

Sistemas de Decisión Celular con Múltiples Posiciones

Germán A. Enciso Ruiz

Departamento de Matemáticas, Universidad de California, Irvine, USA

Correspondencia: enciso@uci.edu

La biología se ha convertido en una disciplina cada vez mas cuantitativa, especialmente con la disponibilidad en años recientes de grandes cantidades de información numérica experimental. Secuencias de ADN, concentraciones de proteínas en organismos vivos, niveles de expresión genética y muchas otras mediciones ahora son posibles de obtener a un costo cada vez mas bajo. Como resultado, muchos modelos matemáticos que antes eran considerados especulativos ahora se pueden validar y usar para hacer predicciones experimentales. En general, como podemos hacer mejores mediciones a nivel molecular, la biología celular ha dejado de ser una ciencia cuasi-descriptiva para acercarse cada vez mas a las ciencias exactas como la física. Es decir, no solamente podemos preguntar *que* pasa, sino también *como* pasa, en términos de reacciones químicas individuales.

DECISION CELULAR

Uno de los sistemas mas estudiados en biología celular es la comunicación de señales a nivel molecular. Las células reciben constantemente información del exterior y reaccionan en respuesta a estos estímulos. Por ejemplo, si una persona sufre una herida menor en la piel, una molécula mensajera llamada EGF (factor de crecimiento epidérmico) es emitida en la herida y viaja hacia células cercanas. EGF se adhiere a receptores sobre la membrana celular, causando una reacción química en cadena que finaliza con la división celular, todo esto para ayudar a cerrar la herida. En este sentido el input del sistema es EGF, y el output es la división celular. Si la concentración de EGF es baja, es importante que el sistema no reaccione y que no haya división - de otro modo cualquier pequeña concentración de EGF causaría división incontrolada, con un resultado equivalente al cáncer.

Este sistema presenta una situación interesante - el input es continuo, porque la concentración de EGF es arbitraria. Pero el output es binario, porque una célula no puede dividirse a medias. De manera que la célula esta transformando una señal continua en una respuesta

binaria, y esta llevando a cabo lo que podría llamarse una *decisión* en respuesta al input: dividirse o no dividirse. Podemos representar la relación input-output gráficamente (figura 1), donde vemos que la función de respuesta es sigmoidea, es decir cambia rápidamente de un valor mínimo a un valor máximo. El eje Y en esta grafica es la concentración de proteínas que dirigen la división celular propiamente, esta es la razón por la que no tiene exactamente valores binarios.

HEMOGLOBINA: SISTEMA DE DECISION MOLECULAR

Otro sistema distinto donde podemos observar elementos similares a nivel molecular es la hemoglobina (H). Esta importante proteína tiene la función de transportar oxígeno (O_2) en la sangre, y constituye alrededor del 95% de la materia seca de la sangre. La hemoglobina tiene la capacidad de adherir O_2 en los pulmones y soltarlo en los órganos periféricos como los brazos y las piernas. En principio H también podría hacer lo contrario y transportar O_2 de los brazos al pulmón, lo que obviamente seria contraproducente. Como veremos en un momento, H tiene la habilidad sorprendente de transportar O_2 principalmente en la dirección correcta.

En la figura 2 representamos la afinidad entre H y O_2 como función de la concentración ambiente de O_2 . Resulta que esta función, medida experimentalmente, es sigmoide. Cuando H se encuentra en el pulmón, la concentración ambiente de O_2 es alta y H captura a O_2 . Cuando H se encuentra en los órganos periféricos la concentración ambiente de O_2 es mas baja, y la afinidad entre las dos moléculas se reduce significativamente, lo que hace que O_2 se libere. Así, la forma sigmoide de esta función hace que el transporte de oxígeno sea mucho mas eficiente.

