

REVISIÓN DEL DOGMA ESTRUCTURA-FUNCIÓN EN LAS PROTEÍNAS

REVISION OF THE STRUCTURE-FUNCTION DOGMA IN PROTEINS

Concepción Hernández-Chinea

RESUMEN

El dogma central de la Bioquímica Estructural considera que la secuencia de una proteína determina su estructura tridimensional y esta a su vez establece su función, por lo que es posible deducir la función de una proteína conociendo su estructura. Nuevas evidencias cuestionan este dogma. Las proteínas conocidas como *Moonlighting* llevan a cabo dos o más funciones diferentes. Las interacciones proteína-proteína, la localización en distintos compartimientos celulares, la localización intra o extracelular o la expresión en tipos celulares o tejidos diferentes influyen notablemente en las funciones alternativas que esas proteínas pueden llevar a cabo sin cambios en su estructura tridimensional. Además, una elevada proporción de secuencias génicas en los organismos eucariotas parecen codificar largos segmentos de aminoácidos que no adoptan en solución una única estructura tridimensional de conformación conocida y predecible. Por sus características, esas proteínas se han denominado intrínsecamente desordenadas o no estructuradas. Las regiones desordenadas de ellas son altamente conservadas entre especies y contrariamente a la visión tradicional del dogma estructura-función, esas regiones carentes de estructura son generalmente funcionales y su estructura depende del ligando al cual se asocian para llevar a cabo su función. Considerando esas evidencias, la relación invariable entre la secuencia, la estructura y la función se cuestiona, haciendo necesario la reevaluación del dogma estructura-función en las proteínas.

ABSTRACT

The central dogma of Structural Biochemistry considers that the sequence of a protein determines its three-dimensional structure, which in turn determines its function, so it is possible to deduce the function of a protein by knowing its structure. New evidence questions this dogma. Proteins known as *Moonlighting* perform two or more different functions. The protein-protein interactions, the location in different cellular compartments, the intra- or extracellular localization or expression in different tissues or cell types significantly influence alternative functions that such proteins may perform without change in three dimensional structures. Additionally, a high proportion of gene sequences in eukaryotes appear to code for long segments of amino acids that do not adopt in solution a single three-dimensional structure of known and predictable shape. Due to such characteristics these proteins have been termed intrinsically disordered or unstructured. The disordered regions of these proteins are highly conserved among species and contrary to the traditional view imposed by the structure-function dogma; these regions without structure are usually functional. They adopt a particular structure depending on the ligand associate with to perform their function. Overall, these evidences question the invariable relationship between sequence, structure and function, a reassess-ment of the structure-function dogma in proteins is necessary.

Palabras Clave: Estructura-función proteica, proteínas moonlighting, dogma estructura-función.

Keywords: Protein structure-function, moonlighting proteins, Structure-function dogma.

INTRODUCCIÓN

Las proteínas son cadenas lineales de aminoácidos, unidos por enlaces peptídicos, que se pliegan formando estructuras tridimensionales complejas para dar origen a una forma tridimensional única. El primer paso en la adquisición de una estructura definida es el plegamiento de la cadena lineal de aminoácidos en dos formas: como una α -hélice o formando una estructura- β , que se estabilizan mediante puentes de hidrógeno. Las estructuras secundarias formadas interactúan mediante enlaces de tipo no covalente dando lugar a lo que se conoce como motivos, los cuales se pliegan tridimensionalmente entre sí para conformar los dominios estructurales. Estos dominios son conservados a lo largo de la evolución y las variaciones que se observan son producto de la adición de otros dominios o cambios en los segmentos interdominio. La forma tridimensional de una proteína es determinada por su estructura primaria. El orden de los aminoácidos en la cadena polipeptídica establece la estructura de la proteína y su función.

Desde principios del siglo pasado se hizo evidente que la función de una proteína depende de su estructura nativa. Los investigadores habían correlacionado claramente las propiedades específicas de las proteínas nativas con configuraciones moleculares específicas y la pérdida de las propiedades específicas que ocurre en la proteína desnaturalizada con los cambios que ocurrían en su configuración (Mirsky y Pauling, 1936).

Los experimentos de Christian Anfinsen con la Ribonucleasa A demostraron que la pérdida de la estructura tridimensional en las proteínas desnaturalizadas se revertía bajo condiciones apropiadas, restituyéndose la funcionabilidad de la proteína. Dado que en el estado desnaturalizado la única estructura que se mantiene en la proteína es la primaria, Anfinsen concluyó que la información para el plegamiento de las proteínas se encontraba en la cadena lineal de aminoácidos (Anfinsen y col., 1961).

Estos hallazgos condujeron a la proposición del dogma central de la biología estructural el cual sentencia que invariablemente, la secuencia de aminoácidos de una proteína determina su estructura tridimensional y esta a su vez determina su función (una secuencia-una estructura-una función).

