

DOS APLICACIONES DE LA TÉCNICA DE ANÁLISIS DE FLUJOS METABÓLICOS

Ms. C. Osmán Fernández-Oliva¹, Dr. C. Julio C. Dustet-Mendoza², Dr. C. Ernesto Chico-Véliz¹



osman@cim.sld.cu

¹Centro de Inmunología Molecular, Ciudad de La Habana, Cuba; ²Facultad de Ingeniería Química, Instituto Superior Politécnico "José Antonio Echeverría", Ciudad de La Habana, Cuba

● Resumen

La Ingeniería Metabólica es la rama de las ciencias biológicas que se encarga del estudio y la evaluación de los cambios que se producen de forma dirigida en el metabolismo. Entre los objetivos que persigue esta metodología integradora está lograr el incremento del rendimiento del producto de interés hasta valores cercanos al máximo teórico. Disponer de un modelo matemático para describir el metabolismo de la célula objeto de estudio es la base para realizar trabajos de Ingeniería Metabólica. En este trabajo se aplica la modelación matemática del metabolismo al caso de estudio de la sobreproducción de lisina con la bacteria *Corynebacterium glutamicum*. Con la aplicación de la técnica de Análisis de Flujos Metabólicos y haciendo uso de la simulación matemática se determinó el máximo rendimiento teórico para la producción de lisina y con estas distribuciones de flujos se identificaron los nodos principales en la red metabólica. Se arriba a las mismas conclusiones que las obtenidas experimentalmente y reportadas en la literatura.

Palabras clave: flujo metabólico, rendimiento teórico, modelación matemática, metabolismo

● Abstract

Metabolic Engineering is the branches of biological science that deals with the study and evaluation of the changes that occurs in a targeted manner in metabolism. Among the objectives of this integrated methodology is to obtain the product's yield of interest to values near the theoretical maximum. A mathematical model to describe the cell metabolism under study is the basis for metabolic engineering. In this paper, mathematical modeling of metabolism is applied to lysine overproduction in *Corynebacterium glutamicum*. With metabolic flux analysis technique application and using mathematical simulation was determined the maximum theoretical yield for the production of lysine and with these flux distributions were identified the principal nodes in the metabolic network. We arrived to the same conclusions as those obtained experimentally and reported in the literature.

Keywords: metabolic flux, theoretical yield, mathematical modeling, metabolism.

● Introducción

Los procesos industriales para cultivar organismos productores de proteínas de interés médico tienen su origen hace más de medio siglo con la producción a gran escala de la penicilina. A partir del descubrimiento del ADN (Ácido Desoxirribonucleico) por Watson y Crick, como el material responsable de la transmisión de la información hereditaria y los avances en la

ingeniería genética para manipular el mismo e insertarlo en organismos vivos, la explosión de productos derivados de la biotecnología ha sido gigantesca. Cuando se calcula el rendimiento teórico de un producto en base a la fuente de carbono y energía por la estequiometría del proceso de fermentación o por la termodinámica del proceso, el valor que se obtiene generalmente es mucho mayor que el valor del rendimiento real que se alcanza en un proceso de

fermentación. Estrategias basadas en la aplicación de la Ingeniería Genética, así como la optimización del proceso como resultado de la aplicación de la metodología de la Ingeniería de Bioprocesos, logran incrementar de manera importante el rendimiento del producto. Entre las variantes usadas con éxito por la Ingeniería Genética se encuentra la incorporación de vectores de expresión con copias múltiples de casetes de expresión /1/.

Sin embargo, la sobreexpresión del producto que no es más que un rendimiento de expresión cercano al teórico, solo se logra con la alteración de algunas de las rutas del metabolismo. A pesar de haberse reconocido desde hace varias décadas la necesidad de conocer mejor el metabolismo de la célula y usar esa información para el diseño de la mejor estrategia de operación, siguen existiendo en la actualidad numerosos estudios basados en el tanteo y error para mejorar los rendimientos del producto de interés en los procesos de fermentación.

La ingeniería en una ruta metabólica para obtener un producto determinado comienza generalmente con el análisis de los pasos finales relacionados con el producto de interés. Sin embargo, tanto la velocidad de formación como el rendimiento están limitados en última instancia por la canalización del carbono desde el metabolismo central hacia la ruta de biosíntesis del producto /2/.

Entre las características de las rutas metabólicas centrales se tiene la existencia de puntos de ramificación o nodos hacia otras rutas de degradación o biosíntesis de otros metabolitos. Se trata entonces de conocer en qué nodo hay que hacer cambios de la regulación enzimática para favorecer el flujo de metabolitos por la ruta que tributa al producto de interés.

Entre las Ciencias Biológicas la Ingeniería Metabólica es la rama que se encarga del estudio y la evaluación de los cambios que se producen de forma dirigida en el metabolismo /3/. La modelación matemática del metabolismo celular es una fuente de generación de nuevos conocimientos, proporciona un ahorro considerable de tiempo y experimentos, y es la base para la realización de trabajos de Ingeniería Metabólica.

La modelación matemática del metabolismo es una tarea compleja debido al número grande de reacciones involucradas en los procesos, los sistemas de regulación de las células para responder ante los cambios del entorno y el número de rutas alternativas para responder ante los cambios. Los modelos matemáticos estequiométricos del metabolismo celular describen la red metabólica como un conjunto de ecuaciones estequiométricas que representan las reacciones bioquímicas que ocurren.

El objetivo de este trabajo es mostrar como mediante la simulación matemática de una red metabólica se pueden determinar los nodos principales de la misma y el rendimiento teórico del producto. Se toma como ejemplo de caso de estudio la sobreproducción de lisina en *Corynebacterium glutamicum* y se comparan los resultados teóricos con los obtenidos experimentalmente por /4/.

Fundamentación teórica

Clasificación de los nodos de una red del metabolismo

Los nodos son puntos en una red metabólica donde una secuencia de reacciones se bifurca entre dos o más rutas metabólicas diferentes /5/.

