

BIO

OPTIMIZACIÓN
DE SISTEMAS
BIOTECNOLÓGICOS

SISTEMAS **BIOTECNO** **LÓGICOS**

Rodolfo Dondo

ediciones UNL



Optimización de sistemas biotecnológicos

Optimización de sistemas biotecnológicos

Rodolfo Dondo

ediciones UNL

CÁTEDRA

**UNIVERSIDAD
NACIONAL DEL LITORAL**



Consejo Asesor
Colección Cátedra
Alicia Camilloni
Daniel Comba
Bárbara Mántaras
Isabel Molinas
Héctor Odetti
Andrea Pacífico
Ivana Tosti

Dirección editorial
Ivana Tosti
Coordinación editorial
María Alejandra Sadrán
Coordinación comercial
José Díaz

Corrección
Lucía Bergamasco
Diagramación interior y tapa
Verónica Rainaud

© Ediciones UNL, 2026.

—

Sugerencias y comentarios
editorial@unl.edu.ar
www.unl.edu.ar/editorial

Optimización de sistemas biotecnológicos/
Rodolfo Dondo; Prólogo de Rodolfo Dondo
– 1ra. ed. – Santa Fe: Ediciones UNL, 2026.
Libro digital, PDF/A – (Cátedra)

Archivo Digital: descarga y online
ISBN 978-987-749-538-6

1. Biotecnología. 2. Matemática. 3. Educación
Superior. I. Dondo, Rodolfo, prolog. II. Título.
CDD 020

© Rodolfo Dondo, 2026.



Índice

NOMENCLATURA Y ACRÓNIMOS / 11

PRÓLOGO / 19

PARTE 1. FUNDAMENTOS MATEMÁTICOS Y HERRAMIENTAS COMPUTACIONALES

1. INTRODUCCIÓN Y FORMALISMOS MATEMÁTICOS / 25

Variables, parámetros y constantes / 27

Estados estacionarios y transientes / 28

Clasificación de modelos / 28

Referencias bibliográficas / 31

2. MÉTODOS DE OPTIMIZACIÓN / 33

Programación lineal / 33

Programación mixta entero-lineal / 44

Programación no lineal / 52

Programación mixta entero-no lineal / 58

Referencias bibliográficas / 67

3. SOFTWARE / 69

Programas de cálculo numérico de alto nivel / 70

Programas de modelado algebraico / 71

Software para resolución de problemas de control óptimo / 73

Referencias bibliográficas / 76

PARTE 2. MODELADO MATEMÁTICO DE SISTEMAS METABÓLICOS, BIOQUÍMICOS Y BIOTECNOLÓGICOS

4. MODELADO A ESCALA MICROSCÓPICA / 79

Representación topológica de redes metabólicas / 79

Modelado matemático del metabolismo / 81

Representación matemática de redes metabólicas / 82

Análisis de balance de flujo (FBA) / 84

Análisis de balance de flujo formulado como un problema
de programación lineal / 90

Análisis de variabilidad de flujo (FVA) / 97

Análisis dinámico de balance de flujo (dFBA) / 98

Análisis termodinámico en redes metabólicas / 101
Modelados cinéticos de redes metabólicas / 111
Referencias bibliográficas / 119

5. MODELADO A ESCALA MACROSCÓPICA / 121

Balances macroscópicos / 122
Ecuaciones de balance de materia total y de componentes / 122
Modelado cinético y tipos de modelos / 124
Integración de ecuaciones de balance con modelos cinéticos / 140
Referencias bibliográficas / 153

6. INTEGRACIÓN DE MODELOS A ESCALA MACROSCÓPICA

CON MODELOS A ESCALA MICROSCÓPICA / 155

Reducción de la dimensión de redes metabólicas a escala genómica / 155
Combinación de modos de flujos elementales y modelos cibernéticos / 162
Reducción dinámica de metabolismos no balanceados / 171
Referencias bibliográficas / 176

PARTE 3. APLICACIONES METABÓLICAS, BIOQUÍMICAS Y BIOTECNOLÓGICAS

7. OPTIMIZACIÓN DE FLUJOS EN REDES METABÓLICAS / 179

Optimización de flujos en redes metabólicas mediante
programación mixta entero-lineal / 182
Identificación óptima de knock-out genéticos / 197
Referencias bibliográficas / 208

8. OPTIMIZACIÓN DINÁMICA DE SISTEMAS MACROSCÓPICOS / 211

Maximización de la producción de biomasa en un reactor
de tipo *fed batch* / 211
Optimización de la producción de un metabolito diferente de la biomasa / 214
Optimización de la producción de un metabolito en presencia de múltiples
sustratos limitantes / 216
Optimización de un cultivo con manipulación de varias
variables de control / 218
Optimización de sistemas modelados mediante redes metabólicas / 221
Referencias bibliográficas / 232

9. ESTUDIO Y MODELADO DE REDES METABÓLICAS

EN MICROALGAS OLEAGINOSAS / 233

Microalgas: definición y generalidades taxonómicas.

Nichos ecológicos. Importancia biotecnológica / 233

Nutrientes: fuentes de carbono y nitrógeno / 234

Síntesis de polisacáridos: rutas biosintéticas de la síntesis de almidón
y otros carbohidratos de alta masa molecular / 245

Síntesis de aceites: rutas biosintéticas de la síntesis de ácidos grasos
y triacilglicéridos / 246

Cultivo autotrófico de microalgas en biorreactores / 247

Aplicaciones del modelado metabólico en la optimización de bioprocesos
con microalgas: etapas generales para la construcción de modelos
a escala genómica / 250

Referencias bibliográficas / 256

SOBRE EL AUTOR / 263

NOMENCLATURA Y ACRÓNIMOS

Símbolos romanos

A	Matriz de coeficientes
a_i	Coefficiente de producción específica del metabolito i
a_i / a_i^-	Variables binarias de eliminación/activación del flujo metabólico i
a_{ij}	ij -ésimo elemento de la matriz A
b	Vector de coeficientes del lado derecho
B	Matriz de coeficientes
B	Biomasa total
b_i	i -ésimo elemento del vector b
C	Vector de metabolitos acumulables
c	Vector de coeficientes de costos
c_j	j -ésimo elemento del vector c
df	Grados de libertad
V_i^+	Cinética agregada de producción del metabolito i
e	Vector de actividades enzimáticas
e_i	i -ésimo elemento del vector de actividades enzimáticas e
$f()$	Función f
F^i	Corriente de entrada a un biorreactor

F^0	Corriente de salida a un biorreactor
G	Energía libre de Gibbs
$g()$	Función g
H	Hessiano
$h()$	Función h
I	Conjunto de metabolitos
$I(t)$	Intensidad de iluminación
J	Conjunto de reacciones químicas
K	Espacio nulo de la matriz S
K_B	Coefficientes cinéticos asociados a la biomasa
K_C	Coefficientes cinéticos asociados a los metabolitos acumulables C
K_K	Coefficientes cinéticos asociados a los productos P
K_{Pi}	Constante específica de inhibición asociada al producto P_i
K_{iS}	Constante específica de inhibición asociada al sustrato i
f_{ji}	Coefficiente de consumo de sustrato i para la producción de metabolito j
z	Función objetivo de un problema de programación matemática
K_S	Coefficientes cinéticos asociados a los sustratos S
K_{Si}	Coefficiente de semi-saturación de la ecuación de Monod en el sustrato S_i

$L()$	Parte dependiente de la trayectoria de la función objetivo de un problema de control óptimo
LB_j	Cota inferior para la variable j
M	Número entero
m_{S_i}	Consumo específico de sustrato S_j para mantenimiento celular
N_{int}	Matriz estequiométrica de metabolitos internos
P	Productos
p	Presión
q_i	Coefficiente cinético específico de producción/consumo del metabolito i
R	Constante universal de los gases
r_i	Velocidad neta de producción del metabolito i
S	Sustratos
S	Matriz estequiométrica
s_{ij}	ij -ésimo elemento de la matriz estequiométrica S
t	Tiempo
T	Temperatura
t_o	Tiempo inicial
t_f	Tiempo final
u	Vector de variables de control
u_i	Variable de control i
UB_j	Cota superior para la variable j

$\mathbf{v}$	Vector de flujos de reacción
V	Volumen
v_i	i -ésimo elemento del vector de flujos de reacción $\mathbf{v}$
V_i^-	Cinética agregada de consumo del metabolito i
w_i	Peso asociado al i -ésimo elemento de un vector
X	Biomasa funcional
$\mathbf{x}$	Vector de variables de estados
$\mathbf{x}_0$	Valor inicial del vector de variables de estado
x_i	i -ésimo elemento del vector $\mathbf{x}$
X_{MAX}	Máxima concentración estacionaria de biomasa
$\mathbf{y}$	Vector de variables enteras
y_i	i -ésimo elemento del vector de variables enteras $\mathbf{y}$
$Y_{i/j}$	Rendimiento en producto i del sustrato j
Y_{P/S_i}	Rendimiento del metabolito P en el sustrato S_i
z	Conjunto de elementos químicos
z_i	Elemento químico i
Z_i^k	Parámetro de fijación del valor de la variable i en la iteración k

Símbolos griegos

μ	Vector de multiplicadores de Lagrange
-------	---------------------------------------

μ	Velocidad específica de crecimiento microbiano
μ_j	Variable dual asociada a la restricción j
δ	Un número pequeño
Δ_r	Variación de la variable r
E	Coefficiente de elasticidad
Θ	Vector de parámetros de un modelo
Λ	Vector de multiplicadores de Lagrange
λ	Variable dual
P	Densidad
Ψ	Función objetivo del problema de control óptimo
Φ	Pare de la función objetivo que depende de los estados al tiempo final
γ	Vector de variables cibernéticas de activación/desactivación
γ_i	i -ésimo elemento del vector de variables cibernéticas γ
γ_{ji}	Coefficiente de producción del metabolito i a partir del sustrato j
ϕ	Función objetivo de un problema de control óptimo

Símbolos matemáticos

$d(\)/dt$	Derivada respecto del tiempo
------------	------------------------------

$ $	Cardinalidad de un conjunto
∇	Operador diferencial
$\int \square$	Integral

Acrónimos

<i>ADP</i>	Adenosina-di-fosfato
<i>ATP</i>	Adenosina-tri-fosfato
<i>BST</i>	Teoría de sistemas bioquímicos
<i>CARB</i>	Carbohidratos
<i>CAT</i>	Catalasa
<i>COBRA</i>	Análisis basado en restricciones
<i>dFBA</i>	Análisis dinámico de balance de flujos
<i>DRUM</i>	Reducción dinámica del metabolismo no balanceado
<i>EDO</i>	Ecuación diferencial ordinaria
<i>EFM</i>	Modo elemental de flujo
<i>EMP</i>	Embden-Meyerhof-Parnas
<i>FBA</i>	Análisis de balance de flujos
<i>FVA</i>	Análisis de variabilidad de flujos
<i>G6P</i>	Glucosa-6-fosfato

<i>GAP</i>	Gliceraldeido-3-fosfato
<i>NADH</i>	Nicotinamida adenina dinucleótido reducido
<i>NADP</i>	Nicotinamida adenina dinucleótido fosfato
<i>NADH₂</i>	Nicotinamida adenina dinucleótido
<i>GOGAT</i>	Glutamina-oxoglutarato aminotransferasa
<i>GPR</i>	Asociación enzima-racción
<i>IP</i>	Programa entero
<i>LP</i>	Programa lineal
<i>MCA</i>	Análisis de control metabólico
<i>MILP</i>	Programa mixto entero lineal
<i>MINLP</i>	Programa mixto entero no lineal
<i>MOMA</i>	Minimización del ajuste metabólico
<i>NAPDPH₂</i>	Fosfato de hidrógeno dinucleótido de nicotinamida y adenina
<i>NLP</i>	Programa no lineal
<i>OC</i>	Control óptimo
<i>PEP</i>	Glucosa-6-fosfato
<i>RUBISCO</i>	Ribulosa-1,5-bifosfato carboxilasa
<i>SA</i>	Análisis de sensibilidad
<i>TAG</i>	Triglicéridos

Prólogo

El objetivo de este libro es proporcionar un tutorial sobre herramientas computacionales que utilizan optimización matemática para el análisis, modelado y rediseño de redes metabólicas así como también de procesos biotecnológicos. La biología de sistemas es la disciplina científica que estudia las propiedades sistémicas y las interacciones dinámicas en un objeto biológico, ya sea una célula o un virus de manera cuantitativa combinando estudios experimentales con modelos matemáticos. Los científicos pueden describir con gran precisión los procesos internos de las estrellas a mil años luz de distancia pero la descripción de cómo crece y se divide una célula sigue siendo desconcertante en muchos sentidos. Usamos levadura para producir cerveza pero, ¿entendemos cómo funcionan los procesos de la vida? Desde el siglo XIX, estos procesos ya no se explican haciendo referencia a «fuerzas vitales» sino mediante leyes de la física y la química. Al estudiar la estructura y la dinámica de los sistemas vivos, investigadores de diferentes disciplinas mostraron cómo los procesos vitales surgen de la estructura y organización funcional de las células, cómo decenas de miles de componentes bioquímicos interactúan de manera organizada y cómo estos sistemas están regulados por información genética y se adaptan continuamente mediante mutaciones y selección. El modelado matemático tiene una larga, aunque relativa marginal tradición en bioquímica y ha influido en este campo de muchas maneras. Los modelos se pueden utilizar para probar hipótesis, realizar predicciones cuantitativas o revelar lagunas o inconsistencias en hipótesis anteriores ayudándonos a mejorar nuestra comprensión de los procesos bioquímicos. Inspirada en las ideas de la cibernética de los años 1960 y 1970, la teoría de los sistemas dinámicos y la teoría de control se aplican con creciente frecuencia en el estudio de las vías metabólicas. Gracias a poderosas técnicas experimentales en genómica y proteómica, se ha acumulado una gran cantidad de datos biológicos y actualmente los modelos computacionales de células están al alcance de la mano. La biología de sistemas es la disciplina dedicada al desarrollo de tales modelos y utiliza a las redes bioquímicas como su concepto principal. Inspirados en la ingeniería eléctrica, los bioquímicos y biotecnólogos describen la estructura de los sistemas bioquímicos mediante gráficos de red que luego pueden

traducirse en modelos dinámicos. Este libro ofrece una visión general de la biología de sistemas, un campo en rápido desarrollo, y proporciona a los lectores métodos establecidos y emergentes. El análisis de las redes de regulación proporciona una perspectiva general sobre la fisiología celular y hace uso de conceptos como modularidad, robustez y optimización. El libro ha sido dividido en tres partes; en la primera se presentan los fundamentos matemáticos y las técnicas computacionales que se utilizarán en la segunda parte. En esta parte se modelan matemáticamente sistemas metabólicos, bioquímicos, y biotecnológicos. Y en la tercera, se presentan casos de aplicación. En esta parte el autor destaca y agradece la presentación del último capítulo realizada por los doctores: Alejandro Beccaría y Vanina Marquez, quienes generosamente decidieron colaborar con sus valiosos aportes.

El libro comienza con un tratamiento formal de las diferentes clases de problemas de optimización que se vuelven poco a poco más complejos, comenzando con programación lineal, siguiendo con programación no lineal, programación mixta entero–lineal, mixta entero–no lineal, concluyendo con problemas de control óptimo. Se explican sintéticamente los fundamentos de las formulaciones de optimización y las técnicas algorítmicas de solución relevantes para cada tipo de problema. Estos se presentan en el contexto del análisis y optimización de redes metabólicas y de optimización de procesos biotecnológicos de interés industrial. A posteriori se presentan metodologías de integración de modelos metabólicos a escala microscópica con modelos a escala macroscópica de procesos biotecnológicos. Este libro proporciona solo un tratamiento condensado de la aplicación de optimización a procesos biológicos y biotecnológicos. La descripción del metabolismo microbiano así como de procesos fermentativos y de interés biotecnológico es limitada por lo cual se proporcionarán referencias a lecturas de consulta relevantes. Se tuvo cuidado para garantizar que la homogeneidad de descripciones, nomenclatura y suposiciones necesarias para la descripción de los biosistemas mencionados. Las notas de la asignatura optativa de grado *Optimización de sistemas metabólicos bioquímicos y biotecnológicos* y del curso de posgrado *Optimización de sistemas biotecnológicos. Herramientas computacionales aplicadas en la obtención de bioproductos microbianos* impartidos en la Facultad de Bioquímica y Ciencias Biológicas de la Universidad Nacional del Litoral forman la base de este libro. El mismo está pensado para un graduado o para un curso avanzado en el uso de la optimización en redes metabólicas y en biotecnología. Nuestra esperanza es que este material sirva como punto de partida para que los estudiantes realicen investigaciones más profundas sobre las técnicas y conceptos relevantes que se presentan en el mismo.

Como autor tengo muchas personas a quienes agradecer. El reconocimiento más grande e importante es para mi esposa, Lía, y para mi hija, Sofía, quienes han tolerado muchas horas de ausencia durante las cuales se concibió, formuló y construyó el contenido de este libro. Sin su paciencia y apoyo durante los últimos años, este libro no habría sido posible.

Rodolfo Dondo

PARTE 1

Fundamentos matemáticos
y herramientas
computacionales

1 Introducción y formalismos matemáticos

Si analizamos a los fenómenos bioquímicos y biológicos, observamos a numerosos procesos que a menudo no pueden explicarse por los principios fundamentales. Aunque los principios generales de la bioquímica están bien establecidos, la bioquímica de las moléculas y sistemas individuales a menudo es desconocida y puede variar considerablemente entre diferentes especies. Los experimentos conducen a hipótesis biológicas sobre procesos individuales, pero a menudo no está claro si estas hipótesis pueden combinarse en una estructura conceptual más amplia y unificada. El modelado matemático y las simulaciones computacionales pueden ayudarnos a comprender la naturaleza interna y la dinámica de estos procesos y sirve para obtener predicciones sobre su desarrollo futuro y el efecto de las interacciones con el entorno.

¿Qué es un modelo?

En un sentido amplio, un modelo es una representación abstracta de objetos o procesos que explica sus características. Una red de reacciones bioquímicas puede representarse mediante un gráfico y la misma red también podría ser descrita por un sistema de ecuaciones diferenciales, que permite simular y predecir el comportamiento dinámico de la misma.

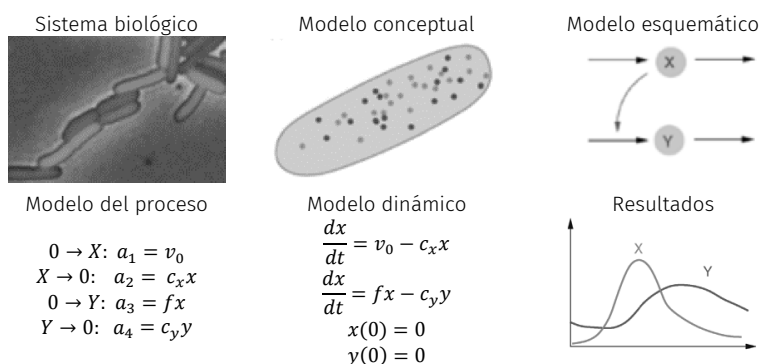


Figura 1.1. Diferentes abstracciones de un modelo matemático

Si se utiliza un modelo para realizar simulaciones, debe asegurarse de que estas predigan el comportamiento del sistema, en al menos aquellos aspectos que se supone que están cubiertos por el modelo. Los modelos de biología de sistemas a menudo se basan en leyes de la termodinámica de las reacciones químicas. Además, en un modelo se necesita hacer postulados específicos que aproten información mediante experimentos y conocimiento bioquímico, y en parte mediante la extrapolación desde otros sistemas. Un modelo resume el conocimiento establecido sobre dicho sistema en una formulación matemática. El modelado es un procedimiento subjetivo y un modelo representa solo algunos aspectos específicos de la realidad. Si se debe dilucidar un mecanismo biológico, entonces la estructura del sistema y las relaciones entre sus partes deben describirse de manera realista. Algunos modelos están destinados a ser aplicables a muchos objetos similares (por ejemplo, la cinética de Michaelis–Menten es válida para muchas enzimas) mientras que otros están específicamente adaptados a un objeto particular. La representación matemática se debe mantener en la forma más simple posible para permitir una implementación fácil y para obtener resultados comprensibles. También se puede modelar de manera muy realista pero muy complicada. Ninguna de las características mencionadas anteriormente hace que un modelo sea incorrecto o correcto, pero determinan si un modelo es apropiado para el caso de estudio a describir.

Ventajas del modelado computacional

Los modelos obtienen su referencia respecto de la realidad mediante la comparación de sus predicciones con los datos experimentales y por lo tanto, su capacidad de predicción depende de la calidad de los experimentos utilizados. Sin embargo, el modelado combinado con la experimentación tiene las siguientes ventajas en comparación con los estudios experimentales:

- El modelado requiere que las hipótesis sean específicas y conceptualmente rigurosas.
- Durante el proceso de formulación de un modelo deben determinarse componentes o interacciones no especificados.
- El modelado proporciona independencia del objeto modelado.
- El tiempo y el espacio se pueden dilatar o comprimir a voluntad.
- Los algoritmos de solución y los programas de computadora pueden usarse independientemente del sistema real.
- El modelado es barato en comparación con la experimentación.

- Los modelos no son perjudiciales para los animales o las plantas y ayudan a reducir los problemas éticos de ciertos experimentos.
- No contaminan el medio ambiente.
- El modelado puede ayudar a la experimentación ya que se pueden evaluar diferentes escenarios previamente a la misma.
- Se pueden imponer perturbaciones que no son factibles de imponer a un sistema real.
- Las simulaciones de modelos se pueden repetir frecuentemente y bajo condiciones diferentes.
- La representación gráfica y la visualización facilitan la comprensión del sistema.
- Se pueden hacer predicciones fundamentadas y verificables.

Además, los modelos computacionales se pueden usar para verificar las hipótesis propuestas para los fenómenos modelados. Los modelos computacionales sirven como «depósitos de conocimientos» tanto establecidos como hipotéticos. Al mismo tiempo, proporcionan a los investigadores descripciones cuantitativas de este conocimiento y permiten su simulación.

VARIABLES, PARÁMETROS Y CONSTANTES

Las cantidades involucradas en un modelo pueden clasificarse como variables, parámetros y constantes. Una constante es una cantidad con un valor fijo tal como la constante universal de los gases R . Los parámetros son cantidades que tienen un valor dado tal como, por ejemplo, el valor de una constante de Monod. Su valor depende del método y de las condiciones experimentales en las que fue determinado y puede cambiar. Las variables son cantidades con un valor que cambia con el tiempo para las cuales el modelo establece relaciones. Las *variables de estado* describen completamente el comportamiento de un sistema modelado y pueden asumir valores independientes o no. Depende de un modelo si una cantidad característica es una variable o un parámetro. Por ejemplo, en una cinética de reacción, la actividad de una enzima aparece como un parámetro pero la concentración de la enzima puede cambiar debido a la expresión génica o la degradación de una proteína, y en un modelo extendido, puede considerarse como una variable.

ESTADOS ESTACIONARIOS Y TRANSIENTES

El concepto de estado estacionario es importante para el modelado de sistemas dinámicos. Los estados estacionarios están determinados por el hecho de que los valores de todas las variables de estado permanecen constantes en el tiempo. El comportamiento asintótico de muchos sistemas dinámicos, es decir el comportamiento después de un tiempo lo suficientemente largo, a menudo es estacionario. Otros tipos de comportamiento asintótico son los regímenes oscilatorios o caóticos. La consideración de que un sistema biológico está en estado estacionario es en realidad una abstracción que se basa en una separación de las escalas de tiempo. Los procesos rápidos y lentos, que van desde la formación y la ruptura de enlaces químicos en nanosegundos hasta el crecimiento de individuos en años, se pueden acoplar en el mundo biológico. Mientras que los procesos rápidos a menudo alcanzan un estado casi estacionario después de un muy corto período de transición, el cambio del valor de las variables lentas a menudo es insignificante en la ventana de tiempo de considerada para el modelado. Por lo tanto, cada estado invariante durante la ventana de tiempo de interés se puede considerar como un estado estacionario de un sistema que está integrado en un entorno no estacionario más grande. El concepto de estados estacionarios es importante en el modelado cinético porque, a menudo, simplifica los modelos matemáticos. Los modelos matemáticos simplificados se pueden usar para probar hipótesis formuladas acerca de la dinámica del sistema, y tales modelos son más fáciles de entender y aplicar. Por otra parte, los estados transientes indican que el sistema asume una serie de estados diferentes y que evoluciona en el tiempo.

CLASIFICACIÓN DE MODELOS

Los modelos matemáticos son, en general, descripciones simplificadas de un proceso real que intentan identificar relaciones entre los factores que influyen en el crecimiento celular y en la generación de metabolitos. Los modelos matemáticos de procesos biotecnológicos suelen clasificarse como *estructurados* y *no estructurados*. Los *modelos no estructurados* son los más simples y consideran a la biomasa como a una masa uniforme sin dinámica interna y cuya velocidad de crecimiento depende de las condiciones medioambientales en las cuales se desarrolla. Por lo tanto, solo contienen información sobre la velocidad de crecimiento celular, el consumo de

substratos y la generación de productos. Los mismos son una buena aproximación al comportamiento real del sistema si los tiempos característicos de respuesta de los mecanismos intracelulares son despreciables o muy grandes con respecto del tiempo de duración del proceso bioquímico. Sin embargo, esta simplificación puede causar errores en las predicciones del modelo si los tiempos característicos de los procesos intracelulares son comparables con los tiempos del proceso. Esto conlleva el uso de los llamados *modelos estructurados* que consideran alguna forma, todavía simplificada, de dinámica intracelular.

Los procesos biotecnológicos generalmente involucran a una fase líquida y una fase gaseosa que en conjunto forman la denominada fase *abiótica* y la población de células es *fase biótica*. Las propiedades de estas fases se caracterizan por medio de variables macroscópicas tal como concentraciones de componentes relevantes y variables intensivas como la temperatura y el pH. Las reacciones catalizadas por microorganismos ocurren en la fase líquida y los componentes relevantes del sistema son:

- La masa celular (X), que es autocatalítica y se sintetiza a partir de los substratos.
- Los substratos (S) que actúan como fuente de nutrientes y de energía para la biomasa.
- Los productos metabólicos (P).

Hay que notar que no siempre los modelos no estructurados son imprecisos. Los mismos muestran buenas capacidades predictivas en casos en los que los tiempos característicos de los procesos metabólicos intracelulares sean muy grandes o muy pequeños respecto de los tiempos característicos de los cambios medioambientales originados en las acciones de control. Por lo tanto se pueden utilizar si se cumple con las condiciones ilustradas en la siguiente figura:

Constantes de tiempo del organismo		
Constantes de tiempo de los cambios medioambientales		
Estado pseudoestacionario	Sistema dinámico	Estado «congelado»
Sistema no dinámico	Retardos, oscilaciones	Sistema no dinámico
Siempre adaptado al medio	Sobreimpulsos	Siempre en estado inicial

Figura 1.2. Comparación de los tiempos característicos de los mecanismos intracelulares con los tiempos característicos de los cambios en las condiciones medioambientales.

Para el caso en el que los modelos no estructurados no sean confiables, el uso de los llamados *modelos estructurados* es una necesidad lógica. En ese sentido, los *modelos cibernéticos* han mostrado una remarcable robustez y capacidad predictiva a pesar de su relativa simplicidad. En este tipo de modelos, una célula puede ser comparada con una «fábrica» que opera en un ambiente competitivo que está representado por la presencia de otras especies. Durante la evolución de la especie, varias estructuras intracelulares fueron optimizadas para funcionar en las competitivas condiciones medioambientales en las que el organismo debe vivir y reproducirse. El principio darwiniano de supervivencia de los más aptos indujo a las especies vivas a desarrollar una serie de respuestas óptimas que le permitan sobrevivir y multiplicarse. Este principio puede representarse matemáticamente por medio de modelos cibernéticos que permiten describir los mecanismos de control intracelulares con un mínimo de conocimiento determinístico. Los principios de ingeniería de control son muy útiles para esta tarea. En ese sentido, la célula desarrolla varias estrategias dependiendo de las influencias que recibe del medio, tales como:

- Sensar al medioambiente para detectar la disponibilidad de materias primas que pueda «procesar».
- Modular las capacidades actuales y potenciales para incrementar «el volumen de procesamiento» de esos materiales.
- Adaptar la «maquinaria celular» para desarrollarse en las condiciones medioambientales actuales.
- Realizar cambios a nivel intracelular que le permitan sobrevivir, crecer y competir con otras especies por las materias primas disponibles. En ese sentido también pueden generar productos que inhiban el crecimiento de especies competitivas (antibióticos) o esporular si las condiciones externas no son propicias para la reproducción de la especie.

La propuesta cibernética suministra una justificación racional y simple para describir a un conjunto de fenómenos dinámicos intracelulares sobre los cuales muy poco es conocido. Este enfoque postula que los organismos han evolucionado bajo diversas presiones medioambientales y la supervivencia de la especie se interpreta como la consecuencia del desarrollo de mecanismos de adaptación que tienen como objetivo la perpetuación de la misma. Dicho objetivo se expresa matemáticamente como la maximización de la velocidad de reproducción de las células bajo las actuales condiciones medioambientales.

REFERENCIAS BIBLIOGRÁFICAS

KLIPP, E.; LIEBERMEISTER, W. (...) KOWALD (2016). *A. Systems Biology: A Textbook* (2nd edition). Wiley-VCH.

RAMKRISHNA, D. (2003). On Bioreactor Models for Control. *J. Process Control*, 13, 581–589. [https://doi.org/10.1016/S0959-1524\(02\)00092-6](https://doi.org/10.1016/S0959-1524(02)00092-6)

ROELS, J. A. (1983). *Energetics and Kinetics in Biotechnology*. Elsevier Biomedical Press.

2 Métodos de optimización

Este capítulo presenta fundamentos de optimización y de modelado matemático. Los problemas resultantes del modelado matemático se agrupan en las siguientes categorías:

1. Programación lineal.
2. Programación mixta entero-lineal.
3. Programación no lineal.
4. Programación mixta entero-no lineal
5. Programación dinámica o control óptimo.

Para cada una de las categorías anteriores se presenta una breve revisión de conceptos y las definiciones necesarias así como también los algoritmos que se utilizan para resolver cada una de las clases de problemas arriba mencionados.

PROGRAMACIÓN LINEAL

La programación lineal (LP) implica la minimización o maximización de una suma ponderada de variables que satisfacen un conjunto de restricciones de igualdad y/o restricciones de desigualdad de tipo lineal.

El término *programación* refleja su procedencia: los primeros programas de tipo LP modelaron problemas de logística. El método Simplex de Dantzig (Dantzig y Orden, 1955) fue el primer algoritmo para resolver de manera eficiente hasta optimalidad problemas lineales (LP). Más información referente a programación lineal puede encontrarse en (Bazaraa *et al.*, 2010; Ignizio y Cavalier, 1994; Rardin, 1998). Entre los software de programación y optimización más utilizado se pueden mencionar a AMPL (Fourer, 1990), GAMS (2021) y PYOMO (Bynum *et al.*, 2021). También hay una serie de otros entornos de software populares como MATLAB y Microsoft Excel que proporcionan fácil acceso a solucionadores de programas lineales.

Formas canónicas y estándar de un problema de optimización lineal

Un problema lineal se puede formular y representar de muchas formas diferentes. Las formas más utilizadas de un problema lineal son las formas canónicas y estándar. Estas dos representaciones son de hecho equivalentes y pueden convertirse una en la otra.

Forma canónica: la forma canónica se usa frecuentemente para introducir relaciones de dualidad. Un problema de minimización en forma canónica contiene solo restricciones de desigualdad de forma mayor o igual ($\geq$) con todas las variables no negativas. De manera similar, un problema de maximización en forma canónica consiste solo en restricciones de desigualdad de forma menor o igual ($\leq$) con todas las variables no negativas:

$$\begin{array}{ll}
 \text{Minimizar } \mathbf{c}^t \mathbf{x} & \text{Maximizar } \mathbf{c}^t \mathbf{x} \\
 \text{sujeto a} & \text{sujeto a} \\
 \mathbf{Ax} \geq \mathbf{b} & \mathbf{Ax} \leq \mathbf{b} \\
 \mathbf{x} \geq \mathbf{0} & \mathbf{x} \geq \mathbf{0}
 \end{array} \quad (2.1)$$

Donde $\mathbf{x}$ es un vector de dimensión $N \times 1$ llamado el vector de variables de decisión. Cada elemento x_j de este vector es una variable de decisión para el problema de optimización; $\mathbf{c}$ es un vector de dimensión $N \times 1$ denominado vector de costos para un problema de minimización y un vector de ganancias para un problema de maximización. Cada elemento c_j de este vector representa el costo (o la ganancia) asociado con cada variable x_j . $\mathbf{A}$ es una matriz $M \times N$ llamada matriz de coeficientes. Cada elemento a_{ij} de esta matriz es el coeficiente de la variable x_j en la restricción i ; $\mathbf{b}$ es un vector de dimensión $M \times 1$ llamado vector del lado derecho. Para problemas de minimización con restricciones de «mayor o igual», este vector a veces se denomina vector de demanda. Las mismas representaciones de un problema LP también se pueden expresar en forma expandida de la siguiente manera:

$$\begin{array}{ll}
 \text{Minimizar } z = \sum_{j=1}^N c_j x_j & \text{Maximizar } z = \sum_{j=1}^N c_j x_j \\
 \text{sujeto a} & \text{sujeto a} \\
 \sum_{j=1}^N a_{ij} x_j \geq b_i \quad i = 1, 2, \dots, M & \sum_{j=1}^N a_{ij} x_j \leq b_i \quad i = 1, 2, \dots, M \\
 x_j \geq 0 \quad j = 1, 2, \dots, N & x_j \geq 0 \quad j = 1, 2, \dots, N
 \end{array} \quad (2.2)$$

Forma estándar: la forma estándar de un problema de LP es un problema de maximización que consiste solo en restricciones de igualdad con todas las variables restringidas para no ser negativas:

$$\text{Maximizar } z = \mathbf{c}^t \mathbf{x} \quad (2.3.1)$$

sujeto a

$$\mathbf{Ax} = \mathbf{b} \quad (2.3.2)$$

$$\mathbf{x} \geq \mathbf{0}$$

O en forma expandida:

$$\text{Maximizar } z = \sum_{j=1}^N c_j x_j \quad (2.4.1)$$

sujeto a

$$\sum_{j=1}^N a_{ij} x_j = b_i \quad i = 1, 2, \dots, M \quad (2.4.2)$$

$$x_j \geq 0 \quad j = 1, 2, \dots, N$$

Cuando se trata de restricciones de igualdad, los elementos b_i del vector $\mathbf{b}$ se denominan coeficientes de requisito. Tanto en la forma estándar como en la canónica, la función objetivo y las restricciones son funciones lineales de las variables x_j . Cualquier problema LP se puede convertir en una forma canónica o estándar usando las siguientes transformaciones:

- Las restricciones de desigualdad se pueden convertir en restricciones de igualdad introduciendo variables de holgura de la siguiente manera:

$$\sum_{j=1}^N a_{ij} x_j \leq b_i \equiv \sum_{j=1}^N a_{ij} x_j + s_i^+ = b_i \quad (2.5)$$

- Las variables reales se pueden expresar como la diferencia de dos variables reales no negativas:

$$x_j = x_i^+ - x_i^- \text{ donde } x_i^+, x_i^- \geq 0 \quad (2.6)$$

- Un problema de minimización se puede convertir en un problema equivalente de maximización del negativo de la función objetivo:

$$\text{Minimizar } z = -\text{Maximizar}(-z) \quad (2.7)$$

Ejemplo:

$$\text{Maximizar } z = x_1 + x_2 \quad (2.8.1)$$

sujeto a

$$x_1 + 2x_2 \leq 2 \quad (2.8.2)$$

$$2x_1 + x_2 \leq 2 \quad (2.8.3)$$

$$x_1, x_2 \geq 0$$

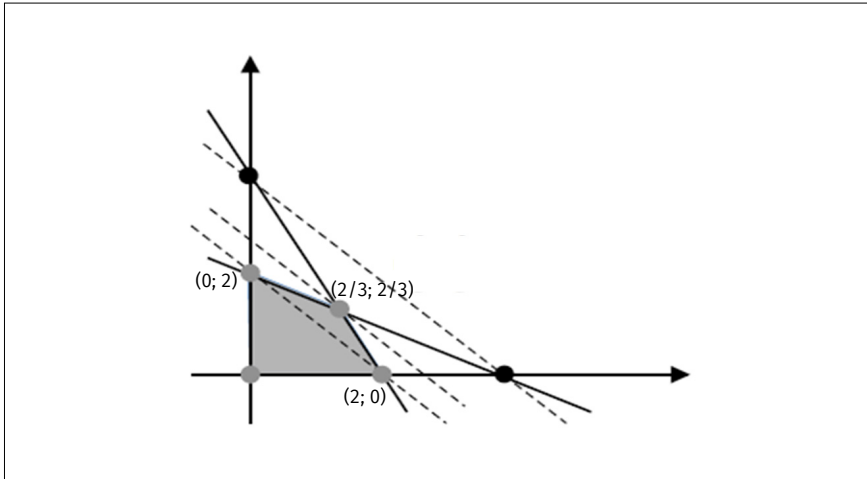


Figura 2.1. El área sombreada indica la región factible definida por las restricciones. Las líneas discontinuas representan los conjuntos de niveles de función objetivo, los puntos grises representan los vértices de región factibles y los puntos negros representan puntos de intersección no factibles. La solución óptima al problema se encuentra en el vértice $(2/3, 2/3)$.

Multiplicidad de soluciones óptimas

En el ejemplo anterior, el problema LP tenía un punto óptimo único debido a que el conjunto de niveles de la función objetivo intercepta a un solo punto de la región factible. Un problema LP también puede tener un número infinito de óptimos si los niveles de la función objetivo se superponen con una faceta completa del poliedro de la región factible. La dimensión de la faceta del poliedro compuesta por todos los puntos óptimos puede variar según la orientación de los conjuntos de niveles de la función objetivo y la ubicación de la región factible.

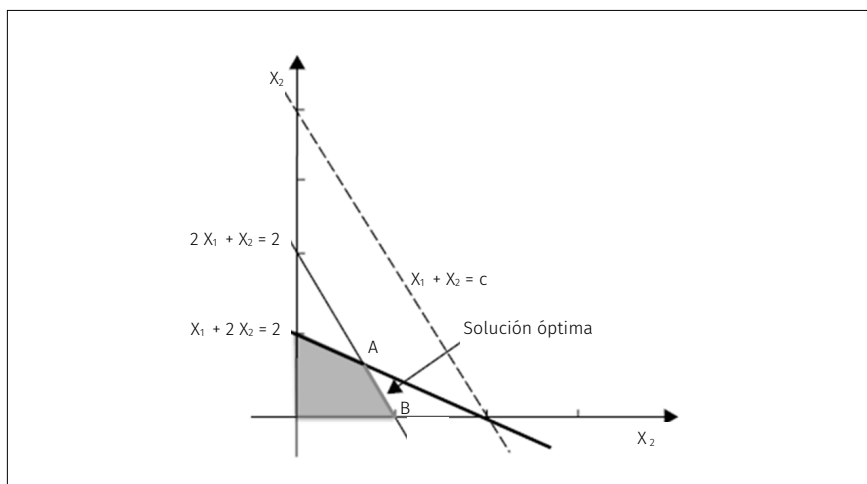


Figura 2.2. Multiplicidad de puntos de solución óptima para el problema de LP. Un número infinito de puntos de solución óptima está definido por el segmento AB.

Soluciones infactibles o no limitadas

Los problemas LP pueden ser infactibles o ilimitados. La infactibilidad implica que el espacio factible está vacío, mientras que la ilimitación significa que se puede aumentar arbitrariamente el valor de la función objetivo, para un problema de maximización, sin violar ninguna de las restricciones. Si reemplazamos las restricciones del ejemplo original por:

$$x_1 + 2x_2 \geq 2 \quad (2.9.1)$$

$$2x_1 + x_2 \geq 2 \quad (2.9.2)$$

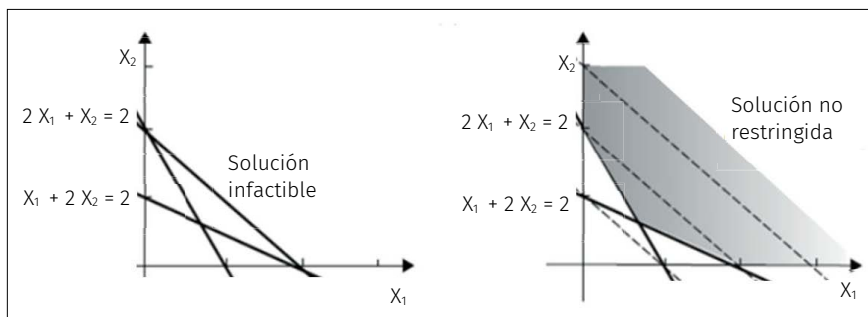


Figura 2.3. Ejemplos de (a) problemas LP inviables y (b) sin limitación.

Dualidad

Un problema LP con N variables y M restricciones está asociado con otro problema LP llamado dual que contiene M variables y N restricciones. La designación de primal y dual es arbitraria, ya que el dual del problema dual es el problema primal. Por lo tanto, se les conoce más apropiadamente como el par primal-dual. El término primal en general denota el problema LP «original».

Formulación del problema dual: considere el siguiente LP en forma canónica:

$$\text{Minimizar } \mathbf{c}^t \mathbf{x} \quad (2.10.1)$$

sujeto a

$$\mathbf{A} \mathbf{x} \geq \mathbf{b} \quad (2.10.2)$$

$$\mathbf{x} \geq \mathbf{0}$$

donde $\mathbf{x}$, $\mathbf{c}$ y $\mathbf{b}$ son vectores de dimensión $N \times 1$, $N \times 1$ y $M \times 1$, respectivamente, y $\mathbf{A}$ es una matriz $M \times N$. El problema dual asociado con el primal se define de la siguiente manera:

$$\text{Maximizar } \mathbf{b}^t \mathbf{y} \quad (2.11.1)$$

sujeto a

$$\mathbf{A}^t \mathbf{y} \leq \mathbf{c} \quad (2.11.2)$$

$$\mathbf{y} \geq \mathbf{0}$$

donde $\mathbf{y}$ es un vector $M \times 1$. A través de la inspección de los problemas primal y dual, se pueden recoger las siguientes observaciones:

- Mientras que el primal implica una minimización, el dual implica una maximización.
- El problema de minimización primal contiene restricciones « $\geq$ », mientras que el problema de maximización dual contiene restricciones « $\leq$ » (es decir, se invierte la dirección de las desigualdades).
- Cada restricción en el problema dual corresponde a una variable en el problema primal y viceversa. Esto significa que el problema principal tiene N variables y M restricciones mientras que el problema dual tiene M variables y N restricciones.
- La matriz de los coeficientes para el dual es la transpuesta de la matriz de coeficientes para el primal y viceversa.
- El vector de costos $\mathbf{c}$ en el primal sirve como el vector del lado derecho en el dual y viceversa.

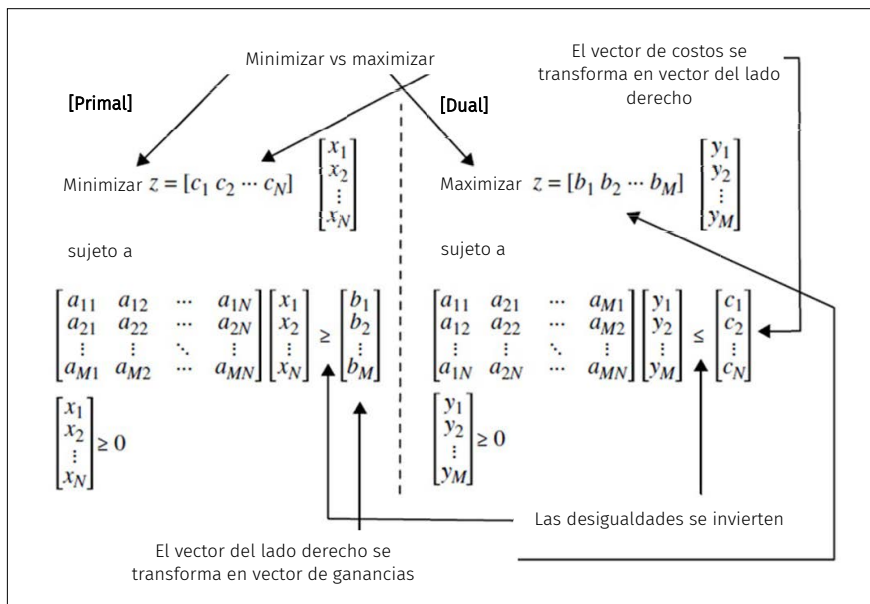


Figura 2.4. Conexiones entre un problema primal y su dual para un LP en forma canónica.

Considere el siguiente problema LP (primal) que corresponde a la estructura matemática de un problema de análisis de balance de flujo (FBA) de redes metabólicas:

$$\begin{array}{ll} \text{Maximizar } z = x_b & \text{para un } b \in J \\ \text{sujeto a} & \end{array} \quad (2.12.1)$$

$$\begin{array}{ll} \sum_{j \in J} S_{ij} x_j = 0 & \forall i \in I \\ LB_j \leq x_j \leq UB_j & j \in J \end{array} \quad (2.12.2)$$

$$x_j \in R \quad (2.12.3)$$

donde:

J	Conjunto de variables.
I	Conjuntos de restricciones de igualdad.
x_j	Variables de decisión.
S_{ij}	Coefficiente de contribución de la variable x_j an la restricción i .
LB_j	Cota inferior para la variable x_j .
UB_j	Cota superior para la variable x_j .

Solución: observe que el problema primal tiene N variables y $M + N \times 2$ restricciones (M igualdades y $2N$ desigualdades). Por lo tanto, el problema dual tendrá $M + N \times 2$ variables y N restricciones. Las variables duales correspondientes a las restricciones primales y se definen de la siguiente manera:

Tabla 2.1. Restricciones duales correspondientes a cada variable primal x_j

Restricción	Índice	Dual
$\sum_{j \in J} S_{ij} x_j = 0$	$\forall i \in I$	$\rightarrow \lambda_i \in R$
$-x_j \leq -LB_j$	$j \in J$	$\rightarrow \mu_j^{LB} \geq 0$
$x_j \leq UB_j$	$j \in J$	$\rightarrow \mu_j^{UB} \geq 0$

Determinación de los términos del lado izquierdo de cada restricción dual			
Restricción primal	Variable dual asociada	Coefficiente de x_j en la restricción	Término de la restricción dual
$\sum_{j \in J} S_{ij} x_j = 0 \quad \forall i \in I$	λ_i	S_{ij}	$\sum_{i \in I} S_{ij} \lambda_i$
$-x_j \leq -LB \quad \forall j \in J$	μ_j^{LB}	-1	$-\mu_j^{LB}$
$x_j \leq UB \quad \forall j \in J$	μ_j^{UB}	1	μ_j^{UB}

Determinación del lado derecho de cada restricción dual		
Variables primales	Coefficiente en la función objetivo primal	Lado derecho de la restricción dual con la variable primal
$x_j \quad \forall j \in J - b$	0	0
x_b	1	1

Determinación de los coeficientes de cada variable dual en la función objetivo dual			
Restricción primal	Variable dual asociada	Lado derecho de la restricción primal	Término de la función objetivo dual
$\sum_{j \in J} S_{ij} x_j = 0 \quad \forall i \in I$	λ_i	0	0 λ_i
$-x_j \leq -LB \quad \forall j \in J$	μ_j^{LB}	-LB	$-LB_j \mu_j^{LB}$
$x_j \leq UB \quad \forall j \in J$	μ_j^{UB}	UB	$UB_j \mu_j^{UB}$

Problemas de optimización no lineal que se pueden transformar en problemas LP

Existen problemas de optimización no lineal que se pueden resolver utilizando una formulación LP por medio de una transformación de su estructura matemática, tal como se muestra a continuación.

Valores absolutos en la función del objetivo: considere el siguiente problema de minimización:

$$\text{Minimizar } z = \sum_{i \in \{1, \dots, N_y\}} |y_i| \quad (2.13.1)$$

$$\begin{aligned} \text{sujeto a} \\ \mathbf{Ax} + \mathbf{y} = \mathbf{b} \\ \mathbf{x} \geq 0, \mathbf{y} \in \mathbf{R} \end{aligned} \quad (2.13.2)$$

El problema puede ser linealizado utilizando la siguiente transformación:

$$|y_i| = y_i^+ \quad \text{donde} \quad y_i^+ \geq 0 \quad \forall i \in \{1, \dots, N_y\} \quad (2.14)$$

Y considerando las siguientes restricciones adicionales:

$$y_i^+ \geq 0 \quad \mathbf{y} \quad y_i^+ \geq -y_i \quad (2.15)$$

Consecuentemente el problema no lineal se transforma en el siguiente problema lineal:

$$\text{Minimizar } z = \sum_{i \in \{1, \dots, N_y\}} |y_i| \quad (2.16.1)$$

sujeto a

$$\mathbf{Ax} + \mathbf{y} = \mathbf{b} \quad (2.16.2)$$

$$\mathbf{y}^+ \geq \mathbf{y} \quad (2.16.2)$$

$$\mathbf{y}^+ \geq -\mathbf{y} \quad (2.16.3)$$

$$\mathbf{x}, \mathbf{y}^+ \geq 0, \mathbf{y} \in \mathbf{R}$$

Esta técnica de transformación es viable solo si la suma de los valores absolutos se minimiza pero no es válida si se maximiza la función objetivo del problema original.

Problemas de optimización de MinMax con restricciones lineales: considere el siguiente problema de optimización de tipo Min–Max:

$$\text{Minimizar } \{ \text{Maximizar}_{i \in \{1, \dots, N_y\}} y_j \} \quad (2.17.1)$$

sujeto a

$$\mathbf{Ax} + \mathbf{y} = \mathbf{b} \quad (2.17.2)$$

$$\mathbf{x} \geq 0, \mathbf{y} \in \mathbf{R}$$

El problema de optimización puede linealizarse de la siguiente manera:

$$\text{Minimizar } z \quad (2.18.1)$$

sujeto a

$$z \geq y_j \quad i \in \{1, \dots, N_y\} \quad (2.18.2)$$

$$\mathbf{Ax} + \mathbf{y} = \mathbf{b} \quad (2.18.3)$$

$$\mathbf{x} \geq 0, \mathbf{y} \in \mathbf{R}$$

El operador de minimización asegurará que al menos una de las variables y_j sea igual a z . Es fácil verificar que los problemas de optimización de *Max–Min* se pueden linealizar de una manera similar. Sin embargo, los problemas *Max–Max* y *Min–Min* no se pueden linealizar tal como se describe anteriormente. Por lo general, requieren la introducción de variables binarias para garantizar que al menos una de las restricciones $z y_j$ (o $j z_i$) se cumpla como igualdad.

Programación fraccional lineal: un problema de programación fraccional lineal consiste en una función objetivo fraccional con numerador denominador y restricciones lineales:

$$\text{Maximizar } \left\{ \text{Minimizar } z = \frac{\mathbf{cx} + p}{\mathbf{dx} + q} \right\} \quad (2.19.1)$$

sujeto a

$$\mathbf{Ax} \leq \mathbf{b} \quad (2.19.2)$$

$$\mathbf{x} \geq 0$$

donde p y q son escalares. Este problema de optimización no es lineal debido a la presencia de un término fraccional en la función objetivo. Este problema puede ser linealizado de la siguiente manera: supongamos que la región factible está delimitada y el denominador de la función objetivo es positiva: $\mathbf{dx} + q \geq 0$. Sea r una nueva variable definida como sigue:

$$r = \frac{1}{\mathbf{dx} + q} \quad (2.20)$$

La función objetivo se puede escribir como $z = \mathbf{c}\mathbf{x} + r p$. El vector original de las variables $\mathbf{x}$ se mapea uno a uno a un nuevo vector de variables $\mathbf{y}$ y se escala por la variable escalar r :

$$\mathbf{y} = r\mathbf{x} \geq 0 \quad (2.21)$$

Al introducir r y reemplazar la variable $\mathbf{x}$ con la variable $\mathbf{y}$, la optimización original del problema se puede reescribir de la siguiente manera:

$$\text{Maximizar } \{\text{Minimizar } z = \mathbf{c}\mathbf{y} + rp\} \quad (2.22.1)$$

sujeto a

$$\mathbf{A}\left(\frac{\mathbf{y}}{r}\right) \leq \mathbf{b} \quad (2.22.2)$$

$$r = \frac{1}{d\mathbf{x} + q} \quad (2.22.3)$$

$$\mathbf{x}, \mathbf{y} \geq 0, r > 0$$

o reformulando:

$$\text{Maximizar } \{\text{Minimizar } z = \mathbf{c}\mathbf{y} + rp\} \quad (2.23.1)$$

sujeto a

$$\mathbf{A}\mathbf{y} \leq r\mathbf{b} \quad (2.23.2)$$

$$d\mathbf{y} + q\mathbf{r} = 1 \quad (2.23.3)$$

$$\mathbf{y} \geq 0, r > 0$$

Este es un programa lineal con respecto a las nuevas variables $\mathbf{y}$. En esencia, esta transformación permite escalar tanto el numerador como el denominador utilizando a r para que el denominador se vuelva igual a uno. El factor de escala sigue siendo una variable hasta que el problema está resuelto.

PROGRAMACIÓN MIXTA ENTERO-LINEAL

Las variables binarias se utilizan frecuentemente en formulaciones de optimización de redes metabólicas para modelar la presencia o ausencia de reacciones y para imponer restricciones al número total de reacciones activas o a las interacciones regulatorias. En general, las variables pueden ser continuas o discretas. Una variable entera es una variable que está restringida para asumir solo valores enteros discretos. Una variable binaria es

una variable entera que puede asumir el valor de solo 0 o 1. Un problema mixto entero–lineal (MILP) es un problema de optimización, donde la función objetivo y todas las restricciones son lineales y un subconjunto de las variables es restringido a ser entero. Un problema MILP se representa de la siguiente manera:

$$\begin{aligned} \text{Minimizar } z &= \mathbf{c}^T \mathbf{x} + \mathbf{d}^T \mathbf{y} & (2.24.1) \\ \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y} &\geq \mathbf{b} & (2.24.2) \\ \mathbf{x} &\in \mathbf{R}^N, \mathbf{y} \in \mathbf{Z}^M \end{aligned}$$

donde $\mathbf{R}$ y $\mathbf{Z}$ representan al conjunto de números reales y enteros, mientras que $\mathbf{x}$ e $\mathbf{y}$ son vectores de variables continuas y enteras, respectivamente. Además, $\mathbf{c}$, $\mathbf{d}$ y $\mathbf{b}$ son vectores de parámetros $N \times 1$, $M \times 1$ y $L \times 1$; y $\mathbf{A}$ y $\mathbf{B}$ son matrices de coeficientes $L \times N$ y $L \times M$, respectivamente. Si todas las variables están restringidas a valores enteros, entonces el problema se conoce como un problema de programación lineal entera (ILP), mientras que si todas las variables pueden asumir valores reales, el problema vuelve a ser un programa lineal (LP). La forma más utilizada de un problema MILP es un caso especial en donde todas las variables enteras están restringidas a ser binarias. Estos problemas son a menudo referidos como 0–1 MILP. Cada MILP con variables enteras acotadas puede transformarse en un 0–1 MILP. Una forma sencilla de reemplazar una variable entera z con variables binarias y_k es como sigue:

$$z = \sum_{k=z^L}^{z^U} k y_k \quad (2.25.1)$$

$$\begin{aligned} \sum_{k=z^L}^{z^U} y_k &= 1 & (2.25.2) \\ y_k &\in \{0,1\} \end{aligned}$$

donde z^L y z^U son el límite inferior y superior en z , respectivamente. Aunque esta es la transcripción más directa, hay una forma más compacta y que usa menos variables binarias. Este método es referido como la representación a nivel de bits de una variable entera (asumiendo $z^L \geq 0$):

$$z = \sum_{k=0}^H 2^k y_k \quad (2.26.1)$$

$$z^L \leq 2^K y_K \leq z^U \quad (2.26.2)$$

donde $K = \lceil \log_2 (z^U - z^L) \rceil$ con $\lceil . \rceil$ denotado la parte entera de un número real.

Modelado con variables binarias

En general, las variables binarias se utilizan en aplicaciones que involucran una decisión de sí/no y refiere a la presencia o ausencia de variables y/o restricciones. A continuación, presentamos algunos de los usos más frecuentes de las variables binarias con particular atención en el análisis de redes metabólicas.

Encendido/apagado de una variable continua: en el análisis de redes metabólicas, las variables binarias se utilizan típicamente como interruptores de encendido y apagado. Se pueden usar para imponer un nocaout o cuantificar un requisito de minimalidad en la función objetivo. Por ejemplo, para apagar el flujo de una reacción j , una variable binaria se utiliza de la siguiente manera:

$$y_j = \begin{cases} 1, & \text{si la reacción } j \text{ está activa} \\ 0, & \text{si la reacción } j \text{ está inactiva} \end{cases} \quad \forall j \in J \quad (2.27.1)$$

$$LB_j \leq x_j \leq UB_j \quad \forall j \in J \quad (2.27.2)$$

Aquí, LB_j y UB_j son límites inferiores y superiores en el flujo de reacción j respectivamente y J es el conjunto de reacciones en la red metabólica. El flujo de la reacción i es forzado a ser cero cuando la variable binaria y_j asume un valor de cero, mientras que puede variar entre sus límites inferior y superior cuando y_j es igual a uno. A veces, el número total de reacciones de apagado o encendido está bajo control. La siguiente restricción asegura que el número total de reacciones inactivadas en una red metabólica es a lo sumo el entero K .

$$\sum_{j \in J} (1 - y_j) \leq K \quad (2.28)$$

Conmutación dependiente de la condición: en algunos casos, apagar o encender una variable continua depende del valor de otra variable continua. Por ejemplo, considere un caso donde la reacción j_1 debe estar inactiva siempre que la reacción j_2 esté inactiva. A menudo es útil tabular todas las combinaciones de valores de variables binarias como se muestra en la Tabla 1 antes de llegar a las restricciones apropiadas para imponer el cambio de la variable condicional.

Tabla 2.2. Análisis de todos los casos para una eliminación de reacción condicional

y_{j_1}	Valores permitidos para y_{j_1}
1	0 o 1
0	0

Al leer las entradas de la Tabla 1, se hace evidente que la siguiente restricción es factible si y solo si se elige la siguiente combinación de valor binario de la Tabla 1.

$$y_{j1} \leq y_{j2} \quad (2.29)$$

Conmutación de restricciones dependiente de la condición: de manera similar a la conmutación de variables, a veces una o más restricciones deben activarse o desactivarse cuando una condición es verdadera o falsa, respectivamente. Considere una restricción de desigualdad $g(x) \leq 0$ que debería estar activa solo si una condición es verdadera. Se puede usar una variable binaria para modelar el caso en el que la condición deseada se cumple o no se cumple:

$$y_j = \begin{cases} 1, & \text{si la reacción } j \text{ está activa} \\ 0, & \text{si la reacción } j \text{ está inactiva} \end{cases} \quad \forall j \in J \quad (2.30)$$

La activación de la restricción condicional se puede capturar en un modelo de optimización de la siguiente manera:

$$g(x) \leq M(1 - y) \quad (2.31)$$

Donde M es un número positivo grande. La imposición condicional de una restricción de igualdad $h(x) = 0$ sigue la misma línea de razonamiento al considerar que una restricción de igualdad puede verse como dos desigualdades con signos de desigualdad opuestos.

$$-M(1 - y) \leq h(x) \leq M(1 - y) \quad (2.32)$$

Modelado de relaciones AND: a veces, se necesita modelar una situación en la que se debe satisfacer simultáneamente múltiples condiciones (operador y/AND) para que se active un resultado. El producto de las variables binarias que representan cada una de estas condiciones captura una relación AND.

Tabla 2.3. Tabla de verdad para el producto de dos variables binarias, $z = y_1 y_2$

y_1	y_2	$z = y_1 y_2$ (y_1 AND y_2)
1	1	1
1	0	0
0	1	0
0	0	0

El producto de las variables binarias da lugar a términos no lineales que convierten un problema MILP en un problema de programación mixto entero no lineal (MINLP). Por lo tanto, deben ser redefinidos como conjuntos equivalentes de restricciones lineales para que los resolvidores MILP puedan solucionar el problema. Como podemos ver en la tabla de verdad, siempre tenemos $0 \leq z \leq 1$. Además, las siguientes restricciones se mantienen en las cuatro combinaciones:

$$z \leq y_1 \quad (2.33.1)$$

$$z \leq y_2 \quad (2.33.2)$$

Estas relaciones proporcionan un límite superior para z ; sin embargo, también se necesita un límite inferior para asegurar que z asuma un valor de uno cuando y_1 e y_2 son iguales a uno. Esto se puede imponer agregando la siguiente restricción:

$$z \geq y_1 + y_2 - 1 \quad (2.34)$$

Tenga en cuenta que no es necesario declarar a z como una variable binaria pues las tres restricciones anteriores junto con la restricción de no negatividad ($z \geq 0$) aseguran que z solo puede tomar un valor de cero o uno. Esta técnica de linealización exacta puede generalizarse fácilmente al producto de más de dos variables binarias de la siguiente manera:

$$z = \prod_{j=1}^M y_j \quad (2.35.1)$$

$$z \leq y_j \quad \forall j \in \{1, \dots, M\} \quad (2.35.2)$$

$$z \geq \sum_{j=1}^M y_j - (M - 1) \quad \forall j \in \{1, \dots, M\} \quad (2.35.3)$$

$$z \geq 0$$

Modelado de relaciones OR: las variables binarias también se pueden utilizar para modelar las relaciones OR. Esta necesidad surge cuando una de las múltiples condiciones es suficiente para activar un proceso. Como antes, una tabla de verdad es útil para identificar todas las restricciones requeridas.

Tabla 2.4. Tabla de verdad para modelar relaciones OR

y_1	y_2	$z = y_1 \text{ OR } y_2$
1	1	1
1	0	1
0	1	1
0	0	0

Se verifica que las siguientes restricciones capturan todos los requisitos impuestos en la tabla de verdad:

$$z \geq y_1 \quad (2.36.1)$$

$$z \geq y_2 \quad (2.36.2)$$

$$z \leq y_1 + y_2 \quad (2.36.3)$$

$$z \leq 1 \quad (2.36.4)$$

Como antes, z puede no ser una variable binaria, ya que estas cuatro restricciones juntas aseguran que solo puede asumir valores de cero o uno. Esta técnica se puede generalizar fácilmente para más de dos condiciones OR de la siguiente manera:

$$z \geq y_j \quad \forall j \in \{1, \dots, M\} \quad (2.37.1)$$

$$z \leq \sum_{j=1}^M y_j \quad (2.37.2)$$

$$z \leq 1 \quad (2.37.3)$$

Linealización exacta del producto de una variable continua y una variable binaria

El producto entre variables binarias y continuas frecuentemente surge en modelos de optimización. Dado que este producto es un término no lineal, debe transformarse en un conjunto de restricciones lineales si se van a utilizar solucionadores MILP. Sea $z = x \cdot y$ el producto de la variable continua $x^L \leq x \leq x^U$ y variable binaria y . La siguiente restricción asegura que z asuma un valor de cero cuando $y = 0$ y z varíe entre x^L y x^U cuando $y = 1$.

$$x^L y \leq z \leq x^U y \quad (2.38)$$

Sin embargo, esta restricción no impone que $z = x$ cuando $y = 1$. Esto se logra mediante la siguiente restricción:

$$x - x^U(1 - y) \leq z \leq x - (1 - y)x^L \quad (2.39)$$

La linealización exacta del producto $x \cdot y$ es una muestra de cómo términos aparentemente no lineales que involucran variables binarias se puede expresar como restricciones lineales que permiten que un modelo de optimización se mantenga lineal y permita el uso de solucionadores MILP como CPLEX y GUROBI. Usando la misma idea, puede lograrse la linealización exacta de términos más complejos como $y_2 x$ o $y_1 y_2 x$. Vale la pena señalar que, en

contraste con el producto de una variable binaria y una variable continua, el producto de dos variables continuas es un término no lineal que no se puede reescribir de manera equivalente como un conjunto de restricciones lineales.

Resolución de problemas MILP

Los problemas MILP son mucho más difíciles de resolver que los problemas de LP. Aunque la región factible de un problema MILP se reduce, ya que algunas de las variables pueden asumir solo variables enteras, se pierden una serie de propiedades que hacen que los problemas de LP sean fáciles de resolver. En particular, la solución óptima, si existe, ya no tiene que ocurrir en un punto de vértice de la región factible. Esto significa que se deben examinar muchos puntos factibles del problema interior para obtener un problema no-polinómico-completo (NP-completo). Los problemas NP-completos pertenecen a una clase de problemas equivalentes cuya solución en el peor de los casos aumenta en orden exponencial con el tamaño del problema.

Procedimiento de ramificación y acotamiento (branch-and-bound): la idea detrás del algoritmo de ramificación y acotamiento es dividir sucesivamente el problema original en subproblemas más pequeños. Se procede resolviendo la relajación lineal de estos subproblemas para eliminar subregiones del espacio factible que no contienen la solución óptima y actualizando la mejor solución entera encontrada hasta el momento. Este proceso iterativo termina cuando solo queda un punto, o una subregión pequeña, que contiene la solución óptima después de excluir todas las demás partes de la región factible. Solo introduciremos algunos conceptos utilizados antes de presentar los pasos algorítmicos para el método de ramificación y acotamiento:

- Relajación.
- Ramificación.
- Orden de precedencia de la ramificación: búsqueda por mejor cota; búsqueda de mayor profundidad; búsqueda de primera amplitud.
- Poda y avance.

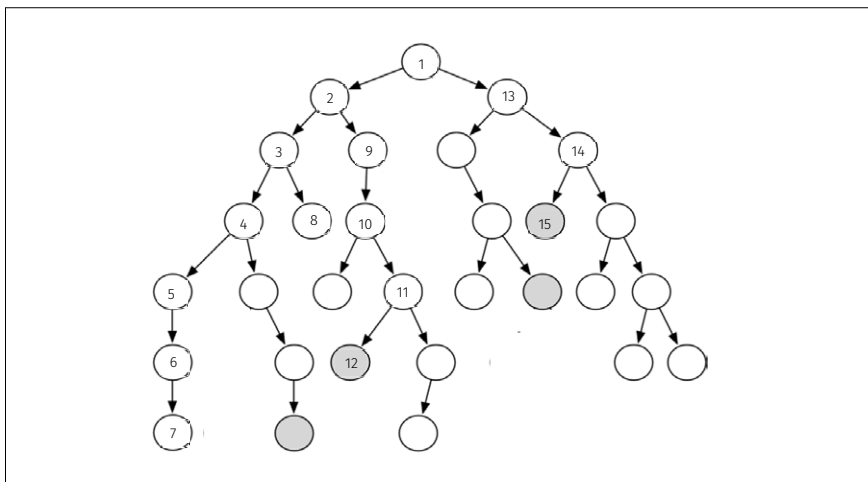


Figura 2.5. Ilustración de un árbol de *branch-and-bound*.

Estrategias de formulación eficiente de problemas MILP: los tiempos de solución de un problema MILP pueden verse afectados drásticamente por la formulación del mismo. A continuación se describen algunas reglas generales para modelar con eficiencia un problema MILP:

- Usar la menor cantidad posible de variables binarias.
- Fijar las variables binarias que no afectan a la solución óptima.
- Agrupar, cuando es posible, todas las variables binarias acopladas.
- Segregar las variables binarias en las restricciones antes que en la función objetivo.
- Usar cotas lo más fuertes posibles para las variables continuas.
- Introducir restricciones redundantes que refuercen las relajaciones lineales.

Ejemplo. Identificación del mínimo conjunto de reacciones necesarias para el crecimiento microbiano: usualmente, solo un subconjunto de genes es esencial para el crecimiento celular. El conjunto más pequeño de genes necesarios para el crecimiento se conoce como *conjunto mínimo de genes*. Si bien muchos de estos genes requeridos no tienen una función metabólica conocida, existe un subconjunto con distintas funciones metabólicas necesarias para la formación de biomasa. La identificación de este subconjunto proporciona información sobre el conjunto mínimo de funciones metabólicas necesarias para la vida. El conjunto más pequeño de genes, y por lo tanto,

de reacciones que se necesitan para sostener el crecimiento y la producción de ATP (para el mantenimiento de las células) se conoce como el conjunto mínimo de genes metabólicos. La tarea de identificar este conjunto mínimo se puede formular como un problema MILP que satisface la restricción que indica a la producción de biomasa, que es una fracción c del máximo teórico para la cepa de tipo salvaje ($0 \leq 1 \leq c$). Consecuentemente, se define una variable binaria para cada reacción en la red de la siguiente manera:

$$y_j = \begin{cases} 1, & \text{si la reacción } j \text{ está activa} \\ 0, & \text{si la reacción } j \text{ está inactiva} \end{cases} \quad (2.40)$$

La siguiente formulación MILP identifica el conjunto de reacción mínimo para el modelo metabólico:

$$\text{Maximizar } z = \sum_{j \in J} y_j \quad (2.41.1)$$

$$\text{sujeto a} \quad (2.41.2)$$

$$\sum_{j \in J} S_{ij} v_j = 0 \quad \forall i \in I \quad (2.41.3)$$

$$LB_j y_j \leq v_j \leq UB_j y_j \quad \forall j \in J \quad (2.41.4)$$

$$v_{biomasa} \geq f v_{biomasa}^{Max WT}$$

La función objetivo minimiza el número de reacciones que deben estar activas en la red para alcanzar objetivo de velocidad de crecimiento de biomasa. Este problema de optimización es computacionalmente intensivo, ya que cientos de variables binarias pueden activarse generando millones de combinaciones capturadas en un modelo a escala genómica. El problema se vuelve desafiante a medida que el objetivo de rendimiento de biomasa se reduce, dado que más combinaciones de reacciones se convierten en soluciones factibles.

PROGRAMACIÓN NO LINEAL

Los problemas de optimización no lineal surgen frecuentemente en el análisis de redes metabólicas debido a funciones objetivos y/o restricciones no lineales. Ejemplos de este tipo de problemas incluyen a la *minimización del ajuste metabólico* (MOMA) y el análisis de flujo metabólico (MFA). A continuación describimos los fundamentos de la programación no lineal (NLP) restringida y no-restringida y los procedimientos de solución relevantes.

