



UNIVERSIDAD NACIONAL DEL SUR

TESIS DOCTORAL EN INGENIERÍA QUÍMICA

ANÁLISIS, MODELAMIENTO Y OPTIMIZACIÓN DE CAMINOS METABÓLICOS

Jimena A. Di Maggio

Bahía Blanca

Argentina

2012

PREFACIO

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Doctor en Ingeniería Química de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra.

La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito del Departamento de Ingeniería Química durante el período comprendido entre el 04/09/2007 y 28/02/2012, bajo la dirección de la Doctora María Soledad Diaz, Profesora Adjunta del Departamento de Ingeniería Química de la Universidad Nacional del Sur y del Doctor Juan Carlos Díaz Ricci, Profesor Asociado de la Facultad de Bioquímica, Química y Farmacia de la Universidad Nacional de Tucumán.



UNIVERSIDAD NACIONAL DEL SUR
Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el .../.../....., mereciendo la calificación de

AGRADECIMIENTOS

A mis directores:

La Dra. María Soledad Díaz, por darme la oportunidad de ingresar al mundo de la investigación, por los conocimientos transmitidos, por el apoyo y la calidez brindados en todo momento.

El Dr. Juan Carlos Díaz Ricci, por brindarme sus conocimientos y por su compañía a lo largo de este trabajo.

A la Planta Piloto de Ingeniería Química, por ser el soporte institucional para realizar las tareas de la presente tesis. Al Consejo Superior de Investigaciones Científicas y Tecnológicas y a la Universidad Nacional del Sur por proveer los medios necesarios para llevar a cabo este trabajo.

A Diana Constenla, Clarisa Cicchelli, Pedro Forbito y Federico Delucchi por la colaboración brindada en el trabajo experimental llevado a cabo en PLAPIQUI.

A la Ing. Nora Perotti, Martín Díaz Alfaro y Daniel Etchegorry por los conocimientos brindados y la enorme colaboración en la realización de los ensayos de cultivos continuos llevados a cabo en PROIMI.

Al Dr. Aníbal Blanco por sus valiosas sugerencias.

A Nora, Any, Adriana, Elsa y Silvia R. por su calidez y por facilitar las tareas en el instituto.

A mis padres, por tanto amor, por haberme apoyado y acompañado siempre en mis proyectos, por darme todas las oportunidades y por los buenos valores transmitidos, los quiero con todo mi corazón.

A mis hermanas, por todos los momentos compartidos y su compañía indispensable, las quiero mucho.

A mi abuela, mis tíos y mis primos, por estar siempre presentes, los quiero.

A mis viejas amigas, Ana, Sole, Caro, Bar y Jor, por tantos años de amistad y por darme siempre su apoyo incondicional, las adoro.

A Vani y Ceci, mis amigas y compañeritas de oficina, por la ayuda brindada, los consejos y por hacer que el día a día sea muy divertido, las quiero.

A Espi y Tici, por la amistad y los momentos lindos compartidos.

A los amigos encontrados en PLAPIQUI: Alexis, Juani, Raquel, Belén, Ivana, Yani, Guille S., Gera, Nico y Andrés, compañeros de grandes momentos.

Al grupo de Procesos, por la calidez de cada uno de sus integrantes que hace que los días laborales sean más amenos.

Y a todos los que, de una u otra manera, hicieron posible este trabajo de tesis.

RESUMEN

El análisis y modificación de caminos metabólicos para mejorar las propiedades de las células usando herramientas genéticas actuales es un objetivo central en Ingeniería Metabólica. La idea básica es integrar la información biológica cuantitativa obtenida de diversas fuentes, especialmente de análisis genómicos, a modelos matemáticos para poder efectuar análisis de datos y adquisición de datos asistida por modelos, a la vez que se pretende sugerir con más efectividad modificaciones a realizar sobre los microorganismos para lograr sobreproducción de compuestos deseados (por ejemplo, biocombustibles o antibióticos), intersección de ciertas moléculas en intervenciones terapéuticas, entre otros.

En la presente tesis se desarrollan modelos cinéticos de redes metabólicas basados en primeros principios, en el marco de problemas de optimización. El objetivo final de proponer modificaciones en dichas rutas metabólicas para la maximización de la producción de un producto de interés biotecnológico, como un aminoácido o un biocombustible. En una primera etapa, se lleva a cabo un estudio de sensibilidad para determinar los principales parámetros a identificar, mediante el desarrollo y aplicación de técnicas de sensibilidad global. Dichas técnicas están basadas en el estudio de varianzas condicionales e incondicionales de las variables de estado. En una segunda etapa, se lleva a cabo el ajuste de los principales parámetros del modelo cinético de rutas metabólicas con datos *in vivo*, obtenidos en la bibliografía para un cultivo continuo de *E. coli*, frente a un pulso de glucosa. Las predicciones del modelo se comparan con otros conjuntos de datos de bibliografía, con resultados aceptables. Asimismo, se llevan a cabo

experiencias con dicho microorganismo en cultivos continuos y se obtienen datos experimentales a nivel bio-reactor, que se emplean para ajustar un modelo cinético de quimiostato. La estimación de parámetros se lleva a cabo formulando un problema de optimización con función objetivo de máxima verosimilitud, sujeto a un sistema de ecuaciones diferenciales algebraicas que representan la red metabólica y el bio-reactor, respectivamente. En la siguiente etapa, el modelo cinético se formula en el marco de un problema de programación no lineal, para el diseño de redes metabólicas que maximizan un objetivo biotecnológico. Finalmente, se formula un problema de diseño óptimo con restricciones que aseguran la estabilidad de la red obtenida, mediante la aplicación de conceptos de estabilidad de sistemas no lineales de Lyapunov.

La metodología presentada en esta tesis, unida a resultados numéricos que concuerdan con datos experimentales intra y extracelulares, permite el estudio sistemático de redes metabólicas y a la vez avanzar hacia la integración de modelos metabólicos complejos para el diseño óptimo de cepas.

ABSTRACT

The main objective in Metabolic Engineering is the analysis and modification of metabolic pathways to improve cell properties, using currently available genetic tools. The basic idea is to couple biological information from different sources, mainly genomic analysis, to mathematical models for computer aided data acquisition and analysis, as well as prediction of modifications on microorganisms to achieve biotechnological products overproduction.

