

## Los posibles caminos desde el gen a la función proteica

Patricio O. Craig<sup>1,2</sup>, Sebastián Klinke<sup>3</sup> y Hernán R. Bonomi<sup>3</sup>

<sup>1</sup> Instituto de Química y Físicoquímica Biológica (IQUIFIB, UBA/CONICET)

<sup>2</sup> Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

<sup>3</sup> Instituto de Investigaciones Bioquímicas de Buenos Aires (IIBBA-CONICET), Fundación Instituto Leloir.

Buenos Aires, Argentina

hbonomi@leloir.org.ar, pocraig@qb.ffyb.uba.ar

Recibido 04/05/2015- Aceptado 30/06/2015

### Resumen

Las proteínas constituyen la maquinaria elemental de todos los seres vivos. Las instrucciones para su producción están escritas en los genes del DNA. Los avances en técnicas de secuenciación de DNA han permitido obtener una inmensa cantidad de información de los genes de una variedad de organismos tales como bacterias, hongos, virus, plantas, animales, etc. En la mayoría de los casos, la función de los genes descritos es desconocida. En este artículo discutimos los desafíos que involucra el análisis funcional de genes y proteínas concentrándonos en el caso particular de la enzima Lumazina Sintasa de *Brucella abortus* como ejemplo.

**Palabras clave:** Proteína, gen, predicción funcional, Lumazina Sintasa, homología

### The possible roads from gene to protein function

#### Abstract

Proteins are the basic components of the machinery of all living beings. The instructions for their production are written in the genes of the DNA. The advances in DNA sequencing techniques have enabled the acquisition of a large amount of information of the genes of a variety of organisms, like bacteria, fungi, virus, plants, animals, etc. In most cases the function of the described genes is unknown. In this article we discuss the challenges involved in the functional analysis of genes and proteins, focusing on the particular case of the enzyme Lumazine Synthase from *Brucella abortus* as example.

**Keywords:** Protein, gene, function prediction, Lumazine Synthase, homology

### **El principio de la función y la función del principio**

El estudio funcional de las proteínas es fundamental para la comprensión de la vida a nivel molecular. Al estudiar procesos biológicos, los científicos, buscamos aislar e identificar los componentes elementales y “principios activos” asociados con la observación de una actividad particular. Esta ruta de estudio desde la función al principio activo, con auge en la era pre-genómica, ha servido históricamente para la asignación de la actividad molecular de una amplia gama de proteínas y genes que las codifican.

Esta información primaria es en la que se apoyan métodos secundarios o indirectos de predicción de función, surgidos de la necesidad de asignar una función probable a proteínas y genes secuenciados en forma sistemática.

Se podría decir que en la era pre-genómica el número de proteínas con función conocida sobre el número de genes descritos era similar, mientras que en la era post-genómica, la secuenciación de genomas completos y metagenomas invirtieron la relación y ahora se sabe la secuencia de infinidad de genes de los cuales se desconoce completamente su función.

Existen varias estrategias para la asignación de función de genes y proteínas. La más utilizada es a través del análisis de la similitud de secuencia. Si uno tiene la secuencia de una proteína y quiere saber su función, lo primero que hace un científico en la actualidad es buscar con una computadora en bases de datos de secuencias la existencia de una proteína similar con función conocida. El razonamiento utilizado aquí tiene base evolutiva. Si existe un gen con una función particular en un organismo, lo más probable es que esa función se conserve en otro organismo codificada por otro gen muy similar en secuencia (homólogo).

Debido a la relación existente entre la estructura y la función de las proteínas, es también posible asignar la función de una proteína en base a la similitud de su estructura con la de una proteína de función conocida. La confianza de la predicción depende del grado de similitud observado. Sin embargo, existen casos de proteínas muy similares con funciones distintas y también proteínas con secuencias o estructuras dispares que tienen la misma función. La confirmación experimental es siempre necesaria en última instancia.

