_c$	k_{3r}
$r_{59} : FASC-FADD [CASP8]_c \xrightarrow{c_{59}} FASC-FADD-CASP8 []_c$	k_{3f}
$r_{60} : FASC-FADD-CASP8 []_c \xrightarrow{c_{60}} FASC-FADD [CASP8]_c$	k_{3r}
$r_{61} : FASC-FADD [FLIP]_c \xrightarrow{c_{61}} FASC-FADD-FLIP []_c$	k_{3f}
$r_{62} : FASC-FADD-FLIP []_c \xrightarrow{c_{62}} FASC-FADD [FLIP]_c$	k_{3r}
$r_{63} : FASC-FADD_2-CASP8_2 []_c \xrightarrow{c_{63}} FASC-FADD_2 [CASP8_2^{P41}]_c$	k_4
$r_{64} : FASC-FADD_3-CASP8_3 []_c \xrightarrow{c_{64}} FASC-FADD_3-CASP8 [CASP8_2^{P41}]_c$	k_4
$r_{65} : FASC-FADD_3-CASP8_2-FLIP []_c \xrightarrow{c_{65}} FASC-FADD_3-FLIP [CASP8_2^{P41}]_c$	k_4
$r_{66} : FASC-FADD_3-CASP8_2 []_c \xrightarrow{c_{66}} FASC-FADD_3 [CASP8_2^{P41}]_c$	k_4
$r_{67} : [CASP8_2^{P41}]_c \xrightarrow{c_{67}} [CASP8_2^*]_c$	k_5
$r_{68} : [CASP8_2^* + CASP3]_c \xrightarrow{c_{68}} [CASP8_2^*-CASP3]_c$	k_{6f}
$r_{69} : [CASP8_2^*-CASP3]_c \xrightarrow{c_{69}} [CASP8_2^* + CASP3]_c$	k_{6r}
$r_{70} : [CASP8_2^* + CASP3^*]_c \xrightarrow{c_{70}} [CASP8_2^*-CASP3]_c$	k_7
$r_{71} : [CASP8_2^* + Bid]_c \xrightarrow{c_{71}} [CASP8_2^*-Bid]_c$	k_{8f}
$r_{72} : [CASP8_2^*-Bid]_c \xrightarrow{c_{72}} [CASP8_2^* + Bid]_c$	k_{8r}
$r_{73} : [CASP8_2^* + tBid]_c \xrightarrow{c_{73}} [CASP8_2^*-Bid]_c$	k_7
$r_{74} : [tBid + Bax]_c \xrightarrow{c_{74}} [tBid-Bax]_c$	k_{9f}
$r_{75} : [tBid-Bax]_c \xrightarrow{c_{75}} [tBid + Bax]_c$	k_{9r}
$r_{76} : [tBid-Bax + Bax]_c \xrightarrow{c_{76}} [tBid-Bax_2]_c$	k_{9f}
$r_{77} : [tBid-Bax_2]_c \xrightarrow{c_{77}} [tBid-Bax + Bax]_c$	k_{9r}
$r_{78} : tBid-Bax_2 [Smac]_m \xrightarrow{c_{78}} Smac^* []_m$	k_{10}
$r_{79} : tBid-Bax_2 [Cyto.c]_m \xrightarrow{c_{79}} Cyto.c^* []_m$	k_{10}
$r_{80} : [Smac^* + XIAP]_c \xrightarrow{c_{80}} [Smac^*-XIAP]_c$	k_{11f}
$r_{81} : [Smac^*-XIAP]_c \xrightarrow{c_{81}} [Smac^* + XIAP]_c$	k_{11r}
$r_{82} : [Cyto.c^* + Apaf]_c \xrightarrow{c_{82}} [Cyto.c^*-Apaf-ATP]_c$	k_{12f}
$r_{83} : [Cyto.c^*-Apaf-ATP]_c \xrightarrow{c_{83}} [Cyto.c^* + Apaf]_c$	k_{12r}
$r_{84} : [Cyto.c^*-Apaf-ATP + CASP9]_c \xrightarrow{c_{84}} [Cyto.c^*-Apaf-ATP-CASP9]_c$	k_{13f}
$r_{85} : [Cyto.c^*-Apaf-ATP-CASP9]_c \xrightarrow{c_{85}} [Cyto.c^*-Apaf-ATP + CASP9]_c$	k_{13r}
$r_{86} : [Cyto.c^*-Apaf-ATP-CASP9 + CASP9]_c \xrightarrow{c_{86}} [Cyto.c^*-Apaf-ATP-CASP9_2]_c$	k_{14f}
$r_{87} : [Cyto.c^*-Apaf-ATP-CASP9_2]_c \xrightarrow{c_{87}} [Cyto.c^*-Apaf-ATP-CASP9 + CASP9]_c$	k_{14r}

Regla	Constante
$r_{88} : [Cyto.c^*-Apaaf-ATP-CASP9_2]_c \xrightarrow{c_{88}} [Cyto.c^*-Apaaf-ATP-CASP9 + CASP9^*]_c$	k_{15}
$r_{89} : [CASP9^* + CASP3]_c \xrightarrow{c_{89}} [CASP9^*-CASP3]_c$	k_{16f}
$r_{90} : [CASP9^*-CASP3]_c \xrightarrow{c_{90}} [CASP9^* + CASP3]_c$	k_{16r}
$r_{91} : [CASP9^*-CASP3]_c \xrightarrow{c_{91}} [CASP9^* + CASP3^*]_c$	k_{17}
$r_{92} : [CASP9 + XIAP]_c \xrightarrow{c_{92}} [CASP9-XIAP]_c$	k_{18f}
$r_{93} : [CASP9-XIAP]_c \xrightarrow{c_{93}} [CASP9 + XIAP]_c$	k_{18r}
$r_{94} : [CASP3^* + XIAP]_c \xrightarrow{c_{94}} [CASP3^*-XIAP]_c$	k_{19f}
$r_{95} : [CASP3^*-XIAP]_c \xrightarrow{c_{95}} [CASP3^* + XIAP]_c$	k_{19r}
$r_{96} : Bax [Bcl2]_m \xrightarrow{c_{96}} [Bcl2-Bax]_m$	k_{20f}
$r_{97} : [Bcl2-Bax]_m \xrightarrow{c_{97}} Bax [Bcl2]_m$	k_{20r}
$r_{96'} : Bid [Bcl2]_m \xrightarrow{c_{96'}} [Bcl2-Bid]_m$	k_{20f}
$r_{97'} : [Bcl2-Bid]_m \xrightarrow{c_{97'}} Bid [Bcl2]_m$	k_{20r}
$r_{96''} : tBid [Bcl2]_m \xrightarrow{c_{96''}} [Bcl2-tBid]_m$	k_{20f}
$r_{97''} : [Bcl2-tBid]_m \xrightarrow{c_{97''}} tBid [Bcl2]_m$	k_{20r}

Las constantes cinéticas asociadas a las reglas son las siguientes:

$k_{1f} = 9,09E - 05 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{11r} = 2,21E - 03 \text{ s}^{-1}$
$k_{1r} = 1,00E - 04 \text{ s}^{-1}$	$k_{12f} = 2,78E - 07 \text{ nM}^{-1} \text{ s}^{-1}$
$k_{2f} = 5,00E - 04 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{12r} = 5,70E - 03 \text{ s}^{-1}$
$k_{2r} = 0,2 \text{ s}^{-1}$	$k_{13f} = 2,84E - 04 \text{ nM}^{-1} \text{ s}^{-1}$
$k_{3f} = 3,50E - 03 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{13r} = 0,07493 \text{ s}^{-1}$
$k_{3r} = 0,018 \text{ s}^{-1}$	$k_{14f} = 4,41E - 04 \text{ nM}^{-1} \text{ s}^{-1}$
$k_4 = 0,3 \text{ s}^{-1}$	$k_{14r} = 0,1 \text{ s}^{-1}$
$k_5 = 0,1 \text{ s}^{-1}$	$k_{15} = 0,7 \text{ s}^{-1}$
$k_{6f} = 1,00E - 05 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{16f} = 1,96E - 05 \text{ nM}^{-1} \text{ s}^{-1}$
$k_{6r} = 0,06 \text{ s}^{-1}$	$k_{16r} = 0,05707 \text{ s}^{-1}$
$k_7 = 0,1 \text{ s}^{-1}$	$k_{17} = 4,8 \text{ s}^{-1}$
$k_{8f} = 5,00E - 03 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{18f} = 1,06E - 04 \text{ nM}^{-1} \text{ s}^{-1}$
$k_{8r} = 0,005 \text{ s}^{-1}$	$k_{18r} = 1,00E - 03 \text{ s}^{-1}$
$k_{9f} = 2,00E - 04 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{19f} = 2,47E - 03 \text{ nM}^{-1} \text{ s}^{-1}$
$k_{9r} = 0,02 \text{ s}^{-1}$	$k_{19r} = 2,40E - 03 \text{ s}^{-1}$
$k_{10} = 1,00E - 03 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{20f} = 2,00E - 03 \text{ nM}^{-1} \text{ s}^{-1}$
$k_{11f} = 7,00E - 03 \text{ nM}^{-1} \text{ s}^{-1}$	$k_{20r} = 0,02 \text{ s}^{-1}$

7.4.2. Análisis de la Apoptosis Inducida por FAS usando Modelos de Sistemas P

De manera similar a como se ha hecho en los casos previos, una vez que se ha desarrollado una especificación de nuestro sistema P para la apoptosis mediatizada por la proteína FAS, Π_{FAS} , es necesario identificar los parámetros asociados a la especificación y asignarles valores que nos permita realizar un análisis del sistema.