PREDICCIÓN DE ESTRUCTURAS PROTEICAS Y FUNCIONALIDAD

El plegamiento depende de la secuencia de aminoácidos de un segmento particular de la cadena polipeptídica. Los estudios de Ramachandran en relación a la conformación que adoptan los polipéptidos dependiendo de su composición y secuencia de aminoácidos, llevaron a la publicación de unas gráficas, conocidas como gráficos de Ramachandran, que establecen cuáles de los 20 aminoácidos que constituyen las proteínas se encuentran con mayor frecuencia en alguna de las estructuras secundarias (Ramachandran y col., 1963). Con base en estos datos y mediante métodos computacionales se puede predecir la estructura secundaria que adoptará una determinada secuencia de aminoácidos. La formación de dominios o estructura terciaria no es posible predecirla y los métodos actuales de proteómica utilizan como base de predicción estructuras proteicas determinadas por cristalografía de rayos X o por resonancia magnética nuclear. Con las bases de datos de estructuras de proteínas se puede predecir con un aceptable grado de certeza, la estructura que adoptará una proteína dependiendo de su secuencia de aminoácidos.

La estructura generalmente está relacionada con una función, manteniéndose dominios definidos y conservados entre proteínas que llevan a cabo una misma función.

Además de los métodos físicos, genéticos y bioquímicos, se han desarrollado métodos *in silico* mediante programas de computación para predecir la función de las proteínas basándose en el análisis de la secuencia y estructura tridimensional de las proteínas. Los métodos de análi-

sis de secuencia predicen la función basados en la homología de secuencia con una proteína de función conocida o utilizando motivos estructurales que corresponden a una función. Sin embargo, la identificación de proteínas que sin cambiar su estructura tridimensional llevan a cabo varias funciones o la incapacidad para predecir una estructura en base a la secuencia de aminoácidos, han complicado la interpretación de los resultados obtenidos por los métodos *in silico* cuyo diseño parte del paradigma estructura-función.

Para estas proteínas, los programas que predicen la estructura utilizando las bases de datos de proteínas con estructura conocida por métodos físicos no aportan información valiosa para predecir todas las funciones en la que puede estar implicada la proteína.

Proteínas Moonlighting

El primer hallazgo que puso en tela de juicio el dogma estructura-función fue el descubrimiento paulatino y cada vez más frecuente, de proteínas que llevaban a cabo funciones diferentes dependiendo de su ubicación celular, de su oligomerización, de la concentración celular de un ligando, sustrato, cofactor o producto o inclusive del tipo celular en el cual se expresa. Debido a la doble función de estas proteínas y aludiendo a la semejanza con la práctica de los estadounidenses de dedicarse a más de un trabajo, práctica que se conoce con el nombre de *moonlighting*, Jeffery (1999) acuñó estas proteínas con el nombre de **Proteínas Moonlighting**. Ellas comprenden un grupo de proteínas multifuncionales en las cuales las diferentes funciones que pueden llevar a cabo se encuentran en una sola cadena polipeptídica que mantiene la misma estructura. Dentro de este grupo se excluyen aquellas proteínas que son multifuncionales debido a fusión de genes, proteínas que pertenecen a familias homólogas y variantes producto del *splicing* diferencial o de la acción de proteasas (Jeffery, 2009).

El primer ejemplo de proteínas multifuncionales se describió a finales de 1980 cuando se

reportó que algunas proteínas estructurales del cristalino del ojo de los vertebrados, eran enzimas conocidas por su participación catalítica en rutas metabólicas como lactato deshidrogenasa y enolasa (Piatigorsky y Wistow, 1989). Desde entonces, las proteínas *moonlighting* fueron creciendo en número llegando a ser un fenómeno general en organismos a todo lo largo de la escala evolutiva. Sin embargo, es mayor su presencia en células eucariotas.

Muchas de las proteínas *moonlighting* actualmente conocidas son enzimas muy conservadas en la evolución como las involucradas en el metabolismo de los azúcares. Por lo menos 7 de las 10 enzimas de la ruta glucolítica y 7 de las 8 involucradas en el ciclo de los ácidos tricarbónicos o ciclo de Krebs, se ha encontrado que poseen una función *moonlighting* (Huberts y van der Klei, 2010)

Se especula que el fenómeno de *moonlighting* podría explicar, el por qué los genomas codifican menos proteínas que las que se pueden predecir, creando un nivel de complejidad totalmente nuevo en la célula que rompe con el dogma estructura-función.

PROTEÍNAS INTRÍNECAMENTE NO ESTRUCTURADAS

Otro hallazgo que transforma los conceptos básicos de la biología estructural y el dogma estructura-función es la existencia de proteínas que no adoptan una estructura definida en largos segmentos de su secuencia de aminoácidos o que carecen totalmente de estructura cuando se encuentran en solución. La abundancia en la naturaleza de proteínas no plegadas o carentes de estructura tridimensional definida, principalmente en eucariotas, ha conducido al descubrimiento de un nuevo grupo de proteínas que no son exactamente desordenadas pero su funcionalidad depende de las zonas o regiones no estructuradas. Estas proteínas se conocen como proteínas intrínsecamente desordenadas o no estructuradas (Tompa, 2002).