MÚLTIPLES POSICIONES

Que hace posible que la hemoglobina tenga esta respuesta sigmoide a la concentración de oxígeno? Hace ya cien años una propuesta fue hecha por A.V. Hill para explicar este fenómeno. La idea es que cada molécula de H no tiene solamente una, sino varias posiciones para adherirse al oxígeno. Ahora se conoce que la hemoglobina en efecto tiene $n = 4$ tales posiciones (figura 2, abajo). Hill

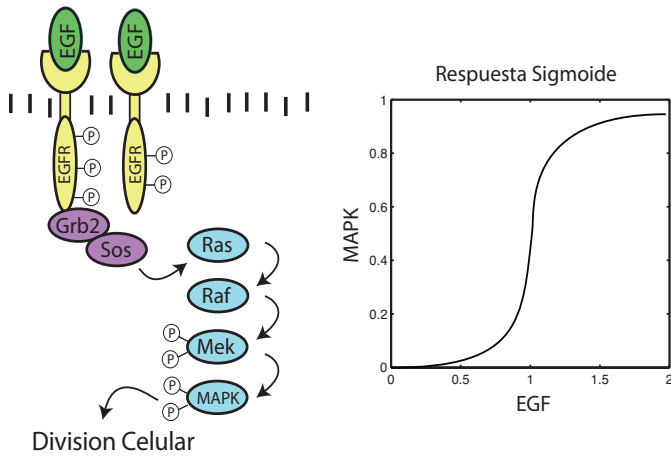
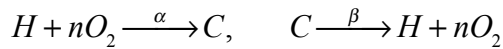


Figura 1 Esquema de la reacción generada por EGF al adherirse a la membrana celular, y ejemplo de una respuesta sigmoide

propuso la siguiente simple reacción que explicaría este fenómeno:



Aquí C representa el complejo formado por n moléculas de oxígeno y una de hemoglobina. Como la hemoglobina alterna entre las dos formas C y H sin cambiar su concentración total, tenemos la ecuación de conservación $H + C = H_{tot}$.

Usando el principio de acción de masa, también común en reacciones químicas, asumimos que la tasa a la que tiene lugar cada reacción es proporcional al producto de sus reactivos:

$$C' = \alpha H O_2^n - \beta C = \alpha (H_{tot} - C) O_2^n - \beta C.$$

En estado de equilibrio $C' = 0$, y podemos calcular la concentración del complejo C como

$$C = H_{tot} \frac{O_2^n}{\beta / \alpha + O_2^n}$$

La fracción C / H_{tot} corresponde a la proporción de hemoglobina adherida a oxígeno y representa una medida de la afinidad entre H y O_2 . Esta función es sigmoide, especialmente a medida que n aumenta. La función $x^n / (K + x^n)$ se conoce como *función de Hill* en honor a

este modelo, y es una de las funciones mas importantes de la biología matemática. El exponente n se conoce como *coeficiente de Hill*.

Este modelo es demasiado simplificado, en el sentido de que las reacciones químicas normalmente tienen lugar entre dos o tres moléculas en un momento dado, no $n + 1$ moléculas en un mismo instante. Sin embargo el modelo establece un principio clave que se observa en muchos otros sistemas: *la presencia de varias posiciones en una misma molécula facilita la construcción de funciones sigmoides y sistemas de decisión celular/molecular.*

MODELAJE USANDO LA FUNCION DE ACTIVACION

La idea de usar moléculas con varias posiciones para crear respuestas sigmoides aparece con gran frecuencia en la literatura. Las modificaciones en cuestión pueden corresponder a la adherencia de una molécula, pero también modificaciones covalentes como la fosforilación, acetilación, etc. Varios modelos se han propuesto para describir en mas detalle las reacciones de proteínas con múltiples posiciones, por ejemplo el modelo de Monod-Wyman-Changeux (1965) y el de Koshland-Nemethy-Filmer (1966). Generalmente se asume que la proteína es *cooperativa*, es decir, que después de modificar una posición, otras posiciones se pueden modificar mas fácilmente.

En trabajo reciente del autor (Enciso, Vargas, Kellogg, 2014) se muestra que aunque la cooperatividad de una proteína es útil, no es necesaria para obtener comportamiento sigmoide. Suponemos en su lugar que las n modificaciones son independientes entre si. En sistemas diferentes a la hemoglobina, la modificación de las posiciones tiene el efecto de cambiar la *actividad* de la proteína. Por ejemplo, si se trata de una enzima, es posible que solamente pueda llevar a cabo su función si esta suficientemente modificada.

