

Análisis teórico de la regulación transcripcional y postranscripcional del operón *hyf* de *Escherichia coli* que tiene un regulador transcripcional codificado entre los genes estructurales

Planteamiento

Una de las características que define la expresión génica en bacterias es la co-transcripción de múltiples genes (en un operón) en un transcrito policistrónico, usualmente controlado bajo una misma secuencia reguladora única o promotor, esto debido a que por lo general participan en un mismo proceso biológico o los productos de estos forman complejos proteicos funcionales (Michael et al, 2014). En principio, los productos de estos operones mantienen una estequiometría formando los complejos funcionales necesarios para un proceso biológico.

Este concepto regulatorio es muy importante para la mayoría de los procesos biológicos en los que se necesitan cantidades equivalentes de los productos génicos de un mismo operón, pero ¿qué sucede en los casos en los que alguno de los productos de ese operón se requiere en una cantidad diferente a la producida como un solo mensajero? Esto sin duda es uno de los problemas más importantes en la regulación de la expresión génica derivada de operones, ya que, en su mayoría, la variación de la transcripción no siempre se ve reflejada en una diferencia drástica en los niveles de proteína. Para este problema, se han descrito algunos mecanismos que solucionan este problema.

Por un lado, un operón puede codificar un mRNA donde cada segmento puede ser traducido con una eficiencia diferente, como es el caso de la saturación diferencial de ribosomas unidos al mRNA, incrementando la traducción de uno de los genes dentro del operón, como es el caso de la síntesis del complejo de la F_1F_0 ATP sintasa, sistemas de toxina y antitoxina, factor sigma y antifactor sigma y los sistemas de dos componentes (Quax et al., 2013). Todos ellos han mostrado una mayor cantidad de proteína de al menos un componente. En el caso de los sistemas de dos componentes, la cinasa se encuentra en menor concentración que el regulador de respuesta, ya que estos factores transcripcionales suelen unirse al DNA como dímeros (Quax et al., 2013).

Otro mecanismo es la regulación por variación en cómo se encuentran reclutados los ribosomas en la secuencia de inicio de la traducción, mediante el incremento del inicio de la traducción por el reclutamiento de las subunidades ribosomales y un plegamiento o estructura secundaria de alta energía del transcrito (Ribosome Binding Site o RBS) (Quax et al., 2013). Así mismo, el trabajo de Quax y colaboradores demostró que aún dentro de los mismos operones, existe una diferencia en el uso de codones, haciendo que algunos genes dentro del mismo operón sean traducidos más o menos eficientemente. Este mecanismo no elimina otros mecanismos de regulación postranscripcionales como lo son la degradación de proteínas o bien la degradación diferencial de transcritos. De este último mecanismo, se tienen algunos ejemplos importantes.

El primero es en relación con la regulación del operón *lac*, el modelo más conocido de regulación genética en bacterias. Este operón está constituido por tres genes estructurales, uno codificante para la enzima β -galactosidasa (*lacZ*), la β -galactosido permeasa (*lacY*) y una β -galactosido transacetilasa, donde los genes *lacZ* y *lacY* son necesarios para el metabolismo de la lactosa (Griffiths et al., 1999). En este operón existe una secuencia tipo tallo-asa entre el gen *lacZ* y el gen *lacY*, el cual es cortado y permite la traducción independiente de ambos transcritos y con una proporción diferencial del mRNA de cada uno de ellos, lo cual determina una abundancia relativa diferencial de cada transcrito como producto final (Murakawa et al., 1991). Esto tiene como reflejo funcional, la abundancia relativa de cada transcrito y del producto enzimático, teniendo sentido biológico en el que, de los tres genes codificados en el mismo operón, la abundancia relativa de cada uno de los genes no es la misma (Wei et al., 2001) y funcionalmente, se requieren cantidades diferentes de cada una de las proteínas involucradas en el metabolismo de la lactosa. Este proceso es independiente de la traducción al demostrarse en un sistema *in vitro* (Chevrier-Miller et al., 1990). Si bien, en estudios con el operón *lac* no han identificado las endoribonucleasas que participan en el procesamiento del transcrito policistronico, hay otros operones en donde se ha demostrado que la RNasa E, la principal endoribonucleasa que inicia el proceso de degradación mediante la liberación de extremos 3' libres, permite el decaimiento diferencial del transcrito de algunos genes dentro de un operón, haciendo a su vez, más sensibles al procesamiento por otras endo y exoribonucleasas (Nilsson y Uhlin, 1991, Båga et al., 1988).