In this thesis work, we formulate metabolic network kinetic models, based on first principles, within an optimization framework. The final objective is to propose modifications in metabolic pathways for aminoacid o bioethanol overproduction. As a first step, we perform global sensitivity analysis (GSA) on the developed metabolic network model to determine identifiable parameters. The applied methodology is a variance-based one, most appropriate for highly nonlinear models like the system under study. The second step includes parameter estimation for the kinetic model, based on *in vivo* collected data from the literature for a glucose pulse inserted in an *E. coli* continuous culture. Model predictions favorably compare to other sets of experimental data from the literature. Fermentations have been carried out under the same conditions for *E. coli* cultures and collected experimental data has been used for parameter estimation in a simple continuous bioreactor kinetic model. Parameter estimation is performed within an optimization framework with a maximum likelihood objective function, subject to a differential algebraic equations system representing the metabolic network or the bioreactor, alternatively.

In the last step, the kinetic model is formulated within a nonlinear programming program for the optimal design of a metabolic network for a desired product overproduction. The model is re-formulated including stability constraints based on Lyapunov stability theory to ensure asymptotic stability of the optimal metabolic network.

The methodology proposed in this thesis, together with numerical results that are in agreement with intra and extracellular experimental data, paves the way to the systematic study of metabolic pathways and the integration of complex metabolic network models for the optimal design of strains of biotechnological interest.

ÍNDICE DE CONTENIDOS

Índice de Figuras	ix
Índice de Tablas	xii

Capítulo 1

Introducción general

1.1. Historia de la biotecnología. Surgimiento de nuevas disciplinas	1
1.1.1. Ingeniería Metabólica	4
1.2. Herramientas computacionales. Modelos y bases de datos	10
1.3. <i>Escherichia coli</i> . Metabolismo central del carbono	17
1.3.1. <i>Escherichia coli</i> . Características	17
1.3.2. Metabolismo central del carbono en <i>Escherichia coli</i>	17
1.3.2.1. Glucólisis	18
1.3.2.2. Ruta de las pentosas fosfato	21
1.3.2.3. Rutas fermentativas	24
1.4. Objetivos de la tesis	27
1.5. Estructura de la tesis	28

Capítulo 2

Modelado de la red metabólica de *Escherichia coli*

2.1. Introducción	30
2.2. Modelado de la red metabólica de <i>Escherichia coli</i>	34
2.2.1 Modelo cinético del metabolismo central del carbono	36
2.2.1.1 Balances de masa	36
2.2.1.2. Cinéticas de reacción	39
2.2.1.2.1. Sistema fosfotransferasa	40
2.2.1.2.2. Glucólisis	42
2.2.1.2.3. Ruta de las pentosas fosfato	47
2.2.1.2.4. Reacciones de biosíntesis y anapleróticas	51
2.2.2. Extensión del modelo metabólico de <i>Escherichia coli</i>	55
2.2.2.1. Balances de masa	55
2.2.2.2. Cinéticas de reacción	57
2.3. Conclusiones	65

Capítulo 3

Análisis de sensibilidad global

3.1. Introducción	66
3.2. Análisis de Sensibilidad Global. Metodología	69
3.3. Resultados y discusión	77
3.4. Conclusiones	94

Capítulo 4

Estimación dinámica de parámetros

4.1. Introducción	96
-------------------------	----

4.2. Métodos	99
4.2.1. Problema de estimación dinámica de parámetros	99
4.2.2. Algoritmo de optimización	102
4.2.3. Estimación dinámica de parámetros: Modelo del metabolismo central del carbono	103
4.2.4. Estimación dinámica de parámetros: Modelo cinético de bio-reactor	105
4.2.4.1. Ensayos de cultivos continuos	107
4.2.4.1.1. Condiciones de cultivo	107
4.2.4.1.2. Ensayos de pulso de glucosa y toma de muestras	108
4.2.4.2. Métodos analíticos	110
4.2.4.2.1. Peso seco	110
4.2.4.2.2. Glucosa y productos de fermentación	110
4.3. Resultados y discusión	111
4.3.1. Modelo del metabolismo central del carbono	111
4.3.2. Ensayos de cultivo continuo y calibración del modelo del bio-reactor	120
4.4. Conclusiones	125

Capítulo 5

Diseño óptimo de redes metabólicas estables

5.1. Introducción	127
5.2. Métodos	131
5.2.1. Problema de diseño óptimo de redes metabólicas	131
5.2.2. Problema de diseño óptimo de redes metabólicas bajo restricciones de estabilidad	137
5.2.2.1. Estabilidad de sistemas no lineales. Fundamentos	138
5.2.2.2. Formulación del problema de diseño óptimo de redes metabólicas	

bajo restricciones de estabilidad	142
5.3. Resultados y discusión	144
5.4. Conclusiones	158
Capítulo 6	
Conclusiones y trabajos futuros	
6.1. Conclusiones generales	160
6.2. Conclusiones particulares	161
6.3. Trabajos futuros	164
Bibliografía citada	166

ÍNDICE DE FIGURAS

Fig. 1.1. Esquema de la jerarquía “ómica”: genómica, transcriptómica, proteómica y metabolómica	8
Fig. 1.2. Reacciones de fermentación en cepa de <i>Escherichia coli</i> modificada (Ingram <i>et al.</i> , 1987). Reacciones de <i>Z. mobilis</i> insertadas al genoma de <i>E. coli</i>	9
Fig. 1.3. Reacciones de la ruta glucolítica en <i>Escherichia coli</i>	20
Fig. 1.4. Reacciones de la ruta de las pentosas fosfato	22
Fig. 2.1. Rutas del metabolismo central del carbono y vías fermentativas de <i>Escherichia coli</i> ..	35
Fig. 2.2. Representación esquemática del sistema fosfotransferasa. Fuente: Kremling <i>et al.</i> (2004)	41
Fig. 3.1. Perfiles de concentración media de piruvato (a) y de glucosa-6-fosfato (b) en función del número de escenarios, luego de la aplicación de un pulso de glucosa de 0,15 g/mL	78
Fig. 3.2. Perfil de S_i para la concentración de piruvato, luego de la aplicación de un pulso de glucosa de 0,15 g/mL	81
Fig. 3.3. Perfil de S_i para la concentración de glucosa extracelular, luego de la aplicación de un pulso de glucosa de 0,15 g/mL	84
Fig. 3.4. Perfil de S_i^{int} para la concentración de glucosa extracelular, luego de la aplicación de un pulso de glucosa de 0,15 g/mL	84
Fig. 3.5. Perfil de S_i para la concentración de 6-fosfogluconato, luego de la aplicación de un	