Optimización no lineal no restringida

Los problemas de optimización no lineal no restringida implican la maximización o minimización de una función no lineal sin ninguna restricción.

$$\text{Minimize}_{x \in \mathbb{R}^N} f(x) \quad (2.42)$$

Métodos de solución para problemas de optimización sin restricciones

Los métodos de solución de problemas de optimización sin restricciones se dividen en dos categorías principales.

1. *Métodos que utilizan evaluaciones de función:* los puntos de solución se determinan mediante el uso de evaluaciones de la función objetivo. Las iteraciones se repiten hasta que no se pueden lograr mejoras adicionales en la función objetivo o se haya alcanzado el número máximo de iteraciones. Estos métodos no siempre son eficientes pero son fáciles de implementar (Método de Newton; Método de Fibonacci).
2. *Métodos que utilizan información de la derivada:* estos métodos explotan la información contenida en las derivadas de primer y/o segundo orden para determinar la dirección de búsqueda. Las iteraciones continúan hasta que se identifica un punto crítico y se deben evaluar las condiciones de optimalidad suficientes para garantizar que el punto crítico identificado al menos sea un mínimo local. Estos métodos se pueden clasificar en *métodos de gradiente* que requieren solo derivados de primer orden para actualizar el punto de solución y métodos de segundo orden que requieren las derivadas tanto de primer orden como de segundo orden para actualizar el punto de solución.

Método de descenso más pronunciado o método de Cauchy: la idea de este método es aproximar la función objetivo mediante una línea, es decir la aproximación mediante una serie de Taylor de primer orden para moverse iterativamente a lo largo de la dirección de descenso más pronunciada, que es ortogonal a la aproximación lineal de la función, hasta que se alcanzan las condiciones de convergencia.

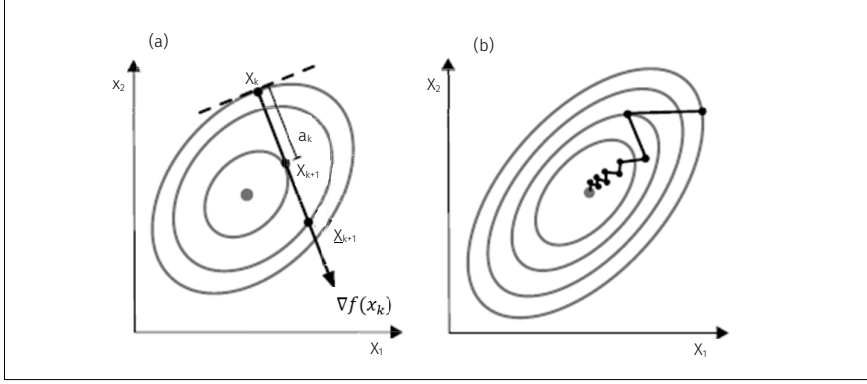


Figura 2.6. Cálculo del tamaño de paso óptimo, α_k en el método de descenso más pronunciado. El objetivo es encontrar α_k que minimice $f(x_{k+1})$. (b) iteraciones oscilantes (zigzag) en el algoritmo de descenso más pronunciado se acerca al mínimo.

El método de actualización para el método de descenso más pronunciado viene dado por la siguiente función en diferencias finitas:

$$x_{k+1} = x_k + \alpha_k d_k = x_k + (-\nabla f(x_k)) = x_k - \nabla f(x_k) \quad (2.43)$$

Método de Newton: el método de descenso más pronunciado utiliza información de las derivadas de primer orden para moverse en la dirección ortogonal a la aproximación lineal de la función objetivo en cada punto de iteración y se mueve a lo largo de la dirección de descenso más pronunciada formada por una aproximación cuadrática de la función objetivo. Asumiendo que f es dos veces diferenciable, sea q la aproximación cuadrática de la función objetivo f en un punto de iteración x_k . De la expansión de Taylor de f en x_k ; al ignorar los derivados superiores a los de segundo orden obtenemos la siguiente expresión:

$$q(x) = f(x_k) + \nabla^T f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k) \quad (2.44)$$

Una condición necesaria y suficiente para el mínimo de la aproximación cuadrática de f es $\nabla q = 0$ que combinado con la ecuación anterior permite deducir que:

$$\nabla q(x) = \nabla^T f(x_k)(x - x_k) + H(x_k)(x - x_k) = 0 \quad (2.45)$$

Esto proporciona la siguiente ecuación recursiva para el método de Newton:

$$x_{k+1} = x_k - \mathbf{H}^{-1}(x_k) \nabla^T f(x_k) \quad (2.46)$$

Otros métodos similares son el Método quasy-Newton y los métodos de gradiente conjugado.

Optimización no lineal restringida

Los problemas de optimización no lineal restringida contienen un objetivo no lineal y/o al menos una restricción no lineal. El problema de optimización se puede describir de la siguiente forma.

$$\text{Minimizar } z = f(\mathbf{x}) \quad (2.47.1)$$

sujeto a

$$g_j(\mathbf{x}) \leq 0 \quad \forall j \in J \quad (2.47.2)$$

$$h_l(\mathbf{x}) = 0 \quad \forall l \in L \quad (2.47.3)$$

$$\mathbf{x} \in \mathbf{R}^n$$

Un problema de optimización no lineal que consiste solo en restricciones de igualdad se puede expresar de la siguiente manera:

$$\text{Minimizar } z = f(\mathbf{x}) \quad (2.48.1)$$

sujeto a

$$h_l(\mathbf{x}) = 0 \quad \forall l \in L \quad (2.48.2)$$

$$\mathbf{x} \in \mathbf{R}^n$$

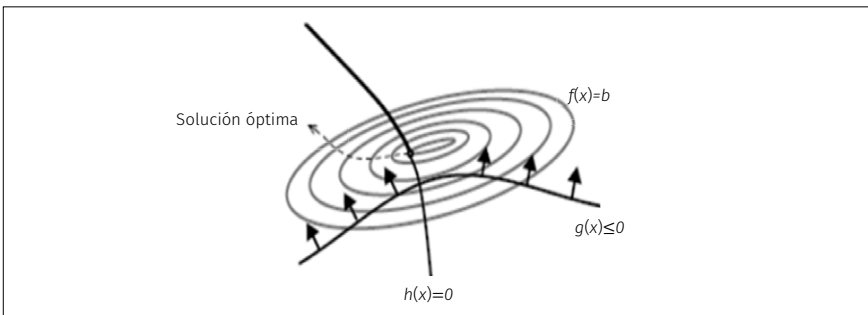


Figura 2.7. Ilustración esquemática de un problema de optimización no lineal con una restricción de desigualdad $g(x) \leq 0$, una restricción de igualdad $h(x) = 0$ y conjuntos de nivel $f(x) = b$.

Problemas no lineales con restricciones de igualdad y de desigualdad

Los problemas no lineales con restricciones de igualdad y desigualdad son la forma más general de un problema de optimización. Las restricciones de desigualdad pueden ser activas o vinculantes, en cuyo caso pueden tratarse como restricciones de igualdad o inactivas o no vinculantes, lo que implica que se pueden ignorar.

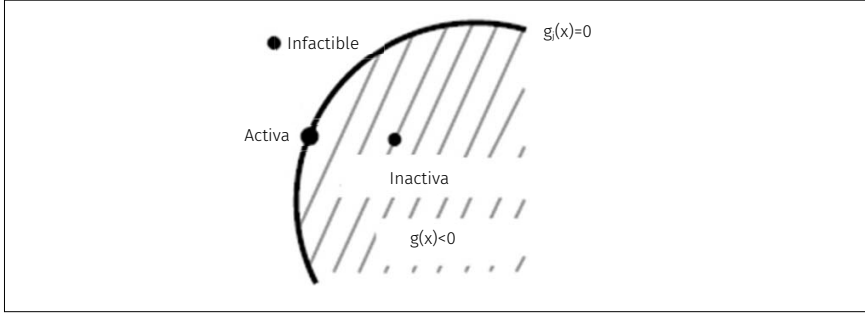


Figura 2.8. Un ejemplo de una restricción de desigualdad infactible, activa e inactiva para diferentes puntos.

$$g_j(\mathbf{x}) \leq \begin{cases} \text{Activa en } \mathbf{x} = \bar{\mathbf{x}} \text{ si } g_j(\bar{\mathbf{x}}) = 0 \\ \text{Inactiva en } \mathbf{x} = \bar{\mathbf{x}} \text{ si } g_j(\bar{\mathbf{x}}) < 0 \end{cases} \quad (2.49)$$

Si la actividad o inactividad de las restricciones de desigualdad en el punto de solución óptima se conociera a priori, entonces las restricciones inactivas podrían simplemente eliminarse y las restricciones activas podrían establecerse como igualdades. Esto transformaría al problema en un problema de optimización no lineal solo con restricciones de igualdad. El problema no lineal general asociado a la función Lagrange se define de la siguiente manera:

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{j=1}^M \mu_j g_j(\mathbf{x}) + \sum_{l=1}^L \lambda_l h_l(\mathbf{x}) \quad (2.50)$$

Aquí el elemento μ_j del vector $\boldsymbol{\mu}$ es el multiplicador de Lagrange asociado con la restricción de desigualdad $g_j(\mathbf{x}) = 0$. Los siguientes dos casos deben considerarse al evaluar la existencia de la contribución a la función de Lagrange por las restricciones de desigualdad en un punto $\mathbf{x}$:

$$\begin{aligned} \text{Si } g_j(\bar{\mathbf{x}}) < 0 \text{ entonces el multiplicador } \mu_j \text{ se hace } 0 \text{ y } \mu_j g_j(\bar{\mathbf{x}}) &= 0 \\ \text{Si } g_j(\bar{\mathbf{x}}) = 0 \text{ entonces } \mu_j > 0 \text{ pero } \mu_j g_j(\bar{\mathbf{x}}) &= 0 \end{aligned} \quad (2.51)$$

Esto implica que para cada punto factible x , el término proporcional a $g_j(x)$ es igual a cero. Esto se conoce como «condición de holgura complementaria».

Condiciones de optimalidad de Karush–Kuhn–Tucker

Las condiciones de Karush–Kuhn–Tucker (κκτ) generalizan el método de los multiplicadores de Lagrange para establecer las condiciones de optimización para problemas que involucran restricciones de igualdad y desigualdad. Sea x un mínimo local del problema no lineal donde $f(x)$, $g_j(x)$ y $h_l(x)$ son continuamente diferenciables en x . Además, suponga que x es un punto regular. Entonces existen multiplicadores únicos μ_j , $j = 1, 2, \dots, M$ y λ_l , $l = 1, 2, \dots, L$ de manera que:

$$\nabla f(\bar{x}) + \sum_{j=1}^M \mu_j \nabla g_j(\bar{x}) + \sum_{l=1}^L \lambda_l \nabla h_l(\bar{x}) = 0 \quad (2.51.1)$$

$$\nabla g_j(\bar{x}) \leq 0 \quad \forall j \in J \quad (2.51.2)$$

$$\nabla h_l(\bar{x}) = 0 \quad \forall l \in L \quad (2.51.3)$$

$$\mu_j g_j(\bar{x}) = 0 \quad \forall j \in M \quad (2.51.4)$$

$$\mu_j > 0 \quad \forall j \in M \quad (2.51.5)$$

Las restricciones (2.51.2) y (2.51.3) requieren que x sea factible para el problema no lineal y se denominan condiciones de factibilidad primaria (PF). Las restricciones (2.51.4) y (2.51.5) se denominan condiciones de doble factibilidad (DF). La restricción (2.51.4) es la condición de holgura complementaria (CS). Estas condiciones juntas forman las condiciones necesarias de κκτ. Cualquier punto $\bar{x}$ que las satisfaga se llama punto κκτ. De manera similar a un punto crítico en programación no lineal sin restricciones, no todos los puntos κκτ son necesariamente un mínimo local. Las condiciones suficientes de κκτ describen las condiciones requeridas para la optimización de un punto κκτ.

Método del gradiente reducido

El método de gradiente reducido generalizado se propuso originalmente para problemas no lineales con restricciones de igualdad y luego se generalizó para problemas con restricciones de igualdad y desigualdad. La idea básica de este método es convertir el problema original en un subproblema de optimización sin restricciones en un espacio dimensional reducido. Para este fin, se utiliza una aproximación lineal de las restricciones no lineales en cada iteración para expresar un subconjunto de variables en función de

otro subconjunto de variables independientes. Se puede demostrar que la solución de este subproblema en el espacio dimensional reducido es un punto KKT del problema original. Este subproblema se resuelve iterativamente hasta que se identifica a un punto KKT. Considere el problema en la forma de matriz de la siguiente manera.

$$\begin{array}{ll} \text{Minimizar } z = f(\mathbf{x}) & (2.52.1) \\ \text{suje } a & \end{array}$$

$$\begin{array}{ll} h_l(\mathbf{x}) = 0 & \forall l \in L \\ \mathbf{x} \in \mathbf{R}^n & \end{array} \quad (2.52.2)$$

donde $\mathbf{h}(\mathbf{x}) = [h_1(x), h_2(x), h_l(x)]$. Se supone que todas las restricciones de desigualdad se han transformado en restricciones de igualdad utilizando variables de holgura.

PROGRAMACIÓN MIXTA ENTERO-NO LINEAL

Los problemas de programación mixtos enteros no lineales (MINLP) involucran variables continuas ($\mathbf{x}$) y enteras ($\mathbf{y}$) con no linealidades en la función objetivo y/o en las restricciones. La forma general de un problema MINLP es la siguiente:

$$\begin{array}{ll} \text{Minimizar } z = f(\mathbf{x}, \mathbf{y}) & (2.53.1) \\ \text{suje } a & \end{array}$$

$$\mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0} \quad (2.53.2)$$

$$\mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{0} \quad (2.53.3)$$

$$\mathbf{x} \in \mathbf{R}^n$$

$$\mathbf{y} \in \mathbf{Z}^p$$

donde $f(\mathbf{x}, \mathbf{y})$, $\mathbf{g}(\mathbf{x}, \mathbf{y})$, y $\mathbf{h}(\mathbf{x}, \mathbf{y})$ son funciones no lineales de $\mathbf{x}$ e $\mathbf{y}$. Esta clase de problemas surgen en aplicaciones de diseño computacional con funciones objetivas no lineales (por ejemplo, en la minimización del ajuste metabólico, MOMA) con modelos que usan expresiones cinéticas para describir algunos de los flujos metabólicos. Aquí nos centramos en problemas de tipo MINLP binarios, es decir con $\mathbf{y} = \{0,1\}$ y describimos dos procedimientos de solución. Para cualquier problema MINLP, uno simplemente puede relajar las variables binarias tratándolas como variables continuas que varíen entre cero y uno y resolver el problema resultante como un problema de programación no lineal. Si los valores óptimos de las variables relajadas son todos cero o uno,

se habrá resuelto el problema MINLP original. Sin embargo, en general, esta relajación no produce soluciones enteras y por lo tanto, se deben invocar estrategias de solución «personalizadas». Los problemas MINLP se resuelven utilizando métodos estocásticos o deterministas. Los métodos estocásticos en general se usan en problemas muy grandes y no lineales que no pueden resolverse de manera óptima dentro del tiempo de resolución asignado utilizando métodos deterministas y/o cuando no se necesita una solución óptima. Los métodos estocásticos populares incluyen recocido simulado (SA), algoritmos genéticos (GA), optimización de colonias de hormigas y optimización de enjambre de partículas. En contraste, los métodos deterministas tienen criterios de finalización y propiedades de convergencia definidos. Estos métodos incluyen, entre otros, descomposición generalizada de Benders (Geoffrion, 1972), aproximación externa (Duran y Grossmann, 1986) y aproximación externa con relajación de igualdad. Todos estos métodos generan una secuencia convergente de los límites superior e inferior hasta alcanzar la solución óptima. El límite superior en todos estos métodos se encuentra fijando las variables binarias y resolviendo el problema no lineal (llamado problema primal) resultante en las variables continuas $\mathbf{x}$. El límite inferior se obtiene resolviendo una relajación del problema original, llamada problema maestro. Los métodos existentes difieren principalmente en la forma en que formulan el problema maestro. A continuación, presentamos una breve descripción de los algoritmos de descomposición de Benders generalizada y de aproximación exterior.

Descomposición de Benders generalizada

El algoritmo de descomposición de Benders generalizada puede usarse para resolver la siguiente clase de problemas MINLP:

$$\text{Minimizar } z = f(\mathbf{x}, \mathbf{y}) \quad (2.54.1)$$

$$\text{sujeto a} \quad (2.54.2)$$

$$\mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0} \quad (2.54.3)$$

$$\mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$$

$$\mathbf{x} \in \mathbf{R}^n$$

$$\mathbf{y} \in \{0,1\}^P$$

El siguiente pseudocódigo describe la lógica de operación de dicho algoritmo:

0. Contador de iteraciones $k = 1$
 Cota superior: $+\infty$
 Cota inferior: $-\infty$
 Iteración factible: $F = 0$
 Iteración infactible: $IF = 0$
 Inicializar y_k
1. Resolver el problema primal $[P^k_{\text{GBD}}]$ en $y = y_k$.
- 1.a. Si el primal es factible:
 Actualizar el conjunto de factibles $F = F \cup \{k\}$.
 Actualizar la cota superior $UB = \min(UB, f(x_k, y_k))$, siendo $x_k = [P^k_{\text{GBD}}]$.
 Usar el valor óptimo de los multiplicadores de Lagrange μ_k y λ_k para construir la función de lagrange $Lk(x, y, \mu_k, \lambda_k)$.
- 1.b. Si el primal es infactible:
 Actualizar el conjunto de infactibles $IF = IF \cup \{k\}$.
 Resolver el de factibilidad problema $[P^k_{\text{GBD}}]$ para obtener x_k, μ_k y λ_k .
 Construir la función de lagrange $Lk(x, y, \mu_k, \lambda_k)$.
2. Resolver el problema master relajado $[P^k_{\text{GBD}}]$ para obtener y_k y u .
 Actualizar la cota inferior $LB = u$.
3. Si $(UB - LB) \leq \epsilon$, donde ϵ es un valor de tolerancia o si $LB > UB$, entonces finalizar. Si no, actualizar el contador de iteraciones $k \leftarrow k + 1$ y retorne a la etapa 1.

Algoritmo de aproximación exterior

El algoritmo de aproximación exterior permite resolver la siguiente clase de problemas de tipo MINLP:

$$\text{Minimizar } z = f(x) + c^T y \quad (2.55.1)$$

sujeto a

$$g(x) + By \leq 0 \quad (2.55.2)$$

$$x \in R^n$$

$$y \in \{0,1\}^P$$

El siguiente pseudocódigo describe la lógica de operación del algoritmo de aproximación exterior:

0. Contador de iteraciones $k = 1$
Cota superior: $+\infty$
Cota inferior: $-\infty$
Iteración factible: $F = 0$
Iteración infactible: $IF = 0$
Inicializar y_k
1. Resolver el problema primal $[P_{OA}^k]$ en $y = y_k$.
 - 1.a. Si el primal es factible:
Actualizar el conjunto de factibles $F = F \cup \{k\}$.
Actualizar la cota superior $UB = \min(UB, f(x_k) + c^T y_k)$, siendo $x_k = [P_{OA}^k]$.
Linealizar $f(x)$ y $g(x)$ en x_k .
 - 1.b. Si el primal es infactible:
Actualizar el conjunto de infactibles $IF = IF \cup \{k\}$.
Resolver el de factibilidad problema $[P_{OA}^k]$ para obtener x_k .
Linealizar $g(x)$ en x_k .
2. Resolver el problema master relajado $[P_{OA}^k]$ para obtener y_{k+1} y u_k .
Actualizar la cota inferior $LB = c^T y_k + u_k$.
3. Si $(UB - LB) \leq \epsilon$, donde ϵ es un valor de tolerancia o si $LB > UB$, entonces finalizar. Si no, actualizar el contador de iteraciones $k \leftarrow k + 1$ y retorne a la etapa 1.

Nótese que si el problema maestro relajado es infactible, entonces el problema original MINLP_{OA} también es infactible.

Optimización dinámica o control óptimo

Los modelos dinámicos describen operaciones y procesos que tienen lugar en varias áreas, que incluyen a ingeniería química, bioquímica, economía, ecología y biotecnología. Muchos procesos y aplicaciones en bioquímica y biotecnología son intrínsecamente dinámicos. Dichos procesos incluyen la operación de reactores batch y semibatch que son de uso intensivo para la producción de productos químicos especializados, productos farmacéuticos de alto valor y polímeros. Para procesos continuos, la optimización dinámica se utiliza en el diseño de sistemas distribuidos, como trenes de reactores de flujo de tapón y columnas de destilación, así como en la determinación de trayectorias óptimas en la transición entre las condiciones de operación.

Los modelos matemáticos que describen problemas de optimización dinámica involucran ecuaciones algebraicas y diferenciales así como restricciones en las variables de control y en las variables de estado que son aquellas variables que evolucionan en el tiempo y caracterizan a un sistema dinámico. El desarrollo de métodos numéricos robustos, junto a una creciente capacidad computacional ha allanado el camino para la formulación y solución de problemas de optimización dinámica con aplicaciones en bioquímica y biotecnología. Consecuentemente, la optimización dinámica se ha convertido en una herramienta importante para la toma de decisiones y la optimización en operaciones biotecnológicas. Los métodos numéricos disponibles para resolver este tipo de problemas se pueden clasificar en enfoques indirectos o variacionales y enfoques directos, que a su vez se pueden dividir en secuenciales y simultáneos (Betts, 2010). En los métodos directos, el problema se discretiza y se transforma en un problema de dimensión finita.

Formulación de problemas de control óptimo

Un problema de optimización sujeto a un sistema de ecuaciones diferenciales y algebraicas (DAE), con o sin restricciones de desigualdad, se conoce como un problema de optimización dinámica o problema de control óptimo. Este problema se puede plantear de la siguiente manera.

$$\text{Minimizar } z = \phi(x(t_f)) \quad (2.56.1)$$

sujeto a

$$\frac{dx}{dt} = f(x(t), y(t), u(t), p) \leq 0 \quad (2.56.2)$$

$$h(x(t), y(t), u(t), p) = 0 \quad (2.56.3)$$

$$g(x(t), y(t), u(t), p) \leq 0 \quad (2.56.4)$$

$$h_p(x(t_f)) = 0$$

$$g_p(x(t_f)) \leq 0 \quad (2.56.5)$$

$$z^{LB} \leq x(t) \leq z^{UB} \quad (2.56.6)$$

$$y^{LB} \leq y(t) \leq y^{UB} \quad (2.56.7)$$

$$u^{LB} \leq u(t) \leq u^{UB} \quad (2.56.8)$$

$$p^{LB} \leq p(t) \leq p^{UB} \quad (2.56.9)$$

$$(2.56.10)$$

El sistema de ecuaciones diferenciales y algebraicas representa al modelo del sistema. Las ecuaciones diferenciales (2.56.2) se deducen a partir de la aplicación de principios de conservación de las cantidades fundamentales, es decir de balances de materia y energía. Las ecuaciones algebraicas (2.56.3) y

(2.56.4) corresponden a ecuaciones constitutivas, que en general se basan en leyes físicas y químicas. Las variables diferenciales de estado y algebraicas están representadas por los vectores $\mathbf{x}(t)$ e $\mathbf{y}(t)$, respectivamente. Las variables diferenciales se denominan variables de estado y corresponden a las cantidades fundamentales que se conservan. Típicamente incluyen concentración de componentes y energía interna en ejemplos de ingeniería (bio)química. Las variables algebraicas en general están relacionadas con las variables diferenciales y corresponden a propiedades físicas y químicas, velocidades de reacción y propiedades termodinámicas. El vector $\mathbf{u}(t)$ depende de t y representa a las variables de control que son aquellas que pueden ser manipuladas. El vector $\mathbf{p}$ corresponde a parámetros, que son valores que no dependen de t . La función objetivo escalar Φ se calcula al final del horizonte de integración. Si el objetivo es una función de varias variables algebraicas y de estado, se agrega una nueva ecuación algebraica al modelo, para definir la nueva variable algebraica como función objetivo. Si la función objetivo Φ incluye un término integral sobre la variable independiente t :

$$\phi = \int_{t_0}^{t_f} \Psi(\mathbf{z}(t), \mathbf{y}(t), \mathbf{v}(t), \mathbf{p}) \quad (2.57)$$

se puede reescribir en forma escalar definiendo una nueva variable diferencial z_0 reemplazando el término integral en la función objetivo y agregando una ecuación diferencial adicional en la que el lado derecho es el término integral con la condición inicial $z_0 = 0$, de la siguiente manera:

$$\phi = z_0(t_f) \quad (2.58.1)$$

$$\frac{dz_0}{dt} = \Psi(\mathbf{z}(t), \mathbf{y}(t), \mathbf{u}(t), \mathbf{p}) \quad (2.58.2)$$

$$z_0(0) = 0 \quad (2.58.3)$$

Los límites inferior y superior para cada una de las variables se denotan con los superíndices LB y UB . Las restricciones en puntos que corresponden a condiciones que deben cumplirse en un determinado punto del horizonte temporal de integración; pueden ser igualdades o desigualdades y se les conoce como restricciones de punto final si se imponen al tiempo final del horizonte t_f . Si estas condiciones adicionales se imponen en cualquier otro punto del horizonte, se denominan restricciones de punto interior. Las restricciones de trayectoria $\mathbf{g}(\cdot)$ representan condiciones que deben cumplirse en todo el horizonte temporal. Estas restricciones de desigualdad aumentan al número de ecuaciones algebraicas $\mathbf{h}(\cdot)$. Se asume que el comportamiento del sistema está definido en términos de ecuaciones diferenciales y algebraicas, en una forma semiexplícita, donde $\mathbf{y}(t)$ puede eliminarse implícitamente a

través de las restricciones algebraicas $h(\cdot)$. Los problemas de optimización dinámica tienen una naturaleza de dimensión infinita y las metodologías de solución los transforman en un problema de dimensión finita. Con ecuaciones diferenciales no lineales, el problema no es convexo y puede tener múltiples óptimos locales, por lo que cualquier solución para el problema de optimización dinámica será localmente óptima.

Métodos de solución para problemas de optimización dinámica

Los problemas de optimización dinámica, también conocidos como problemas de control óptimo, generalmente no son lineales y la mayoría de ellos no tienen soluciones analíticas. Por ello se resuelven mediante métodos numéricos. Existen dos tipos de enfoques numéricos para resolver problemas de optimización dinámica: métodos indirectos y directos. Para una descripción detallada recomendamos el libro de Betts (2010) que describe los detalles de funcionamiento de estos métodos.

Métodos indirectos. Los enfoques indirectos o variacionales se basan en el principio máximo de Pontryagin (1962), en el que las condiciones de optimización se derivan de la aplicación de cálculos de variaciones. Para problemas sin restricciones de desigualdad, las condiciones de optimización pueden escribirse como un conjunto de ecuaciones diferenciales-algebraicas y resolverse como un problema de valor límite de dos puntos. Si hay restricciones de desigualdad para la trayectoria, se requieren condiciones adicionales, y la determinación de los puntos de entrada y salida para las restricciones activas a lo largo del horizonte de integración genera un problema combinatorio, que por lo general es difícil de resolver.

Métodos directos. Los enfoques directos transforman el problema de control óptimo en un problema de programación no lineal (PNL). En estos métodos, las variables de control y de estado se aproximan mediante funciones apropiadas (usualmente polinomios o parametrizaciones constantes por partes). Los coeficientes de las aproximaciones de funciones se tratan como variables de optimización y el problema se transcribe en un problema de programación no lineal.

Métodos de parametrización del vector de control. En estos enfoques, las variables de control se aproximan por constante por partes, mediante funciones lineales o funciones polinómicas, en un número específico de intervalos de control. Se formula un problema NLP a nivel externo con coeficientes de estos polinomios como variables de optimización. Con perfiles fijos para variables de control no hay grados de libertad, y el sistema de ecuaciones

diferenciales se puede resolver con cualquier solucionador estándar de ecuaciones diferenciales.

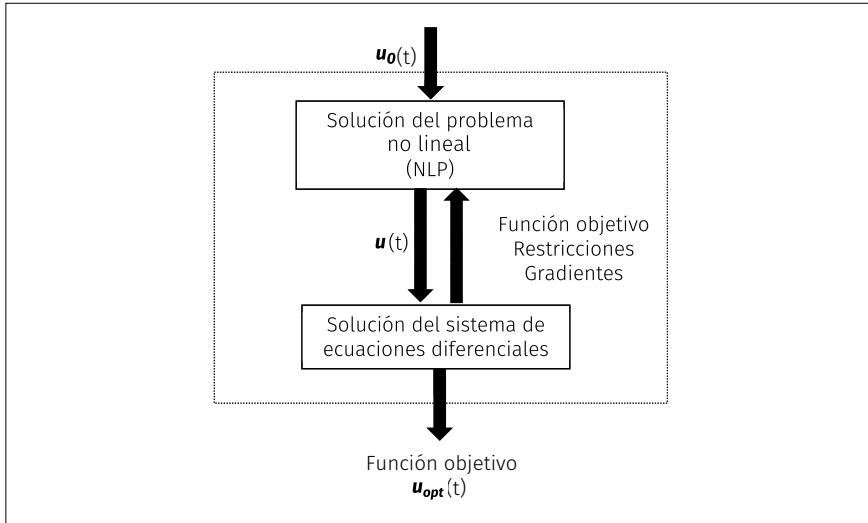


Figura 2.9. Método de parametrización del vector de control.

Como el número de variables de control es por lo general pequeño en comparación con el número total de variables, el problema no lineal resultante tiene un número relativamente pequeño de variables a optimizar. Dependiendo del grado de discretización, se plantearon estrategias de un solo «disparo» o de «múltiples disparos» (Betts, 2010). En el enfoque de disparo único, el sistema de ecuaciones diferenciales con valores iniciales en $t = 0$ se integra a lo largo de todo el horizonte de integración para perfiles fijos de las variables de control. Por lo tanto, los perfiles para las variables algebraicas y diferenciales se determinan junto con el valor de la función objetivo y las restricciones. En esta etapa, las derivadas parciales de variables diferenciales y algebraicas con respecto a las variables de optimización (por ejemplo, $\mathbf{p}$ y coeficientes polinomiales para $\mathbf{u}(t)$) deben determinarse a lo largo del horizonte de integración. Estos perfiles de sensibilidad se obtienen integrando las llamadas *ecuaciones de sensibilidad* junto con el sistema DAE original. La solución de las ecuaciones de sensibilidad es a menudo el paso más costoso en el enfoque de disparo único.

Enfoques simultáneos. En los métodos simultáneos las variables de control y las variables de estado se aproximan en cada intervalo de integración mediante polinomios por partes, cuyos coeficientes son las variables de optimización para un problema no lineal de gran escala. De esta manera, los

sistemas de ecuaciones diferenciales y algebraicas se discretizan y expresan como ecuaciones algebraicas a lo largo del horizonte de integración. Para la discretización se usa comúnmente un método eficiente denominado colocación ortogonal en elementos finitos. Con la colocación en elementos finitos, los perfiles de las variables de estado y de control se aproximan mediante una familia de polinomios en una malla de intervalos de tiempo definida mediante puntos de colocación ($t_0 < t_1 < \dots < t_N = t_f$). Estos polinomios se pueden representar como series de potencia, sumas de polinomios ortogonales o polinomios de interpolación de Lagrange.

Durante la última década, los métodos pseudoespectrales han cobrado importancia en la solución numérica de problemas de control óptimo. Los métodos pseudoespectrales (Rao *et al.*, 2012) son métodos de colocación directa donde el problema de control óptimo se transcribe en un problema de programación no lineal (NLP) mediante la parametrización del vector de estados y el de control usando polinomios globales y colocando las ecuaciones algebraicas diferenciales usando nodos obtenidos de una cuadratura gaussiana. Cabe señalar que algunos investigadores prefieren el término colocación ortogonal pero los términos colocación pseudoespectral y ortogonal tienen el mismo significado. Los tres conjuntos de puntos de colocación más utilizados son los puntos Legendre–Gauss, Legendre–Gauss–Radau y Legendre–Gauss–Lobatto.

Actividades

1. Reformular al siguiente problema de minimización de diferencias entre valores experimentales y predichos para expresarlo como un problema lineal.

$$\text{Minimizar } \sum_{j \in J} |x_j - a_j|$$

Sujeto a:

$$\sum_{j \in J} S_{ij} x_j = 0 \quad \forall i \in I$$

$$LB_j \leq x_j \leq UB_j \quad \forall j \in J$$

$$x_j \in R$$

2. Transformar el siguiente problema de programación no lineal en uno problema lineal:

$$\begin{array}{ll}
 \text{Minimizar} & \frac{x_1}{x_2} \\
 \text{Sujeto a:} & \\
 \sum_{j \in J} S_{ij} x_j = 0 & \forall i \in I \\
 LB_j \leq x_j \leq UB_j & \forall j \in J \\
 x_j \in R & \forall j \in J
 \end{array}$$

Donde los conjuntos, variables y parámetros tienen el mismo significado que en el ejercicio anterior y las cotas LB y UB son no-negativas.

REFERENCIAS BIBLIOGRÁFICAS

- BAZARAA, M.; JARVIS, J. & SHERALI, H. (2010). *Linear programming and network flows* (4th edition). Hoboken. N. J. John Wiley & Sons.
- BETTS, J. (2010). *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming (Advances in Design and Control)* (2nd edition). Series Number 19.
- DANTZIG, G. & ORDEN, A. (1955). *The generalized simplex method for minimizing a linear form under linear inequality restraints*. Pacific J Mathemat., 5 (2):187-195.
- DURAN, M. & GROSSMANN, I. (1986). *An outer-approximation algorithm for a class of mixed-integer nonlinear programs*. Math Program, 36 (3):307-339.
- FOURER, R.; GAY, D. & KERNIGHAN, B. (2003). *AMPL: a modeling language for mathematical programming* (2nd edition). Pacific Grove, CA: Thomson/Brooks/Cole.
- GEOFFRION, A. (1972). *Generalized Benders Decomposition*. J Optim. Theory App., 19 (4): 237-260.
- IGNIZIO, J. & CAVALIER, T. (1994). *Linear programming*. Englewood Cliffs, N.J. Prentice Hall.
- PONTRYAGIN, L.; BOLTANSKII, V. (...) MISHECHENKO, E. (1962). *The Mathematical Theory of Optimal Processes*. Interscience Publishers John Wiley & Sons, Inc. New York-London, Translated from the Russian by K.N., Trirkoff.

RAO, A.; BENSON, D. (...) **HUNTINGTON, G. (2012).** *User's Manual for GPOPS Version 5.0.*
RARDIN, R. (1998). *Optimization in operations research.* Upper Saddle River, N.J. Prentice Hall.

Lecturas adicionales

MARANAS, C. & ZOMORRODI, A. (2016). *Optimization Methods in Metabolic Networks.*
John Wiley & Sons.
BRYSON, A. & HO Y. CH. (1975). *Applied Optimal Control. Optimization, estimation and Control.* John Wiley & Sons.

3 Software

Muchas herramientas computacionales en forma de lenguajes de programación y paquetes de software permiten procesar y visualizar grandes cantidades de datos experimentales y bibliográficos referidos a procesos bioquímicos y biotecnológicos. Las herramientas se pueden dividir a grandes rasgos en programas de propósito general y softwares especializados. Como lenguajes de propósito general podemos mencionar a MATLAB, GNU Octave o SciLab que son paquetes muy poderosos para el análisis numérico, simbólico y visual de datos para problemas matemáticos generales. Sin embargo, la generalidad también puede ser un limitante para su uso porque tienen una curva de aprendizaje pronunciada y requieren un esfuerzo considerable para comenzar a utilizarlos. Por lo tanto, se han desarrollado herramientas especializadas en la aplicación de técnicas de optimización que tienen un rango de aplicación más restringido, pero que son mucho más fáciles de usar. Dentro de esta categoría mencionamos a las herramientas de programación matemática y modelado como GAMS y PYOMO que fueron utilizadas en una plétora de aplicaciones. También existen herramientas computacionales especializadas en la programación de problemas de control óptimo y su transcripción automática en problemas de programación no lineal como GPOPS e ICLOCS. Un problema que surgió con la multitud de herramientas computacionales es la falta de compatibilidad de modelos. Un modelo desarrollado con un programa a menudo tiene que ser reimplementado para su uso con una herramienta computacional diferente. Este problema se aborda con el desarrollo de lenguajes de intercambio de modelos e interfaces. El lenguaje SBML se ha convertido en un estándar de facto en biología de sistemas para este propósito con actualmente más de 280 herramientas que lo respaldan. También existen interfaces que permiten la comunicación entre GAMS y MATLAB.

PROGRAMAS DE CÁLCULO NUMÉRICO DE ALTO NIVEL

MATLAB®

MATLAB (The MathWorks Inc., 1984–2026) es la abreviatura de MATRIX LABORATORY, laboratorio de matrices, es una herramienta de software de cálculo numérico que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). MATLAB está disponible para las plataformas Unix, Windows, Mac OS X y GNU/Linux. Entre sus prestaciones básicas se hallan la manipulación de matrices, la representación de datos y funciones, el desarrollo e implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware para adquisición y procesamiento de datos. MATLAB dispone de dos herramientas adicionales que expanden sus prestaciones: *Simulink* (plataforma de simulación multidominio) y *GUIDE* (editor de interfaces de usuario – GUI). Además, las capacidades de MATLAB se pueden ampliar con las cajas de herramientas o «toolboxes» y «apps» especializadas en ajuste de curvas, optimización procesamiento de señales, biología computacional entre otras; y las de *Simulink* con los paquetes de bloques o «blocksets». MATLAB es un software muy usado en universidades y centros de investigación y desarrollo. En los últimos años ha aumentado el número de prestaciones, como la de programar directamente procesadores digitales de señales o crear código VHDL. El costo de adquisición de MATLAB depende de las prestaciones y cajas de herramientas solicitadas pero existe una versión estudiantil a la que se puede acceder a costo reducido.

GNU Octave

Octave o GNU Octave (Eaton *et al.*, 2023) es un programa libre o *freeware* para realizar cálculos numéricos. Como su nombre indica, es parte del proyecto GNU. Es considerado el equivalente libre de MATLAB. Entre varias características que comparten, se puede destacar que ambos ofrecen un intérprete, permitiendo ejecutar órdenes en modo interactivo. Octave está orientado al análisis numérico. El proyecto fue creado alrededor del año 1988, pero con una finalidad diferente: ser utilizado en un curso de diseño de reactores químicos. Posteriormente, en el año 1992, se decidió extenderlo, y comenzó su desarrollo a cargo de John W. Eaton. La primera versión fue lanzada el 4 de enero de 1993. Un año más tarde, el 17 de febrero de 1994, apareció la versión 1.0. El nombre surge de *Octave Levenspiel*, profesor de uno de los

autores y conocido por hacer buenas aproximaciones, por medio de cálculos elementales, a problemas numéricos de ingeniería química.

SciLab

Scilab (INRIA ENPC, 1989–2024) es un software para análisis numérico, con un lenguaje de programación de alto nivel para cálculo científico. Es desarrollado por Scilab Enterprises, bajo la licencia CecILL, compatible con la GNU General Public License. Las características de Scilab incluyen análisis numérico, visualización 2-D y 3-D, optimización, análisis estadístico, diseño y análisis de sistemas dinámicos, procesamiento de señales, e interfaces con Fortran, Java, C y C++. Mientras que la herramienta Xcos permite una interfaz gráfica para el diseño de modelos.

PROGRAMAS DE MODELADO ALGEBRAICO

GAMS

General Algebraic Modeling System (GAMS Development Corporation, 1976–2026) es un software de alto nivel para el modelado de sistema con el objetivo de desarrollar y resolver fácilmente modelos matemáticos utilizando una serie de resolvers incorporados. GAMS está diseñado para modelar y resolver problemas lineales, no lineales y de optimización entera mixta entero–lineal y mixta entero–no lineal. El sistema está diseñado para aplicaciones de modelado a gran escala y permite al usuario construir grandes modelos. El sistema está disponible para su uso en distintas plataformas informáticas y dispone de numerosas interfaces con programas del paquete *Office* para el intercambio de información. GAMS contiene entornos de desarrollo integrado *IDE* y *STUDIO* que están conectados a un grupo de resolvers de terceros. Entre estos solucionadores se mencionan a BARON, CONOPT, CPLEX, DICOPT, Gurobi, Mosek, SNOPT y XPRESS.

GAMS permite a los usuarios implementar algoritmos híbridos que combinen diferentes solucionadores y algoritmos de descomposición complejos. Los modelos se describen en declaraciones algebraicas concisas y legibles con facilidad. A pesar de que fue diseñado en principio para aplicaciones relacionadas con la economía y ciencias de la administración, tiene una gran comunidad de usuarios dedicados a las ingenierías y las ciencias duras. Los

impulsores del desarrollo de GAMS fueron usuarios de programación matemática que veían a la optimización como un marco útil para la solución de problemas de ciencia e ingeniería la vida real. El impulso de GAMS para el desarrollo surgió de la experiencia de un grupo de economistas del Banco Mundial durante la década de 1970.

PYOMO

(Pyomo—Optimization Modeling in Python (2008–2021; Bynum *et al.*, 2021) es un paquete de software de código abierto basado en Python que contiene diversas capacidades de optimización para formular, resolver y analizar modelos de optimización. Una capacidad central de Pyomo es la de desarrollar modelos y aplicaciones de optimización estructuradas. Pyomo se puede utilizar para definir problemas simbólicos, crear instancias de problemas específicos y resolver estas instancias utilizando solucionadores comerciales y de código abierto. Pyomo integra un lenguaje de programación de alto nivel con todas las funciones que proporciona un conjunto de bibliotecas de soporte, lo que distingue a Pyomo de otros lenguajes de modelado algebraico como AMPL, AIMMS y GAMS. Pyomo puede resolver una amplia gama de problemas, que incluyen programación lineal, programación cuadrática, programación no lineal, programación lineal entera mixta, programación cuadrática de enteros mixtos, programación no lineal de enteros mixtos, programación estocástica, programación disyuntiva generalizada, ecuaciones algebraicas diferenciales, programación binivel y programas matemáticos con restricciones de equilibrio. Pyomo admite capacidades de secuencias de comandos y análisis iterativos dentro de un lenguaje de programación con todas las funciones. Además, Pyomo también ha demostrado ser eficaz para desarrollar herramientas de análisis y optimización de alto nivel. Por ejemplo, el paquete PySP proporciona solucionadores genéricos para programación estocástica. PySP aprovecha el hecho de que los objetos de modelado de Pyomo están integrados dentro de un lenguaje de programación de alto nivel con todas las funciones, lo que permite la paralelización transparente de subproblemas utilizando bibliotecas de comunicación paralela de Python. Pyomo, que se lanzó con anterioridad como la biblioteca de software Coopr Pyomo, fue desarrollado originalmente por investigadores del Centro de Investigación en Computación de los Laboratorios Nacionales Sandía y es un proyecto COIN-OR.

SOFTWARE PARA RESOLUCIÓN DE PROBLEMAS DE CONTROL ÓPTIMO

GPOPS

GPOPS (GPOPS-II: Next-Generation Optimal Control Software; Rao *et al.*, 2012) es un programa desarrollado en MATLAB para resolver problemas muy complejos de control óptimo que involucran a múltiples etapas. Ha sido evaluado y revisado por lo cuál sus resultados pueden ser considerados como muy fiables y ha sido incluido en un paquete de software denominado *Collected Algorithms of the Association of Computing Machinery* (CALGO). Las ventajas de esta herramienta computacional son:

1. Un alto nivel de modularidad que permite plantear funciones objetivo, ecuaciones diferenciales, restricciones y funciones de acople por separado. Esto facilita el planteo, modificación y mejora de los problemas de interés.
2. Utiliza una interface multinivel para especificar la información de cada fase del problema, planteándolo en forma compacta e intuitiva.
3. Acumula los resultados en una estructura de datos que a posteriori puede ser procesada por separado en MATLAB para su posterior interpretación.

Sitio web: <https://www.gpops2.com/>

ICLOCS

ICLOCS (Imperial College London Optimal Control; Falugi *et al.*, 2010) es un software de control óptimo desarrollado por el Imperial College de Londres. ICLOCS proporciona un punto de partida didáctico para resolver problemas de control óptimo de diferente naturaleza, con requisitos mínimos de experiencia del usuario. Otros solucionadores pueden ser más rápidos en algunos problemas. Las características principales de este software son:

1. Transcripción de los problemas de control óptimo en problemas de programación no lineal (NLP) grandes pero dispersos.
2. Amplia variedad de métodos de discretización numérica para una rápida convergencia y alta precisión.
3. Se puede elegir el solucionador favorito para resolver problemas mal condicionados.
4. Permite suministrar información sobre derivada a través de medios analíticos y numéricos.

5. Interface con Simulink para realizar simulaciones.
Sitio web: <http://www.ee.ic.ac.uk/ICLOCS/default.htm>

Actividades

1. Resolver, utilizando GAMS el siguiente problema lineal:

$$\text{Maximizar } 4x_1 + 3x_2$$

sujeto a:

$$x_1 + 2x_2 \leq 7$$

$$-2x_1 - x_2 \leq 5$$

$$5x_1 + 2x_2 \leq 16$$

$$x_1, x_2 \geq 0$$

2. Resolver, utilizando GAMS el siguiente problema lineal:

$$\text{Minimizar } 3x_1 + x_2 + 31x_3$$

sujeto a:

$$6x_1 + 9x_2 + 26x_3 \leq 26$$

$$3x_1 + 2x_2 - 26x_3 \leq 20$$

$$x_1, x_2, x_3 \geq 0$$

- 2.a. Escribir al dual de este problema.
2.b. Resolver al problema dual utilizando GAMS y comparar el valor de la función objetivo.
3. Resolver, utilizando GAMS, el siguiente problema entero:

$$\text{Minimizar } -y_1 - x_2$$

sujeto a:

$$-2y_1 - 2y_2 \leq 1$$

$$16y_1 - 14y_2 \leq 7$$

$$y_1, y_2 \in \mathbb{Z}^+$$

4. Resolver el siguiente problema mixto entero lineal:

$$\text{Minimizar } 4y_1 - 6x_2$$

sujeto a:

$$-y_1 + y_2 \leq 1$$

$$y_1 - 3y_2 \leq 9$$

$$3y_1 + y_2 \leq 15$$

$$y_1, y_2 \in \mathbb{Z}^+$$

5. Resolver los siguientes problemas de optimización no lineal:

$$\begin{aligned} & \text{Maximizar } 4x_1 + 2x_2 - 0.5x_1^2 - 0.25x_2^2 \\ & \text{sujeto a:} \\ & 480x_1 + 300x_2 \leq 24000 \\ & x_1, x_2 \geq 0 \end{aligned}$$

$$\begin{aligned} & \text{Minimizar } 4x_1^2 + 2x_1^2 + x_3^2 + 12(x_1 - 50) \\ & \quad + 12(x_1 + x_2 - 80) \\ & \text{sujeto a:} \\ & x_1 \geq 50 \\ & x_1 + x_2 \geq 100 \\ & x_1 + x_2 + x_3 \geq 150 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

6. Resolver los siguientes problemas de control óptimo.

$$\begin{aligned} & \text{Minimizar } 0.5 \int_0^{t_f} (x^2 + u^2 dt) \\ & \text{sujeto a:} \\ & \frac{dx}{dt} = -x^3 + u \\ & x(0) = 1.5 \\ & x(t_f) = 1 \\ & t_f = 50 \end{aligned}$$

$$\text{Minimizar } x_3(t_f)$$

$$\begin{aligned} & \frac{dx_1}{dt} = x_2 \\ & \frac{dx_2}{dt} = u \\ & \frac{dx_3}{dt} = \frac{1}{2}u^2 \\ & 0 \leq x_1(t) \leq 1/9 \\ & x_1(0) = 0 \\ & x_2(0) = 1 \\ & x_3(0) = 0 \\ & x_1(t_f) = 0 \\ & x_2(t_f) = -1 \end{aligned}$$

REFERENCIAS BIBLIOGRÁFICAS

- BYNUM M.; HACKEBEIL, G. (...) WOODRUFF, D. (2021).** *Pyomo – Optimization Modeling in Python* (3rd. edition), Vol. 67. Springer. <https://www.pyomo.org/>
- EATON, J.; BATEMAN, D. (...) WEHBRING, R. (2023).** GNU Octave version 8.3.0 manual: a high-level interactive language for numerical computations. <https://octave.org/>
- FALUGI, P.; KERRIGAN, E. & VAN WYK, E. (2010).** Imperial College London Optimal Control Software User Guide (ICLOCS). Department of Electrical & Electronic Engineering, Imperial College London, UK. <http://www.ee.ic.ac.uk/ICLOCS/>
- GAMS DEVELOPMENT CORPORATION (1976–2026).** *General Algebraic Modeling System* (GAMS). Release 50.1.0, Fairfax, VA, USA.
- PATTERSON, M. & RAO, A. (2013).** GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using HP-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming. *ACM Transactions on Mathematical Software*. <https://www.gpops2.com/>
- RAO, A.; BENSON, D. (...) HUNTINGTON, G. (2012).** User's Manual for GPOPS Version 5.0.
- SCILAB OPEN SOURCE SOFTWARE FOR NUMERICAL COMPUTATION.** INRIA ENPC. <https://www.scilab.org/>
- THE MATH WORKS, INC. (1984–2026).** MATLAB, Natick, MA: The Math Works, Inc. <https://www.mathworks.com/>

Lecturas adicionales

- ANITA, S.; ARNAUTU, V. & CAPASO, V. (2010).** *An Introduction to Optimal Control Problems in Life Sciences and Economics*. Birkhäuser.
- BURNSTEIN, L. (2011).** *MATLAB in Bioscience and Biotechnology*. Biohealthcare Publishing Limited.
- LENHART, S. & WORKMAN, J. (2007).** *Optimal control applied to Biological Models*. Chapman and Hall SRC.

PARTE 2

Modelado matemático
de sistemas metabólicos,
bioquímicos y biotecnológicos

4 Modelado a escala microscópica

Durante las últimas dos décadas, la formulación de conceptos, matemáticos y de modelos computacionales se han vuelto cada vez más familiares en bioquímica. Los conceptos de la física y la ingeniería se se emplea usualmente en investigación biológica y biotecnológica. Esto implica que la bioquímica va en camino a convertirse en una ciencia cuantitativa tal como la física o la química. Estos conceptos requieren de predicciones matemáticas para complementar las mediciones biológicas y derivar de ellas predicciones útiles. En este capítulo presentaremos una introducción a los conceptos de modelado a escala microscópica y explicaremos la forma en la que se formulan modelos de redes metabólicas. En consecuencia, se brindará una descripción general de los enfoques de modelado utilizados con frecuencia en la biología de sistemas tal como los modelos basados en redes y en reglas lógicas.

REPRESENTACIÓN TOPOLÓGICA DE REDES METABÓLICAS

Los modelos de las células evolucionaron desde modelos compuestos por unas pocas reacciones a modelos bioquímicos de redes de reacciones que involucran a cientos de ellas y recientemente alcanzaron el nivel del genoma completo que comprenden a miles o decenas de miles de reacciones. Se han propuesto varios enfoques matemáticos para integrar las estequiometrias, las cinéticas de reacción y los mecanismos de regulación de estas redes a escala genómica. El poder predictivo que ofrecen los modelos de redes metabólicas es, por una cuestión de escala, superior al de modelos no estructurados y modelos estructurados de baja dimensión.

El metabolismo de un organismo vivo es un sistema muy complejo incluso en organismos unicelulares. El modelado metabólico es particularmente desafiante porque muchas funciones celulares son el resultado de una combinación de diversos factores. Por lo tanto, desentrañar las relaciones causales y las correlaciones entre un fenómeno observado y sus factores asociados requiere un modelo capaz de describir al biosistema subyacente, a la red de interacciones bioquímicas y a su conectividad y regulación.

Las ventajas y desventajas de un modelo están determinadas por su adecuación para describir un proceso biológico y además, es inevitable un compromiso entre la corrección del modelo y la dimensión matemática del mismo. Por lo tanto, se deben considerar suposiciones simplificadoras para hacer que un modelo metabólico sea definido por su objetivo y que sea compatible con respecto a la información proporcionada por los datos experimentales disponibles. Las suposiciones típicas se refieren al nivel de homogeneidad de los compartimentos celulares, de diversidad en las escalas de tiempo biológicas y de la distribución, entre otras. Estos supuestos son los principales factores que determinan la diferencia entre una observación y el resultado de la simulación de un modelo. El modelo metabólico también puede mejorar el conocimiento con respecto al metabolismo celular pues puede evolucionar mediante correcciones. Cuando se requiere intervenir en el funcionamiento celular, un modelo metabólico puede contribuir a descubrir nuevas estrategias de intervención y a organizar la información acumulada en un cuerpo coherente de conocimiento. El modelado metabólico ha ganado importancia en la optimización de bioprocesos. Los modelos metabólicos se dividen en cuatro categorías según la capacidad del modelo para distinguir entre subpoblaciones de biomasa (no segregadas/segregadas) y el grado agregación de los componentes bioquímicos intracelulares:

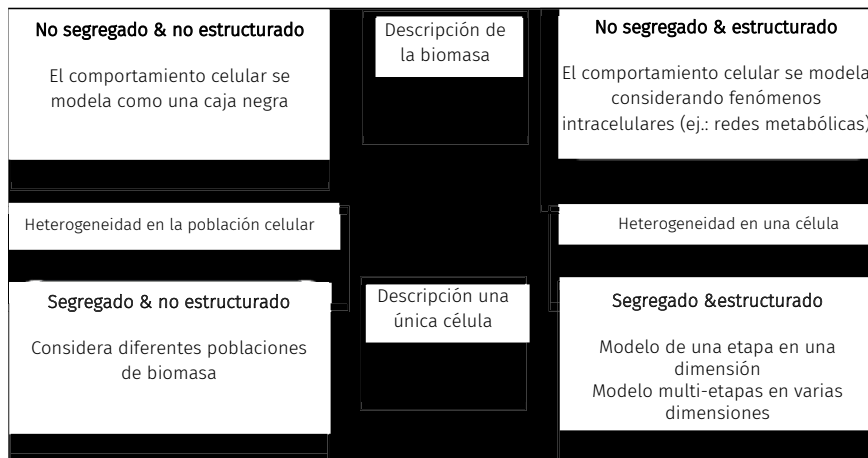


Figura 4.1. Clasificación de modelos matemáticos.

Además, desde una perspectiva dinámica de una red metabólica, los modelos pueden analizarse en estado estacionario o en estado transiente. Consideremos al ejemplo la red de la figura 2.

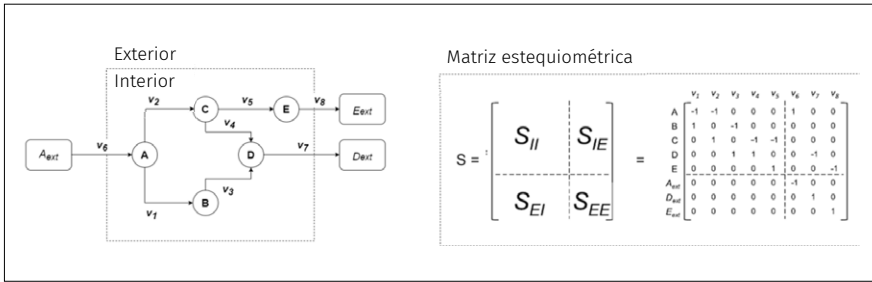


Figura 4.2. Partición de una matriz estequiométrica: (a) Representación gráfica de la red metabólica. (b) Matriz estequiométrica dividida en las cuatro submatrices características. S_{II} : Coeficientes de reacción de metabolitos intracelulares en reacciones intracelulares. S_{IE} : Coeficientes de reacción de metabolitos intracelulares en reacciones de intercambio. S_{EI} : Los coeficientes de reacción de metabolitos extracelulares en reacciones intracelulares son siempre cero. S_{EE} : Los coeficientes de reacción de metabolitos extracelulares en reacciones de intercambio siempre están en la diagonal.

MODELADO MATEMÁTICO DEL METABOLISMO

La gran diversidad funcional de enzimas presentes en los organismos vivos les permite catalizar miles de diferentes reacciones metabólicas convirtiendo a cientos de metabolitos en cientos de productos. El potencial de las funciones enzimáticas da lugar a una compleja red de reacciones en las que se sintetizan los precursores de la biomasa, y se genera ATP. El modelado matemático y las simulaciones computacionales juegan un papel clave para evaluar las posibles distribuciones de flujo metabólico que puede lograr un organismo bajo diferentes condiciones ambientales. Además, estos modelos pueden ser utilizados para identificar metodologías en ingeniería metabólica que consideren la eliminación o adición de genes o la modulación de sus niveles de expresión.

Modelado cinético del metabolismo

El modelado cinético permite cuantificar los flujos de reacción como funciones de concentraciones de metabolitos, de niveles enzimáticos y de parámetros que representan la actividad de las enzimas, la saturación y la

regulación alostérica. Diferentes expresiones tales como las cinéticas de tipo Michaelis–Menten o de Hill se pueden usar para describir a las velocidades de reacción. Al ensamblar las expresiones cinéticas, usualmente se obtiene un sistema de ecuaciones diferenciales no lineales que representan el principio de conservación de la materia para cada metabolito. Con su resolución se obtiene un perfil variante en el tiempo de las concentraciones de los metabolitos y de sus flujos. Hoy, los modelos cinéticos de metabolismos microbianos están disponibles solo para un número muy limitado de casos y cubren generalmente solo a una porción limitada del metabolismo.

Modelado del metabolismo basado en estequiometría

En contraste con el modelado cinético, el modelado basado en estequiometría se basa en un sistema de ecuaciones algebraicas que describe la conservación de la materia para cada metabolito considerando la hipótesis de estado pseudoestacionario. Durante la última década, los modelos estequiométricos metabólicos a escala genómica (gsm) han permitido reconstruir las redes metabólicas de numerosos microorganismos. La naturaleza de estos modelos permite la evaluación de límites teóricos de metabolismo metabólico y la identificación de estrategias de ingeniería genética mediante métodos de análisis de restricciones. Métodos como el análisis de balance de flujo (FBA) utilizan una cantidad mínima de conocimiento biológico y de datos para evaluar el rango de fenotipos metabólicos factibles. Además, la linealidad de este tipo de modelos matemáticos permite un cómputo casi instantáneo de los resultados en comparación con los modelos cinéticos. Esta simplicidad, sin embargo, ocurre a expensas de no poder capturar información sobre concentración intracelular de metabolitos, el efecto de la saturación de enzimas o las no linealidades debidas a la cinética y los efectos regulatorios.

REPRESENTACIÓN MATEMÁTICA DE REDES METABÓLICAS

La matriz estequiométrica de una red metabólica

La matriz estequiométrica de una red metabólica con m metabolitos y n reacciones se define por medio de la siguiente matriz:

$$S_{m \times n} = (S_{i,j})_{m \times n} \quad \forall i \in I = \{1, \dots, m\}, j \in J = \{1, \dots, n\} \quad (4.1)$$

donde la i -ésima fila consta de coeficientes de reacción para el metabolito i con respecto a todas las reacciones de la red y s_{ij} es el coeficiente del metabolito i en la reacción j . Un coeficiente s_{ij} negativo implica que el metabolito i está en el lado izquierdo de la reacción j y por lo tanto se consume. Si es positivo está en el lado derecho y se produce en la reacción j . Si es cero no participa en la reacción j . El balance de materia para cada metabolito i vendrá dado por:

$$\frac{dx_i}{dt} = \sum_{j=1}^n s_{ij} v_j \quad \forall i \in I \quad (4.2)$$

Por lo tanto, tenemos un vector columna para cada reacción y la ecuación anterior puede escribirse en notación matricial–vectorial para toda la red metabólica de la siguiente manera:

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v} \quad (4.3)$$

Donde la matriz estequiométrica contiene las cuatro particiones previamente explicitadas en la figura 4.2.

El metabolismo como un sistema con restricciones

La hipótesis de estado estacionario del metabolismo es la base conceptual para estimar los flujos de las reacciones teóricamente activas y para estimar las velocidades de flujos metabólicos a partir de la estequiometría de la red de reacción y de los datos experimentales de consumo de metabolitos extracelulares. La restricción dada por el principio de conservación de la materia en estado estacionario se formula a través de la matriz estequiométrica para una red de reacciones bioquímicas de cualquier tamaño, y esta restricción tiene un papel central en la determinación del espacio de solución en muchos enfoques de análisis metabólico. La suposición de que los cambios de las velocidades de flujos intracelulares ocurren con una dinámica más rápida que la de los cambios de las velocidades extracelulares respalda la validez de la suposición de estado estacionario. Este supuesto es válido en la fase de crecimiento exponencial a una velocidad específica constante en cultivos batch y en cultivos continuos que operan en estado estacionario. La ventaja del modelado basado en restricciones es que no necesita lidiar con la cinética enzimática y con los mecanismos reguladores. A continuación, revisamos los enfoques basados en restricciones más relevantes.

ANÁLISIS DE BALANCE DE FLUJO (FBA)

En entornos biológicos, las reacciones suelen ser catalizadas por enzimas y las etapas de transporte se llevan a cabo mediante proteínas específicas. Las redes metabólicas involucran la transferencia y conversión de materia de una forma química a otra. Una representación compacta de dicha red es la matriz estequiométrica $\mathbf{S}$ previamente mencionada. Si la red se considera en estado estacionario, se pueden deducir relaciones entre los flujos a partir de ciertos procedimientos matemáticos. El análisis de balance de flujo (Orth *et al.*, 2010) es una aplicación importante para el análisis estructural de redes metabólicas. Se puede derivar una distribución de flujos en toda la red bajo el supuesto de estado estacionario y restricciones derivadas de consideraciones físicas, químicas, así como de una función objetivo tal como la maximización del crecimiento microbiano. Estos supuestos se han aplicado con éxito también a redes de escala genómica que han producido muchos hallazgos científicos tales como el efecto de mutaciones y de knock-outs en la capacidad metabólica, condiciones óptimas de crecimiento, composiciones óptimas de los medios, adaptación evolutiva a los cambios medioambientales y el efecto de fármacos entre otros. Los coeficientes estequiométricos denotan la proporción de moléculas de sustrato y producto involucradas en una reacción. Por ejemplo, para la reacción:



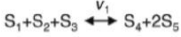
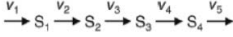
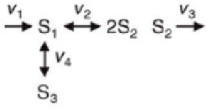
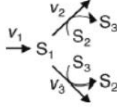
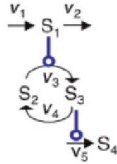
Los coeficientes estequiométricos de S_1 , S_2 y P son -1, -1 y 2, respectivamente. La asignación de los coeficientes estequiométricos no es única. También podríamos argumentar que para la producción de un mol de P , se debe usar la mitad de un mol de cada S_1 y S_2 y, por lo tanto, elegir -1/2, -1/2 y 1. Si cambiamos la dirección de la reacción, entonces podemos elegir 1, 1 y -2. El cambio de concentraciones en el tiempo se puede describir usando ecuaciones diferenciales ordinarias. A partir de la representación de la reacción dada por la ecuación (4.4) y la primera elección de los coeficientes estequiométricos, obtenemos las siguientes ecuaciones:

$$\frac{dS_1}{dt} = -v; \frac{dS_2}{dt} = -v; \frac{dS_3}{dt} = 2v \quad (4.5)$$

Esto significa que la caída de concentración S_1 ocurre a velocidad v y se acompaña con la caída de concentración S_2 con la misma velocidad y con la producción de P al doble de velocidad. Para una red metabólica que consiste en m sustancias y n reacciones, la dinámica del sistema se describe mediante

estas ecuaciones de equilibrio, ya que se considera que existe equilibrio entre la producción de metabolitos y consumo de sustratos.

Tabla 4.1. Ejemplos redes de reacción y su representación mediante matrices estequiométricas y sistemas de ecuaciones diferenciales ordinarias

Red	Matriz estequiométrica	Sistema de ec. diferenciales
S_1 	$\begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \\ 2 \end{pmatrix}$	$\frac{dS_1}{dt} = \frac{dS_2}{dt} = \frac{dS_3}{dt} = -v_1$ $\frac{dS_4}{dt} = v_1$ $\frac{dS_5}{dt} = 2v_1$
S_2 	$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$	$\frac{dS_1}{dt} = v_1 - v_2$ $\frac{dS_2}{dt} = v_2 - v_3$ $\frac{dS_3}{dt} = v_3 - v_4$ $\frac{dS_4}{dt} = v_4 - v_5$
S_3 	$\begin{pmatrix} 1 & -1 & -1 \end{pmatrix}$	$\frac{dS_1}{dt} = v_1 - v_2 - v_3$
S_4 	$\begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	$\frac{dS_1}{dt} = v_1 - v_2 - v_4$ $\frac{dS_2}{dt} = 2v_2 - v_3$ $\frac{dS_3}{dt} = v_4$
S_5 	$\begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$	$\frac{dS_1}{dt} = v_1 - v_2 - v_4$ $\frac{dS_2}{dt} = -v_2 + v_3$ $\frac{dS_3}{dt} = v_2 - v_3$
S_6 	$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\frac{dS_1}{dt} = v_1 - v_2$ $\frac{dS_2}{dt} = v_4 - v_3$ $\frac{dS_3}{dt} = v_3 - v_4$ $\frac{dS_4}{dt} = v_5$

Se asume que las reacciones son la única causa de cambios de concentraciones y que los metabolitos externos no están incluidos en las ecuaciones de balance. Estos metabolitos no se presentan en este modelo porque no se consideran relevantes y se asume que sus concentraciones se mantienen constantes mediante procesos de transferencia.

Las ecuaciones de equilibrio (4.2) también se pueden aplicar si el sistema consta de varios compartimentos. En este caso, cada compuesto en cada compartimento debe considerarse como un compuesto individual y las etapas de transporte se consideran como reacciones de transferencia del compuesto perteneciente a un compartimento al mismo compuesto que pertenece al otro compartimento. Las diferencias de volumen deben tenerse en cuenta, ya que cierto número de moléculas que se mueven de un compartimento a otro cambian las concentraciones en estos compartimentos en el caso de considerar volúmenes diferentes. La Tabla 4.1 proporciona algunos ejemplos de redes de reacción y sus correspondientes matrices estequiométricas. Los metabolitos externos no se incluyen en la red ni en la matriz estequiométrica. Las flechas finas denotan reacciones y las flechas en negrita denotan activación. La matriz estequiometría **S** no contiene información sobre la reversibilidad de las reacciones. Para determinar los signos en **S**, debe asignarse la dirección de las flechas. Por ejemplo, como positiva «de izquierda a derecha» y «de arriba abajo». Si el flujo neto de una reacción se produce en la dirección opuesta a como la flecha lo indica, el valor de la velocidad **v** se considera negativo. La descripción matemática del sistema metabólico involucra a un vector $\mathbf{x} = \{x_1; x_2; \dots; x_n\}^T$ de valores de concentración de metabolitos intracelulares, un vector $\mathbf{v} = \{v_1; v_2; \dots; v_r\}^T$ de velocidades de reacción, y la matriz de coeficientes estequiométricos $\mathbf{S} = \{s_{ij}\}$.

Información codificada en la matriz estequiométrica **S**

La información contenida en **S** permite calcular las combinaciones de flujos individuales que son factibles en el estado estacionario. La condición de estado estacionario viene dada por:

$$\frac{dx}{dt} = \mathbf{S}\mathbf{v} = \mathbf{0} \quad (4.6)$$

Téngase en cuenta que $\mathbf{o}$ es un vector con longitud n , $\mathbf{o} = (0; 0; \dots; 0)^T$. El signo de igualdad en la ecuación (4.6) denota un sistema de ecuaciones lineales para la determinación de las velocidades $\mathbf{v}$. Esta ecuación tiene soluciones no triviales solo para el caso en el que haya menos reacciones linealmente independientes que metabolitos. El rango de una matriz es el número de filas (o columnas) linealmente independientes. Una matriz de kernel $\mathbf{K}$ cumple la siguiente condición:

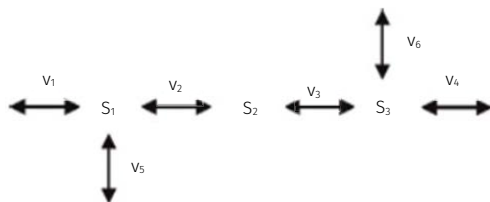
$$\mathbf{SK} = \mathbf{0} \quad (4.7)$$

El núcleo o kernel de una matriz, también llamado espacio nulo, es el núcleo de la combinación lineal definida por la matriz. La ecuación (4.7) expresa las respectivas dependencias lineales entre las columnas de la matriz estequiométrica. $\mathbf{K}$ consiste en r vectores base de rango $\mathbf{S}$ expresados como columnas y puede determinarse utilizando el algoritmo de Gauss. El kernel no está definido de forma única. La multiplicación de $\mathbf{K}$ por una matriz regular $\mathbf{Q}$ de tamaño apropiado ($\mathbf{K}' = \mathbf{KQ}$, equivalente a la combinación lineal de las columnas de $\mathbf{K}$) produce otro núcleo válido $\mathbf{K}'$ de $\mathbf{S}$. Cada conjunto $\mathbf{J}$ de posibles flujos en estado estacionario se puede expresar como una combinación lineal de columnas k_i de $\mathbf{K}$:

$$\mathbf{J} = \sum_{i=1}^{r=\text{Rango}(\mathbf{N})} a_i k_i \quad (4.8)$$

Los coeficientes deben tener unidades correspondientes a las unidades de velocidades de reacción (por ejemplo, mM s⁻¹). Si las entradas en una determinada fila son cero en todos los vectores básicos, entonces se ha encontrado una reacción de equilibrio. En cualquier estado estacionario, la suma de las velocidades de reacción debe ser cero. Por ejemplo, para el sistema de reacción S_4 en la Tabla 4.1, se cumple que $r = 4$ y $\text{Rango}(\mathbf{S}) = 3$. Su kernel consta de una sola columna $\mathbf{K} = (1 \ 1 \ 1 \ 0)^T$. Por lo tanto, $v_4 = \sum_i \alpha_i \mathbf{o} = 0$. En cualquier estado estacionario, las velocidades de producción y de degradación de S_3 deben ser iguales, lo que lleva a un cambio neto igual cero. Si todos los vectores básicos contienen las mismas entradas para un conjunto de filas, esto indica una ruta de reacción no ramificada y en cada estado estacionario, la tasa neta de todas las reacciones respectivas es igual. Hasta ahora, no se ha analizado la reversibilidad de las reacciones en la red. La irreversibilidad de una reacción no afecta a la matriz estequiométrica. Sin embargo, tiene consecuencias para la elección de los vectores de base k_i del kernel $\mathbf{K}$. Se debe elegir un conjunto de vectores de base para satisfacer los signos de flujo cuando se calcula mediante la ecuación (4.8).