En primer lugar se estudia la activación de la caspasa 3 y comparamos los resultados obtenidos en nuestro modelo usando el algoritmo determinista con tiempo de espera, con los resultados experimentales y los deducidos del modelo basad en ODEs en [55].

Los parámetros correspondientes a los multiconjuntos iniciales son calculados de las condiciones iniciales del sistema de ecuaciones diferenciales presentado en [55].

$$\begin{aligned}
M_1 &= (e, FASL^{12500}, \lambda) \\
M_2 &= (s, FAS^{6023}, \lambda) \\
M_3 &= (c, FADD^{10040} + CASP8^{20074} + FLIP^{48786} + CASP3^{120460} + Bid^{15057} + \\
&\quad Bax^{50189} + XIAP^{18069} + Apaf^{60230} + CASP9^{12046}, \lambda) \\
M_4 &= (m, Smac^{60230} + Cyto.c^{60230} + Bcl2^{45172}, \lambda)
\end{aligned} \tag{7.2}$$

Las constantes de las reglas son calculadas a partir de las constantes macroscópicas, excepto para las reglas $r_{96'}$, $r_{97'}$, $r_{96''}$ y $r_{97''}$ cuyas constantes serán puestas a cero en el primer experimento. Este primer conjunto de constantes se denota por C^a .

En el modelo basado en ODEs [55], la caspasa 3 se activó al cabo de cuatro horas. Este resultado era aceptable en relación con los datos experimentales (6 horas). No obstante, en nuestro modelo la activación se consiguió a las 7 horas, mejorando la aproximación del modelo basado en ODEs.

Existen células (como los timocitos y los fibroblastos) que no son sensibles a una sobreexpresión del Bcl2, como se cita en [114]. En estas células, la caspasa 8 activa directamente la caspasa 3.

En [114], Scaffidi y otros sugirieron que el tipo de camino apoptótico regulado por FAS se elegía en función de la concentración de caspasa 8 activada a través del ligando FASL. Si la concentración de la caspasa 8 activada es alta, entonces la caspasa 3 se activa directamente; caso contrario, se elige el camino tipo II a fin de que a través de la mitocondría se amplifique la señal de muerte y se pueda producir la muerte celular.

Para chequear esta hipótesis, el valor de la concentración inicial de la caspasa 8 se incrementó 20 veces en relación con su valor base; el resto de los valores permanecen inalterados. Ello provocó una disminución drástica en el tiempo de activación de la caspasa 3, de acuerdo con los datos experimentales. Los resultados que hemos obtenido en nuestro modelo de sistemas P coincide con los datos experimentales ([114]) y con los obtenidos por la simulación ODE ([55]).

En cambio, analizando la sensibilidad de la activación a la concentración inicial de Bcl2, se comprobó que al incrementar concentración inicial de Bcl2 entre 10 y 100 veces respecto del valor base, el tiempo de activación de la caspasa 3 no sufría variaciones ostensibles, probándose que la activación de la caspasa 3 no es sensible a los cambios de concentración de Bcl2 (cuando el camino tipo I es elegido).

Se sabe que la familia de proteínas Bcl2 permite bloquear el camino apoptótico mitocondrial, sin embargo, no se conoce con exactitud el mecanismo por el cual realiza dicho bloqueo. En [85] y [127] se sugieren distintos mecanismos posibles para bloquear el camino tipo II: mediante enlaces de Bcl2 con (a) Bax, (b) Bid, (c) tBid, o (d) a la vez con Bax y tBid.

A continuación, analizamos la cinética de la activación de la caspasa 3 en los posibles mecanismos de bloqueo de camino tipo II. Para ello, hemos de considerar un modelo de sistemas P para cada uno de los posibles mecanismos. Y aquí se pone de manifiesto uno de los aspectos más interesantes del modelado en el marco de la computación celular con membranas, la *modularidad*: pequeños cambios en la cascada de señales produce pequeños cambios en el diseño del sistema P que lo modela.

Los cuatro sistemas P distintos a considerar, únicamente se diferenciarán en el conjunto de reglas:

- $r_1, \dots, r_{95}, r_{96}, r_{97}$ para modelar el caso (a).
- $r_1, \dots, r_{95}, r'_{96}, r'_{97}$ para modelar el caso (b).

- $r_1, \dots, r_{95}, r''_{96}, r''_{97}$ para modelar el caso (c).
- $r_1, \dots, r_{97}, r''_{96}, r''_{97}$ para modelar el caso (d).

Todas las restantes reglas son iguales para todos los casos (ver apéndice).

La dinámica de la activación de la caspasa 3 se estudia nuevamente variando el valor de la concentración inicial de Bcl2, 10 veces o 100 veces respecto el valor base. Como consecuencia de la simulación ODE se concluyó que el enlace de Bcl2 con, a la vez, Bax y tBid, era el mecanismo más eficiente para bloquear el camino tipo II. Los mismos resultados han sido obtenidos usando el modelado y simulación con sistemas P.

La tabla siguiente muestra un resumen de la sensibilidad de la activación de la caspasa 3 a la concentración de la proteína Bcl2, en función del camino apoptótico seleccionado.

	Activación Caspasa 3 (con sobreexpresión de Bcl2)
Type I (camino receptor de muerte)	<i>Insensible</i>
Type II (camino mitocondrial)	<i>Sensible</i>

Capítulo 8

Un Modelo del Quorum Sensing

Las bacterias se consideran organismos unicelulares independientes. Sin embargo, se ha observado que ciertas bacterias, como la bacteria marina *Vibrio fischeri*, exhiben comportamientos coordinados que permiten a toda una colonia regular la expresión de ciertos genes de forma coordinada, dependiendo del tamaño de la población. Este sistema de regulación de genes dependiente de la densidad celular se denomina *Quorum Sensing* (percepción o detección de quorum) [128]. El *Quorum sensing* es un mecanismo de control de expresión de genes en función de la densidad celular y ha sido descrita como la *historia* más importante de la microbiología molecular de la última década [51]. Este fenómeno permite a una colonia de bacterias dar una respuesta coordinada cuando se verifican ciertas condiciones.

- Emisión de luminiscencia (*Vibrio Fischeri*)
- Secreción de sustancias viscosas provocando infecciones (*Pseudomona aeruginosa*, Gram negativa)
- Liberación de toxinas que matan a organismos que servirán de alimento (*Pfiesteria*)

8.1. Quorum Sensing en *Vibrio Fischeri*

La bacteria marina *Vibrio fischeri* puede vivir libremente formando parte del plancton, o bien en estado de simbiosis en un calamar (en peces o en algunos invertebrados).

Dicha bacteria tiene la capacidad de producir luz mediante la activación de unos genes de luminiscencia. Esto nunca sucede cuando vive en el plancton y puede ocurrir cuando vive en estado de simbiosis. [110].

El Quorum Sensing en *Vibrio fischeri* se basa en la síntesis, acumulación y envío al medio de una señal OHHL. Este fenómeno comienza propiamente con la actuación de una región del ADN de la bacteria (*LuxBox*) que activa la producción de dos proteínas, *LuxR* y *LuxI*. La primera de ellas actúa de receptor intermembranal y la segunda permite la síntesis de una molécula señalizadora, *OHHL*. Estas señales se difunden al entorno a través de la pared celular y pueden entrar en la bacteria desde el entorno por el mismo procedimiento.

Cuando la bacteria percibe que el número de señales existente en su interior supera un cierto umbral, entonces el receptor *LuxR* se une con la señal *OHHL* formando un complejo, *LuxR.OHHL*, que actuará como *factor de transcripción* enlazando con la región *LuxBox*, provocando la *excitación* de la bacteria y dando lugar a una producción masiva de señales y de receptores, a la vez que induce la activación de un pequeño clúster de cinco genes, encargados de la producción de luz (ver figura 8.1).

Si la densidad de bacterias es baja, las señales se producen a niveles basales, caso contrario, la señal se acumula en el entorno siendo capaz de interactuar con la proteína LuxR para formar un complejo, que enlaza con la región Lux Box provocando la transcripción de los genes de luminiscencia.