pulso de glucosa de 0,15 g/mL	85
Fig. 3.6. Perfil de S_i^{int} para la concentración de 6-fosfogluconato, luego de la aplicación de un pulso de glucosa de 0,15 g/mL	86
Fig. 3.7. Perfil de S_i para la velocidad de reacción del sistema fosfotransferasa, luego de la aplicación de un pulso de glucosa de 0,15 g/mL	87
Fig. 3.8. Perfil de S_i para la concentración de fosfoenolpiruvato, luego de la aplicación de un pulso de glucosa de 0,15 g/mL. (a) Horizonte de tiempo completo y (b) Primeros 100 segundos de análisis	88
Fig. 4.1. Lazo de control genético y metabólico. Las líneas punteadas indican interacciones regulatorias, mientras que las flechas sólidas indican transformaciones químicas primarias. Los diferentes cuadros separan el dominio de la regulación metabólica de la transcripcional ...	104
Fig 4.2. Cultivos de <i>E.coli</i> en bio-reactor continuo	109
Fig. 4.3. Perfil simulado y datos observados (Chassagnole <i>et al.</i> , 2002) para la concentración de glucosa extracelular (a), glucosa-6-fosfato (b), fructosa-6-fosfato (c) y 6-fosfogluconato (d)	113
Fig. 4.4. Perfil simulado y datos observados (Chassagnole <i>et al.</i> , 2002) para la concentración de fosfoenolpiruvato (e), piruvato (f) y gliceraldehído-3-fosfato (g)	114
Fig. 4.5. Perfil de flujos simulados en los nodos fosfoenolpiruvato (a) y fructosa-6-fosfato (b) ..	117
Fig. 4.6. Perfiles simulados para concentraciones de tres intermediarios de la ruta de las pentosas: ribulosa-5-fosfato (a), xilulosa-5-fosfato (b) y ribosa-5-fosfato (c)	118
Fig. 4.7. Perfil simulado y datos observados para las concentraciones de fructosa-1,6- difosfato (a), fosfoenolpiruvato (b), glucosa-6-fosfato (c) y piruvato (d)	120
Fig. 4.8. Datos observados para la concentración de glucosa extracelular (a), formiato (b) y lactato (c)	121
Fig. 4.9. Perfiles y datos observados para la concentración de glucosa extracelular (a), acetato (b), formiato (c), etanol (d), lactato (e) y succinato (f)	124
Fig. 5.1. Concentraciones óptimas para los diseños con y sin restricciones de estabilidad para la maximización de la producción de serina	149

Fig. 5.2. Flujos metabólicos óptimos para los diseños con y sin restricciones de estabilidad en la maximización de la producción de serina	150
Fig. 5.3. Distribución óptima de flujos para la maximización de la producción de serina en estado de referencia (\square) y el estado estacionario estable	151
Fig. 5.4. Concentraciones óptimas para los diseños con y sin restricciones de estabilidad para la maximización de la producción de etanol	155
Fig. 5.5. Flujos metabólicos óptimos para los diseños con y sin restricciones de estabilidad para la maximización de la producción de etanol	156
Fig. 5.6. Distribución óptima de flujos para la maximización de la producción de etanol en el estado de referencia (\square) y en el estado estacionario estable	157

ÍNDICE DE TABLAS

Tabla 2.1. Tipos de regulación enzimática	32
Tabla 2.2. Balances de masa para metabolitos intracelulares	38
Tabla 2.3. Correlaciones para las concentraciones de los co-metabolitos (Chassagnole <i>et al.</i> , 2002)	39
Tabla 2.4. Balances de masa para metabolitos intracelulares y extracelulares en modelo extendido	56
Tabla 2.5. Nomenclatura, descripción y unidades de los parámetros del modelo del metabolismo de <i>Escherichia coli</i>	60
Tabla 3.1. Valor nominal de los parámetros inciertos (Chassagnole <i>et al.</i> , 2002)	79
Tabla 3.2. Valores de S_i y S_i^{int} en tres instantes de tiempo para la concentración de glucosa-6-fosfato	90
Tabla 3.3. Valores de S_i y S_i^{int} en tres instantes de tiempo para la velocidad de reacción de la enzima piruvato quinasa	92
Tabla 4.1. Condiciones de cultivo (Chassagnole <i>et al.</i> 2002)	104
Tabla 4.2. Curvas de calibrado para glucosa, etanol, ácido láctico, ácido acético, ácido succínico y ácido fórmico	111
Tabla 4.3. Condiciones iniciales para las variables de estado (Chassagnole <i>et al.</i> , 2002; Degenring 2004)	112

Tabla 4.4. Valor óptimo, nominal y cotas superior e inferior de los parámetros estimados del modelo de la red metabólica de <i>E. coli</i>	115
Tabla 4.5. Mediciones estadísticas para la bondad de ajuste del modelo dinámico del metabolismo celular de <i>Escherichia coli</i>	115
Tabla 4.6. Condiciones de cultivo (Degenring 2004)	119
Tabla 4.7. Valores nominal y optimizado para los parámetros estimados del modelo del bio-reactor	122
Tabla 4.8. Estadísticos para la bondad de ajuste del modelo del bio-reactor	125
Tabla 5.1. Características de los autovalores y comportamiento del sistema	139
Tabla 5.2. Flujos y concentraciones de metabolitos en el estado estacionario de referencia	146
Tabla 5.3. Valores de las variables de optimización y de la función objetivo (Caso 1)	148
Tabla 5.4. Valores de las variables de optimización y de la función objetivo (Caso 2)	153

CAPÍTULO 1

INTRODUCCIÓN GENERAL

1.1. Historia de la biotecnología. Surgimiento de nuevas disciplinas

La demanda actual de productos químicos obtenidos con tecnologías limpias y utilizando recursos renovables como materias primas, ha conducido al desarrollo de los llamados bioprocesos a través, de lo que se conoce como biotecnología blanca o industrial. La biotecnología industrial es la aplicación de los conceptos de la biotecnología en el procesamiento y producción de compuestos químicos, materiales y combustibles a escala industrial (Lorenz & Zinke, 2005). En 1981, la Organización para la Cooperación Económica y el Desarrollo (OECD) definió biotecnología como “la aplicación de principios científicos e ingenieriles para el procesamiento de materiales por la acción de agentes biológicos para proveer bienes y servicios” (Bud, 2009).