Ejemplo. Considérese a la siguiente red de reacciones:



Dicha red contiene a 6 reacciones y su matriz estequiométrica es la siguiente:

$$\begin{pmatrix} 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{pmatrix}$$

Dicha matriz tiene un rango igual a 3 y consecuentemente, su kernel estará representado por tres vectores básicos, por ejemplo $k_1 = [1 \ 1 \ 1 \ 0 \ 0 \ -1]^T$, $k_2 = [1 \ 0 \ 0 \ 0 \ 1 \ 0]^T$ y $k_3 = [-1 \ -1 \ -1 \ -1 \ 0 \ 0]^T$. Las entradas para la segunda y tercera reacción deberán ser siempre iguales y en consecuencia, en el estado estacionario, las reacciones 2 y 3 deberán ser iguales.

El cono de flujo

El análisis estequiométrico de redes bioquímicas se puede realizar considerando solo las reacciones irreversibles expresando a las reacciones reversibles en dos reacciones irreversibles de sentido opuesto. Sobre la base de dicha representación unidireccional, los vectores de la base representada por la ecuación (4.8) forman un cono convexo en el espacio de flujo. Este mapeo relaciona el análisis estequiométrico con los conceptos de geometría convexa de la siguiente manera: la suposición de estado estacionario requiere que un vector de flujo sea un elemento del espacio nulo de la matriz estequiométrica S . Una fila de K puede interpretarse como un hiperplano en el espacio de flujo. La intersección de todos estos hiperplanos forma el llamado «espacio nulo». Siempre que todas las reacciones sean irreversibles, la intersección del espacio semipositivo forma un cono poliédrico llamado «cono de flujo». La intersección de los hiperplanos da como resultado un conjunto de rayos o aristas que comienzan en cero y que describen completamente al cono. Los bordes están representados por vectores y cualquier estado estacionario admisible del sistema es una combinación positiva de estos vectores.

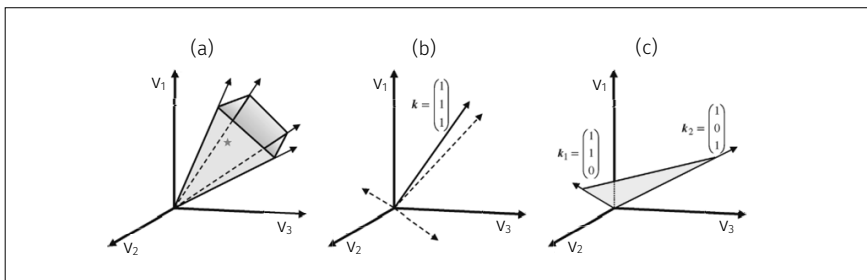


Figura 4.3. El cono de flujo es una representación del subespacio de estados factibles dentro del espacio abarcado por todos los vectores de valores positivos para velocidades de reacciones irreversibles, v_i ($i = 1; \dots; i_r$). Las soluciones factibles son combinaciones lineales de vectores de la base de K . (a) Representación ilustrativa del cono de flujo para un sistema de dimensión superior con rango $r = 4$. Los vectores de la base K son rayos que comienzan en el origen. Los planos que conectan los rayos indican los límites posibles de las distribuciones de flujo real establecidas por las restricciones. El asterisco indica una solución factible en el espacio de flujos. (b) El cono de flujo para una cadena de reacción no ramificada, como la red S_2 de la Tabla 1 es solo un rayo, ya que K está representado por un único vector independiente. (c) El cono de flujo para la red N_3 en la Tabla 4.1 es el plano abarcado por los vectores de base $k_1 = [1 \ 1 \ 0]^T$, $k_2 = [1 \ 0 \ 1]^T$.

Modos de flujo y vías extremas

Una definición rigurosa del término «vía» en una red metabólica no es sencilla. Una «vía metabólica» podría definirse como un conjunto de reacciones que están unidas por metabolitos comunes. Ejemplos típicos incluyen la glucólisis o diferentes vías de síntesis de aminoácidos. Muchísimas vías se conocen desde hace mucho tiempo y es difícil descubrir nuevas vías en mapas metabólicos que se hayan reconstruido a partir de genomas secuenciados de bacterias. El problema de identificación de vías funcionales utiliza el supuesto de estado estacionario y debe considerarse que algunas reacciones son reversibles, mientras que otras son irreversibles. A pesar de estas restricciones, aún se obtienen infinitas soluciones factibles, especialmente en redes de grandes dimensiones. Comenzamos con la definición de un modo de flujo M , que es el conjunto de vectores de flujo que representan rutas directas a través de la red entre los metabolitos externos. En términos matemáticos, un modo de flujo se define como el siguiente conjunto:

$$M = \{v \in R^T : v = \lambda v^*, \lambda = 0\} \quad (4.9)$$

Donde $\mathbf{v}^*$ es un vector r -dimensional que cumple las siguientes condiciones:

1. La condición de estado estacionario $\mathbf{S} \mathbf{v} = 0$.
2. La restricción de signo en $\mathbf{v}^*$, es decir, las direcciones de flujo que cumplen con las condiciones de irreversibilidad que implican que los valores de $\mathbf{v}^*$ contienen flujos no negativos.

Un modo de flujo $\mathbf{M}$ que comprende $\mathbf{v}$ se llama reversible si el conjunto $\mathbf{M}$ que comprende $\mathbf{v}$ es también un modo de flujo. Un modo de flujo es un modo de flujo elemental si utiliza un conjunto mínimo de reacciones y no puede descomponerse más, es decir, el vector $\mathbf{v}$ no puede representarse como una combinación lineal no negativa de dos vectores que cumplen las condiciones (4.8) y (4.9) pero contienen más entradas no nulas que $\mathbf{v}$. Un modo de flujo elemental contiene a un conjunto mínimo de enzimas que podrían funcionar en estado estacionario, con todas las reacciones irreversibles utilizadas en la dirección apropiada. El conjunto de modos elementales de flujo se define de forma única. Un modo de flujo es una vía extrema si:

1. Todas las reacciones son no negativas.
2. El modo de flujo pertenece a la aristas del cono de flujo, lo que también significa que representa un vector base de $\mathbf{K}$.

ANÁLISIS DE BALANCE DE FLUJO (FBA) FORMULADO COMO UN PROBLEMA DE PROGRAMACIÓN LINEAL

FBA (Orth *et al.*, 2010) es el método computacional más utilizado para el análisis de modelos metabólicos estequiométricos a escala genómica. Las células convierten a los sustratos que contienen carbono en metabolitos portadores de energía, como por ejemplo, ATP, precursores de la biomasa y subproductos que se transportan al espacio extracelular. La aplicación de los principios de conservación requiere la evaluación de todas las entradas y salidas, así como las fuentes y los sumideros del sistema. Los metabolitos se pueden «exportar» fuera de la célula o se pueden «importar» desde el medioambiente extracelular. Algunas de estas moléculas de entrada están disponibles en exceso (por ejemplo, el agua), mientras que otras están limitadas por su concentración extracelular (por ejemplo, la fuente de carbono) y/o por su velocidad de difusión (oxígeno). Los productos del metabolismo se exportan fuera de la célula (ácido acético, CO_2). Todas estas entradas y salidas se modelan utilizando reacciones de intercambio, que son reacciones artificiales que transportan metabolitos a través del límite

del sistema (ver figura 4.4). Todas las reacciones de intercambio se consideran como reversibles con la dirección hacia adelante que representa la exportación y la dirección hacia atrás que denota la importación. Por ejemplo, la reacción de intercambio para el metabolito A se define de la siguiente manera:



A diferencia de las reacciones de intercambio que son construcciones artificiales que modelan la transferencia de un metabolito entre el entorno extracelular [e] y el límite del sistema, las reacciones de transporte modelan el transporte real de un metabolito entre el citosol [c] y el entorno extracelular [e] o entre el citosol y otros compartimentos de las células como las mitocondrias en modelos metabólicos de células eucariotas. Una reacción de transporte puede ser pasiva basándose únicamente en la difusión o activa que involucra el consumo de ATP o la translocación simultánea de protones o iones. Por ejemplo, el transporte pasivo del metabolito A que usa la reacción de transporte At se modela de la siguiente manera:



Los metabolitos también pueden ser capturados por los sumideros de un modelo metabólico. Los dos sumideros más utilizados son la reacción de generación de biomasa y de consumo de ATP destinado al mantenimiento celular no asociado al crecimiento. La reacción de producción de biomasa actúa como un sumidero que canaliza los constituyentes necesarios para la biomasa en las proporciones fisiológicamente relevantes. Además del crecimiento, los recursos metabólicos también se usan para fines de mantenimiento celular y esa demanda de energía se considera como un sumidero de ATP equivalente denominado ATP de mantenimiento sin crecimiento (NGAM). La cantidad requerida de NGAM varía entre los organismos. Por ejemplo, el ATP de NGAM para *E. coli* es de 8,39 mmol gDW/h, mientras que para *S. cerevisiae* es de 1,00 mmol gDW/h.

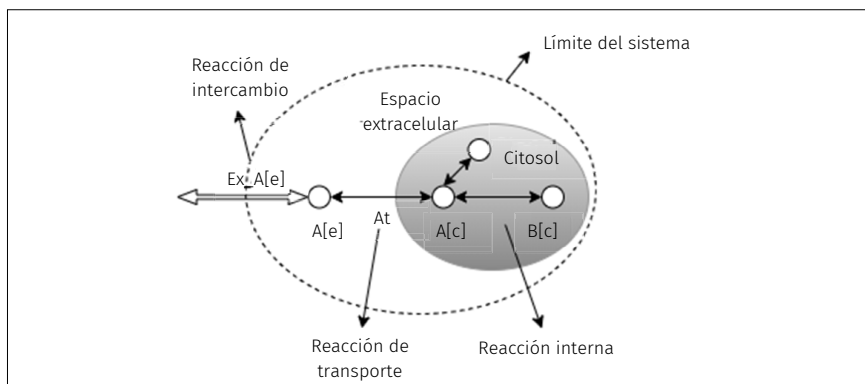


Figura 4.4. Reacciones de intercambio, transporte y reacciones internas.

Balances de componentes

La conservación de un metabolito i que participa en múltiples reacciones j de una red metabólica se puede expresar por medio de la ecuación (4.6), en la que $I = \{1, \dots, M\}$ y $J = \{1, \dots, N\}$ son los conjuntos de metabolitos y reacciones, respectivamente, x_i es la concentración de un metabolito i , v_j es el flujo de la reacción j y S_{ij} es el coeficiente estequiométrico del metabolito i en reacción j . El valor de S_{ij} es negativo, positivo o cero si el metabolito i sirve como reactivo, producto o no participa en la reacción j , respectivamente. Este sistema de ecuaciones diferenciales ordinarias describe tanto el estado estacionario como el comportamiento dinámico de la red metabólica. En los modelos cinéticos, los flujos v_j se reemplazan por expresiones cinéticas que dan como resultado un sistema de ecuaciones diferenciales ordinarias no lineales con concentraciones de metabolitos x_i como variables de estado. En contraste, en los modelos basados en estequiometría, los flujos de reacción no se sustituyen por expresiones cinéticas, sino que se tratan como variables libres. Además, se asume la condición de estado pseudoestacionario para todos los metabolitos intracelulares, dado que las constantes de tiempo para las reacciones metabólicas (milisegundos a decenas de segundos) suelen ser mucho más pequeñas que los de otros procesos celulares como son la regulación transcripcional (minutos) o el crecimiento celular (de decenas de minutos a horas). La ecuación (4.2) se reduce entonces a una ecuación algebraica. Al reformular esta ecuación en forma de matriz, obtenemos:

$$\mathbf{S} \mathbf{v} = \mathbf{0} \quad (4.12.1)$$

$$\begin{bmatrix} S_{11} & \cdots & S_{1N} \\ \vdots & \ddots & \vdots \\ S_{M1} & \cdots & S_{MN} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4.12.2)$$

donde $\mathbf{S}$ y $\mathbf{v}$ se conocen como la matriz estequiométrica y el vector de flujo, respectivamente. La matriz $\mathbf{S}$ captura todas las propiedades estructurales de la red metabólica. Dado que la mayoría de los metabolitos participan en muchas reacciones metabólicas, el número de metabolitos es usualmente menor que el número de reacciones en una red metabólica (es decir, $M < N$). Esto hace que el sistema de ecuaciones algebraicas sea indeterminado, lo que implica que hay un número infinito de vectores de flujo que pueden satisfacer la ecuación de balance materia en estado estacionario. Cualquier vector de flujo $\mathbf{v}$ que satisfaga la ecuación (4.12) está en el espacio nulo de la matriz $\mathbf{S}$. Por lo tanto, la ecuación (4.12) captura el rango de todos los fenotipos metabólicos permisibles del sistema. Sin embargo, no todos estos estados metabólicos son fisiológicamente factibles. Los métodos de análisis y reconstrucción basados en restricciones pretenden acotar el espacio de flujos definidos por la ecuación (4.12) en la medida de lo posible para eliminar fenotipos metabólicos fisiológicamente irrelevantes. Dado que es usual que haya infinitas soluciones, con el fin de obtener al conjunto más pequeño posible de soluciones se puede utilizar un criterio de optimización. El más utilizado es la maximización de la velocidad de crecimiento de biomasa:

$$\begin{aligned} \text{Maximizar } z &= v_{\text{biomasa}} \\ \text{sujeto a} \end{aligned} \quad (4.13.1)$$

$$\mathbf{S} \mathbf{v} = \mathbf{0} \quad (4.13.2)$$

$$v_i \geq 0 \quad \forall v_i \in I_{\text{irreversible}} \quad (4.13.3)$$

Restricciones de capacidad y termodinámicas

Para reducir aún más el espacio de soluciones y comprender mejor el metabolismo en condiciones ambientales específicas, se pueden agregar restricciones tales como velocidades de flujo máximas, flujos medidos experimentalmente o restricciones termodinámicas. Las restricciones de capacidad de flujo y termodinámicas se manifiestan como límites inferiores (LB) y límites superiores (UB) en los flujos de reacción:

$$LB_j \leq v_j \leq UB_j \quad (4.14)$$

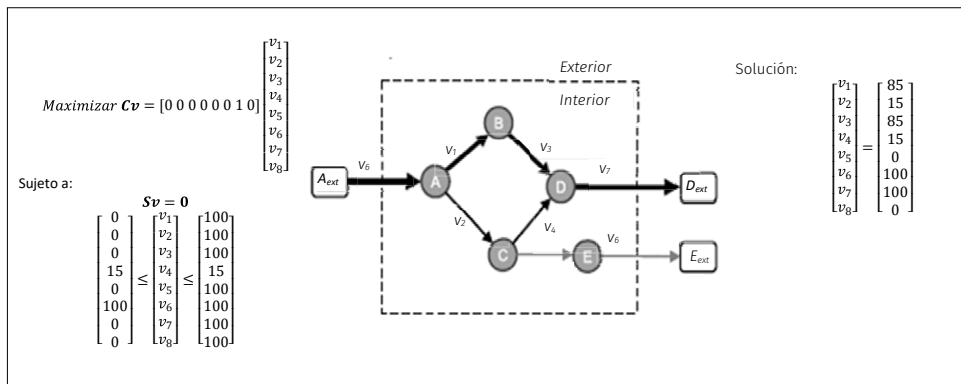


Figura 4.5. Análisis de balance de flujo para una pequeña red usada como ejemplo.

La reversibilidad de una reacción catalizada enzimáticamente se puede deducir en función del cambio en la energía libre de Gibbs (G) de la reacción. El valor exacto de la misma dentro de un entorno celular es muy difícil de evaluar debido a las incertidumbres en las concentraciones de metabolitos. Una reacción generalmente se trata como irreversible si G es negativa. En contraste, una reacción se trata como reversible si el valor de G puede cambiar de signo para diferentes rangos de concentración fisiológicamente relevantes. Esto da lugar a tres casos según los límites de las velocidades de reacción:

- La reacción ocurre solo en la dirección de avance ($LB = 0$, $UB = M$).
- La reacción se produce solo en la dirección inversa ($LB = -M$; $UB = 0$).
- La reacción es reversible ($LB = -M$, $UB = M$).

Donde M es un número positivo grande (por ejemplo, 1000) en comparación con el flujo limitante de transporte de recursos (por ejemplo, 10 mmol gDW/h de glucosa) utilizado para escalar los flujos en la red. Además de consideraciones termodinámicas, una serie de otros factores pueden limitar los flujos de reacción. Por ejemplo, el límite inferior de todas las reacciones de intercambio se establece en cero excepto los correspondientes a los metabolitos que están disponibles en exceso en el medio de crecimiento. El límite inferior se establece en $-M$ para los metabolitos que están disponibles en exceso (agua), mientras que se establece en un valor negativo finito para recursos limitantes como la fuente de carbono. Una práctica común es establecer el recurso limitante que generalmente el sustrato que actúa como fuente de carbono en 10 mmol gDW/h , proporcionando así un escalado

conveniente para todos los flujos en la red. El límite superior para todos los flujos de intercambio se establece en M lo que implica que todos los metabolitos con una reacción de intercambio en el modelo se pueden exportar. Los límites de flujo también se pueden derivar de las mediciones de flujos externos o del análisis de los flujos metabólicos internos. Las limitaciones de capacidad también podrían derivarse en función de la actividad enzimática. Se debe tener cuidado al establecer límites apropiados para todas las reacciones. Los errores en los límites de reacción de intercambio podrían llevar a un modelo metabólico sin distribuciones de flujo factibles o modelos que pueden generar una cantidad ilimitada de ATP a través de un flujo no restringido. Además, límites demasiado permisivos para las reacciones internas podrían llevar a la predicción de ciclos termodinámicamente infactibles. En la Tabla 4.2 se proporciona un resumen de cómo se establecen los límites inferiores y superiores.

Tabla 4.2. Límites típicos inferior (LB_j) y superior (UB_j) para diferentes tipos de reacciones (Maranas y Somorrodí, 2016).

Tipo de reacción	(LB _j , UB _j)
Irreversible	(0, M)
Reversible	($-M$, M)
De intercambio para metabolitos que no están en el medio de cultivo	(0, M)
De intercambio para metabolitos en exceso en el medio de cultivo	($-M$, M)
De intercambio para metabolitos limitantes en el medio de cultivo	($-c$, M)
Consumo de ATP para mantenimiento celular	(c , c)
Reacciones con flujos medidos experimentalmente	(v_j^{min} , v_j^{max})

$c \ll M$ (número positivo grande; por ejemplo, 1000).

El problema de optimización se convierte entonces en:

$$\text{Maximizar } z = v_{biomasa} \quad (4.15.1)$$

$$\text{sujeto a} \quad (4.15.2)$$

$$Sv = 0 \quad (4.15.2)$$

$$LB_j \leq v_j \leq UB_j \quad (4.15.3)$$

$$v_s = m_s \quad (4.15.4)$$

Se pueden encontrar en la literatura varios criterios de optimización. Existen funciones objetivo no lineales como la maximización del rendimiento de biomasa ($Max(v_{biomasa}/v_{glucosa})$) o funciones objetivos múltiples como minimizar los flujos totales y maximizar el rendimiento de biomasa. La

elección de la función objetivo depende de los datos disponibles, las restricciones impuestas y el objetivo del modelo. Por ejemplo, si se mide el crecimiento de la biomasa, el flujo de biomasa $v_{biomasa}$ puede limitarse a la velocidad de crecimiento medida. Por tanto, la función objetivo no será maximizar la velocidad de crecimiento de la biomasa sino, por ejemplo, minimizar la velocidad de consumo de un sustrato limitante. Si también se mide el consumo de sustrato, $v_{substrato}$ puede limitarse al valor medido y la función objetivo puede convertirse en la minimización de la velocidad de utilización de ATP o la minimización de la suma de los valores de flujo. Pero si el modelo tiene la intención de predecir la velocidad de crecimiento de biomasa, la función objetivo será la maximización del crecimiento de la biomasa y la predicción del modelo de $v_{biomasa}$ se comparará con la tasa de crecimiento medida para validar el modelo. Cabe señalar que la solución del problema FBA proporciona flujos elementales con mejor rendimiento en el caso de una función objetivo lineal correspondiente a la maximización de un producto o biomasa. El análisis de equilibrio de flujo se puede utilizar para predecir distribuciones de flujos de reacciones intracelulares en condiciones ambientales dadas. Por extensión, FBA también se puede utilizar para comparar flujos entre diferentes condiciones ambientales y entre diferentes especies. También se puede estudiar el impacto de una reacción determinada en la función objetivo, ya sea que su flujo esté disminuido (inhibición), aumentado (catalizado) o eliminado (supresión de genes) para ayudar a diseñar nuevas cepas. Existen varios programas para calcular FBA, de los cuales el más conocido es la caja de herramientas COBRA (Heirendt *et al.*, 2019) disponible para MatLab.

Análisis de acoplamiento de flujo (FCA)

El análisis de acoplamiento de flujo (Burgard *et al.*, 2004) permite detectar dependencias entre diferentes flujos de reacción de redes metabólicas en estado quasi-estacionario. Las restricciones estequiométricas y de otro tipo no solo determinan todas las posibles distribuciones de flujo de estado estacionario en una red, sino que también permiten deducir relaciones de acoplamiento entre las reacciones. Por ejemplo, si un flujo a través de una reacción en estado estacionario implica un flujo a través de otra reacción, entonces se dice que ambas reacciones están acopladas. Existen varios tipos de acoplamiento de flujo:

Definición 1. Reacción bloqueada: sea $i \in \{1, \dots, n_r\}$ una reacción. Si $v_{sol} = 0$, para i en todo $vsol \in Ssol$, la reacción i se dice bloqueada. De lo contrario, i está desbloqueada.

Definición 2. Relaciones acopladas: sean $i, j \in \{1, \dots, nr\}$ dos reacciones desbloqueadas. Las relaciones de (des)acoplamiento $i \rightarrow j$, $i \leftrightarrow j$, $i \Leftrightarrow j$ se definen de la siguiente manera:

- La reacción i está acoplada direccionalmente con la reacción j , expresada $i \rightarrow j$, si para todo $v_{sol} \in S_{sol}$, $vi = 0 \Rightarrow vj = 0$.
- La reacción i está parcialmente acoplada con la reacción j , expresada $i \leftrightarrow j$, si $i \rightarrow j$ y $j \rightarrow i$.
- La reacción i está completamente acoplada la reacción con j , expresada $i \Leftrightarrow j$, si $i \leftrightarrow j$ y existe una constante λ tal que $\forall v_{sol} \in S_{sol}$, $vj = \lambda vi$.
- La reacción i está desacoplada de la reacción j , si ni $i \rightarrow j$ ni $j \rightarrow i$ ocurren.

FCA se ha utilizado para explorar diversas cuestiones, como la evolución de la red, la esencialidad de los genes, la regulación de genes o el análisis de flujos medidos experimentalmente. Por un lado, FCA puede ayudar en la reconstrucción de redes metabólicas al verificar si el acoplamiento entre reacciones está de acuerdo con los hallazgos experimentales. Por otro lado, FCA puede ayudar a definir estrategias de intervención para eliminar algunas reacciones objetivo.

ANÁLISIS DE VARIABILIDAD DE FLUJO (FVA)

En el FBA, aunque se aplique un criterio de optimización para obtener una solución del sistema, la solución usualmente no es única y en consecuencia se pueden encontrar varias distribuciones de flujo óptimas. Para explorar todas las posibles distribuciones de flujo óptimas, se puede utilizar al denominado «análisis de variación de flujo» (Gudmundsson *et al.*, 2010). Para ello habiendo encontrado el valor óptimo de la función objetivo f_{opt} , cada flujo se maximiza y luego se minimiza para obtener el rango de valores que cada flujo puede tomar y que sea consistente con FBA:

$$\begin{aligned} & \text{Maximizar \& Minimizar } v_k \\ & \text{sujeto a} \end{aligned} \tag{4.16.1}$$

$$Kv = 0 \tag{4.16.2}$$

$$LB_j \leq v_j \leq UB_j \tag{4.16.3}$$

$$v_s = m_s \tag{4.16.4}$$

$$v_{biomasa} = v_{biomasa}^{optimo} \tag{4.16.5}$$

FVA permite deducir rutas óptimas paralelas en una red metabólica y comprender mejor la robustez de la misma. De hecho, los sistemas biológicos a menudo tienen rutas paralelas redundantes que tienen un mismo propósito funcional. Esta redundancia contribuye a la robustez de la red en el contexto de un crecimiento óptimo y conducen a distribuciones de flujo alternativas. Algunas reacciones pueden considerarse opcionales si tienen flujo nulo. Otras pueden resultar esenciales porque deben tener siempre un flujo no nulo. Sin embargo, debemos tener en cuenta el hecho de que estos resultados dependen en gran medida de las restricciones utilizadas y de la elección de la función objetivo del problema FBA.

ANÁLISIS DINÁMICO DE BALANCE DE FLUJO (dFBA)

Aunque las actividades enzimáticas en una célula cambian constantemente como resultado de la regulación metabólica, la validez del supuesto de estado estacionario se justifica en muchos casos como en cultivos celulares continuos y en la fase de crecimiento exponencial de una fermentación de tipo batch. Sin embargo, existen muchas situaciones prácticas, como en biorreactores de tipo *fed batch*, donde las células pueden mostrar un comportamiento dinámico oscilatorio. Recientemente se han propuesto desarrollos de enfoques basados en restricciones para el análisis de balance de flujo dinámico (Mahadevan *et al.*, 2002). El estudio de un biosistema en condiciones transientes mediante métodos basados en restricciones utiliza a modelos de estado estacionario secuencialmente, evaluando el estado del biosistema en cada intervalo de tiempo de muestreo. De esta manera, dFBA se puede usar para estimar los cambios dinámicos de los flujos intracelulares y para estimar el flujo en las principales vías metabólicas activadas en la red de reacción sin la necesidad de un gran número de parámetros cinéticos predefinidos. La extensión dinámica del análisis de balance de flujo se puede formular de las siguientes dos formas. El *enfoque de optimización dinámica* implica la optimización durante todo el horizonte temporal de integración para obtener los perfiles temporales de flujos y de concentración de metabolitos. El problema de optimización dinámica se puede transformar en un problema de programación no lineal (NLP) utilizando algún método de discretización o de colocación ortogonal para facilitar su resolución. Los detalles de la función objetivo y las restricciones en la formulación se presentan en la ecuación (4.18). El *enfoque de optimización estática* implica dividir el horizonte de simulación en varios intervalos de tiempo finitos y resolver el problema FBA instantáneo al comienzo de cada intervalo de tiempo para

luego proceder a la integración numérica en el intervalo con el fin de estimar el consumo de sustratos y producción de metabolitos en dicho intervalo. El problema de optimización se resuelve repetidamente durante cada intervalo para obtener la distribución de flujo en el mismo.

Enfoque dFBA basado en optimización dinámica

Considere una red metabólica con m metabolitos y n flujos. El conjunto de ecuaciones de conservación de materia, para cada metabolito, da como resultado al siguiente conjunto de ecuaciones diferenciales ordinarias:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{v}\mathbf{X} \quad (4.17.1)$$

$$\frac{dX}{dt} = \mu X \quad (4.17.2)$$

$$\mu = \sum_j w_j v_j \quad (4.17.3)$$

donde $\mathbf{x}$ es el vector de concentraciones de metabolitos, X es la concentración de biomasa y $\mathbf{v} = [v_1, \dots, v_j]$ es el vector de flujos metabólicos por gramo de biomasa, $\mathbf{A}$ es la matriz estequiométrica de la red metabólica, μ es la velocidad de crecimiento obtenida como suma de las reacciones que sintetizan los precursores del crecimiento y w_j son las cantidades de los precursores del crecimiento requeridas por gramo de biomasa. Junto con el sistema de ecuaciones dinámicas se deben imponer restricciones adicionales para una predicción realista de las concentraciones de metabolitos y los flujos metabólicos. Estos incluyen niveles de flujo y metabolitos no negativos, límites en la velocidad de cambio de flujos y restricciones adicionales en los flujos de transporte. El problema general de optimización dinámica se formula de la siguiente manera:

$$\text{Max } w_{end} \phi(\mathbf{x}, \mathbf{v}, X) \Big|_0^{t_f} + w_{path} \sum_{j=1}^M \int_{t_0}^{t_f} L(\mathbf{x}, \mathbf{v}, X(t)) \delta(t - t_j) dt \quad (4.18.1)$$

$$\text{sujeto a} \quad (4.18.2)$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{v}\mathbf{X} \quad (4.18.3)$$

$$\frac{dX}{dt} = \mu X \quad (4.18.4)$$

$$\mu = \sum_{j \in J} w_j v_j \quad (4.18.5)$$

$$t = t_0 + j \frac{t - t_0}{M} \quad j = 0, \dots, M \quad (4.18.6)$$

$$c(\mathbf{v}, \mathbf{x}) \leq 0 \quad |\dot{v}| \leq v_{max} \quad t \in [t_0, t_f] \quad (4.18.7)$$

$$\mathbf{x} \geq 0, X \geq 0 \quad t \in [t_0, t_f] \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad X(t_0) = X_0 \quad (4.18.8)$$

donde x_0 y X_0 son las condiciones iniciales para la concentración de metabolitos y la concentración de biomasa, respectivamente, $c(\mathbf{v}, \mathbf{x})$ es una función vectorial que representa restricciones no lineales que podrían surgir debido a la consideración de expresiones cinéticas para flujos, t_0 y t_f son las condiciones iniciales y los tiempos finales, Φ es la parte de la función objetivo que depende del estado final, L es la función objetivo instantánea, δ es la función delta de Dirac, t_j es el instante de tiempo en el que se considera L , w y w_i son los pesos asociados a la función objetivo terminal e instantánea respectivamente, y $\mathbf{v}(t)$ es el perfil temporal de los flujos metabólicos. Si la restricción de no linealidad está ausente, el problema se reduce al de una optimización de un sistema bilineal. El problema de optimización dinámica se resuelve, por ejemplo, parametrizando las ecuaciones dinámicas mediante el método de colocación ortogonal en elementos finitos. El período de tiempo $[t_0, t_f]$ se divide en un número finito de intervalos y los flujos, los niveles de metabolitos y la concentración de biomasa se parametrizaron mediante las raíces de un polinomio ortogonal dentro de cada elemento finito.

Análisis dinámico de balance de flujo basado en ecuaciones en diferencias finitas

En este caso, el período de tiempo se divide en N intervalos. En ausencia de las restricciones no lineales que involucran los flujos, el problema de optimización se reduce a un problema lineal que resuelve el problema al inicio de cada intervalo para obtener los flujos en ese intervalo de tiempo:

$$\text{Maximizar } v(t) \sum_{j=0}^M w_j v_j(t) \quad (4.19.1)$$

$$\text{sujeto a} \quad (4.19.2)$$

$$z(t + \Delta t) \geq 0$$

$$v(t) \geq 0 \quad (4.19.3)$$

$$c(\mathbf{x}(t), \mathbf{v}(t)) \leq 0 \quad t \in [t_0, t_f] \quad (4.19.4)$$

$$|v(t) - v(t - \Delta t)| \leq \dot{v} \Delta T \quad t \in [t_0, t_f] \quad (4.19.5)$$

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + A v \Delta t \quad (4.19.6)$$

$$X(t + \Delta t) = X(t) + \mu X(t) \Delta t \quad (4.19.7)$$

donde T es la duración del intervalo de tiempo. El problema de optimización lineal se resuelve y las ecuaciones dinámicas se integraron asumiendo que los flujos son constantes en el intervalo de tiempo. Luego se formula el problema de optimización en el próximo intervalo de tiempo y se resuelve nuevamente el problema de optimización lineal. Este procedimiento se repitió desde t_0 a t_f .

ANÁLISIS TERMODINÁMICO EN REDES METABÓLICAS

Restricciones termodinámicas en redes metabólicas

El FBA requiere que la distribución de flujos esté en estado estacionario, pero FBA no puede verificar si dichos flujos son termodinámicamente factibles (Henry *et al.*, 2005). De acuerdo con la segunda ley de la termodinámica, las reacciones químicas, a presión y temperatura constantes, deben ser impulsadas por un consumo neto de energía libre de Gibbs $G(p, T)$. La energía libre de Gibbs de un sistema bioquímico está asociada con la cantidad y los tipos de moléculas:

$$G(p, T) = \sum_{i \in \mathcal{I}} m_i \mu_i \quad (4.20)$$

donde μ_i y m_i denotan el potencial químico y la cantidad (en moles) de la sustancia i , respectivamente. Una reacción química cambiará las cantidades de componentes de acuerdo con los coeficientes estequiométricos ij resultando en un cambio de energía libre de Gibbs (en kJ) que se puede expresar por la diferencia de potencial químico:

$$\Delta_r G_j(p, T) = \sum_{i \in \mathcal{I}} m_i \mu_{ij} \quad (4.21)$$

El símbolo Δ_r denota los cambios de energía libre asociados con las reacciones químicas, y el valor negativo $-\Delta_r G_j(p, T)$ también se llama denominada energía de reacción. El vector de diferencias de potencial químico satisface la condición de Wegscheider:

$$(\Delta_r G)^T \mathbf{K} = 0 \quad (4.22)$$

donde $\mathbf{S}$ es el kernel de la matriz estequiométrica $\mathbf{S}$ que satisface la condición $\mathbf{S} \mathbf{K} = \mathbf{0}$. Esto indica que la caída o aumento de la energía libre entre dos compuestos en una red es independiente de la ruta de reacción que se toma para llegar desde un compuesto a otro. Según la segunda ley de la termodinámica, la energía libre de Gibbs debe disminuir en cualquier reacción que ocurra. Para una reacción, la diferencia de potenciales químicos debe ser negativa y la afinidad de reacción debe ser positiva. En general, para una reacción j , se obtiene la siguiente condición:

$$\sum_{i \in \mathcal{I}} \mu_i s_{ij} v_j \leq 0 \quad (4.23)$$

Por lo tanto, un patrón de flujo $\mathbf{v} = (v_1; \dots; v_r)^T$ solo es factible si la condición (4.23) puede ser satisfecha por algún vector $(\mu_1; \dots; \mu_r)^T$ de potenciales químicos. Esta condición no puede ser comprobada usando una matriz estequiométrica. El FBA no requiere que se cumpla la condición anterior y, por lo tanto FBA puede dar lugar a flujos con sentidos termodinámicamente infactibles. Este problema puede evitarse predefiniendo algunas direcciones de flujo que restringirán el espacio de la solución de antemano. El análisis termodinámico de las vías metabólicas puede determinar la direccionalidad y los límites de actividad de las reacciones metabólicas. La termodinámica permite evaluar sistemáticamente la dirección de las reacciones permitidas y permite eliminar de ciclos termodinámicamente inviables (TIC) en las reconstrucciones de redes metabólicas utilizando métodos basados en la optimización. La dirección de una reacción metabólica independiente está determinada por el signo de su cambio de energía libre de Gibbs. Un valor negativo implica que la dirección «hacia adelante» procederá con un flujo más alto que la dirección «hacia atrás», lo que dará como resultado un flujo neto hacia adelante. Las variaciones en la concentración de reactivos y productos pueden conducir a un cambio en el signo del cambio de energía libre de Gibbs y, por lo tanto, a una inversión de su dirección neta. La siguiente expresión define el cambio de energía libre de Gibbs de una reacción j :

$$\Delta_r G_j = \Delta_r G_j^0 + RT \left(\prod_{i \in I: S_{ij} \neq 0} x_i^{S_{ij}} \right) \quad (4.24)$$

O lo que es lo mismo:

$$\Delta_r G_j = \Delta_r G_j^0 + RT \sum_{i \in I: S_{ij} \neq 0} S_{ij} \ln(x_i) \quad (4.25)$$

donde I es el conjunto de metabolitos, S_{ij} es el coeficiente estequiométrico del metabolito i en la reacción j , $\Delta_r G_j$ es el cambio de energía libre de Gibbs de la reacción j , $\Delta_r G_j^0$ es el cambio estándar de energía libre de Gibbs de la reacción j (es decir, a $T = 25^\circ \text{C}$, 1 M concentración de todos los reactivos y productos acuosos y 1 atm es la presión parcial de cualquier gas involucrado en la reacción), R es la constante universal de gases, T es la temperatura y x_i es la actividad del metabolito i . Las estimaciones de mínimos y máximos valores para $\Delta_r G_j$ pueden calcularse para incorporar incertidumbre y desviaciones en las ecuaciones anteriores, pero los límites más amplios posibles se utilizan para incluir todas las distribuciones de flujos termodinámicamente

factibles y para asignar sistemáticamente la reversibilidad y la dirección de las reacciones en una red metabólica de la siguiente manera:

1. Si $\Delta_r G_j^{max} < 0$, entonces la reacción es irreversible en dirección hacia adelante ($v_j > 0$).
2. Si $\Delta_r G_j^{min} > 0$, entonces la reacción es irreversible en dirección hacia atrás ($v_j < 0$).
3. Si $\Delta_r G_j^{min} < 0$ y $\Delta_r G_j^{max} > 0$, entonces la reacción es reversible ($v_j \in R$).

Es importante tener en cuenta datos experimentales pueden usarse para restringir aún más las direcciones identificadas anteriormente.

Ciclos en el metabolismo celular

Los ciclos de conversión están presentes en las redes metabólicas (por ejemplo, el ciclo del ácido tricarboxílico, el ciclo de Calvin o el ciclo de la urea). Los ciclos permiten dividir una conversión química general en una serie de reacciones que generalmente implican barreras de activación más bajas a través de una secuencia de etapas de activación/desactivación que una reacción directa. La figura 4.6 muestra el ciclo TCA que utiliza reacciones separadas para convertir un grupo acetilo en dos moléculas de dióxido de carbono mientras genera tres moléculas de NADH y una molécula de FADH₂ (Maranas y Somorrodí, 2016). A diferencia de los ciclos de conversión, los ciclos fútiles no llevan a cabo una conversión general, sino que acoplan dos vías metabólicas que operan en direcciones opuestas sin ningún impacto general que no sea la disipación de energía en forma de calor. Un ejemplo bien conocido es el ciclo fútil formado al combinar la glucólisis y la gluconeogénesis que consume un ATP por ciclo que conduce a la disipación de energía química en calor. La fotorrespiración también es un ciclo inútil que disipa el exceso de energía de la luz en forma de calor. Los ciclos fútiles son indeseables en la bioproducción ya que reducen la eficiencia energética y, por lo tanto, el rendimiento en el producto. Sin embargo, en ocasiones, cumplen funciones fisiológicas importantes en la regulación del metabolismo o en la preparación de una vía para responder con rapidez a una perturbación y proporcionan un mecanismo para que el metabolismo asuma la postura de suministrar flujo velozmente en cualquier dirección bloqueando el flujo en la dirección opuesta en lugar de tener que ensamblar recursos catalíticos a pedido. Es importante distinguir ciclos fútiles, que son ciclos termodinámicamente factibles de los ciclos termodinámicamente inviables que se encuentran con frecuencia mediante análisis de balance de flujo.

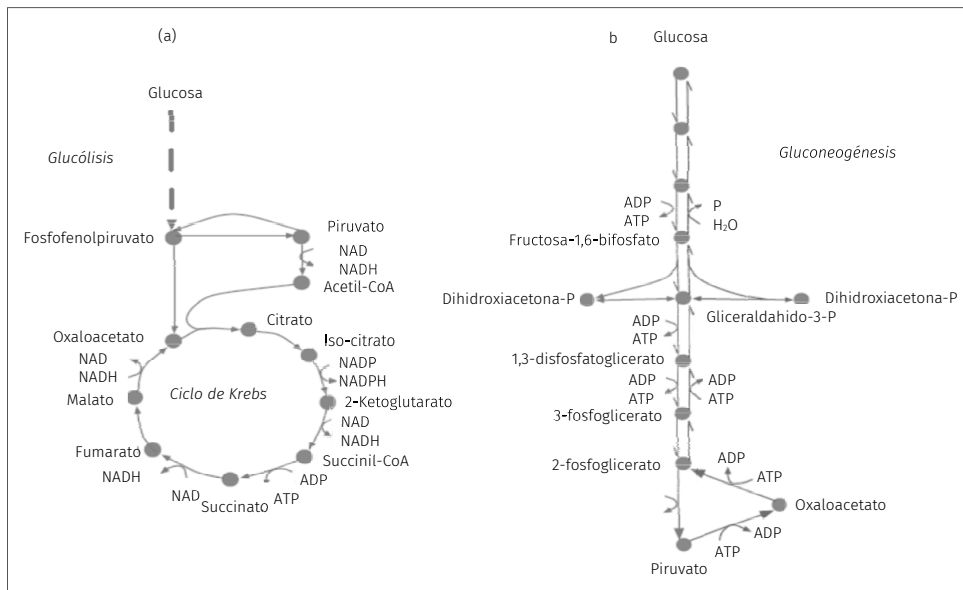


Figura 4.6. Dos ciclos importantes en el metabolismo celular: (a) el ciclo TCA y (b) el ciclo inútil entre la glucólisis y la gluconeogénesis.

Ciclos termodinámicamente inviables

En contraste con los ciclos naturales, los ciclos termodinámicamente inviables (TIC) son causados por la inclusión permisiva de reacciones o direccionalidades de reacción en un modelo metabólico. Estos ciclos consisten en un conjunto de reacciones que forman bucles como $A \rightarrow B \rightarrow C \rightarrow A$ sin que entren ni salgan otros metabolitos. Según la segunda ley de la termodinámica, la fuerza motriz termodinámica general es cero, lo que implica que no puede haber flujo neto alrededor en este ciclo. Los cálculos de análisis de balance de flujo están limitados solo por la estequiometría, por lo que permiten una cantidad ilimitada de flujo a través de las TIC sin entradas externas de energía formando una máquina perpetua del segundo tipo lo cual es termodinámicamente imposible. La identificación y eliminación de TIC es necesaria para evitar flujos metabólicos ilimitados y distribuciones de flujo termodinámicamente inviables. Identificar los TIC en una red es una tarea difícil, ya que podría haber un conjunto de TICs superpuestos que forman un bucle. En cambio, es factible identificar las reacciones que participan en uno o más TIC utilizando el análisis de variabilidad de flujo (FVA).

Un método sistemático denominado análisis de flujo metabólico basado en termodinámica (TMFA), utiliza límites en las energías libres para imponer consistencia termodinámica en los cálculos de FBA. Esto se logra mediante la adición de restricciones termodinámicas al FBA que restringen las direccionalidades de reacción permitidas, eliminando así soluciones termodinámicamente inviables, incluidos los ciclos, de las distribuciones de flujo identificadas. TMFA primero descompone las reacciones reversibles en una red con reacciones con dirección hacia adelante y otra con reacciones hacia atrás para rastrear si una reacción, en la red descompuesta, puede transportar un flujo distinto de cero. Además, las reacciones para las cuales no se conocen los datos de cambio de energía libre de Gibbs se agrupan en un nuevo conjunto de reacciones para las cuales se puede estimar el cambio de energía libre. Es importante tener en cuenta que las reacciones con un cambio de energía libre de Gibbs que es desconocido se retienen en el modelo y el conjunto de reacciones agrupadas se introducen solo para imponer las restricciones de viabilidad termodinámica en tales reacciones. Se define un parámetro α_{kj} para rastrear el conjunto de reacciones que comprende una reacción agrupada k :

$$\alpha_{kj} = \begin{cases} 1 & \text{si la reacción } j \text{ pertenece a la reacción agrupada } k \\ 0 & \text{si no se cumple lo anterior} \end{cases} \quad \forall k \in J^{\text{agregada}}, j \in J^{\text{modelo}}$$

donde J^{modelo} es el conjunto de reacciones en el modelo metabólico con reacciones reversibles descompuestas que contienen ambas reacciones con energías libres estándar de Gibbs conocidas o desconocidas. Se introducen dos conjuntos de variables binarias para capturar si las reacciones en el modelo con $\Delta_r G_j^0$ conocido, así como las reacciones agrupadas, son termodinámicamente factibles:

$$\begin{aligned} z_j &= \begin{cases} 1 & \text{si la reacción } j \text{ es termodinámicamente factible} \\ 0 & \text{si no se cumple lo anterior} \end{cases} & \forall j \in J^{\text{modelo}} \\ y_k &= \begin{cases} 1 & \text{si la reacción agrupada } k \text{ es termodinámicamente factible} \\ 0 & \text{si no se cumple lo anterior} \end{cases} & \forall k \in J^{\text{agregada}} \end{aligned}$$

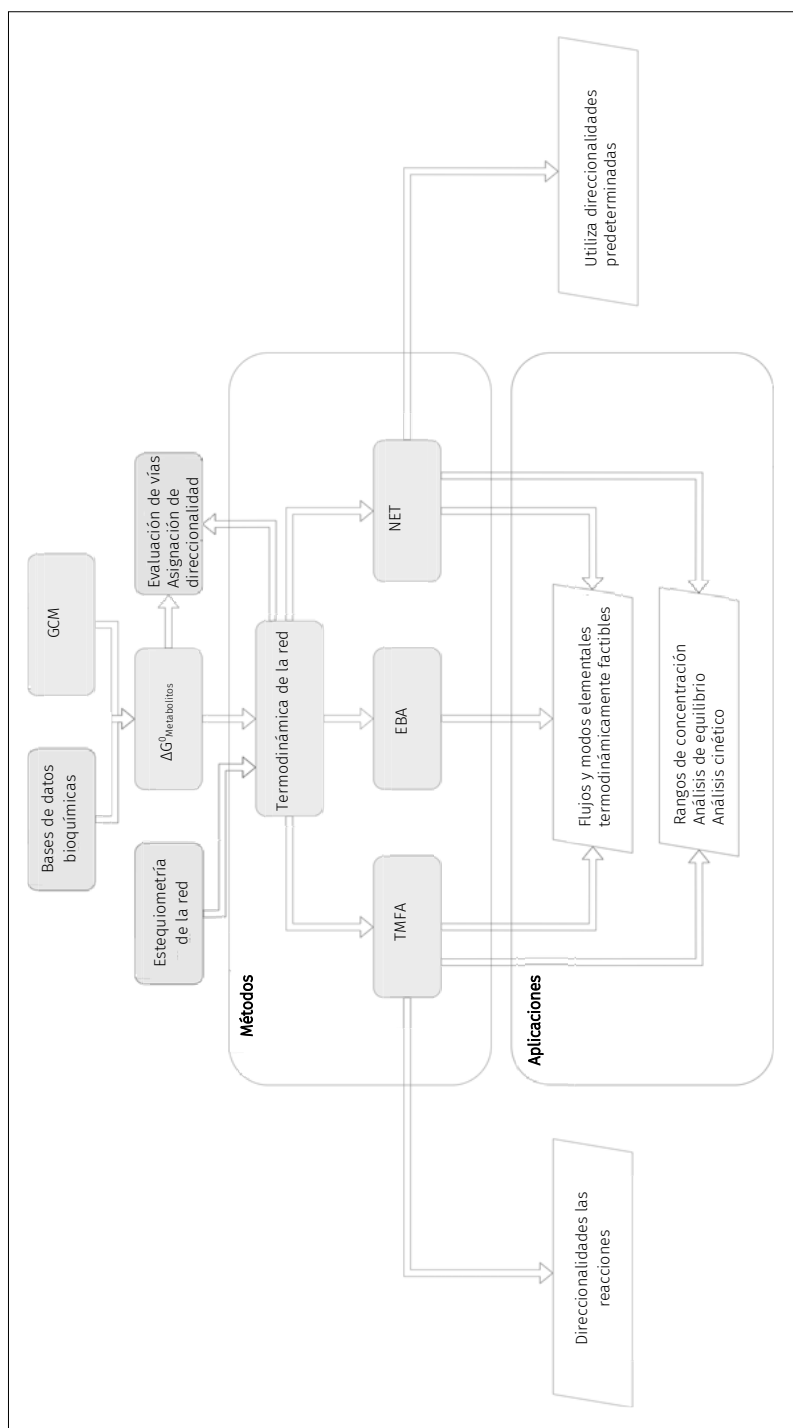


Figura 4.7. Modelado termodinámico de redes metabólicas e integración de la termodinámica en redes metabólicas.

FBA ignora la imposición de la ley de ciclos factibles, que es análoga a la segunda ley de Kirchhoff para circuitos eléctricos, ya que establece que las fuerzas impulsoras termodinámicas alrededor de un bucle metabólico deben sumar cero. Consecuentemente, no puede haber un flujo neto alrededor de un ciclo cerrado en una red en estado estacionario. Se han desarrollado varios métodos (ver Tabla 4.3) para detectar bucles que son intensivos desde el punto de vista computacional y necesitan del conocimiento de la energía libre de Gibbs de cada componente de la red para imponer restricciones adicionales sobre la direccionalidad de la reacción, así como sobre las concentraciones termodinámicamente factibles (Henry *et al.*, 2007). Este enfoque se ha utilizado para calcular ciclos reguladores potenciales, determinar la direccionalidad de una reacción y calcular los rangos de concentración factibles. Todos estos métodos requieren un conocimiento a priori del cambio de energía libre estándar de las reacciones o las energías estándar de formación de todos los metabolitos en la red. Estos valores pueden encontrarse en bibliografía o en bases de datos o pueden estimarse computacionalmente con métodos como la teoría de contribución de grupos. Sin embargo, la falta de precisión y cobertura en algunos casos puede plantear ciertos desafíos por la incertidumbre sobre la verosimilitud de los resultados. Un método más simple para incorporar las restricciones a la ley de ciclos que no requiere datos adicionales como concentraciones de metabolitos y energías libres mediante un modelo MILP denominado FBA sin ciclos. La solución al problema MILP modificado surge de la resolución del problema FBA, con restricciones adicionales para no permitir la inclusión de flujos en la red que contengan bucles.

Tabla 4.3. Modelos basados en restricciones termodinámicas.

Modelo basado en restricciones termodinámicas		
Nombre	Especificación	Referencia
Análisis de balance de energía	Identifica la dirección de reacciones y determina límites a los flujos utilizando las energías libres de Gibbs.	(Beard <i>et al.</i> , 2002)
Análisis termodinámico de redes	Identifica el rango factible de flujos o el rango de concentraciones factibles utilizando las direcciones de flujo y las energías libres de Gibbs.	(Kummel <i>et al.</i> , 2006)
Análisis de flujos metabólicos basados en termodinámica	Identifica direcciones de flujos rangos permisibles de flujos y de concentración utilizando programación MILP y energías libres de Gibbs.	(Henry <i>et al.</i> , 2007)

La condición «sin bucle»

Primero debe considerarse el problema de determinar si una solución de flujo dada, $\mathbf{v}$, contiene un bucle. Para que $\mathbf{v}$ satisfaga la ley del ciclo, las energías de reacción alrededor de cualquier ciclo deben sumar cero. Esta condición se puede escribir de manera concisa como $\mathbf{v}^T \mathbf{G} = 0$, donde $\mathbf{G}$ es un vector de energías libres para cada reacción. En principio, el análisis de vías extremas y de modos elementales se puede utilizar para identificar los ciclos. Estos métodos han demostrado que el número de ciclos crece exponencialmente con el tamaño de la red, y que no es viable enumerar todos los bucles posibles para redes de mediana a gran escala.

Como se muestra en la figura 4.8, todos los bucles se encuentran dentro de la red interna, $\mathbf{S}_{int}$. Cualquier ruta de estado estacionario en $\mathbf{S}_{int}$ es un bucle, y todas esas rutas se pueden expresar como una combinación lineal de la base nula de $\mathbf{S}_{int}$. Todos los bucles se pueden expresar en la forma $\mathbf{v} = \mathbf{N}_{int} \mathbf{a}_i$, donde $\mathbf{N}_{int} = \text{null}(\mathbf{S}_{int})$ y a_i son pesos. Si se puede demostrar que $\mathbf{N}_{int}^T \mathbf{G} = 0$, entonces $\mathbf{v}^T \mathbf{G} = 0$ para todos los bucles $\mathbf{v}$. La condición sin bucles forma un problema lineal. Un vector de variables continuas $\mathbf{G}_i$ indica la fuerza motriz de cada reacción. Se puede pensar que esta cantidad es análoga a la energía libre ΔG_r de cada reacción es tal que el *signo* $\mathbf{G} = \text{signo } \Delta G_r$, aunque numéricamente pueden ser bastante diferentes. Un bucle se define puramente por el signo (dirección) de la distribución de flujo.

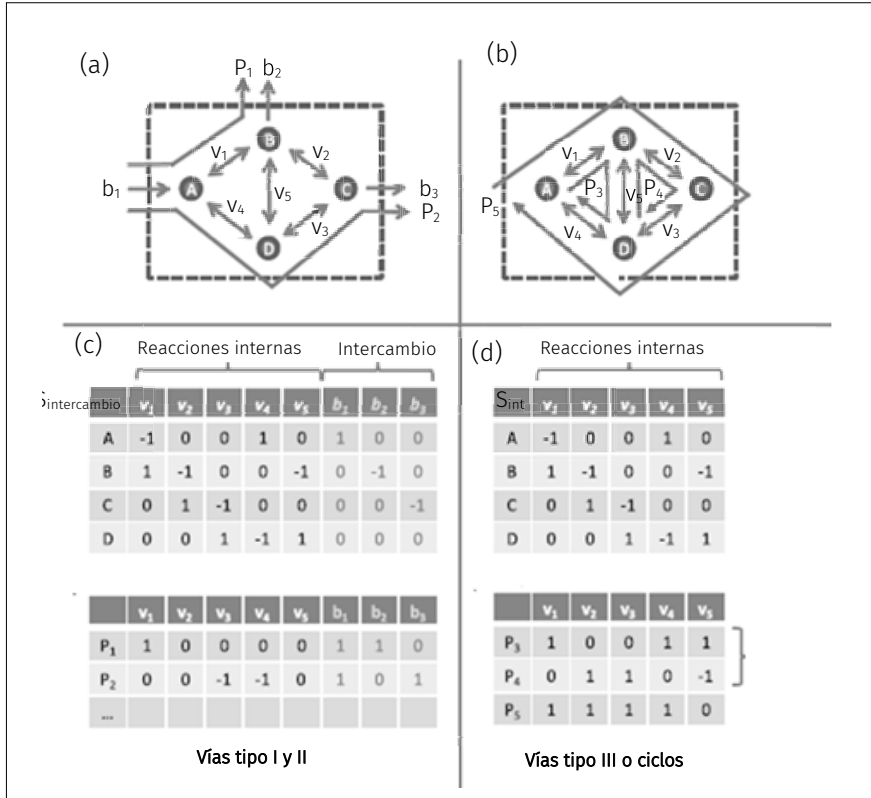


Figura 4.8. Bucles en las redes metabólicas. (a) y (c) presentan una red para ilustrar redes metabólicas con y sin bucles, en las que la red contiene cinco reacciones internas y tres de intercambios. Los cuadros (b) y (d) representan a la parte interna de la red. Las vías (v_1 - v_5) en estado estacionario se superponen y las vías de tipo I y II utilizan reacciones de intercambio. Un subconjunto parcial de estas vías se enumera en el cuadro (c). Las vías de tipo III no contienen reacciones de intercambio y forman un conjunto de bucles. En este ejemplo hay vías de tipo III (d). P_3 y P_4 forman la base del espacio nulo interno (N_{int}), y P_5 se puede escribir como una combinación lineal de las otras dos vías de tipo III.

Por lo tanto, si $N_{int} G_i = 0$, no hay bucle presente. Para verificar que una distribución de flujo no contiene bucles, se debe encontrar una solución a $N_{int} G = 0$ que satisfaga las siguientes restricciones:

$$G_i < 0 \text{ para todo } v_i > 0 \quad (4.26.1)$$

$$G_i > 0 \text{ para todo } v_i < 0 \quad (4.26.2)$$

$$G_i \in R \text{ para todo } v_i = 0 \quad (4.26.3)$$

$$N_{int} G = 0 \quad (4.26.4)$$

En la práctica, es necesario restringir G_i para que sea estrictamente positivo o estrictamente negativo con el fin de evitar una solución degenerada $G_i = 0$ para todo i . Este requisito es otra razón por la cual G_i no puede interpretarse directamente como valores de variación de energías libres ΔG_i aunque sea un proxy de la misma. La siguiente corrección restringe G_i a $[1000, 1]$ o $[1, 1000]$, y es posible que G_i nunca sea con exactitud cero:

$$-1000 < G_i < -1 \text{ para todo } v_i > 0 \quad (4.27.1)$$

$$1 < G_i < 1000 \text{ para todo } v_i < 0 \quad (4.27.2)$$

$$G_i \in R \text{ para todo } v_i = 0 \quad (4.27.3)$$

$$\mathbf{N}_{\text{int}} \mathbf{G} = 0 \quad (4.27.4)$$

Si existe una solución $\mathbf{v}$, entonces esta no contiene bucles. De lo contrario $\mathbf{v}$ contiene un bucle. A diferencia de la mayoría de los problemas lineales, el objetivo ($\max \mathbf{c}^T \mathbf{G}$) no tiene importancia porque solo la factibilidad es relevante. Las restricciones correspondientes a la ley del bucle se pueden agregar a casi cualquier problema de modelado metabólico con restricciones, siempre que este problema contenga una variable, v_i , para cada de los flujos internos en el modelo. La única adición necesaria es la condición que asegura $\text{signo}(\mathbf{v}) = -\text{signo}(\mathbf{G})$. Esto se logra agregando una variable binaria (a_i) para cada reacción interna. El conjunto completo de restricciones se puede expresar de la siguiente manera:

$$y_k = \begin{cases} 1 & \text{si } v_i > 0 \\ 0 & \text{si } v_i < 0 \end{cases} \quad (4.28.1)$$

$$G_i > 0 \text{ si } a_i < 0 \quad (4.28.2)$$

$$G_i < 0 \text{ si } a_i > 0 \quad (4.28.3)$$

$$\mathbf{N}_{\text{int}} \mathbf{G} = 0 \quad (4.28.4)$$

Lo que implica el siguiente problema MILP:

$$-1000a_i + (1 - a_i) \leq G_i \leq -a_i + 1000(1 - a_i) \quad (4.29.1)$$

$$-1000(1 - a_i) \leq v_i \leq 1000 a_i \quad (4.29.2)$$

$$\mathbf{N}_{\text{int}} \mathbf{G} = 0 \quad (4.29.3)$$

$$a_i \in \{0, 1\}; G_i \in R$$

Para evitar soluciones degeneradas, no debe permitirse que G_j sea cero. Estas restricciones se pueden agregar a casi cualquier método basado en restricciones. Por ejemplo, la formulación completa para el análisis de balance de flujo sin bucles (ll-FBA) es la siguiente:

Minimizar

$$c_j v_j \quad (4.30.1)$$

sujeto a

$$\sum_{k \in K} S_{kj} v_k = 0 \quad \forall j \in J \quad (4.30.2)$$

$$lb_j \leq v_j \leq ub_j \quad \forall j \in J \quad (4.30.3)$$

$$-1000a_i + (1 - a_i) \leq G_i \leq -a_i + 1000(1 - a_i) \quad \forall j \in J \quad (4.30.4)$$

$$-1000(1 - a_i) \leq v_i \leq 1000 a_i \cdot N_{int} \quad G = 0 \quad \forall j \in J \quad (4.30.5)$$

$$a_i \in \{0,1\}; G_i \in R, i \in R_{internas}$$

donde en el coeficiente S_{kj} de la matriz estequiométrica; j indica a las reacciones; k indica los metabolitos internos; LB_j y UB_j son los límites inferior y superior, respectivamente, para cada reacción j y c_j es el coeficiente de optimización asociado a cada reacción $j \in J$.

El análisis de variación de flujo sin bucle se realiza con el uso del método ll-FBA descrito anteriormente, seguido de la maximización y minimización secuencial de cada reacción en el modelo. Este cálculo se realiza con y sin las restricciones de la ley de bucle. Las reacciones en las que el rango ($\max v_i - \min v_i$) difiere en $>10^6$ mmol $gDW^{-1}h^{-1}$ entre FVA y ll-FVA se clasificaron como reacciones en bucle afectadas por las restricciones de la ley del bucle.

MODELADOS CINÉTICOS DE REDES METABÓLICAS

Los modelos cinéticos dinámicos intentan proporcionar una descripción mecanística de un biosistema en términos de la actividad enzimática intracelular y de los balances de masa para los metabolitos intracelulares. Las ecuaciones de conservación de materia para cada uno de los componentes forman un sistema de ecuaciones diferenciales ordinarias (EDO) y los valores de los flujos y concentraciones intracelulares se obtienen mediante la integración numérica de dichas ecuaciones de conservación. Además, este tipo de modelos matemáticos pueden simular cambios en la activación relativa de reacciones enzimáticas en función de los parámetros del sistema

y las concentraciones iniciales. La formulación matemática general se da por la siguiente ecuación:

$$\frac{dx}{dt} = Sv \quad (4.31.1)$$

$$v_j(t) = f_j(x, u, \theta) \quad x(t_0) = x_0 \quad (4.31.2)$$

$$x_i(t) \int_{t_0}^{t_f} (\sum_{j=1}^n s_{ij} v_j) dt + x_0 = \int_{t_0}^{t_f} (\sum_{j=1}^n s_{ij} f_j(x, u, \theta)) dt + x_0 \quad (4.31.3)$$

donde el vector $\mathbf{x}(t)$ es un vector de m de variables de estado dependientes del tiempo que representa a la concentración de metabolitos extracelulares e intracelulares, el vector $\mathbf{v}$ es un vector de flujos de reacción que depende de $\mathbf{x}(t)$, y $\mathbf{u}(t)$ representa a un vector de entradas reguladoras, mientras que θ que representa al conjunto de parámetros cinéticos. En algunos casos, se agregan ecuaciones algebraicas adicionales para reflejar a restricciones en las fracciones totales, cambios de volumen y especificación del valor de variables intensivas y complementarias. El vector de flujo es equivalente al vector de velocidades de reacciones bioquímicas y se expresa en unidades de mol de cada componente dividida por el producto de la masa de biomasa por el tiempo. La inclusión de ecuaciones cinéticas introduce no linealidades en el modelo del sistema pero brinda la capacidad de computar cambios de concentración y de flujos intracelulares en modo transiente y no solo en estado estacionario. La función f_j proporciona una descripción mecanística de las velocidades de reacción de la red metabólica. La formulación matemática está basada en teorías cinéticas, termodinámicas y en los principios fundamentales, es decir, en los balances de masa de los metabolitos. Usualmente, la solución numérica de la integración del sistema de ODE es bastante problemática debido a la gran cantidad de estados y parámetros, la naturaleza no lineal de las velocidades de reacción y la rigidez causada por la diferencia entre la escala de tiempo de diferentes variables del sistema y de sus derivadas con respecto del tiempo. Abordar problemas que usan modelos cinéticos en modelos metabólicos a gran escala, es decir, modelos a escala genómica, ha sido un nuevo foco de investigación. A continuación, se presentan las formulaciones matemáticas de modelos cinéticos más relevantes y se discuten sus ventajas e inconvenientes.

Formatos cinéticos aproximados y cuantificación de la regulación metabólica

Los enfoques alternativos al modelado cinético mecanicista han surgido sobre la base de la restricción de la homeostasis. Por lo tanto, el desarrollo de formatos cinéticos aproximados proporciona ecuaciones matemáticas que permiten aproximar la cinética de las reacciones en la vecindad de un

estado metabólico de referencia, con un número reducido de parámetros cinéticos. Las cinéticas aproximadas usualmente tienen cuatro características en común: proporcionalidad entre la velocidad de reacción y la concentración de una enzima, capacidad para alcanzar fase estacionaria, la velocidad de reacción está expresada como función de la concentración de ciertos sustratos e involucra al menor número posible de parámetros cinéticos. También pueden proporcionar una solución analítica en estado estacionario. Los formatos cinéticos aproximados incorporan explícitamente las concentraciones de metabolitos y proporcionan un marco para estimar la regulación metabólica.

Análisis de control metabólico (MCA)

El análisis de sensibilidad respecto de los parámetros de un modelo metabólico tiene el fin de determinar la influencia de los mismos en la simulación del modelo (Torres y Voit, 2009). Los valores iniciales de las concentraciones de metabolitos y los valores de los parámetros cinéticos, así como las limitaciones de las velocidades de flujo pueden considerarse factores que influyen en el comportamiento de un sistema. El análisis de sensibilidad (SA) se divide en dos grandes grupos; el SA local y el SA global. Los métodos locales de SA consideran que solo una entrada varía a la vez y los resultados de una perturbación se observan en las salidas, mientras que el resto de parámetros se mantienen sin cambios. Por el contrario, en un SA global, todas las entradas varían simultáneamente y el efecto resultante en la salida del modelo puede investigarse pero con un gran costo computacional. El análisis de sensibilidad local se usa para caracterizar el efecto de los cambios de parámetros en la solución de modelos dinámicos en la vecindad de un determinado estado de referencia. La linealización sobre un estado de referencia en general se expresa como una aproximación de la serie de Taylor de primer orden:

$$\frac{dx_i}{dt} \approx \frac{x_i(t, p_g^0 + \Delta p_g) - x_i(t, p_g^0 - \Delta p_g)}{2\Delta p_g} \quad (4.32)$$

donde c_i es una propiedad del sistema metabólico, tal como la concentración de metabolito, p^0 es el conjunto inicial de parámetros y p_g son parámetros sujeto a cambios. La diferenciación explícita se puede usar cuando el modelo establece la relación entre la salida y la entrada, en este caso y la función de sensibilidad esté representada por una solución analítica que se puede usar para calcular el coeficiente de sensibilidad. Cuando la salida es una función implícita de los parámetros de entrada, se puede utilizar la diferenciación numérica. En la práctica, los coeficientes de sensibilidad locales se calculan mediante una aproximación numérica. La simulación se usa para

determinar la evolución de una variable de interés como puede ser el flujo metabólico en estado estacionario y luego se usan varias formas numéricas para calcular su valor local absoluto y relativo. Uno de los marcos de SA más conocidos en el contexto del modelado metabólico es el análisis de control metabólico (MCA), que se usa para estudiar las interacciones entre las diferentes vías metabólicas. Mediante el uso de MCA, el efecto de la perturbación de un parámetro en los flujos metabólicos o en las concentraciones de metabolitos.

Tabla 4.4. Coeficientes utilizados en el análisis de control metabólico.

Tabla de coeficientes de MCA		
Nombre	Formulación matemática	Descripción
Coeficientes de control	Flujo: Metabolito intracelular:	Los coeficientes miden el cambio relativo de flujo y concentración debido a un cambio relativo del nivel de actividad enzimática.
Coeficientes de respuesta	Flujo: Metabolito intracelular:	Los coeficientes miden el efecto del cambio de un parámetro externo en los flujos y concentraciones intracelulares.
Coeficientes de elasticidad	Velocidad de reacción intracelular:	La elasticidad es una medida local que cuantifica el cambio relativo de una velocidad de reacción frente a un cambio en la concentración de un metabolito.

Los coeficientes principales en MCA se definen en la Tabla 4.4. En términos teóricos, la estabilidad dinámica de un sistema metabólico implica que tras la perturbación de un sistema en estado estacionario, los estados del sistema, es decir las concentraciones de metabolitos y los flujos intracelulares, eventualmente regresan al mismo estado estacionario. Cuando se perturba un parámetro, la estabilidad estructural asegura que el sistema eventualmente regrese al mismo estado estacionario. Si el sistema es inestable, tales perturbaciones resultan en la divergencia de algunos o todos los estados del sistema. Entonces, el estudio de un sistema metabólico se limita a la vecindad de dicho estado estacionario. Por lo tanto, en la práctica, MCA tiene una capacidad limitada de cuantificación ya que la teoría solo es válida para pequeñas variaciones en los parámetros, en la que la linealización de la ecuación (4.32) es una buena aproximación de comportamientos cinéticos no lineales.

Tabla 4.5. Formatos cinéticos aproximados para MCA y Teoría de sistemas bioquímicos (BST)

Formatos cinéticos aproximados basados en MCA y BST	
Tipo de aproximación	Formulación matemática
Log–lin	$\frac{v_j}{J_j^0} = 1 + \ln\left(\frac{e_j}{e_j^0}\right) + \sum_{i=1}^m \varepsilon_{ji}^0 \ln\left(\frac{x_i}{x_i^0}\right)$
Lin–log	$\frac{v_j}{J_j^0} = \frac{e_j}{e_j^0} \left[1 + \sum_{i=1}^m \varepsilon_{ji}^0 \ln\left(\frac{x_i}{x_i^0}\right) \right]$
Sistema–S Ley general de acción de masa (GMA)	$\frac{dx_i}{dt} = V_i^+ - V_i^- = \alpha_i \prod_j^m x_j^{f_{ij}^-} - \beta_i \prod_j^m x_j^{h_{ij}^-}$ $v_j = \gamma_{+j} \prod_i^m x_i^{f_{ij}^-} - \gamma_{-j} \prod_i^m x_i^{f_{ij}^+}$

La formulación matemática rigurosa de MCA busca en estimar coeficientes para relaciones globales. Estos formatos cinéticos aproximados son denominados *log–lin* y *lin–log*. Sin embargo, por lo general se formulan directamente a partir de la estequiometría de la red metabólica tal como se indica en la Tabla 5.4.

Teoría de sistemas bioquímicos (BST)

La representación mediante la ley de potencia de ecuaciones diferenciales ordinarias es el instrumento clave de la teoría de sistemas bioquímicos (BST) desarrollada por Michael Savageau y Eberhard Voit en la década de 1960 previo al desarrollo de MCA. Ambos formatos se utilizan para representar dinámicas de cambios en procesos bioquímicos y grupos de metabolitos, respectivamente. La teoría comparte gran parte de la filosofía de modelado con MCA. BST brinda la misma solución de estado estacionario y las mismas propiedades de estabilidad local, en el mismo estado de referencia, lo que implica que los parámetros de BST pueden derivarse de los coeficientes utilizados en MCA (órdenes cinéticos en BST a partir de elasticidades en MCA). En BST, los órdenes cinéticos juegan un papel importante y a diferencia de la cinética de acción de masas, pueden adquirir valores negativos o no enteros. Por ejemplo, se puede incluir una variable inhibitoria con un orden cinético negativo. Sin embargo, estas ecuaciones predicen una reacción que puede llevar a infinito a la concentración de ciertos componentes, es decir que no tienen capacidad para simular un fenómeno de saturación. Torres y Voit (2009) sostienen que para el diseño metabólico y para el estudio teórico de los principios de funcionamiento, el formato BST proporciona un mejor

punto de partida porque permite al usuario realizar cálculos algebraicos en estados estacionarios y unos análisis de estabilidad y sensibilidad más sencillos utilizando conceptos de la teoría de control de procesos.