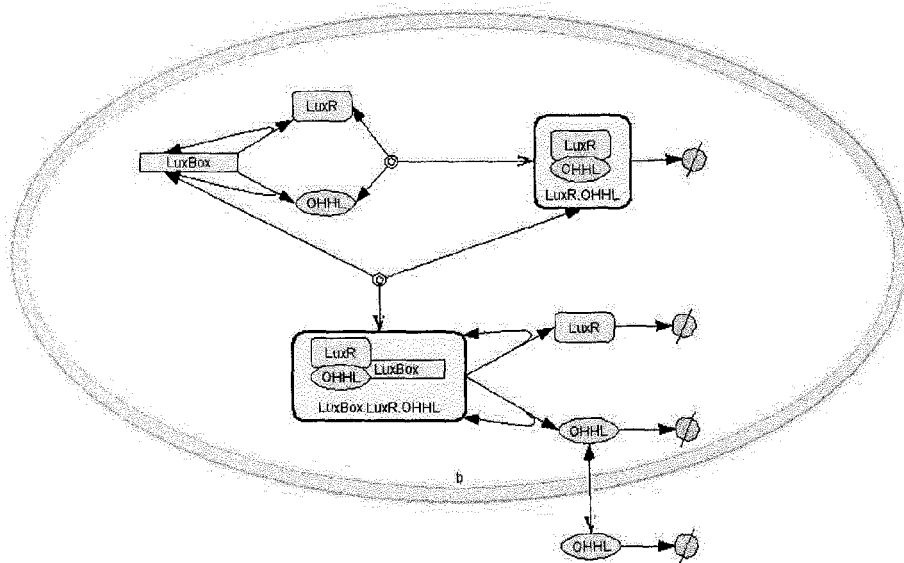


Figura 8.1: Quorum sensing en la bacteria Vibrio Fischeri

8.2. Una Especificación de Sistema P del Quorum Sensing en Vibrio Fischeri

En esta sección se presenta una especificación mediante sistemas P del sistema del quorum sensing en Vibrio fischeri, en el que existen dos regiones relevantes, que denominaremos *entorno* y *bacteria*. Cada bacteria de la colonia se especificará usando una membrana y se introduce una variante de sistemas P en la que el entorno externo se divide en trozos a fin de que cada uno de esos trozos se pueda considerar como una región *bien mezclada*. Esta estructura se denomina *multientorno*.

Definición 8.2.1 (multientorno) *Un multientorno es una colección de membranas denominadas entornos y enlaces de comunicación entre ellas. Formalmente, un multientorno es un grafo, $G = (V, S)$, cuyos nodos V son membranas representando entornos. Estos entornos están conectados de acuerdo con las aristas de S que define dichos enlaces.*

Cada entorno contendrá un multiconjunto de objetos que representan entidades moleculares, y un número de sistemas P de especificación que representan las bacterias individuales localizadas en la región del sistema representada por ese entorno. Además, cada entorno tiene asociado un conjunto de reglas del siguiente tipo:

- Las entidades moleculares del entorno pueden interactuar produciendo nuevas moléculas.

$$[obj_1]_l \xrightarrow{c} [obj_2]_l \quad (8.1)$$

Estas reglas únicamente pueden cambiar el contenido del entorno. Con el fin de poder aplicar la teoría de Gillespie, una propensidad será asociada a cada una de esas reglas. En este caso,

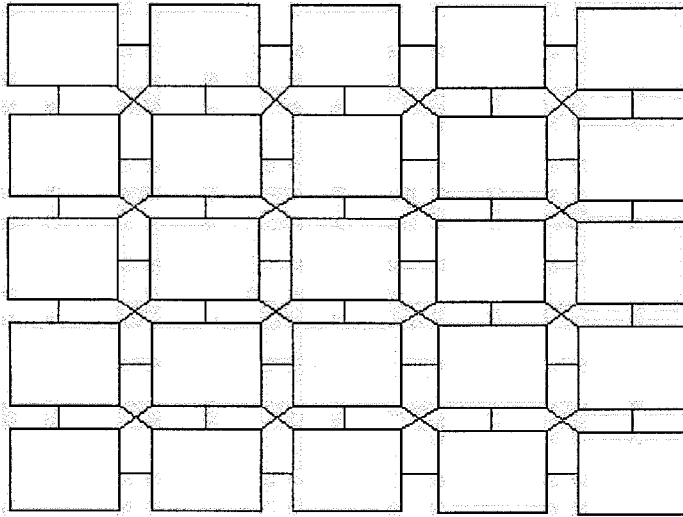


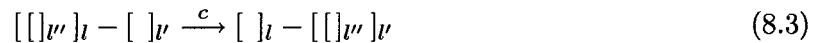
Figura 8.2: Un multientorno con 25 entornos.

la propensidad es calculada multiplicando la constante estocástica c asociada específicamente a esa regla, por el número de posibles combinaciones de objetos en el multiconjunto obj_1 .

- En el caso de movimiento de diferentes sustancias de un entorno a otro (conectado con él) se usa el siguiente tipo de reglas:



- Además de multiconjunto de objetos que representan moléculas, pueden existir en el entorno un cierto número de copias de sistemas P que representan individuos de la colonia. Estos individuos pueden desplazarse de un entorno a otro.



Finalmente, para estudiar una colonia de individuos cada uno de ellos es representado por un sistema P que es distribuido aleatoriamente en los entornos del *multientorno*.

Definición 8.2.2 (Population P system Specification) Una población de sistemas P de especificación es una tupla:

$$P\Pi = (\Sigma, H, G, E_1, \dots, E_m, R_{l_1}, \dots, R_{l_t}, \Pi_1, \dots, \Pi_p, k_1, \dots, k_p) \quad (8.4)$$

donde:

- Σ es un alfabeto finito de objetos.
- $G = (V, S)$ es un grafo cuyos nodos $V = 1, \dots, m$ son membranas que representan entornos etiquetados con elementos de $H = \{l_1, \dots, l_t\}$ y cuyas aristas, S , definen cómo los entornos están conectados.

- $E_j = (l_j, w_j)$, para cada $1 \leq j \leq m$, es la configuración inicial del entorno j , con $l_j \in H$ y $w_j \in \Sigma^*$.
- R_{l_1}, \dots, R_{l_t} son conjuntos finitos de reglas de la forma (8.1), (8.2) y (8.3) asociadas a los entornos l_1, \dots, l_t .
- Π_1, \dots, Π_p son especificaciones de sistemas P que describen los diferentes tipos de individuos de la población.
- $k_1, \dots, k_p \in \mathbb{N}$ es el número de copias de sistemas Π_1, \dots, Π_p que son distribuidos aleatoriamente en los entornos de la configuración inicial del sistema.

A cada población de sistemas P de especificación se le asocia unos conjuntos de parámetros que permiten obtener los distintos modelos de sistemas P .

En el caso de estudio de este capítulo, el Quorum Sensing de la bacteria *Vibrio Fischeri*, se considera:

$$P\Pi_{VF} = (\Sigma, \{e\}, G, E_1, \dots, E_{25}, R_e, \Pi_{VF}, N) \quad \text{en donde:} \quad (8.5)$$

1. El alfabeto Σ contiene un único objeto, OHHL, que representa la señal 3-oxo-C6-HSL.
2. Todos los entornos tienen la misma etiqueta e .
3. El medio en el que está la colonia se representa por un multientorno G que consta de 25 entornos conectados segun la figura 8.2.
4. Los multiconjuntos iniciales, E_1, \dots, E_{25} describirán el número inicial de señales en cada entorno.
5. El sistema Π_{VF} representa una bacteria *Vibrio fischeri*.