Generalmente se asumía que los segmentos no estructurados de las proteínas no desempeñaban una función particular y que solo eran elementos conectores de las regiones funcionales bien estructuradas en las proteínas. Si la estructura está ligada fuertemente a una determinada función, qué sucede con las proteínas que son esencialmente no estructuradas? Paradójicamente la falta de estructura las capacita para llevar a cabo múltiples funciones ya que la cadena polipeptídica al no adoptar una estructura rígida puede fluctuar entre estados alternativos capaces de llevar a cabo una variedad de funciones, generalmente en el ámbito de la regulación y señalización celular (Wright y Dyson, 1999)

Básicamente, la composición de aminoácidos de las proteínas intrínsecamente no estructuradas es la que determina su condición desordenada (Tompa, 2005). Ellas se caracterizan por poseer aminoácidos con carga neta y una baja proporción de aminoácidos hidrofóbicos (Uversky y col., 2000). Basados en esas características se han diseñado diversos programas bioinformáticos (PONDR, DISOPRED, DisProt, DisEMBL, IUPred, entre otros) que predicen la incapacidad de una proteína para adoptar una estructura tridimensional en solución o el desorden estructural. Utilizando estos programas se ha encontrado que el desorden estructural es más prevalente en proteínas de organismos eucariotas (35-51%) que en bacterias (7-33%) o arqueas (9-37%) (Dunker y col., 2000). El hecho de encontrar en organismos multicelulares un mayor porcentaje de proteí-

nas desordenadas podría sugerir que esta condición es una adquisición evolutiva que les permite, con un número limitado de proteínas con largos segmentos desordenados en su cadena polipeptídica o completamente no estructuradas, llevar a cabo múltiples y diversos procesos de regulación y señalización celular. La conformación que adoptan transitoriamente estas proteínas viene determinada por las moléculas con las cuales interactúan y no por su secuencia aminoacídica como en las proteínas globulares, evidenciando la plasticidad de las mismas para asumir distintas funciones regulatorias.

CONCLUSIÓN

El dogma central de la biología estructural no es una verdad incontrovertible, por lo que su validez como dogma queda anulada a raíz de los hallazgos de las últimas décadas en cuanto a la relación estructura función en las proteínas. Si bien no hay dudas en el hecho de que la estructura de una proteína y su función están íntimamente ligadas, la visión actual indica que una determinada estructura es capaz de llevar a cabo más de una función y que no todas las proteínas biológicamente funcionales adoptan espontáneamente estructuras globulares estables. La carencia intrínseca de estructura o alto grado de desorden parece conferirle a una proteína la capacidad de unirse a diferentes ligandos para llevar a cabo diferentes funciones dentro de la célula.

LITERATURA CITADA

- ANFINSEN, C. B., E. HABER, M. SELA Y F. H. WHITE*
1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *PNAS*, 47: 1309-1314.
- DUNKER, A. K., Z. OBRADOVIC, P. ROMERO, E. C. GARNER Y C. J. BROWN*
2000. Intrinsic protein disorder in complete genomes. *Genome Informatic*, 11:161-171.
- HUBERTS, D. H. E I. J. VAN DER KLEI*
2010. Moonlighting proteins: An intriguing mode of multitasking. *Biochimica et Biophysica Acta*, 1803: 520-525.
- JEFFERY, C. J.*
1999. Moonlighting proteins. *Trends Biochem. Sci.*, 24: 8-11.
2009. Moonlighting proteins-an update. *Mol. BioSyst.*, 5: 345-350.
- MIRSKY, E. Y L. PAULIN*
1936. On the structure of native, denatured, and coagulated proteins. *PNAS*, 22: 439-447.
- PIATIGORSKY, J. Y G. J. WISTOW*
1989. Enzyme/crystallins: gene sharing as an evolutionary strategy. *Cell*, 57: 197-199.
- RAMACHANDRAN, G. N., C. RAMAKRISHNAN Y V. SASISEKHARAN*
1963. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, 7:95-99.
- TOMPA, P.*
2002. Intrinsically unstructured proteins *TIBS*, 27: 527-533.
2005. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Letters*, 579: 3346-3354.
- UVERSKY, V. N., J. R. GILLESPIE Y A. L. FINK*
2000. Why are natively unfolded proteins unstructured under physiologic conditions?. *Proteins*, 41, 415-427.
- WRIGHT P. E. Y H. J. DYSON*
1999. Intrinsically Unstructured Proteins: Re-assessing the Protein Structure-Function Paradigm. *J. Mol. Biol.*, 293: 321-331