

22 DE SEPTIEMBRE DE 2005

## La construcción de proteínas estaría regulada por simples reglas

Unos investigadores del Instituto Médico Howard Hughes han descubierto que toda la información necesaria para esculpir una proteína dándole forma y función apropiadas se encuentra en un conjunto relativamente simple de instrucciones codificadas en su secuencia de aminoácidos.

La comprobación se encuentra en dos artículos publicados en el número del 22 de septiembre de 2005, de la revista *Nature*, por investigadores del laboratorio de Rama Ranganathan, investigador del Instituto Médico Howard Hughes en el Centro Médico de la Universidad Texas Southwestern. Ranganathan y sus colegas primero utilizaron un algoritmo para identificar el patrón clave de interacciones entre los aminoácidos de una familia particular de proteínas. Luego, utilizaron este patrón para construir de cero una docena de proteínas sintéticas que reproducían la forma y la función de sus contrapartes naturales, al menos en los tubos de ensayo.

---

"El punto principal de estos artículos es que un conjunto bastante simple de reglas de secuencias es suficiente para especificar una familia de proteínas."

— Rama Ranganathan

---

Los científicos sabían que la secuencia de aminoácidos de una proteína le indica cómo plegarse y qué hacer. Pero los nuevos resultados podrían ayudar a explicar en qué lugar de la secuencia está la información necesaria y podrían ayudar a explicar la forma en la que una proteína puede tolerar simultáneamente muchas mutaciones sin sufrir daño aparente, pero dañarse debido a una sola mutación en ciertos sitios. Y el método sugiere que, si los biólogos siguen las mismas reglas simples que utiliza la evolución, la creación de proteínas complejas podría no ser tan complicada como parecía.

“Los estudios indican que el número de interacciones cruciales de una proteína podría ser más pequeño de lo que se pensaba antes -una ventaja para quienes deseen diseñar de cero proteínas nuevas para satisfacer una función específica-”, escribe Jeffery Kelley del Instituto de Investigación Scripps en un comentario que acompaña al artículo en *Nature*.

“Por muchos años, el diseño de secuencias artificiales que tengan la capacidad de plegarse en proteínas estables con funciones deseadas ha sido, en la ingeniería de proteínas, como la búsqueda del Santo Grial”, escriben Robert Smock y Lila Gierasch de la Universidad de Massachusetts, en Amherst, en un artículo *Preview* del número del 23 de septiembre de 2005, de la revista *Cell*. “Del punto de vista de la ingeniería de proteínas, la metodología es muy prometedora”.

Ranganathan no se proponía construir proteínas artificiales. Su interés consiste en aprender la forma en la que la naturaleza diseña las proteínas naturalmente a través del aleatorio proceso evolutivo de variación y selección. Pero la reconstrucción de una proteína era la mejor forma de evaluar una hipótesis evolutiva notablemente simple que él propuso basándose en datos genómicos.

“Será realmente importante estudiar si estas proteínas artificiales pueden funcionar en células vivas, embebidas dentro de las redes de señalización que fueron seleccionadas a través de la evolución”, dijo Ranganathan. “Pero el punto principal de estos artículos es que un conjunto bastante simple de reglas de secuencias es suficiente para especificar una familia de proteínas”.

Hace seis años, Ranganathan y sus colegas informaron sobre una forma de extraer los aminoácidos claves al observar el registro de productos exitosos de la evolución. Las proteínas modernas son el resultado de módulos que se intercambiaron entre las proteínas con el paso del tiempo. Estos módulos realizan funciones similares que consisten en unirse a dianas específicas, catalizar ciertas reacciones y enviar señales.

Ranganathan razonó que podría comparar una familia de módulos de varias especies para encontrar qué combinaciones de aminoácidos fueron seleccionadas evolutivamente durante muchos milenios después de que se ramificaran en el árbol de la vida. Para lograrlo, escribió un programa computacional que calcula las correlaciones que existen entre los aminoácidos de módulos similares que han pasado la prueba del tiempo.

Fue sorprendente que sólo algunos aminoácidos parecían estar firmemente restringidos y relacionados durante el tiempo evolutivo. Las posiciones altamente restringidas eran biológicamente importantes; la proteína se afectaba perceptiblemente cuando se mutaban esos pocos aminoácidos. Otras partes de la proteína eran diferentes en tanto que parecían haberse desarrollado independientemente unas de otras y casi no habían sido afectadas por mutaciones experimentales, dijo.

Ranganathan y sus colegas encontraron que, a veces, las partes funcionalmente importantes de una proteína plegada que se encuentran lejos del aparente centro de acción son detectadas por el algoritmo basado en la evolución.