La práctica de la biotecnología tanto para satisfacer necesidades y brindar servicios, como para mejorar la calidad de vida de los seres humanos, tiene sus orígenes en la antigüedad. Los microorganismos se utilizaban para preservar leche, frutas y vegetales; producir bebidas, quesos y vinagre; mucho tiempo antes

de que fueran “descubiertos” (Soetaert & Vandamme, 2010). En 7000 a.C., en Babilonia y Sumeria, las levaduras se utilizaban para producir cerveza, mediante el proceso conocido como fermentación. Asimismo, esta técnica se usaba en Asiria en la producción de vino (3500 a.C.). En Asia, la fermentación láctica se utilizaba como un método de preservación; de esta manera, por ejemplo, la leche se fermentaba para producir yogurt, kéfir y kumis. El uso de mohos para sacarificar el arroz en la producción de bebidas alcohólicas, data del año 700 a.C.. En el siglo XIV, la producción de bebidas alcohólicas a partir de granos fermentados, que había tenido su origen en China, ya era un proceso muy común en muchas partes del mundo (Soetaert & Vandamme, 2010).

Un hecho de gran importancia en el desarrollo de las ciencias biológicas fue el uso del microscopio desarrollado por Antonie van Leeuwenhoek en el siglo XVII. Leeuwenhoek observó distintas muestras y reportó la presencia de “animálculos” de tamaño miles de veces menor que un grano de arena (Hulse, 2004; Soetaert & Vandamme, 2010). En 1857, Louis Pasteur postula que la fermentación es un proceso vital de las levaduras, refutando la idea de que éstas eran resultado de la fermentación. Asimismo, los trabajos de Pasteur demuestran la inexactitud del concepto de generación espontánea, probando que la vida microbiana existe si hay vida pre-existente (Hulse, 2004; Soetaert & Vandamme, 2010). Durante la siguiente mitad del siglo XIX, Pasteur, Koch y otros científicos descubrieron que muchas enfermedades de la época eran causadas por microorganismos, lo que permitió grandes avances en la medicina y en el área de la microbiología.

Gregor Mendel identifica los caracteres heredables en el año 1866. A comienzos del siglo XX, específicamente en el año 1929, Alexander Fleming descubre accidentalmente la penicilina, un poderoso antibiótico, en un cultivo de *Staphylococcus aureus* contaminado por el moho *Penicillium notatum*. Este hallazgo fue de fundamental importancia, y la penicilina se convirtió en el primer compuesto terapéutico exitoso producido por microorganismos; allí comienza la “era dorada” de los antibióticos (Hulse, 2004; Soetaert & Vandamme, 2010).

Uno de los inconvenientes en el desarrollo de procesos que utilizan enzimas o microorganismos como agentes para la obtención a escala industrial de productos de interés, reside en el hecho de que la mayoría de los microorganismos que potencialmente se pueden emplear en dichos procesos, producen naturalmente solo trazas de los productos de interés industrial. Estas características determinan procesos biotecnológicos poco factibles y/o competitivos (Stephanopoulos, 2002). Sin embargo, desde mediados del siglo XX, diferentes descubrimientos científicos han permitido el avance de la biotecnología industrial. Según Bud (2009) el “evento fundador” de la biotecnología fue el descubrimiento de la doble hélice del ácido-desoxiribonucleico (ADN) por Francis Crick y James Watson, en el año 1953. Otros hallazgos en el área de la biología molecular con implicancias en la biotecnología fueron: el descubrimiento de la codificación de proteínas mediada por el ácido ribonucleico (ARN) por Nirenberg y Matthei en el año 1961; la síntesis de pequeñas porciones de ADN por Kornberg en 1967; el desarrollo de las enzimas que cortan el ADN en lugares específicos, a cargo de Paul Berg en el año 1971, y el trabajo de Stanley Cohen y Herbert Boyer, en el año 1973, que permitió

la transferencia y expresión de una sección de ADN de un organismo a otro, conocido como técnica de ADN recombinante (Bud, 2009).

Si bien los microorganismos han sido sistemáticamente mejorados por mutagénesis al azar y selección desde hace más de 50 años, el desarrollo de técnicas de ADN recombinante han permitido la modificación dirigida de las cepas a través de la introducción de genes específicos que le confieren a las células propiedades deseables para su aplicación industrial, ambiental y/o medicinal (Stephanopoulos *et al.*, 1998; Stephanopoulos, 2002).

En 1976, Robert Swanson y Herbert Boyer crean Genentech Inc., la primera industria biotecnológica que comenzó a desarrollar productos utilizando la técnica de ADN recombinante. Dirigido por Stanley Cohen, Boyer logra insertar segmentos de ADN recombinado en *Escherichia coli*, para la producción de insulina humana. En 1982, la Food & Drug Administration (FDA) aprueba este producto para su comercialización, convirtiéndose en la primera droga obtenida mediante biotecnología.

1.1.1. Ingeniería Metabólica

La Ingeniería Metabólica surge en la década de los '90, con el desarrollo de las técnicas de ADN recombinante. Se define como la disciplina que se ocupa de mejorar las propiedades celulares a través de la incorporación de transportadores específicos y/o nuevos caminos metabólicos, o bien de la modificación de

reacciones ya existentes, principalmente mediante el uso de tecnología recombinante (Bailey, 1991).

Un concepto importante en la Ingeniería Metabólica es el de *red metabólica*, ya que se puede obtener el mejoramiento del metabolismo y de las funciones celulares considerando el sistema de reacciones en conjunto y no a las reacciones aisladas unas de otras. Las mayores contribuciones de esta disciplina son la medición, determinación y entendimiento del control de los flujos bajo condiciones *in vivo* (Stephanopoulos *et al.*, 1998).