Modelos metabólicos cinéticos con regulación enzimática

La regulación de los flujos que permite actividades celulares continuas se asemeja a ciertas estrategias de control para mantener, aumentar o disminuir la velocidad de producción o consumo de compuestos biomoleculares, es decir, metabolitos, enzimas o proteínas de señalización. La represión y la inducción de retroalimentación son dos ejemplos de estos escenarios de regulación a nivel epigenético, es decir del estudio de interacciones a nivel genético que afectan la expresión génica. Las interacciones a nivel metabólico son la inhibición por retroalimentación y la activación enzimática. La inhibición competitiva y la regulación alostérica son dos formas predominantes, entre otras, que se han utilizado para modelar la regulación de la actividad enzimática. Los modelos cinéticos de redes metabólicas con parámetros específicos como las constantes de afinidad y las velocidades máximas de reacción pueden simular el comportamiento metabólico no lineal de biosistemas en un horizonte de tiempo dado, es decir, que pueden simular la dinámica del crecimiento celular. Dichos modelos dinámicos no consideran regulación transcripcional o la consideran de forma no cinética, por ejemplo, a través de lógica booleana o reglas lógicas multivaluadas. Al modelar un sistema biológico regulado, se debe prestar especial atención a que los tiempos de las reacciones y los procesos involucrados en su totalidad puede abarcar varias escalas de tiempo que van desde nanosegundos a días. Como regla general, las reacciones que ocurren con una constante de tiempo característica que es de un orden de magnitud mayor a la constante de tiempo del sistema se consideran como «congeladas», y la concentración de metabolitos asociados con ellas se observa como variables fijas. Sin embargo, es más complicado manejar reacciones más rápidas y en este caso se deben utilizar dos suposiciones para aproximar las concentraciones. En primer lugar, la suposición de que se alcanza rápidamente el estado estacionario. Esta suposición establece que cuando la enzima está disponible en cantidad catalítica, el complejo enzima-sustrato alcanza un estado estacionario tan rápido como para que se pueda asumir que su concentración permanece constante. La suposición de un equilibrio rápido implica que el sustrato alcanza rápidamente el equilibrio con el complejo enzima-sustrato, siendo las velocidades en ambas direcciones de la reacción de unión más rápidas que la velocidad de conversión del complejo enzima-sustrato en complejo enzima-producto. Cabe mencionar que los tiempos de regulación a nivel metabólico suelen ser notablemente

muy cortos en comparación con las regulaciones impuestas a nivel epigenético. Se cree que el primero tiene lugar en el orden de los minutos, mientras que el segundo ocurre en el orden horas. En consecuencia, la regulación del nivel metabólico de la actividad enzimática puede verse como un ajuste fino, a diferencia del control «grueso» asociado con los mecanismos. En la figura 4.9 se ilustran las escalas de tiempo comparativas de operaciones celulares junto con una representación esquemática de mecanismos reguladores a nivel metabólico y genético.

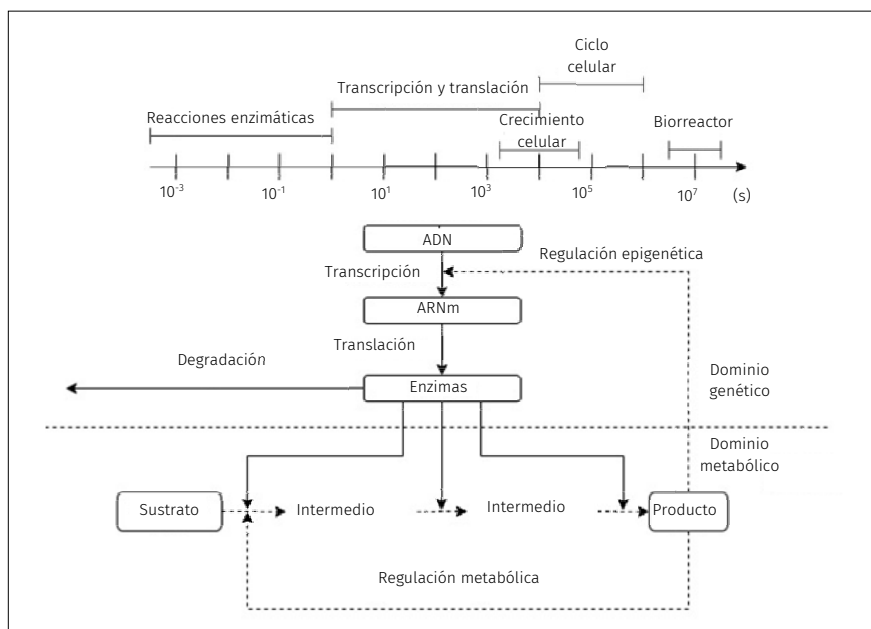


Figura 4.9. Escalas temporales y esquema de regulación. (a) Escala de tiempo de las operaciones celulares. (b) Mecanismo regulador interactivo a nivel genético y metabólico.

Actividades

1. Formule el problema de optimización de Análisis de Balance de Flujo (FBA) para la red metabólica de demostración representada por la siguiente figura. Suponga que A es un recurso limitante que se puede absorber con un flujo máximo de 10 mmol gDW h , B es un recurso ilimitado, F es el producto metabólico exportado al medioambiente y el sumidero «Sink» representa a la formación de biomasa. En este ejemplo, se maximiza el flujo metabólico de formación de biomasa.

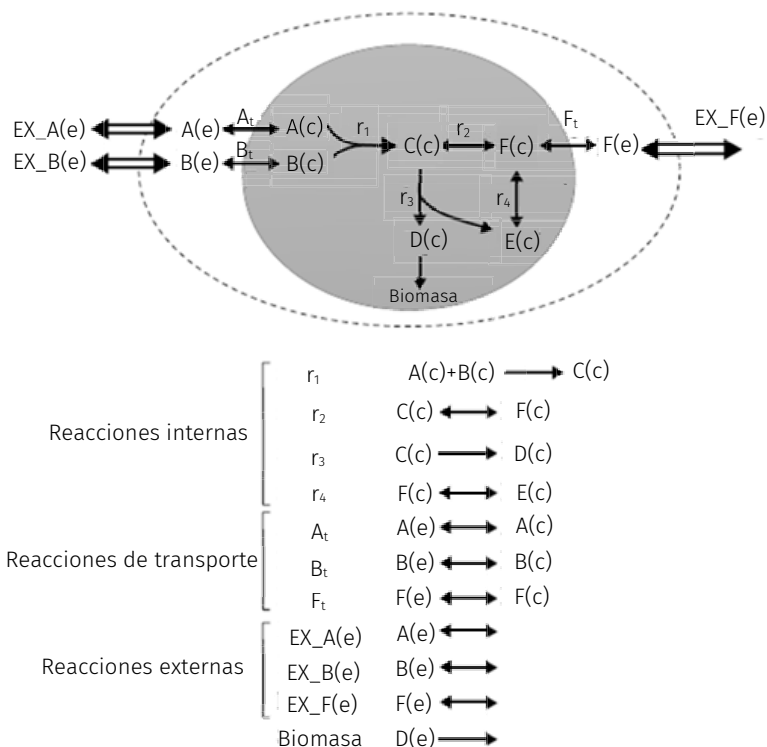


Figura E.4.1.

Construya la matriz estequiométrica que permite expresar en forma matricial las ecuaciones de balances de metabolitos suponiendo que A es el recurso limitante.

Plantee el problema FBA para la red metabólica y preséntelo como un problema de optimización lineal. Resuélvalo para computar el valor de v_{Sink} . A posteriori compute los rangos factibles de flujos para cada reacción mediante un análisis FVA.

- Deduzca las matrices estequiométricas de las siguientes redes metabólicas. En ambos gráficos los componentes de la red están representados por círculos.

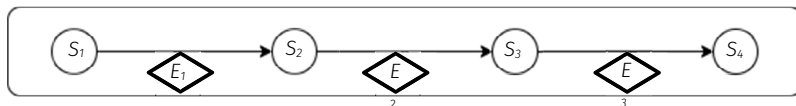


Figura E.4.2.2

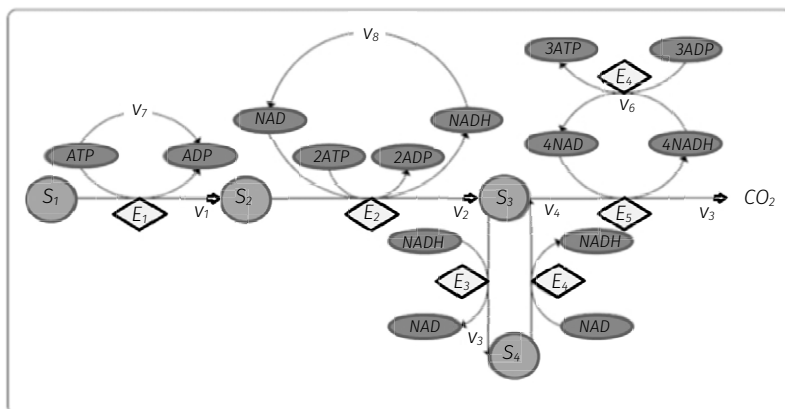


Figura E.4.2.3

REFERENCIAS BIBLIOGRÁFICAS

- BEARD, D.; LIANG, S. & QIAN, H.** (2002). *Energy Balance for Analysis of Complex Metabolic Networks*. Biophys. J., 83. [https://doi.org/10.1016/S0006-3495\(02\)75150-3](https://doi.org/10.1016/S0006-3495(02)75150-3)
- BURGARD, A.; NIKOLAEV, E. (...) MARANAS, C.** (2004). Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res*, 14 (2): 301-312. <https://doi.org/10.1101/gr.1926504>
- GUDMUNDSSON, S. & THIELE, I.** (2010). Computationally efficient flux variability analysis. *BMC Bioinformatics*, 11, 489. <https://doi.org/10.1186/1471-2105-11-489>
- HEIRENDT, L.; ARRECKX, S. (...) FLEMING, R.** (2019). *Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0*. Nature Publishing Group UK.
- HENRY, C.; BROADBELT, L. & HATZIMANIKATIS, V.** (2007). *Thermodynamics-Based Metabolic Flux Analysis*. Biophys J. 1; 92 (5): 1792-1805. <https://doi.org/10.1529/biophysj.106.093138>
- KUMMEL, A.; PANKE, S. & HEINEMANN, M.** (2006). Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Mol. Syst. Biol.*, 2, 2006.0034. <https://doi.org/10.1038/msb4100074>

- MAHADEVAN, R.; EDWARDS, J. & DOYLE III, F.** (2002). Dynamic Flux Balance Analysis of Diauxic Growth in *Escherichia coli*. *Biophysical Journal*, 83 (3), 1331–1340. [https://doi.org/10.1016/S0006-3495\(02\)73903-9](https://doi.org/10.1016/S0006-3495(02)73903-9)
- MARANAS, C. & ZOMORRODI, A.** (2016). *Optimization Methods in Metabolic Networks*. John Wiley & Sons.
- ORTH, J.; THIELE, I. & PALSSON, BØ.** (2010). What is flux balance analysis? *Nature Biotechnology*, 28, 245–248. <https://DOI:10.1038/nbt.1614>
- TORRES, N. & VOIT, E.** (2009). *Pathway Analysis and Optimization in Metabolic Engineering*. Cambridge University Press.

Lecturas adicionales

- KLIPP, E.; LIEBERMEISTER, W. (...) KOWALD, A.** (2016). *Systems Biology: A Textbook* (2nd edition). Wiley-VCH.
- PALSSON, B.** (2005). *Systems Biology. Properties of Reconstructed Networks*. Cambridge University Press.
- PALSSON, B.** (2011). *Systems Biology. Simulation of Dynamic Network States*. Cambridge University Press.
- PALSSON, B.** (2015). *Systems Biology. Constraint-based Reconstruction and Analysis*. Cambridge University Press.

5 Modelado a escala macroscópica

Los procesos biotecnológicos generalmente involucran a las fases líquida y gaseosa que en conjunto forman la denominada fase *abiótica* mientras que la población de células forma la *fase biótica*. Las propiedades de estas fases se caracterizan por medio de variables macroscópicas tal como concentraciones de componentes relevantes y variables intensivas como la temperatura y el *pH*. Las reacciones catalizadas por microorganismos ocurren en la fase líquida y los componentes relevantes del sistema son:

- La masa celular o biomasa (X) que es autocatalítica y se sintetiza a partir de los substratos.
- Los substratos que actúan como fuente de nutrientes y de energía para la biomasa.
- Los productos metabólicos.
- Los gases disueltos (O_2 y CO_2) en la fase líquida que se intercambian con la fase gaseosa.

El vector de concentraciones de componentes relevantes será entonces:

$$\mathbf{x} = [\mathbf{x}_S \ \mathbf{x}_P \ X \ x_{O_2} \ x_{CO_2}] \quad (5.1)$$

En donde $\mathbf{x}_P$ y $\mathbf{x}_S$ pueden involucrar a más de un componente y por lo tanto pueden ser también vectores. Los modelos de una fermentación involucran a tres diferentes tipos de ecuaciones:

- Las ecuaciones de balance de materia para cada componente del sistema.
- Las ecuaciones estequiométricas de rendimiento que describen la relación entre velocidades de consumo y de producción de los componentes que intervienen en el proceso.
- Las ecuaciones que describen la cinética del proceso biotecnológico.

Las ecuaciones cinéticas y de rendimiento son independientes del tipo de biorreactor utilizado y constituyen una descripción del metabolismo microbiano, y de cómo este es afectado por el ambiente fisicoquímico en el que se desarrollan las células. La mayor diferencia entre estos tres conjuntos de ecuaciones es la confiabilidad con la cual estas pueden ser utilizadas. Las ecuaciones de balance de materia dependen del tipo de reactor, pero pueden

ser usadas con confianza debido a que no contienen ningún tipo de suposición previa. Las ecuaciones de rendimiento, basadas en balances de elementos químicos, contienen necesariamente algún tipo de suposición acerca de las composiciones elementales de la biomasa y de algunos productos y substratos que no están definidas con rigurosidad. A diferencia de las ecuaciones de conservación y de rendimiento, que tienen base en los principios fundamentales, las ecuaciones cinéticas son puramente empíricas y son las menos confiables de los tres tipos de ecuaciones. Por lo tanto deberían ser empleadas lo menos posible.

BALANCES MACROSCÓPICOS

Los balances se formulan para una propiedad extensiva de un sistema. La estructura de las ecuaciones de balance es siempre la misma y puede ser expresada en palabras de la siguiente forma:

$$\text{Acumulación en el sistema} = \text{Entrada} - \text{Salida} + \text{Conversión} \quad (5.2)$$

Esta expresión postula que la cantidad total de una propiedad extensiva acumulada en el sistema cambia por dos tipos de procesos; la transformación en el sistema, es decir las reacciones químicas o **conversión** y por el intercambio con el medio ambiente (**Entrada – Salida**). Para aplicar las ecuaciones de balance a un sistema biotecnológico, es conveniente utilizar el concepto de vector de estados que se corresponde al vector de concentraciones de componentes relevantes definido en las ecuaciones (5.1).

ECUACIONES DE BALANCE DE MATERIA TOTAL Y DE COMPONENTES

En el vector de estados $\mathbf{x}$, la concentración de cada compuesto químico presente en el sistema está identificada por cada uno de los elementos del vector. Para su utilización, primeramente hay que plantear un balance de materia total:

$$\frac{d(\rho V)}{dt} = \sum_{\text{Entradas } i} \rho_i F_i - \sum_{\text{Salidas } o} \rho_o F_o \quad (5.3)$$

Luego se plantea un balance de materia para cada uno de los componentes A del vector de estados $\mathbf{x}$:

$$\frac{d(x_A V)}{dt} = \sum \text{Entradas } i x_A^i F_i - \sum \text{Salidas } o x_A^o F_o + r_A V \quad (5.4)$$

$$\mathbf{x}_A = [x_S \ x_P \ x_{O_2} \ x_{CO_2}]$$

donde:

V: Volumen de la fase líquida en el reactor.

F_i : Flujo de entrada al reactor.

F_o : Flujo de salida del reactor.

ρ_i : Densidad de la corriente de entrada i.

ρ_o : Densidad de la corriente de salida o.

ρ : Densidad de la solución acumulada en el reactor.

x_A^i : Concentración del componente A en el flujo de entrada.

x_A^o : Concentración del componente A en el flujo de salida.

x_A : Concentración del componente A en el reactor.

r_A : Velocidades neta de aparición o desaparición del componente A debido a las reacción bioquímicas.

A partir del balance de materia total, se pueden particularizar los siguientes modos de operación del bioproceso:

Modo batch: $F_i = F_o = 0 \Rightarrow V = \text{cte.}$

Modo fed-batch: $F_i \neq 0; F_o = 0 \Rightarrow V \neq \text{cte.}$

Modo continuo: $F_i = F_o = \text{cte.} \Rightarrow V = \text{cte.}$

La velocidad de acumulación de una cantidad dentro de los límites de un sistema es la diferencia entre la velocidad a la que esta cantidad ingresa al sistema y la velocidad a la que sale del mismo, más la velocidad de su producción interna neta. En los procesos (bio)químicos, las cantidades fundamentales que se conservan son la masa, la energía y la cantidad de movimiento.

Las reacciones químicas independientes ($\mathbf{r}$) del sistema se especifican por medio de sus expresiones estequiométricas y la velocidad neta de conversión de cada una de las especies químicas del sistema (r_A) puede ser expresada

en términos del vector de velocidades independientes de reacción $\mathbf{r}$ y de la matriz de estequiometrías de la siguiente forma:

$$\mathbf{r}_A = \mathbf{r}\mathbf{\Lambda} \quad (5.5)$$

Si existen m reacciones independientes, $\mathbf{r}$ es un vector fila m -dimensional $[r_1, \dots, r_m]$ en el cual cada elemento especifica a cada una de las velocidades de reacción. Cada fila de la matriz $\mathbf{\Lambda}$ expresa la estequiometría de la reacción en función del número de moles de cada compuesto convertido por unidad de velocidad de reacción. Por lo tanto $\mathbf{\Lambda}$ es una matriz $m \times n$, siendo n el número de metabolitos que intervienen en el sistema, con la siguiente estructura:

$$\mathbf{\Lambda} = (a_{ij}) \quad \begin{array}{l} j = 1, \dots, m \text{ (} m, \text{ número de reacciones independientes)} \\ i = 1, \dots, n \text{ (} n, \text{ número de metabolitos)} \end{array} \quad (5.6)$$

MODELADO CINÉTICO Y TIPOS DE MODELOS

El modelado cinético expresa la correlación entre las velocidades y las concentraciones de reactivos o productos que, cuando se insertan en los balances de materia, permiten predecir la conversión de sustratos y el rendimiento de productos individuales. Si las ecuaciones que expresan las velocidades están correctamente modeladas, puede ser posible predecir el curso de un bioproceso completo en base a los valores iniciales para los componentes del vector de estado. El modelado cinético consiste en expresar relaciones funcionales entre las velocidades de reacción, las reacciones consideradas en el modelo y las concentraciones de los sustratos, de los productos metabólicos, de los constituyentes de la biomasa y de los metabolitos intracelulares de la biomasa:

$$v_i = f_i(c_s, c_p, X_{macro}, X_{met}, x) \quad (5.7)$$

Si durante el cultivo, la composición de la biomasa permanece constante, entonces las velocidades de las reacciones internas serán proporcionales. Este fenómeno se conoce como «crecimiento equilibrado». En este caso, el proceso de crecimiento puede describirse en términos de una sola variable que es la concentración de la biomasa X (g DW/l). Esta suposición es la base de los llamados modelos no estructurados que han demostrado durante 50 años ser adecuados para modelar y diseñar procesos de cultivo, en particular

cultivos en estado estacionario o batch. Sin embargo, estos modelos no estructurados en general suelen tener poca capacidad predictiva y, como tales, son de poco valor en estudios fundamentales y detallados de funciones celulares.

Modelos cinéticos no estructurados

El crecimiento de las células es el resultado de muchas reacciones químicas, incluidas las reacciones de alimentación, reacciones de biosíntesis y reacciones de ensamblaje. Previamente a la división celular, las células aumentan de tamaño, o extienden sus hifas, en el caso de microorganismos filamentosos, a medida que las macromoléculas se ensamblan para la formación de biomasa. La formación de biomasa se puede cuantificar midiendo el aumento del peso seco o midiendo *in situ* en tiempo real siguiendo el aumento de la turbidez del cultivo en una longitud de onda dada. El crecimiento celular generalmente se cuantifica por la llamada velocidad de crecimiento específica μ (h^{-1}), que para tal cultivo viene dada por:

$$\mu = \frac{1}{X} \frac{dX}{dt} \quad (5.8)$$

Donde X es la concentración de biomasa o el número de células por unidad de volumen. El resultado neto de las reacciones bioquímicas que ocurren dentro de una célula es la conversión de sustratos en biomasa y productos metabólicos. Claramente, el número de reacciones involucradas en la conversión de, por ejemplo, glucosa en biomasa y productos finales es muy grande y por lo tanto es conveniente adoptar una estructura simplificada, que se puede resumir de la siguiente manera:

Las *reacciones de ensamblaje* llevan a cabo modificaciones químicas de macromoléculas y se encargan de su transporte a ubicaciones preespecificadas en la célula y de su ensamblaje para formar estructuras celulares como paredes celulares, membranas, núcleo, etcétera.

Las *reacciones de polimerización* representan un enlace secuencial de moléculas activadas en cadenas poliméricas largas (ramificadas o no ramificadas). Estas reacciones conducen a la formación de macromoléculas a partir de un conjunto de bloques de construcción tales como aminoácidos, nucleótidos y ácidos grasos.

Las *reacciones de biosíntesis* producen los componentes básicos utilizados en las reacciones de polimerización. También producen coenzimas y factores metabólicos relacionados. Además, una gran cantidad de reacciones

biosintéticas ocurren en unidades funcionales llamadas vías biosintéticas, cada una de las cuales consiste en reacciones secuenciales que conducen a la síntesis de uno o más componentes básicos.

Las vías metabólicas se reconocen fácilmente y a menudo se controlan en bloque. Todas las vías biosintéticas comienzan con los metabolitos precursores a partir del cual se pueden sintetizar todos los componentes básicos de la célula. Algunas vías comienzan directamente con dicho metabolito precursor, otras indirectamente ramificándose de un producto intermedio o final de una ruta relacionada.

Las *reacciones de alimentación* producen los metabolitos precursores necesarios para la biosíntesis. Además, generan energía libre de Gibbs en forma de adenosina-trifosfato (ATP), que se utiliza para las reacciones de biosíntesis, polimerización y ensamblaje. Las reacciones de abastecimiento de energía incluyen a las denominadas vías catabólicas.

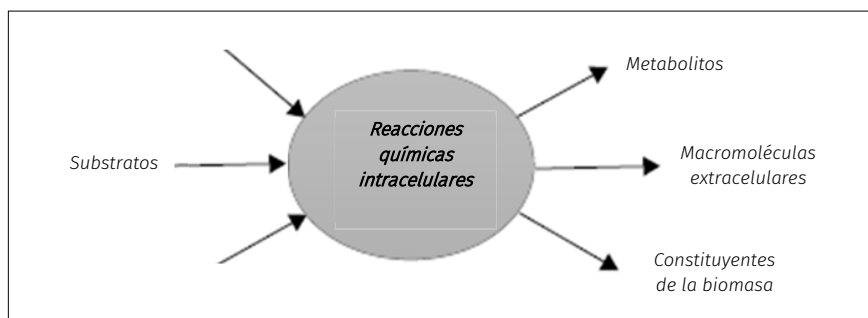


Figura 5.1. Ilustración representativa de la reacciones bioquímicas intracelulares en microorganismos.

Por lo tanto la conversión de, por ejemplo, glucosa en proteína celular se realiza a través de metabolitos precursores formados en las reacciones, a través de bloques de construcción (en este caso, aminoácidos) ensamblados en las reacciones biosintéticas, y finalmente a través de la polimerización de los bloques de construcción que son los aminoácidos. El número de metabolitos celulares es muy grande, pero solo representan una pequeña fracción de la biomasa total.

Tabla 5.1. Composición promedio de una célula de *E. Coli*. (Ingraham *et al.*, 1983)

Composición de una célula promedio de <i>Escherichia Coli</i>		
Macromolécula	% de peso seco	Tipos de moléculas
Proteína	55.0	1050
ARN	20.5	-
rARN	16.7	3
tARN	3.0	60
mARN	0.8	400
ADN	3.1	1
Lípidos	9.1	4
Lipopolisacáridos	3.4	1
Péptidos	2.5	1
Glucógeno	2.5	1
Otros metabolitos	3.9	?

En una descripción cuantitativa del crecimiento celular no es necesario considerar la cinética de todas las reacciones, lo cual conduce a una reducción masiva en el grado de complejidad del modelo matemático. Por lo tanto, la consideración de la cinética enzimática y la cinética de las reacciones individuales es necesaria solo cuando el objetivo del estudio es cuantificar la importancia relativa de una reacción particular en una vía metabólica específica conduciendo al planteo de modelos estructurados. En los modelos no estructurados, las «reacciones» dependen únicamente de las variables macroscópicas que describen las condiciones medioambientales y la única variable biológica considerada es la concentración de biomasa. A pesar de la simplificación extrema que estas suposiciones implican, muchos procesos biotecnológicos pueden ser satisfactoriamente descritos por este tipo de modelos. Sin embargo hay que tener presente que estos modelos no proporcionan una descripción detallada de los mecanismos involucrados en un bioproceso. En los modelos no estructurados, el crecimiento celular se asume en general como autocatalítico:

$$\frac{dX}{dt} = \mu X \quad (5.9)$$

donde la velocidad específica de crecimiento μ (h^{-1}) se modela como una función que depende tanto de la concentración de biomasa como de otras variables macroscópicas medibles.

$$\mu = f(X) \quad (5.10)$$

Usualmente y de acuerdo con observaciones empíricas μ suele mostrar signos de saturación con altas concentraciones de un substrato limitante C_S , en cuyo caso μ es máximo. Al contrario, $\mu = 0$ cuando $C_S \rightarrow 0$. Entonces μ puede plantearse como:

$$\mu = \mu_{MAX} f(x_S) \quad (5.11)$$

donde μ_{max} es una constante. Se han propuesto numerosas cinéticas específicas para la función $f(x_S)$. La Tabla 5.2 compila las funcionalidades más conocidas para las ecuaciones $f(x_S)$. Por otra parte, también se han propuesto cinéticas que tienen en cuenta al efecto inhibitorio de la presencia de sustratos y de productos sobre el crecimiento celular. Las mismas se presentan en la Tabla 5.3. En ciertos casos, la cinética de crecimiento celular puede ser independiente de la concentración de sustratos. Consecuentemente se utilizan, en forma empírica las ecuaciones presentadas en la Tabla 5.4. Estas últimas son útiles cuando el factor limitante es un componente desconocido o que no está considerado en el vector de componentes principales $\mathbf{x}$.

Tabla 5.2. Funciones de saturación para cinéticas dependientes de un substrato limitante (Dondo, 2012)

Nombre	$f(x_S)$
Blackman	$\min(1, K_S S)$
Monod	$\frac{S}{K_S + S}$
Teissier	$1 - e^{-K_S S}$
Moser	$\frac{S^n}{K_S^n + S^n}$
Contois	$\frac{S}{K_S X + S}$
Powel	$\frac{S - K_I \tau}{K_S + S - K_I \tau}$
Mason & Miles	$\frac{S}{K_S + S} + K_D S$
Vavilin	$\frac{S^n}{K_S^n S_0^n + S^n}$

Tabla 5.3. Funciones de inhibición por presencia de sustratos o productos (Dondo, 2012)

Nombre	$f(x_s)$
Inhibición por sustrato	
Andrews & Noack	$\frac{1}{K_S + S + \frac{S^2}{K_{iS}}}$
Webb	$\frac{S \left(1 + \frac{\beta S}{K_{iS}}\right)}{K_S + S + \frac{S^2}{K_{iS}}}$
Aiba <i>et al.</i>	$\frac{S}{K_S + S} e^{-\frac{S}{K_S}}$
Tipo Teissier	$e^{-\frac{S}{K_{iS}}} - e^{-\frac{S}{K_S}}$
Tseng & Wymann	$\frac{S}{K_S + S} K_{iS} (S - S_C)$
Inhibición por producto	
Dagley & Hinshelwood	$\frac{S}{K_S + S} (1 - KP)$
Holzber <i>et al.</i>	$1 - \frac{k_1}{\mu_{Max}} (P - k_2)$
Ghose & Tyagi	$\left(1 - \frac{P}{P_{Max}}\right)$
Aiba & Shoda	$\frac{S}{K_S + S} e^{-KP}$
Jerusalimsky & Neronova	$\frac{S}{K_S + S} \frac{K_P}{K_P + P}$
Levenspiel	$\frac{S}{K_S + S} \left(1 - \frac{P}{P_{Max}}\right)$

Tabla 5.4. Cinéticas de crecimiento independientes de la concentración de sustratos limitantes (Dondo, 2012)

Nombre	$f(x_s)$
Logística	$\left(1 - \frac{X}{X_{Max}}\right)$
Cui and Lawson	$\left(1 - \frac{X}{X_{Max}}\right) / \left(1 - \frac{X}{X_{Lim}}\right)$
Frame and Hu	$\left(1 - e^{\frac{X_{Max}-X}{X}}\right)$

Modelos cinéticos estructurados

En contraste con los modelos no estructurados, los modelos estructurados suministran información sobre el estado fisiológico de los microorganismos e incluyen, por ejemplo, información referida a cambios en la composición de la biomasa y/o en los mecanismos de adaptación a cambios medioambientales. Un modelo estructurado no solo debe incluir variables relacionadas a los procesos intracelulares, sino que también debe ser simple y, en la medida de lo posible, contener un mínimo número de parámetros a identificar. Se han propuesto numerosos tipos de modelos estructurados, los cuales se pueden clasificar en modelos de compartimentos, modelos cibernéticos y enfoques metabólico-regulatorios (Dondo, 2012). De ellos, se prestará particular atención a los modelos cibernéticos pues han mostrado ser muy eficientes en la predicción del crecimiento microbiano en cultivos complejos que involucran a múltiples sustratos y en la predicción de la utilización secuencial o en paralelo de más de uno de ellos; el consumo de sustratos sustituibles; el crecimiento en nutrientes complementarios y en condiciones de limitación de la disponibilidad de fuentes de nitrógeno y/o carbono entre otros escenarios en los cuales los modelos no estructurados son incapaces de proporcionar una descripción adecuada.

El modelo cibernético; las células como estrategias óptimas

Los modelos no estructurados muestran incapacidad predictiva en cultivos con medios que involucran a múltiples sustratos. Esto se debe a la presencia de complejos mecanismos regulatorios que no pueden ser descriptos por medio de ecuaciones cinéticas simples. Los modelos cibernéticos (la palabra

«cibernético» proviene del griego Κυβερνήτης —*kybernetes*— que significa «arte de pilotear un navío») no proponen interrelaciones mecanísticas para describir la utilización de diferentes sustratos y el consumo de cada uno de ellos se asume como dependiente de una cinética propia e independiente. Esta simplificación se compensa asignando una motivación de supervivencia de la especie que está cuantificada por medio de una política de control óptimo de la actividad metabólica de los microorganismos. Esta hipótesis descansa en el supuesto de que la evolución ha seleccionado mecanismos intracelulares que permiten llevar a cabo dicha política. Los modelos cibernéticos son capaces de predecir con notable exactitud el consumo secuencial y/o paralelo de diferentes sustratos a partir de la extrapolación de resultados del crecimiento en un solo sustrato sin suposición alguna acerca de cinéticas interconectadas. La respuesta de los microorganismos a variaciones en la concentración de varios sustratos en el medio de cultivo se especifica para cada especie y se manifiesta como la biosíntesis de las enzimas necesarias para incorporar el sustrato al metabolismo celular. El modelo cibernético (Dhurjati *et al.*, 1985) asume que las células son capaces de asignar en forma óptima los recursos necesarios en cada una de las redes de reacciones involucradas en la síntesis de proteínas necesarias para procesar a cada sustrato. Para explicar la lógica del modelado cibernético, consideremos el crecimiento de un microorganismo en un medioambiente con n sustratos en los que la fuente de carbono actúa como sustrato limitante. Definimos al vector de sustratos presentes en el medio de cultivo como:

$$\mathbf{x}_s = [S_1, \dots, S_p, \dots, S_n] \quad (5.12)$$

siendo S_i la concentración del sustrato limitante i . Además, el vector de estados fisiológicos se define como:

$$\mathbf{x}_p = [X, P_1, \dots, P_p, \dots, P_n] \quad (5.13)$$

siendo P_i la actividad de la enzima necesaria para la metabolización del sustrato S_i y X la concentración de biomasa. El vector de estados del sistema será en consecuencia:

$$\mathbf{x} = [X, S_1, \dots, S_p, \dots, S_n, P_1, \dots, P_p, \dots, P_n] \quad (5.14)$$

Sea $R_i = dP_i/dt$ la velocidad de asignación de recursos críticos para la construcción de la i -enzima P_i . La asignación fraccional de recursos se define como $R_i / R_{max} = u_i$, lo que lleva a definir al vector de asignación de recursos como:

$$\mathbf{u} = [u_1, \dots, u_p, \dots, u_n] \quad (5.15)$$

Dichas asignaciones se consideradas como variables de control manipuladas en forma óptima por el microorganismo. A fin de normalizar las acciones de control, se puede considerar a R_{max} como constante. Entonces debe verificarse que:

$$\sum_{i=1} u_i = 1 \quad (5.16)$$

Análogamente, si la actividad máxima de la enzima P_i es $P_{i\ max}$, se puede definir el vector de actividades adimensionales como:

$$\mathbf{e} = [e_1, \dots, e_p, \dots, e_n] \quad (5.17)$$

donde $e_i = P_i / P_{i\ max}$, lo que conlleva a la definición del siguiente vector de estados con actividades enzimáticas adimensionales:

$$\mathbf{x} = [X, S_1, \dots, S_p, \dots, S_n, e_1, \dots, e_p, \dots, e_n] \quad (5.18)$$

La velocidad de síntesis de e_i se asume como dependiente de los recursos disponibles y se modela con la siguiente ecuación:

$$\frac{de_i}{dt} = a_i f\left(u_i \frac{dP_i}{dt}\right) - b_i e_i \quad (5.19)$$

Donde a_i es una constante, f explicita la dependencia de la velocidad de síntesis de la enzima e_i con respecto de la acción de control u_i y el último término representa la velocidad de inactivación de la enzima siendo b_i una constante. Una vez que la maquinaria celular sintetiza a la enzima clave, se produce crecimiento celular representado por la siguiente pseudoreacción química:



La expresión anterior es una representación macroscópica simplificada que no describe a la estequiometría real de crecimiento. La velocidad de consumo del sustrato i debida a la presencia de la enzima P_i puede ser modelada por una cinética del tipo Monod:

$$\frac{dS_i}{dt} = -V_{Si} e_i \frac{S_i}{S_i + k_{Si}} X \quad (5.21)$$

Donde V_{Si} y k_{Si} son las velocidades específicas de crecimiento y la constante de saturación en el sustrato S_i respectivamente, siendo $V_{Si} = \mu_{max} / Y_{X/S_i}$ donde μ_{max} es la velocidad de crecimiento máxima en dicho sustrato. La producción total de biomasa vendrá dada por:

$$\frac{dX}{dt} = \sum_{i=1}^n Y_{X/S_i} \frac{dS_i}{dt} \quad (5.22)$$

donde Y_{X/S_i} es el rendimiento de biomasa en el sustrato i y n el número de sustratos limitantes que hay en el cultivo. Como se dijo antes, la regulación metabólica se expresa por medio de un criterio de optimización que en este caso viene dado como la maximización de la producción de biomasa:

$$\text{Maximizar } \frac{\int_{t_0}^{t_f} \frac{dx}{dt} dt}{\int_{t_0}^{t_f} dt} = \frac{1}{t_f} \int_{t_0}^{t_f} \frac{dx}{dt} dt \quad (5.23)$$

Siendo t_f el período de fermentación definido por el siguiente criterio de parada:

$$\sum_{i=1}^n \left(S_{i0} + \int_{t_0}^{t_f} \left(-\frac{dS_i}{dt} dt \right) \right) = 0 \quad (5.24)$$

Dicho criterio implica que habrá crecimiento microbiano en tanto haya sustratos limitantes en el medio de cultivo y como consecuencia el crecimiento se detendrá cuando los sustratos han sido totalmente consumidos. La predicción de la evolución de los estados del sistema requerirá entonces de la solución de un problema de control óptimo.

Consideremos, por ejemplo, al crecimiento de *Klebsiella pneumoniae* en un reactor batch con un medio de cultivo en el que hay presente dos sustratos (d -glucosa y d -xilosa) con concentraciones iniciales de 1 g/l y 2 g/l respectivamente (Dhurjati *et al.*, 1985). Definimos al vector de estados será como:

$$\mathbf{x} = [X, S_1, S_2, P_1, P_2] \quad (5.25)$$

donde S_1 es la glucosa y S_2 es la xilosa. La concentración inicial de biomasa es $X(0) = 0.0175$ g/l. Los vectores de control metabólico y de actividades de las enzimas críticas serán:

$$\mathbf{u} = [u_1, u_2] \quad (5.26)$$

$$\mathbf{e} = [e_1, e_2] \quad (5.27)$$

donde $e_1 = P_1 / P_{1max}$ y $e_2 = P_2 / P_{2max}$. La evolución en el tiempo de los estados del sistema estará dada por la solución del siguiente problema:

$$\text{Max} \frac{1}{t_f} \int_{t_0}^{t_f} \frac{dX}{dt} dt \quad (5.28.1)$$

sujeto a

$$\frac{dX}{dt} = \mu_1 e_1 \frac{S_1}{k_{s1} + S_1} + \mu_2 e_2 \frac{S_2}{k_{s2} + S_2} \quad X(0) = 0.0175 \text{ g/l} \quad (5.28.2)$$

$$\frac{dS_1}{dt} = -\mu_1 e_1 \frac{1}{Y_{X/S1}} \frac{S_1}{k_{s1} + S_1} X \quad S_1(0) = 1 \text{ g/l} \quad (5.28.3)$$

$$\frac{dS_2}{dt} = -\mu_2 e_2 \frac{1}{Y_{X/S2}} \frac{S_2}{k_{s2} + S_2} X \quad S_2(0) = 2 \text{ g/l} \quad (5.28.4)$$

$$\frac{de_1}{dt} = a_1 u_1 - a_1 e_1 \quad e_1(0) = 1 \quad (5.28.4)$$

$$\frac{de_2}{dt} = a_2 u_2 - a_2 e_2 \quad e_2(0) = 0.002 \quad (5.28.5)$$

Donde además se consideran las siguientes restricciones:

$$u_1 + u_2 = 1 \quad (\text{Normalización del vector de acciones de control}) \quad (5.29)$$

$$V = 1,5 \text{ l} \quad (\text{Volumen de fermentación}) \quad (5.30)$$

Tabla 4.5. Parámetros del problema (Dhurjati *et al.*, 1985)

Parámetro	Valor	Parámetro	Valor
μ_1	1.07 h ⁻¹	a_1	10.0 h ⁻¹
μ_2	0.51 h ⁻¹	a_2	1.0 h ⁻¹
K_{s1}	0.01 g/l	$Y_{X/S1}$	0.525
K_{s2}	0.01 g/l	$Y_{X/S2}$	0.360

La variación en el tiempo la concentración de biomasa así como la política óptima de control enzimático se ilustran en la figura 4.2, donde la concentración de biomasa predicha por el modelo (---) se compara con la experimental (—) que fue determinada midiendo el consumo instantáneo de oxígeno, el cual se asume como perfectamente correlacionado con la concentración de biomasa al despreciar las acciones de mantenimiento celular. Dicha suposición es válida solo en la fase de crecimiento.

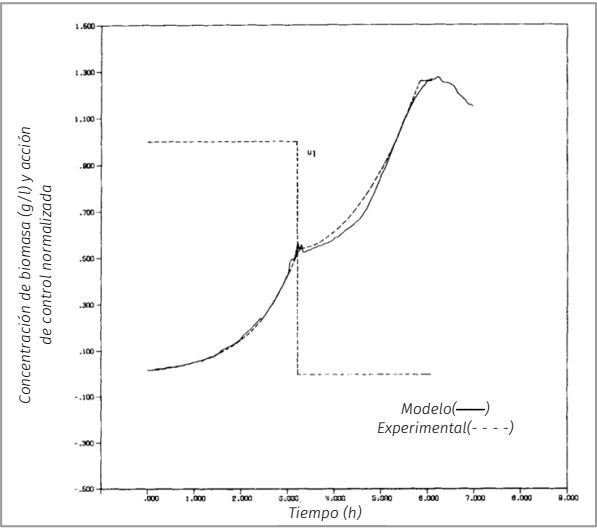


Figura 5.2.a. Estrategia óptima de control enzimático y comparación de las concentraciones celulares predichas y experimentales para el crecimiento de *Klebsiella pneumoniae* (Dhurjati et al., 1985).

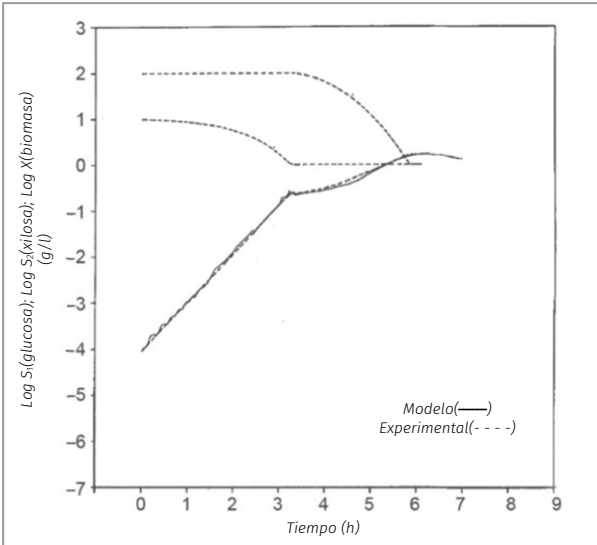


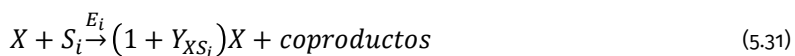
Figura 5.2.b. Evolución en el tiempo de las concentraciones de sustrato y biomasa (Dhurjati et al., 1985).

Como se observa en la primera figura, las células primero asignan totalmente sus recursos a la producción de la enzima involucrada en la utilización de glucosa. Luego, cuando esta se agota, se produce un «switch» y se redireccionan los recursos celulares a la producción de la enzima asociada al consumo de xilosa con el fin de poder consumir el sustrato menos preferido. En esos casos, dichas acciones pueden predecirse por medio de la expresión [$u_i = 1$ para $r_i = \max_j r_j$; $u_i = 0$ para $r_i \neq \max_j r_j$]; siendo r_j la velocidad de procesamiento del sustrato i . Se concluye que el modelo cibernético es capaz de describir a los fenómenos regulatorios que ocurren durante el crecimiento diauxico por medio de una política de asignación óptima de recursos para la producción de las enzimas que son necesarias para el procesamiento de cada sustrato.

Evaluación experimental de la capacidad predictiva de los modelos cibernéticos en cultivos con varios sustratos: sobre la base de los conceptos anteriormente introducidos, se presenta a continuación un modelo cibernético generalizado (Kompala *et al.*, 1986) que sintetiza las características del crecimiento microbiano en cultivos con múltiples sustratos. Dicho modelo se basan en los siguientes postulados:

1. Dado un medio con varios sustratos, los microorganismos prefieren utilizar primero aquel que permite el crecimiento más rápido.
2. En general, el crecimiento varía desde la utilización simultánea, cuando los crecimientos específicos a cada sustrato son muy similares, a la utilización secuencial con fases de latencia intermedias.
3. El crecimiento celular en una mezcla de sustratos nunca es mayor al crecimiento en el medio con el sustrato que implica la mayor velocidad de crecimiento.
4. Si hay crecimiento en un *sustrato lento*, la adición de un *sustrato rápido* inhibe la actividad de las enzimas que permiten consumir el sustrato lento para generar enzimas asociadas al sustrato rápido.

Consideremos que la asimilación del i -ésimo sustrato S_i está catalizada por la enzima E_i , definiéndose una *vía metabólica* representada por la siguiente estequiometría:



La producción de la enzima e_i está inducida por la presencia del sustrato S_i :



Donde X' es la biomasa excluyendo a la enzima e_i . Las velocidades de reacción para estas dos estequiometrías pueden ser las siguientes:

$$r_i = \mu_i e_i \frac{S_i}{k_{Si} + S_i} X \quad (5.33)$$

$$r_{Ei} = \alpha_i \frac{S_i}{k'_{Si} + S_i} X \quad (5.34)$$

En la ecuación (5.33), e_i es la actividad adimensional de la enzima asociada a la metabolización del sustrato i , α_i , en la ecuación (5.34) es la máxima velocidad de síntesis de la enzima e_i . Con el fin de especificar la influencia de los niveles de las enzimas específicas para cada sustrato, la constante máxima velocidad de crecimiento de la biomasa en el sustrato i , μ_i^{MAX} , debe ser reemplazada por el término $\mu_i e_i$. Cuando hay varios sustratos en el medio, los procesos regulatorios de inhibición/activación y de represión/inducción afectan a las ecuaciones (5.33) y (5.34) respectivamente. La velocidad de síntesis de la enzima e_i puede escribirse como:

$$r_{Ei} = r^{\text{max}} u_i \quad \left(0 < u_i < 1; \sum_i u_i = 1 \right) \quad (5.35)$$

donde u_i es la asignación fraccional de recursos intracelulares destinada a la síntesis de la enzima e_i . Si se incorpora el efecto de la dilución del nivel enzimático específico causado por el crecimiento celular y por el decaimiento en la actividad enzimática, la ecuación cinética para e_i se puede escribir como:

$$\frac{dX}{dt} = \alpha_i \frac{S_i}{k_{Si} + S_i} u_i + \frac{d}{dt} (\ln X) e_i - \beta_i e_i \quad (5.36)$$

siendo α_i y β_i las constantes específicas de síntesis y de decaimiento de la actividad enzimática, respectivamente. La variable u_i representa la acción de control de los mecanismos regulatorios intracelulares de represión e inducción. De la misma manera los mecanismos de inhibición y activación que controlan la actividad de las enzimas están representadas por las variables v_i . Consecuentemente, la velocidad de utilización del sustrato S_i será escrita como:

$$\frac{dS_i}{dt} = -\frac{1}{Y_{X/S_i}} r_i v_i = -\frac{1}{Y_{X/S_i}} \mu_i e_i v_i \frac{S_i}{k_{Si} + S_i} X \quad (0 < v_i < 1) \quad (5.37)$$

y la velocidad de crecimiento de biomasa total será:

$$\frac{dX}{dt} = \sum_i r_i v_i = \sum_i \left(\mu_i e_i v_i \frac{S_i}{k_{Si} + S_i} X \right) \quad (5.38)$$

La asignación de recursos metabólicos se considera como proporcional a la cantidad de beneficios, en términos de producción de biomasa, obtenidos en cada alternativa y la asignación fraccional óptima u_i a la i -ésima alternativa vendrá dada por:

$$u_i = \frac{r_i}{\sum_j r_j} \quad (5.39)$$

Adicionalmente, si se normaliza el valor de la variable cibernética v_i , se obtiene la siguiente expresión fraccional:

$$v_i = \frac{r_i}{\max_j r_j} \quad (5.40)$$

Las ecuaciones (5.35) a (5.39) constituyen el modelo cibernético para el crecimiento microbiano en cultivos con múltiples sustratos. Un tipo de crecimiento menos conocido se denomina crecimiento bifásico e involucra a la utilización secuencial de dos sustratos sin la fase de latencia que es común en procesos diaúxicos. Ese tipo de procesos es usual cuando se utilizan medios con sustratos que son isómeros cercanos y cuando la vía metabólica utilizada para consumir a ambos sustratos es la misma. Ejemplos de este tipo de crecimiento lo constituyen el crecimiento de *E.coli* en mezclas de glucosa y galactosa o sacarosa y manosa y de *K. oxytoca* en glucosa y fructosa (Kompala *et al.*, 1986). La ausencia de fase de latencia indica que las enzimas necesarias para la utilización del segundo sustrato ya están presentes en sus máximos niveles al final de la etapa de consumo del primero, aunque el crecimiento en el primer sustrato sea más rápido que en el segundo. El modelo desarrollado para el crecimiento diaúxico no puede simular el crecimiento bifásico debido a la presencia de una enzima común para ambos sustratos. Dado que estos sustratos, S_1 y S_2 , son isómeros que compiten por la misma enzima, hay que introducir la siguiente restricción adicional en la variable v_i .

$$\sum_i v_i = 1 \quad (5.41)$$

La estrategia óptima resultante que maximiza la velocidad de crecimiento microbiano será:

$$v_i \begin{cases} 1 & \text{si } r_i > r_j \\ 0 & \text{si } r_i < r_j \end{cases} \quad (5.42)$$

donde r_i es la velocidad de crecimiento instantánea en el sustrato S_i . Dado que la enzima clave es común a ambos sustratos, resulta que:

$$r_i = \mu_i \frac{S_i}{k_{si} + S_i} X \quad (5.43)$$

El modelo completo para el crecimiento bifásico vendrá dado por:

$$\frac{dS_i}{dt} = -\frac{1}{Y_{X/S_1}} r_i v_i \frac{S_i}{k_{si} + S_i} X \quad (5.44)$$

$$\frac{dX}{dt} = \sum_i r_i v_i = \sum_i \left(\mu_i v_i \frac{S_i}{k_{si} + S_i} X \right) \quad (5.45)$$

La figura 4.3 compara los datos experimentales con las simulaciones basadas en el modelo cibernético de crecimiento bifásico de *K. oxytoca* en glucosa (1 g/l) y fructosa (4 g/l).

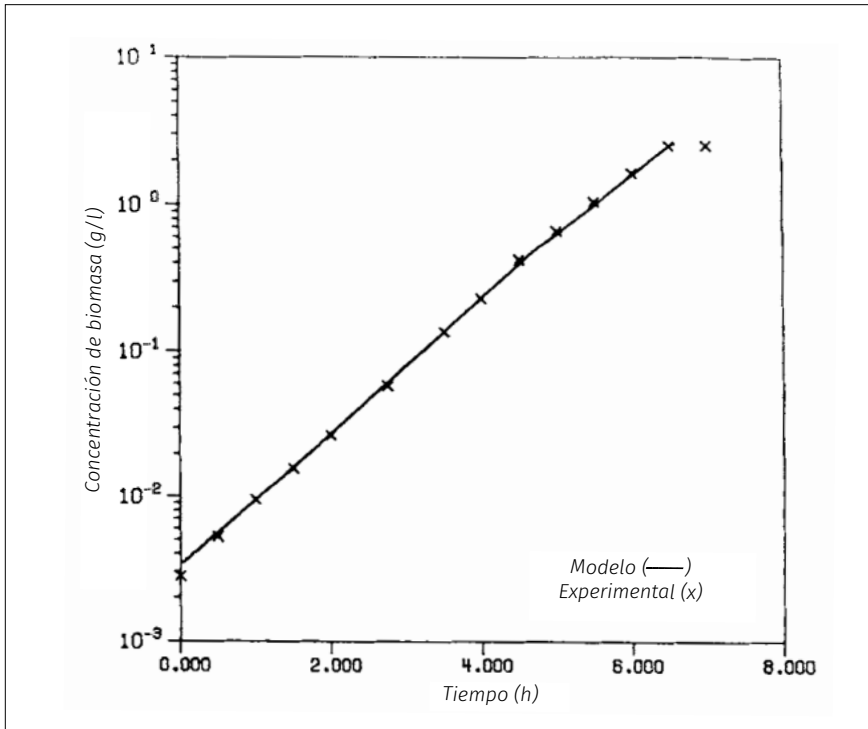


Figura 5.3. Utilización secuencial de glucosa (1 g/l) y fructosa sin latencia diauxica (Kompala et al., 1986).

INTEGRACIÓN DE ECUACIONES DE BALANCE CON MODELOS CINÉTICOS

Estequiometría

El primer paso de una descripción cuantitativa del crecimiento celular consiste en especificar aquellas reacciones que se considerarán en el desarrollo de un modelo. Para este propósito, es importante distinguir entre sustratos, productos metabólicos, metabolitos intracelulares y constituyentes de biomasa. Un *sustrato* es un compuesto presente en el medio, que puede metabolizarse o incorporarse directamente a la célula. Un *producto metabólico* es un compuesto producido por las células y excretado al medio extracelular. Los *constituyentes de la biomasa* son grupos de macromoléculas que forman la biomasa (por ejemplo, ARN, ADN, proteínas, lípidos y carbohidratos), pero también productos macromoleculares que se acumulan dentro de la célula (por ejemplo, un polisacárido o una proteína no secretada). Los *metabolitos intracelulares* son los compuestos restantes que están dentro de la célula (intermedios glucolíticos, metabolitos precursores y componentes básicos). Con el objetivo de especificar una estequiometría para las reacciones bioquímicas, consideramos un sistema en donde N sustratos se convierten en M productos metabólicos y en Q constituyentes de la biomasa. Las conversiones se llevan a cabo en J reacciones en las que K metabolitos intracelulares participan como intermedios en la ruta metabólica. Los sustratos se denominan S_i , los productos metabólicos se denominan P_i , los componentes de la biomasa se denominan $X_{macro,i}$, y los metabolitos intracelulares se denominan $X_{met,i}$. Con estas definiciones, la estequiometría general para la reacción j se puede especificar de la siguiente manera:

$$\sum_{i=1}^N \alpha_{ji} S_i + \sum_{i=1}^M \beta_{ji} P_i + \sum_{i=1}^Q \gamma_{ji} X_{macro,i} + \sum_{i=1}^N g_{ji} X_{met,i} = 0 \quad j = 1, \dots, J \quad (5.46)$$

Aquí, α_{ji} es un coeficiente estequiométrico para el i -ésimo sustrato, β_{ji} es un coeficiente estequiométrico para el i -ésimo producto metabólico, γ_{ji} es un coeficiente estequiométrico para el i -ésimo grupo macromolecular y g_{ji} es un coeficiente estequiométrico para el i -ésimo metabolito intracelular. Todos los compuestos consumidos en la reacción j tienen coeficientes estequiométricos negativos, mientras que todos los compuestos que se producen tienen coeficientes estequiométricos positivos. La estequiometría para todas las reacciones celulares J se puede escribir en forma compacta usando la notación matricial:

$$AS + BP + CX_{macro} + GX_{met} = 0 \quad (5.47)$$

donde **A**, **B**, **C** y **G** son matrices que contienen los coeficientes estequiométricos de las reacciones j para los sustratos, productos metabólicos, constituyentes de biomasa e intermedios de cada reacción, respectivamente. En estas matrices, las filas representan a las reacciones j y las columnas representan a los metabolitos. Es decir que el elemento de la fila j y la columna i -ésima de la matriz **A** especifica al coeficiente estequiométrico para el sustrato i -ésimo en la reacción j -ésima.

Coefficientes de rendimiento

El rendimiento (por ejemplo, cuánta glucosa termina convertida en un metabolito de interés) es un parámetro de diseño muy importante en muchos procesos fermentativos. Este rendimiento normalmente se representa en forma de coeficientes que pueden considerar el flujo relativo hacia el producto de interés con un determinado compuesto como referencia que a menudo es la fuente de carbono o la biomasa. Por lo tanto, los coeficientes de rendimiento tienen unidades de masa por unidad de masa del compuesto de referencia (por ejemplo, moles de penicilina formados por mol de glucosa consumida).

Un coeficiente de rendimiento utilizado frecuentemente en el diseño y operación de fermentaciones aeróbicas es el cociente respiratorio (RQ), que define la cantidad de dióxido de carbono formado por mol de oxígeno consumido. Se pueden encontrar en la literatura varias definiciones de los coeficientes de rendimiento (El-Mansi *et al.*, 2012). Aquí usaremos las siguientes definiciones:

$$Y_{I/J} = \frac{r_J}{r_I} \quad (5.48)$$

$$Y_{I/P} = \frac{r_P}{r_I} \quad (5.49)$$

$$Y_{I/X} = \frac{r_X}{r_i} \quad (5.50)$$

En la descripción clásica del crecimiento celular introducida por Monod en 1942, que considera que el coeficiente de rendimiento Y_{SX} era constante, y que todas las reacciones celulares se agrupan en una única reacción de crecimiento general donde el sustrato se convierte en biomasa. Sin embargo, a fines de la década de 1950 se demostró que el rendimiento de biomasa con respecto al sustrato que actúa como fuente de carbono y energía no es constante. Para describir esto, Herbert (Beefink *et al.*, 1990) introdujo el concepto de metabolismo endógeno y de consumo de sustrato específico para este proceso, además del de síntesis de biomasa. Al mismo tiempo,

Luedeking y Piret (1959) descubrieron que los lactobacilos producen ácido láctico en condiciones de población estacionaria, lo cual es consistente con un metabolismo endógeno de las células. Sus resultados indican una correlación lineal entre la velocidad de producción específica de ácido láctico y la velocidad de crecimiento específica más un término de producción constante:

$$r_p = a\mu + b \tag{5.51}$$

Pirt (1982) introdujo una correlación lineal similar entre la velocidad específica de consumo de sustrato y la velocidad específica de crecimiento y sugirió la existencia de un término constante de mantenimiento celular. Actualmente este es un concepto muy utilizado en el metabolismo endógeno. La correlación lineal de Pirt toma la siguiente forma:

$$r_s = Y_{x/s}^{real} \mu + m_s \tag{5.52}$$

donde el parámetro $Y_{x/s}^{real}$ se conoce como coeficiente de rendimiento real. Por otra parte, m_s es el coeficiente de mantenimiento. Con la introducción de esta correlación lineal, los coeficientes de rendimiento obviamente no pueden ser constantes. Por lo tanto, el rendimiento de biomasa en el sustrato será:

$$Y_{s/x} = \frac{\mu}{Y_{x/s}^{real} \mu + m_s} \tag{5.53}$$

lo que muestra que el rendimiento observado $Y_{s/x}$ disminuye con velocidad específica de crecimiento porque se usa una fracción creciente del sustrato para cumplir con los requisitos de mantenimiento celular. Cuando la velocidad de crecimiento específica es grande, el coeficiente de rendimiento se aproxima al recíproco de $Y_{x/s}$. La siguiente tabla proporciona una compilación de coeficientes de rendimiento y mantenimiento para varias especies microbianas:

Tabla 4.6. Rendimiento verdadero y coeficientes de mantenimiento para diferentes especies microbianas (Nielsen & Villadsen, 1994)

Organismo	$Y_{x/s}$ (g/g DW)	m_s (g/g DW)
<i>Aspergillus nidulans</i>	1.67	0.020
<i>Candida utilis</i>	2.00	0.031
<i>Escherichia coli</i>	2.27	0.057
<i>Klebsiella aerogenes</i>	2.27	0.063
<i>Penicillium chrysogenum</i>	2.17	0.021
<i>Sacharomyces cerevisiae</i>	1.85	0.015
<i>Aerobacter aerogenes</i>	1.79	0.089
<i>Balillus megatarium</i>	1.67	-
<i>Klebsiella aerogenes</i>	2.13	0.074

Las relaciones lineales derivadas son muy útiles para correlacionar datos de crecimiento, en particular en cultivos continuos. La validez general de estas correlaciones lineales se debe a que están fundamentadas en los principios de conservación de materia y energía. El papel del sustrato productor de energía es proporcionar ATP para impulsar las reacciones de biosíntesis y polimerización, así como los procesos de mantenimiento celular de acuerdo con la siguiente relación lineal:

$$r_{ATP} = Y_{x/ATP}\mu + m_{ATP} \quad (5.54)$$

Esta ecuación es análoga a la correlación lineal de Pirt. La cantidad de ATP producido debe ser igual a la suma de las cantidades consumidas para el crecimiento y la cantidad destinada al mantenimiento si el rendimiento de ATP en el sustrato productor de energía es constante (es decir, si r_{ATP} es proporcional a r_s). El equilibrio entre la producción y el consumo de ATP puede extenderse a otros cofactores (por ejemplo, NADH y NADPH), y consecuentemente es posible deducir ecuaciones lineales de velocidad para tres casos diferentes:

- Crecimiento anaeróbico donde el ATP se suministra por fosforilación a nivel de sustrato.
- Crecimiento aeróbico sin formación de metabolitos.
- Crecimiento aeróbico con formación de metabolitos.

Para el crecimiento aeróbico con formación de metabolitos, la velocidad de consumo del sustrato específico toma la siguiente forma:

$$r_s = Y_{x/s}^{real}\mu + m_s \quad (5.55)$$

Esta ecuación puede interpretarse como un modelo agregado con tres reacciones:

- Conversión de sustrato a biomasa con un coeficiente estequiométrico constante en el sustrato y una velocidad de reacción proporcional a la velocidad de crecimiento específica.
- Conversión de un sustrato a un producto metabólico con un coeficiente estequiométrico constante en el sustrato y una velocidad de reacción proporcional a la velocidad de formación del producto específico.
- Consumo de sustrato destinado a cumplir los requisitos de mantenimiento con velocidad m_s .

Las dos primeras reacciones están agregadas en el primer término del lado derecho de la ecuación (5.55). En consecuencia, la estequiometría para estas tres reacciones y sus velocidades de reacción se puede especificar, en forma matricial, de la siguiente manera:

$$-Y_{x/s}^{real} S + X = 0 \quad \mu \quad (5.56)$$

$$-Y_{p/s}^{real} S + P = 0 \quad r_p \quad (5.57)$$

$$-S = 0 \quad m_s \quad (5.58)$$

Por lo tanto, la velocidad de consumo de sustrato es la suma del consumo de sustrato para el crecimiento de biomasa, para la formación de metabolitos y para el mantenimiento celular. Es importante distinguir entre los coeficientes de rendimiento verdaderos, que son coeficientes estequiométricos, y los coeficientes de rendimiento observados, que pueden considerarse como coeficientes estequiométricos de una reacción agrupada. Esta distinción lleva a modelos de caja negra que representan a todos los procesos celulares en forma agregada:

$$-Y_{x/s}^{real} S + Y_{x/p} P + X = 0; \quad \mu \quad (5.59)$$

A pesar del supraíndice «real», los coeficientes de rendimiento verdaderos corresponden a un sistema celular dado, ya que simplemente representan coeficientes estequiométricos de reacciones agrupadas. Si las condiciones ambientales cambian se puede activar un conjunto diferente de rutas metabólicas, y esto puede resultar en un cambio en el patrón de consumo de carbono en cada uno de los tres procesos mencionados con anterioridad llevando a un cambio de coeficientes. Incluso el coeficiente Y_{xATP} fundamental no puede considerarse como verdaderamente constante.

Insertión de ecuaciones cinéticas en balances macroscópicos para el modelado de biorreactores ideales

La estequiometría de las reacciones individuales es la base para realizar cualquier análisis cuantitativo de las velocidades de las reacciones individuales. Normalmente, la velocidad de una reacción química se expresa como la velocidad «hacia adelante» v_i , que define que un compuesto que tiene un coeficiente estequiométrico β en la i -ésima reacción se forma con velocidad βv_i . Con frecuencia, el coeficiente estequiométrico para uno de los compuestos se establece arbitrariamente en 1, por lo que la velocidad de reacción será igual al consumo o la producción de este compuesto que se toma en

referencia en esta reacción particular. Por esta razón, la velocidad de reacción se especifica normalmente con la unidad de moles por unidad de tiempo (o g/h), o si la cantidad total de biomasa se toma como referencia (g DW/h). Para calcular la producción global o la velocidad de consumo global, tenemos que sumar las contribuciones de las diferentes reacciones, es decir, que la velocidad de consumo específica total del sustrato *i*-ésimo sea igual a la suma de los consumos del sustrato en todas las reacciones *j*:

$$r_{s,i} = - \sum_{j=1}^J \alpha_{ji} v_j \quad (5.60)$$

Los coeficientes estequiométricos para los sustratos son negativos, es decir que la velocidad de formación específica del sustrato *i*-ésimo en la reacción *enésima* dada por $\alpha_{ji}v_j$ es negativa, pero la velocidad de generación específica del sustrato normalmente se considera positiva y, por lo tanto, se introduce un signo menos. Para la velocidad de formación específica del *i*-ésimo producto metabólico, de manera similar se tendrá:

$$r_{p,i} = \sum_{j=1}^J \beta_{ji} v_j \quad (5.61)$$

Si un compuesto se consume o se forma en una sola reacción, se puede obtener una medición directa de esta velocidad de reacción. Para los constituyentes de biomasa y los metabolitos intracelulares, podemos definir expresiones similares para la velocidad de formación neta en todas las reacciones *j*:

$$r_{macro,i} = \sum_{j=1}^J \gamma_{ji} v_j \quad (5.62)$$

$$r_{met,i} = \sum_{j=1}^J g_{ji} v_j \quad (5.63)$$

Estas velocidades son velocidades de formación específicas netas, porque un compuesto puede formarse en una reacción y consumirse en otra, y las velocidades determinan los resultados netos de consumo y formación en todas las reacciones celulares *j*. Por lo tanto, si $r_{met,i}$ es positivo, hay una formación neta del metabolito intracelular *i* y si ese valor es negativo, hay un consumo neto de este metabolito. Para finalizar, si $r_{met,i}$ es cero, las velocidades de formación del metabolito *i* equilibran exactamente a las de su consumo. Si las velocidades de reacción para las reacciones celulares *j*

se compactan en el vector de velocidad v , las sumas se pueden formular en notación matricial como:

$$\mathbf{r}_s = -\mathbf{A}^T \mathbf{v} \quad (5.64)$$

$$\mathbf{r}_p = -\mathbf{B}^T \mathbf{v} \quad (5.65)$$

$$\mathbf{r}_{macro} = \mathbf{G}^T \mathbf{v} \quad (5.66)$$

$$\mathbf{r}_{met} = \mathbf{G}^T \mathbf{v} \quad (5.67)$$

Aquí $\mathbf{r}_s$ es un vector de velocidad que contiene a las velocidades de absorción específicas de los N sustratos, $\mathbf{r}_p$ es un vector que contiene a las velocidades de formación específicas de los productos metabólicos M , $\mathbf{r}_{macro}$ es un vector que contiene a las velocidades de formación específica neta de los Q constituyentes de la biomasa y $\mathbf{r}_{met}$ es un vector que contiene a las velocidades de formación específica neta de los K metabolitos intracelulares. La velocidad de crecimiento específica μ de la biomasa total se calcula como la suma de las velocidades de formación de las macromoléculas constituyentes:

$$\mu = \sum_{i=1}^Q r_{macro,i} = \mathbf{1}_Q^T \mathbf{r}_{macro} = \mathbf{1}_Q^T \mathbf{G}^T \mathbf{v} \quad (5.68)$$

Modelos de caja negra

En un modelo de caja negra de crecimiento celular, todas las reacciones celulares se agrupan en una sola reacción. En esta reacción agregada, los coeficientes estequiométricos son idénticos a los coeficientes de rendimiento y por lo tanto la reacción se puede presentar como:

$$X + \sum_{i=1}^M Y_{X/Pi} P_i - \sum_{i=1}^N Y_{X/Si} S_i = 0 \quad (5.69)$$

Debido a que el coeficiente estequiométrico para la biomasa es por convención igual a 1, la velocidad de reacción estará dada por la velocidad de crecimiento, que junto con los coeficientes de rendimiento determinan la estequiometría del sistema. Con frecuencia, los coeficientes de rendimiento no son constantes y, por lo tanto, un modelo de caja negra no se puede aplicar para correlacionar, por ejemplo, la velocidad de consumo de un sustrato específico con la velocidad de crecimiento. Sin embargo, este tipo de modelos es útil para la validación de datos experimentales porque puede formar la base para calcular la conservación de los elementos. Por lo tanto, en el modelo de caja negra existen parámetros $(M + N + 1)$, donde hay M coeficientes de rendimiento para los productos metabólicos, N coeficientes de rendimiento para los sustratos y la velocidad de crecimiento es μ . Debido a que la materia se conserva en la conversión de sustratos a productos metabólicos y biomasa, los $(M + N + 1)$ parámetros del modelo de caja negra no son completamente independientes y deben satisfacer varias restricciones. Por lo tanto, los elementos que fluyen hacia el sistema deben equilibrarse con los elementos que fluyen fuera del mismo. Por ejemplo, el carbono que ingresa al sistema a través de los sustratos debe balancearse con el carbono contenido en los productos metabólicos y la biomasa. Cada elemento considerado en el modelo de caja negra obviamente implica una restricción. Por lo tanto, un balance de carbono permite expresar la siguiente restricción:

$$1 + \sum_{i=1}^M f_{p,i} Y_{x/pi} - \sum_{i=1}^N f_{s,i} Y_{x/si} = 0 \quad (5.70)$$

donde $f_{s,i}$ y $f_{p,i}$ representan el contenido de carbono en el sustrato i -ésimo y el producto metabólico i -ésimo, respectivamente. En la ecuación anterior, la composición elemental de la biomasa se normaliza con respecto al carbono, es decir que está representada por la fórmula $C_1H_aO_bN_c$. La composición elemental de la biomasa depende de su contenido de macromoléculas y, por lo tanto, de las condiciones del cultivo y de la velocidad de crecimiento específica. Por ejemplo, el contenido de nitrógeno es mucho más bajo en condiciones de limitación de nitrógeno que en condiciones de limitación de carbono. Sin embargo, excepto en situaciones extremas, es razonable usar la fórmula de composición general $C_{1.0}H_{1.8}O_{0.5}N_{0.2}$ siempre que la composición de biomasa no se conozca exactamente. A menudo, la composición elemental de sustratos y productos metabólicos se normaliza con respecto a

su contenido de carbono. Por ejemplo, la glucosa se especifica como $C_6H_{12}O_6$. La ecuación anterior se escribe, por mol de C, de la siguiente manera:

$$1 + \sum_{i=1}^M Y_{x/pi} - \sum_{i=1}^N Y_{x/si} = 0 \quad (5.71)$$

En la ecuación anterior, los coeficientes de rendimiento tienen unidades de moles de C por mol de biomasa. Esta ecuación es muy útil para verificar la consistencia de los datos experimentales. Por lo tanto, si la suma del carbono presente en la biomasa y en los productos metabólicos no es igual a la suma del carbono presente en los sustratos, existe una inconsistencia en los datos experimentales. En forma análoga los balances de hidrógeno, oxígeno y nitrógeno serán:

$$a_x + \sum_{i=1}^M a_{p,i} Y_{x/pi} - \sum_{i=1}^N a_{s,i} Y_{x/si} = 0 \quad (5.72)$$

$$b_x + \sum_{i=1}^M b_{p,i} Y_{x/pi} - \sum_{i=1}^N b_{s,i} Y_{x/si} = 0 \quad (5.73)$$

$$c_x + \sum_{i=1}^M c_{p,i} Y_{x/pi} - \sum_{i=1}^N c_{s,i} Y_{x/si} = 0 \quad (5.74)$$

Tabla 4.7. Composición elemental para varios microorganismos (Roels, 1983)

Microorganismo	Composición elemental	Contenido de cenizas (% en peso)	Condiciones de crecimiento
Candida utilis	$\text{CH}_{1.83}\text{O}_{0.46}\text{N}_{0.19}$	7.0	Limitante: glucosa, $D = 0.05 \text{ h}^{-1}$
	$\text{CH}_{1.87}\text{O}_{0.56}\text{N}_{0.20}$	7.0	Limitante: glucosa, $D = 0.45 \text{ h}^{-1}$
	$\text{CH}_{1.83}\text{O}_{0.54}\text{N}_{0.10}$	7.0	Limitante: amoníaco, $D = 0.05 \text{ h}^{-1}$
	$\text{CH}_{1.87}\text{O}_{0.56}\text{N}_{0.20}$	7.0	Limitante: amoníaco, $D = 0.45 \text{ h}^{-1}$
Klebsiella aerogenes	$\text{CH}_{1.75}\text{O}_{0.43}\text{N}_{0.22}$	3.6	Limitante: glicerol, $D = 0.10 \text{ h}^{-1}$
	$\text{CH}_{1.73}\text{O}_{0.43}\text{N}_{0.24}$	3.6	Limitante: glicerol, $D = 0.10 \text{ h}^{-1}$
	$\text{CH}_{1.75}\text{O}_{0.47}\text{N}_{0.17}$	3.6	Limitante: amoníaco, $D = 0.10 \text{ h}^{-1}$
	$\text{CH}_{1.73}\text{O}_{0.43}\text{N}_{0.24}$	3.6	Limitante: amoníaco, $D = 0.80 \text{ h}^{-1}$
Sacharomyces cerevisae	$\text{CH}_{1.82}\text{O}_{0.58}\text{N}_{0.16}$	7.3	Limitante: glucosa, $D = 0.080 \text{ h}^{-1}$
	$\text{CH}_{1.78}\text{O}_{0.60}\text{N}_{0.19}$	9.7	Limitante: glucosa, $D = 0.0255 \text{ h}^{-1}$
	$\text{CH}_{1.94}\text{O}_{0.52}\text{N}_{0.25}$	5.5	No limitado
Escherichia coli	$\text{CH}_{1.77}\text{O}_{0.49}\text{N}_{0.24}$	5.5	No limitado
	$\text{CH}_{1.83}\text{O}_{0.50}\text{N}_{0.22}$	5.5	No limitado
	$\text{CH}_{1.96}\text{O}_{0.55}\text{N}_{0.25}$	5.5	No limitado
	$\text{CH}_{1.93}\text{O}_{0.55}\text{N}_{0.25}$	5.5	No limitado
Pseudomonas fluorescens	$\text{CH}_{1.83}\text{O}_{0.55}\text{N}_{0.26}$	5.5	No limitado
Aerobacter aerogenes	$\text{CH}_{1.64}\text{O}_{0.52}\text{N}_{0.16}$	7.9	No limitado
Penicillium chrysogenum	$\text{CH}_{1.70}\text{O}_{0.58}\text{N}_{0.15}$		
	$\text{CH}_{1.68}\text{O}_{0.53}\text{N}_{0.17}$		Limitante: glucosa, $D = 0.038 \text{ h}^{-1}$
Aspergillus niger	$\text{CH}_{1.72}\text{O}_{0.55}\text{N}_{0.17}$	7.5	Limitante: glucosa, $D = 0.098 \text{ h}^{-1}$
Promedio	$\text{CH}_{1.81}\text{O}_{0.52}\text{N}_{0.21}$	6.0	No limitado

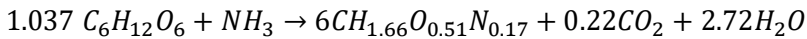
Los cuatro balances elementales se pueden escribir convenientemente conociendo la composición elemental de la biomasa, los sustratos y los productos metabólicos en las columnas de una matriz **E**, donde la primera columna contiene la composición elemental de la biomasa, las columnas en las posiciones 2 a $M + 1$ contiene la composición elemental de los productos metabólicos M , y las columnas de las posiciones $M + 2$ a $M + N + 1$ contienen la composición elemental de los N sustratos. Los cuatro equilibrios elementales se pueden expresar entonces como:

$$\mathbf{A} \mathbf{E} = \mathbf{0} \quad (5.75)$$

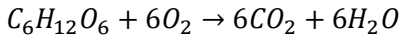
donde **A** es la matriz que contiene los coeficientes estequiométricos. Con las $(N + M + 1)$ variables, los coeficientes de rendimiento $N + M$ y la velocidad de reacción directa y cuatro restricciones, el grado de libertad es $F = M + N + 1 - 4$. Si se miden F variables, es posible calcular las otras velocidades utilizando las cuatro ecuaciones algebraicas dadas pero, en este caso, no queda redundancia para verificar la consistencia de los datos. Por esta razón, es aconsejable esforzarse por medir un número mayor de componentes que los grados de libertad del sistema.