Recientemente, usando RNAseq, se demostró la importancia del decaimiento diferencial del mRNA en varios operones, donde se encuentran elementos que dictan estructuras secundarias que permiten el reconocimiento por endoribonucleasas que protegen el transcrito de degradación, que además se encuentran en las regiones codificantes y que por ende conlleva a la distribución diferencial de ribosomas, generando estequiometrías diferenciales entre las proteínas producidas de cada transcrito (Dar y Sorek, 2018). Este estudio tiene como limitante, el análisis de los operones altamente transcritos y que tienen relevancia metabólica y estequiometrías diferenciales. Adicionalmente, las endoribonucleasas en *E. coli* tienen cierta preferencia por secuencias ricas en A y U (secuencia consenso: [AIU]AUU[A/U]) (Arraiano et al., 1993), lo que permite el análisis fino de las secuencias que se piensan son procesadas.

Todo lo anterior indica que, en bacterias, los operones tienen niveles de regulación adicionales que son importantes para mantener un balance entre las diferentes proteínas que participan en un mismo proceso biológico. Sin embargo, una inspección cuidadosa de algunos operones mostró una característica peculiar, el regulador transcripcional del mismo operón está codificado dentro de los genes estructurales y no en otra posición, como es el caso del represor del operón *lac*, *lacI*. Esta característica no es muy extensiva en el genoma de *E. coli* y algunos

casos, el regulador transcripcional tiene su propio promotor o al menos una secuencia que haciendo un análisis computacional así lo parece.

Un operón que tiene esta propiedad y no tiene promotores internos predichos por métodos bioinformáticos, codifica para un complejo multi enzimático que ensambla una hidrogenasa, la cual en conjunto con la formato hidrogenliasa permite la deprotonación del formato en CO₂ e H₂ durante el crecimiento fermentativo de *E. coli* en pH ácido (Bagramyan et al., 2002). Esta hidrogenasa está codificada en un operón de doce genes, el penúltimo de ellos es un factor transcripcional que activa a este operón. Además, es transcripcionalmente activado en mutantes del regulador que controla los genes necesarios para el metabolismo del formato FhIA (Self et al., 2004), ya que su expresión en la cepa silvestre es muy bajo e inducido solamente en condiciones en las que hay formato y pH ácido en el medio. Análisis bioinformático de este operón indica que solamente tiene un promotor (Figura 1).



Figura 1. Estructura del operón *hyf* de *E. coli*. En la región promotora se muestran los elementos que se han demostrado experimentalmente y por análisis predictivo de regulación. En verde, se muestran los activadores (FhIA, HyfR y Crp) y en rosa el represor (Fnr). El penúltimo gene del operón codifica para el regulador HyfR. En evidencias experimentales, se ha demostrado que FhIA y HyfR compiten por el mismo sitio de activación (Skibinski et al., 2002).

De acuerdo con lo reportado por Skibinski y colaboradores (Skibinski et al., 2002), no hay un producto de amplificación por RT-PCR entre el gen *hyfJ* y *hyfR*. Sin embargo, análisis detallado de la región intergénica entre ambos genes indica que no hay un promotor interno y el análisis de expresión de esta zona del operón se hizo en condiciones basales (Skibinski et al., 2002). Entonces, ¿cómo se regula la cantidad de cada uno de los productos génicos de este operón?

