



Cinvestav

CENTRO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL
Unidad Monterrey

CARACTERIZACIÓN DE LOS MECANISMOS
DEMODULADORES EN LA TRANSMISIÓN DE
INFORMACIÓN EN LAS VÍAS DE SEÑALIZACIÓN

Tesis que presenta

Aarón Vázquez Jiménez

Para obtener el título de

**Doctorado en Ciencias en
Ingeniería y Física Biomédica**

Director de Tesis

Dr. Jesús Guadalupe Rodríguez González.

Apodaca, México

Diciembre 2017

*para aquellos que son el futuro:
Isaías, Nohemí, Jaden y Eva.*

Resumen

Las células para sobrevivir necesitan percibir y responder a las condiciones extracelulares en que se encuentran. Esto lo logran transduciendo la señal exterior en información química a través de las vías de señalización. Para su uso eficiente, la información procesada necesita ser modulada, transmitida y demodulada. La demodulación es crucial ya que es el proceso mediante el cual la información es interpretada para efectuar la correspondiente acción celular. En células eucariotas, la demodulación es llevada a cabo por dos mecanismos: mediante el uso de motivos prealimentadores; y mediante la afinidad del promotor por su factor de transcripción en la expresión de un gen; sin embargo, aún no se entiende cómo funcionan con respecto al procesamiento de la información. Con la intención de explorar los mecanismos demoduladores, se implementó el modelo estocástico de la vía MAP/ERK y el de una vía mínima basado en la interacción con lazos prealimentadores. Usando la teoría de información como marco conceptual y haciendo un símil entre las vías de señalización y un canal de comunicación, se calculó la máxima cantidad de información que es procesada. Los resultados demuestran que los mecanismos demoduladores permiten la extracción de información gracias a que amplifican y adicionan ruido. Aunado a esto, cuando ambos mecanismos están acoplados, la cantidad de información extraída aumenta. Se evaluó los efectos del ruido extrínseco, resultando que cuando la componente intrínseca y la extrínseca del ruido son consideradas es perjudicial, ya que la información disminuye. No obstante, la pérdida de información es controlada por la frecuencia del ruido extrínseco. En conjunto, los resultados establecen los principios de funcionamiento de los mecanismos que demodulan la información.

Abstract

The cells must sense and response to their environment to survive. They transduce stimuli into chemical information using the signaling pathways. Such information needs to be coded, transmitted and decoded in order to be processed. The decoding process is fundamental to ensure the proper information interpretation, in eukaryotes is based on two mechanisms: in the use of feedforward loop motifs and in the gene affinity for their transcription factor. However, it is still unclear how the decoding process works regarding information processing. Therefore, a model of the MAPK/ERK pathway decoding section and a minimal model based on the feedforward loop interaction were examined. The amount of extracted information was calculated using information theory analysis. The results demonstrated that the feedforward loop allows extracting information thanks to amplification and noise addition processes. Additionally, the extraction is enhanced when both decoding mechanisms are coupled. Although the extrinsic noise plays a detrimental role in information transmission, it can be buffered by controlling the noise frequency. All together establish a broader scope of how the decoding process works.

Agradecimientos

Doy las gracias al Consejo Nacional de Ciencia y Tecnología (CONACYT) y al Centro de Investigación y Estudios Avanzados (CINVESTAV) por haberme proporcionado el apoyo económico y la infraestructura que he necesitado para completar el doctorado en ciencias en ingeniería y física biomédica.

Agradezco a mi director de tesis, el Dr. Jesús Rodríguez González por haber compartido conmigo sus conocimientos e instruirme no sólo para la realización de este trabajo, sino para mi formación como investigador.

Igualmente estoy agradecido con todos los investigadores del CINVESTAV Monterrey que, de alguna manera, me han brindado su colaboración.

Gracias a todos aquellos que me cuestionaron e incluso dudaron del trabajo, haciendo pensar en posibilidades que me llevaron a explorar nuevas áreas.

Finalmente, con mucho respeto y humildad, quiero agradecer a mi familia, ya que gracias a esos pequeños detalles son los que me dieron la fuerza para prender la luz en la obscuridad.

Sinceramente,

Aarón Vázquez Jiménez

Diciembre 2017

Índice general

1	Introducción	1
	Motivación	1
	Antecedentes	3
	Justificación	14
	Planteamiento del problema	14
	Hipótesis	15
	Objetivo	15
	Organización	15
2	Fundamentos Teóricos	17
2.1	Características generales de las via de señalización	17
2.2	Expresión Génica	18
2.2.1	Transcripción	19
2.2.2	Traducción	19
2.2.3	Regulación	20
2.3	Vía de señalización MAP/ERK	23
2.3.1	Sección Moduladora	24
2.3.2	Sección Demoduladora	25
2.4	Teoría de Información	26
2.4.1	Entropía	28
2.4.2	Información Mutua	31

ÍNDICE GENERAL

2.4.3	Capacidad del canal	32
2.4.4	Información Mutua Multivariable	33
2.4.5	Desigualdad de procesamiento de datos	34
2.5	Variables Aleatorias	35
2.5.1	Representación Estocástica	36
3	Metodología	39
3.1	Definición de entrada	39
3.2	Modelo	43
3.3	Algoritmo de Gillespie	43
3.4	Distribución de Probabilidad Condicional	44
3.5	Cálculo de la Capacidad del Canal	46
3.5.1	Algoritmo de Arimoto-Blahut	46
3.6	Cálculo de la Información Mutua Multivariable	49
3.7	Ruido Extrínseco	50
3.7.1	Algoritmo Extrínseco	51
3.8	Simulación	52
4	Modelos Matemáticos	55
4.1	Modelo de la vía de señalización MAP/ERK	55
4.2	Modelo Mínimo	58
4.3	Modelo de Genes Reporteros	59
4.4	Modelos estocásticos	60
4.5	Validación	62
5	Resultados & Discusión	67
5.1	Caracterización del FFL como demodulador	67
5.1.1	Vía MAP/ERK	68
5.1.2	Vía mínima	73
5.2	Caracterización de la afinidad de un gen como mecanismo demodulador	78
5.3	Ruido Extrínseco	83
5.4	Resumen	86

ÍNDICE GENERAL

6	Perspectivas	89
7	Conclusión	91
A	Análisis de Estabilidad	93
A.1	Modelo de la sección demoduladora de la vía MAP/ERK	93
A.2	Modelo Mínimo	96
A.3	Modelo de genes reporteros	96
	Bibliografía	99

Introducción

Motivación

Las células captan y procesan una plétora de estímulos con el objetivo de desencadenar respuestas específicas que les permita adaptarse y subsistir. Uno de los mecanismos para efectuar este proceso son las vías de señalización, las cuales son cadenas de interacción que generalmente transmiten información del estímulo del exterior al interior celular. Una forma natural de concebir el procesamiento de todos los estímulos en la célula es idealizando la activación mediante procesos específicos de tal manera que un estímulo es identificado por sólo una vía de señalización, Fig. 1.1(a). Sin embargo, se ha demostrado que múltiples estímulos son reconocidos por vías de señalización en común; aunado a esto, existen nodos o conjuntos de reacciones que se comparten entre múltiples vías, Fig. 1.1(b). A pesar de la intrincada activación, las respuestas celulares desencadenadas son específicas ante cada estímulo. De manera fundamental, en los estudios de señalización celular se trata de entender cómo emerge la especificidad dentro de la activación promiscua dentro de las células [1]. Se postula que la especificidad se origina debido a la organización de la red (la topología de interacción). A partir de la cual se generan códigos modulando las características dinámicas de las señales intermedias en las vías. Tales características son: amplitud, frecuencia

1. INTRODUCCIÓN

y ancho de pulso [2–5]. Posteriormente, la información es transmitida y finalmente debe de ser extraída o demodulada. El bloque demodulador es determinante, ya que la respuesta celular depende de sus propiedades y no del código en sí [6–8]. Se han identificado dos estructuras que llevan a cabo el proceso de demodulación: mediante el uso de topologías prealimentadores (FFL¹) o de la afinidad de un gen por su factor de transcripción². Mediante estos dos mecanismos se ha explicado la generación de respuestas dinámicas que conllevan al encendido de diferentes programas celulares de una manera específica; sin embargo, aún no es claro qué propiedades poseen en la transmisión de la información. En el caso del FFL, se ha demostrado que limita el procesamiento de las señales, lo cual es una característica contraintuitiva para un modulo que procesa información.

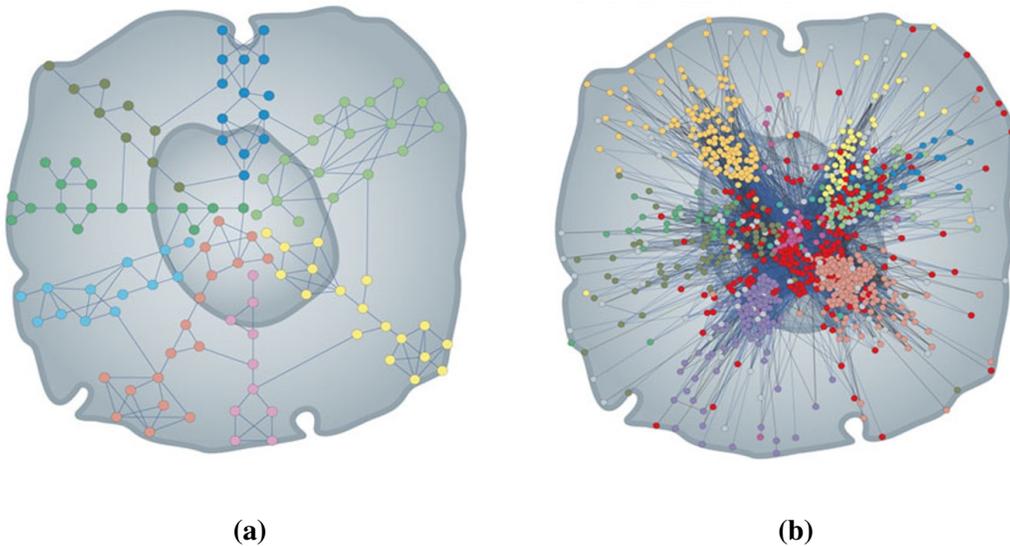


Figura 1.1: Interacción entre vías de señalización. Los nodos de colores diferentes al rojo identifican vías diferentes. Los nodos rojos son puntos que se activan por más de una vía de señalización [9]. (a) Esquema ideal. (b) Esquema real.

¹Por sus siglas en inglés *Feed-Forward Loop*.

²Molécula que promueve la expresión de un gen particular.

Antecedentes

Se estima que la vida surgió hace 3.8 miles de millones de años, aproximadamente 750 millones de años después de la formación de la tierra. El cómo se formaron los primeros organismos es un tema que seduce a investigadores, logrando generar especulaciones; dado que las condiciones medioambientales primitivas no se pueden reproducir en laboratorio. Sin embargo, existen experimentos en los cuales ha sido posible aproximar dichas condiciones, resultando la generación de productos orgánicos básicos para los procesos biológicos como aminoácidos [10].

La explicación de la vida desde el cultivo primario hasta los organismos actuales se basa en la interacción entre las condiciones ambientales y los diversos compuestos; lo cual resalta la importancia del entorno celular. Ya que el medio celular es un factor tan determinante, los organismos deben de tener la capacidad de responder de manera que les permita adaptarse para sobrevivir. La adaptación se refiere a los cambios producidos por la célula en respuesta al entorno [11]. Todas las células reciben y responden a señales de su entorno, las cuales transmiten a su interior. Por lo tanto las células comunican las condiciones medioambientales a su interior o incluso a otras células vecinas.

Señalización Celular

Dentro de los mecanismos de comunicación celular existe la señalización; la cual, es parte de cualquier proceso de comunicación bioquímico que determina las actividades y acciones celulares a través del control en la expresión génica y la función de proteínas con el fin de responder y adaptarse. Por ejemplo, las bacterias poseen proteínas en la membrana que detectan estímulos externos funcionando como sensores, lo que les permite detectar cambios en pH, osmoralidad¹, iluminación, disponibilidad de alimentos, oxígeno y presencia de sustancias nocivas. Tales estímulos inducen respuestas celulares particulares, como el movimiento de la bacteria hacia áreas con altas concentraciones de comida o la interrupción del ciclo celular. En organismo multicelulares se intercambian gran variedad

¹Expresa la concentración de agua dentro y fuera de la célula.

1. INTRODUCCIÓN

de estímulos. Las células vegetales responden a hormonas de crecimiento y variaciones en la intensidad de la luz solar. Las células animales detectan cambios en la concentración de iones y glucosa. En todos estos casos, los estímulos representan *información*; la cual, es detectada por receptores dando como resultado una respuesta celular que involucra un proceso bioquímico. La conversión de información a cambios químicos celulares es una propiedad crucial de todas las células vivas [12]. Se ha determinado que errores en la transmisión de la información son causa de enfermedades como cancer y enfermedades auto-inmunes, entre otras [13, 14].

La señalización celular es un sistema de comunicación en cascada con etapas secuenciales. Se inicia con la detección de un estímulo por un receptor específico, seguida de un proceso de transducción al convertir la señal extracelular en interna mediante un conjunto de reacciones secundarias. La interacción de las moléculas activadas internamente y su interacción con proteínas efectoras permite la regulación de procesos celulares y funciones fisiológicas [15, 16], Fig. 1.2. Conceptualizando el estímulo como un mensaje, la señalización como la transmisión del mensaje y la molécula final como el receptor; se puede concebir el proceso de señalización como la transmisión de información.

Estructuralmente las moléculas que transmiten la información van desde gases simples a proteínas. Algunas moléculas se propagan largas distancias, mientras que otras actúan localmente. Bajo este paradigma la sola actividad del mediador es suficiente para la transmisión del mensaje. Por lo tanto, los mediadores de señalización deben ser únicos para cada estímulo, ya que en caso contrario un solo estímulo podría desencadenar diversas respuestas no específicas. Sin embargo bajo los mapas de interacción obtenidos para diversas vías de señalización (Fig. 1.1) es difícil entender el origen de la especificidad dado que existen múltiples vías o secciones de ellas que son activadas ya sea por otras vías o por un mismo estímulo [9]; una manera de optimizar el manejo de la información es modulándola.

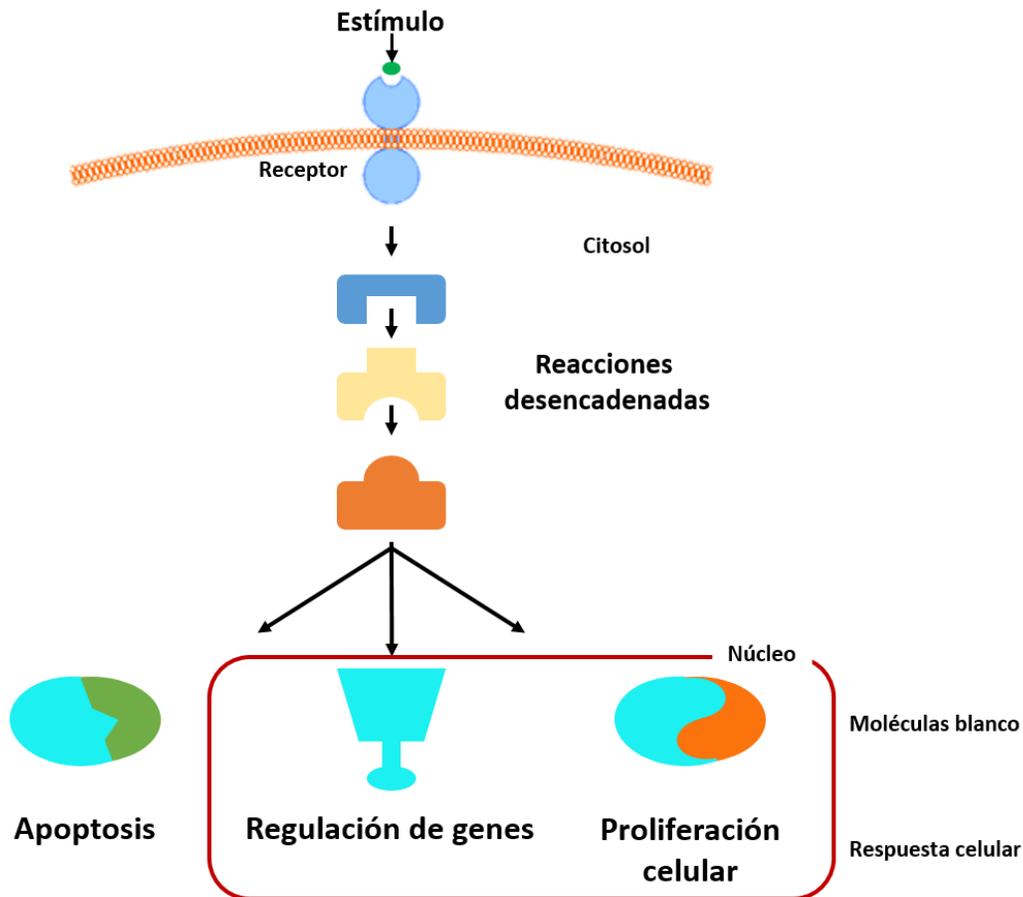


Figura 1.2: Señalización.

Modulación

La modulación en biología, es el establecimiento de relaciones de correspondencia (códigos) entre los estímulos externos y las señales internas. La modulación en el área de telecomunicaciones es el conjunto de técnicas para transportar información, con el fin de aprovechar de mejor manera el canal de información [17]. Biológicamente la modulación optimiza el manejo de recursos al compartir secciones para diversos estímulos y aun así desencadenar respuestas diferentes [18–20].

1. INTRODUCCIÓN

Códigos Dinámicos

Se propone que la información se modula en las características dinámicas de las moléculas¹. Al ser parte intrínseca de la señal [21, 22], y al conjugarse con la identidad química del mediador transmiten la información del estímulo. En general, las características de las señales son: frecuencia², amplitud³, retardo⁴, umbral⁵, transitorio⁶, estado estacionario⁷ y duración, Fig. 1.3.

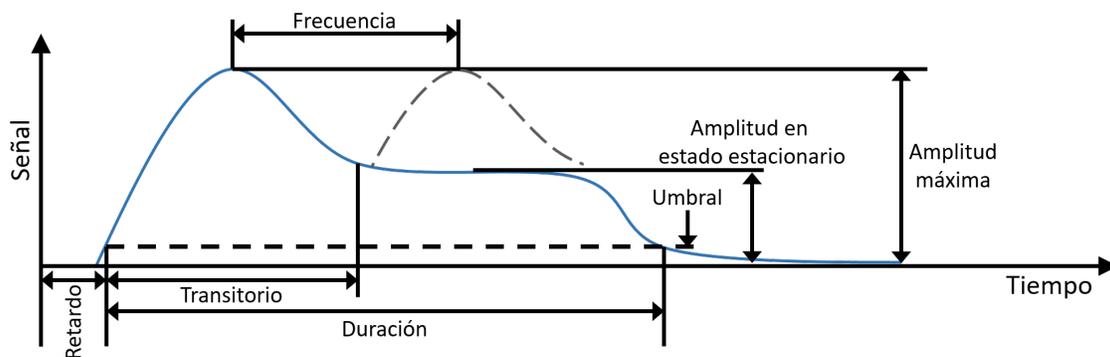


Figura 1.3: Características de las señales.

Por medio de las características de las señales se propaga la información [20], ya que la respuesta depende directamente de ellas; por lo que por medio de las dinámicas se establece un nuevo paradigma de transmisión en el cual las características del estímulo modula la dinámica de los mediadores [7]. Levine et al. (2013) establecen que para entender el código temporal y sus propiedades asociadas, se deben considerar las interacciones que generaron la dinámica específica. A su vez, existen diversos esquemas de modulación ocupados para la transmisión de señales en biología [6, 7, 23, 24]; los cuales, se presentan a continuación.

¹Entiéndase como el patrón espacio-temporal de la actividad.

²Número de repeticiones por unidad de tiempo.

³Valor máximo de la dinámica.

⁴desfase entre la aplicación de la perturbación y el cambio en la dinámica

⁵Valor en amplitud mínimo de la señal para iniciar una actividad específica.

⁶Tiempo necesario para que la dinámica se mantenga constante ante una perturbación.

⁷Valor en el cual la señal ya no cambia en amplitud en función del tiempo

Modulación en Amplitud (AM)

En este esquema de modulación, la información está contenida en la amplitud de la señal. A diferentes magnitudes de estímulo diferentes magnitudes de respuesta [7]. Este mecanismo se presenta en procesos bifásicos en los cuales existe un umbral, el cual al ser superado se produce la activación de la respuesta celular (la vía de señalización se desencadena); siendo un tipo de control todo o nada [25].

Whitehurst et al. (2004) aportan evidencia experimental de la modulación en amplitud. Utilizando la vía de señalización MAP/ERK, observan que la concentración de ERK en el núcleo aumenta en función del aumento en la concentración de factores de crecimiento, Fig. 1.4 (a). La modulación en amplitud es característica de la actividad enzimática y en el estudio de receptores en los cuales se da un comportamiento de dosis-respuesta típico. Además, este fenómeno se da en distintos niveles de organización celular: procariontes y eucariotes¹ [26–28]. Behar y Hoffmann (2010) mencionan que a pesar que la AM es el mecanismo de modulación más estudiado hace falta una caracterización de las limitantes del tipo de información que se pueden transmitir.

Modulación en Frecuencia (FM)

En la modulación de frecuencia, existe una oscilación de la respuesta con una frecuencia asociada, la cual está modulada directamente por el estímulo. La frecuencia de la respuesta es el acarreador de la información del estímulo. Cai et al. (2008) demostraron que la actividad del factor de transcripción Crz1 está modulada vía FM al regular la respuesta a calcio en levaduras; Crz1 es internalizado al núcleo con una frecuencia en función del calcio extracelular. La dinámica de internalización de Crz1 se da a partir del aglomerado de pulsos denominados bursts, Fig. 1.4 (b).

¹Las células se subdividen en dos grandes grupos: eucariotes y procariontes. Dentro de las grandes diferencias son: las eucariotes poseen un núcleo delimitado por membrana celular, además del modo de respirar, reproducción y tamaño. Las células procariontes incluyen a las bacterias y las eucariotes a los animales, hongos, plantas, algas y protozoos [15].

1. INTRODUCCIÓN

El calcio es el segundo mensajero celular más importante ya que regula y desencadena muchos procesos intracelulares [12] por lo que es un ion ampliamente estudiado. Smedler y Uhlén (2014) hacen una compilación en los procesos celulares en los cuales la concentración del iones calcio (Ca^{2+}) oscila¹; de manera interesante establecen los rangos de frecuencia de oscilación para cada proceso. Se observa que existe una yuxtaposición de frecuencias; sin embargo, la especificidad se conserva de cada estímulo. Adicionalmente, existen otros sistemas en donde se ha demostrado la modulación por FM: células T [30], hepatocitos [31], segundos mensajeros IP3 [32], cardiomiocitos [33] y liberación de hormonas [3].

Modulación en Ancho de Pulso (PWM)

Como tercer esquema, la modulación por ancho de pulso establece la relación de la duración del pulso en función del estímulo. Msn2 es un factor de transcripción en levadura activado ante condiciones de estrés, el cual se transloca del citoplasma al núcleo induciendo expresión de genes de acuerdo al tipo de estrés. La permanencia en el núcleo de Msn2 es mayor mientras exista menor cantidad de glucosa pero la amplitud del pico de respuesta es constante; los resultados se repiten para estrés osmótico [2], Fig. 1.4 (c). Este tipo de modulación se observa en otros sistemas como en la cascada de regulación MAPK en hepatocitos [34], la respuesta a factores de necrosis tumoral en fibroblastos de ratón [35] y el efecto provocado por diversos factores de crecimiento en células precursoras de neuronas en rata [18].

Hasta este punto, se han definido diferentes esquemas de transmisión de la información. No obstante, a pesar que el estímulo sea modulado eficientemente, la información transferida está limitada por la precisión con la que es extraída [7, 24].

¹Respuesta inmune, homeostasis, transporte, motilidad, neurotransmisión, producción de energía y cadenas de fosforilación.

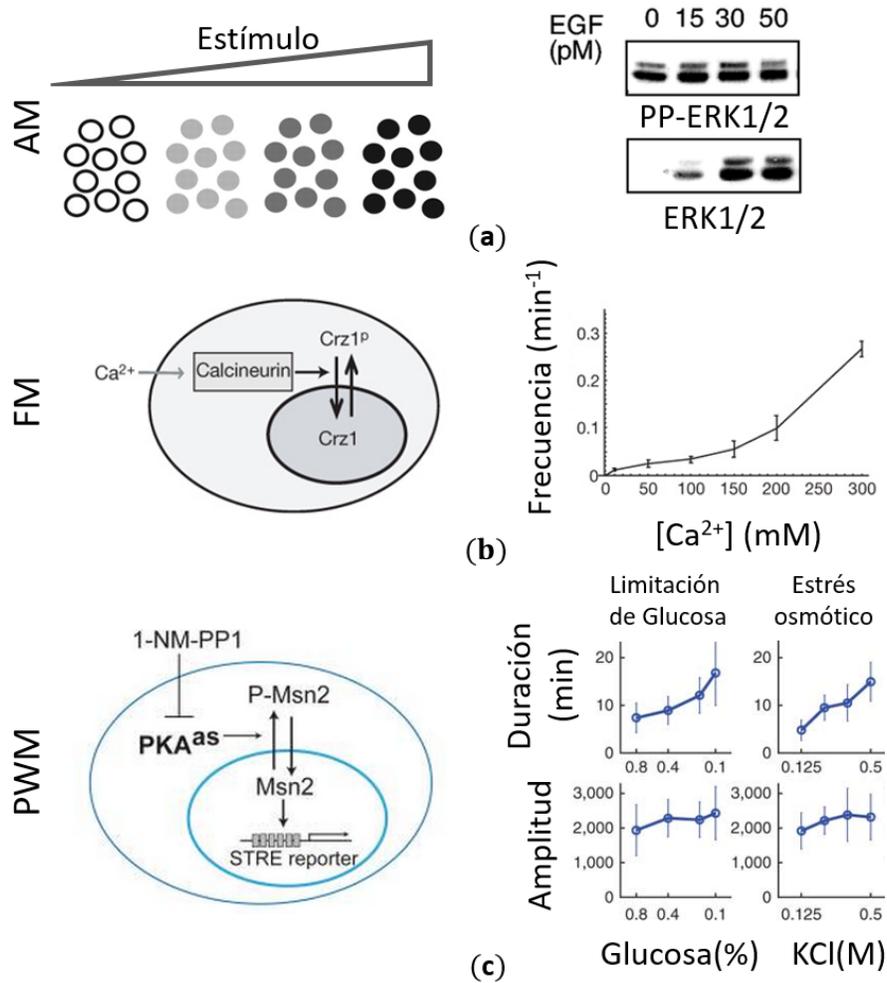


Figura 1.4: Esquemas de modulación. (a) Modulación en Amplitud (AM): la cinasa ERK en sus dos isoformas incrementa su expresión ante incrementos en el factor de crecimiento epidermal, cuando está fosforilado no responde [4]. (b) Modulación en Frecuencia (FM): el factor de transcripción Crz1 en levadura modifica su frecuencia de permanencia en el núcleo en función de la concentración de calcio [29]. (c) Modulación en ancho de pulso (PWM): EL factor de transcripción Msn2 en levadura modifica el ancho de su pulso de actividad en función del nivel de estrés [2].

1. INTRODUCCIÓN

Demodulación

La demodulación dentro del área de las telecomunicaciones es el conjunto de técnicas utilizadas para extraer la información enviada a través de un canal de información. En biología la demodulación se atribuye a entidades dentro de una vía de señalización con la capacidad de interpretar la modulación del estímulo y responder adecuadamente. Esto es de particular importancia ya que la caracterización funcional del código en las vías de señalización requiere centrarse en los mecanismo demoduladores, ya que un mismo código puede leerse por diferentes blancos celulares y controlar diferentes respuestas. Estos mecanismos son la conexión entre los patrones dinámicos y la respuesta celular [7]. Esta idea fue explorada en la oscilación de la concentración de Ca^{2+} intracelular; la cual, oscila en un amplio espectro de frecuencias controlando una gran gama de respuestas. Por lo que el *quid* del reconocimiento está basado en los mecanismo decodificadores [6, 8]. En consecuencia, para entender el flujo de información celular primeramente se debe comprender las capacidades y limitaciones en el proceso de demodulación.

Lazos Prealimentadores (FFL)

Dentro de las vías de señalización se han identificado patrones de regulación recurrentes denominados motivos, un motivo característico son los FFL. Topológicamente, en los FFL interactúan tres moléculas como mínimo en donde la tercera molécula es regulada por las dos anteriores. Se clasifican de acuerdo al tipo de interacción global que posean, si son estructuras con interacción positiva se les denomina coherentes (Fig. 1.5(a)) y si resultan negativas (Fig. 1.5(b)) se les atañe el nombre de incoherente¹[36].

Los FFL coherentes poseen un retraso sensitivo, este se da de acuerdo a la lógica en la activación de la tercera molécula por lo que puede haber un retraso en el encendido si es que Z necesita la activación de X y Y. O en el apagado si es que Z sólo requiere alguna de las anteriores. Los FFL incoherentes se establecen como estructuras aceleradores de respuesta y generadora de picos o impulsos; debido a la existencia de dos ramas (una rápida

¹Existen diversos tipos de FFL, los mostrados son los tipo I, los cuales son de interés y relevancia para el presente proyecto.

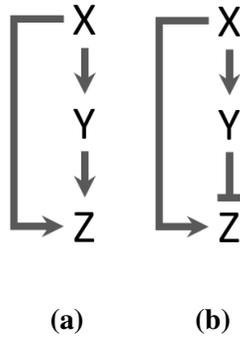


Figura 1.5: Tipos de FFL. Existen diversas combinaciones en los cuales se pueden estructurar, se muestran los tipo 1. **(a)** Coherente. **(b)** Incoherente.

y otra lenta) el sistema puede responder aceleradamente hasta que la rama lenta compense el sobre impulso [36, 37]. Los FFL regulan la expresión de genes en organismos con diferente organización celular: *Escherichia coli* [38, 39], Levaduras [40], *Caenorhabditis elegans* [41], mosca de la fruta [39], erizos de mar [42] y humanos [43]. A su vez, se ha demostrado que la presencia de FFL es determinante en procesos biológicos, tales como: la formación de bloques de memoria [44], detector de estímulos persistentes [41], filtro de señales [45], en mecanismos moleculares que permiten la adaptación a estímulos externos [46] y en procesos en los cuales existe la decodificación.