¿Pero cómo se aborda el análisis de la función de proteínas sin homólogos secuenciales o estructurales de función conocida? En estos casos el desafío es mayor, siendo necesario el análisis de una variedad de datos experimentales para elaborar hipótesis sobre su posible función celular y el diseño de experimentos específicos para su verificación. Esta información abarca la distribución del gen en distintos organismos, el análisis de su contexto genómico, la identidad de genes vecinos, la co-evolución de genes, el patrón de expresión en diversas condiciones celulares, las consecuencias metabólicas o morfológicas de su expresión o silenciamiento, la unión de ligandos, etc.

A pesar de la complejidad del problema, los avances producidos en el análisis funcional de las proteínas tanto desde la vía función-principio como la del principio-función son alentadores. El área posee un interés central y desafía las ideas e imaginación de los más creativos y audaces.

### **Una historia mínima sobre dos isoenzimas y la multiplicación de preguntas**

El Dr. Fernando Goldbaum obtiene su doctorado en el año 1992. Durante la realización de su tesis, el Dr. Goldbaum identifica a una proteína del patógeno *Brucella abortus* por su alta inmunogenicidad, la “pesca” usando los anticuerpos que ella misma genera durante la infección. Esta proteína tiene un peso molecular aparente de 18 kDa en un SDS-PAGE. La primera pregunta que el investigador se hizo fue: **¿qué función cumple?** La secuenciación N-terminal sirvió para identificarla como una secuencia homóloga a la de una enzima, la Lumazina Sintasa (LS) [3]. La LS es la enzima que sintetiza 6,7-dimetil-8-ribitilumazina (lumazina), el cual es el precursor directo de la vitamina B2 (riboflavina) [4]. Esta enzima se encuentra en plantas y microorganismos.

Por esto, la proteína de 18 kDa se renombró como Lumazina Sintasa de *Brucella* (BLS). Se caracterizó de forma biofísica durante bastante tiempo, descubriendo por ejemplo que poseía una organización decamérica, e inclusive se obtuvo su estructura cristalina, pero recién años más tarde se midió su actividad enzimática *in vitro*. Los valores de eficiencia catalítica (kcat/Km) mostraban que era una “pobre” LS. Recién en el año 2005 el genoma completo de *B. abortus* es publicado [7], revelando la existencia de otro gen homólogo a BLS. Surgió entonces la pregunta: **¿es ésta la “verdadera” LS de *Brucella*?** Por cuestiones de nomenclatura y convención, se las renombró como RibH1 a la nueva enzima encontrada *in silico* y RibH2 a BLS. Seguros de que RibH1 iba a poseer una mejor eficiencia catalítica, se midió y para sorpresa...¡era tan baja como la de RibH2! Cuando se resolvió la estructura cristalográfica de RibH1 en nuestro laboratorio [8], se encontró una organización cuaternaria diferente (pentamérica), aunque con el mismo plegamiento monomérico que RibH2.

Entonces nos hicimos nuevas preguntas: **¿son las dos LSs, o ninguna lo es y existe otra codificada en el genoma? Si no son LSs... entonces ¿qué función cumplen?**

Para tratar contestarlas, generamos bacterias *B. abortus* con sus dos genes ribH1 y ribH2 mutados. La primera observación es que esta bacteria no podía sobrevivir a no ser que nosotros le suministráramos de forma externa la riboflavina. También pudimos hacer que la bacteria mutante viviera, sin que le tuviéramos que agregar la vitamina en cuestión, luego de la introducción de un gen que codificaba para una **verdadera enzima LS** perteneciente a la levadura común de la cerveza, el hongo *Saccharomyces cerevisiae*, la cual ya estaba reportada y muy estudiada. Además, si la bacteria tenía una copia sana de ribH1 ó ribH2 podía vivir. Tomando en cuenta estos datos (realizando los controles adecuados), estamos en nuestro derecho de decir que tanto RibH1 como RibH2 poseen la función de LS *in vivo* [9]. De nuevo, aparecieron nuevas preguntas: **¿Por qué *Brucella* codifica para dos enzimas que realizan**