$$\Pi_{VF} = (O, \{b\}, [\], M_1, R_b) \quad \text{donde:} \quad (8.6)$$

- a) $O = \{\text{OHHL}, \text{LuxR}, \text{LuxR-OHHL}, \text{LuxBox}, \text{LuxR-OHHL-LuxBox}\}$
- b) Cada bacteria está etiquetada por b .
- c) El multiconjunto inicial M_1 únicamente contiene objetos LuxBox que permitirán la producción de señales, OHHL, y de moléculas proteínicas, LuxR.
- d) El conjunto de reglas, R_b , es el siguiente:

$$\begin{aligned}
r_1 &: [\text{LuxBox}]_b \xrightarrow{c_1} [\text{LuxBox} + \text{OHHL}]_b \\
r_2 &: [\text{LuxBox}]_b \xrightarrow{c_2} [\text{LuxBox} + \text{LuxR}]_b \\
r_3 &: [\text{LuxR} + \text{OHHL}]_b \xrightarrow{c_3} [\text{LuxR-OHHL}]_b \\
r_4 &: [\text{LuxR-OHHL}]_b \xrightarrow{c_4} [\text{LuxR} + \text{OHHL}]_b \\
r_5 &: [\text{LuxR-OHHL} + \text{LuxBox}]_b \xrightarrow{c_5} [\text{LuxR-OHHL-LuxBox}]_b \\
r_6 &: [\text{LuxR-OHHL-LuxBox}]_b \xrightarrow{c_6} [\text{LuxR-OHHL} + \text{LuxBox}]_b \\
r_7 &: [\text{LuxR-OHHL-LuxBox}]_b \xrightarrow{c_7} [\text{LuxR-OHHL-LuxBox} + \text{OHHL}]_b \\
r_8 &: [\text{LuxR-OHHL-LuxBox}]_b \xrightarrow{c_8} [\text{LuxR-OHHL-LuxBox} + \text{LuxR}]_b \\
r_9 &: [\text{OHHL}]_b \xrightarrow{c_9} \text{OHHL} [\]_b \\
r_{10} &: [\text{OHHL}]_b \xrightarrow{c_{10}} [\]_b \\
r_{11} &: [\text{LuxR}]_b \xrightarrow{c_{11}} [\]_b
\end{aligned}$$

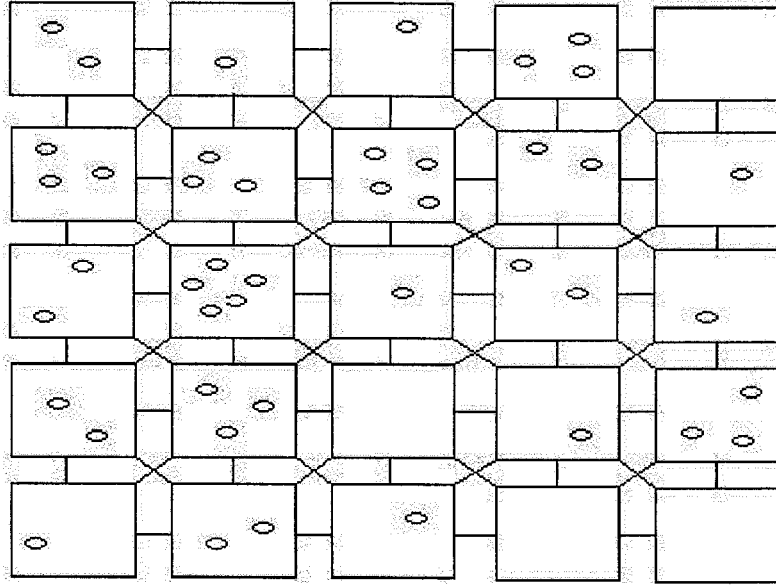


Figura 8.3: Una configuración inicial de una especificación de sistema P

$$r_{12} : [\text{LuxR-OHHL}]_b \xrightarrow{c_{12}} []_b$$

6. Todos los entornos tienen asociados el mismo conjunto de reglas R_e :

$$r_{13} : \text{OHHL} []_b \xrightarrow{c_{13}} [\text{OHHL}]_b$$

$$r_{14} : [\text{OHHL}]_e \xrightarrow{c_{14}} []_e$$

$$r_{15} : [\text{OHHL}]_e - []_e \xrightarrow{c_{15}} []_e - [\text{OHHL}]_e$$

$$r_{16} : [[]_b]_e - []_e \xrightarrow{c_{16}} []_e - [[]_b]_e$$

7. Finalmente, las N copias de Π_{VF} en el multientorno G son distribuidas aleatoriamente. Un ejemplo de una colonia de 44 bacterias se ilustra en la figura 8.3.

8.3. Modelos de Sistemas P del Quorum Sensing en *Vibrio Fischeri*

Los modelos de una especificación de una población de sistemas P se obtienen de la misma forma que en el caso de un sistema P ordinario. Recordemos que los parámetros asociados a una especificación de una población de sistemas P son:

- Los multiconjuntos iniciales asociados al entorno. En los modelos del Quorum Sensing en *Vibrio Fischeri* desarrollados en esta sección, todos los entornos estarán vacíos y etiquetados por e ; es decir, $E_1 = \dots = E_{25} = \mathbb{E}_0 = (e, \lambda)$.

- Las constantes estocásticas de las reglas, $\mathbb{C}_e = \{c_{13}, c_{14}, c_{15}, c_{16}\}$, del conjunto R_e , asociado a los entornos. En nuestros modelos, estas constantes tienen asociados valores fijos, $c_{13} = 1$, $c_{14} = 5$, $c_{15} = 8$ and $c_{16} = 2$.
- Los parámetros correspondientes a la especificación del sistemas P de la bacteria *Vibrio Fischeri*, Π_{VF} :

Recall that according to definition 3.2.2 the parameters of a simple P system specification are the initial multisets and the rule constants.

Regarding the initial multiset associated with the membrane describing a bacterium, recall that we are interested in examining how bacteria communicate to coordinate their behaviours and how the population moves from a downregulated state, where the protein and the signal are produced at basal rates, to an upregulated state where there is a massive production of signals and proteins. Therefore,

En el multiconjunto inicial M_1 sólo existe objetos LuxBox start the production of the signal OHHL and protein LuxR at basal rates, $M_0 = \text{LuxBox}$.

Regarding the stochastic constants associated with the rules in R_b , $\mathbb{C}_b = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9, c_{10}, c_{11}, c_{12}, c_{13}, c_{14}, c_{15}, c_{16}\}$. Where possible we have used estimates derived from the literature [34, 47, 48, 118, 132]. For parameter values not available from these sources we used a "trial-and-error" approach, making an initial "guess" at the values of the missing constants and then comparing the resulting behaviour with known properties of the system. If they did not match the unknown constants were then adjusted systematically, one parameter at a time.

By following this process the following set of parameters were chosen: $c_1 = 2$, $c_2 = 2$, $c_3 = 9$, $c_4 = 1$, $c_5 = 10$, $c_6 = 2$, $c_7 = 250$, $c_8 = 200$, $c_9 = 50$, $c_{10} = 30$, $c_{11} = 20$, $c_{12} = 20$, $c_{13} = 1$, $c_{14} = 5$, $c_{15} = 8$, $c_{16} = 2$. These values have been set such that the degradation rates compensate the basal production of the signal and the protein and such that the production rates when the regulatory region is occupied produce a massive increase in the transcription of the signal and the protein.

- Finally, the last parameter is N , the number of copies of the P system specification Π_{VF} distributed randomly in the environments. This parameter constitutes the only variable parameter of the system in contrast to the previous ones that are assigned fixed values. This parameter represents the size of the colony of *Vibrio fischeri* bacteria and will allow us to study the behaviour of colonies of different sizes.

Summing up the population P system models that will be used in this section consists in the following construct:

En esta sección se estudia la conducta emergente del sistema para dos colonias de diferente tamaño, y se analiza cómo las bacterias pueden percibir la densidad celular del sistema y producir luz sólo cuando esa densidad supera un cierto umbral.

El modelo considerado es el siguiente:

$$PPSM_{VF}(N) = (P\Pi_{VF}; (\mathbb{E}_0, \dots, \mathbb{E}_0), \mathbb{C}_e, (\Pi_{VF}; M_0, \mathbb{C}_b), N) \quad (8.7)$$

Este modelo ha sido representado en SBML (Systems Biology Markup Language) usando CellDesigner, un editor de diagramas estructurado para la representación de sistemas reguladores de genes y redes bioquímicas [37].

Las simulaciones realizadas han sido ejecutadas usando un programa escrito en C con un fichero de entrada SBML que especifica nuestro modelo.

- En primer lugar se ha considerado una colonia de 100 bacterias y hemos analizado la evolución en el tiempo del número de bacterias *quoradas*, así como el número de señales en el entorno.

Nuestro modelo permite realizar un seguimiento a individuos de la colonia. Por ello, se ha estudiado la correlación entre el número de señales dentro de ella y el hecho de que la bacteria esté quorada, en tres bacterias de la colonia.

También se ha estudiado la evolución del número de reglas aplicadas (en particular, las reglas que representan la producción basal y aquellas que representan la producción de señales y proteínas debido al enlace del complejo a la zona LuxBox.

- Para analizar cómo las bacterias perciben la densidad celular en el caso de colonias pequeñas, hemos simulado el comportamiento de una colonia de diez bacterias. Se ha constatado cómo una de esas bacterias ha pasado a un estado de alta regulación, tomando una decisión errónea. No obstante, ella misma se ha dado cuenta de su error y ha pasado a un estado de baja regulación.