La Ingeniería Metabólica identifica cuales son los cambios requeridos en el genoma que producen el cambio deseado en la fisiología del organismo. El estudio de la relación genoma-fisiología ha llevado al desarrollo de diferentes herramientas que permiten un mejor entendimiento de dicha relación (Bailey, 2001). La evaluación de la fisiología celular (fenotipo), constituye un medio para determinar el impacto de los cambios genéticos en una cepa recombinante respecto de la cepa de origen o salvaje (cepa “wild-type”). Según Stephanopoulos *et al.* (1998), los flujos metabólicos determinan el fenotipo del organismo; es por esto que una adecuada cuantificación de los mismos en condiciones *in vivo*, resulta de fundamental importancia. Una herramienta desarrollada para este fin es el *Análisis de Flujos Metabólicos* (MFA), la cual se basa en el cálculo de flujos intracelulares mediante la formulación de modelos estequiométricos, utilizando como datos de entrada el valor de flujos medidos experimentalmente. En un principio, solo se podían cuantificar los flujos extracelulares, como velocidades de consumo de sustrato o de excreción de productos. Sin embargo se estableció que

estas mediciones no eran suficientes para determinar el valor de cada uno de los flujos intracelulares considerados en la red metabólica (Stephanopoulos, 2002). Es por esto que fue necesario cuantificar flujos intracelulares basados en técnicas que utilizan sustratos marcados. La mayoría de las técnicas utilizan sustratos con una cantidad establecida de átomos de carbono marcados, como por ejemplo ^{13}C . El uso de sustratos marcados hace que los metabolitos generados en las rutas metabólicas también se marquen, obteniéndose de esta manera un perfil de isótopos-isómeros (isotopómeros), que pueden ser cuantificados por dos técnicas alternativas: espectroscopía de resonancia magnética nuclear (NMR) o cromatografía gaseosa acoplada a espectrometría de masa (GC-MS). Con esta información, se calculan los flujos intracelulares planteando balances de masa para los isotopómeros (Zupke & Stephanopoulos, 1994; Szyperski, 1995; Schmidt *et al.*, 1997; Sauer *et al.*, 1999; Schmidt *et al.*, 1999; Dauner *et al.*, 2001; Stephanopoulos, 2002; Gosh *et al.*, 2003; Suthers *et al.*, 2007; Teixeira *et al.*, 2008).

Otro de los objetivos importantes de la Ingeniería Metabólica es la identificación de los parámetros responsables del control de los flujos metabólicos, ya que el entendimiento del control de los flujos es indispensable para una correcta modificación de los mismos (Stephanopoulos *et al.*, 1998; Stephanopoulos, 2002). El concepto de *Análisis de Control Metabólico* (MCA) fue desarrollado por Kacser y Burns en 1973 y es, en esencia, un análisis de sensibilidad local que evalúa cómo los flujos metabólicos son afectados por parámetros como la actividad enzimática. Kacser, introdujo también el concepto de *coeficiente de control de flujo*

(FCC) como una medida del grado de control que ejerce la actividad de una enzima sobre el flujo de una reacción metabólica. De una manera más simple, dicho coeficiente puede pensarse como un coeficiente de sensibilidad que evalúa el cambio en el flujo de una reacción respecto del cambio en la concentración o actividad de una enzima específica (Liao & Oh, 1999; Stephanopoulos, 2002).

En los últimos años, los avances en las técnicas experimentales han contribuido al desarrollo y evolución de nuevas áreas como *genómica*, *transcriptómica* y *proteómica*. La *genómica* hace referencia al conjunto de ciencias y técnicas dedicadas al estudio integral del funcionamiento, el contenido, la evolución y el origen de los genomas. El desarrollo de esta área ha permitido la secuenciación del código genético de muchos organismos, entre ellos *Haemophilus influenza*, *Saccharomyces cerevisiae*, *Escherichia coli*, etc. Por otra parte, al estudio del conjunto de genes que se transcriben bajo determinadas condiciones se lo conoce como *transcriptómica*; mientras que al estudio de las proteínas codificadas en el código genético, como así también la determinación de sus funciones y estructuras, se lo denomina *proteómica* (Buchholz *et al.*, 2002). Por otra parte, la necesidad de profundizar el conocimiento del metabolismo desde un punto de vista cuantitativo, llevó al desarrollo de técnicas analíticas sofisticadas que permiten el análisis y la cuantificación de metabolitos intracelulares y extracelulares de bajo peso molecular (*metabolómica*) (Schaefer *et al.*, 1999; Buchholz *et al.*, 2002; Oldiges y Takors, 2005; Steuer, 2006; Bolten *et al.*, 2007). En la Figura 1.1 se muestra la jerarquía establecida entre el genoma, el transcriptoma, el proteoma y el metaboloma, así como también los procesos que

los relacionan. El estudio del metaboloma en diferentes estados fisiológicos provee información de gran utilidad sobre el metabolismo y su regulación, por lo tanto es un complemento a la información que brinda el estudio del genoma, el transcriptoma y el proteoma. Asimismo, dicho conjunto de datos es de gran utilidad para la Ingeniería Metabólica en el desarrollo de nuevas cepas optimizadas para la producción de metabolitos de interés industrial (Buchholz *et al.*, 2002; Bolten *et al.*, 2007).