**una misma función (isoenzimas) si con una sola de ellas puede sobrevivir? ¿Cada enzima estuvo seleccionada para funcionar en momentos diferentes del ciclo de la bacteria? ¿Por qué poseen tan baja eficiencia catalítica comparada con las LSs de otros microorganismos?**

Después de algunos experimentos, la hipótesis actual es que RibH2 funcionaría para suministrar el precursor de la riboflavina cuando la bacteria esta infectando la célula del hospedador mientras que RibH1 lo estaría haciendo en las etapas no infectivas. El porqué de sus bajas eficiencia sigue siendo aún hoy una cuestión meramente especulativa.

### **Corolario**

En ciencia, aunque puede aplicarse como una regla general para la vida de los curiosos, la respuesta a una pregunta **siempre** lleva a muchas más preguntas.

### **Referencias**

1. **Goldbaum FA, Rubbi CP, Wallach JC, Miguel SE, Baldi PC, Fossati CA** (1992) Differentiation between active and inactive human brucellosis by measuring antiprotein humoral immune responses *Journal of Clinical Microbiology* 30: 604-6077.
2. **Goldbaum FA, Leoni J, Wallach JC, Fossati CA** (1993) Characterization of an 18-kilodalton *Brucella* cytoplasmic protein which appears to be a serological marker of active infection of both human and bovine brucellosis *Journal of Clinical Microbiology* 31: 2141-2145.
3. **Goldbaum FA, Velikovsky CA, Baldi PC, Mortl S, Bacher A, Fossati C A** (1999) The 18-kDa cytoplasmic protein of *Brucella* species --an antigen useful for diagnosis--is a lumazine synthase *Journal of Medical Microbiology* 48: 833-839.
4. **Fischer, M, Bacher A** (2005) Biosynthesis of flavocoenzymes *Natural Products Reports* 22: 324-350.
5. **Zylberman V, Craig P O, Klinke S, Braden BC, Cauerhff A, Goldbaum FA** (2004) High order quaternary arrangement confers increased structural stability to *Brucella* sp. lumazine synthase *Journal of Biological Chemistry* 279: 8093-8101.
6. **Klinke S, Zylberman V, Vega DR, Guimaraes BG, Braden BC, Goldbaum FA** (2005) Crystallographic studies on decameric *Brucella* spp. Lumazine synthase: a novel quaternary arrangement evolved for a new function? *Journal of Molecular Biology* 353: 124-137.
7. **Chain PS, Comerci, DJ, Tolmasky ME, Larimer FW, Malfatti SA, Vergez LM, Aguero F, Land ML, Ugalde RA, Garcia E** (2005) Whole-genome analyses of speciation events in pathogenic Brucellae. *Infection and Immunity* 73: 8353-8361.
8. **Klinke S, Zylberman V, Bonomi HR, Haase I, Guimaraes BG, Braden BC, Bacher A, Fischer M, Goldbaum FA** (2007) Structural and kinetic properties of lumazine synthase isoenzymes in the order Rhizobiales *Journal of Molecular Biology* 373: 664-680.
9. **Bonomi HR, Marchesini MI, Klinke S, Ugalde JE, Zylberman V, Ugalde R A, Comerci D J & Goldbaum F A** (2010) An atypical riboflavin pathway is essential for *Brucella abortus* virulence *PLoS One* 5, e9435.

**Los autores son investigadores de CONICET**



ISSN 1666-7948

[www.quimicaviva.qb.fcen.uba.ar](http://www.quimicaviva.qb.fcen.uba.ar)

Revista **QuímicaViva**

Número 2, año 14, Agosto 2015

[quimicaviva@qb.fcen.uba.ar](mailto:quimicaviva@qb.fcen.uba.ar)