Parte IV

Conclusiones

Capítulo 9

Conclusiones

De acuerdo con la motivación inicial [89], el objetivo de los sistemas P no es, propiamente, proporcionar una información comprensible y un modelo exacto de la célula viva, sino explorar la naturaleza computacional de varios rasgos de las membranas biológicas. De hecho, la mayor parte de las investigaciones sobre sistemas P se concentran en el estudio de variantes simplificadas de los sistemas de membranas originales para tratar de mostrar si son *computacionalmente completos* (es decir con la misma potencia computacional que las máquinas de Turing) y/o *computacionalmente eficientes* (es decir, si son capaces de proporcionar soluciones eficientes de problemas computacionalmente duros). En esta memoria hemos trabajado con los sistemas P desde un punto de vista completamente distinto; concretamente como un marco de modelado y simulación de fenómenos biológicos.

Los primeros intentos realizados en este sentido presentaban un marco semi-cuantitativo en el que se tenía en cuenta el carácter discreto de las interacciones moleculares que tienen lugar dentro de un compartimento [6,20], y seguían la estrategia clásica para la evolución de los sistemas basados en el no-determinismo y el paralelismo máximo. A pesar de que se consiguió algún éxito con dichos intentos preliminares, éstos no consiguieron modelar aspectos cuantitativos que son básicos para el funcionamiento de muchos sistemas celulares.

Este enfoque produce dos importantes inconvenientes:

- Las reglas de reescritura no se aplican en la tasa correcta.
- Los tiempos son iguales en todos los pasos y no capturan la evolución temporal del sistema celular real.

Estos dos problemas son interdependientes y han sido solucionados en esta tesis mediante la aplicación a sistemas P de la teoría de Gillespie sobre la cinética estocástica.

Nuestro marco de modelado no se ha restringido a la simple generación de simulaciones de nuestros modelos, como era el caso en trabajos anteriores, sino que hemos dado los primeros pasos hacia el desarrollo de técnicas para analizar modelos de sistemas P basados en *model checking* simbólico y probabilístico.

El punto de partida de la investigación realizada ha sido el estudio de varios casos implicados en redes de señalización y de transcripción. Se ha desarrollado una descripción de cómo especificar y modelar algunas de las interacciones moleculares más importantes que tienen lugar en las células vivas. Además, nuestras investigaciones han puesto de manifiesto el hecho de que los objetos individuales no eran adecuados para especificar ciertas entidades moleculares, tales como operons y clusters de genes, que poseen una estructura interna relevante. Para ello se ha introducido el uso de cadenas para especificar tales moléculas junto con reglas específicas de reescritura en multiconjunto objetos y cadenas.

Recuérdese que un buen marco de modelado enfocado a sistemas celulares, debe satisfacer unos requisitos mínimos, a saber, *relevancia*, *comprensibilidad*, *extensibilidad* y *computabilidad*. A continuación se describen algunos beneficios y limitaciones en el uso de sistemas P como marco de modelado para sistemas biológicos teniendo estas cuatro propiedades deseables.

- **Relevancia:** un buen modelo debe capturar propiedades esenciales de estos sistemas: su organización estructural y su comportamiento dinámico.

Por un lado, la compartimentalización de sistemas celulares está explícitamente especificado cuando se usan estructuras de membrana. Además, las cadenas se usan para describir entidades moleculares con una estructura interna importante. Sin embargo, cuando la estructura interna de entidades moleculares no es relevante, se utilizan objetos individuales.

Por otro lado, el comportamiento dinámico de los sistemas celulares se simula bastante bien con el uso de reglas de reescritura (tanto para multiconjunto de objetos como para multiconjunto de objetos y cadenas), que son similares a las ecuaciones clásicas escritas por biólogos para describir interacciones bioquímicas.

Un beneficio clave de nuestro marco de modelado propuesto es su habilidad para manejar la compartimentalización y el papel clave que desempeñan las membranas biológicas en el funcionamiento de células vivas. Así lo demuestra la colección de casos de estudio que se presentan en la Parte III de la memoria. Nuestro enfoque se ha mostrado muy apropiado para modelar la absorción selectiva de moléculas del entorno, así como en las colonias de bacterias que están interactuando que se comunican enviando y recibiendo señales de difusión. No obstante, no hemos investigado otros procesos importantes donde las membranas biológicas juegan un papel crucial, como la división celular, la adhesión celular, la formación de biocapas, etc. La especificación y la simulación de este tipo de procesos siguen siendo un problema abierto, siendo una línea para investigar en el futuro.

- **Comprensibilidad:** Quizás la propiedad más importante de un buen modelo es la comprensibilidad. Las abstracciones usadas en un buen marco de modelado de sistemas celulares deberían corresponderse bien con los conceptos informales e ideas de la biología molecular. Un modelo debería proporcionar un mejor y más completo conocimiento del sistema celular real, en vez de producir un formalismo complicado y difícil de descifrar. A este respecto, y a pesar de que los sistemas P no se introdujeron como marco de modelado de sistemas celulares comprensible, los conceptos y la terminología que usan son muy próximos a aquellos utilizados en biología molecular celular. Esto hace que las abstracciones desarrolladas dentro de este marco sean comprensibles para el biólogo celular. Como los sistemas P están inspirados en la estructura y el funcionamiento de células vivas, existe una correspondencia directa entre las abstracciones de los sistemas P y las componentes reales de los sistemas celulares.

Las abstracciones de los sistemas P son puramente textuales y hasta el momento carecen de una representación gráfica formal para la visualización de los sistemas modelados. En esta memoria se ha usado muy ocasionalmente una representación gráfica basada en [71] pero la correspondencia formal entre los modelos de sistemas P y esta representación gráfica continúan siendo una cuestión abierta.

- **Computabilidad:** un buen modelo para sistemas celulares debería ser computable a fin de permitir la simulación del comportamiento dinámico de los sistemas y el razonamiento sobre las propiedades de los sistemas. A este respecto, los sistemas P constituyen un marco de modelado adecuado, debido a su fácil computabilidad e implementación. Estas propiedades son en su mayor

parte una consecuencia de que los sistemas P consisten en sistemas basados en reglas distribuidos en estructuras específicas. En particular, las especificaciones y los modelos de los sistemas P son fácilmente implementados usando lenguajes orientados a objetos. Cada membrana puede pensarse como un objeto cuyos atributos son el multiconjunto asociado con él y la etiqueta que representa su clase. El conjunto de reglas asociadas con cada membrana puede ser considerado como funciones o métodos asociados con el objeto que representa la membrana. Más aún, los algoritmos usados en este trabajo para simular los modelos de sistemas P son locales, en el sentido de que casi todas las computaciones sólo consideran un único compartimento. La única computación global que queda es la localización del índice de la regla a aplicar y el compartimento donde va a ser aplicada.

- **Extensibilidad:** otra propiedad interesante de los modelos de sistemas P es su fácil extensibilidad, desde dos puntos de vista:

Por un lado, los modelos específicos de sistemas P pueden ser fácilmente extendidos y modificados para incorporar nuevo conocimiento. En los modelos de sistemas P las interacciones moleculares son representadas separadamente en las reglas de reescritura. A este respecto, cuando nuevas interacciones moleculares necesitan ser integradas en un modelo antiguo, todas las modificaciones están limitadas a añadir nuevas reglas de reescritura que describan las interacciones encontradas. Cuando se prueba que una interacción hipotética ya incorporada en un modelo de sistema P es errónea, el modelo del sistema se cambia simplemente eliminando la regla de reescritura usada para describir esta interacción.

Por otro lado, el marco completo de modelado de sistema P es fácilmente extensible. En esta memoria se presenta un ejemplo de ello. La aplicación de la teoría de Gillespie de cinética estocástica para la estructura compartimentalizada de los sistemas P se consiguió con facilidad. Es más, el uso de cadenas para representar algunas entidades moleculares y el uso de reglas de reescritura en multiconjunto de objetos y cadenas fueron incorporados al marco de modelado de sistemas P con poco esfuerzo.

De hecho, el conocimiento obtenido durante el desarrollo de esta tesis ha puesto de manifiesto la necesidad de ampliar el marco de modelado de sistemas P a fin de incorporar más información cuantitativa de los componentes de sistemas celulares. Por ejemplo, una dirección de trabajo futuro consiste en incorporar algunos atributos a los objetos que representan moléculas y a los compartimentos que describen membranas.

En resumen, en esta memoria se ha intentado adaptar el marco original de los sistemas P al modelado de sistemas celulares, proporcionándoles dinámicas estocásticas y los elementos necesarios para especificar las interacciones en sistemas celulares más relevantes. Se ha probado que este enfoque es un buen marco modelador completando en gran medida las cuatro propiedades generalmente requeridas en marcos de modelaje de sistemas celulares, relevancia, comprensibilidad, computabilidad y extensibilidad.

No obstante, a lo largo del trabajo desarrollado en esta memoria se han descubierto una serie de problemas que abren vías de trabajo futuro, a fin de extender el formalismo propuesto para que sea capaz de especificar y modelar de un modo más preciso los sistemas celulares.