Hipótesis

Los operones donde su propio regulador transcripcional está formando parte del policistrón que regula y carece de un promotor propio, su expresión o abundancia relativa es controlada por un mecanismo de decaimiento del mRNA mediado por la degradación mediada por RNAsas celulares.

Objetivo general

Analizar el mecanismo de regulación en genes que codifican para reguladores transcripcionales dentro de operones en *Escherichia coli*.

Objetivos particulares

- 1) Realizar un análisis *in silico* de la secuencia del operón *hyf* para identificar posibles promotores internos que hayan escapado al análisis global reportado en EcoCyc y otras bases de datos.
- 2) Identificar posibles sitios de procesamiento para la RNasa E en el operón que podrían estar involucrados en el procesamiento del operón y posiblemente en la abundancia relativa de cada elemento de la hidrogenasa.
- 3) Diseñar *in silico* las construcciones que se podrían hacer para evaluar los niveles de expresión del operón, así como, demostrar experimentalmente que no hay promotores internos.

Plan de trabajo

El proyecto involucra el trabajo en bases de datos, principalmente EcoCyc, GeneBank, Kegg, y el uso de un programa muy sencillo de uso, proporcionado por el profesor. El trabajo consiste en dividir en las semanas del verano en hacer primero análisis *in silico* y luego hacer una búsqueda en la literatura (Pubmed) para sustentar los hallazgos que se tengan. Por lo tanto, las semanas de trabajo serán distribuidas de la siguiente forma:

Semana del 22 al 26 de junio: familiarización con las bases de datos, obtener las secuencias del operón *hyf* e iniciar el análisis para buscar promotores internos en el operón.

Semana del 29 de junio al 3 de julio: Continuar con el análisis pendiente de los promotores internos, agrupar si se encuentran secuencias similares a los promotores de *E. coli* y fundamentar si tiene sentido biológico o no. Iniciar con la búsqueda de posibles sitios de reconocimiento para la RNasa E.

En la semana del 6 al 10 de julio, los sitios putativos de reconocimiento para la RNasa E serán analizados por estructura secundaria en el transcrito (mRNA), usando herramientas predictivas de estructura secundaria de RNA (mFold) para analizar si tiene el potencial de ser funcional *in vivo*.

Semana el 13 al 17 de julio: Obtener la secuencia de un plásmido adecuado, así como los componentes necesarios para generar un sistema reportero transcripcional para evaluar los posibles promotores internos. Como se menciona en el fundamento, este operón tiene algunas condiciones fisiológicas en las que se expresa, se buscaría analizar *in vivo* más adelante, si con el diseño de este sistema reportero, poder generar un sistema de análisis para validar (o descartar) los promotores internos identificados.

Semanas del 20 al 31: generar un tríptico informativo sobre la regulación genética de operones incluyendo los sistemas de regulación postranscripcionales para su

uso en el salón de clases a nivel de educación media superior. Generar el reporte de las actividades del verano.

Resultados esperados

Mediante el uso de herramientas bioinformáticas, se espera poder identificar secuencias putativas promotoras en la secuencia del operón *hyf*, las cuales podrían explicar el que el regulador transcripcional se encuentre en entre los genes estructurales. Si no se encuentra ninguno, entonces el proceso podría ser regulado por el reconocimiento de la endoribonucleasa RNasa E, que es la principal enzima que realiza cortes internos en el transcrito e inicia el proceso de decaimiento del transcrito *in vivo*, lo cual abre la puerta a iniciar el análisis más profundo, en este proyecto, se iniciaría el análisis de la estructura secundaria, lo cual implicaría que si estos sitios están accesibles, son más probables de estar involucrados en su procesamiento.