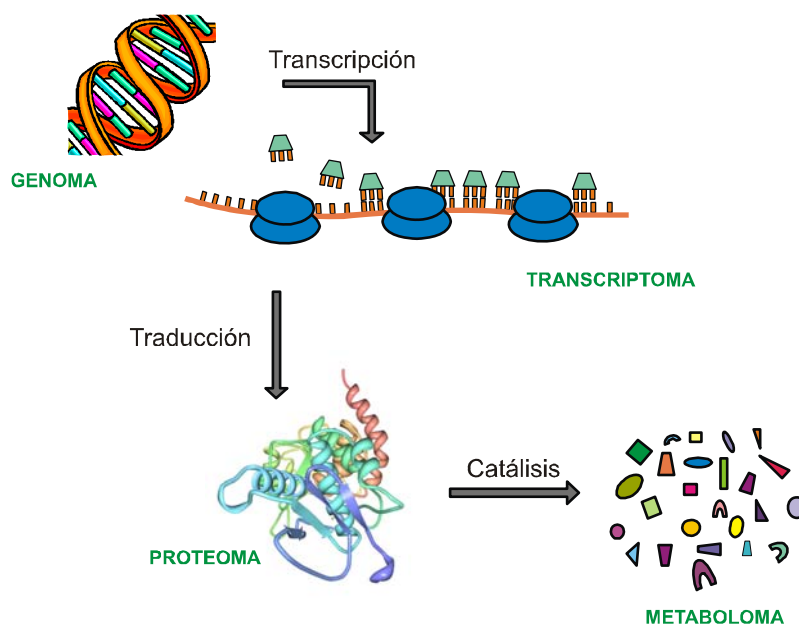


Figura 1.1. Esquema de la jerarquía "ómica": genómica, transcriptómica, proteómica y metabolómica.

La aplicación de la Ingeniería Metabólica ha conducido a la generación de funcionalidades deseadas en las células, que se han traducido en beneficios para las áreas de farmacia, alimentación, química y agrícola, entre otras. Uno de estos ejemplos es la modificación de la bacteria *Escherichia coli* con el objetivo de

incrementar la productividad de etanol de esta cepa. Para tal fin se insertaron los genes heterólogos que codifican las enzimas piruvato descarboxilasa (PDC) y alcohol deshidrogenasa (ADH) de *Zymomonas mobilis* en el genoma de *Escherichia coli* (Ingram *et al.*, 1987), creando los caminos metabólicos que pueden verse en la Figura 1.2. Nickel *et al.* (2006) construyeron una cepa de *Escherichia coli*, capaz de producir poli(3-hidroxiбутирато) (PHB) a partir de suero, mediante la inserción de los genes biosintéticos *pha* correspondientes a *Azotobacter sp.* Otro ejemplo que puede citarse es el de la construcción de una cepa de *Zymomonas mobilis* con una alta tasa de producción de alanina (aminoácido). En este caso, el gen que codifica la enzima alanina deshidrogenasa, proveniente de *Bacillus sphaericus*, fue insertado al genoma de *Z. mobilis* (Uhlenbusch *et al.*, 1991).

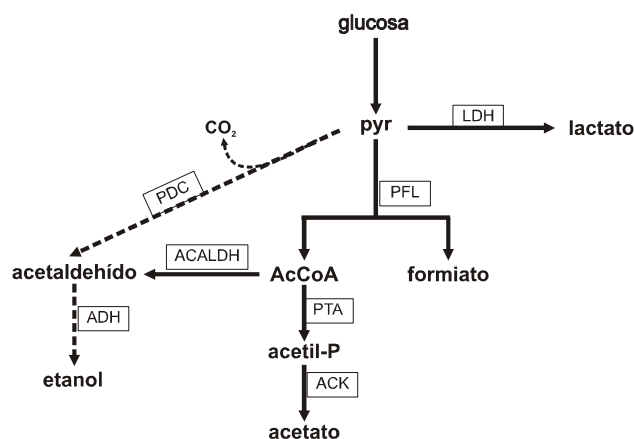


Figura 1.2. Reacciones de fermentación en cepa de *Escherichia coli* modificada (Ingram *et al.*, 1987).
Reacciones de *Z. mobilis* insertadas al genoma de *E. coli* (---)

1.2. Herramientas computacionales. Modelos y bases de datos

La creciente disponibilidad de información experimental tanto a nivel genómico, transcriptómico, proteómico así como también a nivel metabolómico muestran la necesidad de desarrollar herramientas que permitan el estudio y el análisis integral de dicha información. En este sentido, en las últimas décadas se han comenzado a crear bases de datos que recopilan información experimental y la presentan en forma ordenada a través de interfaces amigables. La enciclopedia de los genes y el metabolismo de *Escherichia coli* K12 MG1655, EcoCyc (<http://EcoCyc.org>), es una base de datos generada a partir de información disponible en la literatura científica, que describe: el genoma de *Escherichia coli*, la función y regulación de los genes a nivel transcripcional, post-transcripcional; las enzimas que son codificadas por dichos genes y su regulación; las reacciones catalizadas por las enzimas y la organización de estas reacciones en rutas metabólicas (Karp *et al.*, 1998; Keseler *et al.*, 2011). EcoCyc forma parte de BioCyc (<http://biocyc.org>), una colección de bases de datos del genoma y metabolismo de diferentes organismos secuenciados.

Por otro lado, ExPASy (Expert Protein Analysis System, <http://www.expasy.org>) provee a la comunidad científica acceso a bases de datos y herramientas analíticas dedicadas a lo inherente a las proteínas y el estudio proteómico. Dentro de las bases de datos contenidas en ExPASy, se destacan SWISS-PROT, SWISS-2DPAGE, PROSITE, ENZYMES y SWISS-MODEL (Gasteiger *et al.*, 2003). PROSITE (<http://prosite.expasy.org>) reúne información referente a las

familias y dominios de las proteínas (Sigrity *et al.*, 2009); SWISS-PROT (<http://www.expasy.org/sprot>) es una base de datos que contiene información de la secuenciación de proteínas (Gasteiger *et al.*, 2003); ENZYME (<http://www.expasy.org/enzyme>) recopila información inherente a la nomenclatura de las proteínas (Bairoch, 2000; Gasteiger *et al.*, 2003); SWISS-MODEL (<http://www.expasy.org/swissmod/smrep.html>) es una base de datos de modelos estructurales de proteínas (Gasteiger *et al.*, 2003); SWISS-2DPAGE (<http://www.expasy.org/ch2d/>) reúne información de las proteínas identificadas por la técnica 2D PAGE (electroforesis bidimensional en gel de poliacrilamida) (Gasteiger *et al.*, 2003). BRENDA (BRaunschweig ENzyme DAtabase, <http://www.brenda-enzymes.org>) es una base de datos que contiene información enzimática (cinética, estabilidad, nomenclatura, estructura, etc) y metabólica. Su creación comenzó en 1987, convirtiéndose en una fuente de información fundamental para la investigación en las áreas de biotecnología, medicina e ingeniería metabólica, entre otras (Schomburg *et al.*, 2002).