La regulación de una ruta metabólica puede ocurrir a diferentes niveles, el primero a través de la velocidad de reacción de cada una de las reacciones enzimáticas (dependerá del pH y de las concentraciones intracelulares del sustrato, del producto y del cofactor, que son los elementos primarios para la regulación de la actividad enzimática), el segundo nivel se efectúa a través de la acción de las enzimas reguladoras (muchas enzimas reguladoras son inhibidas por el producto final de la secuencia en la que actúan), el tercer nivel es a través del control genético de la velocidad de síntesis de la enzima, y en los organismos multicelulares que poseen sistemas endocrinos, la regulación metabólica se ejerce aún a otro nivel /6/ (las hormonas elaboradas por las glándulas endocrinas son mensajeros químicos que pasan por la sangre hasta ciertos tejidos que constituyen su objetivo y en los que estimulan o inhiben actividades metabólicas específicas).

La Ingeniería Metabólica se encarga de modificar la regulación de las rutas en los dos primeros niveles, haciendo uso de varios conceptos como rigidez de la red y nodos principales. Los conceptos de rigidez de una red, nodos flexibles y rígidos fueron planteados por /5/.

La sobreproducción de un metabolito particular se puede realizar mediante la desregulación de la ruta directamente asociada con la síntesis del mismo o por la transformación de una célula hospedera mediante la introducción de genes que codifican para la síntesis del producto de interés. Estas modificaciones sin embargo no necesariamente implican altos rendimientos del producto $\left(\frac{\text{moles de producto}}{\text{moles de sustrato}}\right)$, ya que las distribuciones de flujos metabólicos en los puntos ramales o nodos en el metabolismo primario (glucólisis, ciclo de los ácidos tricarboxílicos, ruta de los fosfatos de pentosa) se redireccionan para asegurar un crecimiento balanceado de la célula.

Esta resistencia inherente a la alteración de los flujos metabólicos se conoce como la rigidez de la red /5/. Aunque la actividad enzimática global a través de la rama para la obtención del producto de interés indudablemente gobierna su velocidad de formación, el rendimiento del producto es una función de las particiones o distribuciones de flujos que ocurren en los nodos de la red.

Consecuentemente para lograr la sobreproducción de un metabolito particular se requiere de alteraciones significativas de las distribuciones de flujos en los nodos del metabolismo primario comparadas con las distribuciones de flujos observadas para el crecimiento balanceado. No obstante el rendimiento del producto depende de la partición o distribución de flujos que ocurre solamente en un pequeño número de todos los nodos que contiene la red metabólica.

Estos son los llamados nodos principales y su identificación se realiza observando que en los mismos es donde ocurren los cambios más significativos en las distribuciones de flujos como función del rendimiento del producto deseado. La distribución de flujos requerida para el rendimiento es determinada por la técnica del Análisis de Flujos Metabólicos con o sin datos experimentales.

Cuando la identificación se realiza sin datos experimentales, se toman las distribuciones de flujos obtenidas de la simulación de los rendimientos del producto de interés.

Sobre la base de la rigidez de los nodos es posible clasificar las estructuras de control de la red metabólica en tres categorías. Para comprender estas clasificaciones es importante tener en cuenta que la cinética de la rama está referida a la respuesta local de todas las enzimas que actúan en la rama y que la razón de flujo de la rama está referida al flujo canalizado a través de la misma entre el flujo que entra al nodo.

Los nodos donde las particiones de flujos dentro de cada rama cambian fácilmente para satisfacer la demanda metabólica se conocen como nodos **flexibles**. En un nodo **flexible** las enzimas participantes en cada rama muestran una afinidad similar por el sustrato y las velocidades de reacción de cada rama son de igual magnitud.

El flujo a través de cada rama está controlado por una retroinhibición ejercida por el metabolito terminal correspondiente (figura 1).

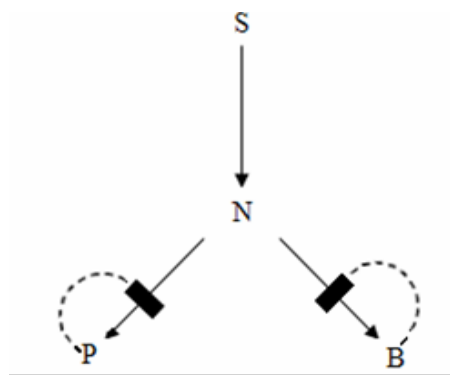


Fig. 1 Esquema de un nodo flexible, donde S es un metabolito intermedio, P el producto deseado y B otro producto que no es de interés. Las líneas discontinuas representan la retroinhibición.

De esta manera la razón de flujo de la rama puede variar entre cero y uno para satisfacer la demanda ubicada en los metabolitos terminales. Por ejemplo, si la retroinhibición ejercida por el metabolito P (figura 1) es eliminada o si P es un precursor de un producto deseado, la razón de flujo de la rama P se acercará a la unidad.

Los nodos **flexibles** son los más fáciles para alterar la distribución de flujos en las ramas y pocas veces requieren modificación.

Un nodo es **débilmente rígido** si la partición de flujos en el mismo está dominada por la cinética de una de sus ramas. Esto puede ser el resultado de una alta actividad enzimática o una alta afinidad enzimática por el sustrato, lo que hace que se necesite una retroinhibición en la rama dominante B (figuras 2 y 3).

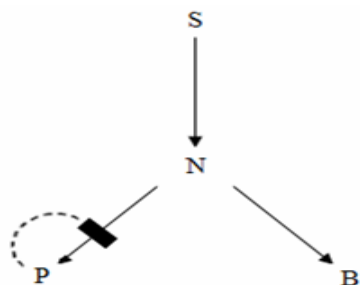


Fig. 2 Esquema de un nodo débilmente rígido, donde S es un metabolito intermedio, P el producto deseado y B otro producto que no es de interés. Las líneas discontinuas representan la retroinhibición. La formación de B ocurre en la rama dominante del nodo. La formación de P ocurre en una rama subordinada.