Bibliografía

- [1] Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J. D. (1994) *Molecular Biology of the Cell*, Garland Publishing, New York, USA.
- [2] Alimonti, J.B., Ball, T.B., Fowke, K.R. Mechanisms of CD4+ T lymphocyte cell death in human immunodeficiency virus infection and AIDS. *Journal of General Virology*, **84** (2003), 1649–1661.
- [3] Alon, U. (2006). *An Introduction to Systems Biology*. Mathematical and Computational Biology Series. Chapman & Hall/Crc.
- [4] Alon, U. (2007) Network Motifs: Theory and Experimental Approaches, *Nature Reviews Genetics*, **8**, 450 – 461.
- [5] Alur, R., Henzinger, T.A. (1999) Reactive Modules, *Formal Methods in System Design*, **15** 7–48.
- [6] Ardelean, I.I., Cavaliere, M. (2003) Modelling Biological Processes by Using a Probabilistic P system Software, *Natural Computing*, **2** (2), 173 – 197.
- [7] Andrei, O., Ciobanu, G., Lucanu, D. (2005) Executable Specifications of P Systems, *LNCS*, **3365** 126 – 145.
- [8] Arkin, A., Ross, J., McAdams, H. H. (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells, *Genetics*, **149**, 1633 – 1648.
- [9] Arnold, L., Theodosopulu, M. (1980) Deterministic limit of the stochastic model of chemical reactions with diffusion, *Adv. Applied Prob.*, **12**, 367 – 379.
- [10] Aziz, A., Sanwal, K., Singhal, V., Brayton, R. (1996) Verifying continuous time Markov chains, *LNCS*, **1102**, 269–276.
- [11] Baier, C., Kwiatkowska (1998) Model checking for a probabilistic branching time logic with fairness, *Distributed Computing*, **11** (3), 125 – 155.
- [12] Barabási, A.L., Oltvai, Z. (2004) Network biology: Understanding the cells functional organization. *Nat. Rev. Genetics*, **5**, 101 – 113.
- [13] Batt, G., Casey, R., de Jong, H., Geiselman, J., Gouzé, J.L., Page, M., Ropers, D., Sari, T., Schneider, D. (2005) Qualitative analysis of the dynamics of genetic regulatory networks using piecewise-linear models. In Pecou, E., Martinez, S., Maass, A., editors, *Mathematical and Computational Methods in Biology*. Paris: Editions Hermann.
- [14] Batt, G., Ropers, D., de Jong, H., Geiselman, J., Mateescu, R., Page, M., Schneider, D. (2005) Validation of qualitative models of genetic regulatory networks by model checking: Analysis of the nutritional stress response in Escherichia coli, *Bioinformatics*, **21**, i19–i28.

- [15] Benner, S.A., Sismour, A.M. (2005) Synthetic biology, *Nature Review, Genetics*, **6**, 533–543.
- [16] Bernardini, F., Gheorghe, M., Krasnogor, N., Giavitto, J.L. (2005) On Self-Assembly in Population P Systems, *Lecture Notes in Computer Science*, **3699**, 46 – 57.
- [17] Bernardini, F., Gheorghe, M., Krasnogor, N. (2006) Quorum sensing P systems, *Theoretical Computer Science*, **371**, (1–2), 20 – 33.
- [18] Bernardini, F., Brijder, R., Rozenberg, G., Zandron, C. (2007) Multiset Based Self-Assembly of Graphs, *Fundamenta Informaticae*, **75**(1–4), 49 – 75.
- [19] Bernardini, F., Manca, V. (2003) P Systems with Boundary Rules, *Lecture Notes in Computer Science*, **2597**, 107–118.
- [20] Besozzi, D., Ciobanu, G. (2005) A P Systems Description of the Sodium-Potassium Pump, *LNCS*, **3365**, 210 –223.
- [21] Bhalla, U.S., Iyengar, R., Emergent properties of networks of biological signaling pathways. *Science*, **283** (1999), 381–387. Not cited
- [22] Bianco, A., Alfaro, L. (1995) Model checking of probabilistic and nondeterministic systems, *LNCS*, **1026**, 499 – 513.
- [23] Bianco, L., Fontana, F., Manca, V. (2006) P systems with Reactions Maps, *Int. J. Found. Comput. Sci.*, **17** (1), 27 – 48.
- [24] Blosssey, R., Cardelli, L., Phillips, A. (2006) A Compositional Approach to the Stochastic Dynamics of Gene Networks, *Transactions on Computational Systems Biology IV, LNBI* **3939**, 99-122.
- [25] Bower, J., Bolouri, H. (2001) *Computational Modelling of Genetic and Biochemical Networks*. MIT Press.
- [26] Cardelli, L. (2005). Brane calculi: interactions of biological membranes, *Lecture Notes in Bioinformatics*, **3082** 257 - 278.
- [27] Cheruku, S., A. Păun, F.J. Romero, M.J. Pérez-Jiménez, O.H. Ibarra. Simulating FAS-induced apoptosis by using P systems. *Proceedings of the First International Conference on Bio-Inspired Computing: Theory and Applications*, Wuhan, China, September, 18–22, 2006.
- [28] Ciobanu, G., Pan, L., Păun, Gh., Pérez-Jiménez, M.J. (2007) P Systems with Minimal Parallelism, *Theoretical Computer Science*, **378** (1), 117–130.
- [29] Dang, Z., Ibarra, O.H. (2005) On One-membrane P systems Operating in Sequential Mode, *Int. J. Found. Comput. Sci.*, **16** (5) 867–881.
- [30] Danos V, Laneve C. (2004) Formal molecular biology. *Theor. Comput. Sci.*, **325**(1), 69 – 110.
- [31] Elf, J., Ehrenberg, M. (2004) Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases, *Systems Bio.*, **2**, 230 – 236.
- [32] El-Masri, H.A., Portier, C.J. (1999) Replication potential of cells via the protein kinase C-MAPK pathway: Application of a mathematical model, *Bulletin of Mathematical Biology*, **61**, 379–398.

- [33] Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S. (2002) Stochastic Gene Expression in a Single Cell, *Science*, **297**, 1183 – 1186.
- [34] Fargerström, T., James, G., James, S., Kjelleberg, S., Nilsson, P. (2000) Luminescence Control in the Marine Bacterium *Vibrio fischeri*: An Analysis of the Dynamics of lux Regulation. *J. Mol. Biol.* **296**, 1127–1137.
- [35] Fontana, F., Bianco, L., Manca, V. (2005) P Systems and the Modelling of Biochemical Oscillations. In Freund, R., Păun, Gh., Rozenberg, G., Salomaa, A. (eds.) *Membrane Computing, Sixth International Workshop, WMC6, Vienna, Austria*, volume 3850 of *Lecture Notes in Computer Science*, pages 199 – 208. Springer Verlag.
- [36] Freund, R. (2005) P systems Working in the Sequential Mode on Arrays and Strings, *Int. J. Found. Comput. Sci.*, **16** (4), 663 – 682.
- [37] Funahashi, A., Morohashi, M., Kitano, H. (2003) Cell Designer: a Process Diagram Editor for Gene-regulatory and Biochemical Networks, *Biosilico*, **1**, 159 – 162.
- [38] Gardiner, C., Steyn-Ross, M. (1984) Adiabatic elimination in stochastic systems. II. Application to reaction diffusion and hydrodynamic-like systems, *Phys. Rev. A.*, **29**, 2823 – 2833.
- [39] Gheorghe, M., Păun, Gh. (2007) Computing by self-assembly: DNA molecules, polynominoes, cells. In N. Krasnogor, S. Gustafson, D. Pelta, and J.L. Verdegay, editors, *Systems Self-Assembly: Multidisciplinary Snapshots*, Studies in Multidisciplinarity, Elsevier.
- [40] Gillespie, D.T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comp. Phys.*, **22**, 403 – 434.
- [41] Gillespie, D.T. (1977) Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, **81**, 2340 – 2361.
- [42] Gillespie, D.T. (1992) A rigorous derivation of the chemical master equation, *Physics A*, **188**, 404–425.
- [43] Gillespie, D.T. (2001) Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, **115**, 1716 – 1733.
- [44] Gillespie, D.T., Petzold, L. (2003) Improved leap-size selection for accelerated stochastic simulation. *J. Chem. Phys.*, **119**, 8229 – 8234.
- [45] Girault, C., Valk, R. (2003) *Petri Nets for Systems Engineering*. Springer Verlag.
- [46] Goss, P.J.E., Peccoud, J. (1998) Quantitative modelling of stochastic system in molecular biology by using stochastic Petri nets, *Proc. Natl. Acad. Sci. USA*, **95**, 6750 – 6755.
- [47] Greenberg, E.P., Kaplan, H.B. (1985). Diffusion of Autoinducer Is Involved in Regulation of the *Vibrio fischeri* Luminescence System. *Journal of Bacteriology* **163**(3), 1210–1214
- [48] Greenberg, E.P., Kaplan, H.B. (1997). Overproduction and Purification of the luxR Gene Product: Transcriptional Activator of the *Vibrio fischeri* Luminescence System. *Proc. Natl. Acad. Sci.* **84**, 6639–6643
- [49] Greenberg, E.P., Lupp, C.; Ruby, E.G., Urbanowski, M. (2003). The *Vibrio fischeri* Quorum-sensing System *ain* and *lux* Sequentially Induce Luminescence Gene Expression and Are Important for Persistence in the Squid Host. *Molecular Microbiology*, **50**(1) 319–331