Se ha identificado al FFL en vías de señalización como estructuras demoduladoras [20]. En la vía MAPK/ERK, el FFL modula el ciclo celular de acuerdo a factores de crecimiento [1, 20, 47]. Además, se ha demostrado su participación en la vía activada del receptor TLR4¹, regulando la respuesta en el reconocimiento de patógenos y en la activación de la respuesta inmune [48]. Se postula además que es el responsable en la codificación de la dinámica de la proteína p53 que induce la respuesta ante daño en el ADN². El FFL se ha denominado el núcleo decodificador debido a su característica de poder generar diversos

¹Proteína transmembranal perteneciente a la familia de receptores que reconocen patrones (PRRs) desencadenando la activación del factor NF-kB responsable de la activación del sistema inmune.

²De acuerdo a la dinámica de p53 la célula repara su ADN o se marca para muerte celular.

1. INTRODUCCIÓN

patrones temporales; a pesar de ello, limita la cantidad de información que puede procesar debido a la forma en que está conectado ya que forma un cuello de botella [5]; característica contraintuitiva para ser el modulo central de extracción de información. A pesar que el análisis sobre el uso del FFL como decodificar ha sido iniciado, su importancia funcional en términos de información aún es desconocida.

Afinidad

Otro mecanismo demodulador probado para células eucariotas, específicamente en levaduras, es el uso de la afinidad (fuerza de asociación) entre el promotor¹ y su factor de transcripción (FT). Una afinidad baja implica que existe una baja probabilidad que el FT se una a al promotor; por lo tanto, la dinámica del FT necesita ser sostenida para lograr la expresión del gen. Comparado con el caso de una afinidad alta, la dinámica del FT puede ser más rápida. Este mecanismo de demodulación se demostró en el factor de transcripción Msn2² activado por estrés en levaduras. Las diferencias en las propiedades de unión y la dinámica de transición del promotor de los genes regulados provocan la respuesta a diferentes patrones dinámicos de Msn2 [2, 24]. Las levaduras transmiten información acerca de la identidad de diferentes estímulos sin embargo no pueden diferenciar intensidades del mismo estímulo, funcionando como un interruptor [49].

Los genes de levadura no son los únicos cuya activación depende de la afinidad; en las vías de señalización en células de mamífero que terminan en la activación de un FT, se promueve la expresión de un gen que desencadena la respuesta final. De tal forma que en las vías que demodulan por FFL también tienen acoplado el segundo mecanismo demodulador, la afinidad. Hasta el momento los estudios se han centrado en caracterizar las propiedades de estos mecanismo en forma desacoplada [5, 50].

¹Región del ADN que controla la expresión de un gen, cuando el promotor es reconocido por un factor de transcripción la expresión se favorece, sección 2.2.

²El mecanismo de modulación fue presentado en la sección anterior.

Ruido

Los procesos biológicos son el resultado de interacciones moleculares que están sujetas a ruido. El ruido es un proceso caracterizado y descrito en muchas áreas del conocimiento, se define como un proceso estocástico que produce variabilidad [51]. Biológicamente se le asocia el nombre de ruido bioquímico. Es originado por la naturaleza aleatoria de las reacciones bioquímicas al transitar de forma discreta entre diferentes estados químicos [52–54]. Está conformado por dos componentes, la intrínseca y extrínseca [55]. La primera es originada debido a la estocasticidad inherente de su naturaleza química. La componente extrínseca se manifiesta en la variación del conteo de moléculas efectoras y de las condiciones ambientales afectando de manera global a los sistemas [56].

Específicamente, en las vías de señalización el ruido extrínseco está asociado a la fluctuación en las moléculas que llevan a cabo el proceso de fosforilación [57, 58]. Por otro lado, existe otra fuente de variabilidad, generada por la activación promiscua¹ de los componentes de las vías de señalización. Ya que diferentes vías son activadas por un mismo estímulo, y moléculas comunes son utilizadas por diversas vías (Fig. 1.1). Uno de los principales retos en el estudio de las vías de señalización es el entendimiento de la especificidad en las respuestas dado que la estocasticidad en las reacciones interfiere en la transducción y transmisión de la señal provocando pérdida de información [50, 59]. De manera limitante, se ha determinado que los mecanismos para disminuir y filtrar el ruido, como los lazos de retroalimentación negativa, provocan una pérdida de información al disminuir el rango de respuesta del sistema [50]. Un modo de incrementar la robustez en la transmisión de información es teniendo múltiples copias de un gen o cuando se considera la información transmitida en una población [49, 50]. A pesar de ello no existe una caracterización de los limitantes y propiedades de los mecanismos demoduladores ante las fuentes de ruido.

¹Referido como una mezcla de activación de diferentes estímulos.

1. INTRODUCCIÓN

Justificación

Todos los trabajos mencionados representan un avance en el entendimiento del manejo de la información en las vías de señalización y en general de los procesos biológicos. Sin embargo, la necesidad de entender las características intrínsecas de los mecanismos demoduladores sigue vigente. Más aún, no se ha estudiado los procesos de demodulación de forma acoplada ni explorado el efecto del ruido, determinado únicamente que es perjudicial. En este trabajo, se reconoce la necesidad de un estudio sistemático de las propiedades y características de los mecanismos demoduladores.

Planteamiento del problema

Comprender el fenómeno de extracción de información en vías de señalización de células eucariotas, tomando en cuenta la interacción entre la señal estímulo y topologías prealimentadoras, además entender el efecto generado en la transmisión de la información acoplando la afinidad del gen por su factor de transcripción como una segunda estructura demoduladora. Todo esto con un enfoque estocástico considerando las diversas fuentes de ruido bioquímico.

Para el desarrollo de esta tesis, se ha cubierto un estudio teórico-computacional de los mecanismos de extracción de la información. Se han combinado conceptos tradicionales del procesamiento de información en telecomunicaciones y se han extendido al campo del estudio de la biología, específicamente de las vías de señalización. Se propuso un método basado en la cuantificación del límite superior de información que se puede procesar por un canal biológico. Se desarrollaron modelos estocásticos *in silico* tanto del estímulo como de los mecanismos demoduladores de forma separada y acoplada. Se implementó un algoritmo que permitió explorar diferentes propiedades de la variabilidad extrínseca. Las respuestas obtenidas de los modelos se procesaron para obtener la máxima cantidad de información que se puede procesar ante diversas condiciones permitiendo disectar las propiedades de los mecanismos demoduladores.

Hipótesis

La funcionalidad de los mecanismos demoduladores en células eucariotas radica en la optimización de extracción de la información y no únicamente en su capacidad de generar diferentes dinámicas temporales en la respuesta celular.

Objetivo

Estudiar los principios de funcionamiento de los mecanismos demoduladores en el proceso de extracción de información en las vías de señalización de células eucariotas considerando la diversas componentes de ruido que afectan el proceso. De forma particular, el objetivo general se ramifica en: 1. Estudio de los motivos FFL y la afinidad en presencia de ruido intrínseco. 2. Estudio de la interacción acoplada de ambos de mecanismos. 3. Interacción y efecto de los mecanismos ante presencia de ruido intrínseco y extrínseco.

Organización

La tesis está dividida en tres secciones: los capítulos dos y tres contienen, respectivamente, el marco teórico y la metodología aplicada en este trabajo; la segunda sección corresponde al capítulo cuatro, el cual describe los modelos matemáticos utilizados para definir el estímulo, el canal de transmisión y los mecanismos demoduladores. La última sección son los capítulos del cinco al siete, que presentan los resultados al aplicar la metodología para procesar la información y obtener una representación de la señal basada en la cantidad de información transmitida; este capítulo está subdividido en tres bloques que exploran las propiedades de: el FFL, la afinidad acoplada al FFL y la interacción con el ruido extrínseco. Se discuten los resultados y se relacionan con el fenómeno de extracción de información. Además, se establecen los alcances y limitaciones del estudio, así como los potenciales trabajos futuros que se deriven de la investigación presentada. Finalmente, se concluye basándose en la generalidad de los resultados.

Fundamentos Teóricos

2.1 Características generales de las via de señalización

Como ya se comentó, mediante las vías de señalización las células tienen la capacidad de detectar y responder a estímulos externos. Lo cual implica una transducción, transmisión y extracción de la información química. Las vías de señalización poseen las siguientes características, las cuales determinan su eficiencia.

Especificidad Las vías de señalización se activan ante estímulos específicos, la especificidad es posible gracias a la interacción complementaria entre el ligando y el receptor mediante fuerzas de interacción no covalentes.

Amplificación Durante las cascadas se activan enzimas asociadas a un receptor, que a su vez catalizan la activación de otras enzimas en un proceso recurrente.

Adaptación La sensibilidad del receptor está sujeta a modificaciones, cuando un estímulo se presenta continuamente ocurre una desensibilización del receptor. Si el estímulo nuevamente es menor que el umbral, el sistema se vuelve sensible otra vez.

2. FUNDAMENTOS TEÓRICOS

Integración Las células responden de manera única y particular de acuerdo a las condiciones sin importar que sean estimuladas por más de una señal simultáneamente.

La iniciación de la transducción de la señal se inicia con el receptor y está directamente regulada por el tipo de receptor. Se conocen seis tipos de receptores: **a)** Receptor de membrana acoplado a proteínas G: el ligando al unirse al receptor activa la subunidad proteica G, la cual regula a otras enzimas¹ desencadenando la activación de segundos mensajeros. **b)** Receptores de membrana con actividad en tirosín cinasa: los receptores al activarse por el ligando forman dímeros y se auto fosforilan²; su activación promueve su función como enzimas cinasas, las cuales fosforilan los residuos de tirosina de las proteínas que activan. **c)** Receptores de membrana con actividad de guanilato ciclasa: al activarse, su dominio enzimático intracelular actúa generando el segundo mensajero guanósín monofosfato cíclico (GMPc). **d)** Compuertas iónicas activadas por ligando: receptor membranal que como su nombre indica se abre y cierra de acuerdo a su activación/desactivación; al estar activado permite el flujo de iones entre el medio extracelular y el citoplasma. **e)** Receptores de adhesión: se unen a macromoléculas en el medio extracelular cambiando de conformación estructural, desencadenando la modificación de su interacción con el interior. **f)** Receptores en el núcleo: receptores localizados en el núcleo que internalizan moléculas modificando la expresión de genes ya sea promoviéndola o restringiéndola.

2.2 Expresión Génica

En biología, una de las funciones básicas de los organismos es la preservación de la especie. La preservación conlleva a que debe de existir cierto grado de conservación de información. La unidad fundamental de información en un sistema vivo es el gen. Bioquímicamente es un segmento de ADN, que contiene la información necesaria para producir una molécula funcional (usualmente una proteína) así como el conjunto de instrucciones necesarias

¹Moléculas de naturaleza proteica que catalizan reacciones químicas termodinámicamente posibles, transcurriendo a mayor velocidad.

²La fosforilación es la adición de grupos fosfato a una molécula, en bioquímica este proceso activa proteínas para que interactuen y puedan desencadenar una respuesta, el proceso es reversible.

para su expresión [12]. En la expresión de un gen los procesos más importantes son la transcripción y traducción, Fig. 2.1.

2.2.1 Transcripción

Proceso de síntesis de ARN a partir del ADN, de los cuales existen de tres tipos: ARNm (mensajero) que codifican la secuencia para la síntesis de proteínas; ARNt (de transferencia) decodifican la información contenida en el ARNm y transfieren el aminoácido apropiado a la cadena polipeptídica durante la síntesis de proteínas; y el ARNr (ribosomal) el cual forma parte estructural de los ribosomas que son la maquinaria que sintetiza las proteínas.

2.2.2 Traducción

Proceso por el cual se sintetizan las proteínas; está regido por 5 procesos intermedios: **a)** activación de aminoácidos, el grupo carboxilo de cada aminoácido es activado para la correcta unión con el ARNt; **b)** iniciación, el ribosoma reconoce y se une al sitio de unión al ARNm maduro, formándose el complejo de iniciación; **c)** elongación, se forma una hebra de aminoácidos (polipéptido); **d)** terminado y liberación, se agrega el codón de terminación y se libera la hebra polipeptídica del ribosoma; **e)** plegamiento y procesamiento post-traducciona, mediante procesos enzimáticos la proteína se vuelve funcional.

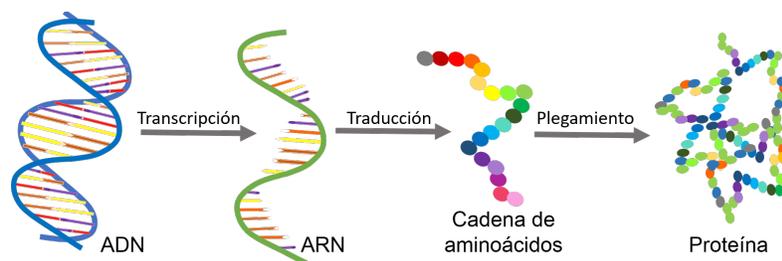


Figura 2.1: Expresión de un gen.

2. FUNDAMENTOS TEÓRICOS

2.2.3 Regulación

La complejidad en la expresión de genes no está determinada por el tipo o tamaño de proteína que se sintetice, sino por sus mecanismos de regulación. Un aspecto fundamental de las vías de señalización es su capacidad de regular la expresión génica, mediante la modificación de la actividad de factores de transcripción. La síntesis, localización o la actividad de estos factores está estrictamente determinada por los sistemas de transducción, mediante mecanismos que incluyen la fosforilación/defosforilación, proteólisis e interacción con segundos mensajeros [60]. Los factores de transcripción regulan la transcripción mediante su unión al promotor [40], zona del gen que controla el inicio de la transcripción, Fig. 2.2. A su vez, esta unión depende de la afinidad que exista entre las dos moléculas. El factor de transcripción aumenta la probabilidad de la unión entre la polimerasa¹ y el gen²[62].

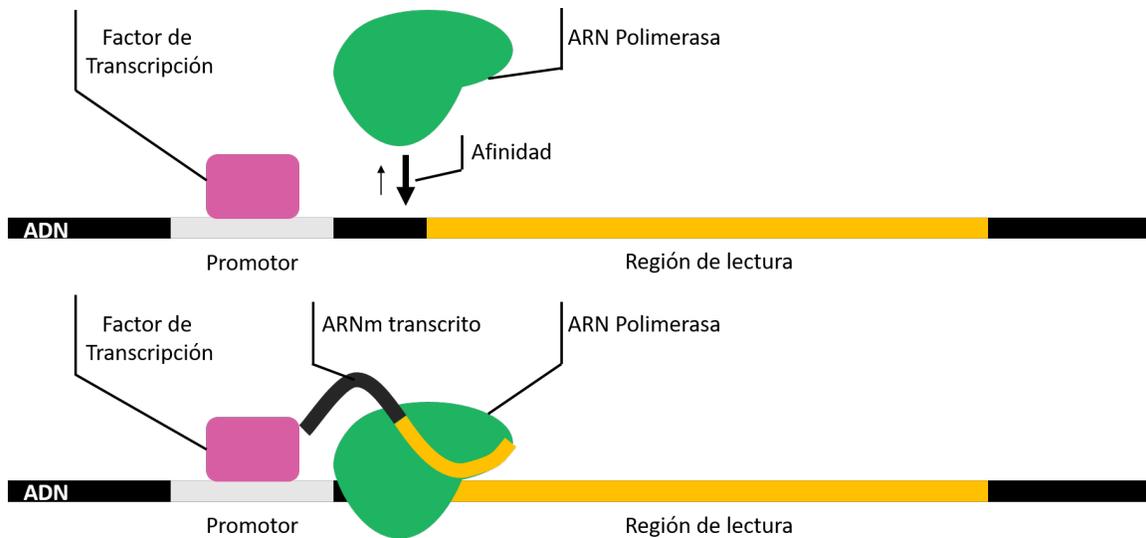


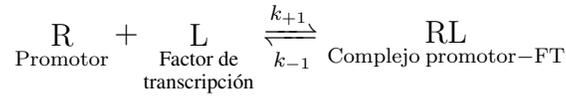
Figura 2.2: Control transcripcional regulado por factores de transcripción.

¹Molécula efectora de la transcripción, en células eucariotas existen varias polimerasas, la polimerasa II es la lleva a cabo la transcripción [61].

²Algunos factores de transcripción inhiben la transcripción al disminuir la probabilidad de unión entre la polimerasa y el gen.

2.2 Expresión Génica

Modelando la relación entre el nivel de expresión dado la activación del promotor y la concentración de un factor de transcripción se tiene que R y L representan respectivamente la sección del promotor en el ADN libre y el factor de transcripción. Cuando hay un reconocimiento entre las dos entidades se unen y forman el complejo promotor-factor de transcripción (RL). La tasa a la cual esta reacción sucede es k_{+1} , sin embargo al ser reversible el complejo se disocia por interacción termodinámicas a una tasa k_{-1} .



En equilibrio químico se tiene que el cambio en la concentración¹ del complejo es cero ($d[RL]/dt = 0$) y las velocidades de reacción se igualan. Esta consideración es plausible dado que el complejo RL alcanza su equilibrio en el orden de segundos; de tal manera que:

$$k_{+1}[R][L] = k_{-1}[RL] \quad (2.1)$$

$$K_d = \frac{k_{-1}}{k_{+1}} = \frac{[R][L]}{[RL]} = \frac{1}{K_a} \quad (2.2)$$

K_d es la constante de disociación representando la tasa a la cual el complejo se disocia, siendo el inverso de la tasa de asociación (K_a). K_a está directamente relacionada a la afinidad que existe entre el factor de transcripción y el promotor, si existe una afinidad alta la tasa de asociación es alta y viceversa. Obteniendo la fracción del promotor ocupado (θ) sabiendo de la Ec. (2.2) que $[RL] = [R][L]/K_d$ se tiene que:

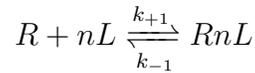
$$\begin{aligned} \theta &\equiv \frac{\text{Fracción ocupada}}{\text{Concentración total}} = \frac{[RL]}{[R_{total}]} \\ &= \frac{[RL]}{[RL] + [R]} \\ &= \frac{[L]}{[L] + K_d} \end{aligned} \quad (2.3)$$

¹Los corchetes cuadrados ([]) denotan concentración.

2. FUNDAMENTOS TEÓRICOS

La ecuación (2.3) se conoce como la ecuación de Hill¹ que relaciona la concentración del factor de transcripción y la fracción del promotor unido. K_d también representa la concentración de FT a la cual se tiene la mitad del promotor unido, $K_d = [L] \Rightarrow \theta = 1/2$.

Existen situaciones en las cuales los factores de transcripción necesitan formar complejos para unirse al promotor. Por lo que se tiene la siguiente reacción.



donde n representa el número de unidades que se necesitan reclutar. Haciendo el mismo procedimiento y bajo las mismas consideraciones anteriores se tiene que la generalización de la ecuación (2.3) es

$$\theta = \frac{[L]^n}{[L]^n + K_d^n} \quad (2.4)$$

Graficando (2.4) se observa que para valores mayores de n el sistema se vuelve más sensible ante pequeños cambios tendiendo la respuesta a una función escalón, Fig. 2.3. Biológicamente n toma valores en el rango de 1 a 4.

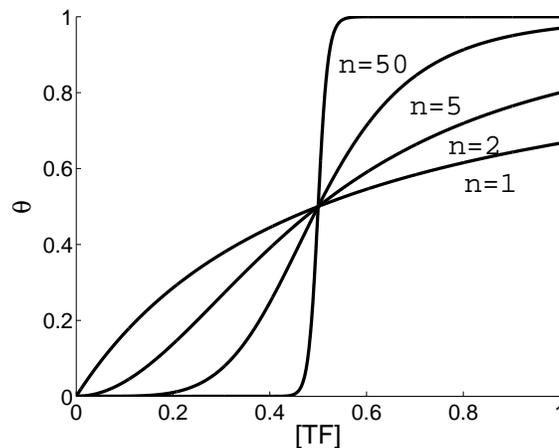


Figura 2.3: Ecuación de Hill.

¹Esta ecuación describe la cinética enzimática y el grado de cooperatividad [63].

2.3 Vía de señalización MAP/ERK

Como ya se estableció en la capítulo 1, las vías de señalización son ocupadas para transmitir información. La cual es modulada, transmitida y demodulada. Dentro de las vías que se ha demostrado que presenta estos procesos se encuentra la vía MAP/ERK.

La vía de señalización MAP/ERK está identificada en mamíferos, donde transduce estímulos desde la superficie de la membrana celular hasta el núcleo con el objetivo de desencadenar respuestas celulares como: proliferación, diferenciación, apoptosis¹, adhesión celular, migración, estabilización de ARNm y efectuar la transcripción, entre otras [64, 65]. El mal funcionamiento de esta vía se asocia al desarrollo de cancer, daño de tejido gástrico, enfermedades congénitas y mutaciones [64, 66, 67]. De manera específica, la proliferación y diferenciación son desencadenadas por factores de crecimiento. El factor de crecimiento epidérmico (EGF)² induce la proliferación celular, mientras que el factor de crecimiento nervioso (NGF)² propicia la diferenciación celular [68]. La proliferación celular es el proceso mediante el cuál una célula madre produce un par de células hijas [69]. Por su parte la diferenciación, es un mecanismo mediante el cual la célula no especializada sufre modificaciones dando lugar a un tipo celular específico, por ejemplo células musculares o hepáticas [69].

La vía de señalización MAP/ERK se divide en dos secciones: la sección moduladora y la sección demoduladora [20]. Cuando es activada por factores de crecimiento el estímulo se codifica por cadenas de interacción diferentes; sin embargo, EGF y NGF activan moléculas en común. Ambos factores de crecimiento activan la cinasa ERK aunque con dinámicas diferentes [70]. En el caso de EGF, la activación de ERK es transitoria mientras que NGF mantiene la activación de forma sostenida. En la sección decodificadora ambos factores de crecimiento comparten la misma vía. De manera notable la dinámica de ERK es suficiente para inducir respuestas diferentes a pesar que el módulo de procesamiento sea el mismo, Fig. 2.4.

¹Muerte celular programada por el mismo organismo, con el fin de controlar su desarrollo y crecimiento.

²Por sus siglas en inglés *Epidermal Growth Factor* y *Nerve Growth Factor*.

2. FUNDAMENTOS TEÓRICOS

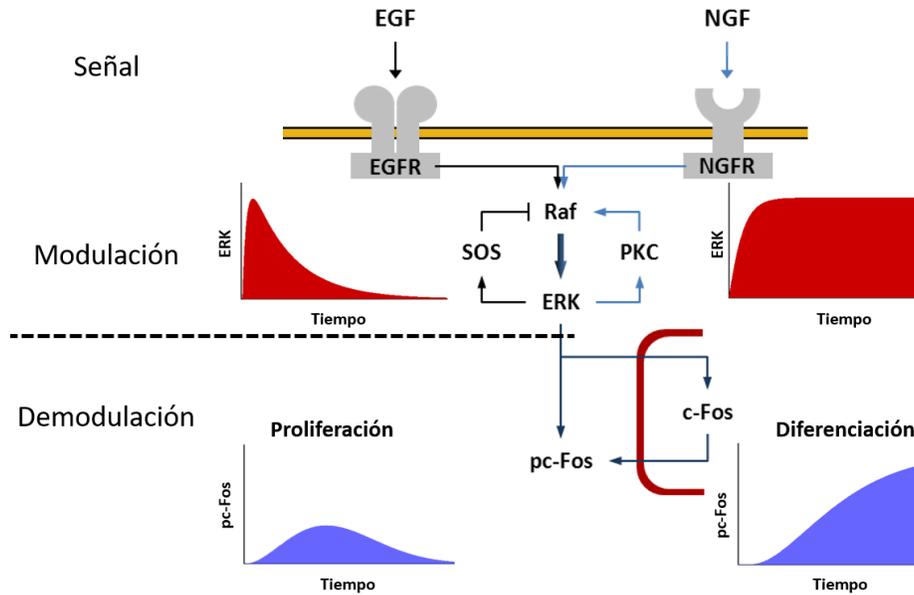


Figura 2.4: Vía MAP/ERK. Esquema de la vía MAP/ERK donde se muestran las diferentes secciones y de manera resumida las cadenas de interacción

2.3.1 Sección Moduladora

Los receptores de EGF y NGF son receptores de membrana con actividad en tirosín cinasa (ver sección 2.1), la activación de la vía induce una cascada de activación de proteínas cinasas que fosforilan los residuos de serina y treonina [65].

De manera particular, EGF mediante su receptor (EGFR) activa a Ras, enzima considerada como una proteína G pequeña que hidroliza GTP¹ [71]. Ras activa a Raf, enzima considerada como un protooncogén², funciona como un interruptor ante la presencia de factores de crecimiento. Raf por su parte activa mediante fosforilación en los residuos de tirosina a la enzima cinasa ERK (por su siglas en inglés *Extracellular signal-regulated kinase*) la cual inhibe la función del intercambiador de nucleótidos SOS al ser fosforilado y disociarlo del complejo [72, 73]. SOS activa a Ras desestabilizando la unión de GDP

¹Proceso por el cual se produce GTP al fosforilar GDT.

²Proteína que promueve el crecimiento y división celular.

2.3 Vía de señalización MAP/ERK

para que se una a GTP y así promover su activación [70]. De manera general, hay un lazo de retroalimentación negativo, lo que genera un pulso transitorio en ERK, Fig. 2.4.

El caso de NGF al ser reconocido por su receptor, existe una internalización del receptor lo que induce una autofosforilación de Raf, produciendo la fosforilación en los residuos de tirosina a ERK (igual que EGF). Existe una retroalimentación positiva mediante la activación de la proteína cinasa C (PKC) [72]. Debido a la retroalimentación positiva y a la internalización del receptor, ERK presenta una activación sostenida, Fig. 2.4.

2.3.2 Sección Demoduladora

Las dos dinámicas diferentes de ERK son procesadas por el mismo modulo demodulador de la vía, cuando ERK es transitorio la respuesta celular es de proliferación; cuando es sostenida la célula se diferencia [70], Fig. 2.4. La molécula que provoca la respuesta es el factor de transcripción c-Fos, siendo una proteína de respuesta rápida que controla la progresión del ciclo celular y la diferenciación [47]. Para ser funcional necesita ser estabilizada y no se degrade, esto se logra mediante su fosforilación (p-c-Fos). Cuando la dinámica de ERK es transitoria induce la expresión de c-Fos; sin embargo, está se degrada rápidamente. Cuando el nivel de ERK es persistente, c-Fos es fosforilado por ERK y así estabilizarse [47, 74]. Topológicamente este proceso esta estructurado mediante el uso de un FFL coherente tipo I. El FFL presenta dos interacciones, una rápida determinada por la interacción con ERK/p-c-Fos y una lenta que tiene como intermediario a c-Fos. Estas interacciones con velocidades diferentes funcionan como un sensor de duración de la activación de ERK [20]. Por lo tanto, la dinámica de ERK es demodulada de acuerdo a su duración.

Hasta aquí, se presentó un panorama resumido de la vía y de la forma en que se da su activación, es de especial interés la sección demoduladora, la cual se analiza con más detalle en la sección 4.1.

2. FUNDAMENTOS TEÓRICOS

2.4 Teoría de Información

Las vías de señalización están dedicadas a procesar información y transmitirla del exterior de la célula hacia el interior para generar una respuesta apropiada. Para el estudio y caracterización del flujo de información es fundamental cuantificar la información. Shannon [75] introdujo el marco teórico necesario para hacerlo al establecer la teoría de información.

La teoría de información es una propuesta teórica que relaciona las leyes y teoremas matemáticos que rigen la transmisión y el procesamiento de la información, con el fin de caracterizar la capacidad de los sistemas de comunicación. Históricamente se desarrolló a partir de la acelerada difusión y especialización que experimentaron los medios de comunicación durante la primera mitad del siglo XX a finales de la segunda guerra mundial. Específicamente, se desarrolló en el área de la telegrafía donde surgió la necesidad de determinar, con la máxima precisión, la capacidad de los diferentes sistemas de comunicación para transmitir información. La primera formulación de las leyes matemáticas que gobiernan dicho sistema fue realizada por Hartley (1928) y sus ideas son consideradas actualmente como la génesis de la teoría de información. Posteriormente, Shannon y Weaver (1949) desarrollaron los principios definitivos de esta teoría. Su trabajo se centró en algunos de los problemas que surgen en los sistemas destinados a manipular información, tales como: ¿cuáles son los mejores métodos para utilizar los diversos sistemas de comunicación?; ¿cómo establecer el mejor método para separar las señales del ruido? y ¿cómo determinar los límites posibles de un canal?.

El concepto de comunicación en el contexto de la teoría de la información es empleado en un sentido muy amplio en el que se incluyen todos los procesos en los cuales un sistema puede influir en otro. En el proceso de comunicación es posible distinguir por lo menos tres niveles de análisis diferentes según las ciencias de la comunicación¹: el técnico, el

¹Las ciencias de la comunicación agrupan disciplinas que estudian y analizan los fenómenos sociales, medios, conjunto semiótico. Entre estas ciencias se hace referencia a la sociolingüística, sociología, cibernética entre otros.

2.4 Teoría de Información

semántico¹ y el pragmático². La teoría de información se desarrolla como una respuesta a los problemas técnicos del proceso de comunicación.

El modelo general de Shannon parte de la existencia de una fuente de información de la cual se genera una señal (mensaje) emitida a través de un transmisor (modulador). La señal viaja por un canal que puede estar sujeto a interferencia (ruido) afectando la información contenida; un receptor (demodulador) capta la señal al final del canal para extraer la información y ser pasada a un destinatario (Fig. 2.5). La información contenida en un mensaje es proporcional a la cantidad de unidades de información³ que se requieren como mínimo para representar al mensaje. En este modelo de comunicación el significado o la identidad del estado es irrelevante. Sólo importa que cada estado sea codificado como un símbolo diferente. Como consecuencia de esto cualquier sistema con la capacidad de tomar múltiples estados puede ser una fuente de información. Matemáticamente la fuente puede ser representada como una variable aleatoria [76].