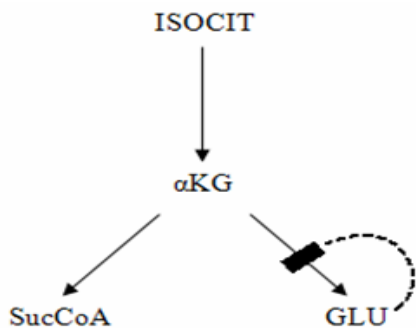


Fig. 3 Ejemplo del nodo α -oxoglutarato presente en el ciclo de los ácidos tricarboxílicos. Las líneas discontinuas representan la retroinhibición. La formación de succinil-coa es la rama dominante del nodo alfa-oxoglutarato.

Aun cuando la rama subordinada P esté desregulada (o sea que se elimine la retroinhibición) o P sea un precursor de un producto deseado, una fracción significativa del flujo seguirá entrando a la rama dominante y limitará el rendimiento del producto.

No obstante, el rendimiento del producto puede ser mejorado si se atenúa la actividad enzimática de la rama dominante B y se amplifica la actividad enzimática de la rama subordinada P lo suficiente para que canalice el incremento del flujo. Los nodos **débilmente rígidos** son generalmente tratados con técnicas de selección-mutación.

Un nodo es **fuertemente rígido** si la razón de flujos de una o más de sus ramas está controlada herméticamente. Esto ocurre comúnmente por una combinación de los controles por retroinhibición y trans-activación de la enzima por un metabolito en la rama opuesta (figuras 4 y 5).

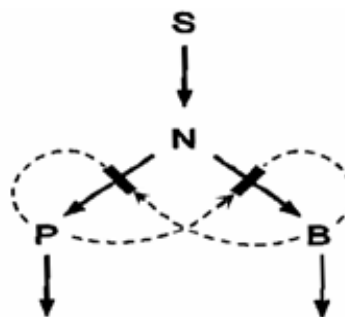


Fig. 4 Esquema de un nodo fuertemente rígido, donde S es un metabolito intermedio, P el producto deseado y B otro producto que no es de interés. Las líneas discontinuas representan la retroinhibición y las flechas discontinuas la activación por el metabolito terminal de la rama opuesta.

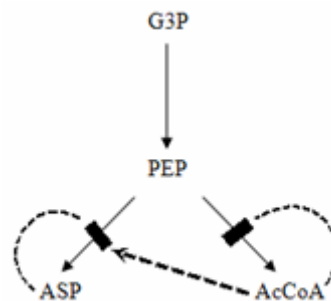


Fig. 5 Ejemplo del nodo Fosfoenolpiruvato presente en la Glucólisis. Las líneas discontinuas representan la retroinhibición y la flecha discontinua la activación por el metabolito terminal de la rama opuesta. El aspartato (ASP) actúa como inhibidor de la secuencia de reacciones en la rama en que se forma y a la vez como activador de la rama para la formación del Acetil-Coa.

En este caso particular además de inhibir su propia síntesis, los metabolitos P y B actúan con un efecto positivo en la rama opuesta. Esta arquitectura estabiliza las particiones de flujos en el nodo, proporcionando concentraciones en el estado estacionario de P y B lo suficientemente altas para inhibir su propia síntesis en ausencia de sus correspondientes activadores B y P respectivamente.

Por ejemplo, si el flujo a través de la rama B disminuye debido al incremento de la razón de flujos en la rama P, la velocidad de síntesis de P va a disminuir por la pérdida de activación por B y por el aumento de la inhibición debido a la concentración remanente de P, dejando la razón de flujos de cada rama prácticamente constante.

Sin embargo, las simulaciones indican que el incremento de la demanda de P, la eliminación de la retroinhibición de P y bloqueando la demanda de B tienen un pequeño efecto en la razón de flujos de la rama P.

La determinación de la rigidez de los nodos se realiza solamente por la vía experimental; y en función de esto último se proponen las modificaciones genéticas a realizar en la célula para lograr el objetivo de la sobreproducción del producto de interés.

Modelación matemática del metabolismo.

Análisis de flujos metabólicos

Los modelos matemáticos más usados para el metabolismo se pueden dividir en dos grupos: Los modelos cinéticos y los modelos estequiométricos o basados en restricciones. Los modelos cinéticos brindan información acerca de la dinámica con que ocurren las reacciones en la célula, pero no han prevalecido sobre los estequiométricos debido a la falta de información sobre las constantes cinéticas de las enzimas para las condiciones *in vivo* [7]. Como una red integrada de reacciones químicas, el metabolismo obedece a las leyes generales de la química como la ley de la conservación de la masa [8, 9].

Para plantear el modelo estequiométrico del metabolismo se toman en cuenta restricciones de balance de masa para cada metabolito, dadas por la estequiometría de las reacciones donde este participa, y se determinan los flujos metabólicos internos o

velocidades de reacción bajo la hipótesis de estado pseudoestacionario, justificada por el hecho de que los cambios metabólicos son muy rápidos (milisegundos a décimas de segundos) en comparación con el crecimiento celular (horas a días)/10/. El modelo que se obtiene contiene una matriz cuyos elementos son los números estequiométricos de los diferentes metabolitos en las distintas reacciones.

Entre las técnicas de análisis más importantes que se aplican en la Ingeniería Metabólica para la modelación matemática del metabolismo se encuentra la de Análisis de Flujos Metabólicos (MFA). Esta técnica se emplea con el fin de describir el estado fisiológico de la célula para las condiciones *in vivo* haciendo uso de mediciones de flujos externos en el medio de cultivo (que son el resultado del intercambio de nutrientes y productos entre la célula y el medio) para el cálculo de los flujos internos (velocidades de las reacciones de la red).

Si se toma como sistema a la célula, que intercambia materia con el medio exterior (medio de cultivo) se tiene para el metabolito *i* el siguiente balance de masa:

$$\frac{dN_i}{dt} = F_{i0} - F_i + \sum_{j=1}^{j=n} \alpha_{ij} \frac{d\varepsilon_j}{dt} \quad (1)$$

donde:

N_i es la cantidad de sustancia del metabolito *i*.

F_{i0} es el flujo de cantidad de sustancia del metabolito *i* que entra al sistema.

F_i es el flujo de cantidad de sustancia del metabolito *i* que sale del sistema.

α_{ij} es el número estequiométrico del metabolito *i* en la reacción *j*. Los valores α_{ij} de tienen signo positivo cuando el metabolito *i* es producto de la reacción y signo negativo cuando es reactivo en la reacción.