Consideremos por ejemplo la producción de levadura de cerveza ($\text{CH}_{1.66}\text{O}_{0.51}\text{N}_{0.17}$), utilizando glucosa ($\text{C}_6\text{H}_{12}\text{O}_6$) como sustrato principal y amoníaco (NH_3) como fuente de nitrógeno. El producto principal es el etanol ($\text{C}_2\text{H}_5\text{OH}$). Dado que el vector de componentes principales involucra a 7 especies y que estas se caracterizan por fórmulas elementales que involucran a 4 elementos; con $(4 - 7) = 3$ «reacciones» independientes es suficiente para la caracterización completa del sistema. Las mismas, que a su vez se especifican mediante sus respectivas estequiometrías pueden ser, por ejemplo, las siguientes:

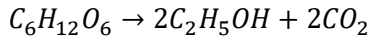
- Crecimiento celular (r_X):



- Respiración (r_C):



- Producción de etanol (r_P):



Entonces, la matriz de coeficientes estequiométricos es:

$$\Lambda = \begin{bmatrix} 6 & -1.037 & -1 & 0 & 0 & 0.22 & 2.72 \\ 0 & -1 & 0 & 0 & -6 & 6 & 6 \\ 0 & -1 & 0 & 2 & 0 & 2 & 0 \end{bmatrix}$$

Y la matriz de composiciones elementales es:

$$E = \begin{bmatrix} 1 & 1.66 & 0.51 & 0.17 \\ 6 & 12 & 6 & 0 \\ 0 & 3 & 0 & 1 \\ 2 & 6 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 1 & 0 \end{bmatrix}$$

Verificándose que:

$$\Lambda E = 0 \quad (5.76)$$

Multiplicando esta igualdad por el vector de reacciones independientes $\mathbf{r}$, se obtiene el vector de producción neta de componentes del sistema $\mathbf{r}_A$:

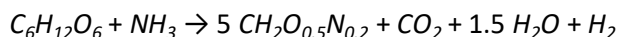
$$\mathbf{r} \mathbf{E} = \mathbf{0} \quad (5.77.1)$$

$$\mathbf{r}_A \mathbf{E} = \mathbf{0} \quad (5.77.2)$$

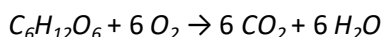
Actividades

1. Se desea obtener altas concentraciones de una levadura en un cultivo que opera en modo batch. Dicha fermentación se puede caracterizar en forma aproximada por medio de las siguientes pseudoreacciones químicas:

- Generación de biomasa

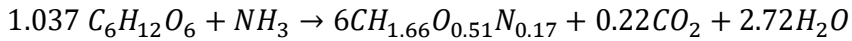


- Respiración

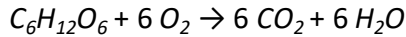


- a. Defina la matriz de estequiometrías de reacción $\mathbf{A}$.
- b. Defina la matriz de composiciones elementales $\mathbf{E}$.
- c. Verifique la conservación de elementos químicos.
- d. Asumiendo que se pueden medir los consumos de oxígeno, de glucosa y la producción de dióxido de carbono, obtenga la expresión que permite calcular la velocidad de aparición y desaparición de las restantes especies.
- e. ¿Se puede calcular la velocidad de aparición o desaparición de las restantes especies si solo se pueden medir la velocidad de consumo de oxígeno y la de producción de dióxido de carbono?
- f. Asumiendo que la cinética de crecimiento microbiano r_x se puede modelar por medio de la ecuación de Monod ($k_S = 20 \text{ g/l}$; $\mu_{max} = 0.4 \text{ h}^{-1}$), y la velocidad específica de respiración es de primer orden en la concentración de biomasa ($k_R = 0.04 \text{ g S/g X h}$), deduzca las cinéticas de consumo de glucosa y de consumo de oxígeno.
- g. Plantee los correspondientes modelos matemáticos si el proceso se lleva a cabo en: (i) un fermentador batch con $X(t=0) = 0.1 \text{ g/l}$; $S(t=0) = 25 \text{ g/l}$; $V = 100 \text{ lt}$. (ii) un reactor *fed batch* con $X(t=0) = 0.1 \text{ g/l}$; $S(t=0) = 25 \text{ g/l}$; $V(t=0) = 100 \text{ lt}$ y $F_I = 1 \text{ lt/h}$.

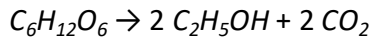
2. Considérese que un cultivo en el que se produce levadura de cerveza se puede representar por medio de las siguientes estequiometrías:
- Crecimiento celular (r_X):



- Respiración (r_C):



- Producción de etanol (r_P):



- Modele la cinética de crecimiento de biomasa por medio de una ecuación del tipo Monod considerando que tanto la fuente de carbono como la de nitrógeno son factores limitantes.
- Introduzca un término que modele la inhibición del crecimiento celular causada por altas concentraciones de etanol.
- Asumiendo que la producción de etanol puede modelarse por medio de la ecuación de Luedeking y Piret y la velocidad de respiración es proporcional a la concentración de biomasa, deduzca, utilizando las relaciones estequiométricas previamente presentadas, las velocidades de producción de CO_2 y de consumo de glucosa, de la sal de amonio y de oxígeno.
- Integre las ecuaciones cinéticas anteriores en ecuaciones de balance de materia para: (i) Obtener las ecuaciones de operación del modo batch. (ii) Obtener las ecuaciones de operación de un reactor en modo *fed batch* con una alimentación estéril que involucra a una solución glucosa, cuya concentración es S_F . (iii) Las ecuaciones de balance de un reactor que opera en modo *fed batch* con dos corrientes de alimentación estéril; una que involucra a una solución concentrada de glucosa (S_F) y otra con una solución de una sal de amonio (S_{NF}). (iv) Un reactor del tipo tanque agitado continuo con las dos corrientes de alimentación mencionadas en punto (iii) operando en estado estacionario. (v) El reactor continuo del punto (iv) operando en régimen transiente.

REFERENCIAS BIBLIOGRÁFICAS

- BEEFTINK, H.; VAN DER HEIJDEN, R. & HEIJNEN, J. (1990). Maintenance requirements: energy supply from simultaneous endogenous respiration and substrate consumption. *FEMS Microbiol Lett*, 73: 203–209. <https://doi.org/10.1111/j.1574-6968.1990.tb03942.x>
- DHURJATI, P.; RAMKRISHNA, D. (...) TSAO, G. T. (1985). A cybernetic view of microbial-growth: modeling of cells as optimal strategists. *Biotechnol Bioeng*, 27 (1): 1–9. <https://DOI:10.1002/bit.260270102>
- DONDO, R. (2012). *Modelado y Optimización Dinámica en Procesos Fermentativos*. Editorial Académica Española.
- EL-MANSI, E. M. T.; EL-MANSI, C. F. A. (...) ALLMAN, A. R. (EDS.) (2012). *Fermentation Microbiology and Biotechnology* CRC press. <https://doi.org/10.1201/b11490>
- INGRAHAM, J.; MAALØE, O. & NEIDHARDT, F. (1983). *Growth of the Bacterial Cells*. Sinauer Associate, Sunderland.
- KOMPALA, D.; RAMKRISHNA, D. (...) TSAO, G. (1986). Investigation of bacterial growth on mixed substrates: experimental evaluation of cybernetic models. *Biotechnology and Bioengineering*, 28, 1044–1055. <https://DOI:10.1002/bit.260280715>
- LUEDEKING, R. & PIRET, E. (1959). A kinetic study of the lactic acid fermentation. Batch process at controlled pH. *J Biochem Microbiol Technol Eng* 1: 393–412. [https://doi.org/10.1002/\(SICI\)1097-0290\(20000320\)67:6<636::AID-BIT3>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-0290(20000320)67:6<636::AID-BIT3>3.0.CO;2-U)
- NIELSEN, J. & VILLADSEN, J. (1994). *Bioreaction Engineering Principles*. Springer. <https://DOI:10.1007/978-1-4757-4645-7>
- PIRT, S. (1982). Maintenance energy: a general model for energy-limited and energy sufficient growth. *Arch Microbiol*, 133: 300–302.
- ROELS, J. (1983). *Energetics and Kinetics in Biotechnology*. Elsevier Biomedical Press.

Lecturas adicionales

- BAILEY, J. & OLLIS, D. (1986). *Biochemical Engineering Fundamentals*. (2nd ed.) McGrawhill.
- EL-MANSI, E.; BRYCE, C. (...) ALLMAN, A. (2012). *Fermentation Microbiology and Biotechnology*.
- RAMKRISHNA, D. & SONG, H. (2018). *Cybernetic Modeling for Bioreaction Engineering*. (1st ed.). Cambridge Series in Chemical Engineering.

6 Integración de modelos a escala macroscópica con modelos a escala microscópica

Los modelos macroscópicos suelen ser de dimensiones reducidas y con frecuencia se utilizan en algoritmos de control automático en línea. Los mismos pueden predecir razonablemente las velocidades de consumo de sustrato, producción de metabolitos y crecimiento de biomasa cuando se cumplen ciertas hipótesis como las de crecimiento equilibrado. Una desventaja del modelado macroscópico reside en que el número y la naturaleza de las reacciones macroscópicas necesarias para representar un bioproceso, sus coeficientes estequiométricos y los coeficientes de su cinética deben determinarse experimentalmente. El modelado macroscópico no tiene en cuenta los mecanismos endocelulares y difícilmente puedan utilizarse para el análisis y optimización de la cinética de metabolitos intracelulares. Por otro lado, el modelado metabólico se basa en el conocimiento de las reacciones bioquímicas que, catalizadas por enzimas, ocurren dentro de una célula. Estos modelos se basan en el conocimiento de las redes metabólicas, transcriptómicas y genómicas que proporcionan una comprensión mecanicista del comportamiento celular y son más apropiados para describir y optimizar bioprocesos que implican compuestos intracelulares. De hecho, el estudio *in silico* del metabolismo ayuda a comprender mejor ciertos mecanismos endocelulares para optimizar la producción de moléculas de interés. En este capítulo se presentan metodologías que permiten conciliar las escalas de modelado macroscópico y el modelado metabólico con el objetivo de predecir el comportamiento de variables macroscópicas así como flujos metabólicos y acumulación de metabolitos intracelulares de interés.

REDUCCIÓN DE LA DIMENSIÓN DE REDES METABÓLICAS DE ESCALA GENÓMICA

Las reconstrucciones de redes metabólicas a escala genómica se utilizan para construir modelos *in silico* del metabolismo celular. Para analizar estos modelos, se ha desarrollado una gran variedad de métodos basados en restricciones. Usualmente se asume que la red metabólica está en estado

estacionario, es decir que las velocidades de producción y consumo de metabolitos internos están equilibradas. Esto conduce a un espacio de flujo de la forma $C = \{v \in R^I \mid Sv = 0, LB \leq v \leq UB\}$. Aquí $S \in R^{I \times J}$ denota la matriz estequiométrica que contiene a un conjunto de metabolitos internos I y un conjunto de reacciones J . El vector v se denomina vectores de flujos factibles y representa a las distribuciones de flujo en estado estacionario de la red metabólica. Los vectores $LB, UB \in R^J$ definen los límites inferior y superior de los flujos, donde $R_{\pm\infty} = R \cup \{\pm\infty\}$. Por $Irrev \subseteq J$ denotamos al conjunto de reacciones irreversibles, que pueden transportar flujo en una sola dirección, es decir, $v^j \geq 0$, para todo $j \in Irrev$. Por simplicidad, asumimos $lb^j \geq 0$, para todo $j \in Irrev$. Las reacciones agrupadas en $Rev = Rxn \setminus Irrev$ se llaman reacciones reversibles. Las redes a escala genómica contienen miles de reacciones y metabolitos y en tales situaciones, a veces es deseable reducir la red dada a una significativamente más pequeña para obtener redes de tamaño práctico. Se han hecho propuestas para reducir grandes redes metabólicas a subredes más pequeñas que conserven las propiedades biológicas de interés. Estas propuestas involucran dos partes: la poda de red y la compresión de red (Röhl y Bockmayr, 2017). En el paso de compresión, las reacciones que pertenecen al mismo subconjunto de enzimas se agrupan y en el de poda se identifican las reacciones removibles y no removibles de tal manera que la red queda reducida a las reacciones no removibles que cumplen con los siguientes cuatro requisitos que debe especificar el usuario:

1. Conjunto de metabolitos protegidos P^I que deben ser retenidos en la red reducida.
2. Conjunto de reacciones protegidas P^J que deben ser retenidas en la red reducida.
3. Conjunto F de funcionalidades protegidas o *fenotipos* para la red reducida. Cada funcionalidad protegida $f \in F$ puede describirse mediante un sistema correspondiente de desigualdades lineales: $Df v \leq df$.
4. Grados de libertad mínimos; $dof \geq dof_{min}$. Aquí, los grados de libertad dof corresponden a la dimensión del espacio nulo de la matriz estequiométrica S , es decir, $dof = |J| - \text{rango}(S)$.

El objetivo es encontrar la subred más pequeña posible que cumpla con todos los requisitos descriptos. En la figura 6.1 se muestra un pequeño ejemplo ilustrativo.

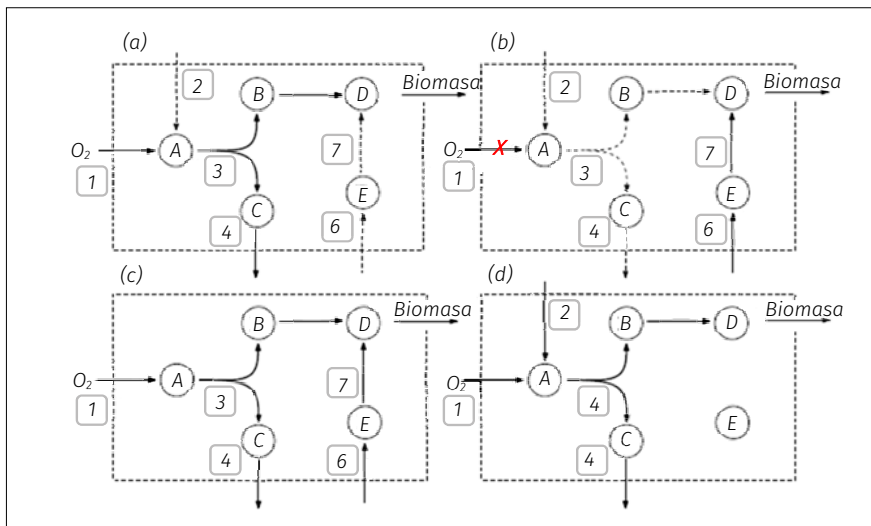


Figura 6.1. Los arcos sólidos corresponden a reacciones activas, los arcos punteados a reacciones inactivas. En (a) el vector de flujo satisface la funcionalidad de transportar el flujo a través de la reacción de la biomasa mientras absorbe oxígeno. En (b) la funcionalidad transporta flujo a través de la reacción de biomasa mientras no hay consumo de oxígeno. La combinación de los dos vectores de flujo conduce a la red en (c) que contiene 7 reacciones activas. En (d) se proporciona una subred mínima que permite ambas funcionalidades con solo 6 reacciones. Las variables binarias correspondientes para (d) tendrán los siguientes valores: $a_1 = 1$, $a_2 = 1$, $a_3 = 1$, $a_4 = 1$, $a_5 = 1$, $a_6 = 0$, $a_7 = 0$, $a_8 = 1$, donde a_i corresponde a la reacción r_i .

Conceptos básicos de desarrollo de un modelo MILP para calcular una subred mínima

Se asume que la red metabólica está en estado estacionario, es decir que $S \mathbf{v} = \mathbf{0}$, con límites en las velocidades de reacción $\mathbf{LB} \leq \mathbf{v} \leq \mathbf{UB}$. Cada funcionalidad $f \in F$ se describe mediante un sistema de desigualdades lineales: $Df \mathbf{v} \leq d\mathbf{f}$. Por ejemplo, podemos requerir que la velocidad de formación de biomasa tiene que ser al menos el 99 % de la velocidad teórica predicha por el análisis de balance de flujos. Se usarán las variables binarias $a^i \in \{0, 1\}$ para indicar si la reacción i tiene un flujo no nulo en la subred. Por lo tanto, necesitamos la relación $a^i = 0$ si y solo si $v^i = 0$. En consecuencia, para una reacción irreversible $i \in Irrev$, esto se puede lograr usando restricciones de la siguiente forma:

$$\delta a_i \leq v_i \leq M a_i \quad (6.1)$$

Para reacciones reversibles, usamos otra variable binaria $\bar{a}^i$ y las siguientes restricciones:

$$\delta a_i - M\bar{a}_i \leq v_i \leq M a_i - \delta \bar{a}_i; a_i + \bar{a}_i \leq 1 \quad (6.2)$$

Este tipo de restricciones se denomina restricción de M -grande (big- M), donde M es una constante suficientemente grande para imponer un límite superior de las velocidades de flujo. Con $\delta > 0$ denotamos un umbral que indica el valor de la velocidad de flujo de una reacción por sobre el cual se considera a la misma como inactiva. En la práctica se elige $\delta = 1e^{-06}$ a $1e^{-04}$. Para obligar a reacciones irreversibles protegidas a transportar flujo, se usan las restricciones $a^i = 1$ para todo $i \in P^{Irr} = P^I \cap Irrev$. Permitir un flujo a través de una reacción reversible protegida se puede realizar de manera similar con las restricciones $a^i + \bar{a}^i = 1$, para todo $i \in P^{Rev} = P^I \cap Rev$. Sea Rxn_m el conjunto de reacciones en las que interviene el compuesto m para cualquier metabolito protegido $m \in P^{Met}$. Por Rxn_m denotamos el conjunto de reacciones reversibles. Si Rxn_m contiene al menos una reacción protegida r , el metabolito m estará protegido por la reacción r . Sin embargo, si $Rxn_m \cap PRX_n = \emptyset$, además se necesitan las siguientes restricciones adicionales para proteger a m :

$$\sum_{i \in Rxn_m} a_i + \sum_{i \in Rev_m} \bar{a}_i \geq 1 \quad \forall m \in p_0^{Met} \quad (6.3.1)$$

donde

$$p_0^{Met} = \left\{ m \in \frac{P^{Met}}{Rxn_m} \cap P^{Rxn} = \emptyset \right\} \quad (6.3.2)$$

Un requisito adicional consiste en especificar un mínimo número de reacciones activas. Por el momento no incluiremos a esta restricción puesto que si existen reacciones que se ven obligadas a transportar flujo, esta restricción es redundante. Para encontrar una subred que contenga el número mínimo de reacciones activas, se minimiza la suma de las variables binarias a^i que indican que una reacción lleva flujo no nulo. El modelo MILP resultante es el siguiente:

$$\text{Minimizar} \quad (6.4.1)$$

$$\sum_{i \in Rxn} a_i + \sum_{k \in Rev} \bar{a}_k$$

sujeto a:

$$Sv = 0 \quad (6.4.2)$$

$$LB \leq v \leq UB \quad (6.4.3)$$

$$D_f v \leq d_f \quad \forall f \in F \quad (6.4.4)$$

$$\delta a_i \leq v_i \leq M a_i \quad \forall i \in Irrev \quad (6.4.5)$$

$$\delta a_i - M \bar{a}_i \leq v_i \leq M a_i - d \bar{a}_i \quad \forall i \in Rev \quad (6.4.6)$$

$$a_i + \bar{a}_i \leq 1 \quad i \in Rev \quad (6.4.7)$$

$$a_i = 1 \quad i \in P^{Rev} \quad (6.4.8)$$

$$a_i + \bar{a}_i = 1 \quad i \in P^{Rev} \quad (6.4.9)$$

$$\sum_{i \in Rxn_m} a_i + \sum_{i \in Rev_m} \bar{a}_i \geq 1 \quad \forall m \in p_0^{Met} \quad (6.4.10)$$

$$v_i \in R \quad i \in Rxn; a_i \in \{0,1\} \quad i \in Rxn; \bar{a}_i \in \{0,1\} \quad i \in Rev$$

Funcionalidades en conflicto

Las subredes resultantes debe mantener todas las funcionalidades deseadas. Por ejemplo, debe mantenerse una velocidad de crecimiento que sea al menos el 99 % de las máximas velocidades de crecimiento tanto en condiciones aeróbicas como anaeróbicas. Estos requisitos no se pueden realizar con el mismo vector de flujo $\mathbf{v}$ porque implican dos estados opuestos de la reacción r_{O_2} que transporta O_2 a la red. Se necesitaría un vector $\mathbf{v}$ con $v_{O_2} \geq \delta$ y $v_{O_2} = 0$ al mismo tiempo, lo cual es imposible porque el modelo anterior calcula solo un vector de flujo factible $\mathbf{v}$ para la red. Para obtener una subred que cumpla con las dos funcionalidades, necesitamos a un vector de flujo que cumpla la condición aeróbica y otro para la condición anaeróbica. Para realizar esto con un solo problema de tipo MILP, debemos modificar el anterior modelo. Primero, debemos buscar un vector de flujo v_o que contenga

los metabolitos protegidos y las reacciones protegidas. Además, para cada funcionalidad $j \in F$ buscamos un vector de flujo v_j que satisfaga $D_j v_j \leq d_j$ con sus correspondientes variables binarias. Por lo tanto, tenemos que usar variables binarias a_i^j para v_i^j con las cuales las ecuaciones (6.1) y (6.2) se convierten en:

$$\delta a_i^j \leq v_i \leq M a_i^j \quad i \in Irrev, i \in Func \quad (6.5)$$

$$\delta a_i^j - M \overline{a_i^j} \leq v_i \leq M a_i^j - \overline{\delta a_i^j} \quad i \in Rev, i \in Func \quad (6.6)$$

$$a_i^j + \overline{a_i^j} \leq 1 \quad i \in Rev, i \in Func \quad (6.7)$$

Usando las nuevas variables a_i^0 , reformulamos las restricciones respecto a las reacciones protegidas: $a_i^0 = 1$, para todo $i \in Plrr$ y $a_i^0 + \overline{a_i^0} = 1$, para todo $i \in P^{Rev}$. Finalmente, las restricciones con respecto a los metabolitos protegidos serán las siguientes:

$$\sum_{i \in Rxn_m} a_i^0 + \sum_{i \in Rev_m} \overline{a_i^j} \geq 1 \quad \forall m \in p_0^{Met} \quad (6.8)$$

Para obtener una subred mínima, tenemos que minimizar el número total de reacciones activas y por lo tanto, necesitamos variables binarias a_i con la propiedad de que $a_i = 0$ si y solo si $i^j = 0$ para todo $j \in F$ o $a_i = 1$ si y solo si $a_i^j = 0$ para algún $j \in F$. Para reacciones irreversibles, esto puede ser codificado mediante las siguientes restricciones:

$$a_i \leq \sum_{j=0}^F a_i^j \leq (1 + |F|) a_i \quad \forall i \in Irrev \quad (6.9)$$

$$a_i \leq \sum_{j=0}^F (a_i^j + \overline{a_i^j}) \leq (2 + s |F|) a_i \quad \forall i \in Rev \quad (6.10)$$

El modelo resultante será el siguiente:

$$\text{Minimizar} \quad \sum_{i \in R \times n} a_i \quad (6.11.1)$$

sujeto a:

$$Sv^j = 0 \quad \forall j \in F \quad (6.11.2)$$

$$lb \leq v^j \leq ub \quad \forall j \in F \quad (6.11.3)$$

$$\delta a_i^j \leq v_i \leq Ma_i^j \quad i \in Irrev, i \in Func \quad (6.11.4)$$

$$\delta a_i^j - \overline{Ma_i^j} \leq v_i \leq Ma_i^j - \overline{\delta a_i^j} \quad i \in Rev, i \in Func \quad (6.11.5)$$

$$a_i^j + \overline{a_i^j} \leq 1 \quad i \in Rev, i \in Func \quad (6.11.6)$$

$$a_i^0; a_i^0 + \overline{a_i^0} \leq 1 \quad i \in Rev, i \in Func \quad (6.11.7)$$

$$\sum_{i \in R \times n_m} a_i^0 + \sum_{i \in Rev_m} \overline{a_i^j} \geq 1 \quad \forall m \in p_0^{Met} \quad (6.11.8)$$

$$a_i \leq \sum_{j=0}^F a_i^j \leq (1 + |F|)a_i \quad \forall i \in Irrev \quad (6.11.9)$$

$$a_i \leq \sum_{j=0}^F (a_i^j + \overline{a_i^j}) \leq (2 + s|F|)a_i \quad \forall i \in Rev \quad (6.11.10)$$

$$v_i \in R \quad i \in R \times n; a_i^j, a_i \in \{0,1\} \quad i \in R \times n; \overline{a_i^j} \in \{0,1\} \quad i \in Rev$$

La resolución de este modelo MILP computa una subred con un número mínimo de reacciones activas que cumple todos los requisitos de funcionalidades especificadas.

Cálculo de todas las subredes mínimas

En ciertos casos debe computarse más de una subred. Por ejemplo, consideremos el caso en que el grado de libertad mínimo es mayor a 1. Si la subred calculada no tiene el grado de libertad requerido, tenemos que calcular una subred diferente. Además, la subred mínima calculada puede no ser única ya que pueden existir diferentes subredes que cumplan con los requisitos especificados por las restricciones y tengan el mismo número de reacciones

activas (Röhl y Bockmayr, 2017). Por lo tanto, es posible que nos interese computar todas las subredes mínimas. Cada vez que se encuentra una subred mínima, formulamos una restricción que excluye esta subred como solución factible y resolvemos el problema nuevamente con la restricción incluida. Para ello formulamos las siguientes restricciones:

$$\sum_{i \in R \times n} (1 - Z_i^k) a_i + \sum_{i \in R \times n} (1 - a_i) Z_i^k \quad k = 1, 2, \dots \quad (6.12)$$

donde $Z_i^k = 1$ si la reacción i transporta flujo en la subred que se calculó en el k -ésimo paso; de lo contrario, $Z_i^k = 0$. Por lo tanto, (6.12) garantiza que al menos una reacción inactiva se volverá activa, o que al menos una reacción activa quedará inactivada en la nueva solución. Resolviendo la formulación iterativamente y agregando las restricciones dadas por la ecuación (6.12) en cada paso, se pueden enumerar todas las subredes de mínima dimensión.

COMBINACIÓN DE MODOS DE FLUJOS ELEMENTALES Y MODELOS CIBERNÉTICOS

El análisis de balance de flujo (Orth *et al.*, 2010) en combinación con la descomposición de redes metabólicas en modos elementales han proporcionado una forma conveniente de modelar el metabolismo celular. Sin embargo, para obtener una estimación de los flujos intracelulares, FBA depende de la disponibilidad de mediciones de flujos externos, tales como los de consumo de sustrato o la velocidad de crecimiento. Este marco no permite el modelado de la regulación metabólica o la formulación de modelos dinámicos, excepto a través de la medición dinámica de flujos de intercambio. El modelado híbrido (Kim *et al.*, 2008) se fundamenta en la descomposición en modos elementales de la red metabólica. Las velocidades de consumo en los diversos modos elementales se combinan utilizando variables cibernéticas globales basadas en la maximización de las respectivas velocidades de consumos de sustratos. La estimación de los flujos de los metabolitos intracelulares se basa en información estequiométrica conjuntamente con mediciones de flujos externos bajo el supuesto de estado estacionario. Por lo tanto, el análisis de balance de flujo, el análisis de variación de flujo y otros enfoques basados en restricciones estequiométricas, tal como los modos elementales de flujo, que emplean la suposición de estado estacionario, se focalizan en predecir flujos de interés como por ejemplo, velocidades de excreción de metabolitos, velocidad de crecimiento microbiano, velocidad de intercambio

de cofactores, etc. Los enfoques dinámicos, por otro lado, requieren mediciones dinámicas de variables extracelulares e intracelulares para la identificación de parámetros cinéticos. El nivel de detalle de un modelo estará dado por el número de mediciones experimentales. La simplificación más común en el modelado del metabolismo proviene de los enfoques de balance de flujo, cuyo objetivo consiste en estimar, en base a medidas de flujos observables, otros flujos que no lo son. Por lo tanto, la suposición más común en el modelado metabólico es que las concentraciones de todos los metabolitos intracelulares están en estado estacionario, lo cual permite representar el metabolismo como una ecuación algebraica de la siguiente forma:

$$\mathbf{S}\mathbf{v} = \mathbf{b} \quad (6.13)$$

donde $\mathbf{S}$ es la matriz estequiométrica, $\mathbf{v}$ es el vector de flujo para todas las reacciones consideradas en el metabolismo y $\mathbf{b}$ representa un vector de flujos de intercambio. Por convención, el último elemento del vector de flujo $\mathbf{v}$ es el flujo de formación de biomasa que representa la contribución de los precursores internos para la síntesis de 1 g de biomasa. Cada fila de $\mathbf{S}$ está asociada con el balance de un metabolito intracelular, mientras que cada columna de $\mathbf{S}$ representa coeficientes estequiométricos de especies que participan en una reacción específica con signo positivo para productos, signo negativo para reactivos y 0 para especies que no participan en la reacción. La ecuación (6.13) se puede utilizar para calcular los flujos desconocidos a partir de las mediciones disponibles. Dado que es habitual medir las concentraciones de los componentes extracelulares para estimar los flujos externos, la ley de conservación de materia en los constituyentes internos de un microorganismo se emplea utilizando la suposición de estado estacionario. La representación matemática de estado estacionario es:

$$\mathbf{S}_m\mathbf{v} = \mathbf{0} \quad (6.14)$$

donde $\mathbf{S}_m$ es la matriz estequiométrica asociada solo con las especies intracelulares. Con la ecuación (6.23) se asume que la velocidad de dilución de los metabolitos debida al crecimiento puede despreciarse en comparación con otros términos en el balance de materia. En la ecuación (6.14), si el número de mediciones de flujo supera el rango de $\mathbf{S}_m$, todos los flujos se estiman mediante criterios de minimización de residuos tal como el método de los mínimos cuadrados. Por otro lado, si $\mathbf{S}$ está subdeterminado, mediante el análisis de balance de flujo se pueden estimar los flujos desconocidos invocando principios de optimización tales como la maximización del crecimiento de biomasa. Sin embargo, no se puede determinar la concentración

de variables intracelulares y los efectos reguladores metabólicos están contenidos en las mediciones de cierto número de concentraciones. Por lo tanto, el modelado del metabolismo celular debe tener en cuenta la regulación metabólica, que ocurre dentro de un marco dinámico. En el desarrollo de un modelo dinámico, primero se distingue entre el vector de flujo $\mathbf{v}$ y un vector de velocidad de reacción $\mathbf{r}$ expresado por célula unitaria mediante cinética apropiada. Denotamos el vector de variables de estado extracelulares mediante la siguiente expresión:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_s \\ \mathbf{x}_p \\ X \end{bmatrix} \quad (6.15)$$

donde $\mathbf{x}_s$ es un vector de n_s sustratos, $\mathbf{x}_p$ el vector de n_p productos de fermentación y X es un escalar que refiere a la biomasa, todos expresados en concentraciones por unidad de volumen del cultivo. Las variables intracelulares están representadas por un vector que comprende n_m componentes, expresados en cantidad por célula. A continuación, reemplazamos el lado derecho de la ecuación (6.13) por derivadas en el tiempo de todas las especies que reaccionan, incluidas las variables extracelulares e intracelulares con lo que dicha ecuación queda expresada de la siguiente manera:

$$S\mathbf{v} = \begin{bmatrix} \frac{1}{X} \frac{dx}{dt} \\ \frac{dm}{dt} \end{bmatrix} \quad (6.16)$$

Donde m representa al vector de metabolitos internos. Se requiere la división por la concentración de biomasa X del vector $\mathbf{x}$ para representar a la concentración en la célula. La ecuación (6.16) representa un modelo dinámico general en el que los efectos de los procesos regulatorios deben reflejarse en el vector de velocidades $\mathbf{v}$. Para el desarrollo de un modelo híbrido-cibernético se utiliza la hipótesis del estado pseudoestacionario en los metabolitos internos:

$$\frac{dm}{dt} = 0 \quad (6.17)$$

A continuación, se introduce la matriz $\mathbf{P}_m$ ($n_x \times n_r$) que representa la proyección del vector de flujo $\mathbf{v}$ sobre el vector $\mathbf{r}_x$ de flujos de intercambio que son medibles experimentalmente, es decir, $\mathbf{P}_m \mathbf{r} = \mathbf{r}_x$. Las ecuaciones (6.15) y (6.16) conducen a:

$$\mathbf{P}_m \mathbf{S} \mathbf{v} \equiv \mathbf{S}_m \mathbf{v} = \frac{d\mathbf{m}}{dt} = 0 \quad (6.18)$$

Y el modelo dinámico vendrá dado por:

$$\frac{1}{X} \frac{dx}{dt} = \mathbf{P}_m \mathbf{S} \mathbf{v} \quad (6.19)$$

donde $\mathbf{v}$ representa los flujos regulados requieren ser identificados completamente en el modelo híbrido. El marco cibernético aborda el modelado metabólico asumiendo que la regulación metabólica es la consecuencia del intento de las células de hacer un uso frugal de sus recursos internos para efectuar la síntesis y controlar las actividades de las enzimas que catalizan a las reacciones químicas con el objetivo de lograr ciertas metas de supervivencia. Por lo tanto, es de interés concebir modelos del metabolismo microbiano de orden reducido que puedan ser útiles computacionalmente. Con este fin, se contempla una síntesis de métodos de análisis de flujos basados en restricciones y bajo el supuesto de estado estacionario con flujos regulados por variables cibernéticas basadas en un objetivo global, formando así un modelo cibernético-híbrido. La hipótesis de estado estacionario para los componentes intracelulares permite el cálculo de los flujos internos en relación con las velocidades de intercambio mediante la explotación de información estequiométrica. Se pueden acoplar los diversos modos de flujo elemental que se incluirán en el metabolismo según lo determinen los pesos relativos de las variables cibernéticas globales. Un modo de flujo elemental muestra las rutas bioquímicas desde los sustratos hasta los productos metabólicos. La ecuación (6.18) representa a un sistema de ecuaciones diferenciales en todas las variables externas del sistema con todas las velocidades de reacción reguladas que se muestran en el lado derecho de la igualdad. Para el modelo híbrido, usamos la descomposición de los modos elementales para expresar el vector de velocidad de reacción de la siguiente manera:

$$\mathbf{Z}^T \mathbf{r}_M = \mathbf{v} \quad (6.20)$$

donde $\mathbf{Z}$ es una matriz que contiene n_z filas, cada una de las cuales contiene a un vector que representa un modo elemental que expresa las reacciones en que vinculan las velocidades de consumo de sustrato con la de producción de metabolitos a través de ese modo. El modelo híbrido incorpora mecanismos de regulación centrándose únicamente en la regulación de las velocidades

de consumo de sustratos a través de los modos elementales. Esto se logra mediante el control de la síntesis y la actividad de una enzima por cada modo elemental utilizando las leyes de correspondencia y proporcionales del modelado cibernético.

$$\text{si } Z_{i,1} \neq 0 \rightarrow u_i = \frac{Z_{i,1}r_{Mi}}{\sum_{j=1}^{n_z} Z_{j,1}r_{Mi}}, \text{ si } Z_{i,1} \neq 0 \rightarrow u_i = 1 \quad (6.21)$$

$$\text{si } Z_{i,1} \neq 0 \rightarrow v_i = \frac{Z_{i,1}r_{Mi}}{\max_j Z_{j,1}r_{Mi}}, \text{ si } Z_{i,1} \neq 0 \rightarrow v_i = 1 \quad (6.22)$$

$$i = 1, 2, \dots, n_z$$

Las variables cibernéticas u_i y v_i están formuladas para maximizar la velocidad de captación de global del sustrato limitante a través de cada modo. Los vectores $\mathbf{u}$ y $\mathbf{v}$ representan a todas las variables cibernéticas.

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \dots \\ u_{n_z} \end{bmatrix} \quad (6.23)$$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \dots \\ v_{n_z} \end{bmatrix} \quad (6.24)$$

El sistema de ecuaciones diferenciales para el vector de estados $\mathbf{x}$ se representa de la siguiente manera:

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}_x \mathbf{Z}^T \mathbf{V} \mathbf{r}_M \mathbf{x} \quad (6.25)$$

donde $\mathbf{V}$ es la matriz diagonal que contiene los componentes de $\mathbf{v}$ a lo largo de su diagonal. El vector de velocidad de captación $\mathbf{r}_M$ presenta n_z enzimas cuyos niveles se incluyen en el vector $\mathbf{e}_M$ que se convierten en variables de estado adicionales. La ecuación diferencial para las actividades de estas enzimas se puede escribir como:

$$\frac{de_{Mi}}{dt} = \alpha + r_{Ei}u_i - (\beta + r_G)e_{Mi} \quad \forall i \in \{j/Z_{j,1} \neq 0\} \quad (6.26)$$

donde α es la constante específica de síntesis de enzimas, β es la velocidad de degradación de la enzima y r_G es la suma de las velocidades de crecimiento en todos los modos elementales.

$$r_G = \sum_{i=1}^{n_z} Z_{i,nf} v_i r_{Mi} \quad (6.27)$$

En la ecuación (6.27), los coeficientes $Z_{i,nf}$ también son una representación del crecimiento de la biomasa en términos de precursores del crecimiento.

El modelo híbrido ahora está definido para todas las reacciones excepto las expresiones cinéticas de las velocidades de consumo de sustrato y las velocidades de síntesis de enzimas. Evidentemente el tamaño del modelo depende del número de modos elementales lo que hace que sea computacionalmente muy costoso.

Formulaciones cinéticas para el modelo híbrido

La cinética cibernética utiliza a la clásica ecuación de Michaelis–Menten que refiere a la catálisis mediante una enzima e con términos regulatorios adicionales implementados mediante las variables cibernéticas u y v . Dicha cinética para una reacción macroscópica i con n sustratos $S_1, \dots, S_n$ se puede escribir como:

$$\alpha_i = v_i e_i \mu_{max,i} \frac{S_1}{K_{S1} + S_1} \cdots \frac{S_n}{K_{Sn} + S_n} \quad (6.28)$$

siendo v_i la variable cibernética que controla la actividad de la enzima e_i , $\mu_{max,i}$ la velocidad máxima de absorción del sustrato i , y $K_s, S_1, \dots, K_{s,n}$ las constantes de semisaturación para cada sustrato $S_1, \dots, S_n$. La dinámica de las enzimas e_i se describe mediante la siguiente ecuación:

$$\frac{de_i}{dt} = A_i + u_i \mu_{max,i} \frac{e_i}{K_{Sei} + e_i} - D_i e_i - \mu e_i \quad (6.29)$$

siendo A_i la velocidad basal de síntesis enzimática, u_i la variable cibernética que controla la actividad de síntesis enzimática, $\mu_{max,i}$ la velocidad máxima de síntesis enzimática sin regulación, K_{s,e_i} la constante de semisaturación, D_i la tasa de degradación mientras que (μe_i) representa la dilución por crecimiento celular. Las variables cibernéticas u_i y v_i describen a los mecanismos reguladores que modulan la síntesis de enzimas y sus actividades, respectivamente. La forma general para u_i y v_i es la siguiente:

$$u_i = \frac{p_i^+}{\sum_i |p_i^+|} \quad (6.30)$$

$$v_i = \frac{p_i^+}{\max_i(p_i^+)} \quad (6.31)$$

donde

$$p_i^+ = \max(p_i, 0) \quad (6.32)$$

siendo p_i el *retorno de la inversión* asociado con la i -ésima reacción macroscópica. La determinación de p_i^+ requiere de la especificación de una función

objetivo. La más utilizada es la maximización de la velocidad de crecimiento celular por medio de la cual se busca asegurar la supervivencia de la especie. La velocidad de consumo de sustrato también se ha utilizado como una función objetivo aunque es menos usual.

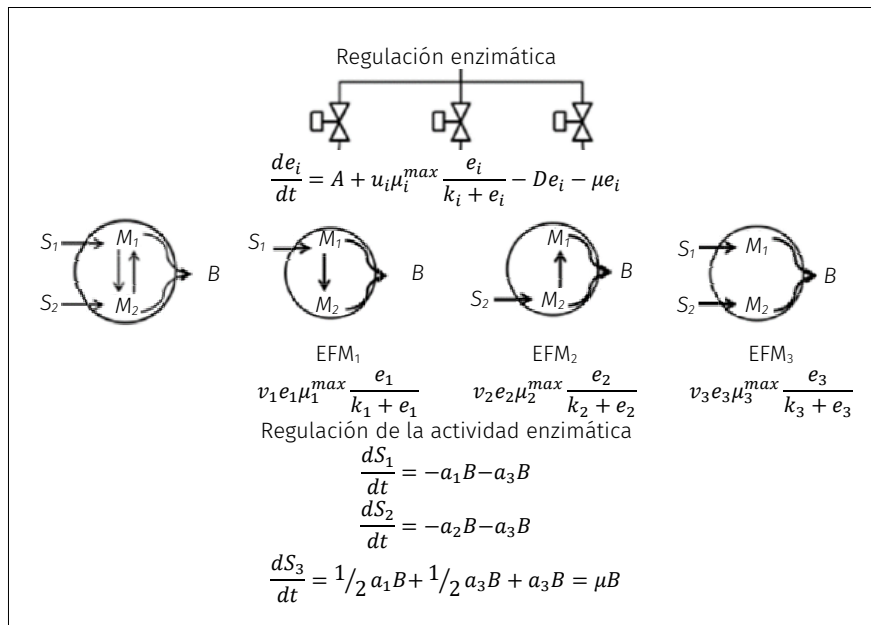


Figura 6.2. Representación esquemática del concepto de modelado cibernético híbrido de Kim *et al.* (2008).

El inconveniente de esta representación es la necesidad de un alto número de parámetros que pueden ser difíciles de estimar. Esta representación a menudo resulta en una parametrización excesiva del modelo con respecto de las mediciones disponibles. Dicha situación puede paliarse parcialmente asignando valores a ciertos parámetros que tienen un impacto limitado en la capacidad de predicción del modelo. Por ejemplo se puede asignar el valor 0.2 a los coeficientes de semisaturación K_S . El uso del enfoque cibernético no es factible con redes metabólicas a escala genómica. Por lo tanto, es necesaria una reducción de los modos elementales de flujo para aplicar este método. Para eso se propuso el llamado *método de análisis de rendimiento* por el cual si un microorganismo consume al sustrato S para generar al producto P y la biomasa X , el espacio de rendimiento está representado por un espacio

bidimensional compuesto por la biomasa X y el rendimiento del producto P con respecto al sustrato S . La proyección de los modos elementales de flujo en el espacio de rendimiento da como resultado una caja convexa, en lugar de un cono poliédrico no acotado. La caja convexa se puede definir por sus vértices, que corresponden a un subconjunto de modos elementales de flujo, que se denominan modos de generación (GM) y su número es menor al número de modos elementales de flujos. Se puede hacer una reducción adicional teniendo en cuenta solo a un porcentaje de la caja convexa en el espacio de rendimiento.

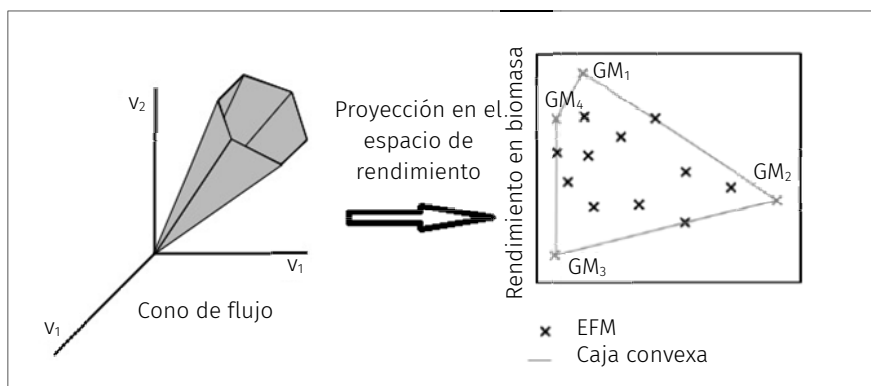


Figura 6.3. Análisis de rendimiento. Los EFM se proyectan en el espacio de rendimiento y se obtiene una caja convexa. Los EFM se reducen a sus vértices compuestos por un subconjunto de EFM.

Modelado cibernético híbrido agrupado

El modelado cibernético híbrido agrupado (Song y Ramkrishna, 2010) consiste en agrupar modos de flujo elementales (EFM) en familias y asignar a cada familia un EFM promedio. Las familias se determinan según los sustratos consumidos. Por ejemplo, si un microorganismo consume los sustratos S_1 o S_2 para producir biomasa X y un producto P , se utilizarán dos familias con sus modos elementales; una para el sustrato S_1 (con o sin sustrato S_2) y otra para el consumo de S_2 . En cada familia, los EFM se normalizan con respecto al sustrato elegido y se dividen en dos subfamilias: una con los EFM que producen biomasa y otra con los EFM que producen energía ignorando al resto. Debe mencionarse que no existe ninguna regla predeterminada para el agrupamiento y también es posible agrupar en familias y subfamilias de acuerdo con otros criterios. Por ejemplo, los productos P excretados también

se pueden usar para crear subfamilias, en lugar de familias de EFM dedicadas solo biomasa y ATP. En caso extremo, el número de familias de EFM será idéntico al número de EFM. Si se agruparan los EFM según la estequiometría de sus flujos, cada EFM corresponde a una familia diferente, por lo que se obtienen tantas familias como EFM lo cual no es deseable. Cada subfamilia $S \in \{F_K^B, F_K^A\}$ de la familia F_K se agrupa en un EFM agregado que se calcula como el promedio ponderado de los EFM individuales $E_j \in S$:

$$E_{SF} = \frac{\sum_{E_j \in SF} E_j n_j^{n_v+2}}{\sum_{E_j \in SF} n_j^{n_v+2}} \quad (6.33)$$

siendo n_j el peso que del EFM $E_j \in S$ y n_v la importancia relativa que se le da a este peso. El parámetro n_v se usa para compensar los efectos de agrupar a los EFM en un EFM promedio. El modelo necesita ser calibrado mediante datos experimentales y su interpretación biológica no está definida con claridad. Los pesos n_j generalmente se definen de la siguiente forma:

$$n_j = \begin{cases} Y_{B,j} & \text{para los EFM de producción de biomasa} \\ Y_{A,j} & \text{para los EFM de producción de ATP} \end{cases} \quad (6.34)$$

siendo $Y_{B,j}$ y $Y_{A,j}$ respectivamente el rendimiento de biomasa y el rendimiento de ATP con respecto al sustrato de la familia F_K elegido como referencia. Luego, para cada familia F_K compuesta por las subfamilias (F_K^B y F_K^A), el EFM agregado se calcula como un promedio ponderado de las dos subfamilias:

$$E_{Fk} = (1 - w)F_K^A + wF_K^B; 0 < w < 1 \quad (6.35)$$

con w determinado de tal manera que se satisfaga el requerimiento de ATP en función de la velocidad de crecimiento que debe determinarse experimentalmente. Una vez que todas las familias se agruparon en el EFM agregado E_{Fk} , se puede aplicar el modelado cibernético híbrido (HCM). En resumen, el enfoque de modelado cibernético híbrido se expresa mediante los siguientes pasos:

1. Calcule los EFM de la red metabólica.
2. Agrupe los EFM en familias según los sustratos disponibles en el medio.
3. Normalice los EFM según el sustrato de referencia para cada familia.
4. Divida cada familia en dos subfamilias, una con EFM que produzcan biomasa y otra con EFM que produzcan ATP e ignore a los restantes EFM.
5. Calcule un EFM agregado para cada subfamilia.
6. Calcule un EFM agregado mediante el promedio ponderado utilizando los EFM de las subfamilias.
7. Asigne una cinética cibernética a cada EFM agregado y deduzca el sistema de ecuaciones diferenciales resultante.

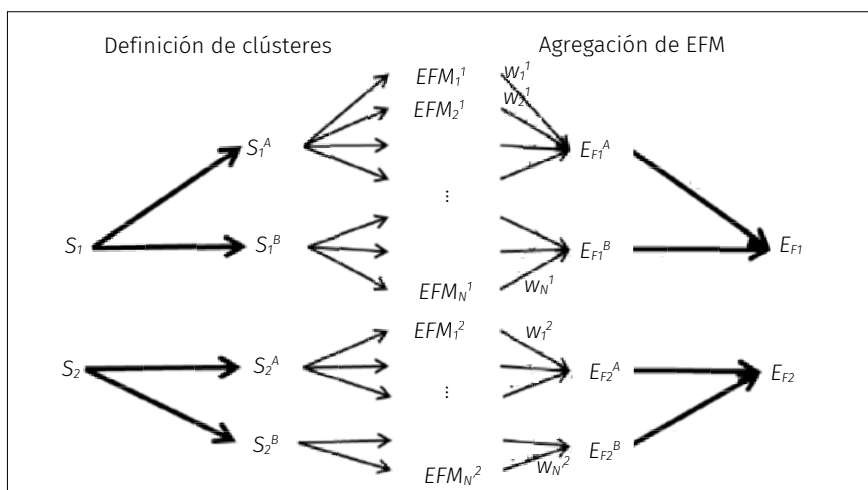


Figura 6.4. Vista esquemática del modelo cibernético híbrido (Song y Ramkrishna, 2010).

Desafortunadamente, la reducción del número de EFM en ambos métodos no siempre es suficiente y por lo general se obtienen demasiadas reacciones macroscópicas; particularmente en presencia de varios sustratos consumidos y varios productos excretados por los microorganismos. Además, para cada reacción macroscópica, se deben estimar al menos seis parámetros cinéticos y los coeficientes n_v y w deben estimarse para cada familia. Como consecuencia,, se produce rápidamente una explosión del número de parámetros a estimar y se limita la aplicabilidad del método.

REDUCCIÓN DINÁMICA DEL METABOLISMO NO BALANCEADOS

Esta propuesta (Baroukh *et al.*, 2014) se basa en considerar un bioproceso batch que involucra a microorganismos que crecen en un reactor de tipo tanque agitado perfectamente mezclado de volumen constante. Los microorganismos consumen sustratos extracelulares representados por el vector $\mathbf{x}_s$ para sintetizar la biomasa X y generar productos excretados representados por el vector $\mathbf{x}_p$. La red metabólica del microorganismo está representada por la matriz estequiométrica $S \in \mathbb{R}^{I \times J}$ que contiene I metabolitos y J reacciones. Al aplicar un balance de materia, el bioproceso se puede representar mediante el siguiente sistema de ecuación diferencial ordinaria:

$$\frac{d\mathbf{m}}{dt} = \frac{d \begin{pmatrix} S \\ C \\ P \\ B \end{pmatrix}}{dt} = \begin{pmatrix} K_S \\ K_C \\ K_P \\ K_B \end{pmatrix} \mathbf{v}B = \mathbf{S}\mathbf{v}B \quad (6.36)$$

donde $\mathbf{m}$ representa al vector de concentración de metabolitos que está integrado por biomasa la X , los sustratos extracelulares $\mathbf{x}_s$, los metabolitos intracelulares $\mathbf{x}_c$ y los productos excretados $\mathbf{x}_p$. Las concentraciones se expresan en términos de concentraciones en solución y no en términos de concentraciones por unidad de célula. El vector cinético $\mathbf{v} \in \mathbb{R}^J$ representa las velocidades de reacción en la red metabólica por unidad de biomasa. Las matrices $\mathbf{K}_S \in \mathbb{R}^{n_S \times |J|}$, $\mathbf{K}_C \in \mathbb{R}^{n_C \times |J|}$, $\mathbf{K}_P \in \mathbb{R}^{n_P \times |J|}$ y $\mathbf{K}_B \in \mathbb{R}^{1 \times |J|}$ son las matrices estequiométricas de la red metabólica para el sustrato, para los metabolitos internos, para los productos y para la biomasa ($n_S + n_C + n_P + 1 = |J|$) respectivamente. Los coeficientes estequiométricos se conocen a priori ya que provienen del conocimiento de la red metabólica y no es necesario determinarlos experimentalmente. La hipótesis de estado cuasi-estacionario implica que los metabolitos internos no se acumulan, es decir que $\mathbf{K}_C \mathbf{v} = 0$. Si se asume que dicha hipótesis es aplicable solo a grupos de reacciones metabólicas denominadas subredes ($\mathbf{S}_N$), los metabolitos restantes que interconectan las subredes, que denominamos A ($A \in C$), no se encuentran en la condición de estado cuasiestacionario. Estos metabolitos pueden acumularse y, por lo tanto sus concentraciones no estarán en estado estacionario.

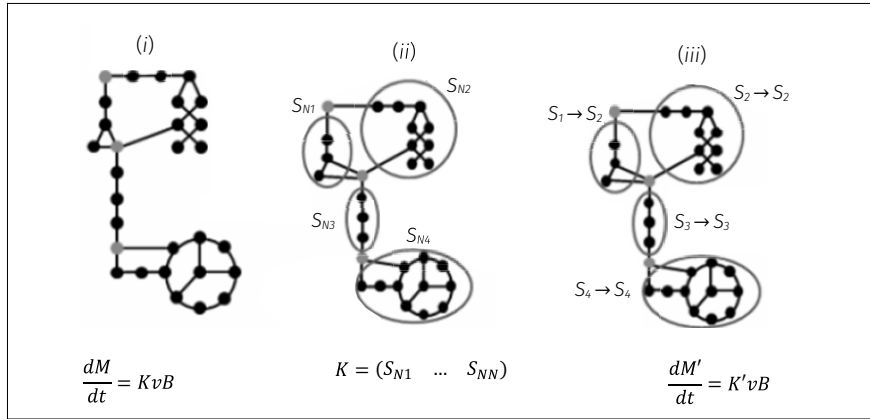


Figura 6.5. La red completa (i) se descompone en subredes ($\mathbf{S}_N$) que están en estado cuasi-estacionario (ii). Las subredes se representan mediante un conjunto de reacciones macroscópicas ($\mathbf{S}_\alpha \rightarrow \mathbf{P}$). (iii) Las reacciones macroscópicas se caracterizan por una cinética. Los metabolitos de enlace que interconectan a las subredes pueden acumularse (círculos rojos) o reutilizarse, lo que permite modelar la dinámica de la red.

A partir del paso (iii), se obtiene un sistema de ecuaciones diferenciales ordinarias (ODE), que representa, a escala macroscópica, al bioproceso, así como a los procesos intracelulares y la acumulación de metabolitos.

Las dimensiones de las matrices en el modelo completo descrito en el paso 1 son $K \in \mathbb{R}^{nm \times |J|}$, $v \in \mathbb{R}^{|J|}$, mientras que para las del modelo resultante proporcionado por este enfoque son $K' \in \mathbb{R}^{nm' \times nE}$ y $\alpha \in \mathbb{R}^{nE}$. La hipótesis de estado cuasi-estacionario para subredes se basa en:

1. La presencia de rutas metabólicas que corresponde a funciones metabólicas conocidas.
2. La presencia de grupos de reacciones reguladas.
3. La presencia de diferentes compartimentos en una célula (por ejemplo, mitocondrias).

Por lo tanto, los grupos de reacciones se determinan teniendo en cuenta estos mecanismos intracelulares. Cabe señalar que algunas reacciones intracelulares pueden pertenecer a varios grupos de reacciones. Matemáticamente, esto está representado por columnas redundantes en la matriz estequiométrica S . Los metabolitos restantes (A) que interconectan las subredes que se forman usando estas reglas generalmente están situados en un punto de ramificación entre varias vías metabólicas o son productos finales de vías metabólicas. Las subredes corresponden matemáticamente a una partición de la matriz estequiométrica K en submatrices $K_{S_{Ni}}$ formadas por reacciones agrupadas:

$$S = (K_{S_{N1}} \dots K_{S_{Nk}}) \quad (6.37)$$

donde $K_{S_{Ni}} \in \mathbb{R}^{nm \times n_{S_{Ni}}} (\sum n_{S_{Ni}} = nr_i)$ representa la subred i compuesta por metabolitos entrantes y salientes $S_{S_{Ni}}$ y $P_{S_{Ni}}$ que se pueden acumular y metabolitos intermedios $S_{S_{Ni}}$ en estado cuasiestacionario. $S_{S_{Ni}}$ y $P_{S_{Ni}}$ son las sub-matrices asociadas a los vectores de sustratos x_s , de productos x_p , a la biomasa X y a los metabolitos intracelulares x_A que se pueden acumular. Se supone que cada subred se encuentra en estado cuasi-estacionario:

$$K_{S_{Ni}} v_{S_{Ni}} = 0 \quad \forall i = 1, \dots, k \quad (6.38)$$

Cada subred se puede representar mediante un conjunto reducido de reacciones macroscópicas:

$$v_{S_{Ni}} = K_{S_{Ni}} \alpha_{S_{Ni}} \quad \forall i = 1, \dots, k (\alpha_{S_{Ni}} \geq 0) \quad (6.39.1)$$

$$(K_{S_{Ni}} E_{S_{Ni}}) S_{S_{Ni}} \rightarrow (K_{P_{Ni}} E_{S_{Ni}}) P_{S_{Ni}} \quad (6.39.2)$$

donde $E_{S_{Ni}}$ es la matriz de los modos de flujo elementales de la subred S_{Ni} y $\alpha_{S_{Ni}}$ es el vector de peso de los modos de flujo elementales. El vector $\alpha_{S_{Ni}}$ puede interpretarse como el vector de cinéticas de las reacciones macroscópicas descritas por la matriz estequiométrica $K_{S_{Ni}}$. Al agrupar todas las subredes se obtiene el siguiente sistema de ecuaciones diferenciales:

$$\frac{d\mathbf{m}}{dt} = (K_{S_{N1}} \dots K_{S_{Nk}}) \begin{pmatrix} v_{S_{N1}} \\ \dots \\ v_{S_{Nk}} \end{pmatrix} B \quad (6.40.1)$$

$$\frac{d\mathbf{m}}{dt} = (K_{S_{N1}} E_{S_{N1}} \dots K_{S_{Nk}} E_{S_{Nk}}) \begin{pmatrix} \alpha_{S_{N1}} \\ \dots \\ \alpha_{S_{Nk}} \end{pmatrix} B \quad (6.40.2)$$

$$\frac{d\mathbf{m}}{dt} = K_E \alpha B \quad (6.40.3)$$

Donde solo los metabolitos $\mathbf{x}_A$ pueden acumularse. Se asume que cualquier otro metabolito $C_j \in C \setminus A$ no se acumula. Por lo tanto:

$$\frac{d\left(\frac{C_j}{V}\right)}{dt} = 0 \quad \forall C_j \in C \setminus A \quad (6.41)$$

A los componentes $C_j \in C \setminus A$ se les asigna una dinámica simple y proporcional a la concentración de B . Por lo tanto, se obtiene un modelo dinámico reducido, definido por el vector de metabolitos $\mathbf{m}' \in \mathfrak{R}^{nm'}$ y la matriz $\mathbf{K}' \in \mathfrak{R}^{nm' \times n_E}$, siendo n_E el número de reacciones macroscópicas:

$$\frac{d\mathbf{m}}{dt} = \frac{d \begin{pmatrix} \mathbf{x}_S \\ \mathbf{x}_P \\ \mathbf{x}_A \\ \mathbf{x}_B \end{pmatrix}}{dt} = \begin{pmatrix} \mathbf{K}'_S \\ \mathbf{K}'_P \\ \mathbf{K}'_A \\ \mathbf{K}_B \end{pmatrix} \alpha B = \mathbf{K} \alpha B \quad (6.42)$$

Este sistema de ecuaciones diferenciales es una versión simplificada del original con la misma estructura pero de dimensión mucho menor, donde se permite la acumulación de algunos metabolitos internos ($\mathbf{x}_A$). Solo es necesario determinar la cinética α de las reacciones macroscópicas resultantes. Pueden utilizarse cinéticas clásicas que se encuentran en la literatura como la ley de acción de masas, la ley de potencia, Michaelis–Menten, Hill y la cinética cibernética. La elección es a menudo arbitraria y el número total de parámetros en los modelos cinéticos debe ser compatible con los datos experimentales disponibles para que se pueda lograr la validación del modelo. Una vez que se determina la cinética α , todos los flujos metabólicos se pueden calcular usando la siguiente expresión:

$$\mathbf{v} = \begin{pmatrix} E_{S_{N1}} \alpha_{S_{N1}} \\ \dots \\ E_{S_{Nk}} \alpha_{S_{Nk}} \end{pmatrix} \quad (6.43)$$

La definición de biomasa X no es convencional pues la biomasa generalmente se representa como una composición promedio y constante de macromoléculas presentes en la célula. Con la hipótesis de estado cuasiestacionario, cualquier elemento del sustrato $\mathbf{x}_S$ termina en biomasa X o en productos excretados $\mathbf{x}_P$. Pero cuando hay acumulación de metabolitos internos, no todos los elementos químicos del sustrato $\mathbf{x}_S$ terminan en biomasa X o en productos $\mathbf{x}_P$. Esto se debe a que pueden acumularse en la biomasa total, indicada como B que puede determinarse gracias a un balance de masa para cada elemento químico:

$$B_Z(t) = \sum_A Z_A A(t) + Z_B X(t) \quad (6.44)$$

donde Z corresponde a un elemento químico $Z \in \{C; N; O; H; P; S; \dots\}$; y Z_A y Z_B corresponde al número de elemento químico Z por mol de metabolitos acumulables x_A y biomasa X ; $A(t)$ y $X(t)$ corresponden a las concentraciones de A y X en el momento t , y $B_Z(t)$ corresponde a la concentración del elemento químico Z en la biomasa total X en el momento t .

En resumen, la metodología, involucra a los cuatro siguientes pasos:

1. Encontrar en la literatura o construir la red metabólica del microorganismo en estudio.
2. Agrupar las reacciones metabólicas en subredes que se asumen en estado cuasiestacionario.
3. Reducir cada subred a un conjunto de reacciones macroscópicas mediante el análisis de modos elementales.
4. Definir la cinética de las reacciones macroscópicas obtenidas y deducir un sistema ODE.

Actividades

1. Utilice la red metabólica IAF1260 descargada de la base de datos BIGG (<http://bigg.ucsd.edu/models>) para obtener una representación matemática reducida aplicando el método de reducción de la dimensión de redes metabólicas detallado en la sección 6.1 al considerar dos funcionalidades alternativas: metabolismo aeróbico y metabolismo anaeróbico. Programe dicho método como un problema de optimización MILP en GAMS y resuélvalo. Informe cuál es la dimensión de la red obtenida y a continuación, obtenga una red alternativa que incremente en dos unidades los grados de libertad de dicha red. Luego, obtenga las representaciones alternativas de la red reducida introduciendo cortes que excluyan a las representaciones anteriores.

2. Utilizando la red reducida que se obtuvo en el punto anterior, aplique la metodología de reducción dinámica del metabolismo no balanceado para obtener una representación macroscópica que considere la potencial acumulación intracelular de citrato y de etanol. Represente gráficamente de la topología la red macroscópica.