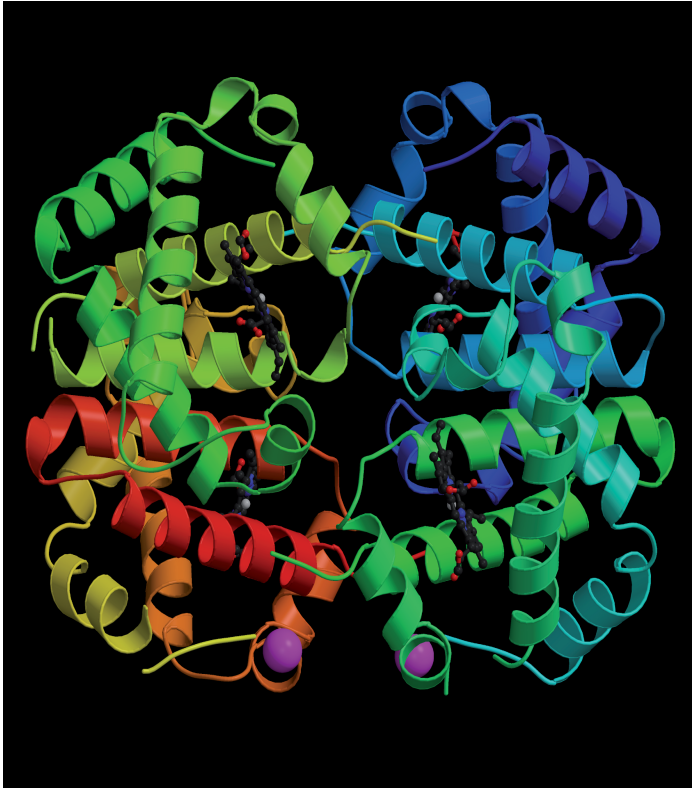
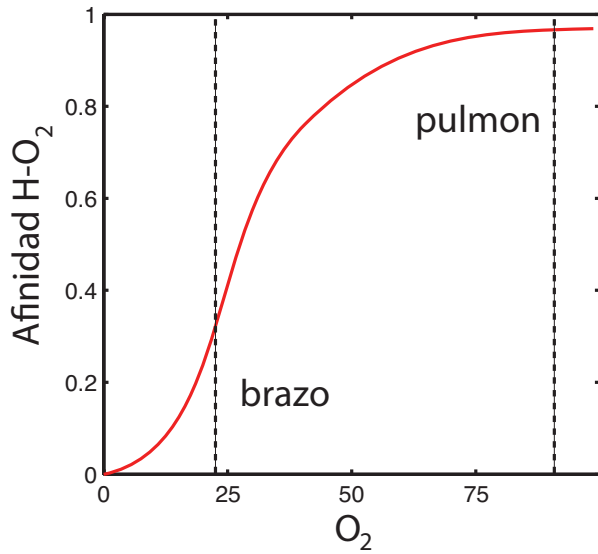


Figura 2: Afinidad estimada entre hemoglobina y oxígeno, medida experimentalmente. La hemoglobina consiste en cuatro copias de una misma estructura, y permite la adhesión de cuatro moléculas de oxígeno.

Fuente: <http://www.bio.davidson.edu/Courses/Molbio/MolStudents/spring2010/Hua/Hemoglobin.html>

Definimos el nivel de activación de una proteína como $a(x)$, donde x es la fracción de modificaciones en una molécula dada. Por ejemplo, en el caso de la hemoglobina se tiene $a(x) = x$, ya que la activación (i.e. el nivel de transporte de oxígeno) es proporcional a la fracción de posiciones ocupadas por O_2 . En el caso de proteínas con 5 o 10 posiciones de fosforilación, es posible que se requieran varias modificaciones antes que la proteína se comience a activar, y la función $a(x)$ misma podría llegar a ser sigmoide. Llamamos a $a(x)$ la *función de activación* del sistema.

Defina p como la fracción colectiva total de modificaciones en una muestra de la proteína. Suponga que I es uno de los 2^n estados posibles, con exactamente i modificaciones, y que P_I y P_{tot} es la concentración de proteína en estado I y la concentración total, respectivamente. Usando la condición de independencia, la probabilidad de que una proteína escogida al azar este en estado I se puede calcular de dos maneras distintas como $P_I / P_{tot} = p^i (1-p)^{n-i}$. Así podemos calcular la concentración P_i de proteína con exactamente i modificaciones:

$$P_i = \binom{n}{i} P_I = P_{tot} \binom{n}{i} p^i (1-p)^{n-i}.$$

Dado que la actividad de una proteína con i modificaciones esta dada por $a(i/n)$, se calcula la actividad total del sistema como

$$P = \sum_{i=0}^n a\left(\frac{i}{n}\right) P_i = P_{tot} \sum_{i=0}^n a\left(\frac{i}{n}\right) \binom{n}{i} p^i (1-p)^{n-i}.$$

Se puede demostrar que el lado derecho de la ecuación converge a la expresión $P_{tot} a(p)$ a medida que $n \rightarrow \infty$.