ε_j es el grado de avance de la reacción *j*.

$\frac{d\varepsilon_j}{dt}$ es la velocidad de la reacción *j* o flujo metabólico de la reacción *j*.

n es el número total de reacciones de la red metabólica.

Si para cada metabolito principal de la ruta metabólica se plantea un balance de masa, incluyendo a la biomasa y el producto de interés y asumiendo un estado pseudoestacionario se obtiene un sistema de ecuaciones lineales que puede ser representado de forma matricial de la siguiente manera:

$$Ax = q \quad (2)$$

donde:

A es una matriz de dimensión $m \cdot n$ que contiene los números estequiométricos de los metabolitos en las reacciones. Siendo m el número de metabolitos balanceados y n el número de reacciones que componen la red metabólica.

x es el vector de dimensión $n \cdot 1$ que contiene los flujos metabólicos o velocidades de reacción de todas las reacciones que componen la red.

q es el vector de dimensión $m \times 1$ que contiene las velocidades netas de salida desde la célula de todos los metabolitos balanceados.

La resolución del modelo se realiza a través del Método de los Mínimos Cuadrados. La ventaja fundamental de la técnica de MFA radica en que permite la determinación exacta de la magnitud de los flujos metabólicos (velocidades de reacción) de una ruta *in vivo* [11]. Los flujos metabólicos constituyen un determinante fundamental de la fisiología de la célula, ya que indican una medida del grado de ajuste de varias rutas metabólicas en las funciones celulares globales y procesos metabólicos.

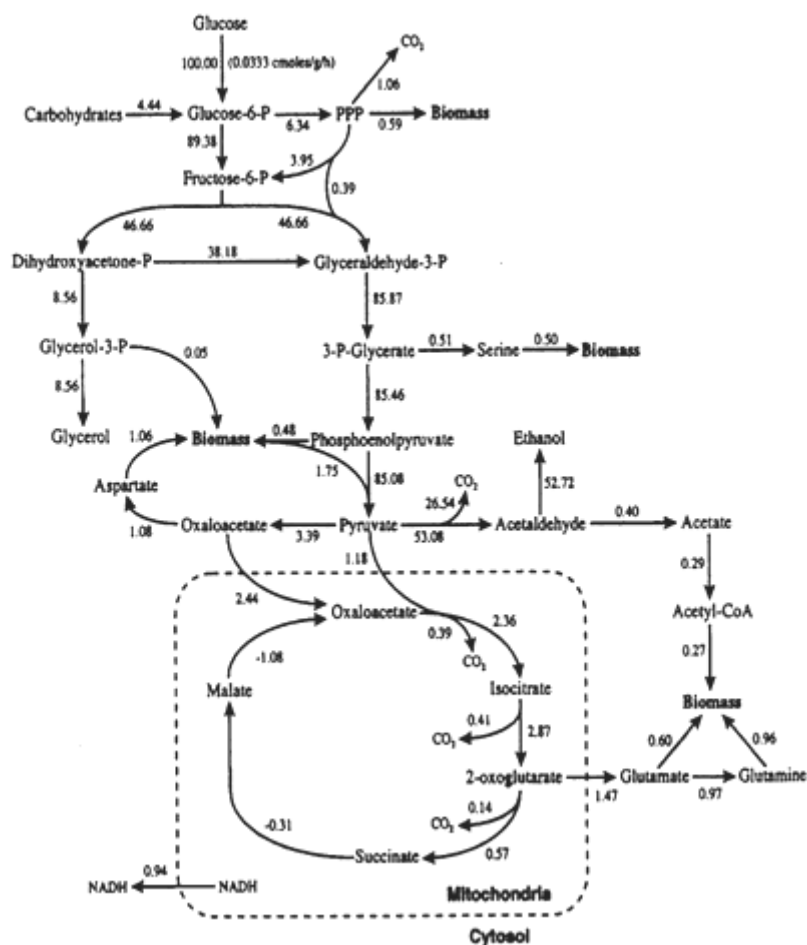


Fig. 6 Mapa de flujos metabólicos para *Saccharomyces cerevisiae* con crecimiento anaerobio. Tomado de [7].

El resultado final que se obtiene con la aplicación de la técnica de MFA es un mapa de flujos metabólicos mediante un diagrama de las reacciones bioquímicas y un estimado del flujo metabólico para cada reacción en el diagrama (figura 6).

Estos mapas de flujos metabólicos, contienen información útil acerca de la contribución de varias rutas a los procesos metabólicos globales de consumo de sustratos y formación de productos.

Como con MFA se caracteriza la red metabólica estudiada *in vivo* bajo diferentes condiciones ambientales, se puede entonces conocer mejor el funcionamiento dinámico de la célula.

Además de la obtención de los flujos metabólicos, el Análisis de Flujos Metabólicos puede proporcionar informaciones adicionales sobre otras características fisiológicas importantes de las células como la identificación de puntos de control (nodos rígidos) en las rutas celulares, el impacto que ha tenido una modificación genética realizada en la célula, la identificación de rutas alternativas para la formación del producto de interés, el cálculo de los flujos extracelulares no medidos experimentalmente, y el cálculo del máximo rendimiento teórico para el producto de interés /11/.

Materiales y métodos

Línea celular. El microorganismo estudiado es la *Corynebacterium glutamicum* ATCC 21253 /4/

Obtención de la matriz de los números estequiométricos

El modelo matemático se construyó siguiendo las reglas generales para aplicar la técnica de MFA descritas en /11/ y para establecer la ecuación para la formación de la biomasa se siguieron las reglas descritas en /12/.

Para poder comparar los resultados teóricos con los reportados por /4/ se tuvo en cuenta la información bioquímica reportada experimentalmente para esta línea celular.

La representación de la red metabólica de *Corynebacterium glutamicum* fue construida principalmente a partir de las redes que presentan las bacterias ácidos glutámicas, aunque pueden existir diferencias entre ellas /13/. En este caso las rutas principales a considerar son la glucólisis, la ruta de los fosfatos de pentosa, el ciclo de ácidos tricarboxílicos y sus rutas anapleróticas, la fosforilación oxidativa y la biosíntesis de glutamato, glutamina, alanina, valina y lisina.