REFERENCIAS BIBLIOGRÁFICAS

- RÖHL, A. & BOCKMAYR, A. (2017). A mixed-integer linear programming approach to the reduction of genome-scale metabolic networks. *BMC Informatics*, 8-2.
- ORTH, J.; THIELE, I. & PALSSON, BØ. (2010). What is flux balance analysis? *Nature Biotechnology*, 28, 245-248. <https://doi.org/10.1038/nbt.1614>
- BAROUKH, C.; MUÑOZ-TAMAYO, R. (...) BERNARD, O. (2014). DRUM: A New Framework for Metabolic Modeling under Non-Balanced Growth. Application to the Carbon Metabolism of Unicellular Microalgae. *Plos One*. <https://doi.org/10.1371/journal.pone.0104499>
- KIM, J. I.; VARNER, J. D. & RAMKRISHNA, D. (2008). A hybrid model of anaerobic E. coli GJT001: combination of elementary flux modes and cybernetic variables. *Biotechnology Progress*, 24, 993-1006. <https://doi.org/10.1002/btpr.73>
- SONG, H.-S. & RAMKRISHNA, D. (2010). Prediction of metabolic function from limited data: Lumped Hybrid Cybernetic Modeling (L-HCM). *Biotechnology and Bioengineering*, 106, 271-284. <https://doi.org/10.1002/bit.2269>

Lecturas adicionales

- RAMKRISHNA, D. & SONG H. (2018). *Cybernetic Modeling for Bioreaction Engineering*. Cambridge Series in Chemical Engineering. <https://doi.org/10.1017/9780511731969>

PARTE 3

Aplicaciones metabólicas,
bioquímicas y biotecnológicas

7 Optimización de flujos en redes metabólicas

Un modelo estequiométrico a escala genómica representa al conjunto de las reacciones metabólicas de un organismo determinado y su vínculo con los genes correspondientes. En principio, un modelo de este tipo puede reconstruirse para cualquier organismo con un genoma anotado. La decisión de incluir una reacción metabólica en el modelo se toma en función de si la enzima que cataliza la reacción está codificada en el genoma del organismo de interés. Además, los modelos a escala genómica contienen información sobre la reversibilidad de las mismas al considerar evidencia experimental y/o información termodinámica. Los modelos estequiométricos a escala genómica contienen vínculos entre reacciones metabólicas y genes que codifican las enzimas relevantes. Esta información se almacena en asociaciones del gen con la enzima con la reacción (GPR). Los GPR se pueden clasificar en términos generales en cuatro categorías diferentes:

1. *Mapeo uno a uno*: una reacción r es catalizada por una única enzima E que está codificada por un gen único G .
2. *Isoenzimas*: una reacción r es catalizada por cualquiera de dos enzimas $E1$ o $E2$ codificadas por los genes $G1$ y $G2$, respectivamente. Cualquiera de los dos o más genes es suficiente para permitir esta reacción. Por lo tanto, se indican con una relación OR en la asociación GPR correspondiente ($G1$ OR $G2$).
3. *Complejo proteico de unidades múltiples*: una reacción r es catalizada por un complejo de dos (o más) enzimas $E1$ y $E2$ codificadas por los genes $G1$ y $G2$, respectivamente. Ambos genes son necesarios para permitir esta reacción y, por lo tanto, se indican con una relación AND ($G1$ AND $G2$) en la asociación GPR.
4. *Combinación de isoenzimas*: con poca frecuencia, una reacción r está codificada por múltiples isoenzimas. Los genes relevantes se vinculan mediante una combinación de relaciones AND y OR en la asociación GPR. Por ejemplo, una relación GPR de la forma ($G1$ AND $G2$) OR ($G3$) OR ($G4$ AND $G5$) significa que la reacción r está codificada por tres isoenzimas, dos de las cuales son complejos proteicos multiunitarios codificados por ($G1$ AND $G2$) AND ($G4$ AND $G5$), respectivamente, mientras que la tercera isoenzima está codificada por el gen $G3$.

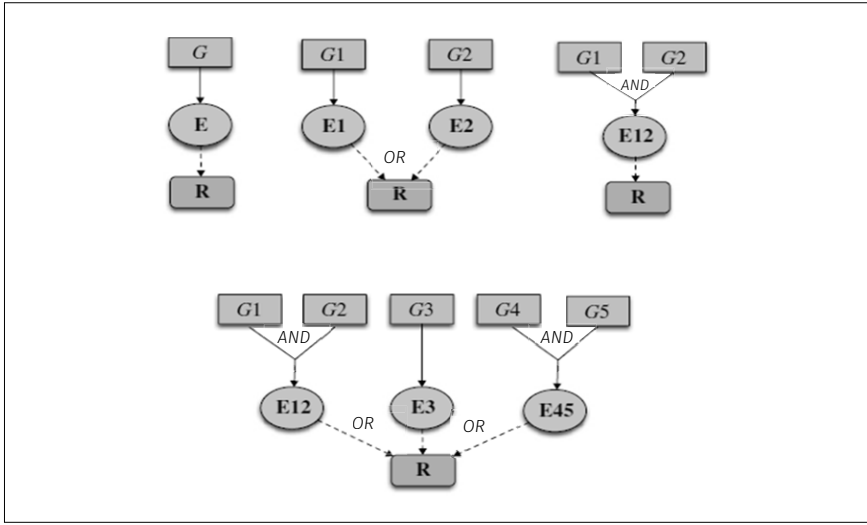
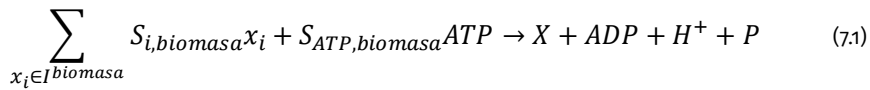


Figura 7.1. Relaciones GPR. G, E y R denotan genes, enzimas y reacciones, respectivamente. Las flechas sólidas representan la producción de enzimas por uno o más genes, mientras que las flechas discontinuas indican la catálisis de una reacción por parte de una o más enzimas.

Un componente fundamental de un modelo GSM es la reacción de formación de biomasa que proporciona una manera conveniente de inventariar todos los componentes que la integran. La «reacción» de producción de biomasa sirve como un sumidero de todos los precursores de biomasa en sus proporciones apropiadas y se puede expresar de la siguiente manera:



Donde x_i es un precursor de biomasa (por ejemplo, aminoácidos, lípidos y carbohidratos) y $S_{i,biomasa}$ es el coeficiente estequiométrico correspondiente del precursor de biomasa i . Estos coeficientes se determinan según la fracción de biomasa aportada por el metabolito x_i medida en base molar y tienen como unidades a $mmol/gDW$ donde gDW denota gramos de peso seco de células. La cantidad de ATP necesaria para el crecimiento $S_{ATP,biomasa}$, biomasa por unidad de biomasa producida, se denomina ATP de mantenimiento asociado al crecimiento. La composición de la biomasa y, por tanto, la reacción de construcción de biomasa varían de un organismo a otro en, por ejemplo, el contenido de clorofila y de varios pigmentos presentes en los organismos

fotosintéticos. El flujo de la reacción de formación de biomasa indica a la velocidad de crecimiento de la célula por cantidad de recurso limitante absorbido del entorno extracelular. Este recurso limitante suele ser la fuente de carbono, pero otros recursos como el oxígeno, la luz o la fuente de nitrógeno pueden volverse limitantes en diferentes condiciones.

Cada metabolito en un modelo a escala genómica puede tener una o más designaciones que denotan su pertenencia a un compartimento celular particular. Por ejemplo, «A[c]» se refiere al metabolito A en el citosol, mientras que «A[e]» se refiere al componente A en el espacio extracelular. Las designaciones [c] y [e] son solo dos de las muchas designaciones de los compartimentos que normalmente se encuentran en los modelos a escala genómica, incluyendo a [m] para mitocondrias y [p] para membrana periplásmica del cloroplasto. Observemos el efecto de una intervención letal en el ejemplo dado por el siguiente ejemplo.

$R_1: g_1 \text{ AND } g_2$	Complejo enzimático
$R_1: g_2 \text{ OR } g_3$	Iso-enzimas
$R_3: g_2 \text{ OR } g_3 \text{ OR } g_4 \text{ OR } g_5$	Varias iso-enzimas

La eliminación simultánea de los genes g_2 y g_3 es necesaria para desactivar la reacción r_2 , pero también se desactivará la reacción r_1 , ya que g_2 ha sido eliminado. Por lo tanto, si bien es posible simular la eliminación de las reacciones r_1 o r_2 del modelo, la reacción r_2 sola nunca puede eliminarse sin eliminar también r_1 . Se puede ver que es muy difícil desactivar la reacción r_3 ya que se requeriría la eliminación simultánea de cuatro genes diferentes g_2 , g_3 , g_4 y g_5 . Además de la eliminación de un solo gen, también es posible realizar eliminaciones de más de un gen. La eliminación simultánea de todos los genes implicará una anulación del crecimiento, es decir que será una intervención letal. La figura 7.3 ilustra una intervención letal. En un triplete letal, solo la eliminación simultánea de los tres genes del triplete resultará en letalidad. La eliminación de cualquier otro subconjunto de dos o menos genes del triplete no afectará el crecimiento. Los conjuntos de letales también se conocen como conjuntos de corte mínimo. El estudio de la letalidad puede arrojar luz sobre qué genes pueden complementarse unos con otros, e incluso cómo diferentes vías pueden compensarse entre sí. Las intervenciones letales también encuentran aplicaciones en la predicción de posibles objetivos farmacológicos combinatorios frente a organismos patógenos, así como en el desarrollo de posibles estrategias terapéuticas novedosas para combatir el cáncer.

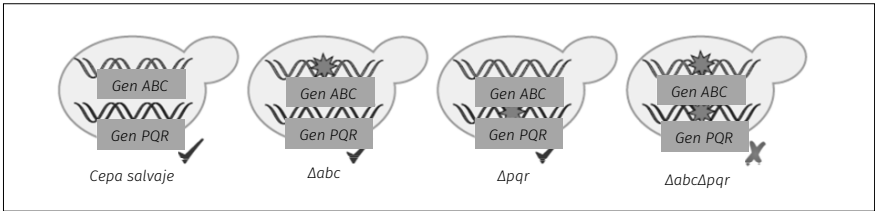


Figura 7.2. Un ejemplo para ilustrar intervenciones letales. Se muestra una célula con dos genes; «abc» y «pqr». Los mutantes con una sola remoción genética «abc» o «pqr» son viables, al igual que la cepa de tipo salvaje. El mutante con la remoción de «abc» y «pqr» no es viable. Por lo tanto, «abc» y «pqr» forman un par letal.

Se han desarrollado numerosos métodos computacionales para identificar intervenciones letales en el marco del análisis basado en restricciones. FBA se ha utilizado para predecir intervenciones letales en redes metabólicas de organismos como la levadura de cerveza. En términos generales, existen tres enfoques principales para identificar intervenciones letales. Una enumeración exhaustiva deberá resolver demasiados problemas lineales para determinar las intervenciones letales sintéticas. La siguiente tabla ilustra el tamaño del espacio de búsqueda según el número de intervenciones letales.

Orden	Resolución exhaustiva de problemas lineales
Simple	2.05×10^3
Doble	1.5×10^6
Triple	9.27×10^8
Cuádruple	4.10×10^{11}

OPTIMIZACIÓN DE FLUJOS EN REDES METABÓLICAS MEDIANTE PROGRAMACIÓN MIXTA ENTERO-LINEAL

El modelado matemático de redes metabólicas permite determinar los cambios en la estructura y en las fuerzas reguladoras que deben considerarse para optimizar un proceso metabólico. El modelado lineal entero mixto de intervenciones en estas redes permite que la solución de la misma proporcione información sobre los niveles de enzimas que deben estar presentes, el alcance de dichos cambios potenciales y las modificaciones de la estructura

reguladora para optimizar a una red (Hatzimanikatis *et al.*, 1996). Dado que los modelos de los procesos metabólicos son una aproximación, la manipulación genética de un organismo puede causar efectos secundarios que no fueron consideradas en dicho modelo. Por lo tanto, el resultado de la resolución del problema debe ser considerado como una sugerencia razonable para cambios en la red metabólica que podrían brindar mejoras útiles en el rendimiento del metabolito deseado. Una orientación y no las reglas cuantitativas precisas, es el resultado que se debe esperar de este tipo de cálculos.

El primer paso para la optimización de una red metabólica es la obtención de un modelo matemático con un número dado de bucles reguladores. Estos bucles reguladores se clasifican como bucles de activación que aumentan la actividad de una enzima reguladora o de inhibición si disminuyen la actividad de la enzima reguladora. El objetivo de la optimización de una red metabólica consiste en determinar cuál de los circuitos reguladores debe conservarse así como también el número, tipo y nivel de manipulación de las actividades enzimáticas para optimizar una determinada función de la vía metabólica, tal como la producción de metabolitos primarios o secundarios. Muchas vías metabólicas son comunes a muchos organismos. Sin embargo, las enzimas que catalizan a la misma reacción en diferentes organismos no tienen necesariamente las mismas propiedades catalíticas y reguladoras. Los métodos de ADN recombinante permiten introducir en un organismo el ADN de otros organismos con la posibilidad de combinar características reguladoras de las vías metabólicas presentes en estos organismos. Los métodos de ingeniería de proteínas permiten modificar las propiedades de las enzimas naturales y diseñar enzimas con nuevas características reguladoras. Por lo tanto, el número posible de bucles reguladores para una determinada vía en un organismo varía desde el número de bucles existentes hasta este número, más el número de bucles reguladores adicionales que pueden introducirse en la misma vía mediante ingeniería genética. De manera general, podemos considerar una superestructura reguladora en la que cada metabolito del sistema puede potencialmente regular a cualquier enzima. La formulación matemática de esa superestructura reguladora conduce a un problema combinatorio y la solución al problema MILP proporcionará el rendimiento máximo o mínimo alcanzable por un sistema determinado. Por lo tanto, proporcionará información valiosa para la ingeniería metabólica y de proteínas.

Específicamente, en esta sección se aborda la cuestión de cómo se deben modificar los niveles de regulación y de catálisis para maximizar el rendimiento de una red metabólica en un metabolito dado. Consideremos a un sistema metabólico que consta de N metabolitos y S parámetros manipulados

externos. Para cada metabolito en el sistema, podemos escribir la forma general de su cinética en representación del sistema-S:

$$\frac{dx_i}{dt} = V_i^+(x_i, \dots, x_N; P_1, \dots, P_S) - V_i^-(x_i, \dots, x_N; P_1, \dots, P_S) \quad i = 1, \dots, N \quad (7.2)$$

donde V_i^+ y V_i^- son velocidades netas agregadas que describen los procesos de producción y consumo, respectivamente del metabolito i y x_i es la concentración de dicho metabolito. Dichas velocidades se describen mediante la siguiente forma de ley de potencia:

$$V_i^+ = \alpha \prod_{j=1}^N X_j^{g_{ij}} \prod_{l=1}^S P_l^{a_{il}} \quad (7.3)$$

$$V_i^- = \beta \prod_{j=1}^N X_j^{h_{ij}} \prod_{l=1}^S P_l^{b_{il}} \quad (7.4)$$

donde x_j ($j = 1, \dots, N$) es la concentración del metabolito j , y P_l ($l = 1, \dots, S$) es la actividad del parámetro manipulado l . Las constantes de velocidad α y β son no-negativas, y los órdenes cinéticos g_{ij} , h_{ij} , a_{il} , b_{il} se definen mediante las siguientes igualdades:

$$g_{ij} = \frac{X_j}{V_i^+} \frac{\partial V_i^+}{\partial X_j} \quad (7.5)$$

$$h_{ij} = \frac{X_j}{V_i^-} \frac{\partial V_i^-}{\partial X_j} \quad (7.6)$$

$$a_{il} = \frac{P_l}{V_i^+} \frac{\partial V_i^+}{\partial P_l} \quad (7.7)$$

$$b_{il} = \frac{P_l}{V_i^-} \frac{\partial V_i^-}{\partial P_l} \quad (7.8)$$

Cada metabolito y cada parámetro del modelo que tienen un efecto sobre cualquiera de los procesos agregados están asociados con un orden cinético distinto de cero. Si un metabolito actúa sobre como sustrato o activador, el orden cinético correspondiente es un real positivo pero si actúa como inhibidor, su orden cinético será un real negativo. Lo mismo se aplica a los órdenes cinéticos de los parámetros manipulados; es decir que un parámetro que actúa como activador (o inhibidor) está asociado a un orden cinético positivo (o negativo). Habiendo construido la representación del sistema S para el proceso de interés, el estado estacionario se puede encontrar mediante la siguiente transformación logarítmica:

$$\mathbf{y} = -\mathbf{L}^{-1} \mathbf{M} \mathbf{q} + \mathbf{L}^{-1} \mathbf{d} \quad (7.9)$$

Donde $\mathbf{y}$ es un vector n -dimensional con los siguientes elementos:

$$y_j = \ln(x_j) \quad (7.10)$$

Y $\mathbf{q}$ es un vector l -dimensional con los siguientes elementos:

$$q_l = \ln(P_l) \quad (7.11)$$

Mientras que las definiciones respectivamente correspondientes a $\mathbf{d}$, $\mathbf{L}$ y $\mathbf{M}$ son las siguientes:

$$d_i = \ln(\beta_i/\alpha_i) \quad (7.12)$$

$$L_{ij} = g_{ij} - h_{ij} \quad (7.13)$$

$$M_{ij} = a_{il} - b_{il} \quad (7.14)$$

Sin embargo, cuando describimos u observamos un sistema metabólico, puede haber algunas salidas que son funciones de la concentración de los metabolitos y de otros parámetros del sistema. El i -ésimo elemento de un conjunto de K productos metabólicos también se puede describir mediante funciones de ley de potencia:

$$\phi_i = \gamma_i \prod_{j=1}^N X_j^{f_{ij}} \prod_{l=1}^S P_l^{c_{il}} \quad \forall i=1, \dots, K \quad (7.15)$$

donde γ_i es un real y los números reales f_{ij} y c_{il} están definidos por:

$$f_{ij} = \frac{X_j}{\phi_i} \frac{\partial \phi_i}{\partial X_j} \quad (7.16)$$

$$c_{il} = \frac{P_l}{\phi_i} \frac{\partial \phi_i}{\partial P_l} \quad (7.17)$$

Si queremos estudiar cómo los cambios en las características catalíticas y reguladoras de ciertas enzimas afectarán el desempeño de una vía metabólica, debemos examinar cómo el desempeño depende de los órdenes cinéticos que cuantifican la afinidad de las diversas enzimas por sus sustratos y sus moduladores. Cuando un proceso está inhibido por el metabolito x_j , entonces el orden cinético correspondiente g_{ij} , tendrá un valor finito real negativo. Igualando este orden cinético a cero, podemos simular y observar cómo respondería el sistema si se hubiera desactivado la innhibición del proceso q . Los cambios en los órdenes cinéticos de los metabolitos alteran el valor de los elementos de la matriz $\mathbf{L}$. Estos cambios determinan los valores de salida en forma no lineal del modelo debido a la inversión de la matriz $\mathbf{L}$. En consecuencia, el análisis de sensibilidad local de las salidas con respecto

a los órdenes cinéticos asume pequeños cambios en estos parámetros para estimar variaciones lineales en el entorno de los puntos de linealización. Por el contrario, cualquier intento de modificar la afinidad de una enzima por sus sustratos y moduladores dará como resultado cambios finitos, a menudo de órdenes de magnitud diferentes a los de las afinidades enzimáticas. Consecuentemente, existe la necesidad de un marco sistemático para optimizar la estructura reguladora con órdenes cinéticos que pueden diferir entre sí en uno o más órdenes de magnitud. La optimización de la estructura regulatoria implica decisiones discretas sobre los bucles regulatorios que deben desactivarse o la introducción de los bucles regulatorios en una vía metabólica. A continuación, se presenta la formulación del problema de optimización de la estructura regulatoria como un problema de programación mixta entero-lineal. La formulación aprovecha la linealidad de las ecuaciones en estado estacionario de la representación del sistema S , así como algunas transformaciones adicionales que permitirán expresar a las variables enteras del modelo en forma lineal.

Optimización del rendimiento de un metabolito en una vía metabólica lineal

La formulación matemática se presentará a través del primer ejemplo utilizado por Hatzimanikatis *et al.* (1996). Considere una vía lineal simple con dos metabolitos dependientes, X_1 y X_2 , con el segundo actuando como inhibidor de la primera reacción de la vía:

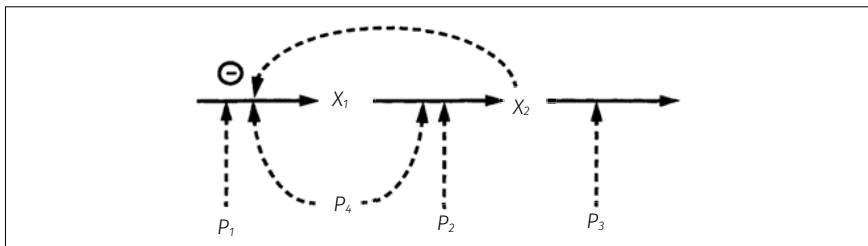


Figura 7.3. Vía metabólica lineal con retroalimentación inhibitoria (adaptado de Hatzimanikatis *et al.*, 1996).

Se consideran cuatro variables manipuladas externas: la cantidad de enzimas P_1 , P_2 y P_3 , que catalizan las tres reacciones y la cantidad de efector, P_4 , que activa la primera y la segunda reacción de la vía metabólica. La representación del sistema- S de la vía es:

$$\frac{dX_1}{dt} = \alpha_1 X_2^{g_{12}} P_1^{a_{11}} P_4^{a_{14}} - \beta_1 X_2^{h_{11}} P_1^{b_{12}} P_4^{b_{14}} \quad (7.18.1)$$

$$\frac{dX_2}{dt} = \alpha_2 X_2^{g_{21}} P_1^{a_{22}} P_4^{a_{24}} - \beta_2 X_2^{h_{21}} P_1^{b_{12}} P_4^{b_{14}} \quad (7.18.2)$$

Con los valores sugeridos por Voit (1992) en su ajuste de datos, la representación de la red ilustrada en la figura 7.4 será la siguiente:

$$\frac{dX_1}{dt} = X_2^{-1} P_1^1 P_4^2 - 0.02 X_1^{0.5} P_1^1 P_4^2 \quad (7.19.1)$$

$$\frac{dX_2}{dt} = 0.02 X_1^{0.5} P_2^1 P_4^2 - 2 X_2^1 P_3^1 \quad (7.19.2)$$

Si se quiere maximizar la concentración del metabolito X_2 , ¿cuál de los parámetros manipulados debe modificarse y cuál debería ser el nivel del parámetro manipulado? ¿Debe desactivarse el ciclo de inhibición por retroalimentación? Las preguntas anteriores deben considerar las siguientes restricciones:

1. El sistema está en estado estacionario.
2. $X_1^{SS} \leq 500$.
3. $V_3^{SS} \leq 10$.
4. Para las tres enzimas, solo se considera sobreexpresión; es decir, $P_l \geq 1$ ($\forall l = 1, \dots, 4$).

En las restricciones anteriores, el superíndice denota el valor de estado estacionario. En general, la concentración de los metabolitos, el valor de las variables manipuladas, las velocidades de reacción y las funciones metabólicas deben estar restringidas por los límites determinados por los rangos fisiológicamente factibles y por el conocimiento biológico disponible. Para ello se introduce un conjunto de nuevas variables:

$$q_l^r + q_l = \ln(P_l) \quad l = 1, \dots, 4 \quad (7.20)$$

donde q_l^r denota el logaritmo del valor de referencia del parámetro l y q_l denota al logaritmo del factor por el cual se multiplican los valores de referencia para dar el valor P_l . En el ejemplo aquí estudiado:

$$q_l^r = \ln(1) \quad (7.21.1)$$

$$q_l = \ln(1) \quad (7.21.2)$$

Se deben tomar dos decisiones; la primera se refiere a los parámetros manipulados, es decir a los niveles de expresión de la enzima y el nivel de concentración del efector externo que deben cambiarse. Es una decisión discreta porque se busca al único parámetro que debe cambiarse mientras que el resto de los mismos se fija en su valor de referencia. La segunda decisión discreta se refiere al ciclo regulatorio. La inactivación o activación del ciclo de inhibición por retroalimentación es equivalente a dar el valor 0 o 1 a la variable de decisión asociada y un orden cinético negativo de V_1^+ con respecto a x_2 y g_2 . Estas decisiones pueden modelarse mediante variables binarias, es decir, variables que pueden ser 0 o 1 solamente. Para la primera decisión discreta, se introduce un conjunto de variables binarias, w_l , para las que tendremos:

$$q_l^r + w_l q_l = \ln(P_l) \quad \forall l = 1, \dots, 4 \quad (7.22)$$

Si w_l es igual a cero, no se permiten cambios en el parámetro manipulado correspondiente P_l porque tenemos que $\ln(P_l) = q$. Además, debe introducirse la siguiente restricción:

$$w_1 + w_2 + w_3 + w_4 = 1 \quad (7.23)$$

La misma indica que solo el cambio de un parámetro w está permitido. La segunda decisión discreta se refiere a la velocidad q , en la cual se introduce la variable binaria z_{12} en el exponente de x_2 :

$$V_1^+ = X_2^{z_{12}g_{12}} P_1 P_4^2 \quad (7.24)$$

Si z_{12} es igual a 0, V_1^+ ya no depende de x_2 y, por lo tanto, se ha modelado la inactivación del ciclo de inhibición por retroalimentación. Si z_{12} es igual a 1, entonces la ecuación anterior se reduce a:

$$V_1^+ = X_2^{g_{12}} P_1 P_4^2 \quad (7.25)$$

Después de aplicar la transformación logarítmica al modelo en estado estacionario, el problema se puede formular matemáticamente de la siguiente manera:

$$\text{Maximizar } (y_2) \quad (7.26.1)$$

sujeto a:

Balances de materia:

$$0.5y_1 + z_{12}y_2 + q_2^r + w_2q_2 - q_1^r - w_1q_1 = \ln(1/0.02) \quad (7.26.2)$$

$$0.5y_1 - y_2 + q_2^r + w_2q_2 + 2q_4^r + 2w_4q_4 - q_3^r - w_3q_3 = \ln(2/0.02) \quad (7.26.3)$$

Cotas:

$$y_1 \leq \ln(500) \quad (7.26.4)$$

$$z_{12}y_2 + q_1^r + w_1q_1 + 2q_4^r + 2w_4q_4 \leq \ln(10) \quad (7.26.5)$$

Cotas en las actividades enzimáticas:

$$P_1 \geq 1 \quad (7.26.6)$$

$$P_2 \geq 1 \quad (7.26.7)$$

$$P_3 \geq 1 \quad (7.26.8)$$

Máximo número de manipulaciones

$$w_1 + w_2 + w_3 + w_4 = 1 \quad (7.26.9)$$

$$(y_2, q_4) \in \mathbb{R}; w_1 \in \{0,1\}$$

Si w_1 , w_2 y w_3 son iguales a 0, w_4 es igual a 1 y z_{12} es igual a 1, entonces el problema es el mismo considerado por Hatzimanikatis *et al.* (1996). Por tanto, resolviendo el problema dos veces, una para $z_{12} = 1$ y otra para $z_{12} = 0$, se considera el caso de maximización de X_2 con y sin regulación por realimentación, respectivamente. Sin embargo, también queremos encontrar cuál de los cuatro parámetros puede ser más eficiente en presencia o ausencia de regulación. Para encontrar la respuesta a la pregunta, debemos resolver el problema $2 \times 4 = 8$ veces. En general, si tuviéramos una ruta con m parámetros y n bucles regulatorios, la respuesta a la misma pregunta requeriría la enumeración de $m2^n$ soluciones para buscar la mejor solución. Por lo tanto, la resolución del problema requiere de un método que lo resuelva de manera eficiente. La formulación mixta entera-lineal del problema implica que cada variable continua y binaria aparecerá linealmente en la función objetivo y en las restricciones. Nótese que en el modelo hay algunos productos bilineales de variables continuas y binarias, tales como $w_l q_l$, $l = 1, \dots, 4$ y $(z_{12} y_2)$, que introducen no linealidades en el modelo. Estas no linealidades pueden

eliminarse mediante las transformaciones apropiadas, introduciendo un conjunto de variables continuas, $t_l = 1, 2, 3, 4$ y s_{12} , que sustituirán a los productos bilineales:

$$t_l = w_l q_l \quad \forall l = 1, \dots, 4 \quad (7.27.1)$$

$$s_{12} = z_{12} y_2 \quad (7.27.2)$$

Para cada una de las nuevas variables continuas, incluiremos cuatro restricciones lineales que garantizarán la consistencia entre el producto bilineal y las variables continuas correspondientes. Después de la introducción de estas variables continuas y de las restricciones obtenemos al siguiente problema MILP:

$$\text{Maximizar } y_2 \quad (7.28.1)$$

sujeto a:

$$0.5y_1 + s_{12} + q_2^r + t_2 - q_1^r - t_1 = \ln(1/0.02) \quad (7.28.2)$$

$$0.5y_1 - y_2 + q_2^r + t_2 - 2q_4^r + 2t_4 - q_3^r - t_3 = \ln(2/0.02) \quad (7.28.3)$$

$$y_1 \leq \ln(500) \quad (7.28.4)$$

$$-s_{12} + q_1^r + t_1 + 2q_4^r + 2t_4 \leq \ln(10) \quad (7.28.5)$$

$$P_1, P_2, P_3 \geq 1 \quad (7.28.6)$$

$$y_2 - s_{12} + y_2^L z_{12} \geq y_2^L \quad (7.28.4)$$

$$y_2 - s_{12} + y_2^U z_{12} \geq y_2^U \quad (7.28.5)$$

$$y_2^L z_{12} - s_{12} \leq 0 \quad (7.28.6)$$

$$-y_2^U z_{12} + s_{12} \leq 0 \quad (7.28.7)$$

$$q_1 - t_i + w_i q_i^L \geq q_i^L \quad (7.28.8)$$

$$q_1 - t_i + w_i q_i^L \leq q_i^U \quad (7.28.9)$$

$$w_i q_i^L - t_i \leq 0 \quad (7.28.10)$$

$$-w_i q_i^U + t_i \leq 0 \quad (7.28.11)$$

$$w_1 + w_2 + w_3 + w_4 \leq 1 \quad (7.28.12)$$

donde $y_i = 1, 2$, s_{12} , q_i y $t_l = 1, \dots, 4$ son variables continuas reales; $q_l^r = 1, \dots, 4$ son parámetros reales, z_{12} y w_i $i = 1, \dots, 4$ son variables binarias. Los superíndices U y L denotan a los límite superior e inferior, respectivamente, de

las variables correspondientes. La resolución del problema proporciona los siguientes valores óptimos para dichas variables:

$$X_1 = 100; X_2 = 5; P_l = 1 \quad \forall l = 1, \dots, 3; P_4 = 5\sqrt{2}$$

Hasta acá la única forma de aumentar la concentración de X_2 es mediante el aumento de P_4 en presencia del bucle regulador de retroalimentación. Aunque la vía metabólica es lineal simple con solo tres reacciones, el resultado no es obvio porque depende de los valores numéricos de los parámetros y de los límites de las variables.

Optimización del rendimiento de la producción de xmp y gmp

La figura 7.4 muestra una forma simplificada de una ruta metabólica de baja dimensión. En Hatzimanikatis *et al.* (1996) se pueden encontrar discusiones y referencias sobre las características bioquímicas de esta red metabólica. La representación del sistema-S de la red es la siguiente:

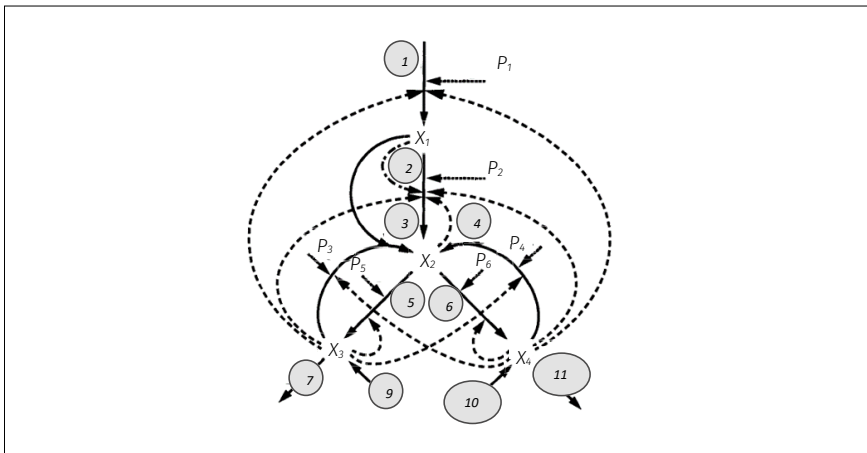


Figura 7.4. Representación de la red metabólica y de la interacción entre los diferentes componentes de la misma (adaptado de Hatzimanikatis *et al.*, 1996).

$$\frac{dx_1}{dt} = 900X_3^{-0.5}X_4^{-0.5}P_1-10X_2^{0.5}X_2^{-0.1}X_3^{-0.2}X_4^{0.2}P_2^{0.6}P_2^{0.4} \quad (7.29.1)$$

$$\frac{dx_2}{dt} = 7.34X_1^{-0.308}X_2^{-0.062}X_3^{-0.162}X_4^{-0.1}P_2^{0.37}P_3^{0.245}P_4^{0.385} \quad (7.29.2)$$

$$\frac{dx_3}{dt} = 2.71X_2^{0.409}X_3^{-0.387}P_5^{0.455}-0.036X_1^{0.14}X_3^{0.43}X_4^{-0.014}P_3^{0.28} \quad (7.29.3)$$

$$\frac{dx_4}{dt} = 13.03X_2^{0.041}X_4^{-0.399}P_6^{0.405}-0.143X_3^{-0.026}X_4^{0.46}P_4^{0.26} \quad (7.29.4)$$

Los órdenes cinéticos se calculan a partir de las velocidades y de sus elasticidades con respecto a los sustratos y moduladores, utilizando las transformaciones lineales. Los parámetros P_1 a P_6 representan la concentración de las enzimas que catalizan las reacciones correspondientes. El objetivo de la optimización es maximizar la concentración en estado estacionario del metabolito X_4 . El problema está sujeto a las siguientes restricciones:

- La concentración de los metabolitos X_1 , X_2 y X_3 no debe desviarse más del 10 % de sus valores de estado estacionario de referencia.
- La concentración de las enzimas no puede aumentar o disminuir más allá de ciertos límites.

Debe determinarse cuál de las actividades enzimáticas debe manipularse, dentro de ± 10 % de su valor de referencia, para maximizar a X_4 . La respuesta se puede encontrar resolviendo seis veces el problema de optimización utilizando un algoritmo de programación lineal. Sin embargo, si se usa la formulación MILP del problema, al incluir dos restricciones adicionales, se puede encontrar la respuesta resolviendo el problema en el que las siguientes restricciones adicionales incluyen a las variables binarias:

$$\sum_{i=1}^6 w_i = 1 \quad (7.30.1)$$

$$z_{13} + z_{14} + z_{21} + z_{22} + z_{23} + z_{24} + z_{34} + z_{43} + z_{53} + z_{64} = 10 \quad (7.30.2)$$

donde la primera restricción permite cambios en las actividades de solo una de las enzimas, y la segunda no permite ninguna modificación de la estructura reguladora. De manera similar, para maximizar a X_4 , puede determinarse cuáles de las dos enzimas deben manipularse, dentro del 10 % de su valor de referencia. La respuesta surge al introducir la siguiente restricción:

$$\sum_{i=1}^6 w_i = 2 \quad (7.32)$$

Siguiendo este procedimiento, se pueden considerar el cambio simultáneo de tres, cuatro, cinco o seis enzimas. Los valores óptimos de la función objetivo, el nivel de enzimas que deben cambiarse se presentan en la Tabla 7.1.

Tabla 7.1. Valores de la función objetivo (X_4), niveles de expresión enzimática y concentraciones de los metabolitos cuando una, dos, tres, cuatro, cinco y seis enzimas cambian simultáneamente, dentro del 10 % de su valor de referencia sin modificaciones en la estructura reguladora. (Hatzimanikatis *et al.*, 1996)

X_4	P_1	P_2	P_3	P_4	P_5	P_6	X_1	X_2	X_3
103.84	100	100	100	100	100	109.17	98.62	90	95.14
104.84	105.68	100	100	100	100	110	110	96.92	96.92
107.83	106.24	100	100	90	100	110	110	90.11	93.49
108.25	110	110	100	90	10	110	103.75	94.97	96.96
108.78	110	110	100	90	90	110	106.31	105.27	95.89
108.82	110	110	94.94	90	90	110	110	105.27	97.04

Una de las conclusiones que se pueden extraer de esta tabla está relacionada con las actividades de las enzimas de control. La enzima 6 parece tener el mayor efecto sobre la concentración de X_4 . Es interesante notar que cuando esta es la única enzima manipulada, la solución óptima no involucra el aumento máximo permitido en P_6 . Solo cuando se cambia otra actividad enzimática, el valor de P_6 aumenta hasta su límite superior. La manipulación simultánea de cuatro o más enzimas no mejora mucho la concentración de X_4 . En la Tabla 7.2 se presentan los coeficientes de control de concentración para los cuatro metabolitos con respecto a las seis enzimas. La indicación sobre qué enzima debe manipularse desde la máxima concentración del coeficiente de control x_4 es consistente con el resultado del cálculo de optimización. Además, las dos enzimas que se deben manipular, según los valores del coeficiente de control mostrados en la Tabla 7.2 son las enzimas P_6 y P_4 .

Tabla 7.2. Coeficientes de control de concentración para los cuatro metabolitos con respecto a las seis enzimas. (Hatzimanikatis *et al.*, 1996)

	C_{Pl}^{Xj}					
	P_1	P_2	P_3	P_4	P_5	P_6
X_1	1.999	-1.343	-0.657	0.101	-0.231	-0.156
X_2	1.504	0.001	-0.003	0.769	-0.977	-1.198
X_3	0.412	0.231	-0.232	0.363	0.106	-0.566
X_4	0.091	0.008	-0.008	-0.274	-0.047	0.427

Esto se debe a la existencia de restricciones en las concentraciones de los metabolitos y de otros parámetros fisiológicos e indica la necesidad de la resolución del modelo mixto entero lineal cuando se consideran restricciones sobre las funciones metabólicas. Si permitimos que los niveles de expresión de la enzima cambien entre 20 % y 500 % de su valor de referencia, los resultados serán bastante diferentes tal como se observa en la Tabla 7.3.

Tabla 7.3. Valores de la función objetivo, niveles de expresión enzimática y concentraciones de metabolitos cuando una, dos, tres, cuatro, cinco y seis enzimas cambian simultáneamente sus actividades entre 20 % y 500 % respecto de su valor de referencia y sin modificaciones en la estructura reguladora. (Hatzimanikatis *et al.*, 1996)

X_4	P_1	P_2	P_3	P_4	P_5	P_6	X_1	X_2	X_3
103.84	100	100	100	100	100	109.17	98.62	90	95.14
104.84	100	100	100	87.18	100	100	98.62	90	95.14
107.95	100	100	100	100	90.68	118.21	99.62	90	90
108.95	100	100	100	76.97	90.68	100	99.62	90	90
226.45	351.22	500	100	100	100	500	90	100	107.79
232.64	352.93	500	100	90.80	100	500	110	90	96.48
240.47	372.65	500	116.08	81.61	100	500	110	90	90
240.47	500	500	242.17	81.61	157.38	500	110	90	90

El valor óptimo de la concentración de X_4 se puede lograr de dos maneras diferentes cuando se consideran una y dos manipulaciones enzimáticas. El valor de la función objetivo, cuando se cambia de enzima, es el mismo que antes porque el factor limitante es en ambos casos la concentración del metabolito X_2 que es igual al máximo valor permitido. Cuando se manipulan dos enzimas existen dos soluciones posibles, y el valor de la función objetivo correspondiente es casi igual al valor del objetivo encontrado en el caso anterior e indicado en la Tabla 7.1. Además, los valores del objetivo en casos que involucran a dos y tres manipulaciones enzimáticas dependen de los límites impuestos a los niveles de expresión de las enzimas. Esta diferencia surge claramente de la naturaleza no lineal del sistema original, así como del carácter restringido del problema. La concentración máxima de X_4 que

se puede lograr mediante cambios en las seis enzimas es del 240,5 % con respecto del valor de referencia. Si consideramos cambios en la estructura regulatoria, al inactivar alguno o algunos de los bucles existentes, podríamos lograr un valor aún mayor de la función objetivo. Resolviendo el problema MILP al permitir cambios en los valores binarios z_{ij} , podemos encontrar el máximo que se puede lograr mediante la manipulación de la actividad de las enzimas y las modificaciones de la estructura reguladora, considerando las restricciones y las propiedades cinéticas dadas del sistema tal como están cuantificadas por las elasticidades de los sustratos y de los moduladores. La solución se puede encontrar cambiando la restricción (7.31) mediante la siguiente restricción:

$$z_{13} + z_{14} + z_{21} + z_{22} + z_{23} + z_{24} + z_{34} + z_{43} + z_{53} + z_{64} \leq 10 \quad (7.33)$$

la cual permite cambios en la estructura regulatoria. Múltiples combinaciones alternativas de estructuras reguladoras y manipulaciones enzimáticas corresponden a este valor óptimo.

Al modificar las estructuras reguladoras, es razonable buscar el mínimo número de bucles que deben inactivarse porque las modificaciones experimentales de las características reguladoras de una enzima son mucho más difíciles de cambiar que cambiar la concentración de dicha enzima. El modelo permite encontrar, entre las estructuras regulatorias alternativas, aquella con el máximo número de bucles regulatorios activos. La función objetivo será, en este caso, el número de bucles reguladores, y el objetivo será la maximización del mismo:

$$\text{Maximizar } z_{13} + z_{14} + z_{21} + z_{22} + z_{23} + z_{24} + z_{34} + z_{43} + z_{53} + z_{64} \quad (7.33.1)$$

sujeto a:

$$y_4 = 5501 \quad (7.33.2)$$

El número máximo de bucles reguladores que pueden permanecer activos en la red de reacción es igual a 5. La estructura reguladora óptima se presenta en la figura 7.6 y los niveles de metabolitos y enzimas se presentan en la Tabla 7.4.

Tabla 7.4. Valores óptimos para la función objetivo, niveles de expresión enzimática y concentraciones de los metabolitos cuando seis enzimas cambian simultáneamente, entre un 20 % y un 2500 % de su valor de referencia con modificaciones en la estructura reguladora (Hatzimanikatis *et al.*, 1996)

i	1	2	3	4	5	6
P_i	500	500	69.53	500	64.70	349.48
X_i	110	90	90	11414		

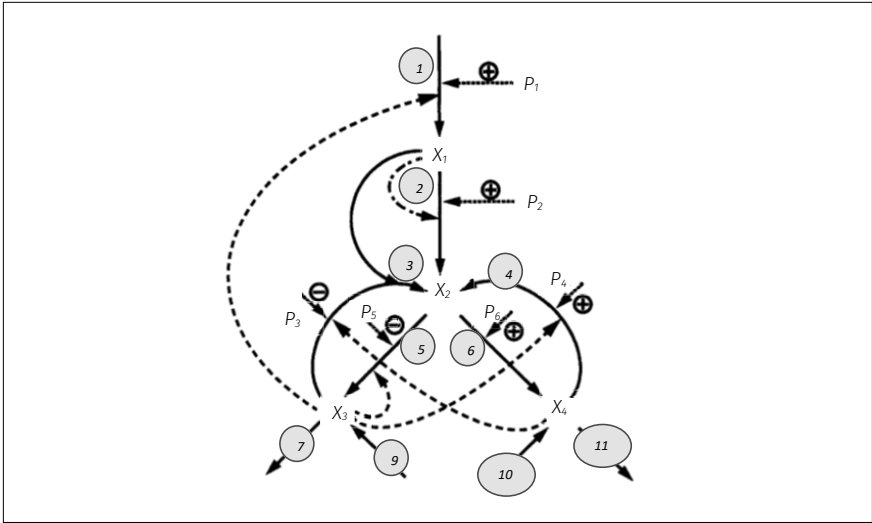


Figura 7.5. Cambios en la estructura reguladora y en los niveles de expresión de enzimas de la vía de síntesis de XMP y GMP que maximizan la concentración de XMP y GMP (X_4). Los signos en círculos junto a las líneas de puntos indicaron aumento (+) o disminución (-) en el nivel de expresión de las enzimas P_i correspondientes (adaptado de Hatzimanikatis *et al.*, 1996).

Se pueden abordar y responder muchas preguntas similares utilizando modelos como los presentados en este ejemplo. Un estudio de los valores de la concentración de metabolitos y de los niveles de actividad de las enzimas, así como de cualquier función metabólica de la red en los estados estacionarios óptimos puede proporcionar a los ingenieros metabólicos información útil antes de cualquier intento de manipular los niveles de enzimas y las arquitecturas reguladoras de la red metabólica.

IDENTIFICACIÓN ÓPTIMA DE KNOCK-OUT GENÉTICOS

La producción sostenible de combustibles y productos químicos ha sido una de las principales motivaciones para el uso de microorganismos como fábricas de productos químicos. El metabolismo microbiano, debido a presiones evolutivas, apunta a obtener un crecimiento rápido, lo cual a menudo está en conflicto directo con el objetivo del ingeniero metabólico de sobreproducción del metabolito de interés. Consecuentemente, la productividad de cepas de tipo salvaje suele estar por debajo de la productividad máxima teórica. Por lo tanto, la producción microbiana de una sustancia química a menudo requiere de ingeniería genética a través de interrupciones de flujos para redirigir el flujo metabólico hacia el producto de interés. El diseño de cepas microbianas es un desafío en biotecnología que se aplica a la producción de biocombustibles, productos de metabolismo secundario y medicamentos y dicho diseño usualmente requiere de múltiples intervenciones metabólicas que a menudo no son intuitivas. Por ejemplo, la sobreexpresión de genes ubicados en la ruta de un metabolito objetivo para lograr un rendimiento de producto satisfactorio a menudo es insuficiente. Es posible que se necesiten manipulaciones adicionales para dirigir el flujo de carbono hacia la ruta de producción y que sea necesario bloquear la producción de productos competitivos. La identificación de intervenciones y la cuantificación de su impacto en la generación de un metabolito requieren de una comprensión del metabolismo celular en todo el sistema. Esta tarea se complica porque los organismos han evolucionado para contrarrestar cualquier perturbación genética o ambiental impuesta externamente. Esto hace que el diseño de cepas estables sea en particular desafiante. En respuesta a estos desafíos, han surgido varios procedimientos computacionales de diseño de cepas que se benefician de los rápidos avances en el desarrollo y la reconstrucción de modelos metabólicos a escala genómica.

Formulación matemática

Los knockouts de genes se modelan con una variable binaria que decide si se debe eliminar una reacción j de la red:

$$y_j = \begin{cases} 0, & \text{si la reacción } j \text{ se debe eliminar} \\ 1, & \text{si la reacción } j \text{ no se debe eliminar} \end{cases} \quad \forall j \in J \quad (7.34)$$

En principio sería suficiente con resolver un problema MILP para identificar la eliminación de reacciones que conducen a la sobreproducción del metabolito de interés. El problema para obtener sobreproducción de un producto P en una cepa microbiana en un medio aeróbico con glucosa como fuente de carbono se puede formular de la siguiente manera:

$$\text{Maximize } z = v_{EX_P(e)} \quad (7.35.1)$$

Sujeto a:

$$\sum_{j \in J} S_{ij} v_j = 0 \quad (7.35.2)$$

$$v_{EX_glc(e)} \geq v_{glc}^{consumo} \quad (7.35.3)$$

$$v_{EX_O_2(e)} \geq v_{O_2}^{consumo} \quad (7.35.4)$$

$$v_{ATP} \geq v_{ATP}^{mantenimiento} \quad (7.35.5)$$

$$v_{biomasa} \geq f v_{biomasa}^{experimental} \quad (7.35.6)$$

$$LB_j y_j \leq v_j \leq UB_j y_j \quad \forall j \in J \quad (7.35.7)$$

$$\sum_{j \in J} (1 - y_j) \leq K \quad (7.35.8)$$

$$y_j \in \{0,1\}$$

donde I es el conjunto de los metabolitos en la red, v_j es el flujo de una reacción j , S_{ij} es el coeficiente estequiométrico del metabolito i en la reacción j , LB_j y UB_j son los límites inferior y superior en el flujo de la reacción j , $v_{EX_P(e)}$, $v_{EX_glc(e)}$ y $v_{EX_O_2(e)}$ denotan el flujo de las reacciones de intercambio para el producto de interés P , la glucosa y el oxígeno, respectivamente, y v_{ATPM} es el flujo correspondiente a la reacción ATP de mantenimiento microbiano no asociada al crecimiento.

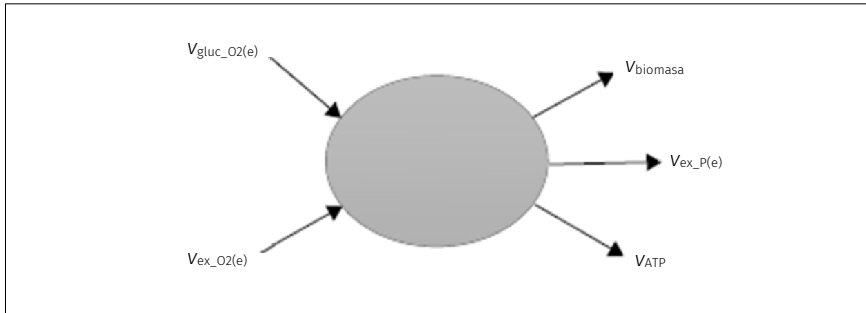


Figura 7.6. Fuentes y sumideros metabólicos en el diseño de cepas.

Además, $v_{biomasa}$ es el flujo de la reacción de biomasa y $v_{biomasa}^{max,WT}$ es el flujo de biomasa teórico máximo calculado para la cepa de tipo salvaje usando FBA. La restricción (7.35.2) representa el equilibrio de sistema en estado estacionario. Las siguientes dos restricciones establecen el consumo máximo permitido de glucosa y oxígeno respectivamente. La (7.35.5) restricción establece la demanda requerida de ATP al valor preespecificado, $v_{ATP-mantenimiento}$, de mantenimiento. Los valores para $v_{glc}^{consumo}$, $v_{O_2}^{consumo}$ y $v_{ATP}^{minimientot}$ para, por ejemplo, *Escherichia coli* se establecen en 10, 20 y 8,39 mmol de gDW/h, respectivamente. La ecuación (7.35.6) requiere que la formación de biomasa en la red alcance un umbral mínimo que se expresa como una fracción (típicamente $f = 0.1$) de la velocidad de crecimiento teórica máxima en la cepa de tipo salvaje. La ecuación (7.35.7) impone cotas a los flujos de reacción j y la ecuación (7.35.8) restringe el número total de intervenciones a un valor K especificado por el usuario. La limitación de esta formulación es que la inactivación de uno o más genes no tiene en cuenta el hecho de que el organismo probablemente reconfigurará los flujos en la red metabólica para restablecer el crecimiento en lugar de producir en exceso el producto deseado. En respuesta a esto, la función objetivo podría modificarse de la siguiente manera:

$$\text{Maximizar} \quad v_{EX_P(e)} + cv_{biomass} \quad (7.36)$$

donde c es un coeficiente positivo. Aún si se selecciona cuidadosamente el peso c , la presencia de una única función objetivo agregada no reconoce la existencia de dos impulsores de flujos separados y a menudo conflictivos. Por un lado, el organismo usa la maquinaria reguladora para restaurar y mantener el crecimiento en respuesta a una perturbación genética pero por otro lado, el ingeniero metabólico elimina ciertas reacciones, mediante desactivaciones genéticas, de la red metabólica para maximizar el flujo hacia el producto objetivo.

Optknock

El procedimiento *OptKnock* (Burgard *et al.*, 2003) reconoce que los organismos tienden a contrarrestar cualquier perturbación genética impuesta externamente a través de la redirección del flujo metabólico para restaurar el máximo crecimiento celular posible. Su objetivo es detectar eliminaciones de reacciones que remodelen la conectividad de la red de tal manera que la producción del metabolito objetivo se maximice mientras el organismo aún mantiene la

máxima velocidad de producción de biomasa bajo las restricciones impuestas por los genes inactivados. A menudo, esto fuerza a que la sobreproducción objetivo sea un subproducto obligatorio de la formación de biomasa lo cual se logra mediante el uso de procedimientos de optimización de dos niveles que involucran a dos funciones objetivos en competencia en lugar de utilizar una única función objetivo agregada. Los problemas de dos niveles son problemas de optimización anidados donde una o más restricciones se expresan implícitamente como un problema de optimización. El problema principal de optimización se conoce como el problema de nivel externo, mientras que los de nivel inferior se denominan problemas de nivel interno. El problema externo en *OptKnock* captura el objetivo de ingeniería metabólica (maximizar el flujo de producción del metabolito objetivo) mientras que el problema interno optimiza un objetivo de aptitud celular (maximización del flujo de producción de biomasa). A continuación, *OptKnock* se plantea como un problema de optimización de dos niveles para la sobreproducción de un metabolito objetivo en un medio aerobio con glucosa como fuente de carbono:

$$\text{Maximize } z = v_{biomasa} \quad (7.37.1)$$

Sujeto a:

$$\sum_{j \in J} S_{ij} v_j = 0 \quad (7.37.2)$$

$$v_{EX_glc(e)} \geq v_{glc}^{consumo} \quad (7.37.3)$$

$$v_{EX_O_2(e)} \geq v_{O_2}^{consumo} \quad (7.37.4)$$

$$v_{ATP} \geq v_{ATP}^{mantenimiento} \quad (7.37.5)$$

$$v_{biomasa} \geq v_{biomasa}^{experimental} \quad (7.37.6)$$

$$LB_j y_j \leq v_j \leq UB_j y_j \quad \forall j \in J \quad (7.37.7)$$

$$\sum_{j \in J} (1 - y_j) \leq K \quad (7.37.8)$$

$$y_j \in \{0,1\}$$

donde la variable binaria y_j indica la activación o desactivación del flujo j y las restricciones son las ya previamente definidas. El problema externo identifica candidatos de reacción para la eliminación que maximizan la producción del metabolito objetivo P , mientras que el problema interno redistribuye los flujos metabólicos para maximizar la formación de biomasa en la red perturbada sujeta a las eliminaciones de reacción impuestas por el problema externo. *OptKnock* sugiere nocaouts que intentan combinar el crecimiento de biomasa con la producción de la cantidad máxima posible del producto objetivo. Las variables binarias y_j son variables para el problema externo pero son parámetros del problema interno. Cualquier función objetivo convexa

alternativa como la minimización del ajuste metabólico (MOMA), la maximización del intercambio de ATP o la minimización del flujo total a través de todas las reacciones se pueden utilizar en el problema interno si se consideran como más representativos de la aptitud celular. Aunque las restricciones presentes en *OptKnock* son lineales, el problema de optimización de dos niveles no es lineal porque una de las restricciones está expresada implícitamente como un problema de optimización. Por lo tanto es necesario convertir esta estructura de dos niveles en un problema estándar de optimización de un solo nivel. Esto se puede lograr agregando las restricciones del problema interno mediante restricciones duales mientras se impone la condición fuerte de dualidad en el problema externo. Las variables duales asociadas con restricciones son las siguientes:

$$\text{Maximizar } z_p = v_{biomasa} \quad (7.38.1)$$

$$\text{sujeto a:} \quad \sum_{j \in J} S_{ij} v_j = 0 \quad \lambda_i \in R \quad (7.38.2)$$

$$-v_{EX_glc(e)} \leq -v_{glc}^{consumo} \quad \mu_{EX_glc}^{LB} \geq 0 \quad (7.38.3)$$

$$-v_{EX_O2(e)} \leq -v_{O2}^{consumo} \quad \mu_{EX_O2}^{LB} \geq 0 \quad (7.38.4)$$

$$-v_{ATP} \leq -v_{ATP}^{mantenimiento} \quad \mu_{ATPm}^{LB} \geq 0 \quad (7.38.5)$$

$$-v_{biomasa} \leq -v_{biomasa}^{maxima} \quad \mu_{biomasa}^{LB} \geq 0 \quad (7.38.6)$$

$$v_{vj} \leq UB_j y_j \quad \forall j \in J - ATP_{mantenimiento} \quad \mu_j^{UB} \geq 0 \quad (7.38.7)$$

$$-v_{vj} \leq UB_j y_j \quad \forall j \in J - \{EX_{glc}, EX_{O2}, ATPm, v_{biomasa}\} \quad \mu_j^{LB} \geq 0 \quad (7.38.8)$$

La representación MILP en un solo nivel es la siguiente:

$$\text{Maximizar } z_p = v_{EX_P} \quad (7.39.1)$$

$$\text{sujeto a:} \quad (7.39.1)$$

$$\sum_{j \in J} S_{ij} v_j = 0 \quad (7.39.2)$$

$$LB_j y_j \leq v_j \leq UB_j y_j \quad \forall j \in J \quad (7.39.3)$$

$$\sum_{i \in I} S_{ij} \lambda_i + \mu_j^{UB} - \mu_j^{LB} = 0 \quad \forall j \in J - v_{biomasa} \quad (7.39.4)$$

$$\sum_{i \in I} S_{ibiomasa} \lambda_i + \mu_{biomasa}^{UB} - \mu_{biomasa}^{LB} = 1 \quad (7.39.5)$$

$$0 \leq \mu_{biomasa}^{LB} \leq \mu_{biomasa}^{LB \max} \quad \forall j \in J \quad (7.39.6)$$

$$0 \leq \mu_{biomasa}^{UB} \leq \mu_{biomasa}^{UB\ max} \quad \forall j \in J \quad (7.39.7)$$

$$\mu_j^{LB} \leq \mu_j^{LB\ max}(1 - y_j) \quad \forall j \in J - v_{biomasa, ATM, EX_{glc}, EX_{O_2} \cup J^{irrev}} \quad (7.39.8)$$

$$\mu_j^{LB} \leq \mu_j^{LB\ max}(1 - y_j) \quad \forall j \in J - ATPM \quad (7.39.9)$$

$$\sum_{j \in J} (1 - y_j) \leq K \quad (7.39.10)$$

$$y_j \in \{0,1\}; \lambda_i, v_j \in R$$

Impacto de los knockouts en la formación de biomasa y la formación de producto

El balance entre las velocidades de formación de biomasa y de formación de producto objetivo se puede ilustrar gráficamente al trazar las soluciones de los problemas de optimización que maximizan y minimizan secuencialmente el flujo de intercambio del producto objetivo para niveles dados de producción de biomasa (Maranas y Zomorodi, 2016). La inactivación de genes conduce a restricciones más estrictas que implican espacios de solución más restringidos con respecto a los de cepas de tipo salvaje en dicho gráfico de compensación. Por lo tanto, el gráfico de compensación para un mutante estará completamente contenido dentro del de la cepa de tipo salvaje para los mismos escenarios. En general, se manifiestan dos posibles comportamientos ante la imposición de knockouts genéticos:

Escenario a: la formación máxima de biomasa y la formación del producto están desacopladas: en este caso, el flujo de formación del producto puede variar entre cero y un valor positivo cuando el flujo de biomasa es máximo, según se ilustra en la figura 7.5.a. Este no es un buen escenario para el diseño de cepas, ya que en presencia de knockouts impuestos, la cepa mutante puede escapar del requisito de sobreproducción.

Escenario b: la formación máxima de biomasa y de producto están acopladas: en este caso, el flujo mínimo de formación del producto es distinto de cero cuando la producción de biomasa se fija al valor máximo, tal como se ilustra en las figuras 7.8.b o 7.8.c. Idealmente, no hay variación en el flujo de formación del producto, lo que implica que el producto es un subproducto directo de la formación de biomasa. Estos escenarios son preferidos para el diseño de cepas ya que un nivel dado de crecimiento fija la velocidad de formación del producto, lo que implica que los esquemas de selección de crecimiento pueden usarse para aumentar la productividad del metabolito objetivo.

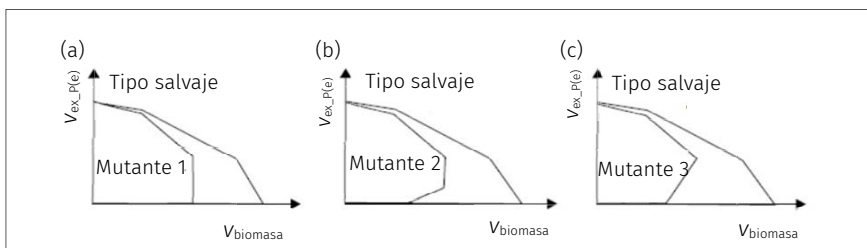


Figura 7.7. Balances entre formación de biomasa y formación de productos en cepas mutantes y de tipo salvaje. (a) La formación del producto se desacopla de la formación de biomasa en el Mutante 1. (b) La formación del producto se combina con la formación de biomasa en el Mutante 2, pero puede variar entre un valor mínimo y máximo positivo. (c) La formación del producto está completamente acoplada con la formación de biomasa en el Mutante 3.

RobustKnock

Una posible limitación de *OptKnock* surge cuando el mutante involucra un rango de rendimientos de formación de producto en condiciones de máxima velocidad de formación de biomasa (ver figuras 7.5.a y 7.5.b). Esto se debe a que *OptKnock* siempre selecciona la mayor velocidad de formación del producto bajo la condición de máxima velocidad de formación de biomasa aunque el flujo de formación del producto real puede ser menor que este máximo o incluso cero. Esto quiere decir que la formación del producto está desacoplada del crecimiento de la cantidad de biomasa. *RobustKnock* (Tepper & Shlomi, 2009) es una versión modificada de *OptKnock* que se desarrolló para abordar este problema optimizando al peor de los casos de formación del producto y mientras se maximiza la velocidad de formación de biomasa. Para este fin, la función objetivo del problema externo en *RobustKnock* se expresa como:

$$\text{Maximizar (Minimizar } v_{EX_P}) \quad (7.40)$$

Esto conlleva un problema de tres niveles:

$$\text{Maximizar}(\text{Minimizar } v_{EX_P}) \quad (7.41.1)$$

sujeto a:

$$\sum_{j \in J} S_{ij} v_j = 0 \quad \forall i \in I \quad (7.41.2)$$

$$LB_j y_j \leq v_j \leq UB_j y_j \quad \forall j \in J \quad (7.41.3)$$

$$\sum_{j \in J} (1 - y_j) \leq K \quad (7.41.4)$$

$$y_j \in \{0,1\}; \lambda_i, v_j \in R$$

El problema externo aquí es un problema de optimización de máximo–mínimo que busca identificar un conjunto de supresiones que maximicen un flujo de producción mínimo del metabolito objetivo. El problema de nivel interno es similar a *OptKnock* y maximiza los objetivos celulares sujetos a eliminaciones de reacción identificadas por el problema externo. Este problema de optimización se puede reformular como un problema MILP estándar en dos pasos mediante la transformación del problema en un problema estándar de máximo–mínimo reemplazando el problema de nivel interno con restricciones regulares y la posterior transformación del problema resultante de máximo–mínimo en un problema MILP estándar al reemplazar el problema de minimización interna con restricciones regulares. Ambos pasos implican reemplazar los problemas de optimización interna con restricciones primales y duales e imponer la condición de dualidad fuerte de manera similar a *OptKnock*. Esto implica un problema de optimización más complejo y generalmente más exigente desde el punto de vista computacional.

Inclinación de la función objetivo

Una alternativa más simple y más manejable computacionalmente implica el uso de una función objetivo «inclinada» (Feist *et al.*, 2010) para simular el peor de los casos para la formación del producto. En esencia, este método modifica la función objetivo interna en *OptKnock* para maximizar la formación de biomasa con una pequeña penalización debida a la formación del producto objetivo.

$$\text{Maximizar } v_{biomasa} - \varepsilon v_{EX_P} \quad (7.42)$$

Aquí ε es un peso pequeño que garantiza que, al maximizar la función objetivo, identifica la velocidad de formación de producto más pequeña posible de acuerdo con la maximización de la velocidad de formación de biomasa.

Minimización del ajuste metabólico

El método de minimización del ajuste metabólico (MOMA) se basa en la hipótesis de que la distribución del flujo metabólico en una cepa mutante sufre una redistribución mínima tras la perturbación con respecto a los flujos originales de una cepa de tipo salvaje (Segré *et al.*, 2002). Las distribuciones de flujo pronosticadas por MOMA tienen una concordancia generalmente mayor con los datos experimentales de flujos y velocidades de crecimiento de cepas mutantes de *Escherichia coli* que las predichas por FBA. MOMA se puede formular como un problema de programación cuadrática (QP), donde se minimiza la suma de las diferencias cuadráticas entre los flujos estacionarios a posteriori de una desactivación genética y los flujos de estado estacionario en la cepa de tipo salvaje.

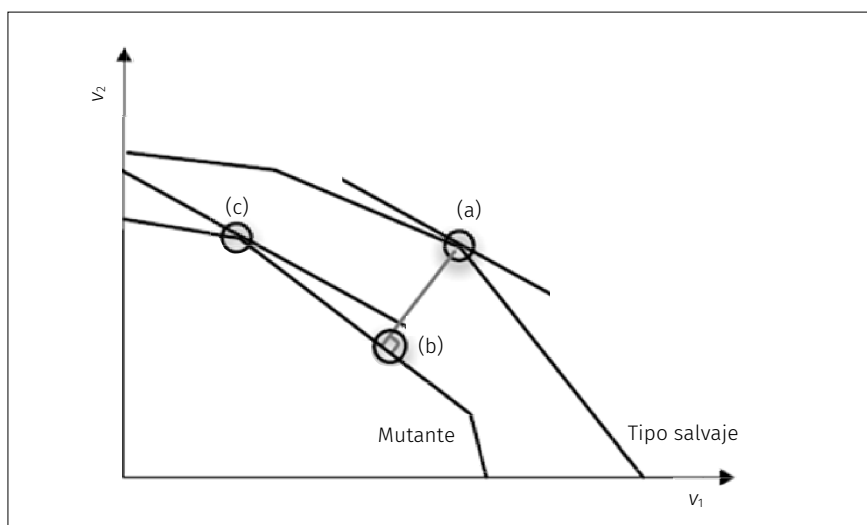


Figura 7.8. Comparación de FBA y MOMA para una red simple con solo dos reacciones. El área debajo de las líneas representa las regiones factibles para una cepa de tipo salvaje y una cepa mutante. Los puntos A y B denotan la solución de máxima producción de biomasa según FBA para la cepa mutante y de tipo salvaje, respectivamente. El punto C, derivado de MOMA, es la proyección del punto A sobre la región factible de la cepa mutante.

Las restricciones del problema son las mismas que las del FBA, lo cual implica la siguiente formulación:

$$\text{Minimizar } \sum_{j \in J} (v_j - v_j^{WT})^2 \quad (7.43.1)$$

sujeto a:

$$\sum_{j \in J} S_{ij} v_j = 0 \quad \forall i \in I \quad (7.43.2)$$

$$LB_j y_j \leq v_j \leq UB_j y_j \quad \forall j \in J \quad (7.43.3)$$

Aquí, I y J denotan a los conjuntos de metabolitos y reacciones, respectivamente. Además, v_j y v_j^{WT} son los flujos de reacción j en las cepas mutantes y de tipo salvaje, respectivamente. MOMA puede verse como una extensión de FBA ya que la mayoría de los valores de flujo permanecen muy cerca de los valores calculados para la cepa de tipo salvaje usando FBA. La primera restricción representa el balance de materia en estado estacionario para cada metabolito y la segunda ecuación impone límites inferior y superior en los flujos de reacción.

Incorporación de expresiones cinéticas en modelos estequiométricos

Los modelos metabólicos estequiométricos no pueden capturar cuantitativamente el efecto de los niveles de concentración y de actividad de las enzimas. Esto puede conducir a predicciones de distribución de flujo que son inconsistentes con los rangos permitidos para las concentraciones intracelulares de metabolitos, con las actividades enzimáticas y con la regulación a nivel de sustratos. Los modelos cinéticos del metabolismo describen la evolución temporal de las concentraciones de metabolitos, las actividades enzimáticas y los flujos de reacción utilizando un sistema de ecuaciones diferenciales ordinarias. Sin embargo, en contraste con los modelos estequiométricos, la reconstrucción de modelos cinéticos a gran escala se ve obstaculizada por la escasez de datos experimentales sobre flujos y concentraciones lo cual impide la parametrización del modelo. Los modelos cinéticos están disponibles actualmente para una pequeña lista de organismos y abarcan solo una fracción del metabolismo conocido. Se han realizado esfuerzos en los últimos años para mejorar la predicción de fenotipos metabólicos utilizando modelos estequiométricos a escala genómica complementados con información cinética. En estos modelos híbridos, el flujo de reacciones con información cinética disponible se identifica resolviendo el sistema de ecuaciones diferenciales que representan los equilibrios de componentes y

FBA se utiliza para inferir el flujo para el resto de las reacciones en la red. El conjunto de reacciones J en la red se divide en dos subconjuntos: reacciones con información cinética, J^{kin} , y reacciones sobre las cuales solo se tiene información estequiométrica, J^{stoic} . Para las reacciones pertenecientes a J^{kin} , se permite que la velocidad de reacción enzimática máxima v_j^{max} varíe entre cero (lo que denota la eliminación de la actividad de la reacción) a una velocidad máxima permisible por la actividad de la enzima en una cepa tipo salvaje $v_j^{max, ref}$. El siguiente programa no lineal permite identificar el flujo máximo para un producto P :

$$\begin{aligned}
 & \text{Maximizar } z = v_{EX_P} & (7.44.1) \\
 & \text{sujeto a:} \\
 & \quad v_j = v_j(v_j^{max}, C_i, P_j) & (7.44.2) \\
 & \quad 0 \leq v_j^{max} \leq f v_j^{max, ref} \quad \forall j \in J & (7.44.3) \\
 & \quad C_i^{LB} \leq C_i \leq C_i^{UB} \quad \forall i \in I & (7.44.4) \\
 & \quad \sum_{j \in J} S_{ij} v_j = 0 & (7.44.5) \\
 & \quad LB_j \leq v_j \leq UB_j \quad \forall j \in J^{stoic} & (7.44.6) \\
 & \quad C_i \geq 0, v_j \in R
 \end{aligned}$$

Aquí, v_{EX_P} (e) es el flujo de intercambio del producto objetivo P , C_i es la concentración de un metabolito i en la parte cinética del modelo (J^{kin}), C_i^{LB} y C_i^{UB} denotan los límites inferior y superior en la concentración del metabolito i . P_j es el conjunto de parámetros cinéticos para la reacción $j \in J^{kin}$. La función objetivo de este problema maximiza el flujo de intercambio de un metabolito objetivo en forma consistente con la cinética de reacción y con la estequiometría. La primera restricción determina las velocidades de reacción utilizando las ecuaciones cinéticas para reacciones en J^{kin} . La segunda restricción impone un rango permitido de manipulaciones enzimáticas en la parte cinética de la red, mientras que la tercera restricción impone límites inferior y superior en la concentración de metabolitos en J^{kin} . Finalmente, las últimas dos restricciones imponen la condición de equilibrio en estado estacionario y los límites inferior y superior para los flujos de reacción computados mediante FBA. Obsérvese que, si bien los flujos de reacciones en J^{stoic} están limitados solo por el equilibrio en estado estacionario, los límites los de J^{kin} están determinados por las expresiones cinéticas. Las expresiones cinéticas son típicamente funciones no lineales de las concentraciones de metabolitos.