Una de las bases de datos de mayor alcance es: Kyoto Enciclopedia de Genes y Genoma (KEGG, <http://www.genome.jp/kegg>), cuya creación comenzó en el año 1995 con el objetivo de vincular la información genómica con información funcional de los procesos celulares. KEGG es una herramienta que consiste en tres bases de datos: “Rutas” (Pathways) para la representación de las funciones celulares como redes de moléculas que interactúan, dentro de esta base, en la sección “Metabolismo”, se pueden encontrar gráficos interactivos de redes metabólicas para distintos organismos; “Genes” (Genes), donde se encuentran los catálogos

de los genes para los organismos total o parcialmente secuenciados; “Ligando” (Ligand), en donde se recopila la información de los compuestos químicos presentes en las células, así como también de las enzimas y de las reacciones que éstas catalizan. Asimismo, desde esta base de datos se puede acceder a bases externas, como las descritas anteriormente (Kanehisa y Goto, 2000; Masoudi-Nejad *et al.*, 2007). LIGAND, es una base de datos compuesta por tres secciones: Compuestos, Reacciones y Enzimas. Esta base de datos es utilizada en la reconstrucción de las vías metabólicas de los organismos secuenciados descritos en la base de datos KEGG (Goto *et al.*, 2000; 2002).

Al igual que en otras áreas de la Ingeniería, la predicción, asistida por modelos matemáticos, es fundamental en el diseño de una cepa. Si no fuera de esta manera, la Ingeniería Metabólica debería basarse en un razonamiento meramente intuitivo y los experimentos deberían llevarse a cabo por prueba y error (Liao & Oh, 1999). En este sentido, la formulación de modelos matemáticos de redes metabólicas constituye un camino para lograr la interpretación de numerosos conjuntos de datos obtenidos experimentalmente a distintos niveles (genómico, transcriptómico, proteómico y metabolómico) (Lander, 1999; Schmidt, *et al.*, 1999; Buchholz *et al.*, 2002; Chassagnole *et al.* 2002; Reed & Palsson, 2003; Degenring *et al.* 2004; Schwender, 2008). Existe una gran variedad de enfoques de modelado que han sido utilizados para la formulación de modelos matemáticos de redes metabólicas, entre los que se pueden destacar modelos cinéticos, cibernéticos, estocásticos, teoría de sistemas bioquímicos y modelos basados en restricciones (Reed & Palsson, 2003). De todos ellos, los enfoques más utilizados

en el desarrollo de estas herramientas matemáticas han sido los basados en restricciones y los cinéticos (Varner & Ramkrishna, 1999).

Los modelos basados en restricciones o modelos estequiométricos, son modelos formulados en estado estacionario e incluyen restricciones de tipo estequiométricas (balances de masa de los metabolitos involucrados en la red), termodinámicas (reversibilidad de las reacciones) y de capacidad enzimática (velocidades máximas de reacción). La hipótesis de estado estacionario, se basa en el hecho de que los transitorios metabólicos son muy rápidos (milisegundos a décimas de segundos) en comparación con el tiempo de crecimiento celular (horas a días) (Stephanopoulos *et al.*, 1998). Estos modelos brindan un conjunto de soluciones factibles dentro del espacio generado por las restricciones impuestas. Para determinar cuál de dichas soluciones es la apropiada, se han propuesto diferentes estrategias, entre las que se destacan: análisis de rutas extremas (“*extreme pathways analysis*”), análisis de modos elementales (“*elementary mode analysis*”) y análisis de balance de flujos (“*flux balance analysis*”, FBA) (Varma & Palsson, 1994; Schuster *et al.*, 1999; Schilling *et al.*, 2000; Edwards *et al.*, 2002; Kauffman *et al.*, 2003; Klamt & Stelling, 2003; Reed & Palsson, 2003; Papin *et al.*, 2004).

Tanto el análisis de modos elementales como el de rutas extremas utilizan el análisis convexo para describir el espacio de soluciones. Los modos elementales y las rutas extremas son conjuntos de vectores que representan por sí mismos distribuciones de flujos factibles, desde el punto de vista bioquímico, a través de una red metabólica definida. Según Papin *et al.* (2004) los “modos elementales”

son conjuntos de vectores derivados del análisis convexo de la matriz estequiométrica, que tienen la propiedad de constituir el mínimo número de enzimas que se necesitan para que el microorganismo exista como una unidad funcional. Es decir, si una reacción es eliminada del modo elemental, éste no podría operar como una unidad funcional. A esta propiedad se la conoce como “independencia genética”. Al set de vectores linealmente independientes obtenido de la matriz estequiométrica se lo denomina “rutas extremas” (Schilling *et al.*, 2000). Las rutas extremas no son un conjunto de reacciones metabólicas genéticamente independientes, representan los límites del espacio convexo de soluciones de la red metabólica (Papin *et al.*, 2004).

El análisis de balance de flujos (FBA) es un enfoque basado en los principios de la programación lineal y ha sido la metodología más utilizada para la determinación de fenotipos “óptimos” (Varma y Palsson, 1994; Bonarius *et al.*, 1997; Pramanik & Keasling, 1997; Sauer *et al.*, 1998; Ramakrishna *et al.*, 2001; Reed y Palsson, 2003; Rodriguez & Infante, 2009). Este enfoque consiste en determinar, entre las infinitas soluciones que posee el sistema de ecuaciones lineales subdeterminado que describe la red metabólica, aquella que maximiza un cierto objetivo biotecnológico, mediante la resolución del siguiente problema de Programación

Lineal:

$$\begin{aligned} & \max f \\ & \text{s.a.} \\ & \mathbf{S} \cdot \mathbf{r} = \mathbf{b} \\ & r_i \leq r_i^{\max} \end{aligned}$$

donde \mathbf{S} es la matriz de coeficientes estequiométricos de la red metabólica, \mathbf{r} es el vector de flujos, r_i^{max} corresponde al valor máximo que puede alcanzar un flujo y f es el objetivo biotecnológico a maximizar.