También se toman en consideración las formaciones de trealosa y de ácido acético como productos colaterales del metabolismo central del carbono y la formación de biomasa.

Análisis de sensibilidad de la matriz de los números estequiométricos

El análisis de la matriz que representa una red bioquímica cualquiera se hace a través de los conceptos de Álgebra Lineal, lo que permite traducir los conceptos y conocimientos matemáticos a conceptos biológicos y viceversa /14/.

Es importante medir la sensibilidad o el grado de propagación de los errores de la matriz. El análisis de sensibilidad consiste en conocer si la matriz está bien-condicionada. Una matriz mal-condicionada con pequeñas variaciones en las mediciones experimentales provoca grandes cambios en los flujos metabólicos calculados /11/. Una medida de esta sensibilidad está dada por el número de condición, definido como:

$$CN = \|A\| \|A^T A\|^{-1} \|A^T\| \quad (3)$$

donde:

$\| \cdot \|$ representa alguna norma matricial.

A representa la matriz de los números estequiométricos.

A^T representa la matriz transpuesta de los números estequiométricos.

Si el número de condición se encuentra entre 0 y 100 se dice que la matriz está bien-condicionada, o sea que es poco sensible. Por el contrario si el número de condición es mayor

que 1 000 se dice que la matriz está mal-condicionada.

Método de los Mínimos Cuadrados

La solución de mínimos cuadrados v del sistema $Ax = q$ satisface la ecuación normal $A^T Ax = A^T q$ y recíprocamente cualquier solución de la ecuación normal es solución de mínimos cuadrados. Además, si $A^T A$ es invertible, la solución es única y viene dada por $v = (A^T A)^{-1} A^T q$. La matriz $(A^T A)^{-1} A^T$ es análoga con A^{-1} , y se denomina indistintamente inversa generalizada de A o matriz pseudoinversa de A . La condición necesaria para que $A^T A$ sea invertible es que el rango de A sea igual al número de columnas.

El rango de una matriz no es más que el número máximo de filas linealmente independientes, por lo que para cumplir la condición necesaria el número de balances de metabolitos linealmente independientes tiene que ser igual al número de reacciones de la red metabólica (número de incógnitas del sistema de ecuaciones).

Debido a que la mayoría de las células son capaces de utilizar una gran variedad de compuestos como fuentes de carbono, nitrógeno y energía; existen muchas rutas complementarias que operan al mismo tiempo y que incluidas en la matriz de los números estequiométricos crean una dependencia lineal responsable de la variación en el rango de la matriz /15/.

● Resultados y discusión

Selección del conjunto de reacciones que representan la red del metabolismo y el modelo matemático

Para obtener la matriz de los números estequiométricos para aplicar la técnica MFA se necesita hacer primeramente la selección de las rutas metabólicas principales que aportan la energía y los precursores requeridos para la formación de la biomasa, así como analizar las reacciones que forman la ruta biosintética del producto de interés.

En el caso de *Corynebacterium glutamicum* los análisis genéticos y bioquímicos indican que esta

bacteria posee al menos tres sistemas fosfotransferasas habilitados para el consumo de glucosa, fructosa y sacarosa /13/. Para el caso de estudio la glucosa es el único azúcar presente en el medio de cultivo, por lo que el sistema de la glucosa-fosfotransferasa es el único de los tres que pudiera estar activo.

J. Vallino, 1991 /13/ demostró también que aunque *Corynebacterium glutamicum* posee la enzima glucoquinasa, es el sistema catalizado por la glucosa-fosfotransferasa el que está involucrado en el transporte y la fosforilación de la glucosa extracelular, ya que la glucosa-fosfotransferasa presenta una mayor actividad que la otra enzima.

Entre las cuatro rutas anapleróticas sólo la fosfoenolpiruvatocarboxilasa ha sido detectada con una alta actividad en *Corynebacterium glutamicum* /13/. Por otra parte, aunque es indiscutible la existencia del ciclo del glioxilato en *Corynebacterium glutamicum*, su grado de inducción cuando hay glucosa en el medio de cultivo es bastante ambiguo (Vallino, J., 1991 /13/ determinó experimentalmente que la actividad de la enzima que inicia el ciclo isocitrato liasa es inhibida por la glucosa presente en el medio de cultivo).

Como la enzima es necesaria para el ciclo, si esta no tiene actividad no se tiene en cuenta en el modelo matemático las reacciones que dependen de la misma. En la tabla 1 aparecen las reacciones que componen el modelo matemático para el metabolismo de la bacteria en las condiciones de estudio.

Determinación de la sensibilidad de la matriz de los números estequiométricos

La matriz obtenida que representa la red metabólica (tabla 1) tiene una dimensión de $37 \cdot 34$ con un rango de 34 y un número de condición de 59 por lo que es funcional matemáticamente y poco sensible (el número de condición es menor que 100. Cumplida esta condición el modelo puede ser aplicado para la solución de los problemas de interés.

TABLA 1. REACCIONES DE LA RED DEL METABOLISMO PARA LA SOBREPDUCCIÓN DE LISINA POR *Corynebacterium glutamicum* QUE FORMAN EL MODELO MATEMÁTICO