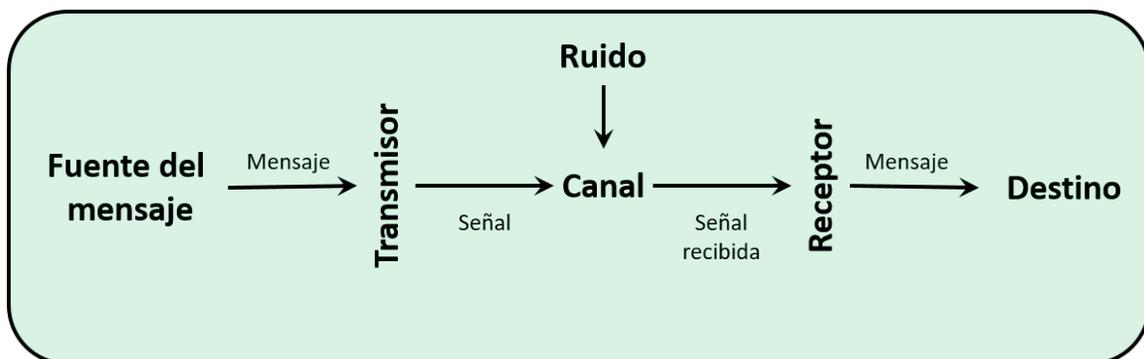


Figura 2.5: Esquema de transmisión. Esquema ideado por Claude E. Shannon

A partir de la teoría de información, sistemas complejos de comunicación pueden ser reducidos a cajas negras, importando únicamente la relación entre la entrada y la salida o la relación estímulo-respuesta para poder caracterizar la cantidad de información que procesa

¹Se refiere al significado del mensaje y su interpretación

²Efectos conductuales generados a partir de la recepción de un mensaje

³En la tecnologías de información se asume que tales unidades son *bits*.

2. FUNDAMENTOS TEÓRICOS

el sistema. Algunos ejemplos en biología en los cuales se han utilizado las herramientas y conceptos descritos en la teoría de información son: redes de procesamiento neuronal [77], redes de regulación genética [78] y vías de señalización [50, 79, 80].

Bajo el paradigma básico de la teoría de información, la información (mensaje modulado) es transmitida en un canal ruidoso de una fuente a un receptor; donde es demodulada y extraída. Aplicando esta idea a las vías de señalización, cualquier señal que estimule a la vía puede ser considerada una fuente, el mensaje es la información sobre el estímulo. El receptor es la molécula final de la cadena de interacción y las reacciones bioquímicas intermedias son el canal ruidoso en donde se transmite la información. Explícitamente la arquitectura de fuente/canal/receptor (Fig. 2.5) está dada por ligando-receptor/vía de señalización/cambios en la expresión de un gen, Fig. 1.2.

Es sabido que la actividad y concentración de las moléculas que intervienen en los sistemas biológicos están bajo efecto de aleatoriedad intrínseca y extrínseca [51, 52, 81]. Intrínsecamente la variabilidad es inducida en los procesos de transcripción [82], en la aleatoriedad de las reacciones bioquímicas y en las múltiples interacciones generadas por otras vías [58]. Extrínsecamente la aleatoriedad viene generada por la influencia del medio como: temperatura, pH y fluctuaciones en las concentraciones de las moléculas efectoras [53, 57]. Particularmente en estos sistemas, la estocasticidad no está restringida a estar localizada en el canal, el estímulo mismo está sujeto a fluctuaciones.

A continuación se desarrollan los conceptos de teoría de información y su intersección en los procesos biológicos en los cuales está involucrada la transmisión de información.

2.4.1 Entropía

La entropía, como un concepto termodinámico, fue introducido por el físico alemán Rudolf Clausius en la segunda mitad del siglo XVIII. Fue originalmente definida como:

$$dS = \frac{Q_{1 \rightarrow 2}}{T}, \quad (2.5)$$

2.4 Teoría de Información

donde dS es un cambio fundamental de entropía, $Q_{1\rightarrow 2}$ es la cantidad de calor intercambiado entre el sistema y el entorno y T es la temperatura absoluta ¹. Con los años, la concepción de entropía de la ecuación (2.5) fue teniendo modificaciones. El físico austriaco Ludwig Boltzmann le dio una interpretación estadística, estableciendo una relación entre la entropía y la probabilidad termodinámica [83]:

$$S = k \cdot \ln(\Omega). \quad (2.6)$$

Donde S representa la entropía, k la constante de Boltzmann y Ω el número de microestados posibles, es decir, la variable aleatoria del sistema (\ln es la función logaritmo neperiano o natural). La ecuación asume que todos los microestados tienen la misma probabilidad de aparecer ², por lo tanto tienen una probabilidad de aparición determinado por una distribución uniforme.

En la termodinámica se estudian partículas con microestados Ω que tienen cierta probabilidad $P(\Omega)$ de aparición. La entropía en el campo de la teoría de la información tiene una estrecha relación con ella, donde los microestados Ω son reemplazados por cualquier otra variable aleatoria. A partir de este momento, se utilizará la concepción de entropía propuesta por Shannon (1948), donde se considera la entropía como un criterio cuantitativo para analizar el grado de desorden de cualquier distribución de probabilidad, el cual puede proveer de información sobre la dinámica de la variable (X) con una distribución de probabilidad $P(x)$ asociada. En otras palabras, una medición del grado de incertidumbre de una variable aleatoria.

Si X es una variable aleatoria discreta con función de distribución de probabilidad $P(x)$, entonces su entropía $H(X)$ es:

$$H(X) = - \sum_x P(x) \log_b(P(x)). \quad (2.7)$$

¹La ecuación (2.5) implica que cuando un sistema termodinámico pasa, en un proceso reversible e isotérmico, del estado 1 al estado 2, el cambio en su entropía es igual a la cantidad de calor intercambiado entre el sistema y el medio dividido por su temperatura absoluta.

²La cantidad de entropía de un sistema es proporcional al logaritmo natural del número de microestados posibles.

2. FUNDAMENTOS TEÓRICOS

La base del logaritmo b indica las unidades que tiene la entropía. Si $b = n$ entonces la entropía estará en *nats*¹. En este trabajo, se definirá la entropía usando base dos ($b = 2$) por lo que todas las unidades de entropía estarán en *bits*. En la ecuación anterior, se considera que $0\log(0) = 0$; dando como resultado la función descrita en la Figura 2.6.

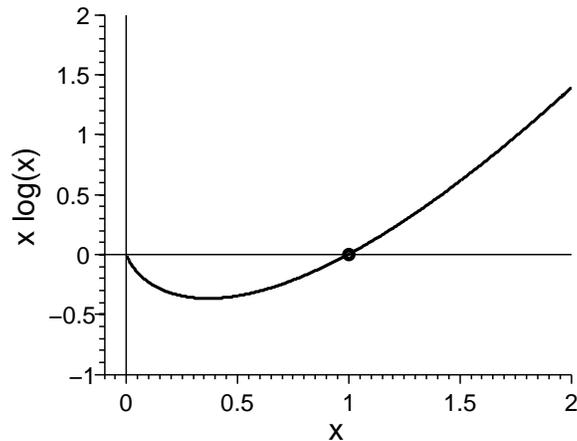


Figura 2.6: Gráfica de $y = x \log(x)$.

La entropía, es una función de la distribución de probabilidad de X . No depende de los valores tomados por la variable aleatoria X , sino de sus probabilidades [84]. De manera que, la entropía puede caracterizar el comportamiento de largas secuencias de variables aleatorias permitiendo estimar las probabilidades de eventos aislados. Para cualquier distribución de probabilidad, es posible definir la entropía.

Extendiendo el concepto de entropía, en un experimento en donde se registre la respuesta en una población celular. Dependiendo del sistema, la respuesta generada se describe con una distribución con parámetros específicos. Asociando el grado de dispersión de la distribución, una distribución ancha posee alta entropía y de manera contraria con una distribución angosta [80], Figura 2.7(a).

¹Natural unit of information. Un nat es la información contenida de un evento si la probabilidad de que dicho evento ocurra es $1/e$.

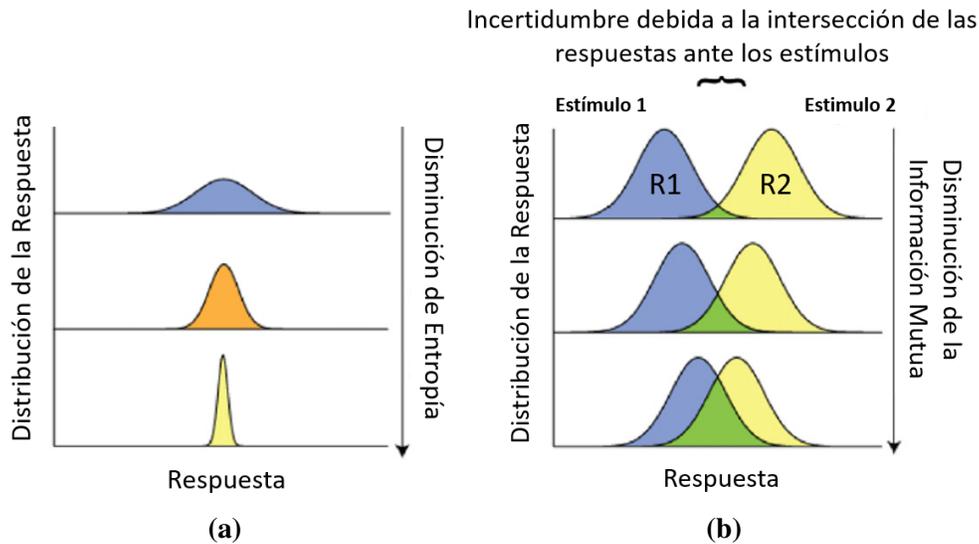


Figura 2.7: Representación de métricas definidas en teoría de información. (a) Entropía. (b) Información Mutua.

2.4.2 Información Mutua

La información mutua de dos variables aleatorias establece su dependencia estadística¹, mide la reducción de la entropía (Ec. (2.7)) debido al conocimiento del valor de otra variable aleatoria [84]. Por ello, la información mutua es igual a la información contenida por las dos variables. A través de la teoría de información se define como información al grado de dependencia estadística entre dos o más variables aleatorias. La información mutua (I) se describe matemáticamente como:

$$\begin{aligned}
 I(R; S) &= H(R) - H(R|S) \\
 &= \sum_{i,j} P(R = r_i)P(S = s_j|R = r_i) \log_2 \frac{P(R = r_i|S = s_j)}{P(R = r_i)}. \quad (2.8)
 \end{aligned}$$

Donde $P(R)$ es la distribución de probabilidad marginal de la respuesta y $P(R|S)$ es la distribución de probabilidad condicional de la respuesta R dado el estímulo S . Por lo tanto, para poder calcular la información mutua es necesario conocer la distribución marginal

¹La información mutua es cero cuando las variables son independientes entre ellas.

2. FUNDAMENTOS TEÓRICOS

del estímulo ($P(S)$) lo cual biológicamente no es viable dado que los experimentos se hacen bajo condiciones controladas fijando el estímulo.

Conceptualmente, la información mutua en la respuesta celular mide la exactitud con la que una población celular responde con una distribución de respuesta única para cada estímulo. De esta manera, si se escoge aleatoriamente un valor en la distribución de las respuestas es posible determinar cuál es el grado de certidumbre de asociar la respuesta con un estímulo específico. Si las distribuciones se superponen, la certidumbre asociada a diferenciar los estímulos decrece, y por consiguiente la información mutua entre el estímulo y la respuesta celular también disminuye [76, 80], Fig. 2.7(b).

2.4.3 Capacidad del canal

Es posible explorar la capacidad máxima de transmitir información por un canal específico, esta cantidad es conocida como la capacidad del canal $CC(R; S)$, matemáticamente está definida como:

$$CC(R; S) = \max_{P(S)} I(R; S) \quad \text{tal que} \quad \begin{cases} \sum_i P(S_i) = 1, \\ P(S_i) \geq 0 \end{cases} \quad (2.9)$$

La capacidad del canal es la información mutua maximizada sobre todas las posibles distribuciones de probabilidad del estímulo, siendo efectivamente el ancho de banda¹ que es una propiedad intrínseca del canal. En comparación, la información mutua es la cantidad de información transmitida en un canal de comunicación para una distribución de estímulo específica ($P(S)$), lo cual hace que la métrica no sea completamente una característica intrínseca del canal. La capacidad del canal obtiene una gran relevancia en el teorema de codificación de canales con ruido. El teorema establece que es posible transmitir un mensaje a través de un canal con ruido y ser diferenciado de otro mensaje con un error arbitrario, con la restricción de que la entropía del estímulo en la entrada sea menor que la capacidad del canal [85]. El teorema delimita el límite superior de la cantidad de información para

¹Cantidad de estímulos o entradas que un canal puede procesar [84].

obtener una transmisión exitosa.

Para los canales de comunicación en las vías de señalización, la probabilidad condicional de la respuesta $P(R|S)$ se puede determinar experimentalmente en condiciones controladas midiendo la distribuciones de las respuestas ante estímulos de interés; esto implica fijar el estímulo artificialmente. Por otro lado, $P(S)$ determina la frecuencia de ocurrencia natural de los estímulos. La cual no es fácil obtener experimentalmente. Sin embargo, para sistemas biológicos la capacidad del canal establece una noción sobre la máxima cantidad de información transmitible. Bajo la hipótesis de codificación eficiente¹ se asume que los sistemas biológicos cuya función primaria está la transmisión de información, se consideran que evolutivamente son óptimos para responder ante las fuentes que los estimulan [76, 89].

2.4.4 Información Mutua Multivariable

En las secciones anteriores se expone la manera de obtener la cantidad de información transmitible por un canal, las formulaciones presentadas con anterioridad son para dos variables. Aún así, la información mutua se puede expandir a múltiples dimensiones o variables sin que pierda su esencia. Para ello, se introduce el concepto de información mutua multivariable (MMI²), Ec. (2.10), la cual cuantifica la información estadística que es transferida entre las múltiples variables de interés.

$$\text{MMI}(R_1, \dots, R_n; S) = \sum_{r_1, \dots, r_n, s} P(r_1, \dots, r_n, s) \log_2 \frac{P(r_1, \dots, r_n, s)}{P(r_1, \dots, r_n)P(s)} \quad (2.10)$$

¹La hipótesis de eficiencia de código fue propuesta por Horace Barlow en 1961 como un modelo teórico de la codificación neuronal [86]. En el cerebro, las neuronas se comunican entre ellas a través de impulsos eléctricos generados por los potenciales de acción. La hipótesis de Barlow establece que las neuronas usan el código más eficiente posible para representar la información de tal manera que se disminuya el número de impulsos necesarios para cada señal. Tal propuesta ha sido ampliamente explorada en códigos neuronales para la visión y sonidos [87]. A su vez, el estudio en neuronas se ha extendido a la activación de sus vías de señalización y no únicamente al registro de los potenciales de acción [88].

²Por sus siglas en inglés *Multivariate Mutual Information*.

2. FUNDAMENTOS TEÓRICOS

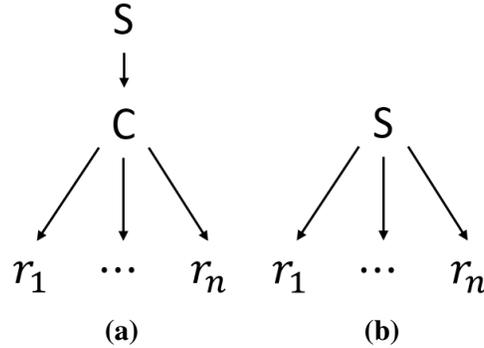


Figura 2.8: Modelos de activación. (a) Múltiples salidas (r) comparten el mismo tronco de procesamiento (C) ante un estímulo (S). (b) Modelo ramificado, el estímulo activa las respuestas por vías y mecanismos diferentes.

Para realizar el cálculo de MMI es necesario conocer la probabilidad conjunta de todas las variables. Para ello, se utiliza la regla de la cadena de probabilidad, Ec. (2.11).

$$P(r_1, \dots, r_n, s) = \prod_{k=1}^n P\left(r_k \mid \bigcap_{j=1}^{k-1} r_j, s\right) \quad (2.11)$$

En el caso particular del estudio de vías de señalización y del presente estudio, se utiliza la MMI para cuantificar la información que múltiples respuestas (R) poseen sobre un estímulo específico (S) de manera no redundante; es decir, la cantidad de información única y diferente que la respuesta en conjunto tiene sobre el estímulo. Las múltiples respuestas pueden ser generadas de dos maneras: la primera cuando un estímulo activa una vía específica con un tronco común de procesamiento para después diferenciarse en las múltiples respuestas, Fig. 2.8(a). La segunda se da por medio de vías independiente, Fig. 2.8(b).

2.4.5 Desigualdad de procesamiento de datos

Partiendo de una suposición, en la cual se tiene un evento X , que corresponde a la unión de dos moléculas. Con fundamento en el evento, se quiere predecir la cantidad de tiempo

que ambas moléculas permanecen unidas. La predicción únicamente se puede hacer analizando la colección de datos que se registraron del evento (Y). A su vez, la información en Y necesita ser procesada y transmitida para poder ser evaluada, siendo Z la representación de estos eventos, Fig. 2.9. Dependiendo de las técnicas de registro y fiabilidad del procesamiento de los datos es la certeza con que se podrá hacer la predicción.

Existe una pérdida de información debido a que para el registro se debe discretizar la información. A su vez, al ser procesada y transmitida la información necesita una nueva representación, por ejemplo ser representada en bits, acarreado una nueva pérdida de información. En consecuencia, se tiene menos información sobre la medición original y por lo tanto del evento de interés. El proceso es representado gráficamente en la figura 2.9.

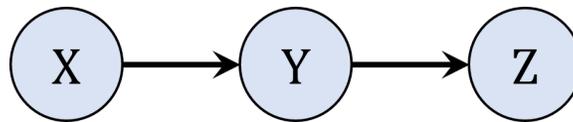


Figura 2.9: Desigualdad de procesamiento de datos.

Donde Y contiene toda la información de X para determinar Z . Z no puede tener más información de X que Y . Matemáticamente se describe por la Ec. 2.12, donde $I(X : Y)$ es información mutua entre X y Y . $I(X : Z)$ es la información que comparten X y Z .

$$I(X : Y) \geq I(X : Z) \quad (2.12)$$

Lo anterior implica que se pierde información al momento de procesarla. En el mejor de los casos, se conserva la misma cantidad de información llevando la ecuación anterior a una igualdad. En procesos de transmisión, la igualdad se asocia a sistemas que no están bajo influencia de ruido y con una alta precisión de procesamiento. Sin embargo, en la mayoría de los casos hay una desigualdad.

2.5 Variables Aleatorias

Los canales de comunicación están sujetos a ruido inevitablemente, esto ocasiona la pérdida de información [50, 84]. A pesar de todas las fuentes de variabilidad, las vías de

2. FUNDAMENTOS TEÓRICOS

señalización desencadenan respuestas celulares específicas y únicas. Para caracterizar y estudiar el proceso de transmisión de información en vías de señalización celular; y entender la manera en que las células lidian con el ruido, es indispensable considerar el ruido en el sistema.

Como resultado de la variabilidad en la realidad biológica, las variables matemáticamente son descritas como variables aleatorias. Una variable aleatoria es una función definida en un espacio probabilístico, dentro de sus propiedades principales es que tienen asociada una distribución de probabilidad. Una manera de describir las propiedades de una variable aleatoria es caracterizar la distribución de probabilidad por la que es gobernada. La caracterización se puede hacer mediante los parámetros estadísticos, entre ellos la esperanza o promedio (μ) y la varianza (σ^2), entre otros. Utilizando estos dos parámetros, mediante el coeficiente de variación (CV) es posible cuantificar la dispersión normalizada con respecto a la media, Ec. 2.13.

$$CV = \frac{\sqrt{\sigma^2}}{\mu} \quad (2.13)$$

La dispersión en una colección, establece el rango de valores posibles que puede tomar. Específicamente, el rango dinámico es la relación entre el valor menor y mayor. A menor valor de CV existe una dispersión menor, por lo tanto la distribución es más estrecha estableciendo un rango dinámico menor. De manera contraria para valores grandes del CV.

2.5.1 Representación Estocástica

Particularmente, en bioquímica las variables aleatorias son discretas¹ ya que las reacciones determinan un cambio en el número de moléculas presentes, ya sea que se transformen, produzcan o degraden. En la figura 2.10 se ejemplifica este proceso en el cual la variable transita entre varios estados de acuerdo a las reacciones que la afecten. Partiendo de un

¹Una variable aleatoria discreta es una variable que dentro de su función de probabilidad sólo está definida en ciertos valores del conjunto, biológicamente tiene sentido ya que no es posible concebir parcialidades de moléculas.

2.5 Variables Aleatorias

estado inicial cada flecha de color diferente es una reacción que puede llevar a la molécula a cierto estado discreto en el espacio.

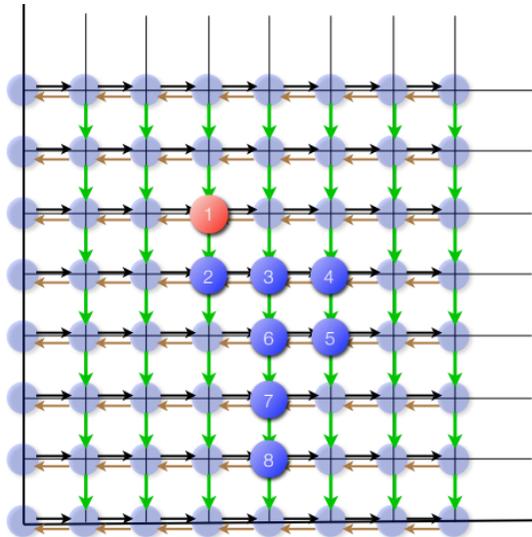


Figura 2.10: Proceso aleatorio.

La probabilidad de que suceda una u otra reacción depende de la función de transición asociada a la reacción. Para ejemplificar, en la figura 2.11 se muestra una variable aleatoria A que modifica su conteo molecular denotado por n . A puede salir del estado donde existe un número n_A mediante la producción de una molécula llevando a un nuevo estado $n_A + 1$ o puede degradarse disminuyendo su número una unidad $n_A - 1$. La función de transición para que se produzca y degrade una molécula A es $f_1(\alpha)$ y $f_2(\alpha)$, respectivamente¹. Estas funciones reciben el nombre de propensiones.

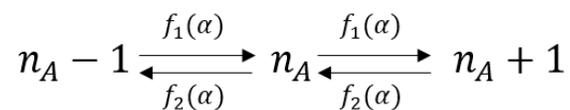


Figura 2.11: Transición aleatoria.

¹Las funciones de transición pueden tener dependencia únicamente paramétrica o de otras variables aleatorias.

2. FUNDAMENTOS TEÓRICOS

Dado lo anterior, la probabilidad de que el sistema se encuentre en un conteo n_A en el tiempo $t + 1$ depende directamente de la evolución temporal del sistema, de manera general la razón de cambio de la probabilidad en el tiempo para un conteo molecular n_A está determinada por la Ec. (2.14).

$$\frac{dP(n_A, t)}{dt} = [(n_A - 1) P(n_A - 1, t) - n_A P(n_A, t)] f_1(\alpha) + [(n_A + 1) P(n_A + 1, t) - n_A P(n_A, t)] f_2(\alpha) \quad (2.14)$$

La ecuación anterior se compone de cuatro términos, el primero se asocia a la probabilidad de tener un conteo $n_A - 1$ y que se produzca una molécula con el fin de que se llegue al estado deseado n_A . Más aún, existe la probabilidad de que al tiempo t ya se encuentre el sistema con un conteo n_A y que se produzca una molécula llevando al estado $n_A + 1$. Los siguientes dos términos son los asociados a que se degrade una molécula llevando al conteo n_A o $n_A - 1$.

La ecuación antes descrita tiene el nombre de la ecuación química maestra, se ha logrado resolver esta ecuación para sistemas sencillos. Sin embargo existen algoritmos estocásticos que permiten dar solución de manera numérica. De manera particular, Gillespie (1977) estableció un algoritmo que converge a la solución exacta de la ecuación química maestra, esto se explica con mayor detalle en la Sección 3.3.

Metodología

En esta sección se describe la metodología utilizada para caracterizar las propiedades de los mecanismos demoduladores en vías de señalización. Se incluyen los diferentes algoritmos ocupados en la modelación, resolución y establecimientos de las condiciones para los diferentes experimentos *in silico*¹.

La Figura 3.1 muestra de manera condesada los pasos empleados para realizar la caracterización. Se inicia con la definición de una entrada o estímulo para bloque demodulador, el cual se modeló estocásticamente registrando la respuesta ante cada estímulo. Los registros ante varias condiciones se ocuparon para obtener métricas que definan la calidad del proceso de transmisión de información. En las siguientes secciones, cada etapa se describe detalladamente.

3.1 Definición de entrada

Nakakuki et al. (2010) estudiaron la vía de señalización MAP/ERK. En sus experimen-

¹Experimentos implementados bajo simulación computacional.

3. METODOLOGÍA

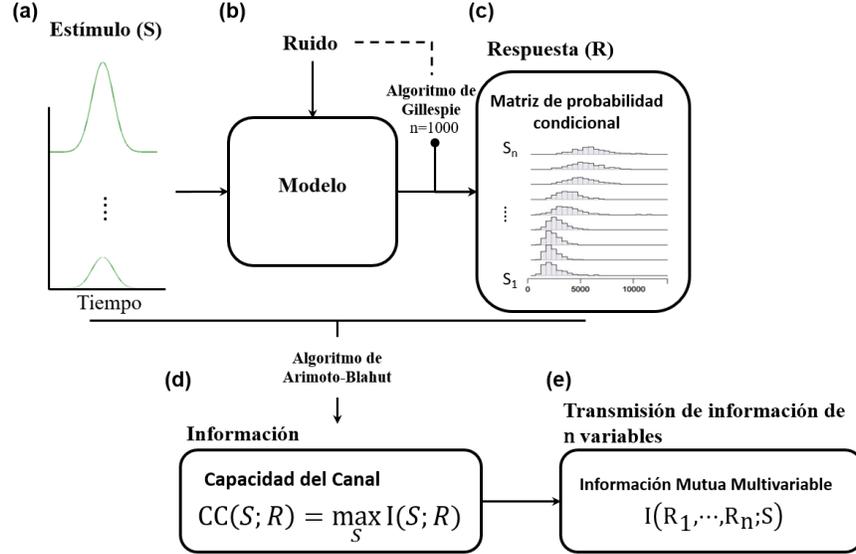


Figura 3.1: Metodología. (a) Definición de la entrada. (b) Modelo estocástico de la vía de señalización MAP/ERK o el modelo mínimo. (c) Estimación de la probabilidad condicional de la respuesta. (d) - (e) Estimación de métricas definidas en la teoría de información.

tos, estimularon células con diferentes factores de crecimiento registrando la dinámica de $ppERK_{cyt}$. La cual es la salida del bloque modulador y la entrada del bloque demodulador, de ahí la relevancia para el presente proyecto dado que uno de los principales objetivos es estudiar las propiedades del bloque demodulador (Fig. 2.4). A su vez, en el mismo estudio, basados los datos experimentales caracterizaron matemáticamente la dinámica de la vía de señalización MAP/ERK. El modelo matemático como resultado del trabajo de Nakakuki et al. es presentado con mayor detalle en el capítulo 4. La descripción de $ppERK_{cyt}$ en su forma determinista representa su concentración promedio en función del tiempo (Ec. (3.1), la cual se deriva de resolver la Ec. (4.1)). A su vez, la caracterización de $ppERK_{cyt}$ en el trabajo citado fue utilizando dosis diferentes (0.1nM, 1nM y 10nM) del factor de crecimiento epidermal.

$$ppERK_{cyt} = L \left[k_2 \left(e^{-\frac{t}{\tau_2}} - 1 \right) - k_1 \left(e^{-\frac{t}{\tau_1}} - 1 \right) \right] \quad (3.1)$$

La dinámica de $ppERK_{cyt}$ está modelada por la diferencia de dos funciones exponen-

3.1 Definición de entrada

Tabla 3.1: Valores de parámetros para definir la entrada.

	Parámetro	Valor		
		I_1	I_2	I_3
EGF	κ_1	2.19	1.23	0.257
	κ_2	1.97	0.949	0.101
	τ_1	3.16	1.0	9.68
	τ_2	8.68	3.69	6.3

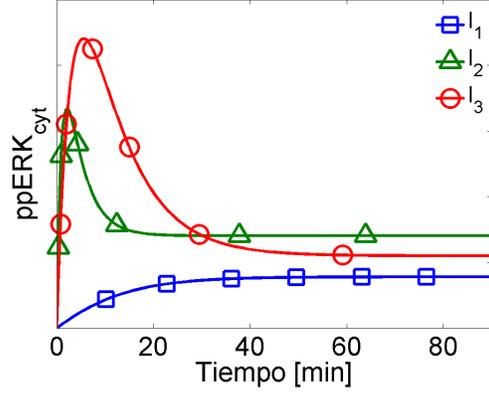


Figura 3.2: Definición de entradas con $L = 100$.

ciales con diferente factor de amplificación (k_i) y tasa de decaimiento (τ_i). L es el factor de escalamiento, este se ocupa dado que la caracterización experimental del sistema [1] está basada en los registros de expresión a partir de fluorescencia, por lo que las unidades son relativas y arbitrarias. Los parámetros para obtener cada entrada son presentados en la Tabla 3.1. Las diferentes dinámicas ante cada concentración se presentan en la figura 3.2, donde I_1 , I_2 y I_3 corresponden a las concentraciones de 0.1nM, 1nM y 10nM, respectivamente. Dinámicamente $ppERK_{cyt}$ tiene un estado estacionario (A_N) determinado por la Ec. (3.2), todos los parámetros son positivos; dado que biológicamente la concentración tiene que ser positiva $k_1 \geq k_2$.

$$A_N = \lim_{t \rightarrow \infty} ppERK_{cyt} = L(k_1 - k_2) \quad (3.2)$$

Dado que la descripción de $ppERK_{cyt}$ es fenomenológica, implícitamente se tiene el efecto del bloque modulador. Se sabe que el ruido bioquímico afecta a este bloque también. Por lo que al modelar estocásticamente a $ppERK_{cyt}$ se está considerando la variabilidad inducida en el primer bloque. La descripción estocástica de $ppERK_{cyt}$ se caracterizó de forma inicial al derivar la Ec. (3.1) dando como resultado la siguiente ecuación:

$$\frac{dppERK_{cyt}}{dt} = L \left[\frac{k_1}{\tau_1} e^{-\frac{t}{\tau_1}} - \frac{k_2}{\tau_2} e^{-\frac{t}{\tau_2}} \right]. \quad (3.3)$$

3. METODOLOGÍA

En la cual hay dos términos con signo opuesto. Se asume cada término como la representación de procesos de activación, y desactivación de $ppERK_{cyt}$ de acuerdo al signo. Con esta descripción la entrada se modela en función de sus propensiones (Tabla 3.2) y se resuelve a través del algoritmo de Gillespie (Sección 3.3).