El diseño de un sistema reportero para este operón requiere de hacer una construcción con el posible promotor principal, el que se encuentra río arriba del primer gen, más las construcciones con las secuencias internas identificadas en el presente trabajo. En ese sentido, se planea hacer un diseño en el que se genere un bicistrón entre el regulador transcripcional y 10 aminoácidos iniciales del siguiente gen fusionados con un gen reportero o bien, al revés, primero el gen río arriba del regulador y 10 o 15 aminoácidos del regulador transcripcional, para evaluar *in vivo* si hay procesamiento, se vería afectada la abundancia del reportero.



Dr. Bernardo Franco Bárcenas

Profesor, Departamento de Biología

División de Ciencias Naturales y Exactas

Universidad de Guanajuato, campus Guanajuato

bfranco@ugto.mx

Referencias

Arraiano C, Yancey SD, Kushner SR. Identification of endonucleolytic cleavage sites involved in decay of *Escherichia coli* trxA mRNA. J Bacteriol. 1993 Feb;175(4):1043-52.

Båga M, Göransson M, Normark S, Uhlin BE. Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression. Cell. 1988 Jan 29;52(2):197-206.

- Bagramyan K, Mnatsakanyan N, Poladian A, Vassilian A, Trchounian A. The roles of hydrogenases 3 and 4, and the F₀F₁-ATPase, in H₂ production by *Escherichia coli* at alkaline and acidic pH. FEBS Lett. 2002 Apr 10;516(1-3):172-8.
- Chevrier-Miller M, Jacques N, Raibaud O, Dreyfus M. Transcription of single-copy hybrid lacZ genes by T7 RNA polymerase in *Escherichia coli*: mRNA synthesis and degradation can be uncoupled from translation Nucleic Acids Res. 1990 Oct 11;18(19):5787-92.
- Dar D, Sorek R. Extensive reshaping of bacterial operons by programmed mRNA decay. PLoS Genet. 2018 Apr 18;14(4):e1007354. doi: 10.1371/journal.pgen.1007354.
- Griffiths Anthony JF, Gelbart William M, Miller Jeffrey H, Lewontin Richard C. Regulation of the Lactose System. Modern Genetic Analysis. New York: W. H. Freeman;1999.
- Michael T, Madigan Martinko JM, Bender Kelly S, Buckley Daniel H, David A, Stahl TB. Brock Biology of Microorganisms (14va Edición). Pearson; 2014.
- Murakawa GJ, Kwan C, Yamashita J, Nierlich DP Transcription and decay of the lac messenger: role of an intergenic terminator. J Bacteriol. 1991 Jan;173(1):28-36.
- Nilsson P, Uhlin BE. Differential decay of a polycistronic *Escherichia coli* transcript is initiated by RNaseE-dependent endonucleolytic processing. Mol Microbiol. 1991 Jul;5(7):1791-9.
- Quax TEF, Wolf YI, Koehorst JJ, Wurtzel O, van der Oost R, Ran W, et al. Differential translation tunes uneven production of operon-encoded proteins. Cell Rep. 2013; 4: 938–44. <https://doi.org/10.1016/j.celrep.2013.07.049>.
- Self WT, Hasona A, Shanmugam KT. Expression and regulation of a silent operon, hyf, coding for hydrogenase 4 isoenzyme in *Escherichia coli*. J Bacteriol. 2004 Jan;186(2):580-7.
- Skibinski DA, Golby P, Chang YS, Sargent F, Hoffman R, Harper R, Guest JR, Attwood MM, Berks BC, Andrews SC. Regulation of the hydrogenase-4 operon of *Escherichia coli* by the sigma(54)-dependent transcriptional activators FhIA and HyfR. J Bacteriol. 2002 Dec;184(23):6642-53.
- Wei Y, Lee JM, Richmond C, Blattner FR, Rafalski JA, LaRossa RA. High-density microarray-mediated gene expression profiling of *Escherichia coli*. J Bacteriol. 2001 Jan;183(2):545-56.