- [50] Hansson, H., Jonsson, B. (1994) A logic for reasoning about time and reliability. *Formal Aspects of Computing*, **6** (5), 512 – 535.
- [51] Hardie, K.R., Williams, P., Winzer, K. (2002). Bacterial cell-to-cell communication: sorry, can't talk now, gone to lunch. *Current Opinion in Microbiology*, **5**, 216–222
- [52] Harel, D. (1987). Statecharts: A visual formalism for Complex Systems. *Science of Computer Programming*, **8** (3), 231–274.
- [53] Hartwell, L.H., Hopfield, J.J., Leibler, S. Murray, A.W. (1999) From Molecular to Modular Cell Biology, *Nature*, **402**, C47 – C52.
- [54] Holcombe, M., Gheorghe, M., Talbot, N. A hybrid machine model of rice blast fungus, *Magnaporthe Grisea*. *BioSystems*, **68**, 2–3 (2003), 223–228.
- [55] Hua, F., Cornejo, M., Cardone, M., Stokes, C., Lauffenburger, D. Effects of Bcl-2 Levels on FAS Signaling-Induced Caspase-3 Activation: Molecular Genetic Tests of Computational Model Predictions. *The Journal of Immunology*, **175**, 2 (2005), 985–995 and correction **175**, 9 (2005), 6235–6237.
- [56] Hucka, M., A. Finney, H. Sauro, H. Bolouri, J. Doyle, H. Kitano, A. Arkin, B. Bornstein, D. Bray, A. Cornish-Bowden, A. Cuellar, S. Dronov, E. Gilles, M. Ginkel, V. Gor, I. Goryanin, W. Hedley, T. Hodgman, J. Hofmeyr, P. Hunter, N. Juty, J. Kasberger, A. Kremling, U. Kummer, N. Le Novere, L. Loew, D. Lucio, P. Mendes, E. Minch, E. Mjolsness, Y. Nakayama, M. Nelson, P. Nielsen, T. Sakurada, J. Schaff, B. Shapiro, T. Shimizu, H. S. and J Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, **19**, 524 – 531.
- [57] Huang, K.C., Meir, Y., Wingreen, N.S. (2003) Dynamic structures in Escherichia coli: Spontaneous formation of MinE rings and MinD polar zones. *Proc. Natl. Acad. Sci. USA*, **100**, 12724 – 12728.
- [58] Huber, R.E., Wallenfels, K., Kurz, G. (1975). Action of β -galactosidase on allolactose. *Canad. J. Biochem.*, **53**, 1035–1038.
- [59] Ideker, T., T. Galitski, and L. Hood (2001). A new approach to decoding life: systems biology. *Annual Review of Genomics and Human Genetics*, **2**, 343 – 372.
- [60] Ionescu, M., Sburlan, D. (2004) On P systems with Promoters/Inhibitors, *J. UCS*, **10** (5) 581 – 599.
- [61] Ionescu, M., Păun, Gh., Yokomori, T. (2006) Spiking Neural P systems, *Fundam. Inform.* **71** (2-3) 279 – 308.
- [62] Jaatela, M. (2004) Multiple cell death pathways as regulators of tumour initiation and progression. *Oncogene*, **23**, 2746–2756.
- [63] Jackson, D., Holcombe, M., Ratnieks, F. (2004) Trail geometry gives polarity to ant foraging networks. *Nature* **432**, 907–909.
- [64] Kell, D.B. (2002) Genotype-phenotype mapping: genes as computer programs, *Trends in Genetics*, **18**, 11, 555 – 559.

- [65] Kell, D.B. (2005) Metabolomics, machine learning and modelling: towards an understanding of the language of cells, *Biochemical Society Transactions*, **33**, 520 – 524.
- [66] Kennell, D., Riezman, H. (1977). Transcription and Translation Initiation Frequencies of the Escherichia coli lac Operon. *J. Mol. Biol.*, **114**, 1–21.
- [67] Kerr, J.F., Wyllie, A.H., Currie, A.R. (1972) Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *British Journal Cancer*, **26**, 239.
- [68] Kholodenko, B.G, Demin, O.V., Moehren, G., Hoek, J.B. (1999) Quantification of short term signalling by the Epidermal Growth Factor Receptor, *J. Biol. Chem.* **274**, 30169 – 30181.
- [69] Kierzek, A.M., Zaim, J., Zielenkiewicz, P. (2001). The Effect of Transcription and Translation Initiation Frequencies on the Stochastic Fluctuations in Prokaryotic Gene Expression. *The Journal of Biological Chemistry*, **276**, 11, 8165–8172.
- [70] Kitano, H. (Ed.) (2001). *Foundations of Systems Biology*. MIT Press.
- [71] Kitano, H., Matsuoka, Y., Funahashi, A., Oda, K. (2005) Using process diagrams for the graphical representation of biological networks, *Nature Biotechnology*, **23**(8), 961 – 966.
- [72] Koch, A.J., Minhardt, H. (1994) Biological pattern formation: from basic mechanism to complex structures, *Rev. Mod. Phys.*, **66**, 1481 – 1507.
- [73] Kiatkoska, M., Norman, G., Parker, D. (2005) Probabilistic Model Checking in Practice: Case Studies with PRISM, *ACM SIGMETRICS Performance Evaluation Review*, **32** (4), 16 – 21.
- [74] Kuttler, C. (2006) Simulating Bacterial Transcription and Translation in a Stochastic Pi-calculus, *Transactions on Computational Systems Biology VI, LNBI*, **4220**, 113 – 149.
- [75] Lipton, P. (2005) Testing hypotheses: Prediction and prejudice, *Science*, **307**, 219 – 221.
- [76] Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J. (2003) *Molecular Cell Biology*. Fifth Edition, Freeman.
- [77] Lolkema, J.S., Carrasco, N., Kaback, R. (1991). Kinetic analysis of lactose exchange in proteoliposomes reconstituted with purified lac permease. *Biochemistry*, **30**, 1284 – 1290.
- [78] McAdams, H., Arkin, A. (1997) Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA*, **94**, 814 – 819.
- [79] McAdams, H., Arkin, A. (1999) It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Gnet.*, **15**, 65 – 69.
- [80] Milner, R. (1999) *Communication and Mobile Systems: The π -calculus*. Cambridge University Press.
- [81] Moehren G., Markevich, N., Demin, O., Kiyatkin, A., Goryanin, I., Hoek, J.B., Kholodenko, B.N. (2002) Temperature Dependence of the Epidermal Growth Factor Receptor Signaling Network Can Be Accounted for by a Kinetic Model, *Biochemistry* **41**, 306–320.
- [82] Moghal, N., Sternberg, P.W. (1999) Multiple positive and negative regulators of signaling by the EGFR receptor. *Curr. Opin. Cell Biology*, **11**, 190–196. not cited
- [83] Newman, M. (2003) The structure and function of complex networks. *SIAM Rev.*, **45**, 167 – 256.