De esta manera se obtiene la aproximación $P \approx P_{tot} a(p)$, es decir

$$f(E) \approx P_{tot} a(p(E)) \quad (1)$$

donde E es el input que lleva a cabo las modificaciones (O_2 , enzimas, etc.), $p = p(E)$ es la fracción colectiva de posiciones modificadas dada la concentración de E ,

$a(x)$ es la función de activación, y $f(E)=P$ es la actividad total, es decir el output del sistema. Esta aproximación permite estimar la función de respuesta del sistema usando mediciones independientes, dado que $p(E)$, $a(x)$, P_{tot} se pueden medir experimentalmente.

Otra ventaja de esta ecuación se encuentra en su uso para el modelaje práctico de proteínas con varias posiciones. Un modelo biológico puede contener tres o cuatro proteínas con una docena de posiciones cada una. Es demasiado oneroso modelar todos los posibles estados de cada proteína separadamente en este contexto. Por otro lado, si se asume que cada proteína tiene solo dos estados, activo e inactivo, se está prácticamente asumiendo que $n=1$ y se elimina la complejidad aportada por el número de posiciones. Si en lugar de eso se usa la aproximación (1), se puede mantener un número limitado de variables (en este caso solo p, P) y al mismo tiempo hacer uso del número de posiciones. En particular, si la función $a(x)$ es sigmoide, es fácil ver que $f(E)$ también puede serlo.

QUIMIOTAXIS

Uno de los sistemas biológicos más estudiados es la capacidad que tienen muchos microorganismos para nadar en busca de nutrientes, llamada *quimiotaxis*. Bacterias como *E. coli* no tienen ojos ni oídos para encontrar comida, pero tienen lo que podría llamarse una excelente nariz: receptores en su membrana que detectan cambios minúsculos en la concentración del nutriente. Nadando tentativamente en una dirección al azar, si la bacteria siente que la concentración de este nutriente está aumentando, entonces continúa nadando en esa dirección; de lo contrario, la bacteria da vueltas al azar buscando una nueva dirección. El resultado es un algoritmo altamente eficiente para encontrar nutrientes y sobrevivir.

En el diagrama de la figura 3 (línea azul) se representa la sensibilidad de la 'nariz' responsable por la detección de nutrientes. Se trata de la concentración de CheA, proteína que representa el nivel de activación de los receptores, como función de la concentración de nutriente L . Esta función es sigmoide, aunque en este caso la concentración de CheA baja, en lugar de subir, a medida que L aumenta. Supongamos que la concentración promedio del

nutriente en un momento dado es L_0 . Se puede ver que si la concentración es un poco menor o mayor que L_0 , se genera un gran cambio en la activación de CheA. Esta sensibilidad es esencial para la quimiotaxis, porque los cambios relevantes de concentración pueden ser extremadamente pequeños.

La razón que esta nariz puede ser tan sensible a cambios en la concentración del input L es, como se podrá adivinar, consecuencia de la modificación en múltiples posiciones (figura 4). Los receptores capaces de adherir al nutriente están juntos unos a otros formando matrices hexagonales altamente regulares, esencialmente formando un solo objeto con un gran número n de posiciones de adherencia. Basados en esta información, mi estudiante Shane Ryerson y yo propusimos una función de activación $a_M(x)$ del receptor, y desarrollamos un modelo de este sistema.

¿Que pasa si la concentración promedio de L no es cercana a L_0 ? Esta pregunta es importante, porque en ese caso se pierde buena parte de la sensibilidad del sistema. Resulta que la bacteria es capaz de adaptarse a diferentes concentraciones a mediano plazo: si la concentración cambia consistentemente para estar cerca de L_1 en lugar de L_0 , la función de activación de CheA eventualmente cambia para detectar cambios pequeños alrededor de L_1 y así mantener la sensibilidad deseada.