“Una gran parte del trabajo demuestra que las proteínas tienen un patrón denso y local de interacciones estructurales entre aminoácidos”, dijo Ranganathan. “Pero en los datos evolutivos, vemos una arquitectura de poca

densidad pero distribuida de aminoácidos”. Esta nueva forma de examinar el problema podría contribuir a los trabajos de diseño de proteínas al proporcionar la arquitectura natural de las interacciones de aminoácidos. Actualmente, la mayoría de los algoritmos de diseño de proteínas intentan optimizar todas las relaciones atómicas de la estructura cristalina de una proteína, dijo.

Para estudiar las implicaciones globales -que no todos los aminoácidos de una proteína tienen la misma significancia para su estructura y función- el grupo de Ranganathan encontró las redes de aminoácidos críticas en una familia de proteínas y las utilizó para construir proteínas sintéticas funcionales de una secuencia de aminoácidos, que de otra forma sería aleatoria. Prometía ser un proyecto de intenso trabajo y que requeriría mucho tiempo. Al principio, nadie en su laboratorio quería intentar superar el desafío.

Para comenzar, Ranganathan seleccionó la familia del dominio WW, pequeño módulo utilizado por proteínas de muchas especies distintas para unirse a sitios dianas en otras proteínas y para mediar interacciones entre proteínas. Todos los dominios WW se pliegan de una forma simple que se asemeja a una mano ahuecada para sostener agua. En las personas, una mutación en el dominio WW obstaculiza la capacidad de una proteína llamada *nedd4* de unirse a un canal iónico y ponerlo donde se lo necesita en una célula renal, lo que puede llevar a una enfermedad seria llamada síndrome de Liddle, dijo.

En experimentos descritos en el primer artículo en *Nature*, Michael Socolich, técnico senior, y Steve Lockless, estudiante de doctorado, dirigieron el análisis de 120 dominios WW en organismos que iban desde la levadura hasta los seres humanos, utilizando todas las secuencias disponibles en las bases de datos genómicas de ese entonces. A partir de este análisis evolutivo, encontraron el patrón global de aminoácidos conservados y firmemente unidos.

A partir de allí, crearon cuatro bibliotecas que contenían los planos del dominio WW, sintetizaron todas las proteínas de las bibliotecas y probaron cuáles se plegaban correctamente.

La primera biblioteca contenía un surtido aleatorio de dominios WW naturales para probar el índice de plegamiento innato de su método experimental. Cerca de dos tercios de las proteínas se plegaron correctamente. La segunda biblioteca contenía secuencias totalmente aleatorias de WW. Ningunas de las proteínas resultantes se plegaron, según lo esperado.

En una tercera biblioteca, hicieron que la computadora preservara todos los aminoácidos que habían sido conservados con el paso del tiempo, pero retuvieron la información sobre qué aminoácidos necesitaban ser unidos. Ningunas de estas proteínas se plegaron correctamente. Pero en la cuarta biblioteca, que contenía las secuencias originadas en la computadora que preservaron unas pocas relaciones claves de aminoácidos por mucho tiempo,

cerca de un tercio de las proteínas se plegaron.

El primer artículo probó que algunos aminoácidos claves pueden cooperar para controlar la forma de una proteína. “Comienza a explicar la forma en la que una secuencia puede divergir y conservar la esencia de una estructura”, dijo Ranganathan. “Pero la naturaleza construye la función, no la estructura”.

En el segundo artículo de *Nature*, el estudiante postdoctoral William Russ y sus colegas ampliaron su análisis a 292 dominios WW de las bases de datos genómicas que ahora son cada vez mayores y probaron la habilidad de los dominios WW artificiales plegados de unirse al mismo espectro de péptidos que las secuencias naturales de WW.

Cuando el equipo examinó la red de aminoácidos que coevolucionaba del dominio WW, vieron que contenía residuos en la parte posterior de una molécula unidos al sitio de unión activo. Al estar tan lejos del sitio de unión activo, el residuo parece participar en la unión del péptido, según demostraron los experimentos.

En conjunto, los artículos muestran que es necesaria y suficiente la conservación más la información de unión para recrear los miembros de una familia específica de proteínas. En esta etapa, los investigadores no pueden distinguir las reglas para la estructura de las de la función, o si se diferencian. Ni conocen los mecanismos físicos subyacentes a las reglas estadísticas de las secuencias.

¿Así que, qué sucede con todos los aminoácidos excluidos de la red de coevolución? “Los considero como el motor para que la evolución avance”, dijo. “La independencia de las interacciones locales le dan a la proteína una capacidad de variación nueva. La proteína conserva una red central pequeña de interacciones claves que por sí mismas garantizan su función. Eso libera a las otras posiciones para que varíen independientemente. De esta forma, las proteínas creadas por mutación aleatoria se hacen fuertes a la perturbación aleatoria. Es bueno tener tolerancia hacia el proceso que te creó”.