Tabla 3.2: Propensiones de la entrada.

Parámetro	Propensión	Descripción
$\emptyset \rightarrow ppERK_{cyt}$	$\frac{\kappa_1 L}{\tau_1} e^{-\frac{t}{\tau_1}}$	Activación de $ppERK_{cyt}$
$ppERK_{cyt} \rightarrow \emptyset$	$\frac{\kappa_2 L}{\tau_2} e^{-\frac{t}{\tau_2}}$	Desactivación de $ppERK_{cyt}$

Al modelar de esta manera $ppERK_{cyt}$, se está suponiendo que la activación y la desactivación están determinadas por una reacción, respectivamente. Es sabido que ambos procesos son llevados a cabo en múltiples etapas. Sin embargo, la suposición hecha es plausible dado lo siguiente: en la sección 2.3 se describió el proceso por el cual $ppERK_{cyt}$ es activada, el proceso se resume como una cadena fosforilación; Raf regula la actividad de ERK mediante la cinasa MEK, se ha reportado experimentalmente que la concentración de Raf es menor comparada con la de MEK [91, 92] por lo que en el paso de Raf a MEK existe amplificación de la respuesta en ERK. A su vez, en cada proceso de la cadena de fosforilación cada enzima siguiente está presente en exceso en comparación con su enzima que la activa [93]. Además de la implicación en la amplificación de la respuesta, existe una respuesta más rápida en la activación de ERK, que comparada con el tiempo de escala del bloque demodulador (bloque del cual ERK es la entrada) es ordenes magnitud superior, esto se ha demostrado experimentalmente [1].

Por su parte, Nakakuki et al. (2010) determinaron que el proceso de defosforilación de ERK tanto en el núcleo como en el citoplasma se resume en 10 reacciones regidas por Dusp; a su vez, compararon el efecto de tener las 10 reacciones y sólo una que las englobara. Determinaron que la descripción con las dos aproximaciones son representativamente iguales, esto debido a las altas concentraciones de Dusp y a que las velocidades de las diferentes reacciones son rápidas, por lo que el sistema se puede considerar en equilibrio químico.

3.2 Modelo

Los modelos matemáticos ocupados en el presente trabajo se presentan de forma detallada en el capítulo 4, principalmente se modeló estocásticamente el bloque demodulador (Fig. 2.4) de la vía de señalización MAP/ERK basados en la caracterización determinista realizada por Nakakuki et al. (2010). Para explorar a más detalle las propiedades de los mecanismos demoduladores se implementó un modelo mínimo con interacciones topológicas tipo FFL. Finalmente se modeló la expresión de genes reporteros para evaluar la cantidad de información transmitida en diferentes puntos de la vía de señalización. Por otro lado, a partir de estos genes se evaluó el segundo mecanismo demodulador basado en la afinidad del promotor por su factor de transcripción.

3.3 Algoritmo de Gillespie

En reacciones en donde el conteo molecular son del orden del número de Avogadro, las tasas de reacción representan el número de veces que la reacción ocurre por unidad de tiempo. Sin embargo, para conteos moleculares pequeños, a dichas tasas (llamadas propensiones) se les asocia una interpretación probabilística: la probabilidad dada de ocurrencia de la reacción correspondiente por unidad de tiempo. Para propensiones pequeñas, el número de reacciones ocurridas en diferentes realizaciones del mismo experimento puede variar drásticamente. Debido a eso, una forma de estudiar la dinámica del sistema es hacer uso de simulaciones estocásticas tipo Montecarlo por medio del algoritmo de Gillespie [90].

A través del algoritmo de Gillespie se recrea la dinámica probabilística en la ocurrencia de reacciones químicas, dicha aleatoriedad es de naturaleza intrínseca. Uno de los requisitos del algoritmo es abstraer el sistema a un conjunto de reacciones químicas y se puede determinar sus respectivas propensiones de ocurrencia para cada reacción. Dentro de los fundamentos del algoritmo se basa en suponer que el sistema se define como un proceso de Markov, indicando que el estado futuro depende únicamente del estado actual del sistema y no del pasado; por lo tanto, el sistema no tiene memoria.

3. METODOLOGÍA

El conjunto de reacciones derivadas del proceso de interés se establece en la Tabla 4.1, con sus respectivas propensiones de ocurrencia dependiendo el modelo usado.

Algoritmo de Gillespie

1. Inicializar el conteo molecular del sistema, y calcular las propensiones.
2. Aleatoriamente calcular el tiempo de espera para la ocurrencia de la siguiente reacción tomando en cuenta que obedece una distribución exponencial con parámetro igual a la propensión total.
3. Se escoge aleatoriamente cuál es la siguiente reacción a ocurrir, tomando en consideración que la probabilidad de que suceda la reacción es proporcional a su propensión.
4. Actualización del tiempo del sistema y del conteo molecular de acuerdo a los resultados obtenidos en los pasos anteriores.
5. Se regresa al paso 2 si el tiempo de simulación no ha concluido.

3.4 Distribución de Probabilidad Condicional

Dado la naturaleza estocástica de la realidad biológica y de la forma en que se está modelando, las respuestas no son deltas de Dirac que corresponden a una situación sin variabilidad asociando un único posible valor de respuesta, Fig. 3.3(a). Al ser variables aleatorias tienen asociadas una distribución de probabilidad, que asigna un valor de probabilidad a cada estado posible, Fig. 3.3(b). Una forma de aproximar la distribución de probabilidad es a través de una distribución de frecuencia¹. Una distribución de frecuencia es la agrupación de datos en categorías excluyentes indicando el número de observaciones en cada categoría². La cual, tiene la propiedad de que la suma de todos los elementos en las categorías es igual al total de número de datos n , donde m_j es el conteo dentro de cada categoría j ,

¹La representación gráfica de la distribución de frecuencia comúnmente se llama histograma

²Las categorías se establecen como rangos numéricos de agrupación.

3.4 Distribución de Probabilidad Condicional

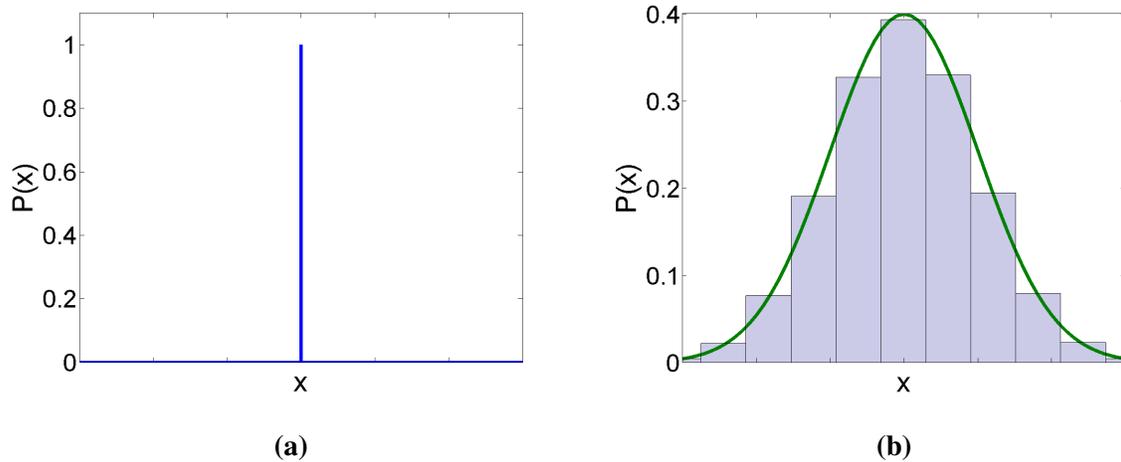


Figura 3.3: Diferentes distribuciones de probabilidad. (a) Delta de Dirac que asocia un único valor de respuesta, este tipo de distribuciones está asociado en variables sin variabilidad y representa a los sistemas deterministas. (b) Representación de una distribución de probabilidad.

siendo k la cantidad total de categorías, Ec. (3.4).

$$n = \sum_{j=1}^k m_j \quad (3.4)$$

Las distribuciones de probabilidad se pueden aproximar mediante la distribución de frecuencia; sin embargo, es necesario normalizar en función de la densidad de la distribución tomando en cuenta el rango de las categorías. La densidad se obtiene integrando la distribución de frecuencias, numéricamente se calculó mediante la regla de los trapecios¹; como resultado se obtiene la distribución de probabilidad discreta.

La caracterización previa se utiliza para registrar la respuesta del sistema ante un estímulo específico. Por lo tanto, la distribución de respuesta es dependiente del estímulo, matemáticamente se le llama distribución condicional $(P(R|S))^2$. Por lo que para cada es-

¹ $\int_a^b f(x)dx \sim \frac{b-a}{n} \left[\frac{f(a)+f(b)}{2} + \sum_{k=1}^{n-1} f(a + k\frac{b-a}{n}) \right]$.

²Representa la probabilidad de la respuesta R dado que ocurre el estímulo S .

3. METODOLOGÍA

tímulo se obtiene una distribución condicional. El conjunto de distribuciones condicionales se agruparon en una matriz que se le denominó matriz de distribución condicional, donde cada reglón es una condición diferente.

La matriz resultante es rectangular lo cual implica que tiene el mismo número de categorías en cada distribución de frecuencia. Para determinar los valores extremos de las distribuciones se obtuvo el rango mayor ante cada estímulo. El valor mayor se ocupó para establecer el tamaño de la matriz, la categorías vacías se llenan con cero. Gráficamente la matriz se representa como el acomodo de cada histograma de R ante S , Fig. 3.1 (c). Por otro lado, una distribución de probabilidad discreta se aproxima a una continua cuando en límite del ancho del rango tiende a cero. El ancho de cada categoría ocupado es de una unidad, ya que es el valor mínimo que puede tomar dado que se trabaja con conteos moleculares; siendo la mejor aproximación a la distribución continua de la respuesta.

3.5 Cálculo de la Capacidad del Canal

En la sección 2.4 se presentan los fundamentos teóricos sobre la manera de caracterizar la información que se transmite en un canal. La información mutua es la reducción de la incertidumbre del estímulo dado que se conoce la salida; por lo tanto, determina la cantidad de información que se transfirió del estímulo a la respuesta. La información mutua está descrita por la ecuación (2.8), para su cálculo se necesita el conocimiento de la distribución de probabilidad $P(S)$ del estímulo S . En el área de la biología esta información es limitada ya que implica la medición de la frecuencia con que ocurren los estímulos. Una alternativa a esto es calcular la capacidad del canal siendo la maximización de la I sobre todas las distribuciones de probabilidad de S , por lo que el cálculo de la CC se vuelve un problema de optimización que puede ser resultó de forma computacional, Ec. (2.9).

3.5.1 Algoritmo de Arimoto-Blahut

El algoritmo de Arimoto-Blahut resuelve el problema de encontrar numéricamente la distribución de probabilidad del estímulo $P(S)$ que maximiza la información mutua. Se parte

3.5 Cálculo de la Capacidad del Canal

de que la información mutua es una función concava en R , por lo tanto el problema se reduce en encontrar el valor del máximo. En la Figura 3.4 se esquematiza esta situación. Mediante un proceso iterativo se aproxima una función que converge monotónicamente al valor máximo.

A través del algoritmo de Arimoto-Blahut se asume que en la iteración l se encuentra la distribución del estímulo $P(S)^{(l)}$ con una correspondiente información mutua de $I(S^{(l)}; R)$ (Fig. 3.4). En la iteración $l + 1$ se encontrará una $P(S)^{(l+1)}$ que se aproxime al máximo global. Por lo que $I(S^{(l+1)}; R) \geq I(S^{(l)}; R)$. Para optimizar la búsqueda se introduce la función $\Psi(R^{(l)}, R, W)$ que localmente (alrededor de $P(S) = P(S)^{(l)}$) aproxima $I(S, R)$ [84].

Se espera que $P(S)$ parta de un estado de alta entropía o incertidumbre (Fig. 3.5(a)) e iterativamente evolucione hacia un estado de baja entropía que permita optimizar la cantidad de información transmitida entre el estímulo y la respuesta (Fig. 3.5(b)).

3. METODOLOGÍA

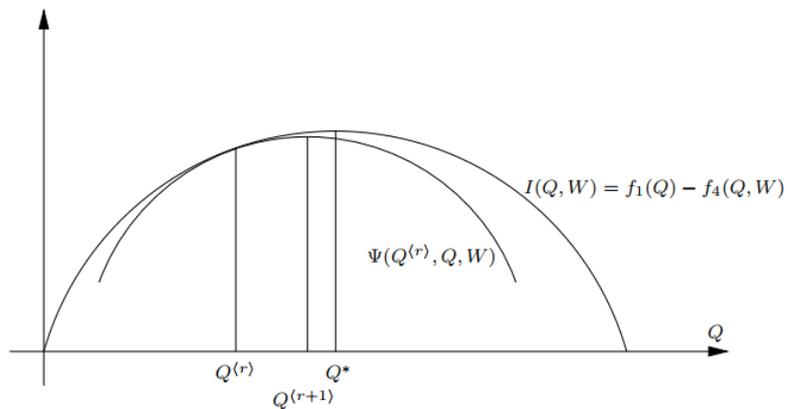


Figura 3.4: Algoritmo de Arimoto-Blahut. Información mutua de un canal genérico $I(R; S)$ y una función de aproximación $\Psi(R^{(l)}, R, W)$

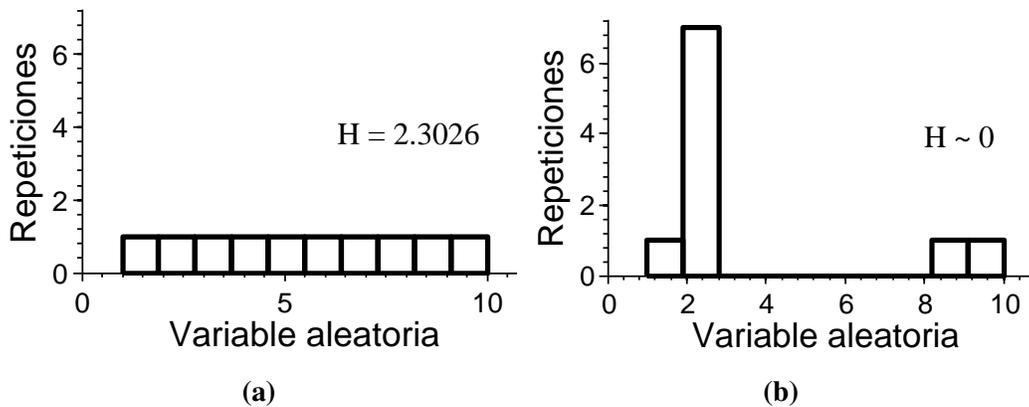


Figura 3.5: Evolución de la distribución S dado el algoritmo de Arimoto-Blahut. (a) Distribución uniforme, alta entropía. (b) Distribución focalizada, la entropía tiende a cero.

3.6 Cálculo de la Información Mutua Multivariable

Algoritmo de Arimoto-Blahut para calcular CC

1. Se establece

$$\begin{aligned} P(r) &\rightarrow q(r) \\ P(r|s) &\rightarrow Q(r|s) \end{aligned}$$

de tal forma que expandiendo las ecuaciones (2.8) y (2.9), con las sustituciones anteriores se obtiene que

$$CC(R; S) = \max_{P(S)} \sum_{i,j} q(r_i) p(r_i|s_j) \log_2 \frac{Q(r_i|s_j)}{q(r_i)}$$

2. Se establece $l = 0$ y se escoge una distribución de probabilidad para $q^{(0)}$ de tal manera que $q^{(0)}(i) > 0$ para todo i . Es recomendable usar una distribución uniforme ya que posee la máxima entropía 3.5(a).

3. Calcular

$$\begin{aligned} Q^{(l)}(r|s) &= \frac{q^{(l)}(r) p(s|r)}{\sum_{r'} q^{(l)}(r') p(s|r')} \\ q^{(l+1)}(r) &= \frac{\prod_s Q^{(l)}(r|s)^{p(s|r)}}{\sum_{r'} \prod_s Q^{(l)}(r'|s)^{p(s|r')}} \end{aligned}$$

4. Establecer $l = l + 1$ y regresar al paso 3 de tal forma que

$$q^{(1)}(r) \rightarrow Q^{(1)}(r|s) \rightarrow q^{(2)}(r) \rightarrow Q^{(2)}(r|s) \rightarrow \dots$$

5. Si $|q^{(n)} - q^{(n-1)}| \leq \varepsilon$ se considera que existe convergencia y la iteración termina; donde $\varepsilon = 10^{-6}$.

3.6 Cálculo de la Información Mutua Multivariable

A partir de la información mutua multivariable se calcula la cantidad de información no redundante contenida en un conjunto de variables (Sección 2.4.4), matemáticamente es-

3. METODOLOGÍA

tá definida de acuerdo a la ecuación (2.10). Su cálculo genera la necesidad de conocer complejos términos de probabilidad conjunta y condicional (Ec. (2.11)). Sin embargo utilizando la naturaleza de la estructura de la vías de señalización y la independencia entre las respuestas (Fig. 2.8) los términos se simplifican ya que los múltiples sistemas (en el caso particular del trabajo, genes) son independientes entre sí pero dependientes con respecto al estímulo. De tal forma que únicamente se necesita el conocimiento de las probabilidades condicionales (Ecs. (3.5)- (3.6)).

$$P(r_1, \dots, r_n, s) = P(s) \prod_{k=1}^n P(r_k | s) \quad (3.5)$$

$$P(r_1, \dots, r_n) = \prod_{k=1}^n P(r_k) = \prod_{k=1}^n \left(\sum_{s'} P(s') P(r_k | s') \right) \quad (3.6)$$

Se aprecia que aún se necesita la probabilidad del estímulo $P(S)$, sin embargo está se obtiene del cálculo de la capacidad del canal a través del algoritmo de Arimoto-Blahut, sección 3.5.1.

3.7 Ruido Extrínseco

Dado que a partir del algoritmo de Gillespie se representa únicamente el efecto del ruido intrínseco, y dentro de los intereses del presente trabajo fue considerar la componente extrínseca, se modeló la fuente de variabilidad externa. Como ya se mencionó en la introducción, el ruido extrínseco afecta las tasas a las cuales ocurre la fosforilación induciendo fluctuaciones en las mismas en función del tiempo [53, 57], por lo que al variar la tasa se emula la variación en las condiciones del entorno biológico, Fig. 3.6(a). Se supuso que la tasa de fosforilación en el modelo de la vía MAP/ERK (representado por el parámetro k_{11}) fluctúa aleatoriamente en el tiempo con una frecuencia específica. De manera simple, se calcula recursivamente el valor del parámetro respectivo de forma aleatoria; la distribución de probabilidad que rige el cálculo, se establece que es log-normal centrada en el valor nominal del parámetro [94], el valor del parámetro se muestra en la Tabla 4.1.

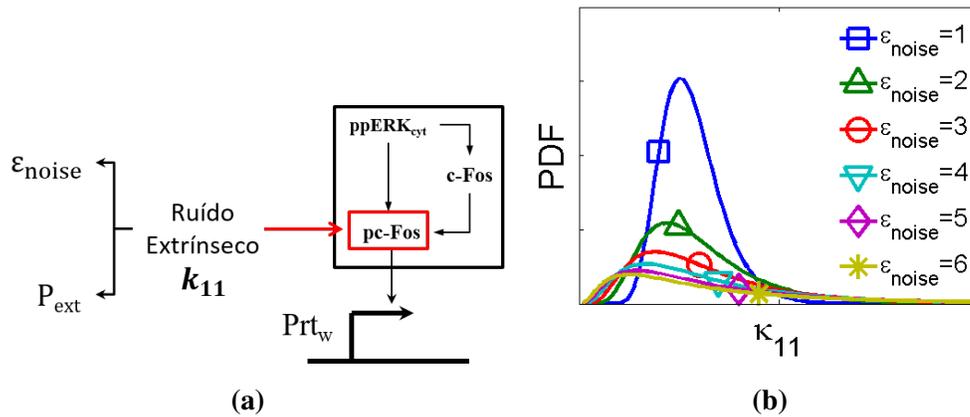


Figura 3.6: Fuente de ruido extrínseco. (a) La fuente de ruido extrínseco controla la tasa de fosforilación de *pc-Fos*; tiene dos parámetros de control: $\varepsilon_{\text{noise}}$ y P_{ext} . (b) Función de densidad de probabilidad log-normal en función de $\varepsilon_{\text{noise}}$ para determinado el valor de k_{11} .

Se desea controlar los rangos de fluctuación de k_{11} ; sin embargo, no hay evidencia experimental de los rangos de variación para los parámetros que representan el proceso de fosforilación, para el presente trabajo se supuso que los rangos de fluctuación eran de dos ordenes de magnitud superiores e inferiores al valor nominal. Se asoció de manera directa la intensidad de fluctuación y los rangos de variación. A mayor intensidad inducida se espera que el rango de valores en la distribución de probabilidad sea mayor. La distribución de probabilidad log-normal posee un parámetro que controla la forma de la distribución. Tal parámetro se definió como $\varepsilon_{\text{noise}}$, con el cual al tener un valor mayor la distribución resultante es más amplia controlando la inducción de ruido extrínseco al sistema, Fig. 3.6(b). Por otro lado, se mencionó que se recalculan los parámetros de forma recursiva, por lo que se controla el periodo (P_{ext}) con el cual un nuevo valor se genera para k_{11} .

3.7.1 Algoritmo Extrínseco

Para incluir el efecto del ruido extrínseco, el bloque (b) de la Fig. 3.1 se cambió por el descrito en la Fig. 3.7. La fuente extrínseca está determinada por dos parámetros de control $\varepsilon_{\text{noise}}$ y P_{ext} , el primero determina la forma de la distribución con la que k_{11} varía; el

3. METODOLOGÍA

segundo establece la frecuencia con la cual un nuevo valor de las tasa es recalculado. Ya determinada la forma de la distribución y la frecuencia de estímulo, se generan N ventanas de tiempo de acuerdo a la Ec. (3.7) del tamaño de P_{ext} , en función del tiempo de simulación (t_f) igual a 200 min.

$$N = \frac{t_f}{P_{\text{ext}}} \quad (3.7)$$

Para cada ventana se calcula un nuevo valor de k_{11} a partir del computo de un número aleatorio con la distribución ya determinada. Dentro de la ventana el sistema se simula por medio del algoritmo de Gillespie.

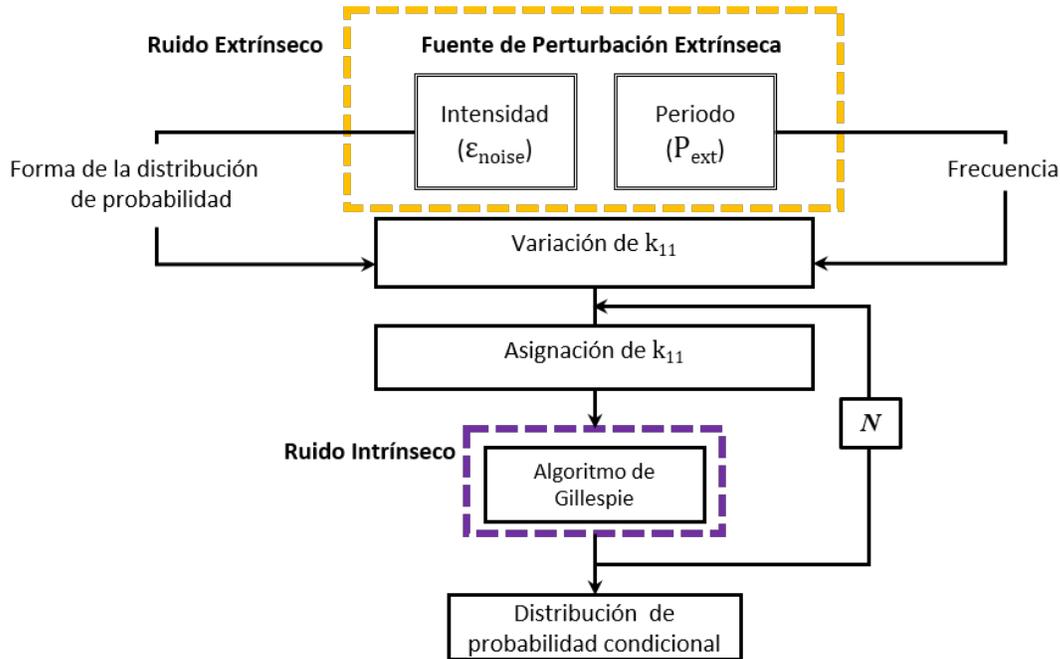


Figura 3.7: Algoritmo Extrínseco.

3.8 Simulación

Para resolver los modelos estocásticos en el dominio temporal se usó el algoritmo de Gillespie [90]. La simulación estocástica tuvo una duración de 200 min. con 1000 realizaciones

3.8 Simulación

ante cada condición. Las condiciones iniciales para todas las especies fueron igual a cero en todos los modelos y condiciones. Para el cálculo de las distribuciones de frecuencia, se registró el estado del sistema cada minuto de simulación, y se normalizan¹. Todas las distribuciones se agrupan en la matriz de probabilidad condicional, donde cada renglón es la distribución condicional de la respuesta dado el estímulo $P(S|R)$ para cada valor de L teniendo un rango de $[10, 120]$ con doce intervalos. Para cada condición en cada modelo se obtienen 200 matrices de probabilidad condicional.

Para obtener la cantidad de información que se transmite, se usó el algoritmo de Arimoto-Blahut [84]. El cual, utiliza la matriz de probabilidad condicional. Como condición inicial del algoritmo de Arimoto-Blahut, se establece que la distribución de probabilidad inicial del estímulo $P(S)$ es uniforme, siendo la distribución con máxima entropía. Se obtiene un valor de CC por minuto de cada condición experimental y modelo. Para calcular la MMI se utiliza $P(S)$ obtenida a través del algoritmo de Arimoto-Blahut en el cálculo de la CC. Todos los algoritmos de resolución y análisis fueron implementados en Python en su versión 2.7.

¹Dado que la temporalidad de los fenómenos de interés es del orden de decenas de minutos, el periodo de muestreo de un minuto es suficiente para capturar las propiedades del dinámicas del sistema.

Modelos Matemáticos

En esta sección se describen en detalle los modelos matemáticos en su representación determinista y estocástica. Incluye tres modelos: modelo de la sección demoduladora de la vía MAP/ERK, modelo de una vía mínima y el modelo de genes reporteros. El análisis de estabilidad de cada modelo se presenta en el apéndice A.

4.1 Modelo de la vía de señalización MAP/ERK

La vía de señalización MAP/ERK transduce estímulos de la superficie de la membrana celular hasta el núcleo con el objetivo de desencadenar una respuesta celular específica, por ejemplo: proliferación o apoptosis [65]. La vía de señalización se divide en dos secciones: la sección moduladora y la demoduladora [20]. La sección moduladora tiene múltiples segmentos de vías que a pesar de compartir moléculas ante diferentes estímulos, son consideradas independientes; la sección demoduladora está acoplada a la salida de la primera. Es el punto de convergencia de la información generada por los diferentes estímulos, Fig. 2.4. Ver sección 2.3 para obtener información más detallada.

4. MODELOS MATEMÁTICOS

En el presente trabajo se estudia la sección demoduladora. Nakakuki et al. (2010) caracterizaron experimental y matemáticamente esta sección. La Figura 4.1 muestra la representación de la estructura interna. El modelo generado está disponible en el repositorio de modelos computacionales para procesos biológicos bajo el número de identificación 1003170000¹. En el trabajo citado describieron la dinámica temporal de las especies moleculares involucradas a través de diez ecuaciones diferenciales, Ecs. (4.1)-(4.10).

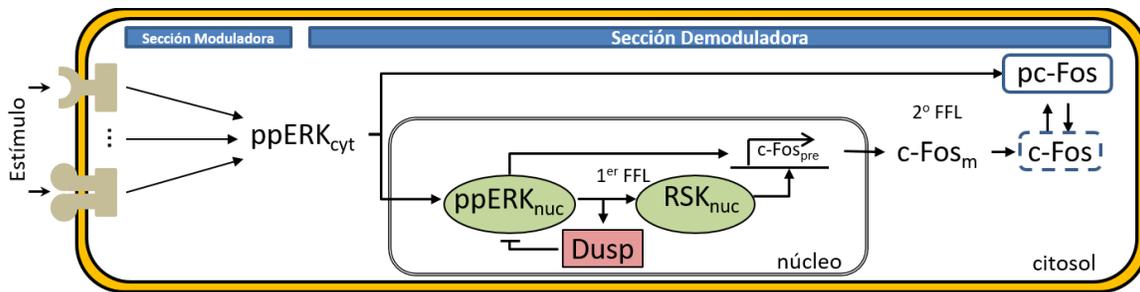


Figura 4.1: Vía MAP/ERK con la descripción explícita de la sección demoduladora.

La entrada de sección decodificadora es la cinasa doblemente fosforilada $ppERK_{cyt}$ ², la cual fue descrita fenomenológicamente ante varias concentraciones de diferentes factores de crecimiento como estímulos (Sección 3.1). Al modelar de esta manera la entrada, ambas secciones de la vía se consideran, ya que la dinámica resultante de $ppERK_{cyt}$ es consecuencia de las interacciones generadas en el bloque modulador. Siguiendo la cadena de reacciones, la cinasa se internaliza al núcleo ($ppERK_{nuc}$) con una tasa k_1 ; se puede desfosforilar e inactivarse por fosfatasas constitutivas y debido a la acción de la cinasa de acción dual $Dusp$ ³ con tasas k_2 y k_3 , respectivamente. La internalización en el núcleo de $ppERK$ activa de manera lineal (k_4) a la cinasa ribosomal RSK_{nuc} , la cual coadyuva al proceso de traducción al volverse una subunidad de la molécula efectora, el ribosoma [95]. RSK_{nuc} es degradada (k_5) por fosfatasas internas. $ppERK_{nuc}$ y RSK_{nuc} activan la transcripción del gen del factor de transcripción $c-Fos$. El transcrito primario ($c-Fos_{pre}$)

¹<https://www.ebi.ac.uk/biomodels-main/BIOMD0000000251>

²Por su siglas en inglés *Extracellular signal-Regulated Kinases*.