Actividades

1. Utilice *OptKnock* para identificar dos genes a noquear con el fin de forzar la sobreproducción de acetato en *Escherichia coli* utilizando el modelo metabólico iAF1260 bajo condición aeróbica en un medio con glucosa como fuente de carbono. Utilice un flujo de 10 mmol gDW h de glucosa como base y el 10 % de la máxima velocidad teórica de crecimiento microbiano como el nivel mínimo requerido de formación de biomasa en la red. Determine si la producción de acetato está acoplada con la formación de biomasa en la cepa mutante.
2. Formule y resuelva *OptKnock* para identificar eliminaciones de reacción que conducen a la sobreproducción de succinato utilizando el modelo iAF1260 de *Escherichia coli* bajo las mismas condiciones de absorción que en el punto anterior pero asumiendo condiciones anaeróbicas.
3. Identifique las mejores combinaciones de eliminación de reacción, incluidas estrategias de eliminación de reacción alternativas, si están disponibles.
4. Vincule las eliminaciones de las reacciones identificadas con los knockouts de genes correspondientes utilizando las asociaciones proporcionadas en el modelo.
5. Comente sobre su viabilidad y compárelas con la evidencia experimental de la literatura.
6. Genere el gráfico de compensación de succinato vs. biomasa para todos los mutantes y elija un diseño que considere más prometedor.

REFERENCIAS BIBLIOGRÁFICAS

- HATZIMANIKATIS, V.; FLOUDAS, C. & BAILEY, J. (1996). Analysis and design of metabolic reaction networks via mixed-integer linear optimization. *AIChE Journal*, 42 (5), 1277–1292. <https://doi.org/10.1002/aic.690420509>
- BURGARD, A.; PHARKYA, P. & MARANAS, C. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng.*, 84 (6): 647–657. <https://DOI:10.1002/bit.10803>
- SEGRÈ, D.; VITKUP, D. & CHURCH, G. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci USA*, 99 (23): 15112–15117. <https://DOI:10.1073/pnas.232349399>
- TEPPER, N. & SHLOMI, T. (2010). Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. *Bioinformatics*, 26 (4): 536–543. <https://doi.org/10.1093/bioinformatics/btp704>

- FEIST, A.; ZIELINSKI, D. (...) PALSSON, B.** (2010). Model driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metab Eng*, 12 (3): 173–186. [https://DOI:10.1016/j.ymben.2009.10.003](https://doi.org/10.1016/j.ymben.2009.10.003)
- HATZIMANIKATIS, V.; FLOUDAS, C. & BAILEY, J.** (1996). Optimization of regulatory architectures in metabolic reaction networks. *Biotechnol Bioeng.*, 52 (4): 485–500. [https://DOI:10.1002/\(SICI\)1097-0290\(19961120\)52:4<485::AID-BIT4>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1097-0290(19961120)52:4<485::AID-BIT4>3.0.CO;2-L)
- VOIT, E.** (1992). Optimization in integrated biochemical systems *Biotechnol. Bioeng.* 40, 5, 572–582. <https://doi.org/10.1002/bit.260400504>
- MARANAS, C. D. & ZOMORRODI, A. R.** (2016). *Optimization Methods in Metabolic Networks*. John Wiley & Sons.

Lecturas adicionales

- RAMAN, K.** (2021). *An Introduction to Computational Systems Biology: Systems-Level Modelling of Cellular Networks*. Chapman & Hall/CRC Computational Biology Series.
- TORRES, N. & VOIT, E.** (2002). *Pathway Analysis and Optimization in Metabolic Engineering*. Cambridge University Press.
- MARANAS, C. D. & ZOMORRODI, A. R.** (2016). *Optimization Methods in Metabolic Networks*. John Wiley & Sons.

8 Optimización dinámica de sistemas macroscópicos

Consideremos un sistema biológico o biotecnológico cuya dinámica está modelada por un sistema de ecuaciones diferenciales ordinarias (EDOs) o por un sistema de ecuaciones en diferencias discretas y supongamos que dicho sistema tiene variables que pueden ser manipuladas. Es natural cuestionarse cómo controlar a dichas variables para obtener la «mejor solución» en términos de un objetivo predeterminado. Para ello presentamos una serie de aplicaciones de la teoría de control óptimo, cuyos fundamentos ya fueron expuestos en la sección 2.5. Estas aplicaciones se presentan con un creciente grado de complejidad a medida que se progresa en su presentación. El objetivo de este capítulo es presentar la utilidad de la teoría de control óptimo en aplicaciones biológicas y biotecnológicas.

Con frecuencia, en procesos biotecnológicos que operan en modo batch, los substratos se cargan inicialmente en el reactor y el mismo se opera de acuerdo con un curso de acción predeterminado durante el cual no se adicionan ni remueven componentes del sistema, excepto el intercambio de O_2 y CO_2 entre las fases líquida y gaseosa del sistema. En operaciones de tipo semi-batch se pueden adicionar substratos o remover cierto volumen de cultivo en forma secuencial o simultánea. La principal diferencia de ambos modos de operación con respecto del modo de operación continuo es que no existe un estado estacionario de operación. En ambos casos, determinar el curso óptimo de acción requiere de la aplicación de conceptos de control óptimo. Si bien los modelos no estructurados presentan limitaciones predictivas en las condiciones antes mencionadas, se han utilizado con éxito en varias aplicaciones, algunas de las cuales se presentan a continuación.

MAXIMIZACIÓN DE LA PRODUCCIÓN DE BIOMASA EN UN REACTOR DE TIPO FED BATCH

Consideremos un ejemplo relativamente simple que utiliza a un sistema de tres EDOs en el que se busca maximizar la cantidad de biomasa que se obtiene al tiempo final del proceso. En este caso el problema se plantea de la siguiente manera:

$$J = X(t_f)V(t_f) \quad (8.1)$$

sujeto a las restricciones dadas por el modelo dinámico del modo de operación *fed batch*:

$$\frac{dX}{dt} = \mu X - u \frac{X}{V} \quad (8.2.1)$$

$$\frac{dS}{dt} = -q_S X + u \frac{S_F - S}{V} \quad (8.2.2)$$

$$\frac{dV}{dt} = \quad (8.2.3)$$

donde X es la concentración de biomasa, S del sustrato limitante, V es el volumen de cultivo y t_f es el tiempo final de la misma. Los parámetros μ y q_S son las velocidades específicas de crecimiento y de producción del metabolito de interés mientras que el flujo u de la solución con sustrato limitante cuya concentración es S_F . Consideremos el ejemplo presentado por (Levišauskas and Staniškis, 1991) en el que las condiciones iniciales son las siguientes:

$$\begin{bmatrix} X \\ S \\ V \end{bmatrix}_{t_0} = \begin{bmatrix} 1.0 \text{ g/l} \\ 16.7 \text{ g/l} \\ 5.0 \text{ l} \end{bmatrix} \quad (8.3)$$

Además la concentración de la solución de sustrato en la corriente de alimentación es $S_F = 200 \text{ g/l}$, el tiempo de fermentación es una variable de diseño, el volumen máximo de fermentación es $V(t_f) = 6.0 \text{ l}$ y la dependencia de los parámetros cinéticos μ y q_S respecto de las variables de estado viene expresada por las siguientes ecuaciones:

$$\mu = 0.713S - 0.069X \text{ [h}^{-1}\text{]} \quad (8.4.1)$$

$$q_S = 0.157 - 0.0107 + 0.0163S^2 \text{ [g S/g X h]} \quad (8.4.2)$$

Las restricciones de diseño para la variable controlada son:

$$0 \leq u \left[\frac{\text{l}}{\text{h}} \right] \leq 1.5 \quad (8.5)$$

La solución para este problema se ilustra en la figura 8.1. en ella se grafica la política óptima de alimentación y la consecuente evolución de las variables de estado.

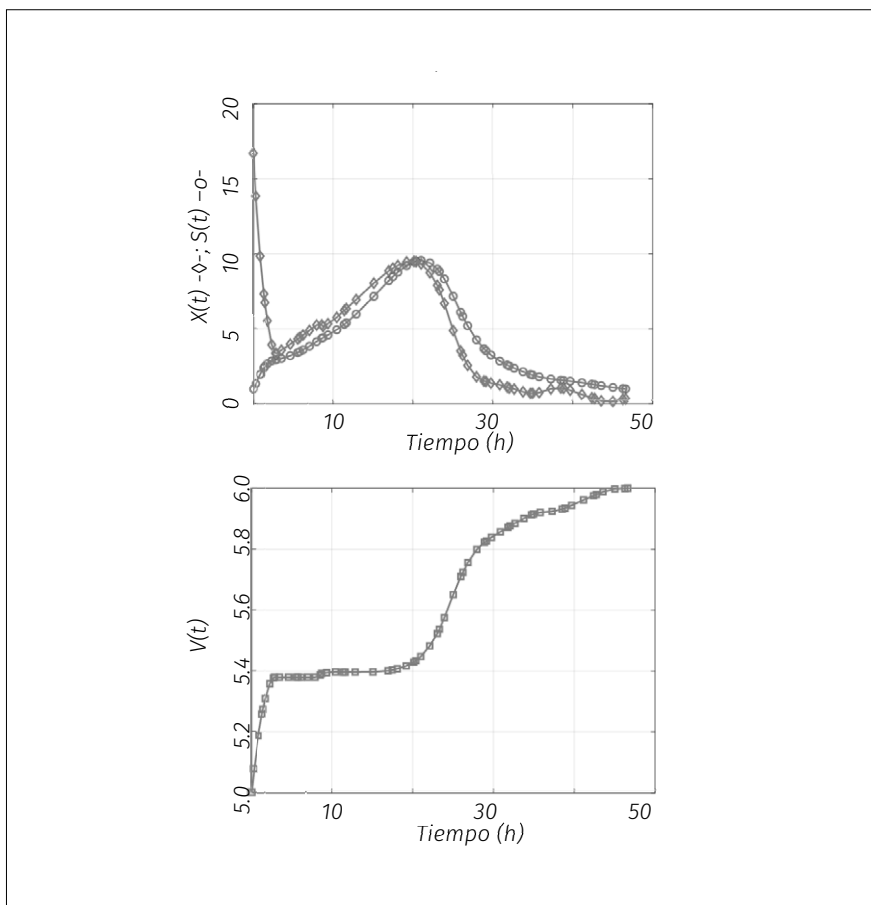


Figura 8.1. Perfiles temporales para la velocidad de adición de sustrato, de la concentración de biomasa y sustrato limitante y del volumen de cultivo.

OPTIMIZACIÓN DE LA PRODUCCIÓN DE UN METABOLITO DIFERENTE DE LA BIOMASA

En numerosos casos la biomasa no es el producto de interés, sino que actúa como «fábrica» de dicho producto. Hay que mencionar que las condiciones que implican un máximo crecimiento celular suelen ser diferentes a las de máxima productividad del metabolito deseado por lo que la determinación heurística de la política óptima de control no es sencilla y la mejor opción es acudir a la optimización vía modelos. Este tipo de problemas es de un nivel de complejidad mayor al anterior pues debe considerarse una ecuación diferencial adicional que involucra a la variación de la concentración del producto.

Consideremos el siguiente ejemplo de optimización de la producción de etanol en un reactor de tipo *fed batch*. El objetivo del problema suele ser maximización de la cantidad de etanol producido mediante, por ejemplo, *S. cerevisiae* utilizando a la velocidad de adición de sustrato como variable de control sin un tiempo de fermentación especificado lo que implica un problema de tiempo final libre. La formulación matemática del problema es la siguiente:

$$\begin{aligned} & \text{Maximizar} \\ & J = P(t_f)V(t_f) \end{aligned} \quad (8.6.1)$$

sujeto a las restricciones dadas por el modelo dinámico de modo de operación *fed batch* con las cinéticas correspondientes al crecimiento microbiano, el consumo de sustrato limitante y la producción del metabolito de interés, además de la variación del volumen del cultivo:

$$\frac{dX}{dt} = 0.408 \frac{S}{0.22 + S_1 + P/16} X - u \frac{X}{V} \quad (8.6.2)$$

$$\frac{dS}{dt} = -10 \frac{S}{0.22 + S_1 + P/16} X + u \frac{150 - S}{V} \quad (8.6.3)$$

$$\frac{dP}{dt} = \frac{S}{0.44 + S_1 + P/71.5} X - u \frac{P}{V} \quad (8.6.4)$$

$$\frac{dV}{dt} = u \quad (8.6.5)$$

Las cinéticas corresponden al ejemplo presentado en (Chen & Hwang, 1990) y en general este tipo de cinética depende de la cepa empleada y de los efectos de sustratos y metabolitos. Obsérvese que en este ejemplo existe inhibición de crecimiento por presencia del metabolito de interés.

Para encontrar el perfil de la velocidad de adición de la solución rica en sustrato limitante sustrato $u(t)$ y el tiempo de fermentación t_f , para el intervalo de tiempo $[t_0, t_f]$ se resolvió el problema utilizando GPOPS y considerando las siguientes condiciones iniciales:

$$\begin{bmatrix} X \\ S \\ P \\ V \end{bmatrix}_{t_0} = \begin{bmatrix} 1.0 \text{ g/l} \\ 150 \text{ g/l} \\ 0 \text{ g/l} \\ 10 \text{ l} \end{bmatrix} \quad (8.7)$$

Mientras que la variable de control debe estar restringida por una cota superior y otra inferior:

$$0 \leq u \left(\frac{l}{h} \right) \leq 1 \quad (8.8)$$

Finalmente, el volumen máximo permitido es $V(t_f) = 200 \text{ l}$. La solución a este problema se ilustra en los siguientes gráficos.

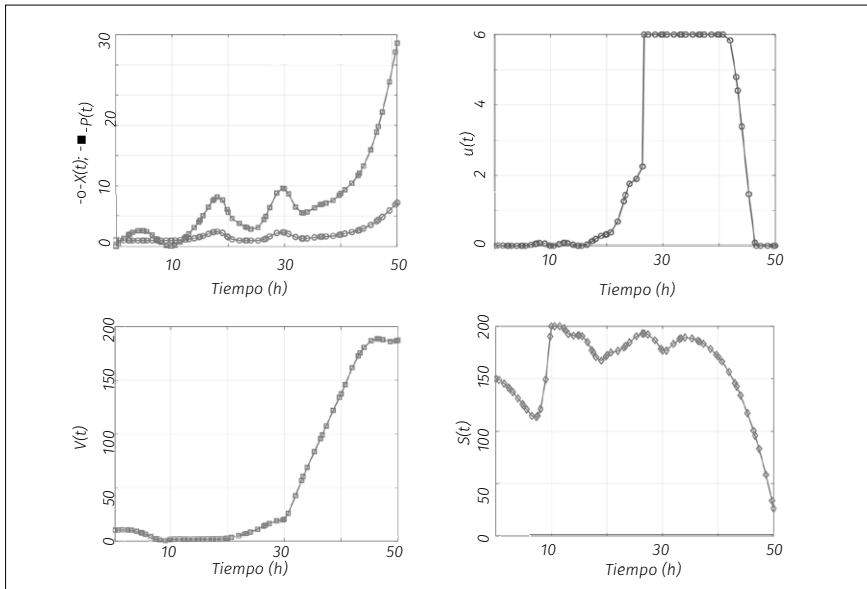


Figura 8.2. Perfil óptimo de la velocidad de adición de sustrato para la producción de etanol y evolución temporal óptima de las variables de estado.

OPTIMIZACIÓN DE LA PRODUCCIÓN DE UN METABOLITO EN PRESENCIA DE MÚLTIPLES SUSTRATOS LIMITANTES

Este tipo de problemas involucra la maximización de la producción de un metabolito y emplea un sistema de ecuaciones cinéticas relativamente más complejo que en los ejemplos anteriores. Consideremos al siguiente ejemplo (Levišauskas & Staniškis, 1991) en el que se ilustra la maximización de la producción de un biosurfactante, en el cual el sistema el de EDO₅ involucra a seis estados y el crecimiento celular se ve afectado en una forma bastante compleja por la presencia de biomasa y sustratos. La variable de control es el flujo de la solución adicionada al reactor. Dicha solución contiene cierta cantidad de sustratos esenciales (S_1 , glucosa; S_2 , nitrato de amonio; y S_3 , fosfatos). El modelo matemático tiene la siguiente formulación:

Maximizar

$$J = P(t_f)V(t_f) \quad (8.9.1)$$

sujeto a:

$$\frac{dX}{dt} = \mu X - u \frac{X}{V} \quad (8.9.2)$$

$$\frac{dS_1}{dt} = -q_{S1}X + u \frac{S_{F1} - S_1}{V} \quad (8.9.3)$$

$$\frac{dS_2}{dt} = -q_{S2}X + u \frac{S_{F2} - S_2}{V} \quad (8.9.4)$$

$$\frac{dS_3}{dt} = -q_{S3}X + u \frac{S_{F3} - S_3}{V} \quad (8.9.5)$$

$$\frac{dP}{dt} = q_P X - u \frac{P}{V} \quad (8.9.6)$$

$$\frac{dV}{dt} = u \quad (8.9.7)$$

donde X y P son la concentración de biomasa y de biosurfactante respectivamente. Las variables S_1 , S_2 y S_3 representan las concentraciones de glucosa, nitrato de amonio y fosfato respectivamente. Las ecuaciones diferenciales (8.9.2)–(8.9.6) representan a la dinámica de la concentración de la biomasa,

la glucosa, el nitrato de amonio, las sales de fosfato y el producto respectivamente, siendo los coeficientes μ la velocidad específica de crecimiento microbiano, q_{S1} , q_{S2} y q_{S3} respectivamente las velocidades específicas de consumo de glucosa, nitrato de amonio, y sales de fosfato y siendo q_P la velocidad específica de producción de biosurfactante. La ecuación (8.9.7) representa a la variación del volumen de cultivo en función del flujo de la corriente de entrada al biorreactor $u(t)$. Para la resolución numérica del presente ejemplo consideremos a las siguientes condiciones iniciales:

$$\begin{bmatrix} X \\ S_1 \\ S_2 \\ S_3 \\ P \\ V \end{bmatrix}_{t_0} = \begin{bmatrix} 1.4 \text{ g/l} \\ 22.5 \text{ g/l} \\ 0.616 \text{ g/l} \\ 0.0292 \text{ g/l} \\ 0.08 \text{ g/l} \\ 3 \text{ l} \end{bmatrix} \quad (8.10)$$

También se considera que las concentraciones de los sustratos esenciales adicionados por medio de la corriente de entrada son $S_{1F} = 200 \text{ g/l}$, $S_{2F} = 3.42 \text{ g/l}$ y $S_{3F} = 1.14 \text{ g/l}$. El tiempo de fermentación es $t_f = 11 \text{ h}$, el volumen máximo de fermentación es libre y la dependencia de los parámetros cinéticos con respecto de las variables de estado viene expresada por las siguientes ecuaciones:

$$\mu = 1.388 \frac{S_1}{3.2337X + S_1 + \frac{S_1^2}{24.1029}} - 0.0219X \quad (8.11.1)$$

$$q_{S1} = 0.1492 \mu \quad (8.11.2)$$

$$q_{S2} = 0.0032 + 0.00079 \mu + 0.00022 q_P \quad (8.11.3)$$

$$q_{S3} = 0.0012 + 0.004 \mu \quad (8.11.4)$$

$$q_P = 0.0014 \quad (8.11.5)$$

La restricción de diseño para la variable de control es:

$$0 \leq u \left[\frac{l}{h} \right] \leq 0.5 \quad (8.12)$$

Las siguientes figuras ilustran soluciones a dicho problema.

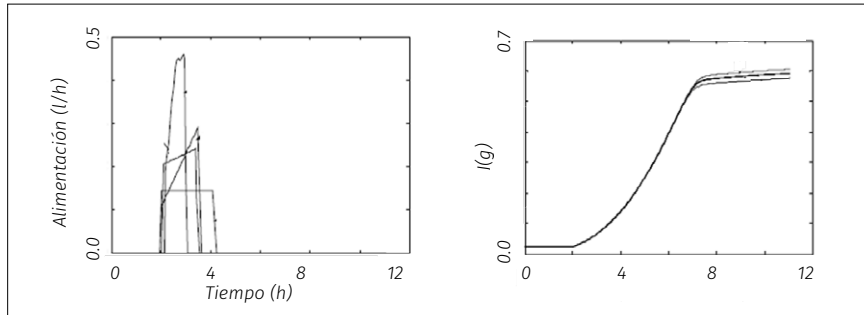


Figura 8.3. Perfil óptimo de la velocidad de adición de la solución con sustratos y evolución temporal óptima de las variables de estado.

Nótese que, por problemas de convergencia e inestabilidad numérica, se obtienen diferentes perfiles de control que conllevan a una evolución diferenciada del valor de la función objetivo.

OPTIMIZACIÓN DE UN CULTIVO CON MANIPULACIÓN DE VARIAS VARIABLES DE CONTROL

La obtención de productos de alto valor por medio de microorganismos modificados genéticamente está convirtiéndose en práctica común en numerosas empresas farmacéuticas. Los procesos de este tipo necesitan de modelos fieles que por lo general suelen ser de una complejidad considerable e involucran a la manipulación de más de una variable controlada. Para ilustrar este tipo de procesos consideremos al ejemplo, que fue propuesto por Lee y Ramirez (1994) que involucra la determinación de trayectorias óptimas para dos variables de control en un proceso *fed batch* cuyo objetivo es la producción de una proteína por medio de una bacteria recombinante. El objetivo es la maximización de la ganancia del proceso y las variables a manipular en forma óptima son la velocidad de adición del nutriente principal (u_1) y de adición de un inductor de excreción específico (u_2). La formulación matemática del proceso es la siguiente:

Maximizar

$$J = P(t_f)V(t_f) \quad (8.13.1)$$

sujeto a las restricciones dadas por el sistema de ecuaciones diferenciales que modelan al proceso:

$$\frac{dV}{dt} = u_1 + u_2 \quad (8.13.2)$$

$$\frac{dX_1}{dt} = g_1 X - (u_1 + u_2) \frac{X}{V} \quad (8.13.3)$$

$$\frac{dS}{dt} = 100 \frac{u_1}{V} - (u_1 + u_2) \frac{S}{V} - g_1 \frac{X}{0.51} \quad (8.13.4)$$

$$\frac{dP}{dt} = g_2 X - (u_1 + u_2) \frac{P}{V} \quad (8.13.5)$$

$$\frac{dI_1}{dt} = 4 \frac{u_2}{V} - (u_1 + u_2) \frac{I_1}{V} \quad (8.13.6)$$

$$\frac{dI_2}{dt} = -0.09 \left(\frac{I_1}{0.034 + I_1} \right) I_2 \quad (8.13.7)$$

$$\frac{dI_3}{dt} = -0.09 \left(\frac{I_1}{0.034 + I_1} \right) (1 - I_2) \quad (8.13.8)$$

Donde g_1 y g_2 son las siguientes funciones:

$$g_1 = \left(\frac{S}{14.35 + S + \frac{S^2}{111.5}} \right) \left(I_1 + I_2 + \frac{0.22}{0.22 + I_1} \right) \quad (8.14.1)$$

$$g_2 = \left(\frac{S}{14.35 + S + \frac{S^2}{111.5}} \right) \left(\frac{0.0005 + I_1}{0.022 + I_1} \right) \quad (8.14.2)$$

Las variables de estado del sistema están especificadas por el siguiente vector de estados:

$\begin{bmatrix} V \\ X \\ S \\ P \\ I_1 \\ I_2 \\ I_3 \end{bmatrix}$	$=$	$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}$	Volumen de fermentación Concentración celular Concentración del nutriente Concentración de la proteína recombinante Concentración del inductor Factor de shock del inductor en el crecimiento celular Factor de recuperación del crecimiento celular por el inductor
---	-----	---	---

Mientras que el vector de las variables de control es:

$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$	Velocidad de adición del nutriente Velocidad de adición del inductor
--	--

La función objetivo maximiza la ganancia de la fermentación, que viene expresada por la diferencia entre el valor del producto obtenido ($P(t_f)V(t_f)$) y el costo del inductor consumido en la fermentación ($Q \int u_2 dt$), donde Q es el cociente entre el costo de una unidad de inductor y el valor unitario de la proteína excretada.

Para la resolución numérica consideramos un horizonte de fermentación definido por el intervalo $[t_o = 0 \text{ h}; t_f = 10 \text{ h}]$ con las siguientes condiciones iniciales:

$$\begin{bmatrix} V \\ X \\ S \\ P \\ I_1 \\ I_2 \\ I_3 \end{bmatrix}_{t_0} = \begin{bmatrix} 1 \text{ l} \\ 0.1 \text{ g/l} \\ 40 \text{ g/l} \\ 0 \text{ g/l} \\ 0 \text{ g/l} \\ 1 \text{ g/l} \\ 0 \text{ g/l} \end{bmatrix} \quad (8.15)$$

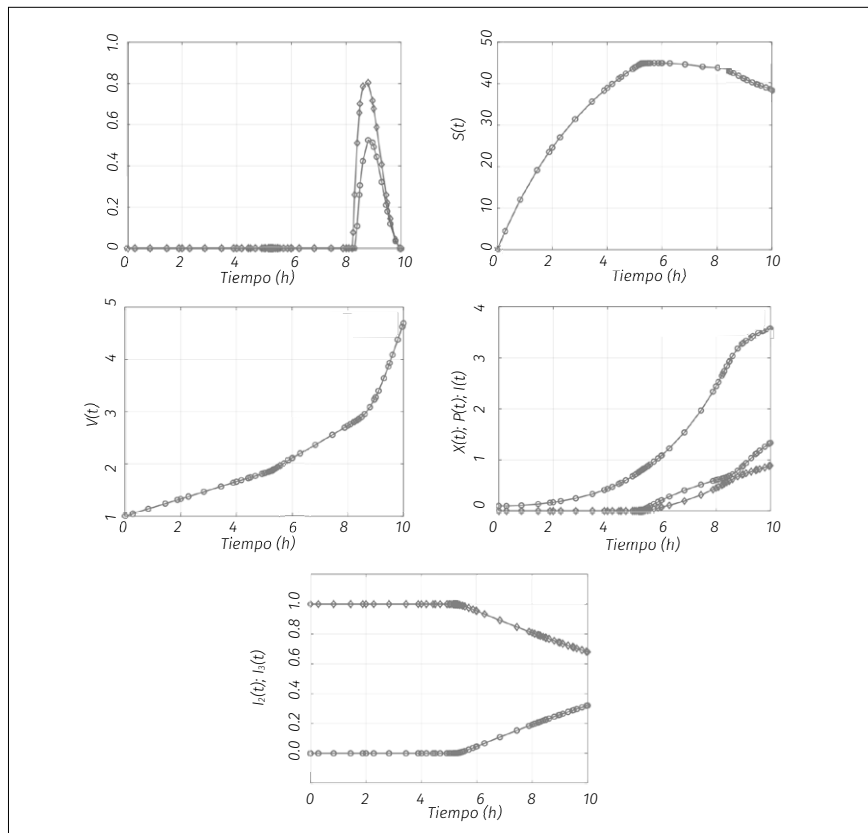


Figura 8.4. Velocidad óptima de adición de sustrato e inductor y la correspondiente evolución óptima de las variables de estado.

Además se consideran las siguientes restricciones en el valor de las variables controladas:

$$0 \leq u_1 \left(\frac{l}{h} \right) \leq 1 \quad (8.16.1)$$

$$0 \leq u_2 \left(\frac{l}{h} \right) \leq 1 \quad (8.16.2)$$

La solución obtenida, mediante el uso de GPOPS, para este problema se muestra en los gráficos que ilustran la política óptima de adición de sustrato e inductor y sus efectos en las variables de estado.

OPTIMIZACIÓN DE SISTEMAS MODELADOS MEDIANTE REDES METABÓLICAS

En la producción de lípidos para biocombustibles a partir de microalgas, se debe optimizar la composición del cultivo cosechado para obtener la cantidad máxima factible de triglicéridos (TAG). Bioprocesos complejos como este necesitan de modelos basados en redes metabólicas que, por ser de naturaleza dinámica, deben ser de escala reducida e involucran a un alto número de variables de estado y variables de control. Estos son, desde el punto de vista del control óptimo, problemas extremadamente complejos (Marquez *et al.*, 2024).

En el desarrollo de un modelo para optimizar la recolección de lípidos de algas sujetas a un ciclo circadiano, el modelo cinético–metabólico del metabolismo del alga debe integrarse en equilibrios macroscópicos correspondientes, en este ejemplo, al modo de operación de un estanque de canal abierto (ORP). Además se debe definir el objetivo, los estados y las variables de control. El modo de funcionamiento continuo suele considerarse como uno de los más eficientes, pero no es adecuado para sincronizarse con reacciones autotróficas alineadas con el ciclo circadiano. La operación cíclica de alimentación a un reactor de tipo batch puede sincronizar las condiciones de cultivo y las variaciones de volumen con los ciclos día–noche. Debido a la alimentación de una solución con sustratos el volumen de cultivo aumenta y, si es posible controlar el flujo de salida y retener una porción de un cultivo al final de un ciclo para usar la parte residual como inóculo para el siguiente ciclo, este modo de operación se denomina reactor continuo alimentado por lotes repetidos o de volumen variable —vvcr— (Asenjo y Merchuk, 1995). De este modo se puede sincronizar las variaciones de volumen con el ciclo

circadiano mediante el control de los flujos de entrada y salida. Además, la falta de una fuente de nitrógeno permite redireccionar la energía de la fotosíntesis en la red metabólica hacia moléculas que no contienen nitrógeno como lípidos e hidratos de carbono. Los niveles de nitrógeno se pueden controlar manipulando la concentración de la fuente de nitrógeno en el flujo de entrada. En resumen, el modelo macroscópico correspondiente a un VVCR puede describirse mediante el siguiente sistema de EDOs:

$$\frac{d}{dt} \begin{pmatrix} NO_3 \\ CARB \\ TAG \\ PEP \\ G6P \\ GAP \\ B \end{pmatrix} = \begin{pmatrix} r_{NO_3} \\ r_{CARB} \\ r_{TAG} \\ r_{PEP} \\ r_{G6P} \\ r_{GAP} \\ r_B \end{pmatrix} B - \frac{F_i}{V} \begin{pmatrix} NO_3 \\ CARB \\ TAG \\ PEP \\ G6P \\ GAP \\ B \end{pmatrix} + \frac{F_i}{V} \begin{pmatrix} NO_3^{in} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (8.17.1)$$

La dinámica del volumen de cultivo viene dada por:

$$\frac{dV}{dt} = F_i - F_o \quad (8.17.2)$$

En el modelo dado por estas ecuaciones, *CARB* (carbohidratos), *TAG* (triglicéridos), *PEP* (glucosa-6-fosfato), *G6P* (glucosa-6-fosfato) y *GAP* (gliceraldeído-3-fosfato) son las concentraciones macroscópicas de compuestos intracelulares; *B* es la concentración macroscópica de biomasa funcional, NO_3/NO_3^{in} son la concentración de nitratos en el cultivo y flujo de alimentación respectivamente y F_i/F_o son los flujos entrada y salida al cultivo. Para el caso en el que haya excreción de *PEP* y que este fenómeno no tenga efecto en la cinética intracelular, el balance de *PEP* estará dado por la siguiente ecuación:

$$\frac{dPEP_{excr}}{dt} = r_{excr}B - \frac{F_i}{V}PEP_{excr} \quad (8.17.3)$$

El vector $\mathbf{r} = [r_{NO_3}, r_{CARB}, r_{TAG}, r_{PEP}, r_{G6P}, r_{GAP}, r_B]'$ representa las velocidades de producción neta de los respectivos componentes y se calcula de la siguiente manera:

$$\mathbf{r} = \mathbf{S}_{stoich}\mathbf{\Lambda} \quad (8.18)$$

donde $\mathbf{S}_{stoich}$ es la matriz de coeficientes estequiométricos construida tomando filas asociadas al componente $NO_3(t)$, $CARB(t)$, $TAG(t)$, $PEP(t)$, $G6P(t)$, $GAP(t)$, $B(t)$, de la matriz $\mathbf{S}'$ presentada en la Tabla 8.1.

Tabla 8.1. Red metabólica de baja dimensión para microalgas autótrofas unicelulares genéricas.

		MR ₁	MR ₂	MR ₃	MR ₄	MR ₅	MR ₆	MR ₇	MR ₈
K'_S	Luz	-30							
	CO ₂	-3						16.61	11.67
	O ₂	3					1	-2.15	-7.37
	Pi	-1	1	1		-1		14.51	4.23
	SO ₄								-0.03
	NO ₃								-1.31
	MG								-0.025
	H ₂ O	-2	-1	-1		1		4.31	6
	H		1		1	1		-34.38	-4.46
	ATP		-1		-1	1		2	
K'_A	ADP		1		1	-1		-2	
	NADH					1		13.46	
	NAD							-13.46	
	NADP					-1		-29.3	
	NADPH							29.3	
	GAP	1		-2	2	-1		-1	
	G6P			1	1		1		-1.14
	PEP					1		-16.61	-3.13
	CARB						1		
	TAG							1	-0.11
K'_B	B								1

El vector de reacciones intracelulares independientes se denota por $\Lambda = [r_{MR1}, \dots, r_{MR8}]$. El flujo de fotones hacia las células se calcula mediante interpolación entre puntos proporcionados por una tabla expresada de la siguiente manera:

$$I_t = I(t) \quad (8.19)$$

La misma resume los gráficos de datos de iluminación como el que se muestra en la figura 8.5, que representa 5 días ideales de verano sin nubes, donde t es el tiempo y $t = 0$ h corresponde a la salida del sol.

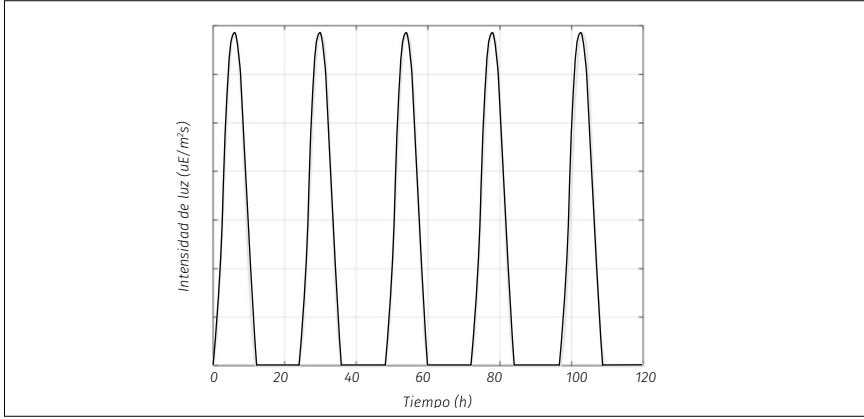


Figura 8.5. Intensidad de luz diaria típica.

La cantidad de biomasa total, que tiene una composición variable según el momento del día, se calcula mediante la siguiente expresión:

$$X(t) = \sum_e \sum_A (e_A A(t) + e_B B(t)) \quad \begin{array}{l} A \in \{CARB; TAG; PEP; G6P; GAP\} \\ e \in \{C; H; O; N; P; S\} \end{array} \quad (8.20)$$

donde e_A y e_B corresponden al número de átomos del elemento e por mol de A y mol de biomasa respectivamente, mientras que $A(t)$ y $B(t)$ corresponden a la concentración de A y biomasa funcional B en el tiempo t . Las condiciones de cultivo deben controlarse de manera óptima para maximizar la productividad. Como el biodiesel se produce a partir de lípidos, la función objetivo implica la maximización de la cantidad de TAG recolectados a lo largo de todo el tiempo de cultivo:

$$\text{Maximize} \int_{t=0}^{t_f} (F_o(t) TAG(t)) dt \quad (8.21)$$

La dinámica de los componentes intracelulares y la biomasa funcional en condiciones de abundancia de nitrógeno está definida por las ecuaciones (8.17.1) y la dinámica del volumen de cultivo está representada por la ecuación (8.17.2) mientras que la ecuación (8.17.3) define la dinámica de excreción de PEP . Además, r_{PEP} debe reemplazarse por $(r_{PEP} - r_{EXCR})$ en la ecuación (8.17.1) para considerar el efecto de la dilución intracelular causada por el proceso de excreción. Para modelar la hipótesis de disipación, la ec. (8.17.3) se elimina y la ecuación (8.22) reemplaza la ecuación cinética correspondiente a la reacción MR1 que se muestra en la Tabla 8.1.

$$r_{MR1} = k'_{MR1} \left(1 - \frac{X_C/B}{Q_r^{max}} \right) I \quad (8.22)$$

El vector de estados viene dado por $x(t) = \{NO_3(t), CARB(t), TAG(t), PEP(t), G6P(t), GAP(t), B(t), V(t), PEP_{EXCR}(t)\}$ si consideramos la hipótesis de excreción o por $x(t) = \{NO_3(t), CARB(t), TAG(t), PEP(t), G6P(t), GAP(t), B(t), V(t)\}$ si consideramos la hipótesis de la disipación para ventear el exceso de energía producido en la fotosíntesis. En un vvcr, se pueden manipular los flujos de entrada $F_i(t)$ y de salida $F_o(t)$ al cultivo y para inducir la falta de nitrógeno, se debe controlar la concentración de nitrato, $NO_3^{IN}(t)$, en el flujo de entrada. El objetivo de la escasez inducida de fuente de nitrógeno es canalizar el flujo de carbono capturado por la fotosíntesis hacia la producción del metabolito de interés (TAG). Así, el vector de variables de control estaría dado por $u(t) = \{F_i(t), F_o(t), NO_3^{IN}(t)\}$. Los límites para las variables de control se definen de la siguiente manera:

$$0 \leq F_i(t) \leq F_i^{max} \quad (8.23.1)$$

$$0 \leq F_o(t) \leq F_o^{max} \quad (8.23.2)$$

$$0 \leq NO_3^i(t) \leq NO_3^{i max} \quad (8.23.3)$$

La siguiente restricción, que impone el volumen del ORP como límite superior del volumen de cultivo, debe aplicarse durante todo el horizonte temporal de optimización:

$$0 \leq V(t) \leq V^{max} \quad (8.24)$$

Se resolvió numéricamente el problema utilizando GPOPS con los parámetros cinéticos presentados en la Tabla 8.2 considerando dos condiciones iniciales necesarias dada la naturaleza cíclica de la producción de TAG en un ORP bajo ciclos fotoperiódicos:

1. el volumen inicial y final y así como las concentraciones inicial y final de biomasa deben ser los mismos.
2. El flujo de fotones por unidad de superficie de cultivo iluminada corresponde al flujo representado en la figura 8.5. Los valores de estado mínimos y los límites de control se resumen en la Tabla 8.2.

Se resolvieron problemas correspondientes a dos «estados fisiológicos» iniciales. Primero se consideró que la producción comienza con la biomasa sin contenido de CARB y TAG asumiendo que los valores iniciales para estos compuestos eran nulos. Para determinar los perfiles óptimos de biomasa

«no agotada» se resolvió otro conjunto de problemas considerando como estados iniciales a los correspondientes de los estados a $t = 48$ h (que se extrajeron de la optimización en la que la biomasa inicialmente estaba agotada) considerándose que la biomasa ha alcanzado un estado fisiológico estacionario correlacionado con el ciclo circadiano. Las condiciones iniciales para ambos casos también se resumen en la Tabla 8.3.

Tabla 8.2. Representación estequiométrica y cinética del núcleo de la red metabólica de una micro-alga (Barouck, 2014)

Red	Reacción macroscópica	Estequiometría	Cinética
Fotosíntesis	MR ₁	30 Light + 3 CO ₂ + 2 H ₂ O + Pi --> GAP + 3 O ₂	$r_{MR1} = k_{MR1} \cdot I$
Glucólisis superior	MR ₂	ATP + H ₂ O <--> ADP + Pi + H	$r_{MR2} = 0$
	MR ₃	2 GAP + H ₂ O --> G6P + Pi	$r_{MR3} = k_{MR3} \cdot \text{GAP}$
	MR ₄	G6P + ATP --> H + ADP + 2 GAP	$r_{MR4} = k_{MR4} \cdot \text{G6P}$
Glucólisis inferior	MR ₅	GAP + ADP + Pi + NAD <--> PEP + ATP + NADH + H ₂ O + H	$r_{MR5} = k_{MR5} \cdot \text{GAP} - k'_{MR5} \cdot \text{PEP}$
Síntesis de carbohidratos	MR ₆	G6P <--> CARB + Pi	$r_{MR6} = k_{MR6} \cdot \text{G6P} - k'_{MR6} \cdot \text{CARB}$
Síntesis de lípidos	MR ₇	GAP + 16.61 PEP + 2 ADP + 13.46 NAD + 29.3 NADPH + 34.48 H + 2.15 O ₂ <--> TAG + 14.61 Pi + 2 ATP + 13.46 NADH + 29.3 NADP + 4.31 H ₂ O + 16.61 CO ₂	$r_{MR7} = k_{MR7} \cdot \text{PEP} \cdot \text{GAP} - k'_{MR7} \cdot \text{TAG}$
Síntesis de biomasa	MR ₈	3.13 PEP + 7.37 O ₂ + 4.46 H + 1.31 NO ₃ + 1.14 G6P + 0.11 TAG + 0.03 SO ₄ + 0.0025 Mg --> B + 11.67 CO ₂ + 4.23 Pi + 6 H ₂ O	$r_{MR8} = k_{MR8} \cdot \text{PEP} \cdot \text{G6P} \cdot \text{NO3}$

Tabla 8.3. Parámetros cinéticos, restricciones en las variables de estado y de control y condiciones iniciales

Parametro	Valor	Unidades		
k _{MR1}	1.1*10 ⁻³ (Excretion)/1.5*10 ⁻³ (Dissipation)	μE (m ² s mol h M B) ⁻¹		
k _{MR3}	223.53	(h mol B) ⁻¹		
k _{MR4}	10.30	(h mol B) ⁻¹		
k _{MR5}	436.95	(h mol B) ⁻¹		
k' _{MR5}	10.00	(h mol B) ⁻¹		
k _{MR6}	80. 00	(h mol B) ⁻¹		
k' _{MR6}	6.50	(h mol B) ⁻¹		
k _{MR7}	1.50 * 10 ³	(M h M B) ⁻¹		
k' _{MR7}	0.80	(h mol B) ⁻¹		
k _{MR8}	1.00*10 ⁴	(M ² h mol B)		
k _{excr}	8.00	(h mol B) ⁻¹		
Q _{min} excreción	13.00	mol/mol		
Q _r ^{Max} disipación	48.00	mol /mol		
Estados	iniciales			
Estados	Agotado	No agotado	Mínimo	Máximo
NO ₃	Libre	Libre	0.0 g/l	2.0 g/l
CARB	0.0 mol /l	0.0397 mol /l	0.0 mol/l	0.2 mol /l
TAG	0.0 mol /l	0.0047 mol /l	0.0 mol /l	0.2 mol /l
PEP	Libre	Libre	0.0 mol /l	0.2 mol /l
G6P	Libre	Libre	0.0 mol /l	0.2 mol /l
GAP	Libre	Libre	0.0 mol /l	0.2 mol /l
B	0.004 mol /l	0.005 mol /l	0.0 mol /l	1 mol /l
V	1000 l	1605.1 l	0 l	2000 l
PEP _{exr}	0 mol /l	0 mol /l	0	
Variable de control	Mínimo		Máximo	
F _i	0 l/h		100 l/h	
F _o	0 l/h		400 l/h	
NO ₃ ⁱⁿ	0 g/l		1 M/l	

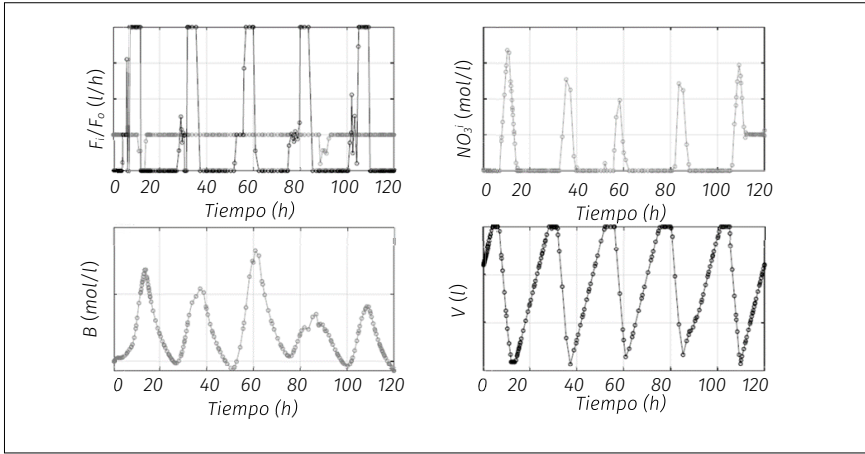


Figura 8.6. Perfiles óptimos de las variables de control (F_i ; F_o ; NO_3^i) y dinámica óptima de la biomasa y el volumen B (—) and V (—) para la biomasa inicialmente vacía de CARB y TAG (hipótesis de disipación).

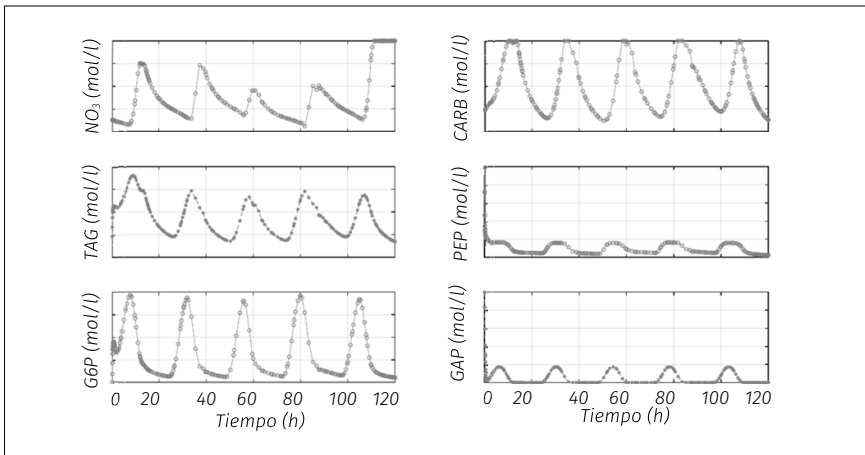


Figura 8.7. Dinámica óptima de los estados intracelulares para la biomasa que inicialmente carece de CARB y TAG (hipótesis de disipación).

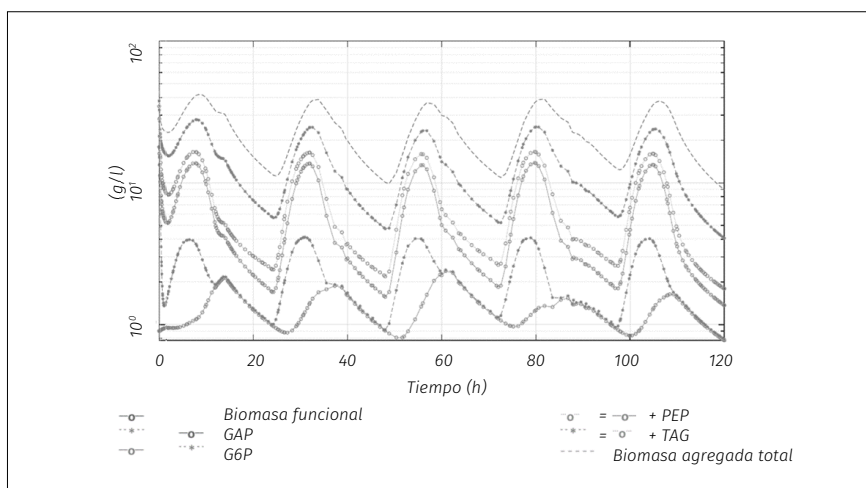


Figura 8.8. Dinámica de la biomasa funcional, componentes intracelulares agregados y biomasa total.

La figura 8.6 representa la dinámica de las variables de control y de los estados macroscópicos. Para ilustrar la dinámica de la biomasa total y de sus componentes agregados, la figura 8.7 representa la evolución de la concentración de biomasa funcional y componentes intracelulares. Debe tenerse en cuenta que la concentración de biomasa funcional $B(t)$ varía entre 0,7 g/l y 3 g/l, mientras que la concentración de biomasa agregada total $X(t)$ se encuentra en el rango de 10 g/l a 40 g/l. La biomasa funcional constituye aproximadamente el 7 % de la biomasa total agregada. En la figura 8.8 se ilustra la evolución de la concentración de la biomasa funcional, de los componentes intracelulares y de la biomasa en el tiempo. Se hace evidente la naturaleza cíclica de la composición de biomasa y su dependencia del momento del día.

Actividades

1. El control óptimo es una técnica de gran utilidad para desarrollar estrategias de quimioterapia. Usualmente se busca desarrollar un tratamiento, considerado minimizar la densidad tumoral y los efectos secundarios de los fármacos durante un intervalo de tiempo determinado. Existen varios modelos de muerte celular por quimioterapia pero consideraremos la hipótesis de muerte logarítmica que establece que la muerte celular debida a fármacos quimioterapéuticos es

proporcional a la población tumoral. Por lo tanto, si $N(t)$ es la densidad normalizada del tumor en el momento t , se considera el siguiente modelo dinámico:

$$\frac{dN}{dt} = rN(t) \ln \left(\frac{1}{N(t)} \right) - u(t) \delta N(t)$$

donde r es la tasa de crecimiento del tumor, es la magnitud de la dosis del químico anti-tumoral y $u(t)$ describe la farmacocinética del fármaco, donde $u(t) = 0$ implica que no hay efecto del fármaco. La condición inicial se toma como $N(0) = N_0$, donde $0 < N_0 < 1$. La función objetivo utilizada es cuadrática y el costo del control, que representa los posibles efectos secundarios y la densidad tumoral $N(t)$ se minimizan en un intervalo de tiempo. Finalmente se requiere que $u(t) \geq 0$ para todo t . Por lo tanto el problema se plantea de la siguiente manera:

$$\begin{aligned} & \text{Min } \int_0^{t_f} aN(t)^2 + u(t)^2 \\ & \text{Sujeto a} \\ & \frac{dN}{dt} = rN(t) \ln \left(\frac{1}{N(t)} \right) - u(t) \delta N(t) \\ & u(t) \geq 0 \end{aligned}$$

Donde a es un parámetro positivo. Resuelva el problema con los siguientes valores de los parámetros:

$$r = 0.3; a = 3.0; \delta = 0.45; N_0 = 0.975; t_f = 20.$$

2. El suelo contaminado contiene bacterias que, a través de su metabolismo son capaces de eliminar contaminantes y un método para gestionar y limpiar las áreas contaminadas es aumentar el nivel de estas bacterias, ya que una mayor cantidad de ellas dará como resultado una degradación más rápida del contaminante. La inyección de nutrientes necesarios para el metabolismo y el crecimiento de las colonias ha demostrado ser una técnica exitosa para aumentar la población bacteriana. El modelado del proceso de remediación no es trivial. Un modelo realista necesitaría incluir efectos espaciales y heterogeneidades en el entorno. Por esta razón, nos centramos en el entorno más controlado de un biorreactor, como los que se utilizan para el tratamiento de aguas residuales. Se considera a un biorreactor con mezcla ideal, donde están presentes un contaminante y una

bacteria que degrada al contaminante a través de su metabolismo. Las bacterias y el contaminante tienen concentraciones espacialmente uniformes y variables en el tiempo $X(t)$, $Z(t)$ respectivamente. El medio de cultivo en el biorreactor es rico en los nutrientes necesarios para el crecimiento de las bacterias, excepto uno, cuya inyección en el reactor tiene una concentración espacialmente uniforme y variable en el tiempo $u(t)$. La velocidad de crecimiento bacteriano está dada por $G u(t)x(t)$ y la tasa de muerte por $D x(t)^2$, de modo que:

$$\frac{dX}{dt} = G u(t)x(t) - D x(t)^2 \quad x(0) = x_0$$

La constante positiva G representa la velocidad máxima de crecimiento, mientras que la constante positiva D es la velocidad de muerte natural. Nótese que, como la degradación del contaminante ocurre a través del metabolismo, el contaminante no es necesario para el crecimiento bacteriano y por lo tanto no aparece en la ecuación. Se supone que la tasa de degradación es proporcional a las concentraciones bacterianas y de contaminantes:

$$\frac{dZ}{dt} = -K x(t)z(t) \quad z(0) = z_0$$

donde $K > 0$ representa la velocidad de degradación de las bacterias. Se supone que el suelo tiene limitaciones físicas que requieren los límites $0 < u(t) < u_{\max}$. El objetivo es minimizar el uso del nutriente y los niveles finales de contaminantes. Consideremos el siguiente funcional objetivo:

$$\ln(z(t_f)) + \int_0^{t_f} A u(t)$$

Plantee y resuelva el problema de control óptimo para el siguiente conjunto de parámetros y condiciones iniciales:

$$K = 2; G = 1; D = 1; x_0 = 0.5; z_0 = 0.1; A = 1; u_{\max} = 1; t_f = 2$$

REFERENCIAS BIBLIOGRÁFICAS

- ASENJO, J. & MERCHUK, J.** (Eds.) (1995). *Bioreactor system design*. CRC Press. <https://doi.org/10.1201/9781482277470>
- BAROUKH, C.; MUÑOZ-TAMAYO, R. (...) BERNARD, O.** (2014). DRUM: A new framework for metabolic modeling under non-balanced growth. Application to the carbon metabolism of unicellular microalgae. *PLoS One*, 9 (8). <https://doi.org/10.1371/journal.pone.0104499>
- CHEN, C. & HWANG, C.** (1990). Convergence analysis of a computational method for optimal control of non-linear differential-algebraic systems. *Int. J. Syst. S.* 21, 11, 2337–2350.
- LEE, J. & RAMIREZ, W.** (1994). Optimal *fed batch* control of induced foreign protein produced by recombinant bacteria. *AIChE J.*, 40, 5, 899–907. <https://doi.org/10.1002/aic.690400516>
- LEVIŠAUSKAS, D. & STANIŠKIS, J.** (1991). Mathematical modeling, optimization and experimental investigation of fed-batch biomass growth process. *Bioautomation*, Bulgarian Academy of Science. *Sofia*, 8, 23–30.
- MARQUEZ, V.; BECCARIA, A. & DONDO, R.** (2024). Optimal harvesting of lipids-rich microalgae subject to a circadian cycle under limited nitrogen supply. *Latin American Applied Research*, 54 (2): 149–157. <https://doi.org/10.52292/j.laar.2024.1144>

Lecturas adicionales

- LENHART, S. & WORKMAN, J.** (2007). *Optimal control applied to Biological Models*. Chapman and Hall SRC. <https://doi.org/10.1201/9781420011418>
- ANITA, S.; ARNAUTU, V. & CAPASSO, V.** (2010). *An Introduction to Optimal Control Problems in Life Sciences and Economics*. Birkhäuser.

9 Estudio y modelado de redes metabólicas en microalgas oleaginosas

En este capítulo, confeccionado por el doctor Alejandro Beccaría y la doctora Vanina Marques, se aborda el tema del modelado metabólico en microalgas, como una herramienta para la obtención eficiente de diferentes bioproductos. Para esto se presentan generalidades del grupo microbiano y su relevancia biotecnológica, la diversidad metabólica en microalgas de importancia industrial y los diferentes biorreactores usados para su cultivo. Con estos elementos, se analizan algunos ejemplos de modelado, y se presentan de las bases de datos más tradicionales.

MICROALGAS: DEFINICIÓN Y GENERALIDADES TAXONÓMICAS. NICHOS ECOLÓGICOS. IMPORTANCIA BIOTECNOLÓGICA

Las microalgas conforman un grupo polifilético de microorganismos que son generalmente fotosintéticos, pero a diferencia de las plantas superiores no poseen hojas, tallos ni raíces. Este grupo es heterogéneo respecto a su morfología, metabolismo, modos de reproducción, evolución y ecología, entre otros aspectos.

Son organismos generalmente acuáticos, pero pueden encontrarse también en ambientes muy diversos, incluso en ambientes extremos en relación con la temperatura, disponibilidad de agua, pH, etcétera.

En cuanto a su estructura celular, pueden ser eucariotas o procariotas, este último es el tipo celular de las cianobacterias o algas verde azuladas. Pueden ser unicelulares o pluricelulares (coloniales).

Desde una perspectiva biotecnológica, estos organismos son relevantes por su capacidad de producir y acumular compuestos con aplicaciones actuales y potenciales, en el área de los alimentos, la industria cosmética, la agricultura y la alimentación animal. Entre esos compuestos se pueden mencionar: ácidos grasos poliinsaturados (como los ácidos araquidónico, eicosapentaenoico y docosahexaenoico), polisacáridos (beta-glucanos, exopolisacáridos sulfatados, etc.), pigmentos (betacaroteno, astaxantina, ficobiliproteínas, clorofila). Por su composición bioquímica, en términos de porcentajes de proteínas y

carbohidratos de calidad nutricional, la biomasa microalgal de algunas especies se comercializa con fines nutracéuticos, para la alimentación humana o para la cría de animales (ganadería, acuicultura, etc.). Además, se ha demostrado que varias especies de microalgas son candidatas para contribuir en la sustitución de los combustibles fósiles, ya que pueden acumular almidón que puede ser convertido en etanol, mientras que otras pueden acumular lípidos con un perfil de ácidos grasos adecuado como materia prima para la elaboración de biodiesel (Chen *et al.* 2013, Li *et al.* 2008).

NUTRIENTES: FUENTES DE CARBONO Y NITRÓGENO

La producción de biomasa y productos biotecnológicos derivados de las microalgas es afectada de manera directa por las condiciones de cultivo, entre ellas se encuentran las fuentes de carbono, energía y nitrógeno empleadas. Atendiendo a las fuentes de carbono (C) y energía empleadas, el metabolismo puede ser autotrófico, mixotrófico o heterotrófico.

Las microalgas pueden emplear fuentes de nitrógeno (N) inorgánicas, principalmente NO_3^- , NH_4^+ y en algunos casos NO_2^- . Algunas también son capaces de usar fuentes orgánicas de N, como la urea, aminoácidos, purinas, pirimidinas, glicina betaína, péptidos y peptonas (Neilson y Lewin, 1974). Diferentes sistemas de transportes de estas formas del N que pueden ser incorporadas, han sido encontrados y estudiados en microalgas (Raven y Giordano, 2016).

Metabolismo del carbono: autotrofia, heterotrofia y mixotrofia

Acorde a las fuentes de carbono y energía que pueden utilizar, las microalgas son en su mayoría fotoautotróficas, es decir, emplean el carbono desde el CO_2 y la energía proveniente de los fotones de la luz (ya sea solar o artificial).

El crecimiento en condiciones autotróficas es uno de los modos más explotado de cultivo de estos organismos, ya que implica el consumo de CO_2 , por lo cual constituye una estrategia económica y que, además, contribuye a la mitigación este gas de efecto invernadero. Sin embargo, las eficiencias fotosintéticas y la productividad en biomasa continúan siendo objeto de optimización, en particular para que puedan ser comercialmente viables (Masojidek *et al.*, 2013). El cultivo autotrófico puede llevarse a cabo en medios minerales o empleando aguas residuales, valiéndose de la capacidad de las microalgas de crecer empleando compuestos que contienen fósforo, nitrógeno, azufre, etc., como fuente de esos elementos químicos, empleando el

CO₂ atmosférico y la luz solar como fuente de energía, siendo capaces de tolerar, e incluso biorremediar mediante diferentes mecanismos físicos y bioquímicos, sustancias cuyo vuelco se encuentra regulado por las diferentes legislaciones y que no podrían ser depuradas por otros grupos microbianos por ser potencialmente inhibitorias o tóxicas para algunas formas de vida.

Una gran cantidad de microalgas pueden utilizar HCO₃⁻ como fuente de C, además del CO₂. Para poder usarlo, deben tener un transportador específico, porque el HCO₃⁻ no puede difundir por la membrana plasmática. Esta capacidad posiblemente responda al hecho de que la concentración de CO₂ disuelto en agua suele ser en muchos cuerpos de agua menor respecto a la concentración de HCO₃⁻ (Beardall y Raven, 2016). Esta situación es diferente del caso de las plantas terrestres, que están expuestas al CO₂ en la fase gaseosa.

En microalgas, tanto el HCO₃⁻ como el CO₂ pueden ser transportados activamente al interior de las células, dependiendo de la microalga y las condiciones eco y fisiológicas. Anhidrasas carbónicas que pueden estar presentes en varios espacios de los compartimentos subcelulares, dependiendo también de la microalga, transformarían el HCO₃⁻ en CO₂. Esto permite aumentar localmente la concentración de CO₂ para favorecer la fotosíntesis sobre la fotorrespiración, como consecuencia de las actividades opuestas de la RUBISCO, como se comentará más adelante.

Pero algunas microalgas son capaces de utilizar también compuestos orgánicos como fuentes de carbono y energía, es decir, son heterótrofas (Perez Garcia *et al.*, 2011).

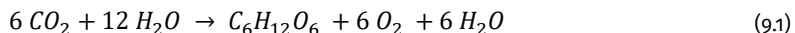
Algunas microalgas pueden crecer bajo un régimen mixotrófico, empleando simultáneamente tanto CO₂ como compuestos orgánicos asimilados a través del metabolismo fotosintético como respiratorio (Neilson y Lewin, 1974). La energía puede ser obtenida tanto de la luz como de los compuestos orgánicos. La fotoheterotrofia constituye un caso particular, en el cual la luz es la fuente de energía y el carbono es obtenido desde los compuestos orgánicos.

En esta diversidad trófica, algunas microalgas son fotoautótrofas obligadas, algunas pueden crecer de manera mixotrófica y unas pocas especies pueden alternativamente crecer bajo un régimen con exclusividad heterotrófico, en ausencia de luz.

Conceptos básicos de fotosíntesis, fotorrespiración, respiración aeróbica y fermentación

La fotosíntesis es una vía metabólica, es decir un conjunto de reacciones que se llevan a cabo en el interior de las células y en las que ocurren transformación de compuestos químicos. En la fotosíntesis, los fotones de luz aportan

la energía para que el CO_2 presente en la atmósfera se transforme en carbohidratos. La ecuación 9.1 representa este proceso de manera generalizada.



Estas transformaciones ocurren en dos etapas o fases que comprenden reacciones dependientes de luz (fase luminosa) y reacciones independientes de la luz (fase oscura).

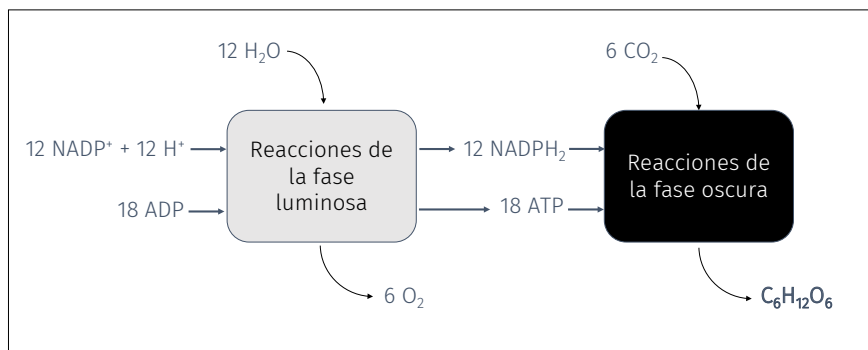


Figura 9.1. Esquema del proceso de la fotosíntesis oxigénica, principales productos de las fases que la comprenden.

Tanto las reacciones dependientes como las independientes de la luz cesan en la oscuridad porque la síntesis de ATP y la reducción de NADP^+ requieren luz. Por otra parte, las reacciones de la fase oscura pueden ocurrir en presencia de luz, pero no la requieren directamente.

La fotosíntesis se lleva a cabo como resultado de la captación de fotones y la transferencia de electrones a nivel de una serie de complejos y proteínas que se encuentran en las membranas tilacoides, en el interior de los cloroplastos de las células eucariotas fotosintéticas. En el caso de las cianobacterias, donde no hay cloroplastos, las membranas tilacoides se encuentran sueltas y distribuidas en el citoplasma.

Los procesos que se describen a continuación, implican a las membranas tilacoides, y ocurren tanto en microalgas eucariotas y cianobacterias, con algunas diferencias en los pigmentos involucrados en la captación de la energía luminosa. Las microalgas realizan fotosíntesis, tomando la luz como fuente de energía en el rango visible del espectro (400–700 nm), aunque las longitudes de onda más adecuadas son las próximas al azul (450 nm) y al rojo (650 nm), ya que los pigmentos presentes en la mayoría de las microalgas tienen una alta absorción en tales longitudes (Su, 2021).

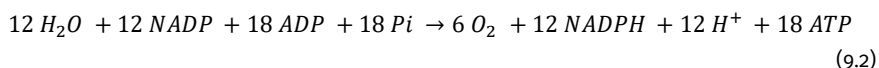
En condiciones de suficiencia de nutrientes, la actividad fotosintética y en consecuencia la velocidad específica de crecimiento, aumenta al incrementarse la intensidad luminosa hasta alcanzar un valor de saturación (que es dependiente de la especie), a partir del cual comienza el fenómeno de fotoinhibición (Masojídek *et al.*, 2013).

En la fase luminosa la energía de los fotones de la luz es transformada en energía química en la forma de ATP (un compuesto con alta energía) y NADPH_2 (una coenzima en su forma reducida que actúa como transportador de electrones en varias vías metabólicas). En esta fase, se consume H_2O que las microalgas obtienen de su entorno acuoso, y se producen O_2 y energía de los fotones es canalizada por el movimiento de electrones a través de una serie de complejos proteicos. La energía de los fotones es la que impulsa este transporte, excitando los electrones de los pigmentos. Esta energía es captada por dos complejos formados por proteínas y pigmentos (clorofilas, carotenoides y ficobilinas), los fotosistemas (FS) I y II. En los FS los pigmentos fotosintéticos absorben la energía de los fotones y esa energía es capaz de excitar electrones que pasan a través proteínas móviles y citocromos, para producir NADPH_2 .

En esa transferencia ocurre el pasaje de protones al interior, al lumen de los sacos tilacoides, contra gradiente de concentración. En el FSII, ocurre la fotólisis del agua. Las moléculas de H_2O ceden electrones en una reacción que genera oxígeno y protones, en el interior del espacio tilacoide.

El transporte pasivo de H^+ desde el lugar más concentrado (espacio tilacoide) al lugar de menor concentración (estroma del tilacoide), es decir a favor de gradiente, impulsa la síntesis de ATP, catalizada por la ATP sintasa. Esta enzima de transmembrana cataliza dicha síntesis a partir de ADP y fósforo inorgánico (Pi), proceso que se denomina foto-fosforilación.

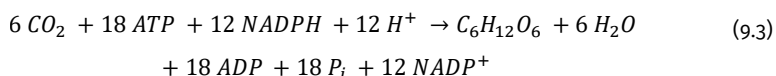
La ecuación (9.2) representa la reacción global de la fase luminosa:



Durante la fotosíntesis, también ocurre un proceso que se denomina transporte cíclico de electrones, que ocurre en paralelo al descrito anteriormente. Este proceso involucra al FSII, los electrones son reciclados con intervención de la ferredoxina y el complejo de citocromos. En este proceso cíclico no se forma NADPH_2 , pero sí hay contribución al gradiente de H^+ que permiten la generación de ATP, que es requerido en mayor proporción en las reacciones de la fase oscura, que se describe a continuación.

En las reacciones de la fase oscura se emplea la energía del ATP y el NADPH_2 generado en las reacciones de la fase luminosa, para sintetizar glucosa a

partir del CO_2 obtenido desde el entorno. En las microalgas, esta fijación del carbono ocurre a través del ciclo de Calvin, en el interior del cloroplasto de las microalgas eucariotas, o en el carboxisoma de las cianobacterias. La glucosa generada puede ser exportada del cloroplasto al citosol y ser utilizada, como se detalla más adelante. La ecuación 3 sintetiza las reacciones que ocurren en la fase oscura de la fotosíntesis.



La fijación del C ocurre en las siguientes etapas, como se esquematiza en la figura 2:

- i) Carboxilación: en esta fase el CO_2 se combina con ribulosa-1,5-bifosfato (un azúcar de 5 átomos de C), para formar 3-fosfoglicerato. Esta reacción está catalizada por la enzima ribulosa-1,5-bifosfato carboxilasa o RUBISCO. Esta posee actividad carboxilasa, pero también actividad oxigenasa, sobre la que se comentará más adelante.
- ii) Reducción: el 3-fosfoglicerato se fosforila y se reduce a gliceraldehído-3-fosfato, utilizando el poder reductor del NADPH_2 y con la energía suministrada en la forma de ATP que provienen de la fase lumínica previa. Se generan NADP^+ , ADP y P_i .
- iii) Regeneración: una parte de las moléculas de gliceraldehído-3-fosfato se combinan para formar la ribulosa-1,5-bifosfato para volver a ser empleada en la etapa i) del ciclo. Esta regeneración requiere suministro de ATP . Por otra parte, dos de las moléculas de gliceraldehído-3-fosfato que representan la ganancia neta del ciclo de Calvin se condensan para producir fructosa-1-fosfato que luego se puede transformar en glucosa u otras hexosas.

La fotorrespiración es un proceso que se da en las células fotosintéticas, como consecuencia de la actividad oxigenasa que posee la enzima RUBISCO. En el proceso de fotorrespiración el carbono orgánico se convierte en CO_2 , sin ganancia de ATP . Se trata de un proceso diferente del de la respiración mitocondrial, que se describe más adelante. Las actividades carboxilasa y oxigenasa son competitivas y dependen de la relación de concentraciones entre O_2 y CO_2 . Si bien las enzimas RUBISCO de diferentes organismos fotosintéticos tienen diferentes afinidades por el CO_2 , la actividad oxigenasa se ve más favorecida cuanto menor es la relación $\text{CO}_2:\text{O}_2$ en la región del sitio activo de la enzima (Raven *et al.*, 2020).

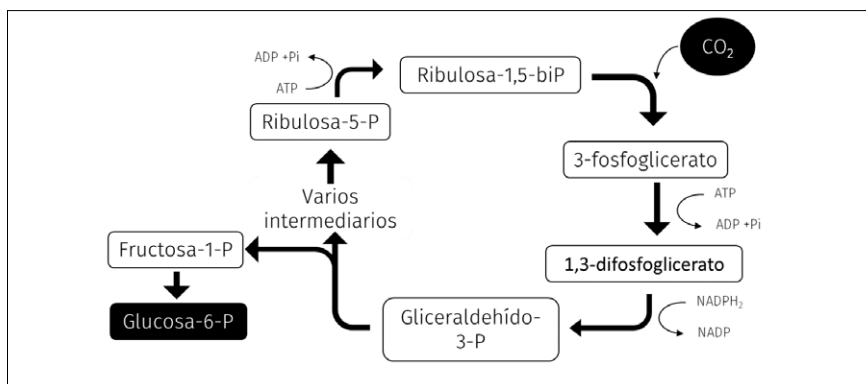


Figura 9.2. Esquema simplificado del ciclo de Calvin, vía a través del cual ocurre la fijación del carbono en las microalgas y otras células fotosintéticas.