- [84] Nijhawan, D., Honarpour, N., Wang, X. (2000) Apoptosis in neural development and disease. *Annual Reviews Neuroscience*, **23**, 73–87.
- [85] Oltavi, Z.N., Milliman, C.L., Korsmeyer, S.J. (1993) Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. *Cell*, **74**, 4, 609–619.
- [86] Orton, R.J., Sturm, O.E. et al. (2005) Computational modelling of the receptor-tyrosine-activated MAPK pathway, *Biochem. J.*, **392**, 249 – 261.
- [87] Palsson, B. O. (2006). *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press.
- [88] G. Păun (2002) *Membrane Computing: An Introduction*. Springer-Verlag.
- [89] Păun, Gh. (2000) Computing with Membranes, *Journal of Computer and System Sciences*, **61**(1) 108 – 143.
- [90] Păun, A., Păun, Gh. (2002) The Power of Communication: P systems with Symport/Antiport, *New Generation Comput.*, **20** (3) 295 – 306.
- [91] G. Păun (2002) *Membrane Computing: An Introduction*. Springer-Verlag.
- [92] Pérez-Jiménez, M.J., Romero-Campero, F.J. (2006) P Systems, a new computational modelling tool for systems biology, *Transactions on Computational Systems Biology VI, LNBI*, **4220**, 176 – 197.
- [93] Pescini, D., Besozzi, D., Mauri, G., Zandron, C. (2006) Dynamical probabilistic P systems, *International Journal of Foundations of Computer Science*, **17** (1) 183–195.
- [94] Pogson, M., Smallwood, R., Qwarnstrom, E., Holcombe, M. (2006) Formal agent-based of intracellular chemical interactions. *BioSystems*, **85**, 1 37–45. not cited
- [95] Priami, C., Regev, A., Shapiro, E., Silvermann, W. (2001) Application of a stochastic name-passing calculus to representation and simulation of molecular processes, *Information Processing Letters*, **80**, 25 – 31.
- [96] Ptashne, M. (2004). *A Genetic Switch, Third Edition Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press.
- [97] Ptashne, M., Gann, A. (2002). *Genes and Signals*. Cold Spring Harbor Laboratory Press.
- [98] Ray, T. S. (1994) An evolutionary approach to synthetic biology: Zen and the art of creating life. *Artificial Life*, **1**(1/2): 195 – 226. MIT Press.
- [99] Reddy, V.N., Liebman, M.N., Maverovouniotis, M.L. (1996) Qualitative Analysis of Biochemical Reaction Systems, *Comput. Biol. Med.*, **26**, 1, 9 – 24.
- [100] Regev, A., Panina, E.M., Silvermann, W., Cardelli, L., Shapiro, E. (2004) BioAmbients: an abstraction for biological compartments, *Theoretical Computer Science* **325** 141–167.
- [101] Regev, A., Shapiro, E. (2004) The π -calculus as an abstraction for biomolecular systems, In Gabriel Ciobanu and Grzegorz Rozenberg, editors, *Modelling in Molecular Biology*. Springer Berlin.
- [102] Regev, A., Shapiro, E. (2002) Cellular abstractions: Cells as computation, *Nature* **419** 343.

- [103] Reisig, W. (1998) *Elements of Distributed Algorithms. Modelling and Analysis with Petri Nets*. Springer Verlag.
- [104] Reisig, W., Rozenberg, G. (eds.) (1998) *Lectures on Petri Nets I: Basic Models*, volume 1491 of *Lecture Notes in Computer Science*. Springer Verlag.
- [105] Reisig, W., Rozenberg, G. (eds.) (1998) *Lectures on Petri Nets II: Applications*, volume 1492 of *Lecture Notes in Computer Science*. Springer Verlag.
- [106] Rohwer, J.M., Meadow, N.D., Roseman, S., Westerhoff, H. (2000). Understanding Glucose Transport by the Bacterial Phosphoenolpyruvate: Glycose Phosphotransferase System on the Basis of Kinetic Measurements in Vitro. *The Journal of Biological Chemistry*, **275**, 45, 34909–34921.
- [107] Romero-Campero, F.J., Pérez-Jiménez, M.J. Modelling Gene Expression Control Using P Systems: The Lac Operon, A Case Study, submitted.
- [108] Romero-Campero, F.J., Pérez-Jiménez, M.J. A Model of the Quorum Sensing System in *Vibrio fischeri* Using P Systems, submitted.
- [109] Rosenfeld, N., Elowitz, M.B., Alon, U. (2002) Negative Autoregulation Speeds the Response Times of Transcription Networks, *J. Mol. Biol.*, **323**, 785 – 793.
- [110] Ruby, E.G., McFall-Ngai, M.J. (1999) Oxygen-utilizing reactions and symbiotic colonisation of the squid light organ by *vibrio fischeri*, *Trends in Microbiology*, **7**, 414 – 420.
- [111] Rutten, J., Kwiatkowska, M., Norman, G., Parker, D. (2004) *Mathematical Techniques for Analyzing Concurrent and Probabilistic Systems*, volume 23 of *CRM Monograph Series*, AMS.
- [112] Savageau, M.A. (1976) *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology*, Addison-Wesley, Reading, MA.
- [113] Savageau, M.A. (1974) Comparison of classical and auto-genous systems of regulation in inducible operons, *Nature*, **252**, 546 – 549.
- [114] Scaffidi, C., Fulda, S., Srinivasan, A., Friesen, C., Li, F., Tomaselli, K.J., Debatin, K.M., Kramer, P.H., Peter, M.E. (1998) Two CD95 (APO-1/Fas) signaling pathways. *The Embo Journal*, **17**, 1675–1687.
- [115] Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., Muller, G. (2002) Computational Modeling of the Dynamics of the MAP Kinase Cascade Activated by Surface and Internalized EGF Receptors, *Nature Biotechnology*, **20**, 4 370–375.
- [116] Shmulevich, I., Dougherty, E., Kim, S., Zhang, W. (2002) Probabilistic Boolean Networks: A rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, **18**, 261–274.
- [117] Starbuck, C., Lauffenburger, D.A. (1992) Mathematical model of the effects of Epidermal Growth Factor Receptor trafficking dynamics on Fibroblast proliferation responses, *Biotechnol. Prog.*, **8**, 132 – 143.
- [118] Stevens, A.M., Greenberg, E.P. (1997). Quorum Sensing in *Vibrio fischeri*: Essential Elements for Activation of the Luminescence Genes. *Journal of Bacteriology* **179**(2) 557–562.
- [119] Sydor, J.R. (1998) Transient kinetic study on the interactions of Ras-binding domain of c-Raf-1 reveals rapid equilibrium of the complex, *Biochemistry* **37**, 14292 – 14299.

- [120] Szallasi, Z., J. Stelling, and V. Periwal (Eds.) (2006). *System modeling in cellular biology*. MIT press.
- [121] Terrazas, G., Krasnogor, N., Gheorghe, M., Bernardini, F., Diggle, S., Camara, M. (2005) An Environment Aware P-System Model of Quorum Sensing. *Lecture Notes in Computer Science*, **3526**, 479 – 485.
- [122] Thieffry, D., Huerta, A.M., Perez-Rueda, E., Collado-Vides, J. (1998) From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays*, **20**, 433 – 440.
- [123] Turing, A.M. (1952) The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc.*, **237**, 37 – 72.
- [124] Vayttaden, S., Ajay, S., Bhalla, U. (2004) A spectrum of models of signalling pathways. *ChemBioChem*, **5**, 1365 – 1374.
- [125] Van Kampen, N.G. (1992) *Stochastic processes in Physics and Chemistry*, Elsevier Science B.V., Amsterdam, The Netherlands.
- [126] Walker, D.C., Southgate, J., Hill, G., Holcombe, M., Hose, D.R., Wood S.M., MacNeil, S., Smallwood, R.H. (2004) The epitheliome: modelling the social behaviour of cells. *BioSystems*, **76**, 1–3, 89–100.
- [127] Wang, K., Yin, X.M., Chao, D.T., Milliman, C.L., Korsmeyer, S.J. (1996) BID: a novel BH3 domain-only death agonist. *Genes & Development*, **10**, 2859–2869.
- [128] Waters, C.M., Bassler, B.L. (2005) Quorum sensing: Cell-to-cell communication in bacteria. *Annual Review of cell and Developmental Biology*, **21**, 319 – 346.
- [129] Wells, A. (1999) EGFR–receptor. *Int. Journal Biochem. Cell Biology*, **31**, 637–643.
- [130] Wiley, H.S., Shvartsman, S.Y., Lauffenburger, D.A. (2003) Computational modeling of the EGFR–receptor system: A paradigm for systems biology. *Trends in Cell Biology*, **13**, 1, 43–50.
- [131] Wong, P., Gladney, S., Keasling, J.D. (1997). Mathematical Model of the lac Operon: Inducer Exclusion, Catabolite Repression, and Diauxic Growth on Glucose and Lactose. *Biotechnol. Prog.*, **13**, 132–143
- [132] Nottingham Quorum Sensing Web Site <http://www.nottingham.ac.uk/quorum/>
- [133] PRISM Web Site: www.cs.bham.ac.uk/~dxd/prism/
- [134] P Systems Modelling Framework Web Site:
www.dcs.shef.ac.uk/~marian/PSimulatorWeb/P_Systems_applications.html
- [135] SBML Site: <http://sbml.org>



UNIVERSIDAD DE SEVILLA

Reunido el tribunal en el día de la fecha, integrado por los abajo firmantes, para evaluar la tesis doctoral de D. FRANCISCO JOSÉ ROMERO CAMPERO titulada P SYSTEMS, A COMPUTATIONAL MODELLING FRAMEWORK FOR SYSTEMS BIOLOGY acordó otorgarle la calificación de Sobresaliente cum Laude por unanimidad

Sevilla, a 6 de febrero de 2008

Vocal,

Presidente,

Vocal,

Secretario,

Vocal,

Doctorando,