³Actúa en los residuos de tirosina y serina/treonina, de ahí su acción dual.

4.1 Modelo de la vía de señalización MAP/ERK

es sintetizado considerando una dinámica tipo Hill (Sección 2.2) con constantes n y k_6 , $c-Fos_{pre}$ es procesado (k_7) a ARNm maduro; a su vez, es degradado a una tasa k_8 . $c-Fos_m$ se traduce en la proteína inestable y defosforilada $c-Fos$ (k_9), la cual es degradada (k_{10}) con una cinética lineal. Para que sea funcional $c-Fos$ necesita ser estabilizada por $ppERK$ mediante fosforilación (k_{11}). Sin embargo el proceso de fosforilación es reversible al estar presente fosfatasas (k_{13}). Finalmente, la proteína $pc-Fos$ es degradada con cinética de primer orden (k_{12}). Topológicamente las variables descritas interactúan mediante dos FFL coherentes del tipo I. El FFL interno determina la transcripción de $c-Fos$, mediante el externo se lleva a cabo la estabilización y fosforilación del factor de transcripción. Todos los parámetros son positivos y su valor se muestran en la tabla 4.1.

$$ppERK_{cyt} = x_1 - x_2 \quad (4.1)$$

$$\frac{dppERK_{nuc}}{dt} = k_1 ppERK_{cyt} - ppERK_{nuc} (k_2 + k_3 Dusp) \quad (4.2)$$

$$\frac{dRSK_{nuc}}{dt} = k_4 ppERK_{nuc} - k_5 RSK_{nuc} \quad (4.3)$$

$$\frac{dc-Fos_{pre}}{dt} = \frac{(ppERK_{nuc} RSK_{nuc})^n}{(ppERK_{nuc} RSK_{nuc})^n + k_6^n} - k_7 c-Fos_{pre} \quad (4.4)$$

$$\frac{dc-Fos_m}{dt} = k_7 c-Fos_{pre} - k_8 c-Fos_m \quad (4.5)$$

$$\frac{dc-Fos}{dt} = k_9 c-Fos_m + k_{13} pc-Fos - c-Fos (k_{10} + k_{11} ppERK_{cyt}) \quad (4.6)$$

$$\frac{dDusp}{dt} = k Dusp \quad (4.7)$$

$$\frac{dpc-Fos}{dt} = k_{11} c-Fos ppERK_{cyt} - pc-Fos (k_{12} + k_{13}) \quad (4.8)$$

$$\frac{dx_1}{dt} = \frac{-x_1 + \kappa_1 L}{\tau_1} \quad (4.9)$$

$$\frac{dx_2}{dt} = \frac{-x_2 + \kappa_2 L}{\tau_2} \quad (4.10)$$

El modelo propuesto por Nakakuki et al. (2010) es un modelo diferencial donde cada variable representa la media de las mediciones en población. A su vez, el trabajo citado no fue realizado con el fin de caracterizar las propiedades de transmisión de la información de

4. MODELOS MATEMÁTICOS

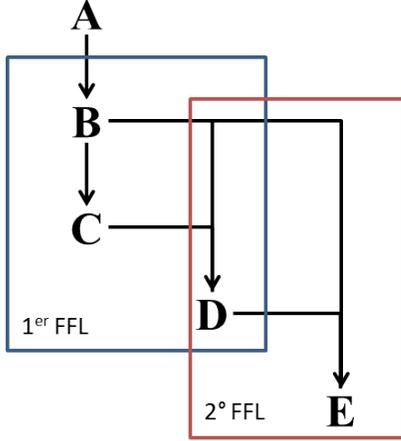
la vía MAP/ERK; el trabajo se centro en establecer el mapa de interacción entre las variables de la sección demoduladora de la vía MAP/ERK. Esto es de particular utilidad para el presente estudio, porque mediante este modelo se tiene conocimiento de las variables que participan y como interactúan entre ellas, permitiendo estudiar el flujo de la información a lo largo de la vía de señalización y no únicamente la relación de entrada-salida. Sin embargo, el modelo presentado se tuvo que reestructurar ya que fue necesario trabajar en el dominio estocástico considerando la aleatoriedad generada en cada sub-proceso. La representación estocástica de la vía de señalización MAP/ERK se presenta en la sección 4.4.

Como resultado del análisis de estabilidad (Apéndice A), el modelo de la vía MAP/ERK es monoestable con un valor de SS igual a cero para todas las variables exceptuando D_{usp} , sección A.1.

4.2 Modelo Mínimo

A través del modelo de una vía de señalización mínima se exploró la relación entre la amplificación y la transmisión de la información en el FFL. Debido a esto, el modelo presentado tiene una ganancia uniforme, todas las variables consideradas tienen el mismo valor de estado estacionario, el cual está determinado directamente por la entrada. El modelo mínimo contiene las principales características de la sección decodificadora de la vía MAP/ERK, compuesto por las interacciones de cuatro proteínas ($B-E$) interconectadas mediante dos FFL anidados, Fig. 4.2. El número de parámetros fue intencionalmente reducido asumiendo que todas las proteínas son las mismas. El modelo matemático que describe la dinámica de las proteínas B , C , D y E es presentado a continuación.

4.3 Modelo de Genes Reporteros



$$\frac{dB}{dt} = kA(t) - gB \quad (4.11)$$

$$\frac{dC}{dt} = k\frac{B}{G_N} - gC \quad (4.12)$$

$$\frac{dD}{dt} = k\frac{BC}{A_N G_N^2} - gD \quad (4.13)$$

$$\frac{dE}{dt} = k\frac{BD}{A_N G_N^2} - gE \quad (4.14)$$

$$G_N = k/g \quad (4.15)$$

Figura 4.2: Modelo mínimo.

La evolución temporal de cada proteína está dada por su respectiva ecuación diferencial. Los parámetros k y g representan las razones de activación o fosforilación y de degradación y defosforilación, respectivamente. El rango de los parámetros es mayor e igual a cero. El modelo mínimo fue evaluado para la misma entrada que el modelo de la vía MAP/ERK (Sección 3.1), donde $A(t) = ppERK_{cyt}(t)$ descrita por la ecuación (4.1). Como se mencionó, la amplificación está determinada por la entrada y todas las proteínas modeladas tienen el mismo valor de SS (Ec. (4.16)) que está determinado por el producto de A_N y G_N (Ecs. (3.2) y (4.15)). Al igual que el modelo anterior, se realizó la representación estocástica del modelo mínimo (sección 4.4).

$$SS = A_N * G_N \quad (4.16)$$

El modelo mínimo es monoestable para cualquier valor positivo de los parámetros, determinado directamente por el valor de g , sección A.2.

4.3 Modelo de Genes Reporteros

Para evaluar y comparar el efecto que tienen ciertas condiciones se utilizan genes reporteros acoplados a distintos niveles de las vías. El proceso de expresión del gen se resumió

4. MODELOS MATEMÁTICOS

en tres grandes subprocesos: **a)** activación del promotor **b)** transcripción y **c)** traducción. (Sección 2.2). Cada uno se modeló como se muestra a continuación.

$$\frac{dP_m}{dt} = \kappa_{pm} \frac{FT^n}{FT^n + K_{rp}^n} (1 - P_m) - \gamma_{pm} P_m \quad (4.17)$$

$$\frac{dm}{dt} = \kappa_m P_m - \gamma_m m \quad (4.18)$$

$$\frac{dP}{dt} = \kappa_p m - \gamma_p P \quad (4.19)$$

donde P_m , m y P representan el promotor en estado activo, la concentración de ARNm y la concentración de proteínas, respectivamente. La evolución temporal de cada especie bioquímica está representada por su derivada. La dinámica del promotor (Eq. (4.17)) se modeló con una ecuación tipo Hill, n es la constante de Hill y el inverso de K_{rp} es la afinidad del gen por su factor de transcripción. Cada proceso tiene una tasa particular de activación/producción (κ) y una de desactivación/degradación (γ).

Del resultado del análisis de estabilidad, al igual que los modelos anteriores, el modelo de los genes acoplados es monoestable en todo el espacio paramétrico siempre y cuando sean positivos, sección A.3.

4.4 Modelos estocásticos

Los modelos deterministas presentados en las secciones anteriores, representan la media estadística en una población. Sin embargo, estas descripciones descartan las fluctuaciones generadas a partir de las componentes de ruido en los sistemas modelados. Se ha demostrado que el ruido determina de manera fundamental la respuesta ya sea en las vías de señalización [58] o en la expresión de genes [52, 96]. Aunado a esto, los modelos ocupados representan canales por los cuales se procesa información, y para ser caracterizados como tales, se debe estudiar la repercusión del ruido. Por lo que fue necesario reestructurar las representaciones deterministas en estocásticas al considerar que las variables son aleatorias.

4.4 Modelos estocásticos

En la sección 2.5.1 se describió el formalismo teórico para describir el cambio de la distribución de probabilidad en función del tiempo para una variable aleatoria, mediante la ecuación química maestra. Para resolver la ecuación química maestra de cada modelo se ocupó el algoritmo de Gillespie, lo cual implicó definir el sistema como un conjunto de reacciones químicas y establecer sus propensiones, tabla 4.1. Los valores de los parámetros del modelo de la vía de señalización MAP/ERK fueron tomados del trabajo de Nakakuki et al. [1]. Para el modelo mínimo y los genes acoplados se utilizaron parámetros que representaran los rangos biológicos sin ser particulares de algún sistema.

Para establecer las propensiones del modelo de la vía de señalización MAP/ERK, se ocupó la caracterización determinista presentada (sección 4.1). Es importante señalar que ciertas propensiones se modelaron sin considerarse los pasos intermedios. Específicamente, la propensión para la producción del transcrito primario $c-Fos_{pre}$ la cual engloba de manera condensada las siguientes reacciones: **a)** unión de $ppERK_{nuc}$ y $pRSK_{nuc}$ al promotor del gen $c-fos$ y **b)** transcripción del gen. Dentro de las diferencias entre las reacciones, existe diferencia del tiempo de escala. De manera global el proceso está determinado y limitado por la reacción más lenta, que es la transcripción de $c-fos$. Por lo que se asume que la que primera se encuentra en equilibrio químico, esto es posible dado que $ppERK_{nuc}$ y $pRSK_{nuc}$ se presentan en conteo moleculares del orden de cientos de miles de moléculas [1] y tienen una velocidad de reacción alta.

Otra propensión que conjunta diversas reacciones es la activación del promotor del gen acoplado. Tal reacción está determinada por la presencia del factor de transcripción $pc-Fos$. El cual se encuentra en una concentración de 113 nM [97], siendo una concentración alta. A su vez, $pc-Fos$ tiene un tiempo de vida medio de 4 horas [1], por lo que su concentración cambia lentamente.

La Ec. (4.20) ejemplifica la ecuación obtenida para el modelo mínimo, de la misma manera se obtuvo la respectiva ecuación para el modelo de la vía MAP/ERK y sistema de expresión. A través de las representaciones determinista y estocástica se estudian dife-

4. MODELOS MATEMÁTICOS

rentes propiedades de los sistemas por lo que el análisis es complementario para ambas representaciones.

$$\begin{aligned}
 \frac{dP(n_B, n_C, n_D, n_E, t)}{dt} = & [P(n_B - 1, n_C, n_D, n_E, t) - P(n_B, n_C, n_D, n_E, t)] \frac{k}{A_N} n_A \\
 & + [P(n_B + 1, n_C, n_D, n_E, t)(n_B + 1) - P(n_B, n_C, n_D, n_E, t)n_B]g \\
 & + [P(n_B, n_C - 1, n_D, n_E, t) - P(n_B, n_C, n_D, n_E, t)]g n_B \\
 & + [P(n_B, n_C + 1, n_D, n_E, t)(n_C + 1) - P(n_B, n_C, n_D, n_E, t)n_C]g \\
 & + [P(n_B, n_C, n_D - 1, n_E, t) - P(n_B, n_C, n_D, n_E, t)] \frac{k}{G^2} n_B n_C \\
 & + [P(n_B, n_C, n_D + 1, n_E, t)(n_D + 1) - P(n_B, n_C, n_D, n_E, t)n_D]g \\
 & + [P(n_B, n_C, n_D, n_E - 1, t) - P(n_B, n_C, n_D, n_E, t)] \frac{k}{G^2} n_B n_D \\
 & + [P(n_B, n_C, n_D, n_E + 1, t)(n_E + 1) - P(n_B, n_C, n_D, n_E, t)n_E]g
 \end{aligned} \tag{4.20}$$

4.5 Validación

Para validar el modelo estocástico, los resultados numéricos descritos por Nakakuki et al. (2010) fueron replicados. Como ya se mencionó, ellos caracterizan las vía MAP/ERK de forma experimental y desarrollan el modelo matemático de la vía que reproduce esos resultados. El modelo estocástico se describió utilizando las propensiones de la Tabla 4.1 y fue resuelto numéricamente a través del algoritmo de Gillespie; se hicieron 1000 realizaciones. Para el caso determinista se resolvió el sistema de ecuaciones que describen a la vía, Ecs. (4.1)- (4.10). Se despreció la componente de ruido y se resolvió el sistema utilizando el método de Runge-Kutta de quinto orden. Para ambos modelos, el tiempo de simulación fue de 90 minutos con una frecuencia de muestreo de un minuto y condiciones iniciales para todas las moléculas iguales a cero.

La figura 4.3 muestra los resultados de la solución de ambos modelos. Se presentan la dinámica temporal de tres variables: la entrada de la vía ($ppERK_{cyt}$), la salida del primer FFL ($c-Fos_m$) y la última molécula a la salida del segundo FFL ($pc-Fos$), Figura 4.1. Se

4.5 Validación

Tabla 4.1: Reacciones, propensiones, y valor de los parámetros de los modelos estocásticos.

	Reacción	Propensión	Valor	Descripción
Vía de señalización MAP/ERK	$ppERK_{cyt} \xrightarrow{k_1} ppERK_{nuc}$	$k_1 ppERK_{cyt}$	15.0 min^{-1}	Transporte de molécula al núcleo
	$ppERK_{nuc} \xrightarrow{k_2} DUSP$	$k_2 ppERK_{nuc}$	1.0 min^{-1}	Producción de DUSP
	$ppERK_{nuc} \xrightarrow{k_3} \emptyset$	$k_3 ppERK_{nuc}$	50.0 min^{-1}	Defosforilación de la ERK doblemente fosforilada por fosfatasa endógena
	$ppERK_{nuc} \xrightarrow{k_4 DUSP} \emptyset$	$k_4 DUSP ppERK_{nuc}$	14.0 min^{-1}	Defosforilación mediada por DUSP
	$ppERK_{nuc} \xrightarrow{k_5} pRSK_{nuc}$	$k_5 ppERK_{nuc}$	0.1 min^{-1}	Activación de RSK en el núcleo
	$pRSK_{nuc} \xrightarrow{k_6} \emptyset$	$k_6 pRSK_{nuc}$	0.15 min^{-1}	Dilución e inactivación de RSK en el núcleo
	$ppERK_{nuc} + pRSK_{nuc} \xrightarrow{n, k_6} c\text{-Fos}_{pre}$	$\frac{(ppERK_{nuc} pRSK_{nuc})^n}{k_6^n + (ppERK_{nuc} pRSK_{nuc})^n}$	$n = 1.1$ $k_6 = 0.13$	Producción del transcrito primario $c\text{-Fos}_{pre}$
	$c\text{-Fos}_{pre} \xrightarrow{k_7} c\text{-Fos}_m$	$k_7 c\text{-Fos}_{pre}$	0.5 min^{-1}	Maduración del transcrito primario
	$c\text{-Fos}_m \xrightarrow{k_8} \emptyset$	$k_8 c\text{-Fos}_m$	0.08 min^{-1}	Degradación del ARN_m
	$c\text{-Fos}_m \xrightarrow{k_9} c\text{-Fos}$	$k_9 c\text{-Fos}_m$	0.3 min^{-1}	Transcripción de la proteína inestable y defosforilada $c\text{-Fos}$
	$c\text{-Fos} \xrightarrow{k_{10}} \emptyset$	$k_{10} c\text{-Fos}$	0.3 min^{-1}	Degradación de la proteína $c\text{-Fos}$
	$ppERK_{cyt} + c\text{-Fos} \xrightarrow{k_{11}} pc\text{-Fos}$	$k_{11} ppERK_{cyt} c\text{-Fos}$	$0.11 N^{-1} \text{ min}^{-1}$	Activación y estabilización de la proteína $c\text{-Fos}$
	$pc\text{-Fos} \xrightarrow{k_{12}} \emptyset$	$k_{12} pc\text{-Fos}$	0.001 min^{-1}	Degradación de la proteína fosforilada $c\text{-Fos}$
	$pc\text{-Fos} \xrightarrow{k_{13}} c\text{-Fos}$	$k_{13} pc\text{-Fos}$	0.6 min^{-1}	Inactivación de la proteína fosforilada $c\text{-Fos}$
Modelo mínimo	$A \xrightarrow{k} B + A$	kA	20.0 min^{-1}	Activación de la proteína B
	$B \xrightarrow{g} \emptyset$	gB	0.1 min^{-1}	Dilución y degradación de la proteína B
	$B \xrightarrow{k} C + B$	kB	20.0 min^{-1}	Activación de la proteína C
	$C \xrightarrow{g} \emptyset$	gC	0.1 min^{-1}	Dilución y degradación de la proteína C
	$B + C \xrightarrow{k} D + B + C$	$k \frac{BC}{A_N G_N^2}$	20.0 min^{-1}	Activación de la proteína D por las proteínas B y C
	$D \xrightarrow{g} \emptyset$	gD	0.1 min^{-1}	Dilución y degradación de la proteína D
	$C + D \xrightarrow{k} E + C + D$	$k \frac{CD}{A_N G_N^2}$	20.0 min^{-1}	Activación de la proteína E por las proteínas C y D
	$E \xrightarrow{g} \emptyset$	gE	0.1 min^{-1}	Dilución y degradación de la proteína E
Gen acoplado	$pc\text{-Fos} + P_m^* \xrightarrow{\kappa_{P_m}, n, K_{rp}} P_m + pc\text{-Fos}$	$\kappa_{P_m} \frac{pc\text{-Fos}^{n_g} P_m^*}{pc\text{-Fos}^{n_g} + K_{rp} P_m^*}$	$\kappa_{P_m} = 0.1 \text{ min}^{-1}$ $n_g = 2.0$ $K_{rp} = 100.0 \text{ min}^{-1}$	Activación del promotor
	$P_m \xrightarrow{\gamma_{P_m}} P_m^*$	$\gamma_{P_m} P_m$	0.2 min^{-1}	Inactivación del promotor
	$P_m \xrightarrow{\kappa_m} m + P_m$	$\kappa_m P_m$	14.0 min^{-1}	Transcripción del gen reportero
	$m \xrightarrow{\gamma_m} \emptyset$	$\gamma_m m$	1.0 min^{-1}	Degradación del ARN_m
	$m \xrightarrow{\kappa_p} P + m$	$\kappa_p m$	20.0 min^{-1}	Traducción y maduración del gen reportero
	$P \xrightarrow{\gamma_P} \emptyset$	$\gamma_P P$	0.1 min^{-1}	Pérdida del conteo de proteína debido a la dilución y degradación

4. MODELOS MATEMÁTICOS

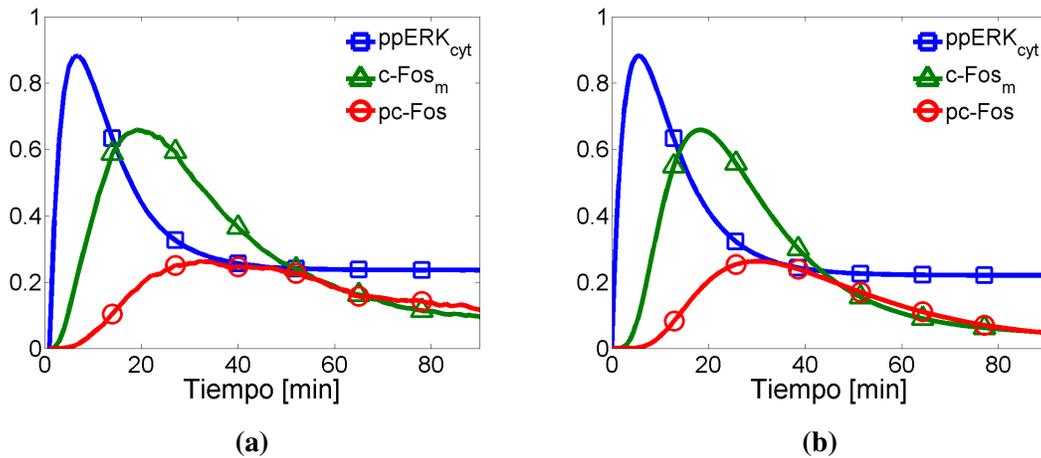


Figura 4.3: Convergencia entre la representaciones estocásticas y deterministas. Se muestran tres variables de la vía MAP/ERK: $ppERK_{cyt}$, $c-Fos_m$ y $pc-Fos$. (a) Modelo estocástico (b) Modelo determinista.

presenta la media de las mil realizaciones estocásticas, Fig. 4.3(a). La cual contrasta en su convergencia con la solución del modelo determinista, Fig. 4.3(b). Los parámetros para la entrada ($ppERK_{cyt}$) son los descritos para I_3 en la tabla 3.1, que corresponde al estímulo de 10nM de factor de crecimiento epidermal (EGF). Numéricamente se comprueba la convergencia entre la media poblacional de la ecuación química maestra y el modelo determinista. Dado que el modelo estocástico reproduce los resultados deterministas y consecuentemente los resultados experimentales es confiable usar la representación para caracterizar el proceso de la transmisión de la información en esta vía.

La validación presentada, se basa en la convergencia entre la media estadística del modelo estocástico con la dinámicas generadas a partir del modelo determinista. Sin embargo, no se compara el nivel de ruido presente en el sistema. Dado que el modelo base que se ocupó para establecer la representación estocástica fue un estudio teórico-experimental basado en registros de fluorescencia, las dinámicas registradas son cualitativas, por otro lado no existen datos experimentales que establezcan las concentraciones ni la resolución de

4.5 Validación

respuesta de todas las variables modeladas. Debido a esto, se utilizó un factor de escala para simular diferentes condiciones. Se sabe la concentración del factor de transcripción *pC-Fos*, que es la salida funcional de la vía MAP/ERK. En ciertas condiciones se alcanzaban los conteos moleculares reportados sin embargo en otras condiciones el sistema se expresaba por debajo de esos límites. Esto impactó al tener una posible sobre estimación del ruido intrínseco. Sin embargo, los resultados se expanden utilizando el modelo mínimo en el cual se trata de controlar la sobre estimación.

Resultados & Discusión

En este capítulo se presentan y discuten los resultados obtenidos al estudiar el proceso de transmisión de información en los mecanismos demoduladores. Para evaluar la cantidad de información se acoplaron genes reporteros en distintas secciones de la vía de señalización; además, es incluido el efecto del ruido extrínseco en la transmisión de la información. El capítulo está dividido en tres bloques: uno para cada mecanismo demodulador y el tercero considerando el ruido extrínseco. Los modelos ocupados para los distintos experimentos son descritos en el capítulo anterior. En cada experimento se evalúa el efecto de la variabilidad al momento de procesar la información y su relevancia. Los conceptos teóricos son retomados y relacionados con los resultados obtenidos con el fin de expandir su interpretación y poder entender las propiedades de los mecanismos demoduladores.

5.1 Caracterización del FFL como demodulador

Las vías de señalización captan y transducen los estímulos del exterior al interior de la célula. Se considera que estas señales son información bioquímica que se propaga [7]. La información es percibida, modulada, transmitida y demodulada con el objetivo de desencadenar una acción celular específica. Un mecanismo demodulador es la utilización de

5. RESULTADOS & DISCUSIÓN

topologías tipo FFL. Sin embargo no se ha establecido las características y propiedades del FFL en la transmisión de la información, ver sección 2.3.2.

5.1.1 Vía MAP/ERK

Para iniciar la caracterización se ocupó el modelo estocástico de la vía de señalización MAP/ERK, la cual contiene una sección demoduladora (SD) compuesta por dos FFL anidados [1, 20], Fig. 4.1. Para caracterizar las propiedades de la SD para transmitir información se acoplaron dos genes reporteros idénticos (Sección 4.3). El primero interactuando completamente con la vía (Prt_w) siendo activado por el factor de transcripción $pc-Fos$. El segundo gen (Prt_{wo}) es activado sin interactuar con la SD, Fig. 5.1(a)¹. La correspondiente descripción estocástica para el gen acoplado está dado en la tabla 4.1, y la determinista en las ecs. (4.17)-(4.19).

Para evaluar la participación de la SD compuesta por dos FFL anidados se obtuvo la cantidad de información que se trasmite de $ppERK_{cyt}$ a cada gen. Esto se hizo cuantificando la capacidad del canal, la cual determina la máxima cantidad de información transmitible por un canal (Sección 2.4.3). La figura 5.1(b) muestra la CC en función del tiempo para ambos genes. Se observa que el gen Prt_w presenta un valor mayor de CC comparado con el gen Prt_{wo} . Este resultado resalta la presencia del FFL en el proceso de transmisión de información; sin embargo, es un resultado contraintuitivo. En largas cadenas de interacción de especies moleculares una cierta cantidad de ruido es adicionada en cada paso [51, 53] lo que conlleva a que la información se pierda [98]. A su vez, en cada paso la información sufre una nueva representación o procesamiento al verse involucrada una nueva variable. Según la desigualdad de procesamiento de datos (sección 2.4.5), la información debe de ir disminuyendo. Por otro lado, se ha demostrado que el FFL promueve la pérdida de información, esto debido a la manera en que está estructurado [5], siendo un cuello de botella. Sin embargo, los resultados indican que la presencia del FFL incrementa la cantidad de información. Adicionalmente, la CC para ambos genes presenta una dinámica temporal

¹El segundo gen es activado por $ppERK_{nuc}$ siendo la cinasa que se internaliza al núcleo celular, lugar donde se inicia el proceso de expresión del gen.

5.1 Caracterización del FFL como demodulador

descrita por tres fases: de crecimiento, un valor máximo y de decaimiento. Esto tiene sentido ya que el gen no está activado en todo momento, el sistema tiende a desactivarse con el fin de poder reaccionar ante otro estímulo.

Con el fin de tener un mejor entendimiento del flujo de la cantidad de información a lo largo de la vía, se caracterizó la máxima cantidad de información en cada etapa. La figura 5.1(c) muestra la CC en función del tiempo para tres variables: $ppERK_{nuc}$, $c-Fos$ y $pc-Fos$. Las cuales se localizan al inicio, y a las salidas del primero y segundo FFL en la sección demoduladora de la vía MAP/ERK, Fig 4.1. En las figuras 5.1(b) y 5.1(c) se muestra que la CC incrementa con respecto a la cadena de interacción avanza, $pc-Fos$ tiene más información que su molécula antecesora; incluso Prt_w posee más información que su factor de transcripción $pc-Fos$.

La ganancia de información en las últimas moléculas, implica la violación de la desigualdad de procesamiento de datos¹. De acuerdo con dicha desigualdad, las vías de señalización con un menor número de interacciones intermedias son mejores, ya que la pérdida de información es menor. No obstante, se ha demostrado que la CC puede ser incrementada a través de la amplificación y adición de ruido. Con esto en mente, se evaluó el conteo molecular promedio en función del tiempo para ambos genes, Fig. 5.1(d). La expresión del gen Prt_w es un orden de magnitud mayor que Prt_{wo} . Comparando las trazos de la CC y el conteo molecular promedio (Figs. 5.1(b) y 5.1(d)) se observa que el gen con mayor expresión es el que tiene mayor CC, por lo tanto mayor información.

La figura 5.2 muestra la capacidad del canal para las variables de la vía MAP/ERK ante tres entradas (I_1 , I_2 y I_3). Se aprecia que los resultados son equivalentes para todas las variables y para diferentes entradas. La CC incrementa para las últimas variables. Analizando el conteo molecular promedio de la SD (Fig. 5.3), las variables que tienen un conteo mayor tienen un valor de CC mayor. Estos resultados no dependen de la entrada, dependen

¹Recordando, la desigualdad establece que la información debe ser perdida progresivamente en cada paso de la vía [84, 99]. Sección 2.4.5

5. RESULTADOS & DISCUSIÓN

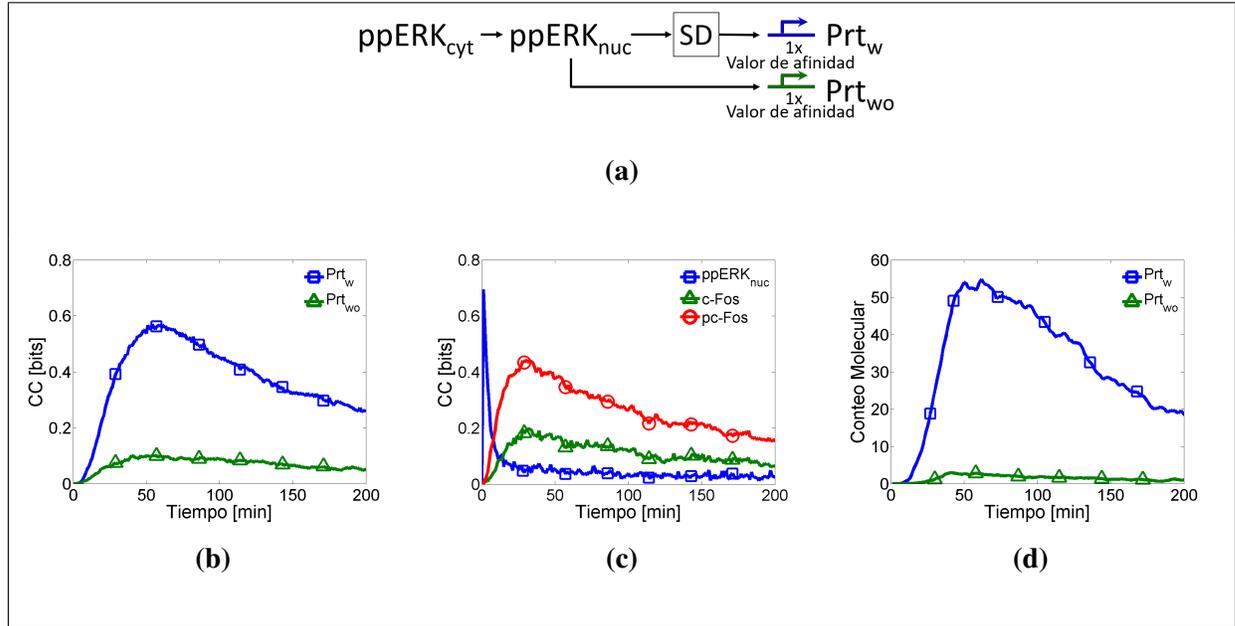


Figura 5.1: Exploración de las propiedades del FFL usando la vía MAP/ERK. (a) Genes acoplados a distintos niveles de la sección decodificadora (SD) de la vía MAP/ERK interactuando (Prt_w) y sin interactuar (Prt_{wo}) con la vía. (b) Capacidad del canal (CC) en función del tiempo de los genes acoplados. (c) CC de moléculas dentro de la vía MAP/ERK: al principio ($ppERK_{nuc}$) a la salida del primer FFL ($c-Fos$) y a la salida del segundo FFL ($pc-Fos$). (d) Trazo temporal del conteo molecular promedio de los genes acoplados para un valor de $L = 100$.

directamente de la estructura de la vía de señalización.