OTRAS APLICACIONES

Una gran cantidad de sistemas en biología molecular dependen de respuestas sigmoideas para su función. Algunas veces esta función se puede interpretar como una decisión celular. Otros sistemas biológicos requieren funciones sigmoideas para crear comportamientos dinámicos complejos, como por ejemplo oscilaciones periódicas en la concentración de proteínas, o la presencia de varios estados de equilibrio en el sistema. En el trabajo (Enciso Vargas Kellogg 2014) se desarrolló una aplicación de las funciones de activación $a(x)$ en el contexto de la regulación celular en la levadura. En este modelo varias de las proteínas tienen múltiples posiciones, y su análisis fue posible gracias a la reducción en la ecuación (1).

A medida que se conoce más sobre sistemas específicos en biología molecular, se han descubierto principios que se aplican igualmente en sistemas totalmente diferentes. Estos principios de diseño de sistemas moleculares dan

orden a lo que de otro modo parecería inaccesible, y el uso de múltiples posiciones para crear funciones sigmoides es un ejemplo particular. Las matemáticas son especialmente útiles para estudiar principios de diseño porque permiten abstraer propiedades comunes y describir comportamientos posibles, independientemente del contexto.

German A Enciso Ruiz

German A. Enciso Ruiz completó sus estudios de pregrado en la Universidad de los Andes, Departamento de Matemáticas, en el año 2000. Bajo la supervisión del profesor Xavier Caicedo Ferrer, durante su tesis de pregrado estudio la fuerza de elección del teorema de Hahn-Banach en la ausencia del axioma de elección. En Estados Unidos, Enciso se enfocó en el estudio de las matemáticas aplicadas a la biología y obtuvo su doctorado bajo el profesor Eduardo Sontag en Rutgers University. Completo estudios postdoctorales en Ohio State University y Harvard Medical School, antes de trabajar en su posición actual como profesor investigador en California. Sus intereses son el estudio de la comunicación celular a nivel molecular, los sistemas dinámicos en biología, y el modelaje de sistemas biológicos con colaboradores experimentales.

FIGURAS

Figura 1: Esquema de la reacción generada por EGF al adherirse a la membrana celular, y ejemplo de una respuesta sigmoide.

Figura 2: Afinidad estimada entre hemoglobina y oxígeno, medida experimentalmente. La hemoglobina consiste en cuatro copias de una misma estructura, y permite la adhesión de cuatro moléculas de oxígeno.

Fuente:

<http://www.bio.davidson.edu/Courses/Molbio/MolStudents/spring2010/Hua/Hemoglobin.html>

Figura 3: La bacteria *E. coli* nada a través de sus flagelos, y busca comida usando receptores en su membrana como nariz. Diagrama de la activación de CheA como función de la concentración ambiental del nutriente.

Figura 4: Diagrama de las reacciones básicas en la quimiotaxis. Experimentos recientes muestran que los receptores están organizados en forma hexagonal (Briegel et al, 2014). Modelo matemático, función de activación

$a_M(x)$ usada, y solución del sistema como función de tiempo, bajo un aumento y reducción de la concentración de nutriente de 20% en $t=10$ y $t=50$, respectivamente (Ryerson y Enciso, 2014).

REFERENCIAS

U. Alon U, An Introduction to Systems Biology: Chapman and Hall/CRC, 2007

A. Briegel et al., Structure of bacterial cytoplasmic chemoreceptor arrays and implications for chemotactic signaling, eLife 2014(3):e02151:1-16, doi 10.7554/eLife.02151, 2014

G. Enciso, Multisite mechanisms for ultrasensitivity in signal transduction, en C. Poetschke and P. Kloeden (eds.), Nonautonomous and Random Dynamical Systems in Life Sciences, Lecture Notes in Mathematics 2102 (Mathematical Biosciences Subseries), Springer Verlag, 2013

G. Enciso, D.R. Kellogg, A. Vargas, Compact modeling of allosteric multisite proteins: Application to a cell size checkpoint, PLoS Computational Biology 10(2), 1-12, 2014

A. Hill, The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves, Proceedings of the Physiological Society 40(suppl.), iv-vii, 1910

D. Koshland, G. Nemethy, D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits, Biochemistry 5(1), 365-385, 1966

J. Monod, J. Wyman, J.P. Changeux On the nature of allosteric transitions: a plausible model, Journal of Molecular Biology 12: 88-118, 1965.

S. Ryerson, G. Enciso, Ultrasensitivity in independent multisite systems, Journal of Mathematical Biology 69(4):977-99, doi: 10.1007/s00285-013-0727-x, 2014.

Z. Serber, J. Ferrell, Tuning bulk electrostatics to regulate protein function, Cell 128, 441-444, 2007