La actividad oxigenasa de la RUBISCO cataliza la reacción del O_2 con ribulosa-1,5-bisfosfato, para formar 3-fosfoglicerato y 2-fosfoglicolato. El 2-fosfoglicolato puede ser desfosforilado y convertido tras varios pasos en serina y luego amonio y CO_2 , o ser transportado y metabolizado por mecanismos que involucran asimilación de O_2 . La fotorrespiración es un mecanismo compensatorio que permite reducir la limitación de aceptores de electrones y la disipación de energía que podría generar daño en el FSII y reducir la generación de especies reactivas del oxígeno particularmente en condiciones de exceso de luz (Raven *et al.*, 2020).

En la naturaleza las microalgas están expuestas a un ciclo de luz y oscuridad, por lo que las actividades metabólicas y fisiológicas están reguladas por un ritmo circadiano (Suzuki y Johnson, 2001). Por lo general, los procesos más sensibles a la radiación UV están programados endógenamente para ser llevados a cabo durante las horas de oscuridad. En el caso de iluminarse artificialmente de manera continua, el crecimiento ocurre, pero los procesos celulares no se desarrollarían acorde a un patrón dependiente de la luz.

Los procesos que ocurren durante las horas de oscuridad, requieren del metabolismo respiratorio de una parte del C fijado durante las horas de luz. De esta manera y según algunas estimaciones, durante la fase oscura hay una disminución de hasta un 30 % de la masa microalgal generada en las horas de luz (Edmundson y Huesemann, 2015), aunque la proporción entre tasa fotosintética y respiratoria suele ser variable entre microalgas de diferentes taxones y en diferentes estados fisiológicos. Estos aspectos deben ser tenidos en cuenta cuando se realizan estimaciones de la productividad de la biomasa, y en el planteo de los modelos metabólicos, como se detallará más adelante.

En las microalgas, la asimilación oxidativa de sustratos puede satisfacer dos requerimientos fundamentales, el suministro energía para el mantenimiento y biosíntesis en condición de oscuridad y provee los esqueletos carbonados necesarios para la biosíntesis, en cualquier condición (oscuridad o luz) (Perez-García *et al.*, 2011). Es decir, además de ATP, NADH o NADPH₂, los compuestos carbonados que se producen en las vías oxidativas de la glucosa, son materia prima para la síntesis de componentes celulares, proteínas (estructurales y catalíticas), lípidos, carbohidratos, ácidos nucleicos. En la figura 9.3 se esquematizan las vías metabólicas centrales que operan en microalgas eucariotas. La glucosa formada como producto de la fotosíntesis, así como la glucosa que es consumida en el caso de los cultivos mixo o heterotróficos, puede ser metabolizada en el citosol por dos vías diferentes en la mayoría de las microalgas: la glicólisis clásica por la vía Embden-Meyerhof-Parnas (EMP) y la vía de las pentosas fosfato (VPF). En el caso de algunas diatomeas, se ha demostrado además a nivel genómico y funcional la existencia de las enzimas de la vía Entner-Doudoroff en la matriz mitocondrial (Fabris *et al.*, 2012). La existencia de esta vía es muy rara en organismos eucariotas y no está representada en el esquema de la figura 9.3.

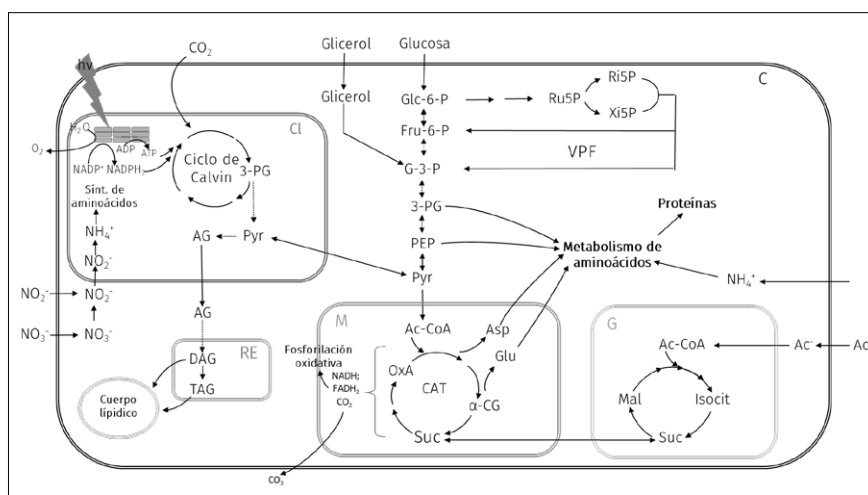


Figura 9.3. Esquema simplificado del metabolismo central en microalgas eucariotas, en los principales compartimentos intracelulares: C, citosol; M, mitocondria; RE, retículo endoplasmático; G, glioxisoma; CL, cloroplasto. Glc-6-P, glucosa-6-fosfato; Fru-6-P, fructosa-6-fosfato; G-3-P: gliceraldehído-3-fosfato; 3-PG, 3-fosfoglicerato; PEP, fosfoenolpiruvato; Pyr, piruvato; Ac-CoA, acetil-coenzima A; OxA, oxalacetato; αCG, alfa-cetoglutarato; Suc, succinato; Glu, glutamato; CAT, ciclo de los ácidos tricarboxílicos; Asp, aspartato; Mal, malato; Isocit, isocitrato; Ac-, acetato; Ru5P, ribulosa-5-fosfato; Ri5P, ribosa-5-fosfato; Xi5P, xilosa-5-fosfato; VPF, vía de las pentosas fosfato; AG, ácidos grasos; DAG diacilglicéridos; TAG, triacilglicéridos; hv, energía lumínica.

Como resultado de la glicólisis por la vía EMP, se produce piruvato, ATP y NADH. El piruvato se descarboxila, formando acetilo que se une a la coenzima A (CoA), dando AcCoA, que es el intermediario que puede ingresar al CAT o a la síntesis de lípidos. Del CAT se obtiene fundamentalmente poder reductor, en forma de NADPH₂ y FADH₂, ATP y CO₂. En el caso de la cadena respiratoria en microalgas, al igual que otros organismos, el producto principal es el ATP, necesario para diversas funciones (mantenimiento, síntesis, transporte, etcétera).

Las condiciones del entorno influyen en la tasa de flujo a través de las vías. La VPF opera de manera más preponderante en condición de oscuridad, cuando el crecimiento es exclusivamente heterotrófico (Perez-García *et al.*, 2011), o cuando por alguna condición la fotosíntesis se encuentra inactivada, aún en presencia de luz (Jeon *et al.*, 2017). Como resultado de la VPF se genera NADPH₂, una serie de pentosas (entre las que se destaca a la ribosa-5-P, necesaria para la síntesis de ácidos nucleicos) y, además, fructosa-6-P y gliceraldehído-3-P.

Cabe destacar que, en algunas microalgas como cianobacterias y ciertos dinoflagelados, existen diferencias particulares en intermediarios del CAT y la cadena respiratoria, que no serán detalladas en este capítulo y que son analizadas por Raven y Beardall (2016).

En sus entornos naturales, las microalgas pueden encontrarse transitoriamente e incluso sobrevivir de manera más o menos prolongada, en condición de oscuridad y anoxia. El piruvato generado durante la glicólisis puede tener diferentes destinos, en distintos compartimentos subcelulares y dependiendo de las condiciones del entorno.

El proceso de respiración y fermentación presentan varias diferencias sustanciales. La más destacable es que en la primera, la generación de ATP ocurre por fosforilación dependiente de una cadena transportadora de electrones, con generación de gradiente de protones y actividad ATPasa, mientras en la segunda, el ATP se genera por fosforilación a nivel de sustrato. La respiración será aeróbica cuando el aceptor final de electrones en la cadena transportadora es el O₂, pero si este no está presente, algunos organismos eucariotas pueden utilizar aceptores finales diferentes, tomados del medioambiente (como por ejemplo el nitrato) o producido endógenamente como el fumarato (Müller *et al.*, 2012). La cantidad de ATP producido también es diferente en cada vía metabólica, siendo la respiración aeróbica la que mayor cantidad de moléculas rinde por cada mol de glucosa que se oxida.

Microalgas eucariotas como algunas especies de los géneros *Chlamydomonas*, *Chlorella*, *Scenedesmus* o *Chlorogonium* (Atteia *et al.*, 2013) y cianobacterias, como por ejemplo especies de los géneros *Anabaena*, *Oscillatoria* o *Nostoc* (Stal y Moezelaar, 1997) pueden metabolizar fermentativamente diferentes

sustratos. Las rutas metabólicas empleadas con más frecuencia son las fermentaciones ácido-mixtas, mientras que la respiración anaeróbica, que suele ser típica en ciertos heterótrofos eucariotas, no se encuentra comúnmente entre las microalgas.

En sus ambientes naturales, en condición de anoxia y oscuridad, las microalgas (tanto eucariotas como procariotas) pueden generar lactato, acetato, etanol, formiato, CO_2 , en algunas especies H_2 , como productos de la fermentación de carbohidratos endógenos, y solo en algunos casos también los exógenos (Kreuzberg, 1984; Gaffron y Rubin, 1942; Skjånes *et al.*, 2008; Stal y Moezelaar, 1997). Por otra parte, en algunas especies de *Euglena*, se acumulan ceras —ácidos grasos de 10 a 18 átomos de C esterificados con alcoholes alifáticos— en el citoplasma (Inui *et al.*, 1982), como productos de la fermentación de paramilón, que es polisacárido de reserva en este género.

En algunas cianobacterias, el metabolismo involucrado en oscuridad y anoxia es homofermentativo mientras que en otros casos son heterofermentativos, involucrando enzimas de las vías de las pentosas fosfatos y de la vía EMP, dependiendo de la especie (Stal y Moezelaar, 1997).

La mayor parte del conocimiento acerca de las vías involucradas en el metabolismo fermentativo y la producción de hidrógeno proceden del estudio a través de análisis proteómicos y bioquímicos de *Chlamydomonas reinhardtii* que es considerada una microalga eucariota modelo. Las enzimas involucradas en la transformación fermentativa del piruvato han sido encontradas en mitocondrias, cloroplasto y citosol (Terashima *et al.*, 2011; Catalanotti *et al.*, 2013) por lo que se especula que esas serían las localizaciones más probables para el desarrollo de tal catabolismo. En la figura 9.4 se encuentran representadas las reacciones de las vías fermentativas descritas en *Chlamydomonas*, a modo ilustrativo. Sin embargo, cabe mencionar que para otras microalgas las vías de fermentación podrían ser diferentes, ya que los productos de fermentación detectados varían entre géneros e incluso entre especies del mismo género (Atteia *et al.*, 2013).

La síntesis de H_2 en *Chlamydomonas* ha sido asociada al estroma del cloroplasto por ser el sitio donde se localizó a la hidrogenasa involucrada en la catálisis de la reacción. Esta síntesis puede también ocurrir en anoxia y en presencia de luz, mediante el proceso conocido como biofotólisis directa. En esta condición, los electrones excitados por la energía luminosa en el PSI y PSII se transfieren secuencialmente hacia una ferredoxina (FDX), que a su vez es capaz de ceder electrones a la hidrogenasa que se ocupa de la síntesis de hidrógeno, volviéndola a su estado reducido. Este proceso involucra al PSII, donde ocurre la fotólisis del H_2O , generando O_2 , que actúa como supresor de la actividad hidrogenasa (Kim y Kim, 2011). Este proceso no produce CO_2 y depende de la eficiencia fotosintética de la microalga en la que ocurre.

En *Chlamydomonas*, se ha descrito una condición particular para favorecer la síntesis de H_2 . En condición de restricción o deprivación de S, ocurre un marcado decremento de la actividad del PSII que conlleva a la reducción de la generación de O_2 por ruptura de moléculas de H_2O , manteniendo bajos niveles de O_2 que favorecen la síntesis de H_2 (Kim y Kim, 2011). En ausencia de luz, la FDX recibe equivalentes de reducción a partir de los procesos de oxidación de sustratos derivados de los compuestos de reserva (almidón, glucógeno, etc.) por vía fermentativa (figura 9.4).

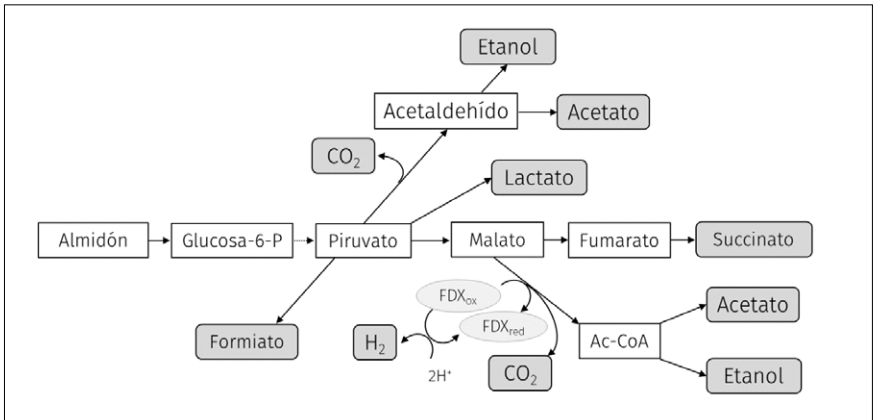


Figura 9.4. Representación esquemática del metabolismo de *Chlamydomonas reinhardtii* en condiciones de anaerobiosis y oscuridad. Las vías representadas se encuentran en varias localizaciones subcelulares de la microalga.

Metabolismo del nitrógeno

Como se comentó anteriormente, la incorporación de nitrógeno a las células está mediada por transportadores. Para el caso del NO_2^- , una vez en el interior de la célula será transportado como tal al cloroplasto mientras que el NO_3^- será previamente reducido a NO_2^- . En esta organela será reducido a NH_4^+ para su posterior asimilación. La reducción de estos compuestos demanda energía y otros recursos a la célula. Por otra parte, cuando la fuente de N en el medio es el NH_4^+ , este es transportado al interior de la célula como tal. Su asimilación es mediada por las enzimas que componen el ciclo de la glutamina: glutamina sintasa (GS) y glutamina-oxoglutarato aminotransferasa (GOGAT), o en algunos taxones de microalgas y en determinadas condiciones también por la glutamato deshidrogenasa (GDH) (Hellebust y Ahmad, 1989). Por la actividad GS se incorpora otro NH_4^+ al glutamato para dar glutamina.

En la mayoría de las microalgas existen dos isoformas de GS que tienen localizaciones diferenciales, una de ellas cloroplastídica, que está activa en presencia de luz, mientras que la otra isoforma se localiza fuera del cloroplasto y su actividad es preponderante en condición de oscuridad (Vega, 2020). La GOGAT, genera dos moléculas de glutamato a partir de glutamina y alfa-cetoglutarato. También existen dos isoformas que actúan en condición de luz u oscuridad, y sus actividades se encuentran acopladas a la oxidación de ferredoxina o NADPH_2 , respectivamente (Hellebust y Ahmad, 1989).

Por otra parte, en algunas clorofitas como *Chlorella sorokiniana* por ejemplo, una GDH cataliza la síntesis de glutamato a partir de NH_3 y alfa-cetoglutarato, una reacción dependiente de NADPH_2 , y también existen isoformas con localizaciones variables, dependiendo de la microalga. Esta enzima es inducible por las condiciones del entorno, especialmente por las concentraciones de NH_4^+ .

Una vez incorporado el amonio como glutamato, ya sea por acción de las enzimas del ciclo GS/GOGAT, o por la GDH, el nitrógeno se incorpora y distribuye a otros aminoácidos, por ejemplo: i) aspartato, por transaminación del oxalacetato, ii) asparagina por transferencia del amonio de la glutamina al aspartato. Los aminoácidos mencionados anteriormente constituyen insumos necesarios en las vías de la síntesis de otros aminoácidos y compuestos nitrogenados como las proteínas, clorofilas, nucleótidos y poliaminas (Inokuchi *et al.*, 2002).

El metabolismo del carbono y del nitrógeno están relacionados ya que la asimilación del N inorgánico requiere esqueletos carbonados que provienen de la respiración del CO_2 fijado o asimilado como orgánico, además de energía como ATP y NADPH_2 que provienen del TCA, de la fotosíntesis y el transporte mitocondrial de electrones para sintetizar glutamato, glutamina y aspartato. Dado que en general las enzimas que intervienen en estos metabolismos están reguladas por condiciones que dependen del entorno, resulta indispensable dimensionar que ambos metabolismos tienen una fuerte influencia en la distribución del C en otras vías como las de acumulación de compuestos de reserva.

SÍNTESIS DE POLISACÁRIDOS: RUTAS BIOSINTÉTICAS DE LA SÍNTESIS DE ALMIDÓN Y OTROS CARBOHIDRATOS DE ALTA MASA MOLECULAR

Durante el período de exposición a la luz, el CO₂ se fija en moléculas de glucosa, a partir de las cuales se construyen polisacáridos, como por ejemplo el almidón.

El almidón, al igual que en otros organismos fotosintéticos, sirve como molécula de reserva de carbono y fuente de energía química en estos tipos celulares (Busi *et al.*, 2014) y su síntesis constituye un proceso fisiológico. El almidón, así como los lípidos, se acumulan en las microalgas en condiciones de agotamiento de algún macronutriente, usualmente N, P o S (Ran *et al.*, 2019), es decir en respuesta a una situación de estrés. Algunas microalgas acumulan otros tipos de polisacáridos de reserva, como por ejemplo crislaminarina o paramilón. Las microalgas producen también polisacáridos estructurales, generalmente considerados exopolisacáridos, que constituyen la pared celular o que se secretan al medio. Muchos de los exopolisacáridos son heteropolisacáridos, es decir compuestos por diversos tipos de monosacáridos.

Entre los diferentes linajes de microalgas también existen diferencias en el compartimento subcelular donde ocurre la síntesis de los polisacáridos. Incluso el mismo polisacárido, como por ejemplo el almidón puede sintetizarse en sitios diferentes en los distintos linajes, en algunos en el cloroplasto, en otros en el citosol.

Para el caso del almidón, la síntesis es un proceso que ocurre en las siguientes etapas (Busi *et al.*, 2014): i) activación a nucleótido azúcar (ADP-glucosa o UDP-glucosa, en función de la microalga), ii) elongación de la cadena de α -1,4-glucano; iii) ramificación y iv) desramificación que facilita la cristalización en gránulos.

En lo que respecta a los polisacáridos de tipo estructural, la pared celular de las microalgas (en los casos en los que está presente), presenta una amplia diversidad en cuanto a las dimensiones y su composición, en general suele estar conformada por una capa de celulosa fibrilar de espesor variable y otras capas de otros polisacáridos. La síntesis de celulosa es un proceso complejo, mediado por una celulosa sintasa asociada a la membrana plasmática. El proceso de polimerización ocurre del lado externo de dicha membrana, donde se van depositando las microfibrillas de celulosa a la vez que va ocurriendo la polimerización de la celobiosa (el disacárido que constituye la unidad que se repite en la molécula del polisacárido) (Tsekos, 1999).

La síntesis de otros exopolisacáridos estructurales, generalmente heteropolisacáridos, en las microalgas eucariotas ocurre en el complejo de Golgi, involucrando vías particulares en virtud de los monosacáridos, en muchos

casos a través de mecanismos de elongación de una cadena en crecimiento por transferencia de los residuos de azúcares desde azúcares nucleótidos (Rossi y De Philippis, 2016).

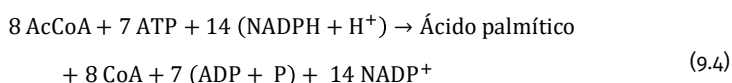
SÍNTESIS DE ACEITES: RUTAS BIOSINTÉTICAS DE LA SÍNTESIS DE ÁCIDOS GRASOS Y TRIACILGLICÉRIDOS

Muchas especies de microalgas son capaces de sintetizar y acumular elevadas cantidades de aceite (químicamente triacilgllicéridos, TAGS), conocido como aceite microbiano o unicelular. Cuando la proporción de aceite es superior al 20 % en base seca de la biomasa, se considera que es una cepa oleaginosa (Thevenieau y Nicaud, 2013).

Diversos ácidos grasos pueden estar contenidos en esos aceites, destacándose perfiles compatibles para su aplicación en la síntesis de biodiésel y, cuando el aceite presenta elevadas cantidades de ácidos grasos poliinsaturados de cadena larga, para la formulación de alimentos funcionales o fármacos (Nazir *et al.*, 2020).

Cuando la biosíntesis de TAGS en microalgas ocurre por síntesis *de novo* de ácidos grasos, que sería el mecanismo principal en estos microorganismos (Li *et al.*, 2013), las rutas metabólicas se distribuyen entre los cloroplastos y el retículo endoplasmático. En la primera de estas organelas, como se muestra en las figuras 2 y 3, el 3-fosfoglicerato es un intermediario del ciclo de Calvin que puede ser sustrato de la enzima fosfogliceromutasa para generar el 2-fosfoglicerato. Este puede ser transformado en fosfoenolpiruvato mediante la enzima enolasa, compuesto que se convierte a piruvato mediante la enzima piruvato quinasa, en una reacción que produce ATP. Finalmente, el piruvato puede ser descarboxilado en una reacción bioquímica que involucra a la enzima piruvato deshidrogenasa y a la coenzima A como cofactor, generando Ac-CoA, que constituye el precursor para la síntesis de los ácidos grasos; además del ATP y el NADPH₂, todos formados en el cloroplasto.

El grupo acetilo del Ac-CoA formado, será usado en su mayoría para la síntesis de malonil-Coenzima A, fijando CO₂ y consumiendo ATP y NADPH₂. Luego, se liberan tantas moles de CO₂ como las fijadas, resultando el balance de la reacción global que se representa en la ecuación 4, para la biosíntesis del ácido palmítico (C16:0).



Los ácidos palmítico y esteárico (C18:o) son los ácidos grasos mayormente producidos mediante la síntesis *de novo*. Dependiendo de las cepas y condiciones de cultivo, podrá incrementarse la cantidad de átomos de carbono en los ácidos grasos, mediante elongasas, así también como generarse insaturaciones en ciertos sitios específicos de la cadena carbonada, mediante desaturasas (Chen y Wang, 2021).

Una vez sintetizados los ácidos grasos son exportados al citoplasma y desde allí al retículo endoplasmático. En esta organela, se van formando los ésteres glicéricos de manera secuencial, a través de lo que se conoce como ruta de Kennedy (Scodelaro Bilbao *et al.*, 2017; Lenka *et al.*, 2016), usando glicerol-3-fosfato como primer cosustrato. Los TAGs formados son movilizados al citoplasma donde se acumulan en estructuras conocidas como cuerpos lipídicos. La biogénesis de los cuerpos lipídicos puede incluir otras vías diferentes a las aquí brevemente comentadas (Guéguen *et al.*, 2021; Scodelaro Bilbao *et al.*, 2017; Fan *et al.*, 2011).

CULTIVO AUTOTRÓFICO DE MICROALGAS EN BIORREACTORES

Como ya se comentó en la sección 8.2.1, muchas especies de microalgas pueden crecer heterotróficamente, empleando fuentes de energía orgánicas, no lumínicas. Este metabolismo permite recurrir a biorreactores tradicionales para el desarrollo de cultivos microbianos, donde la iluminación no es un aspecto a considerar.

Cuando se desarrollan cultivos autotróficos, la calidad y cantidad de luz que incide al cultivo son aspectos críticos para obtener buenos rendimientos y productividades.

Para estas formas de cultivo se han desarrollado biorreactores específicos que, por su transparencia, se los suele llamar también fotobiorreactores. Los cultivos de microalgas que recurren al desarrollo de un metabolismo mixotrófico, también emplean fotobiorreactores. En estos dispositivos para el cultivo, se destacan entre sus propiedades constructivas la de ser permeables a luz y alojar a los cultivos en espesores limitados, a fin de permitir la llegada de la fuente de energía al seno del cultivo.

Tipos y modos de operación de fotobiorreactores

Se han desarrollado una gran diversidad de fotobiorreactores, que pueden inicialmente agruparse en «abiertos» o «cerrados» (Ramírez-Mérida *et al.*, 2015). Esta clasificación resuelve la manera en que los cultivos se vinculan con el ambiente. Los sistemas abiertos lo hacen de manera directa, sin interponer barreras filtrantes que permitan una selectividad en el intercambio de masa. Estos son, en general, los fotobiorreactores más económicos de construir, operar y mantener (Borowitzka, 2005). El éxito de los cultivos desarrollados en estos sistemas depende fuertemente del empleo de medios de cultivo selectivos (que en función de la especie de microalga podrán ser alcalinos, hiperosmóticos, etc.) sumado a las condiciones autotróficas. Entre estos, los más comunes son los estanques agitados, construidos como un canal que permite una circulación del cultivo en su interior. En la figura 5 se muestra un esquema de este tipo de sistema de cultivo. El agitador mecánico, tipo rueda de paletas, mantiene al cultivo en agitación, permitiendo la incorporación del CO₂ atmosférico y la remoción del oxígeno fotosintético.

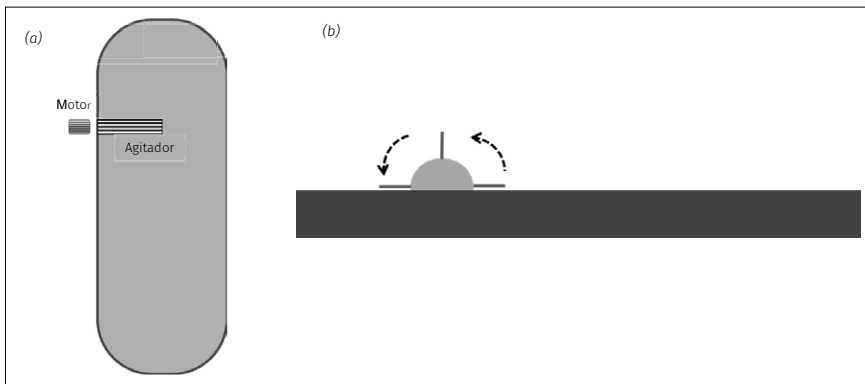


Figura 9.5. Esquemas de un estanque agitado, tipo raceway pond: A) vista superior. B) vista lateral. La circulación del cultivo (en el esquema A) y del impulsor (en el esquema B), se representa con flechas de trazo discontinuo.

Por otra parte, los fotobiorreactores cerrados, que autores como Behrens (2005) reservan exclusivamente el término fotobiorreactor a los sistemas cerrados, se diseñan de tal manera que para el intercambio gaseoso entre el cultivo y el ambiente atmosférico se interponen filtros hidrofóbicos. Estos hacen de barrera para la contención de los microorganismos del cultivo, evitando su diseminación en el ambiente, y también esterilizando la corriente

de gases que se introducen al cultivo. El empleo de filtros, lleva entonces a adoptar otras formas constructivas de los fotobiorreactores. Entre las diferentes alternativas, como pueden consultarse en algunos trabajos de revisión (Zittelli *et al.*, 2013; Carvalho *et al.*, 2006; Behrens, 2005) aparecen los paneles y los tubulares. En la figura 9.6 se presentan esquemas de estos dispositivos.

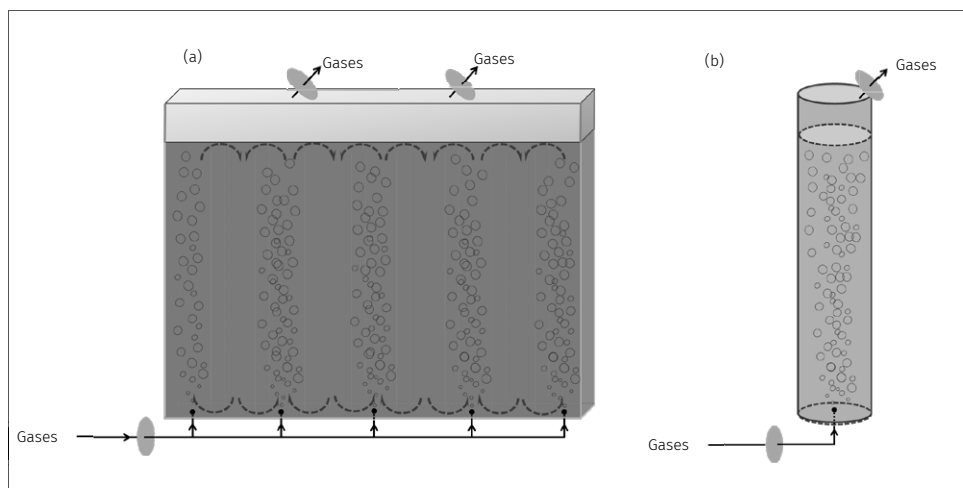


Figura 9.6. Esquemas de fotobiorreactores cerrados. A) Panel, tipo air-lift. B) Tubular tipo columna de burbujeo. El gas inyectado y evacuado atraviesa filtros esterilizantes (representados en los dibujos como óvalos grises), ingresando al fotobiorreactor en forma de burbujas. La circulación del cultivo (en el esquema A) se representa con flechas de trazo discontinuo.

Los paneles funcionan normalmente con agitación neumática y, dependiendo de los arreglos internos que tuviese el fotobiorreactor, podrán ser columnas de burbujeo o *air-lift*. Los tubulares, pueden funcionar también con ese tipo de agitación, como columnas de burbujeo o *air-lift*, o con agitación hidráulica. Para esto último, una bomba mecánica recircula el cultivo para cumplir con el tiempo de residencia necesario (Sirohi *et al.*, 2022). Los reactores tubulares pueden disponerse en diferentes arreglos para aprovechar mejor el espacio y extenderse en longitud (Carvalho *et al.*, 2006).

En todos los fotobiorreactores el espesor y densidad de los cultivos es una variable a considerar para optimizar la penetración de la fuente de energía.

Si bien hay tipos de fotobiorreactores más adecuados para cada aplicación, todos permiten la operación en modo *batch*, *fed-batch* y continua; el uso de aire enriquecido en CO₂ y la incorporación de diferentes instrumentos para la monitorización de los cultivos y la automatización del bioproceso.

**APLICACIONES DEL MODELADO METABÓLICO EN LA
OPTIMIZACIÓN DE BIOPROCESOS CON MICROALGAS:
ETAPAS GENERALES PARA LA CONSTRUCCIÓN DE MODELOS
A ESCALA GENÓMICA**

Los modelos metabólicos a escala genómica describen el metabolismo de un organismo, integrando la información biológica (genética, bioquímica, fenotípica, etc.), obtenida de diferentes fuentes, como bases de datos, experimentos y otros modelos (Aite *et al.*, 2013; Gudmundsson *et al.*, 2017). El proceso de elaboración de ese modelo es iterativo, no completamente automatizado (Juneja *et al.*, 2016) y requiere del uso de herramientas y plataformas computacionales, y de la toma de decisiones particulares relacionadas al organismo para el cual se está tratando de construir el modelo. Estos modelos son de utilidad no solo para comprender el metabolismo, sino también para realizar predicciones e introducir mejoras racionales a los procesos de obtención de bioproductos a través de la ingeniería genética y metabólica. Esto ha sido demostrado en varias publicaciones en las que se han reconstruido modelos metabólicos de microalgas eucariotas como *Chlamydomonas* (Boyle y Morgan 2009) o *Chlorella* (Juneja *et al.*, 2016), y en cianobacterias como *Synechocystis* (Shastri y Morgan, 2005) y *Arthrospira* (Klanchui *et al.*, 2012), consideradas modelos, entre otras.

Tabla 9.1. Estudios de modelado metabólico a escala genómica en microalgas

Género de microalga	Referencias
<i>Chlamydomonas</i>	Boyle y Morgan, 2009 Gomes de Oliveira Dal’Molin <i>et al.</i> , 2011 Imam <i>et al.</i> , 2015
<i>Chlorella</i>	Zuñiga <i>et al.</i> , 2016 Juneja <i>et al.</i> , 2016
<i>Nannochloropsis</i>	Loira <i>et al.</i> , 2017 GOShah <i>et al.</i> , 2017
<i>Phaeodactylum</i>	Kim <i>et al.</i> , 2016 Levering <i>et al.</i> , 2016
<i>Tetraselmis</i>	Lim <i>et al.</i> , 2017
<i>Monoraphidium</i>	Bogen <i>et al.</i> , 2013
<i>Ostreococcus</i>	Krumholz <i>et al.</i> , 2012
<i>Synechocystis</i>	Nogales <i>et al.</i> , 2012 Knoop <i>et al.</i> , 2013
<i>Synechococcus</i>	Broddrick <i>et al.</i> , 2016 Hendry <i>et al.</i> , 2016

En la Tabla 9.1 se muestran, a modo de ejemplos destacados, casos de modelados metabólicos a escala genómica de microalgas. Muchos de ellos han sido ampliamente validados y usados como plantillas para organismos relacionados. Para la construcción de un modelo metabólico a escala genómica, las actividades se estructuran en una serie de etapas (Thiele y Palsson, 2010; Gudmundsson *et al.*, 2017). Estas son: i) construir un borrador inicial de la red metabólica, ii) curaduría y refinamiento manual (o parcialmente automatizado) del modelo, iii) conversión a un formato matemático, y iv) verificación y evaluación o validación.

La construcción del borrador inicial de la red metabólica para elaborar un modelo a escala genómica, implica la obtención del genoma completo anotado, es decir la información relacionada a las secuencias que codifican para cada gen y la función que se le atribuye en ese organismo. En esta etapa se obtiene una colección de funciones metabólicas de ese organismo, y esta colección puede complementarse con la información de bases de datos bioquímicas. Muchas veces se emplean también modelos metabólicos de organismos relacionados y previamente publicados, a modo de plantillas. Para la construcción de este borrador inicial se pueden usar recursos computacionales que emplean diferentes lenguajes como por ejemplo ModelSEED (Henry *et al.*, 2010); RAVEN 2.0 (Wang *et al.*, 2018), entre otras. Por ejemplo, para la construcción del borrador inicial de la red metabólica de *Chlorella variabilis* NC64A, Juneja *et al.* (2016), descargaron los datos de secuencia del genoma completo de esta microalga, que está disponible en la base de datos JGI (<https://genome.jgi.doe.gov/portal/>).

En el caso de las microalgas, como en otros organismos, la asignación de funciones a ciertos genes se realiza por homología con los de organismos para los que dicha información ya se encuentra confirmada. Mucha de la información obtenida en ese primer borrador debería ser confirmada y/o refinada mediante evidencia experimental o datos de transcriptómica, de regulación, y mediante información que puede ser incluida. Además, las anotaciones de los genomas suelen ser revisadas y actualizadas. Por otra parte, no toda la información obtenida del genoma será de utilidad a los fines de la reconstrucción metabólica. Estas son algunas de las razones que motivan la segunda etapa que consiste en la curaduría y refinamiento de la red metabólica. Es necesario revisar todas las reacciones metabólicas incluidas en el primer borrador en lo que respecta a balances de masa y carga, asociación de genes a funciones, direccionalidad, compartimentación intracelular, etc. La compartimentación intracelular de las reacciones requiere de la información de localización intracelular de las enzimas involucradas en la vía. En el caso de las microalgas, una particularidad es la existencia de

cloroplastos, que no se encuentran en organismos no fotosintéticos. De hecho, más del 30 % de las reacciones en la red metabólica de *Chlamydomonas* están localizadas en los cloroplastos (Chang *et al.*, 2011).

En esta etapa también se revisan otros aspectos como el uso de cofactores, que puede ser un organismo específico. Para la curaduría y refinamiento se emplean servidores y plataformas como KEGG (Kanehisa y Goto, 2000), BRENDA (Schomburg *et al.*, 2002), MetaCyc (Caspi *et al.*, 2012), Expasy (Duvaud *et al.*, 2021) o BiGG models (King *et al.*, 2016). Esta última plataforma ya contiene varios modelos publicados, incluyendo algunos que corresponden a microalgas eucariotas modelo y cianobacterias. Entre otras plataformas específicas para microalgas eucariotas o cianobacterias, se encuentran Cyanobase (Nakao *et al.*, 2010), ChlamyCyc (May *et al.*, 2009) y las bases de datos Cyanocyc y Algae, incluidas en la plataforma BioCyc (Karp *et al.*, 2019).

En algunos casos, durante el proceso de reconstrucción de la red metabólica se suelen identificar «gaps» en las vías, enzimas que faltan, porque en algunos casos los genes están presentes, pero no tienen la función asignada, es decir no se encuentran anotados como tales. Ese análisis aplicado a la red metabólica de *Chlamydomonas reinhardtii* permitió encontrar funciones para genes que no estaban identificadas con anterioridad, encontrando nuevas vías o enzimas no conservadas (Boyle y Morgan, 2009).

La reacción de formación de la biomasa debe ser incluida a la red y constituye una función objetivo. Para esto se requiere conocer o aproximar la composición de la biomasa (en términos de porcentajes de proteínas, lípidos, carbohidratos, ADN y ARN) con el mayor detalle posible, y esto se determina experimentalmente (Thiele y Palsson, 2010). Para algunos organismos no modelados, puede faltar alguna información específica, pero la misma puede aproximarse a partir de uno filogenéticamente cercano. En los casos de las microalgas, se suele acudir a los de los géneros enumerados en la Tabla 9.1.

La red consiste en una lista de reacciones estequiométricamente balanceadas. La etapa de conversión a un lenguaje matemático es una tarea automatizada e implica la conversión de esas reacciones en una matriz estequiométrica **S**, en la que las filas están constituidas por los coeficientes de los metabolitos y las columnas contienen los coeficientes estequiométricos de las reacciones incluidas en el modelo.

Se deben definir límites para cada sistema añadiendo una reacción de intercambio para todos los metabolitos que pueden ser consumidos o secretados por el organismo que se está modelando. Además, se añadirán restricciones a la reconstrucción, convirtiéndola así en un modelo específico para cada condición. Si se asumen condiciones estacionarias, en las que la concentración de metabolitos internos permanece constante, existe un vector

de flujos de reacción. $\mathbf{v}$ tal que $\mathbf{S}\mathbf{v}=\mathbf{0}$. Una de las herramientas más empleadas en esta etapa es COBRA toolbox.

Uno de los enfoques computacionales más usados para el estudio de las redes metabólicas es el análisis de balance de flujos (FBA, que deriva de *Flux Balance Analysis*), que permite simular los flujos intracelulares en el contexto de la red metabólica. FBA fue utilizado para la predicción de la distribución de flujos metabólicos en condiciones de crecimiento auto, mixto y heterotrófico en *Chlamydomonas reinhardtii* por Boyle y Morgan (2009). La reconstrucción usada por estos autores incluyó las vías del metabolismo primario (por ejemplo, las de fotofosforilación, glicólisis, gluconeogénesis, ciclo de los ácidos tricarboxílicos, fotorrespiración, síntesis de aminoácidos, lípidos, polisacáridos, entre otras relevantes). En este caso, el metabolismo primario, la red consistió en 458 metabolitos y 484 reacciones. La red metabólica fue usada para el desarrollo del modelo estequiométrico, en la forma de la matriz $\mathbf{S}$ y el vector de flujos. Se aplicó análisis del balance de flujos acoplado a la programación lineal para estimar los flujos intracelulares en cada condición de crecimiento. Estos resultados y los datos experimentales obtenidos por estos autores demostraron ser de utilidad para seleccionar las condiciones más adecuadas para el crecimiento y la obtención de productos en estas microalgas.

Más adelante, Chang *et al.* (2011) obtuvieron una reconstrucción metabólica más detallada para *C. reinhardtii*, que consistió en 2190 reacciones y 1068 metabolitos. Esta nueva reconstrucción permitió explicar la actividad de un número mayor de genes involucrados en el metabolismo respecto a la anterior. En este trabajo se incluye además una intensiva verificación transcriptómica y experimental, que aportan evidencias novedosas para la validación de la red metabólica de esta microalga.

La secuencia del genoma de *Chlorella variabilis* NC64A, junto con las bases de datos como KEGG, y la reconstrucción metabólica de una microalga relacionada como es *Chlamydomonas* (Chang *et al.*, 2011) fueron usados en el proceso de curaduría de la red metabólica de *C. variabilis* por Juneja *et al.* (2016). Estos autores obtuvieron un modelo que incluye 1455 reacciones y 1236 metabolitos, y realizaron el análisis de la red por FBA. Los autores utilizaron el modelo para predecir las velocidades de crecimiento en diferentes condiciones, que luego midieron de manera experimental. Si bien las velocidades predichas fueron mayores a las medidas, fue posible identificar cuáles podrían haber sido las limitaciones del modelo que condujeron a las deficiencias en las predicciones.

Un modelo a escala metabólica para *Chlorella vulgaris* fue obtenido por Zuñiga *et al.* (2016), utilizando los datos de secuencia genómica para esta

microalga y la metodología basada en la curación del borrador, como se mencionó anteriormente. El modelo obtenido incluyó 2294 reacciones y 1770 metabolitos, su funcionalidad se validó simulando el crecimiento en diferentes condiciones mediante el FBA. Los autores informaron que las velocidades obtenidas por la simulación fueron consistentes con las que se obtuvieron de manera experimental.

Las actividades de las enzimas que componen las redes metabólicas están sujetas a regulaciones tanto a nivel de la expresión génica como alostérico. Los fenómenos regulatorios son complejos, interconectan diferentes vías metabólicas y están vinculados con las condiciones del entorno. Estas condiciones son dinámicas puesto que las concentraciones de sustratos y metabolitos, el pH y otros parámetros del entorno celular fluctúan, dando como resultado heterogeneidades espaciales y temporales en un bioproceso, que inducen adaptaciones en las células a diferentes escalas de tiempo. Las magnitudes de esas fluctuaciones intracelulares dependen a su vez de factores externos a la célula como los fenómenos de mezcla y transferencia que en biorreactores dependen de parámetros diseño y operación (Wang *et al.*, 2020). Si los modelos pueden contemplar esas heterogeneidades espaciales y temporales en los biorreactores, las predicciones que puedan realizarse con esos modelos serán mejoradas. En los cultivos de microalgas, donde el crecimiento puede ser auto o mixotrófico, el modelado metabólico debe contemplar el hecho de que las fuentes de carbono y energía son diferentes entre sí, por lo que suele diferenciarse de los modelos estándar de crecimiento heterotrófico que ocurren en otro tipo de organismos. La luz, como fuente de energía, representa una complejidad extra, ya que la intensidad con la que llega a cada célula disminuye, conforme la densidad del cultivo aumenta, lo que contribuye a la heterogeneidad temporal del sistema.

El FBA implica la aplicación de programación lineal para analizar los flujos metabólicos en la red bajo condiciones balanceadas en cuanto a los metabolitos, para lo que deben hacerse dos supuestos. Uno de ellos es que las células están en estado estacionario. Este supuesto desestima cualquier cambio transitorio en los metabolitos y es el que permite asumir que $Sv=0$, como ya se explicó. Esto equivale a asumir que el cultivo en el biorreactor existe teóricamente en un estado estacionario, lo cual debe contemplarse como una idealización.

El otro supuesto es que todas las células están siguiendo el mismo objetivo mientras crecen (por ejemplo, la producción de biomasa), lo que implica una disposición óptima de los flujos para ese fin. Esto puede no ser correcto en particular para las microalgas, cuya composición bioquímica puede variar dependiendo de las condiciones ambientales. Específicamente para el crecimiento autotrófico en cultivos con fotoperíodos, existirán variaciones en

la composición de la biomasa en los períodos de luz y de oscuridad, los objetivos de las células serán diferentes en estos períodos, por ejemplo. El objetivo del metabolismo celular alterna entre períodos de luz y oscuridad, entre la maximización de la absorción de luz y la movilización de compuestos de reserva. El FBA es un método muy versátil para resolver el cálculo de los flujos, útil para un análisis estándar, pero no se ajusta adecuadamente al entorno real en las microalgas, que representan sistemas dinámicos.

Uno de los enfoques que abordan el supuesto de los cambios fenotípicos que implican el cultivo en condiciones fluctuantes de luz y oscuridad (entre otras fluctuaciones de los cultivos), es el análisis de balance de flujos dinámico (dFBA). Este enfoque trata de la resolución del sistema, como un problema de programación no lineal. Barouk *et al.* (2014) presentaron un marco teórico al que denominaron DRUM (*Dynamic Reduction of Unbalanced Metabolism*) adecuado para procesos intracelulares en los que ocurre acumulación de metabolitos. En este enfoque, la red completa debe agruparse en subredes, en las que puede asumirse un estado cuasiestacionario, que se reducen luego a un conjunto de reacciones macroscópicas para las cuales se define posteriormente su cinética. Se obtiene un sistema de ecuaciones diferenciales, que representa la evolución a escala macroscópica del bioproceso, así como los procesos intracelulares y la acumulación de metabolitos.

Este enfoque fue aplicado particularmente para analizar una red metabólica reducida, incluyendo las vías centrales del carbono en la microalga *Tysochrisis lutea* cultivada en un biorreactor tubular con fotoperíodo. Esta aplicación resultó en un modelo que permitió explicar los datos experimentales obtenidos en términos de acumulación de lípidos y carbohidratos en el cultivo de la microalga sometido a ciclos de luz y oscuridad.

En síntesis, un modelo metabólico consiste en una red de reacciones bioquímicas interconectadas, estequiométricamente balanceadas, representadas en una matriz en la que los metabolitos constituyen las filas y las reacciones las columnas. La suma de los flujos individuales de los metabolitos de que componen la biomasa, permite la predicción de tasas de crecimiento, en las condiciones en las que esos flujos fueron calculados. Si bien las microalgas son capaces de acumular diversos productos de interés biotecnológico, su potencial para acumular ácidos grasos en condiciones autotróficas ha sido un notable motor para el desarrollo de modelos metabólicos a escala genómica en estos organismos. El análisis de esos modelos permite la predicción de datos útiles en la optimización de cepas y condiciones de cultivo. En estos organismos fotosintéticos, uno de los principales inconvenientes radica en el hecho de que los supuestos para la aplicación de los métodos menos complejos y más difundidos para el cálculo de los flujos intracelulares,

son hipótesis que difícilmente se ajustan a la realidad. Una alternativa para superar este desafío, es aplicar estrategias como dFBA, más adecuadas para microorganismos sometidos a condiciones fluctuantes de su entorno, como son la limitación de nutrientes y fotoperíodos, en el caso de las microalgas.

De la experiencia aprendida se desprende, a modo de perspectiva a futuro, que el creciente conocimiento a nivel de genomas, funcionalidad y regulación de genes, enzimas y vías metabólicas, y el desarrollo de plataformas computacionales avanzadas que permitan simplificar el cálculo complejo, permitirán profundizar la comprensión sobre el metabolismo, mejorar la posibilidad de predecir su comportamiento y optimizar las condiciones de cultivo de estos microorganismos de particular interés biotecnológico.

REFERENCIAS BIBLIOGRÁFICAS

- AITE, M.; CHEVALLIER, M. (...) Cortés, M. (2018). Traceability, reproducibility and wiki-exploration for «à-la-carte» reconstructions of genome-scale metabolic models. *PLoS Computational Biology*, 14 (5): e1006146. <https://doi.org/10.1371/journal.pcbi.1006146>
- ATTEIA, A.; VAN LIS, R. (...) MARTIN, W. (2013). Anaerobic energy metabolism in unicellular photosynthetic eukaryotes. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1827 (2), 210–223. <https://doi.org/10.1016/j.bbabi.2012.08.002>
- BAROUKH, C.; MUÑOZ-TAMAYO, R. (...) BERNARD, O. (2014). DRUM: a new framework for metabolic modeling under non-balanced growth. Application to the carbon metabolism of unicellular microalgae. *PLoS One*, 9 (8), e104499. <https://doi.org/10.1371/journal.pone.0104499>
- BEARDALL, J. & RAVEN, J. (2016). Carbon acquisition by microalgae. En *The Physiology of Microalgae, Developments in Applied Phycology* 6, Borowitzka, M., et al. (Eds.). Springer. https://doi.org/10.1007/978-3-319-24945-2_4
- BEHRENS, P. (2005). Photobioreactors and Fermentors: The Light and Dark Sides of Growing Algae. En *Algal Culturing Techniques*. Andersen, R. A. (Ed.). Elsevier Academic Press.
- BOGEN, C.; AL-DILAIMI, A. (...) RUPP, O. (2013). Reconstruction of the lipid metabolism for the microalga *Monoraphidium neglectum* from its genome sequence reveals characteristics suitable for biofuel production. *BMC genomics*, 14, 1–18. <https://doi.org/10.1186/1471-2164-14-926>
- BOROWITZKA, M. A. (2005). Culturing Microalgae in Outdoor Ponds. En Andersen, R. A. (Ed.). *Algal Culturing Techniques*. Elsevier Academic Press.

- BOYLE, N. & MORGAN, J.** (2009). Flux balance analysis of primary metabolism in *Chlamydomonas reinhardtii*. *BMC Systems Biology*, 3, 1–14. <https://doi.org/10.1186/1752-0509-3-4>
- BRODDRICK, J.; RUBIN, B. (...) DIAMOND, S.** (2016). Unique attributes of cyanobacterial metabolism revealed by improved genome-scale metabolic modeling and essential gene analysis. *Proceedings of the National Academy of Sciences*, 113 (51), E8344–E8353. <https://doi.org/10.1073/pnas.1613446113>
- BUSI, M.; BARCHIESI, J. (...) GOMEZ-CASATI, D.** (2014). Starch metabolism in green algae. *Starch–Stärke*, 66 (1–2), 28–40. <https://doi.org/10.1002/star.201200211>
- CARVALHO, A.; MEIRELES, L. & MALCATA, F.** (2006). Microalgal Reactors: A Review of Enclosed System Designs and Performances. *Biotechnology Progress*, 22 (6): 1490–1506. <https://aiche.onlinelibrary.wiley.com/doi/full/10.1021/bp060065r>
- CASPI, R.; ALTMAN, T. (...) KESELER, I.** (2012). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*, 40 (D1), D742–D753. <https://doi.org/10.1093/nar/gkr1014>
- CATALANOTTI, C.; YANG, W. (...) GROSSMAN, A.** (2013). Fermentation metabolism and its evolution in algae. *Frontiers in Plant Science*, 4, 150. <https://doi.org/10.3389/fpls.2013.00150>
- CHANG, R.; GHAMSARI, L. (...) FU, W.** (2011). Metabolic network reconstruction of *Chlamydomonas* offers insight into light-driven algal metabolism. *Molecular Systems Biology*, 7 (1), 518. <https://doi.org/10.1038/msb.2011.52>
- CHEN, C.; ZHAO, X. (...) CHANG, J.** (2013). Microalgae-based carbohydrates for biofuel production. *Biochemical Engineering Journal*, 78, 1–10. <https://doi.org/10.1016/j.bej.2013.03.006>
- CHEN, H. & WANG, Q.** (2021). Regulatory mechanisms of lipid biosynthesis in microalgae. *Biological Reviews*, 96 (5): 2373–2391. <https://doi.org/10.1111/brv.12759>
- DUVAUD, S.; GABELLA, C. (...) DURINX, C.** (2021). Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Research*, 49 (W1), W216–W227. <https://doi.org/10.1093/nar/gkab225>
- EDMUNDSON, S. & HUESEMAN, M.** (2015). The dark side of algae cultivation: characterizing night biomass loss in three photosynthetic algae. *Chlorella sorokiniana*, *Nannochloropsis salina* and *Picochlorum* sp. *Algal Research*, 12, 470–476. <https://doi.org/10.1016/j.algal.2015.10.012>
- FABRIS, M.; MATTHIJS, M. (...) BAART, G.** (2012). The metabolic blueprint of *Phaeodactylum tricornutum* reveals a eukaryotic Entner–Doudoroff glycolytic pathway. *The Plant Journal*, 70 (6), 1004–1014. <https://doi.org/10.1111/j.1365-313X.2012.04941.x>
- FAN, J.; ANDRE, C. & XU, C.** (2011). A chloroplast pathway for the de novo biosynthesis of triacylglycerol in *Chlamydomonas reinhardtii*. *FEBS Letters* 585: 1985–1991. <https://doi.org/10.1016/j.febslet.2011.05.018>

- GAFFRON, H. & RUBIN, J.** (1942). Fermentative and photochemical production of hydrogen in algae. *The Journal of General Physiology*, 26 (2), 219–240. <https://doi.org/10.1085/jgp.26.2.219>
- GOMES DE OLIVEIRA DAL'MOLIN, C.; QUEK, L. (...) PALFREYMAN, R.** (2011). Alga-GEM– a genome–scale metabolic reconstruction of algae based on the *Chlamydomonas reinhardtii* genome. *BMC Genomics* 12 (Suppl 4), S5. <https://doi.org/10.1186/1471-2164-12-S4-S5>
- GUDMUNDSSON, S.; AGUDO, L. & NOGALES, J.** (2017). Applications of genome–scale metabolic models of microalgae and cyanobacteria in biotechnology. En *Microalgal–Based Biofuels and Bioproducts*, Gonzalez–Fernandez, C., Muñoz, R. (Eds). Woodhead Publishing. <https://doi.org/10.1016/B978-0-08-101023-5.00004-2>
- GUÉGUEN, N.; LE MOIGNE, D. (...) MARÉCHAL, E.** (2021). Lipid Droplets in Unicellular Photosynthetic Stramenopiles. *Frontiers in Plant Science*, 12: 639276. <https://doi.org/10.3389/fpls.2021.639276>
- HELLEBUST, J. & AHMAD, I.** (1989). Regulation of nitrogen assimilation in green microalgae. *Biological oceanography*, 6 (3–4), 241–255. <https://doi.org/10.1080/01965581.1988.10749529>
- HENDRY, J.; PRASANAN, C. (...) WANGIKAR, P.** (2016). Metabolic model of *Synechococcus* sp. PCC 7002: Prediction of flux distribution and network modification for enhanced biofuel production. *Bioresource Technology*, 213, 190–197. <https://doi.org/10.1016/j.biortech.2016.02.128>
- HENRY, C.; DEJONGH, M. (...) STEVENS, R.** (2010). High–throughput generation, optimization and analysis of genome–scale metabolic models. *Nature Biotechnology*, 28 (9), 977–982. <https://doi.org/10.1038/nbt.1672>
- IMAM, S.; SCHÄUBLE, S. (...) BALIGA, N.** (2015). A refined genome–scale reconstruction of *Chlamydomonas* metabolism provides a platform for systems–level analyses. *The Plant Journal*, 84 (6), 1239–1256. <https://doi.org/10.1111/tpj.13059>
- INOKUCHI, R.; KUMA, K. (...) OKADA, M.** (2002). Nitrogen–assimilating enzymes in land plants and algae: phylogenic and physiological perspectives. *Physiologia Plantarum*, 116 (1), 1–11. <https://doi.org/10.1034/j.1399-3054.2002.1160101.x>
- INUI, H.; MIYATAKE, K. (...) KITAOKA, S.** (1982). Wax ester fermentation in *Euglena gracilis*. *FEBS Letters*, 150 (1), 89–93. [https://doi.org/10.1016/0014-5793\(82\)81310-0](https://doi.org/10.1016/0014-5793(82)81310-0)
- JEON, S.; JEONG, B. & CHANG, Y.** (2017). Chemicals and fuels from microalgae. En *Consequences of microbial interactions with hydrocarbons, oils, and lipids: production of fuels and chemicals*, S.Y. Lee (Ed.). Springer International Publishing. https://doi.org/10.1007/978-3-319-50436-0_384
- JUNEJA, A.; CHAPLEN, F. & MURTHY, G.** (2016). Genome scale metabolic reconstruction of *Chlorella variabilis* for exploring its metabolic potential for biofuels. *Bioresource Technology*, 213: 103–10. <https://doi.org/10.1016/j.biortech.2016.02.118>

- KANEHISA, M. & GOTO, S.** (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28 (1), 27–30. <https://doi.org/10.1093/nar/28.1.27>
- KARP, P.; BILLINGTON, R. (...) KOTHARI, A.** (2019). The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics*, 20 (4), 1085–1093. <https://doi.org/10.1093/bib/bbx085>
- KIM, D. & KIM, M.** (2011). Hydrogenases for biological hydrogen production. *Bioresource Technology*, 102 (18), 8423–8431. <https://doi.org/10.1016/j.biortech.2011.02.113>
- KIM, J.; FABRIS, M. (...) LUN, D.** (2016). Flux balance analysis of primary metabolism in the diatom *Phaeodactylum tricornutum*. *The Plant Journal*, 85: 161–176. <https://doi.org/10.1111/tpj.13081>
- KING, Z., LU, J. (...) LEWIS, N.** (2016). BiGG Models: A platform for integrating, standardizing, and sharing genome-scale models. *Nucleic Acids Research* 44 (D1): D515–D522. <https://doi.org/10.1093/nar/gkv1049>
- KLANCHUI, A.; KHANNAPHO, C. (...) PHODEE, A.** (2012). iAK692: A genome-scale metabolic model of *Spirulina platensis* C1. *BMC Systems Biology*, 6, 1–15. <https://doi.org/10.1186/1752-0509-6-71>
- KNOOP, H.; ZILLIGES, Y. (...) STEUER, R.** (2010). The metabolic network of *Synechocystis* sp. PCC 6803: systemic properties of autotrophic growth. *Plant Physiology*, 154 (1), 410–422. <https://doi.org/10.1104/pp.110.157198>
- KREUZBERG, K.** (1984). Starch fermentation via a formate producing pathway in *Chlamydomonas reinhardtii*, *Chlorogonium elongatum* and *Chlorella fusca*. *Physiologia Plantarum*, 61: 87–94. <https://doi.org/10.1111/j.1399-3054.1984.tb06105.x>
- KRUMHOLZ, E.; YANG, H. (...) LIBOUREL, I.** (2012). Genome-wide metabolic network reconstruction of the picoalga *Ostreococcus*. *Journal of Experimental Botany*, 63 (6), 2353–2362. <https://doi.org/10.1093/jxb/err407>
- LENKA, S.; CARBONARO, N. (...) LI, Y.** (2016). Current advances in molecular, biochemical, and computational modeling analysis of microalgal triacylglycerol biosynthesis. *Biotechnology Advances*, 34 (5): 1046–1063. <https://doi.org/10.1016/j.biotechadv.2016.06.004>
- LEVERING, J.; BRODDRICK, J. (...) BEERI, K.** (2016). Genome-Scale Model Reveals Metabolic Basis of Biomass Partitioning in a Model Diatom. *PLOS ONE*, 11 (5): e0155038. <https://doi.org/10.1371/journal.pone.0155038>
- LI, Y.; HAN, D. (...) HU, Q.** (2013). Molecular and Cellular Mechanisms for Lipid Synthesis and Accumulation in Microalgae: Biotechnological Implications. En *Handbook of Microalgal Culture Applied Phycology and Biotechnology* (2da. Ed.). Richmond, A., Hu, Q. (Eds.). Wiley-Blackwell. <https://doi.org/10.1002/9781118567166.ch28>
- LI, Y.; HORSMAN, M. (...) DUBOIS-CALERO, N.** (2008). Biofuels from microalgae. *Biotechnology Progress*, 24 (4), 815–820. <https://doi.org/10.1021/bp070371k>

- LIM, D.; SCHUHMANN, H. (...) AGUILERA, F.** (2017). RNA-Seq and metabolic flux analysis of *Tetraselmis* sp. M8 during nitrogen starvation reveals a two-stage lipid accumulation mechanism. *Bioresource Technology*, 244, 1281–1293. <https://doi.org/10.1016/j.biortech.2017.06.003>
- LOIRA, N.; MENDOZA, S. (...) GENOVA, A.** (2017). Reconstruction of the microalga *Nannochloropsis salina* genome-scale metabolic model with applications to lipid production. *BMC Systems Biology*, 11, 66. <https://doi.org/10.1186/s12918-017-0441-1>
- MAY, P.; CHRISTIAN, J. (...) WALTHER, D.** (2009). ChlamyCyc: an integrative systems biology database and web-portal for *Chlamydomonas reinhardtii*. *Bmc Genomics* 10, 209. <https://doi.org/10.1186/1471-2164-10-209>
- MASOJÍDEK, J.; TORZILLO, G. & KOBLÍŽEK, M.** (2013). Photosynthesis in microalgae. En *Handbook of Microalgal Culture: Applied Phycology and Biotechnology* (second edition). Richmond, A. & Hu, Q. (Eds.). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118567166.ch2>
- MÜLLER, M.; MENTEL, M. (...) MARTIN, W.** (2012). Biochemistry and Evolution of Anaerobic Energy Metabolism in Eukaryotes. *Microbiology and Molecular Biology Reviews*, 76 (2), 444–495. <https://doi.org/10.1128/mmr.05024-11>
- NAKAO, M.; OKAMOTO, S. (...) SATO, S.** (2010). CyanoBase: the cyanobacteria genome database update 2010. *Nucleic Acids Research* 38, D379–D381. <https://doi.org/10.1093/nar/gkp915>
- NAZIR, Y.; HALIM, H. (...) SONG, Y.** (2020). Microalgae single cell oil. En *Handbook of Microalgae-Based Processes and Products. Fundamentals and Advances in Energy, Food, Feed, Fertilizer, and Bioactive Compounds*. Jacob-Lopes, E.; Manzoni Maroneze, M. (...) Queiroz Zepka, L. (Eds.). Elsevier. <https://doi.org/10.1016/B978-0-12-818536-0.00016-6>
- NEILSON, A. & LEWIN, R.** (1974). The uptake and utilization of organic carbon by algae: an essay in comparative biochemistry. *Phycologia*, 13 (3), 227–264. <https://doi.org/10.2216/i0031-8884-13-3-227.1>
- NOGALES, J.; GUDMUNDSSON, S. (...) THIELE, I.** (2012). Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proceedings of the National Academy of Sciences*, 109 (7), 2678–2683. <https://doi.org/10.1073/pnas.1117907109>
- PEREZ-GARCIA, O.; ESCALANTE, F. (...) BASHAN, Y.** (2011). Heterotrophic cultures of microalgae: metabolism and potential products. *Water Research*, 45 (1), 11–36. <https://doi.org/10.1016/j.watres.2010.08.037>
- RAMÍREZ-MÉRIDA, L.; ZEPKA, L. & JACOB-LOPES, E.** (2015). Current Status, Future Developments and Recent Patents on Photobioreactor Technology. *Recent Patents on Engineering*, 9 (2): 80–90. <https://dx.doi.org/10.2174/1872212109666150206235316>
- RAN, W.; WANG, H. (...) LAN, X.** (2019). Storage of starch and lipids in microalgae: Biosynthesis and manipulation by nutrients. *Bioresource Technology*, 291, 121894. <https://doi.org/10.1016/j.biortech.2019.121894>

- RAVEN, J. & BEARDALL, J.** (2016). Dark respiration and organic carbon loss. En Borowitzka, M. et al. (eds.). *The Physiology of Microalgae, Developments in Applied Phycology* 6. https://doi.org/10.1007/978-3-319-24945-2_6
- RAVEN, J.; BEARDALL, J. & QUIGG, A.** (2020). Light-Driven Oxygen Consumption in the Water-Water Cycles and Photorespiration, and Light Stimulated Mitochondrial Respiration. En Larkum, A.; Grossman, A. & Raven, J. (Eds). *Photosynthesis in Algae: Biochemical and Physiological Mechanisms. Advances in Photosynthesis and Respiration*, vol 45. Springer. https://doi.org/10.1007/978-3-030-33397-3_8
- RAVEN, J. & GIORDANO, M.** (2016). Combined nitrogen. The physiology of microalgae. En Borowitzka, M. et al. (eds.). *The Physiology of Microalgae, Developments in Applied Phycology* 6. <https://doi.org/10.2216/i0031-8884-13-3-227.1>
- ROSSI, F. & DE PHILIPPIS, R.** (2016). Exocellular Polysaccharides in Microalgae and Cyanobacteria: Chemical Features, Role and Enzymes and Genes Involved in Their Biosynthesis. En En Borowitzka, M. et al. (eds.). *The Physiology of Microalgae, Developments in Applied Phycology* 6. https://doi.org/10.1007/978-3-319-24945-2_21
- SCHOMBURG, I.; CHANG, A. & SCHOMBURG, D.** (2002). BRENDA, enzyme data and metabolic information. *Nucleic Acids Research*, 30 (1), 47–49. <https://doi.org/10.1093/nar/30.1.47>
- SCODELARO BILBAO, P.; SALVADOR, G. & LEONARDI, P.** (2017). Fatty Acids from Microalgae: Targeting the Accumulation of Triacylglycerides. En Fatty Acids. Catala, A. (Ed.). *InTech*. <http://dx.doi.org/10.5772/67482>
- SHAH, A.; AHMAD, A. (...) ALI, B.** (2017). Reconstruction and analysis of a genome-scale metabolic model of *Nannochloropsis gaditana*. *Algal Research*, 26, 354–364. <https://doi.org/10.1016/j.algal.2017.08.014>
- SHASTRI, A. & MORGAN, J.** (2005). Flux balance analysis of photoautotrophic metabolism. *Biotechnology Progress*, 21 (6), 1617–1626. <https://doi.org/10.1021/bp050246d>
- SKJÅNES, K.; KNUTSEN, G. (...) LINDBLAD, P.** (2008). H₂ production from marine and freshwater species of green algae during sulfur deprivation and considerations for bioreactor design. *International Journal of Hydrogen Energy*, 33 (2), 511–521. <https://doi.org/10.1016/j.ijhydene.2007.09.040>
- SIROHI, R.; PANDEY, A. (...) SIM, S.** (2022). Design and applications of photobioreactors– a review. *Bioresource Technology* 349, 126858. <https://doi.org/10.1016/j.biortech.2022.126858>
- STAL, L. & MOEZELAAR, R.** (1997). Fermentation in cyanobacteria. *FEMS Microbiology Reviews*, 21 (2), 179–211. <https://doi.org/10.1111/j.1574-6976.1997.tb00350.x>
- SU, Y.** (2021). Revisiting carbon, nitrogen, and phosphorus metabolisms in microalgae for wastewater treatment. *Science of the Total Environment*, 762, 144590. <https://doi.org/10.1016/j.scitotenv.2020.144590>

- SUZUKI, L. & JOHNSON, C.** (2001). Algae know the time of day: circadian and photoperiodic programs. *Journal of Phycology*, 37 (6), 933–942. <https://doi.org/10.1046/j.1529-8817.2001.01094.x>
- TERASHIMA, M.; SPECHT, M. (...) HIPPLER, M.** (2011). The chloroplast proteome: a survey from the *Chlamydomonas reinhardtii* perspective with a focus on distinctive features. *Current Genetics*, 57, 151–168. <https://doi.org/10.1007/s00294-011-0339-1>
- THEVENIEAU, F. & NICAUD, J.** (2013). Microorganisms as sources of oils. *OCL* 20, D603. <https://doi.org/10.1051/ocl/2013034>
- THIELE, I. & PALSSON, B.** (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, 5 (1), 93–121. <https://doi.org/10.1038/nprot.2009.203>
- TSEKOS, I.** (1999). The sites of cellulose synthesis in algae: diversity and evolution of cellulose-synthesizing enzyme complexes. *Journal of Phycology*, 35 (4), 635–655. <https://doi.org/10.1046/j.1529-8817.1999.3540635.x>
- VEGA, J.** (2020). Nitrogen and sulfur metabolism in microalgae and plants: 50 years of research. *Progress in Botany*, 81, 1–40. https://doi.org/10.1007/124_2018_26
- WANG, H.; MARCIŠAUSKAS, S. (...) HERMANSSON, D.** (2018). RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. *PLOS Computational Biology*, 14 (10): e1006541. <https://doi.org/10.1371/journal.pcbi.1006541>
- WANG, G.; HARINGA, C. (...) ZHANG, S.** (2020). Coupled metabolic-hydrodynamic modeling enabling rational scale-up of industrial bioprocesses. *Biotechnology and Bioengineering*, 117 (3), 844–867. <https://doi.org/10.1002/bit.27243>
- ZITTELLI, G. C.; BIONDI, N. (...) TREDICI, M. R.** (2013). Photobioreactors for Mass Production of Microalgae. En *Handbook of Microalgal Culture: Applied Phycology and Biotechnology* (2nd. ed.). Richmond, A. & Hu, Q. (Eds.). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118567166.ch13>
- ZUÑIGA, C.; LI, C-T. (...) MCCONNELL, B.** (2016). Genome-scale metabolic model for the green alga *Chlorella vulgaris* UTEX 395 accurately predicts phenotypes under autotrophic, heterotrophic, and mixotrophic growth Conditions. *Plant Physiology*, 172:589–602. <https://doi.org/10.1104/pp.16.00593>

SOBRE EL AUTOR

RODOLFO DONDO. Profesor adjunto en la Facultad de Bioquímica y Ciencias Biológicas de la Universidad Nacional del Litoral e Investigador Independiente del CONICET. Obtuvo el doctorado en Ingeniería Química en la Facultad de Ingeniería Química en el año 2000 y trabaja en el Grupo de Ingeniería de Procesos y Sistemas del Instituto de Desarrollo Tecnológico para la Industria Química (INTEC). Su interés de investigación se centra en el campo del modelado y la optimización de procesos, bioprocesos y de cadenas de suministro. Ha publicado más de 40 artículos en revistas científicas internacionales.

**OPTIMIZACIÓN
DE SISTEMAS
BIOTECNOLÓGICOS**
Rodolfo Dondo

C Á T E D R A

Los recientes avances tecnológicos han permitido la modelización del metabolismo celular como redes de interacciones entre moléculas, lo cual ha permitido el desarrollo de redes bioquímicas a escala genómica que permiten computar sus funciones fisiológicas. Por otra parte, el campo de la biotecnología está en constante evolución y presenta una perspectiva macroscópica de modelización. Este libro presenta un intento de reconciliar ambas escalas integrando el modelado macroscópico con el modelado a escala genómica utilizando no solo matemáticas sino también software de cálculo numérico y de optimización. El texto familiariza a los lectores con las capacidades de varios paquetes de software de cálculo y optimización para modelar, simular y optimizar bioprocesos. Se busca presentar con rigurosidad los conceptos matemáticos y las herramientas computacionales a estudiantes de las carreras relacionadas a la bioquímica y la biotecnología.

**UNIVERSIDAD
NACIONAL DEL LITORAL**