Orden en la matriz	Reacción
1	GLC + PEP = PYR + G6P
2	G6P + 0.5 ATP = TRE + 0.5 ADP + 0.5 P _i
3	G6P = F6P
4	F6P + ATP = 2 GAP + ADP
5	GAP + P _i + NAD ⁺ + ADP = G3P + NADH + H ⁺ + ATP
6	G3P = PEP + H ₂ O
7	PEP + ADP = PYR + ATP
8	PYR + NADH + H ⁺ = LAC + NAD ⁺
9	PEP + H ₂ O + CO ₂ = OAA + P _i
10	PYR + CoA + NAD ⁺ = AcCoA + CO ₂ + NADH + H ⁺
11	AcCoA + OAA + H ₂ O = ISOCIT + CoA
12	ISOCIT + NAD(P) ⁺ = α-KG + CO ₂ + NAD(P)H + H ⁺
13	α-KG + CoA + NAD ⁺ = SucCoA + CO ₂ + NADH + H ⁺
14	SucCoA + GDP + P _i = SUC + CoA + GTP
15	SUC + FAD + H ₂ O = MAL + FADH ₂
16	MAL + NAD ⁺ = OAA + NADH + H ⁺
17	AcCoA + ADP + P _i = AC + ATP + CoA
18	α-KG + NH ₃ + NADPH + H ⁺ = GLU + NADP ⁺ + H ₂ O
19	GLU + NH ₃ + ATP = GLU + NADP ⁺ + H ₂ O
20	PYR + GLU = ALA + α-KG
21	2 PYR + NADPH + H ⁺ + GLU = VAL + CO ₂ + NADP ⁺ + H ₂ O + α-KG
22	G6P + H ₂ O + 2 NADP ⁺ = Ru5P + 2 NADPH + 2 H ⁺ + CO ₂
23	Ru5P = R5P
24	Ru5P = X5P
25	R5P + X5P = S7P + GAP
26	S7P + GAP = F6P + E4P
27	E4P + X5P = F6P + GAP
28	2 NADH + 2 H ⁺ + 4 ADP + 4 P _i + O ₂ = 2 H ₂ O + 4 ATP + 2 NAD ⁺
29	2 FADH ₂ + 2 ADP + 2 P _i + O ₂ = 2 H ₂ O + 2 ATP + 2 FAD
30	OAA + GLU = ASP + α-KG
31	PYR + 2 NADPH + 2 H ⁺ + GLU + ATP + ASP + SucCoA = LYS + CO ₂ + 2 NADP ⁺ + SUC + α-KG + CoA + ADP + P _i
32	LYS = LYS _e
33	0,09 R5P + 3,82 ATP + 0,025 GLN + 0,08 ASP + 0,312 NAD ⁺ + 0,054 ALA + 0,043 LEU + 0,033 LYS + 0,015 MET + 0,052 THR + 0,04 VAL + 0,332 AcCoA + 0,476 H ⁺ + 0,476 NADPH + 0,021 G6P + 0,007 F6P + 0,036 E4P + 0,013 GAP + 0,15 G3P + 0,052 PEP + 0,03 PYR + 0,446 GLU = BIOMASA + 3,82 ADP + 3,82 P _i + 0,143 CO ₂ + 0,312 NADH + 0,312 H ⁺ + 0,476 NADP ⁺ + 0,332 CoA + 0,364 α-KG
34	ATP = ADP + P _i

Aplicación 1: Determinación del máximo rendimiento teórico de lisina a partir de glucosa en *Corynebacterium glutamicum*

Un primer estimado del rendimiento máximo teórico para el producto de interés, se obtiene a partir de los balances de materiales por elementos (generalmente C, H, O, N) de la ecuación global de un proceso de fermentación en la que no se tiene en cuenta la formación de biomasa: $C_r H_g O_h + a NH_3 + b O_2 = z C_k H_r O_s N_t + c H_2 O + d CO_2$.

En este caso se tienen cinco números estequiométricos desconocidos (a, b, z, c, d) y

cuatro balances elementales, de manera que para darle solución al sistema es necesario añadir una ecuación más que sería la restricción \geq .

El máximo rendimiento teórico calculado por la estequiometría de la ecuación global de fermentación es de $0,857 \frac{\text{moles de lisina}}{\text{moles de glucosa}}$ o sea el 85,7 % de la glucosa es transformada a lisina. El problema principal asociado al cálculo del rendimiento de esta manera, es la alta probabilidad de que exista una distribución de flujos que no sea factible debido a que no se tienen en cuenta las restricciones bioquímicas de la célula.

Es por esto que se realiza el cálculo mediante la distribución de flujos teóricos. La distribución teórica de flujos es generada si se eliminan los balances de metabolitos que hacen a la matriz de los números estequiométricos sobredeterminada. Así, las velocidades para la biomasa (reacción 33 en las tablas 1 y 2), la glucosa (reacción 1 en las tablas 1 y 2) y el producto de interés son fijadas en 0, -100 y Y_p respectivamente, donde Y_p es el rendimiento molar del producto (mol lisina/mol glucosa en la tabla 2), y se va simulando su valor hasta que se tenga una distribución de flujos que no sea factible. Es en este

punto donde se obtiene el máximo rendimiento teórico. De esta manera, se eliminan los balances de CO_2 , NH_3 y O_2 de la matriz de los números estequiométricos, lo que provoca un aumento del número de condición a 881. Este valor es aceptable (el número de condición es menor que 1 000) teniendo en cuenta que se encuentra en el intervalo establecido y especialmente si se considera que no existe incertidumbre en el valor del vector de las mediciones. En la tabla 2 se observan los resultados de los flujos metabólicos de las diferentes reacciones obtenidos para los diferentes valores de rendimiento.

TABLA 2. DISTRIBUCIÓN DE FLUJOS TEÓRICOS PARA DIFERENTES VALORES DE RENDIMIENTO DE LISINA A PARTIR DE GLUCOSA EN *Corynebacterium glutamicum*. LOS FLUJOS ESTÁN EXPRESADOS EN MMOL/Lh

Reacción	Rendimiento de Lisina Y_p (mol Lisina/mol glucosa)					
	0%	20%	35%	60%	64%	75%
	Flujo metabólico de la reacción (mmol/L h)					
1	100	100	100	100	100	100
2	0	0	0	0	0	0
3	220	148	94	4	-9	-50
4	140	116	98	68	64	50
5	240	216	198	168	164	150
6	240	216	198	168	164	150
7	140	96	63	8	0	25
8	0	0	0	0	0	0
9	0	20	35	60	64	75
10	240	176	128	48	36	0
11	240	176	128	48	36	0
12	240	176	128	48	36	0
13	240	176	128	48	36	0
14	240	156	93	-12	-27	-75
15	240	176	128	48	36	0
16	240	176	128	48	36	0
17	0	0	0	0	0	0
18	0	40	70	120	127	150
19	0	0	0	0	0	0
20	0	0	0	0	0	0
21	0	0	0	0	0	0
22	-120	-48	6	96	109	150
23	-40	-16	2	32	36	50
24	-80	-32	4	64	73	100
25	-40	-16	2	32	36	50
26	-40	-16	2	32	36	50
28	480	372	291	156	136	75
29	120	88	64	24	18	0
30	0	20	35	60	64	75
31	0	20	35	60	64	75
32	0	20	35	60	64	75
33	0	0	0	0	0	0
34	2640	1996	1513	708	591	225

Como puede observarse, la primera distribución de flujos teóricos que no es factible ocurre a un valor de rendimiento de lisina de 64 %. En esta situación el flujo metabólico de la reacción 7 alcanza el valor de cero, y de seguir aumentando el rendimiento implicaría valores de flujos negativos. Esto no sería posible debido a que dicha reacción es irreversible.