De manera interesante, el FFL no únicamente incrementa la máxima cantidad de información, también induce un incremento en la expresión o activación de las moléculas. Como resultado del análisis anterior, se derivaron las siguientes preguntas: ¿Porqué la CC aumenta en lugar de disminuir en las últimas moléculas? ¿Este incremento es debido a la presencia del FFL o a la ganancia y amplificación? Usando el modelo de la vía MAP/ERK no es posible diferenciar entre el efecto de la amplificación y la presencia del FFL, ya que cualquier modificación a las propiedades de la vía modifica la ganancia.

5.1 Caracterización del FFL como demodulador

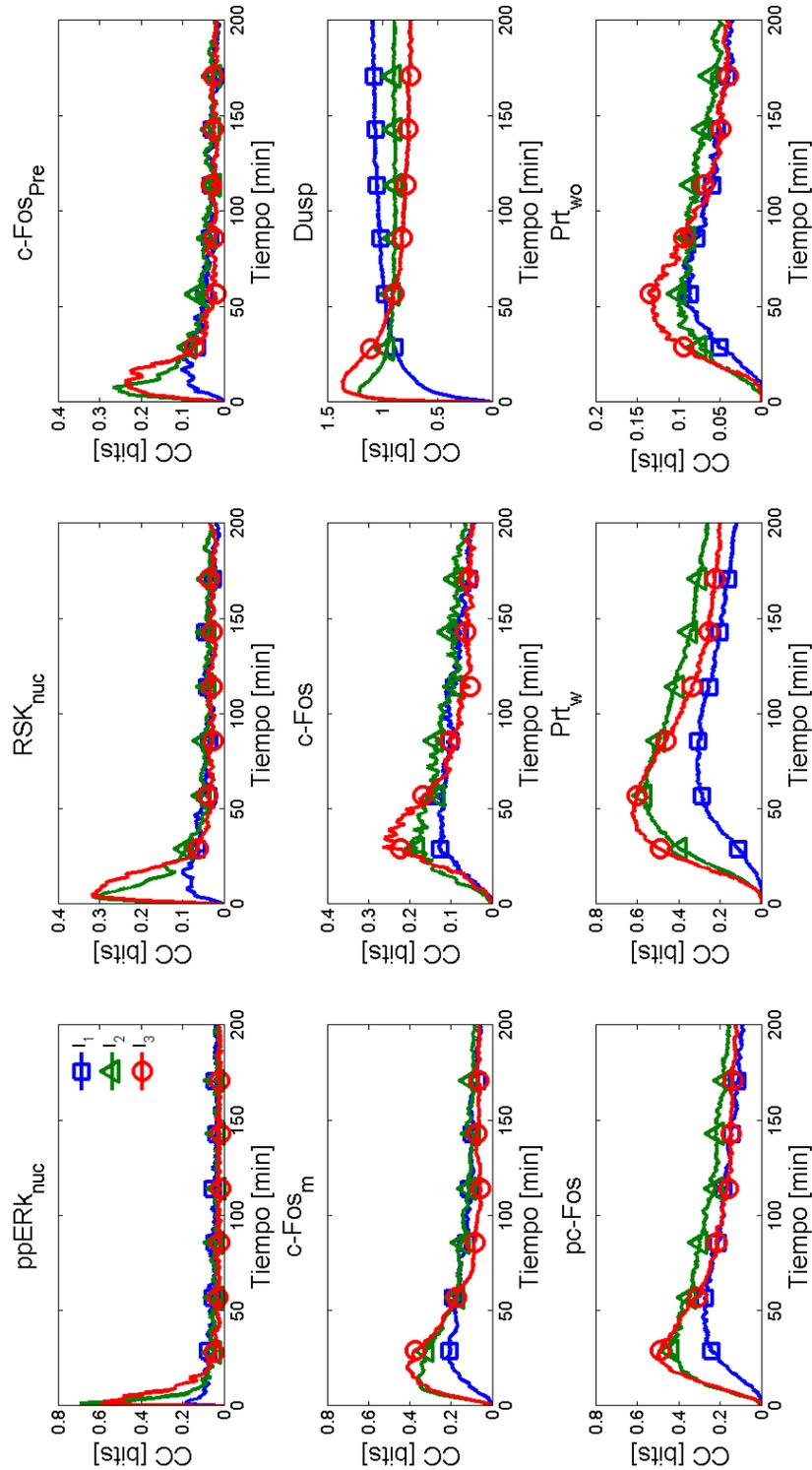


Figura 5.2: Capacidad del Canal para todas las variables de la sección decodificadora de la vía MAPK/ERK para tres entradas I_1, I_2 y I_3 .

5. RESULTADOS & DISCUSIÓN

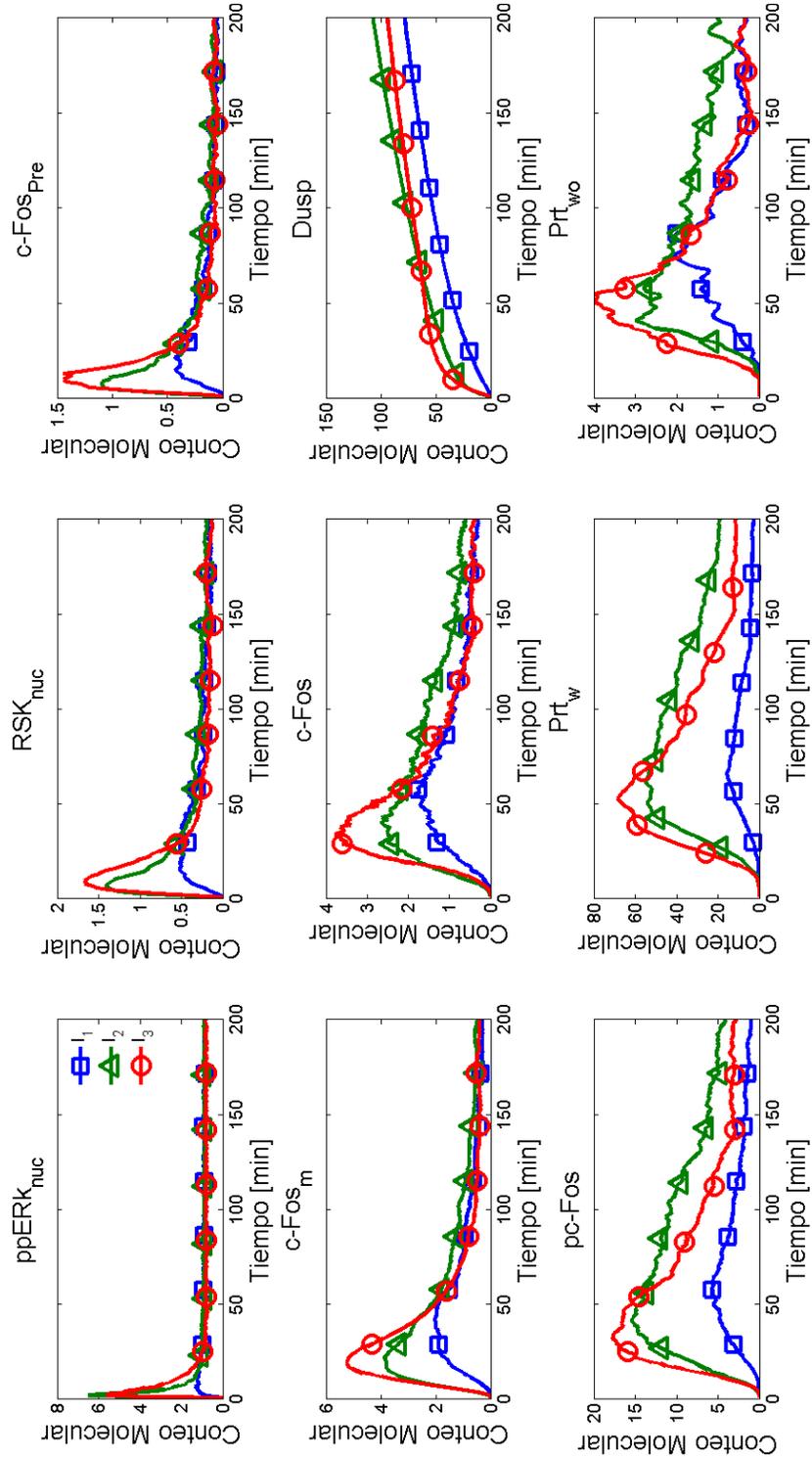


Figura 5.3: Conteo molecular promedio para las variables de la sección decodificadora de la vía MAP/ERK estimulada con las entradas I_1 , I_2 y I_3 con $L = 100$.

5.1 Caracterización del FFL como demodulador

5.1.2 Vía mínima

Para estudiar las implicaciones de la amplificación y la presencia del FFL, se ocupó el modelo de una vía genérica mínima que incluya las principales características de la sección demoduladora de la vía MAP/ERK, Sección 4.2. El modelo está construido por la interacción de cuatro proteínas (B, C, D y E) interconectadas por dos FFL anidados, Fig. 5.4(a), donde $A = ppERK_{cyt}$. La descripción estocástica para cada gen está dada en la tabla 4.1 y la determinista por las ecs. (4.11)-(4.14). La figura 5.4(b) muestra la dinámica temporal promedio de las proteínas. El modelo tiene la característica de tener el mismo estado estacionario para $B-E$ lo cual implica que la amplificación a lo largo de la vía es la misma.

Evaluando la cantidad de información que se transmite en cada paso de la vía, se obtuvo la CC en función del tiempo para cada variable, Fig 5.4(c). La CC para todas las variables alcanza el mismo valor en el equilibrio ($CC(\infty)$). Consecuentemente, la cantidad de información no cambia a lo largo de la vía mínima. Al igual que los resultados en la vía MAP/ERK, hay una contradicción con la desigualdad de procesamiento de datos para sistemas ruidosos, ya que la información debe disminuir y en este caso se conserva. Para valorar el grado de variabilidad presente que pudiera explicar que la información no se degrade, se calculó el coeficiente de variación en cada especie en función del parámetro de escala L a un tiempo igual a 150 min, Fig 5.4(d). El CV incrementa con respecto la cadena de reacciones avanza ($B \rightarrow E$). Dado que la CC se conserva y la variabilidad incrementa, da pie a postular que la adición de variabilidad compensa la pérdida natural de información debida a la propagación y activación de moléculas subsecuentes. Los resultados anteriores fueron obtenidos describiendo la entrada $ppERK_{cyt}$ con los parámetros para I_2 . Para evaluar el efecto de la amplificación, se repitió el análisis con las entradas I_1 e I_3 .

Los resultados son cualitativamente similares para las otras entradas; la $CC(\infty)$ no es afectada a lo largo de la vía, Figs. 5.4(e) y 5.4(f). Las diferentes entradas, poseen perfiles dinámicos distintos, dentro de esas características cada una tiene un valor de SS diferente, Fig. 3.2. Se aprecia que I_2 tiene un valor de SS mayor que las entradas restantes. Bajo la presunción de que la amplificación incrementa el valor de CC se esperaría que

5. RESULTADOS & DISCUSIÓN

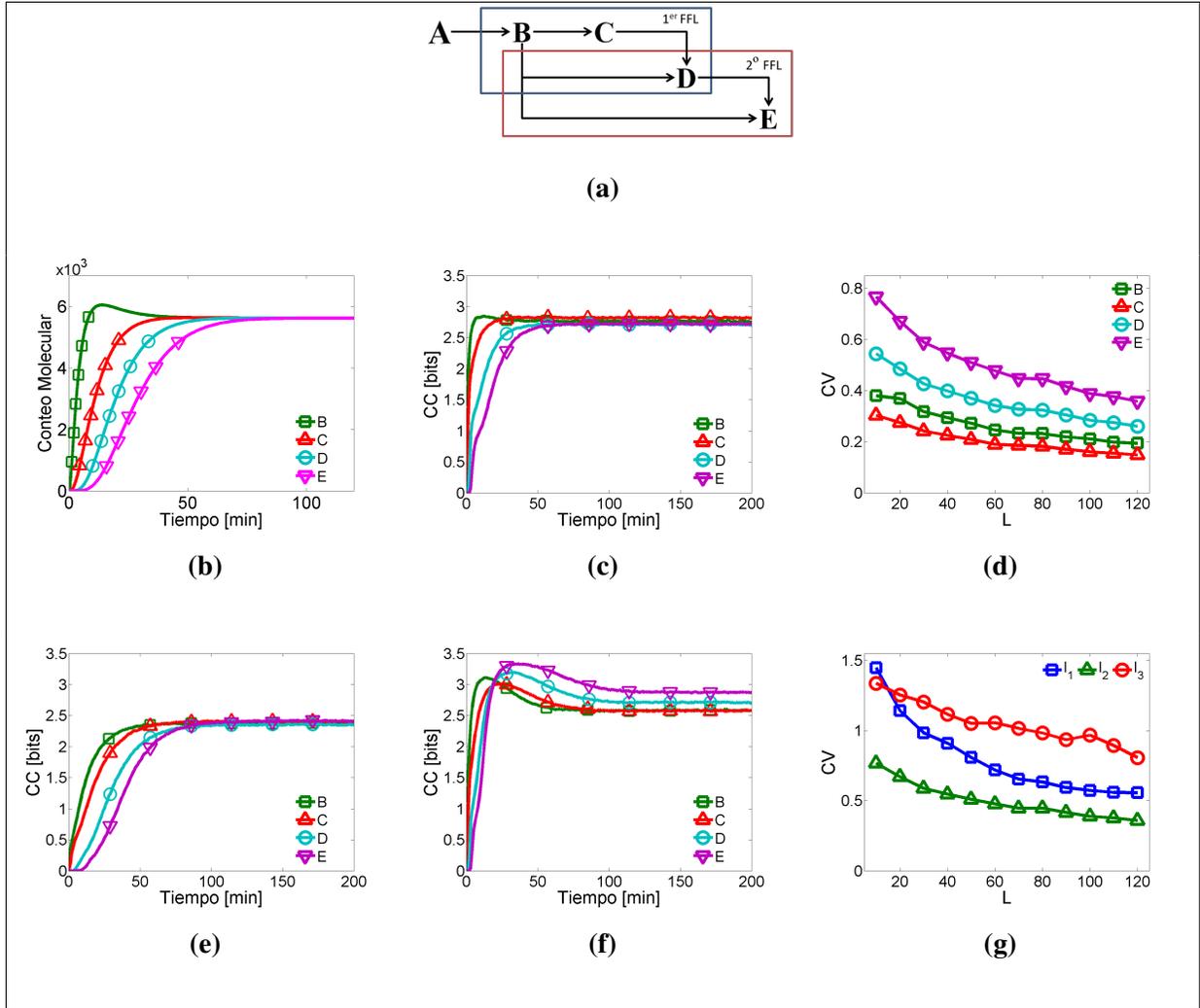


Figura 5.4: Exploración de las propiedades del FFL usando la vía mínima. (a) Modelo de la vía mínima, compuesta por cuatro proteínas ($B - E$) interaccionando en dos FFL anidados. La entrada A es igual a $ppERK_{cyt}$ con los parámetros correspondiente a I_2 . (b) Dinámica promedio temporal de las proteínas para $L = 100$. (c) Evolución temporal de la CC para todas las proteínas, para la entrada I_2 . (d) Coeficiente de variación (CV) en función del parámetro de escala L para $t = 150$ min para las proteínas $B - E$. (e) - (f) Capacidad del canal en función del tiempo para cada proteína del modelo mínimo para las entradas I_1 e I_3 . (g) Coeficiente de variación para la proteína E en función del parámetro de escala L a $t = 150$ min.

5.1 Caracterización del FFL como demodulador

$CC_{I_2}(\infty) > CC_{I_3}(\infty) > CC_{I_1}(\infty)$ dado que $SS_{I_2} > SS_{I_3} > SS_{I_1}$. Comparando las figuras 5.4(c), 5.4(e) y 5.4(f) se observa que la capacidad del canal para I_3 es mayor respecto a los demás entradas, resultado contrario a lo esperado. La figura 5.4(g) muestra el CV de la proteína E en función del parámetro de escala a un tiempo de 150 min para las tres entradas. De manera interesante, calculando la variabilidad inducida por cada entrada a la vía se encontró que $CV_{I_3} > CV_{I_2} > CV_{I_1}$. De manera general, I_3 induce mayor variabilidad. Por lo tanto, la inducción de variabilidad intrínseca ayuda o mejora las capacidades de transmitir la información.

Cuando la información se transmite en condiciones sin variabilidad, las distribuciones de probabilidad de la respuesta siguen la función delta de Dirac, Fig. 3.3(a). La variabilidad induce que la respuesta tenga un grado de incertidumbre provocando que las distribuciones se ensanchen, como resultado de esto puede existir un empalme entre las distribuciones de las respuestas, provocando una reducción en la capacidad de discernir las respuestas de manera separada. Sin embargo; observando el fenómeno desde otro punto de vista, el incremento en el ancho de la distribución incrementa el rango dinámico, por lo que existen mayores estados accesibles para el sistema, lo cual puede conllevar a la capacidad de discernir entre diferentes respuestas implicando un aumento de la CC [50, 100–103]. Por lo tanto, la capacidad de identificar está ligada al rango dinámico que está directamente relacionado al perfil de la distribución de probabilidad de la respuesta [104, 105], y el perfil está determinado por la variabilidad inducida. Lo cual implica que la adición de ruido intrínseco es benéfico en el proceso de transmisión de la información.

Sin embargo, ha sido reportado que la variabilidad disminuye la calidad en la transmisión de información [50, 106, 107]. De manera contrastante, teórica [99] y experimentalmente [50] se ha reportado que el proceso de filtrar la señal para evitar la propagación de ruido, provoca una mayor pérdida de la información limitando la CC. Por lo tanto, los resultados obtenidos sugieren que la célula a través de la evolución ha desarrollado los mecanismos necesarios para transmitir la información utilizando la variabilidad como un potenciador para mejorar este proceso.

5. RESULTADOS & DISCUSIÓN

De acuerdo con Tikhonov et al.(2015), la cantidad total de información no es relevante; lo que realmente importa es la cantidad de información a la cual las moléculas acceden. Ellos, establecieron teórica y experimentalmente la diferencia entre estos dos tipos de información; la total disminuye en cada paso como lo dictamina la desigualdad de procesamiento de datos. Sin embargo, la información accesible está contenida dentro de la información total pero esta si puede incrementar en cada paso pero nunca es mayor que la total. La forma en que puede darse el incremento es a través de procesos de amplificación ruidosa y una correcta sintonización de los parámetros. La amplificación ruidosa como su nombre lo indica, consiste en darle una ganancia a la respuesta de salida e inducir variabilidad. Esto es congruente con los resultados obtenidos, cuando se calcula la CC se está obteniendo la cantidad de información a la que puede acceder la variable y no toda la que está disponible. Los resultados presentados muestran que hay dos correlaciones positivas con respecto al incremento en la CC, la primera con respecto a la amplificación y la segunda con respecto a la inducción de variabilidad, que están directamente relacionadas con la presencia del FFL. Por lo tanto, se concluye que el FFL como mecanismo demodulador funciona como un amplificador ruidoso, propiedad que no se le había atribuido. A su vez, si lo que importa es la amplificación y adición de ruido el FFL puede ser remplazado por otro tipo de topologías.

Para evaluar la participación del FFL en la amplificación y adición de ruido, se realizó un estudio comparativo. EL doble FFL anidado se eliminó resultando una cadena de interacción lineal de cuatro proteínas y se obtuvo la CC, Fig. 5.5(a). Para llevar a cabo la comparación se tuvieron que reestructurar las ecs. (4.13) y (4.14) como las ecuaciones que a continuación se muestran:

$$\frac{dD}{dt} = k \frac{C}{G_N} - gD, \quad (5.1)$$

$$\frac{dE}{dt} = k \frac{D}{G_N} - gE. \quad (5.2)$$

Es importante señalar que el modelo de interacción lineal y el modelo con el doble FFL anidado tienen los mismo parámetros y el mismo valor de estado estacionario, Fig. 5.5(b). La figura 5.5(c) muestra la CC para todas las proteínas del modelo lineal siendo estimulado

5.1 Caracterización del FFL como demodulador

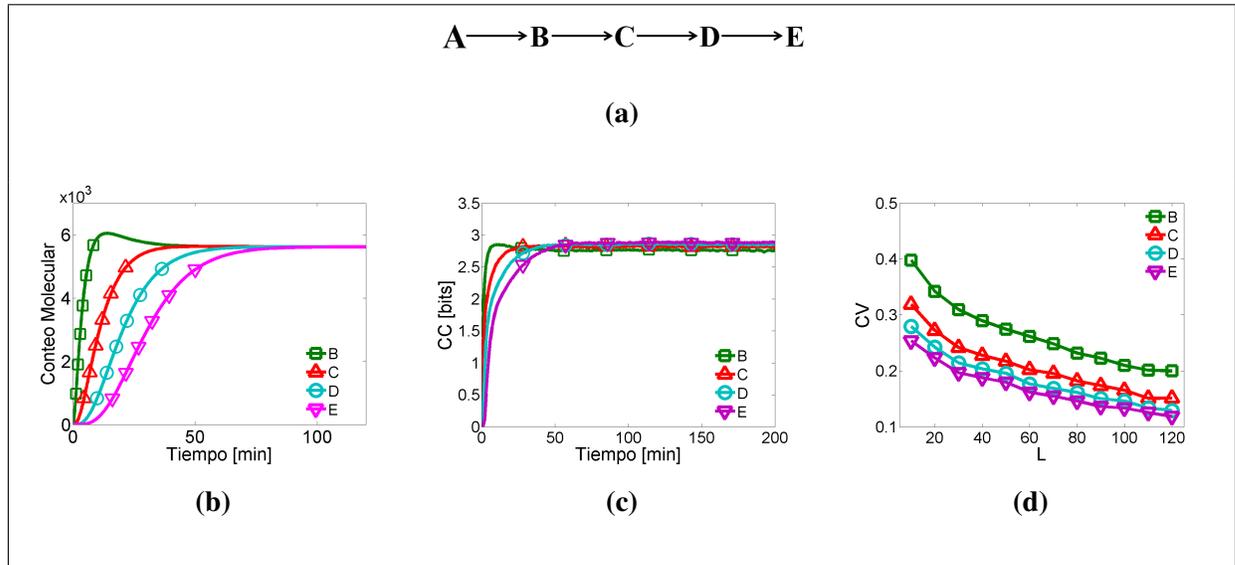


Figura 5.5: Evaluación de la propiedades para transmitir información de una cadena de interacción lineal. (a) Modelo mínimo de una cadena de interacción lineal, formado de cuatro proteínas ($B - E$) estimulado por $A = ppERK_{cyt}$. (b) Trazo temporal del conteo molecular promedio, la entrada está descrita por I_2 . (c) CC para el modelo de interacción lineal. (d) CV en función del parámetro de escala L para $t = 150$ min para las proteínas $B - E$.

por I_2 , comparando la dinámica temporal con el resultado obtenido con el modelo mínimo, Fig. 5.4(c). Se observa que el valor de CC es el mismo en estado estacionario para todas las proteínas ($B - E$). Adicionalmente, es el mismo valor para ambos modelos. Sin embargo, cuando se calcula el coeficiente de variación se observa que la variabilidad disminuye con respecto la cadena de interacción avanza, Fig. 5.5(d). De resultados anteriores, se determinó que la adición de variabilidad es fundamental para la extracción de información, por lo que la disminución de la variabilidad pudiera ser una situación no deseable.

Como ya se mencionó en la introducción (Capítulo 1), el FFL tiene diversas propiedades temporales que le permiten detectar estímulos persistentes [36], ser un módulo que separa e identifica la naturaleza del estímulo [72], entre otras. Estas características sumadas a que amplifica y adiciona ruido lo hacen único. Continuando con el análisis, el FFL no es

5. RESULTADOS & DISCUSIÓN

el único mecanismo demodulador en células eucariotas, otro mecanismo es la afinidad del promotor por su factor de transcripción, por lo que se procedió a analizar sus características de forma aislada y acoplado con el FFL.

5.2 Caracterización de la afinidad de un gen como mecanismo demodulador

Para investigar la manera en que la afinidad del promotor afecta la transmisión de la información, se acoplaron dos conjuntos de genes interactuando (W) y sin interactuar (WO) con la sección decodificadora de la vía MAP/ERK. Cada conjunto está conformado por tres genes con la única diferencia de tener valores diferentes de afinidad, Fig. 5.6(a). Para obtener diferentes valores de afinidad se modificó el parámetro que determina la constante de disociación del factor de transcripción con el promotor (K_{rp} , Ec. (4.17)) que está inversamente relacionado a la constante de afinidad; por lo que un valor alto para K_{rp} establece una baja afinidad y viceversa, los cambios se hicieron incrementando o decrementando un orden de magnitud el valor nominal del parámetro. Los sufijos **B**, **N** y **A** etiquetan al gen con baja, normal y alta afinidad. Se ocupó la misma descripción estocástica para todos los genes, con la diferencia en el valor de la propensión que es afectada por K_{rp} , tabla 4.1.

Con el objetivo de evaluar la relación entre la afinidad y la cantidad de información, se calculó el valor de CC para cada gen. De manera inicial, evaluando el efecto de la afinidad únicamente; se evalúa el conjunto de genes que no interacciona con la SD. Se observa que el gen con la afinidad mayor presenta el valor mayor de capacidad de canal; mientras más grande sea el valor afinidad mayor la cantidad de información que se puede transmitir, Fig. 5.6(b). Para determinar el efecto que se tiene al acoplar los dos mecanismos demoduladores, se obtuvo la CC para el conjunto de genes interaccionando con la SD. Se observa que el resultado anterior se conserva: mayor afinidad mayor CC, Fig. 5.6(c). Aunado a esto, la presencia del FFL potencia la cantidad de información transmitible independientemente del valor de afinidad. Por otra parte, si se quisiera obtener un valor alto de CC es posible hacerlo con una afinidad elevada; sin embargo, es posible tener el mismo valor al

5.2 Caracterización de la afinidad de un gen como mecanismo demodulador

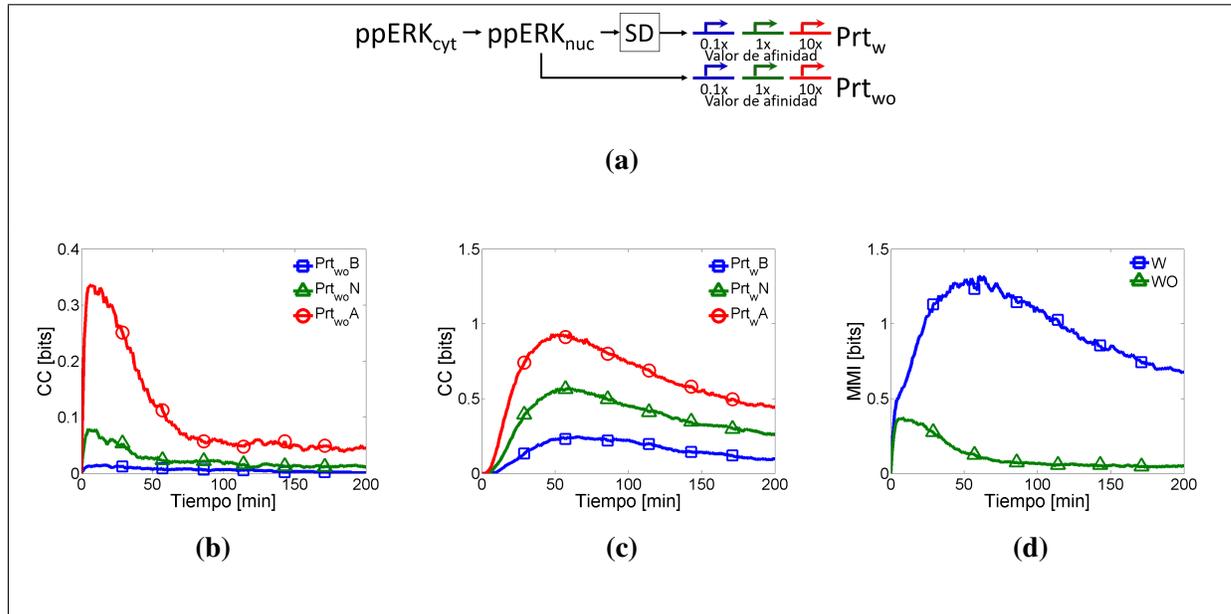


Figura 5.6: Análisis del mecanismo de demodulador por afinidad del promotor y del acoplamiento de ambos mecanismos. (a) Dos conjuntos de tres genes fueron acoplados interactuando (W) y sin interactuar (WO) con la sección decodificadora de la vía MAP/ERK. El número debajo de cada gen indica el escalamiento para cada valor de afinidad con respecto al valor nominal. (b) - (c) Series de tiempo de la CC para los grupos de genes, los sufijos **B**, **N** y **A** de cada gen indican el valor de la afinidad como bajo, normal y alto, respectivamente. (d) Información mutua multivariable (MMI) para cada grupo de genes en función del tiempo.

acoplar ambos mecanismos con un valor de afinidad menor.

Se postula que el acople de los mecanismos demoduladores ayuda a incrementar el valor de la información sin necesidad de exceder los límites energéticos o estructurales de la célula. La regulación de genes posee muchas restricciones biofísicas como la programabilidad, la velocidad de la respuesta, el rango dinámico de respuesta, la energía de unión entre el FT y el ADN por nombrar algunas; todas estas características se ven afectadas por los parámetros cinéticos del sistema [109], por lo que al modificarse un parámetro como la afinidad, puede ocasionar un desbalance o simplemente que no sea biológicamente posible, lo que resalta la importancia de la presencia del FFL.

5. RESULTADOS & DISCUSIÓN

El incrementar el valor de la afinidad no es la única manera en aumentar la CC, se ha demostrado experimentalmente que la información se incrementa cuando se considera múltiples copias de un gen [50]. Aunado a esta situación, en una población de células genómicamente idénticas existen diferencias intrínsecas, lo que induce diferencias entre los diversos genes [52, 81] como en su valor de afinidad entre el FT y el promotor. Por consiguiente, resalta la pregunta si múltiples genes con diferente valor de afinidad como un conjunto incrementan el valor de la CC.

Para evaluar la información total contenida en cada conjunto de genes se calculó la cantidad total de información no redundante contenida en los conjuntos de genes mediante la información mutua multivariable (Sección 2.4.4). La figura 5.6(d) muestra la MMI en función de tiempo para el conjunto interactuando con la SD (W) y para el conjunto sin interactuar (WO). De las primeras cosas que se aprecian es que el conjunto W tiene el valor de MMI mayor comparado con el grupo WO, lo cual es congruente con los resultados mostrados y discutidos con anterioridad. A su vez, se resalta que la información contenida en ambos grupos de genes es mayor que la CC en los genes por separado; de tal manera, que la heterogeneidad del valor de afinidad incrementa la CC. Por otro lado mayor afinidad conlleva una mayor expresión, para evaluar el efecto entre la afinidad y la amplificación se acoplaron grupos de genes al modelo mínimo.

Un grupo de genes se acoplaron con la proteína B por lo que este grupo no interacciona con el doble FFL anidado. El segundo conjunto de genes es activado por la molécula E . Los sufijos **A**, **N** y **B** son las etiquetas que al igual que el caso anterior denotan un valor de afinidad alto, normal y bajo, respectivamente. Los diversos genes acoplados se modelaron de la siguiente manera.

$$\frac{dFT_i}{dt} = k \left(\frac{FT}{FT + K_{rp}} \cdot \frac{A_N G_N + K_{rp}}{A_N G_N} \right) - gFT_i \quad (5.3)$$

donde FT es la proteína que activa el gen, ya sea B o E y el subíndice i es la etiqueta que determina el valor de la afinidad. A_N y G_N están determinados por las Ecs. (3.2) y (4.15),

5.2 Caracterización de la afinidad de un gen como mecanismo demodulador

respectivamente. Bajo las mismas condiciones que en el análisis en la vía MAP/ERK, los diferentes valores de afinidad se obtuvieron al modificar la constante de disociación del FT por el promotor (K_{rp}) un orden de magnitud, Fig. 5.7(a). Determinando el valor de estado estacionario para la Ec. (5.3) se tiene que es $FT_i^* = k/g$ para cualquier valor de K_{rp} . Se aprecia que ambos conjuntos presentan el mismo valor de estado estacionario dado que la vía mínima no presenta ganancia en ningún punto, Fig. 5.7(b).

La figura 5.7(c) muestra las series de tiempo de la capacidad del canal para los genes de los distintos conjuntos. Cabe resaltar que, los genes con menor valor de afinidad presentan la mayor cantidad de CC; estos resultados se oponen a los obtenidos en la vía MAP/ERK. Como ya se mostró, la variabilidad promueve el aumento en la cantidad de información. Por lo tanto, se prosiguió a calcular el CV, Fig. 5.7(d). Cuando la amplitud es limitada, mientras menor sea el valor de afinidad, el gen posee mayor variabilidad. Ya se probó que la variabilidad incrementa la cantidad de información, estableciendo el beneficio de tener un proceso estocástico y susceptible a ruido, lo cual es congruente con los resultados previos.

En conjunto, los resultados obtenidos en la evaluación de la afinidad concuerdan con Cheong et al. (2011). Ellos demostraron que múltiples copias de un gen incrementan la información en la vía de señalización del NF- κ B¹. Como resultado del incremento de la información, el sistema es capaz de discriminar diferentes concentraciones de TNF². Un conjunto de genes extrae más información que un gen solo. Dado que la información no redundante no es la suma de las contribuciones individuales, lo cual implica que cada gen extrae cierta cantidad de información única. A su vez, también se extrae información compartida entre todos los genes, la cual se ha demostrado que provee robustez al sistema [111]. La vía MAP/ERK tiene la capacidad de identificar estímulos y desencadenar diferentes respuestas [20]; sin embargo, no se ha demostrado que puede discriminar intensidades.

¹Vía involucrada en la respuesta inmune, activada frente a estímulos de estrés, radiación, citosinas, antígenos, entre otros [110].

²Factor de necrosis tumoral, citocina liberada por las células del sistema inmune que interviene en la inflamación, la apoptosis y la destrucción celular.

5. RESULTADOS & DISCUSIÓN

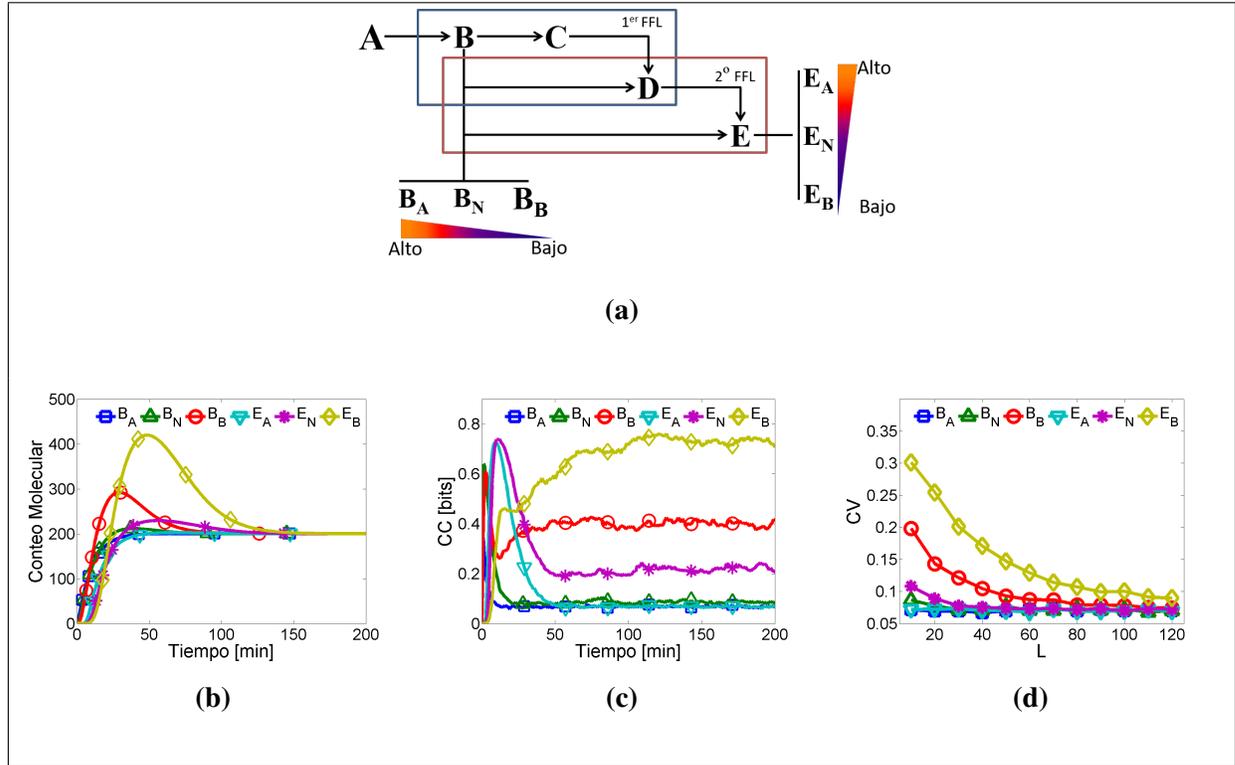


Figura 5.7: Estudio de afinidad con la amplificación controlada mediante el modelo mínimo. (a) Se acoplaron dos grupos de tres genes a distintos niveles de la vía mínima, el primer grupo se acopló directamente con la primera proteína (B) y el segundo grupo con la última (E). Los genes de cada grupo tienen diferentes valores de afinidad, los sufijos A , N y B corresponden a valores de afinidad en el promotor del gen para alto, nominal y bajo, respectivamente. Los incrementos y decrementos en este parámetro son de un orden de magnitud. (b) Conteo molecular de los genes acoplados. (c) Capacidad del canal en función del tiempo para los seis genes acoplados a diferentes niveles de la vía mínima. (d) CV para los seis genes con respecto al parámetro de escala L a $t = 150$ min.

En levaduras se ha demostrado que sólo es posible la identificación de la identidad del estímulo pero no la intensidad, funcionando como un interruptor [49]. Bajo este paradigma, el proceso de demodulación es un sistema digital que cambia entre diferentes estados, donde cada estado es un estímulo diferente. No obstante, Al acoplar ambos mecanismo demoduladores la información se incrementa, conllevando a una posible identificación de

intensidades en los estímulos.

Como último bloque de resultados, se procedió a estudiar la contribución del ruido extrínseco en el proceso de transmisión de la información. El análisis anterior fue realizado considerando sólo el ruido intrínseco generado por las interacciones moleculares.

5.3 Ruido Extrínseco

El ruido permea en todos los niveles de los procesos biológicos, de manera general se agrupa en dos categorías: intrínseco y extrínseco. Se ha demostrado que el ruido extrínseco provee de características interesantes que resultan benéficas en los procesos celulares como correlacionar respuestas, la resonancia estocástica, evolución, entre otras [112]. Sin embargo, en la transmisión de la información cualquier fuente de ruido es perjudicial [50, 106, 107]. De los resultados anteriores se ha demostrado que la presencia de variabilidad intrínseca es benéfica en el proceso de transmisión incrementando la CC. Por consiguiente, resulta interesante evaluar el efecto y contribución del ruido extrínseco.

Para explorar el efecto del ruido extrínseco en la transmisión de la información, se modeló una fuente de variabilidad extrínseca. De acuerdo con Huh y Paulsson (2011) y Ladbury y Arold (2013) esta fuente en las vías de señalización está dada en los procesos de fosforilación. Por ende, se escogió la fosforilación de *c-Fos* determinado por k_{11} (Sección 3.7). Este parámetro se consideró una variable aleatoria, por lo tanto tiene una distribución de probabilidad asociada. La forma de la distribución está controlada por el parámetro $\varepsilon_{\text{noise}}$ a valores mayores, la distribución se hace más ancha. Se calcula un número aleatorio con la distribución asignada y se asigna a k_{11} ; cada cierto tiempo P_{ext} se recalcula y asigna un nuevo valor para la tasa de fosforilación. $\varepsilon_{\text{noise}}$ y P_{ext} están asociados a la intensidad y al inverso de la frecuencia del ruido extrínseco. Se estimó el valor de la CC en diferentes condiciones.

La figura 5.8 muestra la capacidad del canal del gen acoplado a la salida del modelo de la vía MAP/ERK (Fig. 3.6(a)) fijando el valor de intensidad ante ocho diferentes valores de

5. RESULTADOS & DISCUSIÓN

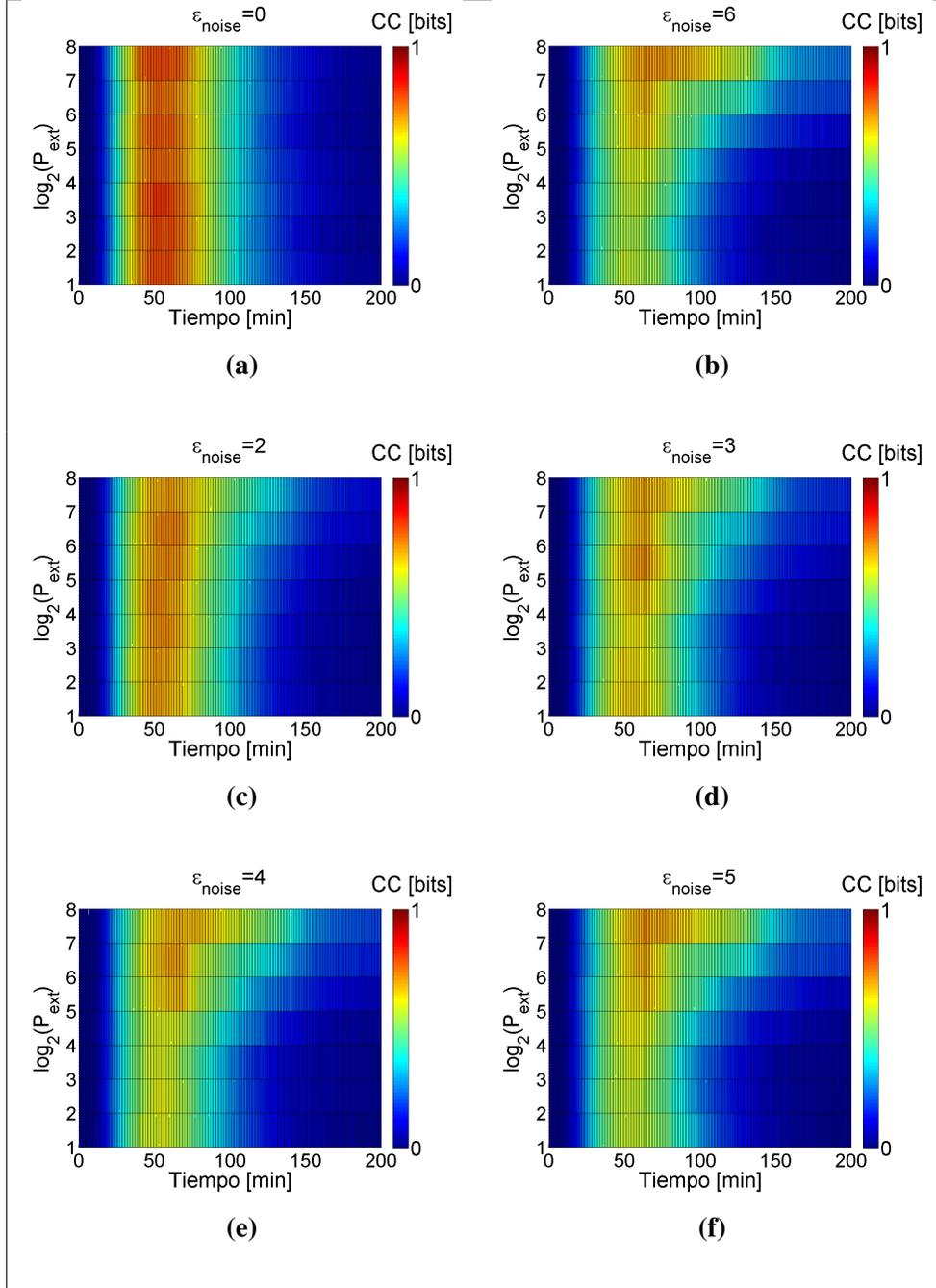


Figura 5.8: Evaluación de la capacidad del canal en función del tiempo ante ruido intrínseco y diferentes condiciones de ruido extrínseco. Se evaluó el efecto de la intensidad (ϵ_{noise}) y el tiempo de residencia de las condiciones de ruido extrínseco (P_{ext}), la barra de color representa el valor de CC. (a) $\epsilon_{noise} = 0$. (b) $\epsilon_{noise} = 6$. (c) $\epsilon_{noise} = 2$. (d) $\epsilon_{noise} = 3$. (e) $\epsilon_{noise} = 4$. (f) $\epsilon_{noise} = 5$.

5.3 Ruido Extrínseco

frecuencia para la fuente de ruido extrínseca, la barra de color representa el valor de la CC. Como experimento control se evaluó la CC ante una intensidad cero; se observa que la CC tiene el mismo comportamiento temporal ante los diferentes valores de P_{ext} , Fig. 5.8(a). Comparando este perfil con el obtenido para el máximo valor de $\varepsilon_{\text{noise}}$ (Fig. 5.8(b)), se observa al analizar cualquier línea horizontal, fijando P_{ext} , la CC disminuye. Por lo tanto, el ruido extrínseco es perjudicial en el proceso de transmisión de información. Sin embargo, cuando la frecuencia empieza a disminuir (mayores valores de P_{ext}) la CC se recupera; además, el lapso en que se tiene un alto valor de la capacidad del canal se ensancha. En consecuencia, el sistema puede tolerar ruido en altas intensidades siempre y cuando sea en frecuencias bajas del orden de las decenas de mHz, ya que de lo contrario hay pérdida de información. Los mismos resultados son observados para diferentes valores de $\varepsilon_{\text{noise}}$, Figs. 5.8(c)-5.8(f).

De acuerdo con Selimkhanov et al. (2014), la exclusiva participación del ruido extrínseco aumenta la cantidad de información transmitible debido al incremento del rango dinámico de la respuesta. Por el contrario, cuando las dos fuentes de ruido son consideradas, la información disminuye dramáticamente; resultado que concuerda con los obtenidos en el presente trabajo. El ruido intrínseco induce variabilidad decorrelacionando a la respuesta del sistema, en contraste, el ruido extrínseco produce fluctuaciones atenuadas a la dinámica de la respuesta [52, 113]. Cuando interactúan ambas fuentes de ruido inducen la pérdida de información como dos fuerzas opuestas. Existen otros mecanismos para mitigar el efecto del ruido, se ha demostrado que las respuestas dinámicas son más robustas al poder incluir mayor cantidad de ruido y verse menos afectadas [113]. Adicionalmente, un mecanismo para proveer robustez a la pérdida de información es mediante el uso de lazos de retroalimentación negativa [107], sin embargo se ha demostrado que el uso de estos lazos de control limita y restringe la información [50]. Lestas et al. (2010) determinó matemáticamente que la transmisión de información química está fundamentalmente limitada por el número de reacciones integradas en el tiempo de escala del fenómeno. Ellos establecieron que el error de transmisión decrece como la raíz cuarta del número de reacciones; para reducir significativamente el error se necesitan un gran número de reacciones lo cual es energéticamente costoso y probablemente ineficiente. En el presente trabajo se demuestra

5. RESULTADOS & DISCUSIÓN

que en lugar de incrementar el número de reacciones, la célula utiliza estructuras de interacción topológica (como el FFL) que le permita mejorar la transmisión de información. Además, este proceso es robusto a la pérdida de información si el sistema se encuentra afectado por ruido de baja frecuencia.

5.4 Resumen

El presente trabajo se basó en caracterizar y estudiar los mecanismos demoduladores mediante el uso de modelos estocásticos de la vía de señalización MAP/ERK y de una vía mínima. El análisis se efectuó al separar las propiedades de cada mecanismo con el fin de poder estudiar la relevancia y participación de cada característica por separado. Los modelos se evaluaron ante diferentes condiciones haciendo uso de herramientas matemáticas; se obtuvo la cantidad de información total que es posible transmitir ante cada condición. Esto se logró utilizando la teoría de información como marco conceptual al estudiar el proceso biológico como un canal de comunicación y caracterizarlo como tal.

Los resultados obtenidos no se particularizan a la vía de señalización MAP/ERK. Esta vía se ocupó como modelo inicial ya que se ha comprobado experimentalmente que posee los mecanismos demoduladores. Los resultados pueden extenderse a cualquier vía que contenga dentro de su estructura un FFL. Esto es posible ya que se ocupó el modelo mínimo que es la representación genérica de cualquier vía estructurada en una topología tipo FFL.

Por consiguiente, el presente trabajo elucidó que la función del FFL como mecanismo demodulador es: acoplar, amplificar y añadir variabilidad al sistema con el objetivo de aumentar los límites de transmisión de información. De forma general la amplificación y adición de variabilidad intrínseca es benéfica y la forma en que está estructurado el FFL potencia el efecto. Por otra parte cuando se acoplan ambos mecanismos, la información aumenta dando la posibilidad de reducir el efecto de cuello de botella que por naturaleza el FFL lleva a cabo. A su vez, el efecto de tener una población celular es benéfico permitiendo captar mayor información del mensaje transmitido y que en conjunto permita al sistema responder a un mayor número de estímulos o hacerlo de una manera más precisa.

5.4 Resumen

Todo estos resultados se ven afectados en presencia de ruido extrínseco, sin embargo si las fluctuaciones son de frecuencia baja (mHz) independientemente de la intensidad es posible que el comportamiento del sistema sea correcto y eficiente.

CAPÍTULO

6

Perspectivas

El presente trabajo plantea un enfoque teórico-computacional originado en el área de telecomunicaciones aplicado al campo de la biología. Tal perspectiva es un contexto natural de estudio si se plantea la existencia de un alfabeto o lenguaje químico ocupado por las células. A través de ese lenguaje se modula la información con el objeto de que exista coordinación de estímulo-respuesta. Es importante señalar la diferencia entre información y significado; el significado involucra un proceso cognitivo. La información es un conjunto organizado de datos de cualquier naturaleza que inducen un cambio. El estudiar el flujo de información es de importancia para esclarecer los nodos en los cuales la información es críticamente afectada ya sea brindando robustez o su pérdida. De tal motivación surgió el interés de estudiar los mecanismo demoduladores, determinándose que funcionan como amplificadores ruidosos; resultados interesantes dada la implicación evolutiva de utilizar la variabilidad a favor en lugar ocupar recursos para borrarla del panorama, otro ejemplo de este proceso es la resonancia estocástica. Sin embargo, la presente caracterización sólo establece una relación limitada con respecto a todos los procesos que afectan ya sea de manera positiva o negativa a la información.

6. PERSPECTIVAS

Se postula que biológicamente la gran mayoría de la información está contenida en las dinámicas temporales de las señales, incluyendo cambios en amplitud y frecuencia de la respuesta. En el trabajo presentado, se ignoró la componente temporal de la respuesta debido a que no existe la formulación teórica para llevar a cabo tal cuantificación, especialmente para sistemas no estacionarios. A su vez, dado que las células ocupan diferentes esquemas de modulación, es interesante plantear las propiedades de cada uno de ellos y para qué casos cada esquema es óptimo.

Se necesitan establecer nuevos sistemas teórico-experimentales que permitan expandir los resultados, como el control de las fuentes de ruido o mediante el uso de biología sintética acoplar genes reporteros en diferentes puntos de las vías. Por otro lado, se modeló implícitamente la participación del bloque modulador, sin embargo el estudiar este bloque de forma aislada y acoplada al bloque demodulador de forma explícita establecería relevancia ya que es la primera sección por donde la información transita, que a su vez podría determinar los puntos clave del procesamiento. Con la misma idea de caracterizar cuáles son las interacciones con mayor relevancia, hace falta un estudio de robustez paramétrica, sin embargo se debe de realizar de tal manera que no se enmascare otros fenómenos como el cambio en la amplificación o adición de ruido ya que como resultado de este trabajo se demostró su gran importancia. Tal que, todo parámetro que afecte la amplificación, va a afectar la cantidad de información.

Como resultado de estos estudios, se espera que se genere una descripción más detallada de los puntos y características claves en la transmisión de la información, que en combinación con la extensión y formulación de los conceptos de la teoría de información a la biología significaría un paso adelante en el entendimiento de la eficiencia de las células en la regulación de sus procesos; por eficiencia se entiende el uso óptimo de los recursos que están a su alcance, incluso la misma presencia de ruido.

Conclusión

La demodulación, sin importar si se refiere al área de telecomunicaciones o biología, restringe o aumenta la calidad de la comunicación. Si las entidades que transfieren información están bien coordinadas y sintonizadas en un ambiente con bajos niveles de ruido, el mensaje es transmitido correctamente. La realidad biológica demuestra que el ruido es un factor inamovible y en algunos casos determinante. En el área de ingeniería y desarrollo tecnológico, se usan dispositivos con la capacidades de filtrar y recuperar la información con la fidelidad necesaria para que el proceso sea robusto. A nivel celular, la información inevitablemente se perderá debido a la presencia de estocasticidad en las interacciones. No obstante, la célula hace uso de estructuras y recursos para sacar ventaja de la presencia del ruido, tales como la presencia de FFL y utilizar la variabilidad para aumentar los límites de la cantidad de información a extraer.

En conclusión, el lazo pre-alimentador tiene la propiedad de extraer más información que otras topologías viables para llevar a cabo la señalización celular como una cadena de interacción lineal. Amplifica e incrementa el rango dinámico añadiendo variabilidad en las

7. CONCLUSIÓN

interacciones moleculares, estos efectos se ven potenciados cuando es acoplado y considerado un gen al cual se regula su afinidad. Por consiguiente, la principal función de los mecanismo demoduladores en el proceso de transmisión de información es amplificar, acoplar y añadir variabilidad. Asimismo, si el sistema se encuentra en un ambiente fluctuante con baja frecuencia, sin importar la intensidad, el proceso de transmisión de la información es robusto.

APÉNDICE

A

Análisis de Estabilidad

En esta sección se presenta el análisis matemático para determinar la estabilidad de los tres modelos descritos en el capítulo 4. El análisis parte de la descripción determinista de cada modelo. Para determinar la estabilidad, primero se obtiene el valor de estado estacionario del sistema. Matemáticamente el estado estacionario se calcula al igualar la razón de cambio de la molécula (derivada) a cero, esto implica que la molécula ya no cambia con respecto al tiempo. El valor de estado estacionario se denota con el símbolo * sobre cada variable.

A.1 Modelo de la sección demoduladora de la vía MAP/ERK

De manera inicial, se analiza la estabilidad del modelo de la vía MAP/ERK, Ecs.(4.1)-(4.8). Se determina el valor de SS para la entrada ($ppERK_{cyt}$), en la sección 3.1 se muestra que $ppERK_{cyt}$ tiene un estado estacionario positivo siempre y cuando $\kappa_1 \geq \kappa_2$. Los valores de estado estacionario para las moléculas restantes se muestran en las Ecs. (A.3)-(A.8). De manera particular para obtener el SS de $Dusp$ se resuelve la ec. (4.7) en el dominio

A. ANÁLISIS DE ESTABILIDAD

temporal y se evalúa el límite cuando el tiempo tiende a infinito. Resultando que:

$$Dusp(t) = Dusp(0) e^{kt} \quad (\text{A.1})$$

$$Dusp(t)^* = \lim_{t \rightarrow \infty} Dusp(t) = \infty \quad (\text{A.2})$$

Donde $Dusp(0)$ representa su condición inicial, se observa que esta variable diverge en el tiempo de forma exponencial por lo que en términos prácticos nunca llega al estado estacionario¹ implicando la no existencia del estado estacionario. Analizando el SS para $ppERK_{nuc}$, Ec. (A.3); se observa que depende inversamente de $Dusp^*$, y ya que esta tiende a infinito $ppERK_{nuc}^* = 0$. De manera general, observando las ecuaciones que definen SS para las variables restante se encuentra que $SS \propto ppERK_{nuc}^*$ implicando que todas la variables tengan un único estado estacionario igual a cero. Esto tiene sentido de manera biológica, ya que la vía MAP/ERK se activa de acuerdo a estímulos externos. Por lo que después de activarse necesita regresar a sus condiciones basanales para poder ser activado por un nuevo estímulo. $Dusp$ es el mecanismo por el cual la vía se desactiva.

$$ppERK_{nuc}^* = \frac{k_1}{k_2 + k_3 Dusp^*} ppERK_{cyt}^* \quad (\text{A.3})$$

$$pRSK_{nuc}^* = \frac{k_4}{k_5} ppERK_{nuc}^* \quad (\text{A.4})$$

$$c-Fos_{pre}^* = \frac{k_4^n}{k_7 k_4^n + k_7 (k_5 k_6)^n} ppERK_{nuc}^{*2n} \quad (\text{A.5})$$

$$c-Fos_m^* = \frac{k_4^n}{k_8 k_4^n + k_8 (k_5 k_6)^n} ppERK_{nuc}^{*2n} = A ppERK_{nuc}^{*2n} \quad (\text{A.6})$$

$$c-Fos^* = \frac{k_9 A}{k_{10} + k_{11} ppERK_{cyt}^* - k_{13} \frac{k_{11}}{k_{12} + k_{13}} ppERK_{cyt}^*} ppERK_{nuc}^{*2n} \quad (\text{A.7})$$

$$pc-Fos^* = \frac{k_{11}}{k_{12} + k_{13}} ppERK_{cyt}^* c-Fos^* \quad (\text{A.8})$$

Ya obtenido el estado estacionario de las variables, se procede a determinar la estabilidad mediante el cálculo de los eigenvalores (λ) del sistema. Esto se hace igualando a cero

¹El mismo resultado se obtiene al igualar su derivada a cero, resultando que $kDusp^* = 0$ expandiendo está ecuación con la Ec. (A.1) se tiene que: $kDusp(0) e^{kt} = 0$. La única manera en que se cumple la ecuación dado que los parámetros son positivos es cuando $t \rightarrow -\infty$

A.1 Modelo de la sección demoduladora de la vía MAP/ERK

el determinante de la diferencia entre la matriz identidad y el Jacobiano (J) evaluado en el SS de las variables, Ec. (A.9).

$$\det|\lambda I - J_{SS}| = 0 \quad (\text{A.9})$$

$$J_{SS} = \begin{bmatrix} \frac{\partial f_1}{\partial ppERK_{nuc}} & \cdots & \frac{\partial f_1}{\partial pc-Fos} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_7}{\partial ppERK_{nuc}} & \cdots & \frac{\partial f_7}{\partial pc-Fos} \end{bmatrix}_{SS} \quad (\text{A.10})$$

El jacobiano se obtiene mediante la ecuación (A.10), donde $f_1 \dots f_7$ son las ecuaciones diferenciales definidas en (4.2)-(4.8). Evaluando la Ec. (A.9) se obtiene que :

$$\det \begin{bmatrix} \lambda + k_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ -k_4 & \lambda + k_5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda + k_7 & 0 & 0 & 0 & 0 \\ 0 & 0 & -k_7 & \lambda + k_8 & 0 & 0 & 0 \\ 0 & 0 & 0 & -k_9 & \lambda + k_{10} & 0 & -k_{13} \\ 0 & 0 & 0 & 0 & 0 & \lambda - k & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda + k_{12} + k_{13} \end{bmatrix} = 0 \quad (\text{A.11})$$

La matriz descrita por la Ec. (A.11) tiene la propiedad de ser una matriz triangular inferior, lo que implica que sus elementos por arriba de la diagonal principal son cero. A pesar que existe un elemento diferente a cero ($-k_{13}$) debido a la simetría con su elemento transpuesto la propiedad se conserva. Tal característica determina que el resultado del determinante es la multiplicación de los elementos de la diagonal principal, Ec. A.12.

$$(\lambda + k_2)(\lambda + k_5)(\lambda + k_7)(\lambda + k_8)(\lambda + k_{10})(\lambda - k)(\lambda + k_{12} + k_{13}) = 0 \quad (\text{A.12})$$

Resolviendo el sistema de séptimo orden se obtienen los eigenvalores, se determina la estabilidad mediante su signo. Si es negativo el sistema es estable; en caso contrario para valores positivos se establece que esa variable es inestable. Todos los valores de los parámetros al ser definidos como tasas de reacción son positivos.

$$\begin{aligned} \lambda_1 &= -k_2 & \lambda_5 &= -k_{10} \\ \lambda_2 &= -k_5 & \lambda_6 &= k \\ \lambda_3 &= -k_7 & \lambda_7 &= -(k_{12} + k_{12}) \\ \lambda_4 &= -k_8 \end{aligned}$$

Cada eigenvalor (λ) se asocia a cada variable de las ecuaciones diferenciales (4.2)-(4.8). El

A. ANÁLISIS DE ESTABILIDAD

único eigenvalor positivo es λ_6 siendo el correspondiente a $Dusp$, lo cual comprueba nuevamente su divergencia. Como resultado del análisis mostrado, se determinó que el modelo de la vía MAP/ERK es estable y converge a cero.