Vallino, J., 1991/13/ demostró experimentalmente que en la bacteria *Corynebacterium glutamicum* existe actividad de la enzima PEP-sintetasa que cataliza la siguiente reacción: $\text{PYR} + \text{ATP} + \text{H}_2\text{O} = \text{PEP} + \text{AMP} + \text{P}_i$. Por tanto, para rendimientos mayores que 64 % esta reacción sustituye a la reacción 7, de manera que los flujos negativos con la nueva estequiometría den positivos, indicando que esta última reacción es la que lleva a cabo la formación del fosfoenolpiruvato.

La próxima distribución de flujos que no es factible ocurre para un valor de rendimiento de lisina de 75 %, en donde aquellas reacciones del ciclo de los ácidos tricarbónicos (reacciones 11 a 16 en las tablas 1 y 2) que son irreversibles alcanzan el valor de cero (reacción 11), además para ese rendimiento la reacción de formación de Acetil-CoA a partir de Piruvato que también es irreversible, su flujo metabólico toma el valor de cero.

Por tanto, bajo estas condiciones se alcanza un máximo de rendimiento teórico de lisina de 75 %. Este valor, aunque es menor que el calculado por el balance de materiales por elementos de la ecuación estequiométrica que representa el proceso de fermentación, es más confiable debido a que tiene en cuenta las restricciones bioquímicas. Así se puede

concluir que el concepto de máximo rendimiento teórico no es tan concreto como se puede pensar. Para tener precisión, el máximo rendimiento teórico reportado debe siempre estar acompañado por la bioquímica usada para determinarlo y la limitación en que se incurrió, como se hizo en este caso.

Aplicación 2: Identificación de los nodos principales de la red metabólica para la sobreproducción de lisina

Para la identificación de los nodos principales se utilizan los valores de flujos obtenidos en la simulación para conocer el rendimiento teórico máximo. De la figura 7 se puede observar que los nodos principales de la red metabólica son glucosa-6-fosfato, piruvato y fosfoenolpiruvato por mostrar los mayores cambios en su razón de flujos metabólicos en función del rendimiento de lisina.

Los cambios en el nodo oxalacetato se deben a los del nodo fosfoenolpiruvato, por lo que este tipo de nodo se conoce como principal trivial. Estos resultados (identificación de los nodos principales) coinciden con los reportados en un trabajo experimental /13/. Se demuestra entonces que la simulación matemática es una vía posible para la identificación de los nodos principales de una red metabólica.

Una vez identificados los nodos principales debe determinarse por la vía experimental el tipo de rigidez de cada uno; y en función de esto último se proponen las modificaciones genéticas a realizar en la célula para lograr el objetivo de la sobreproducción del producto de interés. Esta es la esencia de la Ingeniería Metabólica.

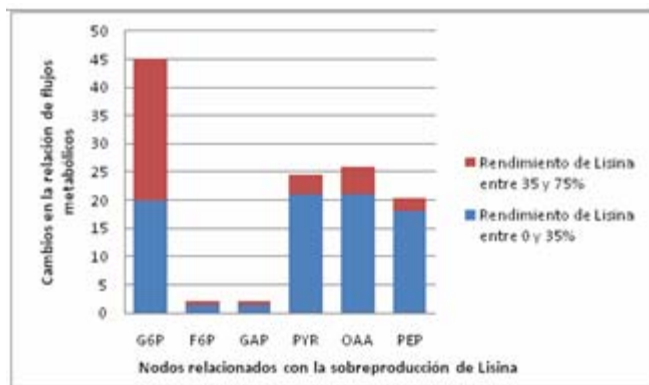


Fig. 7 Cambios en la relación de flujos metabólicos en los nodos presentes en la red metabólica para la sobreproducción de lisina.



Conclusiones

La metodología de Análisis de Flujos Metabólicos permite estimar el máximo rendimiento teórico de un producto y a partir de las distribuciones de flujo determinar los nodos principales de la red metabólica mediante la simulación del funcionamiento de la red con el modelo matemático que la describe. Esto se demuestra en el caso de estudio de la sobreproducción de lisina en Corynebacterium glutamicum a partir de glucosa.

La importancia del uso de la simulación radica en el ahorro de tiempo y además permite dirigir mejor los experimentos a realizar para la determinación de la rigidez de los nodos identificados como principales.



Bibliografía

- MANSUR, M. "Sistema de obtención de insulina humana a partir de *Pichia pastoris*". Tesis de Doctorado en Ciencias Biológicas. Centro de Ingeniería Genética y Biotecnología. Universidad de La Habana. Ciudad de La Habana, 2007.
- LIAO, J. C.; S. HOU; Y. CHAI. "Pathway Analysis, Engineering, and Physiological Considerations for Redirecting Central Metabolism". *Biotechnology and Bioengineering*. Vol. 52, pp. 129-140, 1996.
- SHIMIZU, H. "Metabolic Engineering-Integrating Methodologies of Molecular Breeding and Bioprocess Systems Engineering". *Journal of Bioscience and Bioengineering*. Vol. 94, 563-573, 2002.
- VALLINO, J.; G. STEPHANOPOULOS. "Metabolic Flux Distribution in *Corynebacterium glutamicum* During Growth and Lysine Overproduction". *Biotechnology and Bioengineering*. Vol. 41, pp. 633-646, 1993.
- STEPHANOPOULOS, G.; J. VALLINO. "Network Rigidity and Metabolic Engineering in Metabolite Overproduction". *Science*. Vol. 252, pp. 1675-1681, 1991.
- LEHNINGER, A. *Principios de Bioquímica*. Omega, 2005.
- PATIL, K. R.; M. AKESSON; J. NIELSEN. "Use of Genome-Scale Microbial Models for Metabolic Engineering". *Current Opinion in Biotechnology*. Vol. 15, 64-69, 2004.
- CORTASSA, S.; A. IGLESIAS; D. LLOYD. *An Introduction to Metabolic and Cellular Engineering*. World Scientific, 2002.
- KAUFFMAN, K. J.; P. PRAKASH; J. S. EDWARDS. "Advances in Flux Balance Analysis". *Biochemical Engineering*. Vol. 14, 491-496, 2003.
- FELL, D. "Metabolic Control Analysis: a Survey of its Theoretical and Experimental Development". *Biochemical Journal*, Vol. 286, pp. 313-330, 1992.
- STEPHANOPOULOS, G.; A. ARISTIDOU; J. NIELSEN. *Metabolic Engineering: Principles and Methodologies*. Academic Press, 1998.
- ZUPKE, C.; G. STEPHANOPOULOS. Intracellular Flux Analysis in Hybridomas Using Mass Balances and *in vitro* ¹³C NMR, *Biotechnology and Bioengineering*, Vol. 45, pp. 292-303, 1995.
- VALLINO, J. "Identification of Branch-Point Restrictions in Microbial Metabolism Through Metabolic Flux Analysis and Local Network Perturbations". PhD Thesis, Massachusetts Institute of Technology, 1991.
- SCHILLING, C.; S. SCHUSTER; B. PALSSON. "Metabolic Pathway Analysis: Basic Concepts and Scientific Applications in the Post-genomic Era". *Biotechnology Progress*. Vol. 15, pp. 296-303, 1999.
- FERNÁNDEZ, O. "Modelación matemática del metabolismo, una base para trabajos de Ingeniería Metabólica. Aplicación a casos de estudio". Tesis de Diploma. Instituto Superior Politécnico "José Antonio Echeverría". Ciudad de La Habana, 2008.

Nomenclaturas, símbolos y letras griegas

- A: Adenosina
- AC: Ácido Acético
- AcCoA: Acetil-CoA
- ACE: Acetaldehído
- ADP: Adenosina- Difosfato
- ALA: Alanina
- AMP: Adenosina-Monofosfato
- ARG: Arginina
- SN: Asparagina
- ASP: Aspartato
- ATP: Adenosina-Trifosfato
- CH: Carbohidratos
- CIT: Citrato
- CO₂: Dióxido de Carbono
- CoA: Coenzima A
- CTP: Citidina-Trifosfato
- CYS: Cisteína
- DHAP: Dihidroxiacetona
- DHF: Dihidrofolato
- dTTP: Timidina- Trifosfato
- E4P: Eritrosa-4-Fosfato
- F1,6DP: Fructosa-1,6-Difosfato

- F6P: Fructosa-6-Fosfato
 - FAD: Flavina Adenina Dinucleótido
 - FADH₂: Flavina Adenina Dinucleótido reducida
 - FOR: Formiato
 - FUM: Fumarato
 - G1,3DP: 1,3 Difosfoglicerato
 - G2P: 2- Fosfoglicerato
 - G3P: 3-Fosfoglicerato
 - G6P: Glucosa-6-Fosfato
 - GAP: Gliceraldehído-3-Fosfato
 - GDP: Guanosina- Difosfato
 - Glc: Glucosa
 - GLN: Glutamina
 - GLU: Glutamato
 - GLY: Glicina
 - GTP: Guanosina-Trifosfato
 - H⁺: Ion Hidronio
 - H₂O: Agua
 - HCO₃⁻: Ion Hidrógeno Carbonato
 - HIS: Histidina
 - ILE: Isoleucina
 - ISOCIT: Isocitrato
 - LAC: Lactato
 - LEU: Leucina
 - LYS e: Lisina extracelular
 - LYS: Lisina
 - MAL: Malato
 - MET: Metionina
 - N¹⁰- Formil- THF: N¹⁰- Formil- Tetrahidrofolato
 - N⁵- Formimino- THF: N⁵- Formimino- Tetrahidrofolato
 - N⁵, N¹⁰- Metilén- THF: N⁵, N¹⁰- Metilén- Tetrahidrofolato
 - NAD⁺: Nicotinamina Adenina Dinucleótido
 - NADH: Nicotinamina Adenina Dinucleótido reducida
 - NADP⁺: Nicotinamina Adenina Dinucleótido Fosfato
 - NADPH: Nicotinamina Adenina Dinucleótido Fosfato reducida
 - NH₃: Amoníaco
 - NH₄: Amonio
 - O₂: Oxígeno
 - OAA: Oxalacetato
 - OLC: Ácido Oleico
 - PEP: Fosfoenolpiruvato
 - PHE: Fenilalanina.
 - P_i: Fósforo Inorgánico
 - PRO: Prolina
 - PRPP: 5- Fosfo- Ribosa-Pirofosfato
 - PYR: Piruvato
 - QUI r: Quinona reducida
 - QUI: Quinona
 - R5P: Ribosa-5-Fosfato
 - Ru5P: Ribulosa-5-Fosfato
 - S7P: Sedoheptulosa-7-Fosfato
 - SER: Serina
 - SO₃²⁻: Ion Sulfito
 - SUC: Succinato
 - SucCoA: Succinil-CoA
 - THF: Tetrahidrofolato
 - THR: Treonina
 - TRE: Trealosa
 - TRP: Triptófano
 - TYR: Tirosina
 - UTP: Uridina- Trifosfato
 - VAL: Valina
 - X5P: Xilulosa-5-Fosfato
 - αKG: α-Oxoglutarato
- α_i número estequiométrico del metabolito i en la reacción j
- ε_j grado de avance de la reacción j (cantidad de sustancia/volumen)
- $\frac{dc_j}{dt}$ velocidad de la reacción j o flujo metabólico de la reacción j (cantidad de sustancia/volumen tiempo)