A.2 Modelo Mínimo

Realizando el mismo análisis para determinar la estabilidad del modelo mínimo, el polinomio característico para obtener los eigenvalores resultado de las Ecs. (A.9) y (A.10) se muestra a continuación:

$$(\lambda + g)^4 = 0 \quad (\text{A.13})$$

Los eigenvalores para todas las variables en el modelo mínimo son negativos e iguales ($\lambda = -g$), dado que g sólo toma valores positivos el sistema es estable para todo valor.

A.3 Modelo de genes reporteros

Se obtiene el estado estacionario del sistema de expresión génica igualando las Ecs. (4.17)-(4.19) a cero. El valor de SS está determinado por las Ecs. (A.14)-(A.16). El sistema de expresión se modeló con la intención de acoplarlo como gen reportero en los modelos de la vía MAP/ERK y el modelo mínimo. El sistema tiene un único SS ; como se aprecia, P_m^* depende del SS de factor de transcripción (FT^*). Si el gen es acoplado a la vía MAP/ERK, el estado estacionario P_m^* , m^* y P^* es cero, ya que el $FT = pc-Fos$ y $pc-Fos^* = 0$ (Sección 4.1). Dado que el SS del modelo mínimo es mayor a cero, cuando es sistema de expresión es acoplado a tal, el SS es positivo.

$$P_m^* = \kappa_{pm} \frac{FT^{*n}}{FT^{*n} (\gamma_{pm} + \kappa_{pm}) + K_{rp}^n \gamma_{pm}} \quad (\text{A.14})$$

$$m^* = \frac{\kappa_m}{\gamma_m} P_m^* \quad (\text{A.15})$$

$$P^* = \frac{\kappa_p \kappa_m}{\gamma_p \gamma_m} P_m^* \quad (\text{A.16})$$

Realizando el mismo análisis para determinar la estabilidad del sistema de expresión génica, el polinomio característico para obtener los eigenvalores resultado de las Ecs. (A.9)

A.3 Modelo de genes reporteros

y (A.10) se muestra a continuación:

$$\left(\lambda + \frac{FT^{*n}}{FT^{*n} + K_{rp}^n} + \gamma_{pm} \right) (\lambda + \gamma_m) (\lambda + \gamma_p) = 0 \quad (\text{A.17})$$

Los eigenvalores para todas las variables son negativos indicando que el sistema es estable en todo el espacio paramétrico.

$$\lambda_{P_m} = - \left(\frac{FT^{*n}}{FT^{*n} + K_{rp}^n} + \gamma_{pm} \right)$$

$$\lambda_m = -\gamma_m$$

$$\lambda_P = -\gamma_p$$

Bibliografía

- [1] Takashi Nakakuki, Marc R. Birtwistle, Yuko Saeki, Noriko Yumoto, Kaori Ide, Takeshi Nagashima, Lutz Brusch, Babatunde A. Ogunnaike, Mariko Okada-Hatakeyama, y Boris N. Kholodenko. Ligand-specific c-fos expression emerges from the spatiotemporal control of ErbB network dynamics. *Cell*, 141(5):884–896, 2010. 1, 11, 39, 40, 41, 42, 43, 56, 57, 61, 62, 68
- [2] Nan Hao y Erin K O’Shea. Signal-dependent dynamics of transcription factor translocation controls gene expression. TL - 19. *Nature structural & molecular biology*, 19 VN - r(1):31–39, 2012. 2, 8, 9, 12
- [3] J. J. Walker, J. R. Terry, K. Tsaneva-Atanasova, S. P. Armstrong, C. A. McArdle, y S. L. Lightman. Encoding and Decoding Mechanisms of Pulsatile Hormone Secretion. *Journal of Neuroendocrinology*, 22(12):1226–1238, 2010. 8
- [4] Angelique Whitehurst, Melanie H Cobb, y Michael A White. Stimulus-Coupled Spatial Restriction of Extracellular Signal-Regulated Kinase 1/2 Activity Contributes to the Specificity of Signal-Response Pathways. *Molecular and Cellular Biology*, 24(23):10145–10150, dec 2004. 7, 9
- [5] Hirenkumar K. Makadia, James S. Schwaber, y Rajanikanth Vadigepalli. Intracellular Information Processing through Encoding and Decoding of Dynamic Signaling Features. *PLoS Computational Biology*, 11(10), 2015. 2, 12, 68

BIBLIOGRAFÍA

- [6] C Salazar, A Z Politi, y T Hofer. Decoding of calcium oscillations by phosphorylation cycles: analytic results. *Biophys J*, 94(4):1203–1215, 2008. 2, 6, 10
- [7] Marcelo Behar y Alexander Hoffmann. Understanding the temporal codes of intracellular signals. *Current Opinion in Genetics and Development*, 20(6):684–693, 2010. 6, 7, 8, 10, 67
- [8] Erik Smedler y Per Uhlén. Frequency decoding of calcium oscillations. *Biochimica et Biophysica Acta - General Subjects*, 1840(3):964–969, 2014. 2, 8, 10
- [9] A. Maáyan. Formation of Regulatory Patterns During Signal Propagation in a Mammalian Cellular Network. *Science*, 309(5737):1078–1083, 2005. 2, 4
- [10] S. L. Miller. A Production of Amino Acids Under Possible Primitive Earth Conditions. *Science*, 117(3046):528–529, 1953. 3
- [11] Sue E Huether y Kathryn L McCance. *Understanding pathophysiology*. Mosby, Elseiver, St. Louis, Missouri, fifth ed. edition, 2012. ISBN 9780323354097 0323354092 0323431240 9780323431248. 3
- [12] David L Nelson, David L Nelson, Albert L Lehninger, y Michael M Cox. *Lehninger principles of biochemistry*. W.H. Freeman, New York, fifth edition, 2008. ISBN 071677108X 9780716771081 9781429208925 1429208929. 4, 8, 19
- [13] Spiros A Vlahopoulos, Osman Cen, Nina Hengen, James Agan, Maria Moschovi, Elena Critselis, Maria Adamaki, Flora Bacopoulou, John A Copland, Istvan Boldogh, Michael Karin, y George P Chrousos. Dynamic aberrant NF- κ B spurs tumorigenesis: a new model encompassing the microenvironment. *Cytokine & growth factor reviews*, 26(4):389–403, aug 2015. 4
- [14] Kepeng Wang, Sergei I Grivennikov, y Michael Karin. Implications of anti-cytokine therapy in colorectal cancer and autoimmune diseases. *Annals of the Rheumatic Diseases*, 72(suppl 2):ii100 LP – ii103, mar 2013. 4

BIBLIOGRAFÍA

- [15] Geoffrey M Cooper y Robert E Hausman. *The cell : a molecular approach*. ASM Press ; Sinauer Associates, Washington, D.C.; Sunderland, Mass., seventh edition, 2009. ISBN 9780878933006 087893300X. 4, 7
- [16] Julian Downward. The ins and outs of signalling. *Nature*, 411(6839):759–762, jun 2001. 4
- [17] H Baher. *Analog & Digital Signal Processing*. Jhon Wiley & Sons, LTD, New York, 2nd editio edition, 2001. ISBN 0471 62354 7. 5
- [18] C.J Marshall. Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell*, 80(2):179–185, jan 1995. 5, 8
- [19] Arnab Chatterjee y Parongama Sen. Agent dynamics in kinetic models of wealth exchange. *Physical Review E*, 82(5):6, nov 2010.
- [20] Jeremy E. Purvis y Galit Lahav. Encoding and decoding cellular information through signaling dynamics. *Cell*, 152(5):945–956, 2013. 5, 6, 11, 23, 25, 55, 68, 81
- [21] Roger Brent. Cell signaling: What is the signal and what information does it carry? *FEBS Letters*, 583(24):4019–4024, 2009. 6
- [22] Roger P Alexander, Philip M Kim, Thierry Emonet, y Mark B Gerstein. Understanding Modularity in Molecular Networks Requires Dynamics. *Science signaling*, 2(81):pe44–pe44, jul 2009. 6
- [23] Joe H Levine, Yihan Lin, y Michael B Elowitz. Functional roles of pulsing in genetic circuits. *Science (New York, N.Y.)*, 342(6163):1193–200, 2013. 6
- [24] Anders S. Hansen y Erin K. O’Shea. Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression. *Molecular Systems Biology*, 9(1):704, 2013. 6, 8, 12

BIBLIOGRAFÍA

- [25] Savas Tay, Jacob J Hughey, Timothy K Lee, Tomasz Lipniacki, Stephen R Quake, y Markus W Covert. Single-cell NF- κ B dynamics reveal digital activation and analogue information processing. *Nature*, 466(7303):267–271, jul 2010. 7
- [26] J W Black, P Leff, N P Shankley, y J Wood. An operational model of pharmacological agonism: the effect of E/[A] curve shape on agonist dissociation constant estimation. *British Journal of Pharmacology*, 160(Suppl 1):S54–S64, jul 2010. 7
- [27] Jeffrey P Mackeigan, Leon O Murphy, a Christopher, John Blenis, y Christopher a Dimitri. Graded Mitogen-Activated Protein Kinase Activity Precedes Switch-Like c-Fos Induction in Mammalian Cells Graded Mitogen-Activated Protein Kinase Activity Precedes Switch-Like c-Fos Induction in Mammalian Cells. *Society*, 25(11):4676–4682, 2005.
- [28] Richard C Yu, C Gustavo Pesce, Alejandro Colman-Lerner, Larry Lok, David Pincus, Eduard Serra, Mark Holl, Kirsten Benjamin, Andrew Gordon, y Roger Brent. Negative feedback that improves information transmission in yeast signalling. *Nature*, 456(7223):755–761, dec 2008. 7
- [29] Long Cai, Chiraj K Dalal, y Michael B Elowitz. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature*, 455(7212):485–90, 2008. 7, 9
- [30] Clemens Utzny, Mustapha Faroudi, y Salvatore Valitutti. Frequency encoding of T-cell receptor engagement dynamics in calcium time series. *Biophysical journal*, 88(1):1–14, jan 2005. 8
- [31] Ann Zahle Larsen, Lars Folke Olsen, y Ursula Kummer. On the encoding and decoding of calcium signals in hepatocytes. *Biophysical Chemistry*, 107(1):83–99, jan 2004. 8
- [32] Akihiko Tanimura, Takao Morita, Akihiro Nezu, Akiko Shitara, Noboru Hashimoto, y Yosuke Tojyo. Use of fluorescence resonance energy transfer-based biosensors for

BIBLIOGRAFÍA

- the quantitative analysis of inositol 1,4,5-trisphosphate dynamics in calcium oscillations. *Journal of Biological Chemistry*, 284(13):8910–8917, 2009. 8
- [33] M J Berridge. Remodelling Ca^{2+} signalling systems and cardiac hypertrophy. *Biochemical Society Transactions*, 34(2):228 LP – 231, mar 2006. 8
- [34] M Robert Tombes, L Kelly Auer, Ross Mikelsen, Kristopher Valerie, P Mathias Wyman, J Christopher Marshall, Martin McMahon, y Paul Dent. The mitogen-activated protein (MAP) kinase cascade can either stimulate or inhibit DNA synthesis in primary cultures of rat hepatocytes depending upon whether its activation is acute/phasic or chronic. *Biochemical Journal*, 330(3):1451 LP – 1460, mar 1998. 8
- [35] Juan-Jose Ventura, Anette Hübner, Chao Zhang, Richard A Flavell, Kevan M Shokat, y Roger J Davis. Chemical Genetic Analysis of the Time Course of Signal Transduction by JNK. *Molecular Cell*, 21(5):701–710, mar 2006. 8
- [36] U Alon. Network motifs: theory and experimental approaches. *Nat Rev Genet*, 8(6):450–461, 2007. 10, 11, 77
- [37] Michael E. Wall, Mary J. Dunlop, y William S. Hlavacek. Multiple functions of a feed-forward-loop gene circuit. *Journal of Molecular Biology*, 349(3):501–514, jun 2005. 11
- [38] R Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii, y U Alon. Network Motifs: Simple Building Blocks of Complex Networks. *Science*, 298(5594):824 LP – 827, oct 2002. 11
- [39] Ron Milo, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, y Uri Alon. Superfamilies of Evolved and Designed Networks. *Science*, 303(5663):1538 LP – 1542, mar 2004. 11
- [40] Sarah A Teichmann y M Madan Babu. Gene regulatory network growth by duplication. *Nature Genetics*, 36(5):492–496, 2004. 11, 20

BIBLIOGRAFÍA

- [41] S Mangan y U Alon. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21):11980–11985, 2003. 11
- [42] Eric H Davidson, Jonathan P Rast, Paola Oliveri, Andrew Ransick, Cristina Calestani, Chiou-Hwa Yuh, Takuya Minokawa, Gabriele Amore, Veronica Hinman, César Arenas-Mena, Ochan Otim, C Titus Brown, Carolina B Livi, Pei Yun Lee, Roger Revilla, Alistair G Rust, Zheng jun Pan, Maria J Schilstra, Peter J C Clarke, Maria I Arnone, Lee Rowen, R Andrew Cameron, David R McClay, Leroy Hood, y Hamid Bolouri. A Genomic Regulatory Network for Development. *Science*, 295(5560):1669 LP – 1678, mar 2002. 11
- [43] Duncan T Odom, Nora Zizlsperger, D Benjamin Gordon, George W Bell, Nicola J Rinaldi, Heather L Murray, Tom L Volkert, Jörg Schreiber, P Alexander Rolfe, David K Gifford, Ernest Fraenkel, Graeme I Bell, y Richard A Young. Control of Pancreas and Liver Gene Expression by HNF Transcription Factors. *Science*, 303(5662):1378 LP – 1381, feb 2004. 11
- [44] Zhangang Han, Thomas M. Vondriska, Ling Yang, W. Robb MacLellan, James N. Weiss, y Zhilin Qu. Signal transduction network motifs and biological memory. *Journal of Theoretical Biology*, 246(4):755–761, 2007. 11
- [45] W H de Ronde, F Tostevin, y P R ten Wolde. Feed-forward loops and diamond motifs lead to tunable transmission of information in the frequency domain. *Physical Review E*, 86(2):21913, aug 2012. 11
- [46] Wenzhe Ma, Ala Trusina, Hana El-Samad, Wendell A. Lim, y Chao Tang. Defining Network Topologies that Can Achieve Biochemical Adaptation. *Cell*, 138(4):760–773, 2009. 11
- [47] Leon O Murphy, Jeffrey P MacKeigan, y John Blenis. A network of immediate early gene products propagates subtle differences in mitogen-activated protein kinase signal amplitude and duration. *Molecular and cellular biology*, 24(1):144–53, 2004. 11, 25

BIBLIOGRAFÍA

- [48] Vladimir Litvak, Stephen A Ramsey, Alistair G Rust, Daniel E Zak, Kathleen A Kennedy, Aaron E Lampano, Matti Nykter, Ilya Shmulevich, y Alan Aderem. Function of C/EBP δ in a regulatory circuit that discriminates between transient and persistent TLR4-induced signals. *Nature Immunology*, 10(4):437–443, 2009. 11
- [49] Anders S. Hansen y Erin K. O’Shea. Limits on information transduction through amplitude and frequency regulation of transcription factor activity. *eLife*, 4(MAY): 1–19, 2015. 12, 13, 82
- [50] R. Cheong, a. Rhee, C. J. Wang, I. Nemenman, y a. Levchenko. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science*, 334(6054): 354–358, 2011. 12, 13, 28, 35, 75, 80, 81, 83, 85
- [51] Clive G Cg Bowsher y Peter S Ps Swain. Identifying sources of variation and the flow of information in biochemical networks. *Proceedings of the National Academy of Sciences of the United States of America*, 109(20):E1320–8, 2012. 13, 28, 68
- [52] Michael B Elowitz, Arnold J Levine, Eric D Siggia, y Peter S Swain. Stochastic Gene Expression in a Single Cell. *Science*, 297(5584):1183 LP – 1186, aug 2002. 13, 28, 60, 80, 85
- [53] JE; Ladbury y ST; Arold. Noise in cellular signaling pathways : causes and effects. *Trends Biochem Sci*, 37(5):173–178, 2013. 28, 50, 68, 83
- [54] Jonathan M Raser y Erin K O’Shea. Noise in Gene Expression: Origins, Consequences, and Control. *Science*, 309(5743):2010 LP – 2013, sep 2005. 13
- [55] Johan Paulsson. Summing up the noise in gene networks. *Nature*, 427(6973):415–418, jan 2004. 13
- [56] Tyler M Earnest, Elijah Roberts, Michael Assaf, Karin Dahmen, y Zaida Luthey-Schulten. DNA looping increases the range of bistability in a stochastic model of the lac genetic switch. *Physical Biology*, 11(2):029501, 2014. 13

BIBLIOGRAFÍA

- [57] Dann Huh y Johan Paulsson. Random partitioning of molecules at cell division. *Proceedings of the National Academy of Sciences of the United States of America*, 108(36):15004–9, 2011. 13, 28, 50, 83
- [58] John E. Ladbury y Stefan T. Arold. Noise in cellular signaling pathways: Causes and effects. *Trends in Biochemical Sciences*, 37(5):173–178, 2012. 13, 28, 60
- [59] Alex Rhee, Raymond Cheong, y Andre Levchenko. Noise decomposition of intracellular biochemical signaling networks using nonequivalent reporters. *Proceedings of the National Academy of Sciences of the United States of America*, 111(48):17330–5, 2014. 13
- [60] David S Latchman. Transcription factors: An overview. *The International Journal of Biochemistry & Cell Biology*, 29(12):1305–1312, 1997. 20
- [61] Patrick Cramer, David A Bushnell, Jianhua Fu, Averell L Gnatt, Barbara Maier-Davis, Nancy E Thompson, Richard R Burgess, Aled M Edwards, Peter R David, y Roger D Kornberg. Architecture of RNA Polymerase II and Implications for the Transcription Mechanism. *Science*, 288(5466):640 LP – 649, apr 2000. 20
- [62] Stephen T Smale y James T Kadonaga. The RNA Polymerase II Core Promoter. *Annual Review of Biochemistry*, 72(1):449–479, jun 2003. 20
- [63] Melanie I. Stefan y Nicolas Le Novère. Cooperative Binding. *PLoS Computational Biology*, 9(6):2–7, 2013. 22
- [64] Lufen Chang y Michael Karin. Mammalian MAP kinase signalling cascades. *Nature*, 410(6824):37–40, mar 2001. 23
- [65] Robert Roskoski Jr. ERK1 / 2 MAP kinases : Structure , function , and regulation. *Pharmacological Research*, 66(2):105–143, 2012. 23, 24, 55
- [66] Bei Zheng, Paolo Fiumara, Yang V Li, Georgios Georgakis, Virginia Snell, Mamoun Younes, Jean Nicolas Vauthey, Antonino Carbone, y Anas Younes. MEK/ERK pathway is aberrantly active in Hodgkin disease: a signaling pathway shared by CD30,

BIBLIOGRAFÍA

- CD40, and RANK that regulates cell proliferation and survival. *Blood*, 102(3):1019 LP – 1027, jul 2003. 23
- [67] James A McCubrey, Linda S Steelman, William H Chappell, Stephen L Abrams, Ellis W T Wong, Fumin Chang, Brian Lehmann, David M Terrian, Michele Milella, Agostino Tafuri, Franca Stivala, Massimo Libra, Jorg Basecke, Camilla Evangelisti, Alberto M Martelli, y Richard A Franklin. Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1773(8):1263–1284, 2007. 23
- [68] Raymond C Harris, Eunkyung Chung, y Robert J Coffey. EGF receptor ligands. *Experimental Cell Research*, 284(1):2–13, 2003. 23
- [69] Bruce Alberts. *Molecular biology of the cell*. Garland Science, New York, 4th ed. edition, 2002. ISBN 9780815332183. 23
- [70] Silvia D M Santos, Peter J Verveer, y Philippe I H Bastiaens. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nature cell biology*, 9(3):324–330, 2007. 23, 25
- [71] Maria G Khrenova, Bella L Grigorenko, Anatoly B Kolomeisky, y Alexander V Nemukhin. Hydrolysis of Guanosine Triphosphate (GTP) by the Ras-GAP Protein Complex: Reaction Mechanism and Kinetic Scheme. *The Journal of Physical Chemistry B*, 119(40):12838–12845, oct 2015. 24
- [72] Satoru Sasagawa, Yu-ichi Ozaki, Kazuhiro Fujita, y Shinya Kuroda. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nature cell biology*, 7(4):365–373, 2005. 24, 25, 77
- [73] Stéphane Pierre, Anne-Sophie Bats, y Xavier Coumoul. Understanding SOS (Son of Sevenless). *Biochemical Pharmacology*, 82(9):1049–1056, 2011. 24
- [74] Ido Amit, Ami Citri, Tal Shay, Yiling Lu, Menachem Katz, Fan Zhang, Gabi Tarcic, Doris Siwak, John Lahad, Jasmine Jacob-Hirsch, Ninette Amariglio, Nora Vaisman,

BIBLIOGRAFÍA

- Eran Segal, Gideon Rechavi, Uri Alon, Gordon B Mills, Eytan Domany, y Yosef Yarden. A module of negative feedback regulators defines growth factor signaling. *Nat Genet*, 39(4):503–512, apr 2007. 25
- [75] Claude E Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(July 1928):379–423, 1948. 26, 29
- [76] Alex Rhee, Raymond Cheong, y Andre Levchenko. The application of information theory to biochemical signaling systems. *Physical Biology*, 9(4):045011, 2012. 27, 32, 33
- [77] A Borst y F E Theunissen. Information theory and neural coding. *Nature neuroscience*, 2(11):947–57, 1999. 28
- [78] Gašper Tkačik, Curtis G Callan, y William Bialek. Information flow and optimization in transcriptional regulation. *Proceedings of the National Academy of Science of the United States of America*, 105(34):12265–12270, 2008. 28
- [79] G Tkačik, AM Aleksandra M Walczak, William Bialek, y Gasper Tkacik. Optimizing information flow in small genetic networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 80(3 Pt 1):031920, 2009. 28
- [80] Christian Waltermann y Edda Klipp. Information theory based approaches to cellular signaling. *Biochimica et biophysica acta*, 1810(10):924–932, 2011. 28, 30, 32
- [81] Peter S Swain, Michael B Elowitz, y Eric D Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12795–800, 2002. 28, 80
- [82] Johan Paulsson. Models of stochastic gene expression. *Physics of Life Reviews*, 2(2):157–175, 2005. 28

BIBLIOGRAFÍA

- [83] Werner Ebeling y Igor M. Sokolov. *Statistical Thermodynamics and Stochastic Theory of Nonlinear Systems Far from Equilibrium (Advanced Series in Statistical Mechanics: Volume 8)*. World Scientific, Humboldt-Universität, Germany, 2005. ISBN 978-981-02-1382-4. 29
- [84] Thomas M. Cover y Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, second edition, 2006. ISBN 9780471241959. 30, 31, 32, 35, 47, 53, 69
- [85] David J C MacKay. *Information Theory, Inference, and Learning Algorithms David J.C. MacKay*, volume 100. Cambridge University Press, Cambridge, England, 2005. ISBN 9780521642989. 32
- [86] H. B. Barlow. Possible Principles Underlying the Transformations of Sensory Messages. *Sensory Communication*, pages 216–234, 1961. 33
- [87] Michael S. Lewicki. Efficient coding of natural sounds. *Nature Neuroscience*, 5(4): 356–363, 2002. 33
- [88] Braden A.W. Brinkman, Alison I. Weber, Fred Rieke, y Eric Shea-Brown. How Do Efficient Coding Strategies Depend on Origins of Noise in Neural Circuits? *PLoS Computational Biology*, 12(10):1–34, 2016. 33
- [89] Toby Berger. Living Information Theory. *IEEE Inform. Theory Soc. Newsl.*, 53(1): 6–13, 2003. 33
- [90] Daniel T Gillespie. Exact Stochastic Simulation of couple chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977. 38, 43, 52
- [91] James E Ferrell. Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends in Biochemical Sciences*, 21(12):460–466, 1996. 42

BIBLIOGRAFÍA

- [92] C Y Huang y J E Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19):10078–10083, sep 1996. 42
- [93] Deborah K. Morrison. MAP kinase pathways. *Progress in Biophysics & Molecular Biology*, 4(71):479–500, 2012. 42
- [94] Vahid Shahrezaei, Julien F Ollivier, y Peter S Swain. Colored extrinsic fluctuations and stochastic gene expression. *Mol Syst Biol*, 4(196):196, 2008. 50
- [95] Morten Frödin y Steen Gammeltoft. Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. *Molecular and Cellular Endocrinology*, 151(1):65–77, 1999. 56
- [96] Michael L Simpson, Chris D Cox, y Gary S Saylor. Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *Journal of Theoretical Biology*, 229(3):383–394, aug 2004. 60
- [97] Nikoletta Szalóki, Jan Wolfgang Krieger, István Komáromi, Katalin Tóth, y György Vámosi. Evidence for Homodimerization of the c-Fos Transcription Factor in Live Cells Revealed by Fluorescence Microscopy and Computer Modeling. *Mol Cell Biol*, 35(August):3785–3798, 2015. 61
- [98] Etay Ziv, Ilya Nemenman, y Chris H. Wiggins. Optimal signal processing in small stochastic biochemical networks. *PLoS ONE*, 2(10), 2007. 68
- [99] Ioannis Lestas, Glenn Vinnicombe, y Johan Paulsson. Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467(7312):174–178, 2010. 69, 75, 85
- [100] Siobhan S Mc Mahon, Oleg Lenive, Sarah Filippi, y Michael P H Stumpf. Information processing by simple molecular motifs and susceptibility to noise. *Journal of the Royal Society, Interface / the Royal Society*, 12(110):0597, 2015. 75

BIBLIOGRAFÍA

- [101] Gašper Tkačik, Aleksandra M Walczak, Home Search, Collections Journals, About Contact, My Iopscience, y I P Address. Information transmission in genetic regulatory networks : a review. *Journal of physics. Condensed matter : an Institute of Physics journal*, 153102(15):153102, 2011.
- [102] J. O. Dubuis, G. Tkacik, E. F. Wieschaus, T. Gregor, y W. Bialek. Positional information, in bits. *Proceedings of the National Academy of Sciences*, 110(41):16301–16308, 2013.
- [103] Tamar Friedlander y Naama Brenner. Adaptive response and enlargement of dynamic range. *Mathematical Biosciences and Engineering*, 8(2):515–528, 2011. 75
- [104] Andre Levchenko y Ilya Nemenman. Cellular noise and information transmission. *Current Opinion in Biotechnology*, 28:156–164, 2014. 75
- [105] Clive G. Bowsher y Peter S. Swain. Environmental sensing, information transfer, and cellular decision-making. *Current Opinion in Biotechnology*, 28:149–155, 2014. 75
- [106] Shinsuke Uda, Takeshi H Saito, Takamasa Kudo, Toshiya Kokaji, Takaho Tsuchiya, Hiroyuki Kubota, Yasunori Komori, Yu-ichi Ozaki, y Shinya Kuroda. Robustness and compensation of information transmission of signaling pathways. *Science (New York, N.Y.)*, 341(6145):558–61, 2013. 75, 83
- [107] Margaritis Voliotis, Rebecca M Perrett, Chris McWilliams, Craig a McArdle, y Clive G Bowsher. Information transfer by leaky, heterogeneous, protein kinase signaling systems. *Proceedings of the National Academy of Sciences of the United States of America*, 111(3):E326–33, 2014. 75, 83, 85
- [108] Mikhail Tikhonov, Shawn C Little, y Thomas Gregor. Only accessible information is useful: insights from gradient-mediated patterning. *Royal Society open science*, 2(11):150486, 2015. 76

BIBLIOGRAFÍA

- [109] Tamar Friedlander, Roshan Prizak, Călin C. Guet, Nicholas H. Barton, y Gašper Tkačik. Intrinsic limits to gene regulation by global crosstalk. *Nature Communications*, 7:12307, 2016. 79
- [110] T D Gilmore. Introduction to NF- κ B: players, pathways, perspectives. *Oncogene*, 25(51):6680–6684, 2006. 81
- [111] M. D. Brennan, R. Cheong, y A. Levchenko. How Information Theory Handles Cell Signaling and Uncertainty. *Science*, 338(6105):334–335, 2012. 81
- [112] Mark Viney y Sarah E Reece. Adaptive noise. *Proceedings. Biological sciences / The Royal Society*, 280(1767):20131104, 2013. 83
- [113] Jangir Selimkhanov, Brooks Taylor, Jason Yao, Anna Pilko, John Albeck, Alexander Hoffmann, Lev Tsimring, y Roy Wollman. Accurate information transmission through dynamic biochemical signaling networks. *Science (New York, N.Y.)*, 346(6215):1370–3, 2014. 85