FBA se emplea para calcular, interpretar y predecir las distribuciones de flujos metabólicos tales que maximicen o minimicen un cierto objetivo. Es decir, formulando un problema de programación lineal sujeto a las restricciones estequiométricas, termodinámicas y de capacidad, es posible hallar la distribución óptima de flujos. Por otra parte, se han aplicado técnicas de programación lineal mixto entera para hallar múltiples soluciones óptimas (si existiesen) en modelos de redes metabólicas basados en restricciones (Lee *et al.*, 2000; Phalakornkule *et al.*, 2001; Lin *et al.*, 2005; Gosh *et al.*, 2005; Motamediam & Naeimpoor, 2011). El análisis de balances de flujos es una de las herramientas más utilizadas, junto con la determinación experimental de flujos metabólicos intracelulares y extracelulares, en lo que se conoce como Análisis de Flujos Metabólicos (Sección 1.1.1).

Por su parte, los modelos cinéticos de redes metabólicas describen cuantitativamente la dinámica de las concentraciones de metabolitos y de las velocidades de reacción de las rutas bioquímicas implicadas en la red. Estos modelos están basados en la estructura de la red, es decir contemplan la estequiometría de las reacciones, las cinéticas e información regulatoria, en este punto se incluye la expresión cinética correspondiente a cada reacción teniendo en cuenta la presencia de inhibidores y activadores de las enzimas (Wolfgang & Noack, 2011). Tradicionalmente los modelos cinéticos incluyen ecuaciones cinéticas mecanísticas que son derivadas de ensayos *in vitro* realizados sobre las

enzimas aisladas. Sin embargo, debido a la complejidad que exhiben estas expresiones, algunos autores proponen el uso de ecuaciones cinéticas no-mecanísticas de tipo lin-log (*linear-logarithmic*), “cinética más conveniente” y expresiones basadas en la ley de acción de masas (Hatzimanikatis *et al.*, 1996; Hatzimanikatis & Bailey, 1997; Nielsen, 1997; Hatzimanikatis *et al.*, 1998; Visser *et al.*, 2000; Degenring *et al.*, 2004; Visser *et al.*, 2004; Liebermeister & Klipp, 2006a,b).

En una primera instancia, los modelos de redes se basaban sólo en el metabolismo del organismo bajo estudio. Sin embargo, los progresos en la secuenciación del genoma de algunos organismos y en el estudio de los procesos de transcripción, traducción y su regulación, han llevado al desarrollo de modelos que describen estos mecanismos, así como también a la ampliación de los modelos del metabolismo incorporando nuevas reacciones. En este sentido, se han desarrollado modelos a escala genómica acoplados a modelos estequiométricos de la red metabólica para distintos organismos (Edwards & Palsson, 2000; Price *et al.*, 2003; Duarte *et al.*, 2004; Patil *et al.*, 2004; de Figueredo *et al.*, 2009; Chung *et al.*, 2010; Sohn *et al.*, 2010), y modelos dinámicos del metabolismo y el mecanismo de transcripción para la producción de aminoácidos (Usuda *et al.*, 2010), para la predicción de constantes cinéticas del fenómeno de transcripción (Hardiman *et al.*, 2010) y para el estudio de la dinámica de la regulación genética y metabólica (Asenjo *et al.*, 2006) en *Escherichia coli*.

1.3. *Escherichia coli*. Metabolismo central del carbono

1.3.1. *Escherichia coli*. Características

Escherichia coli es una bacteria gram-negativa que pertenece a la familia *Enterobacteriaceae*; es un bacilo no formador de esporas, anaerobio facultativo. Las células pueden medir desde 1 a 2 μm de longitud y de 0,1 a 0,5 μm de diámetro. Esta bacteria se multiplica vegetativamente por fisión binaria.

Escherichia coli es uno de los microorganismos procariotas mejor caracterizados y ha sido considerado como organismo modelo de un sinnúmero de estudios bioquímicos, biológicos y biotecnológicos. Asimismo, la secuenciación del genoma de *Escherichia coli* K12, finalizada en 1997, estimuló la continuidad de la investigación para lograr un mejor entendimiento de este organismo, de gran importancia en el ámbito experimental, médico e industrial (Blattner *et al.*, 1997; Han *et al.*, 2006).

1.3.2. Metabolismo central del carbono en *Escherichia coli*

El metabolismo se define como el conjunto de todas las reacciones bioquímicas catalizadas por enzimas que ocurren en las células (Nelson & Cox, 2005). Las funciones específicas del metabolismo celular son:

- Obtención de energía de moléculas biodegradables o de luz solar absorbida,
- Conversión de los precursores de los componentes macromoleculares de las células,
- Síntesis de proteínas, lípidos, ácidos nucleicos y otros componentes celulares,
- Degradación de biomoléculas necesarias para el mantenimiento celular.

El metabolismo central del carbono está formado por un conjunto de reacciones que tienen importancia vital ya que participa en la síntesis de metabolitos precursores y en la obtención de energía. Dentro del metabolismo central del carbono se encuentran la vía glucolítica, la ruta de las pentosas fosfato, el ciclo de los ácidos tricarboxílicos o ciclo de Krebs y las rutas fermentativas.

1.3.2.1. Glucólisis

En la glucólisis (vocablo proveniente del griego *glykys*, que significa azúcar y *lysis*, que significa romper) la glucosa (hexosa) es oxidada en una serie de diez reacciones enzimáticas, obteniéndose dos moléculas de piruvato (compuesto tricarbonado). En la secuencia de reacciones, parte de la energía libre que cede la glucosa se almacena en forma de ATP y NADH.

Según Nelson y Cox (2005) la glucólisis consta de dos fases: una fase preparatoria, en la cual se consume ATP y la glucosa es convertida en

gliceraldehído-3-fosfato; y una fase de beneficios, en la cual se obtiene energía en forma de ATP y NADH mientras que las moléculas de gliceraldehído-3-fosfato continúan siendo oxidadas hasta la obtención de piruvato.

La Figura 1.3 muestra las reacciones que forman parte del camino glucolítico. La fase preparatoria de la glucólisis consta de cinco reacciones, la primera de ellas es la fosforilación de la molécula de glucosa (glcext) para dar glucosa-6-fosfato (g6p); siendo el fosfoenolpiruvato (pep) el dador del grupo fosforilo. Esta reacción irreversible, es catalizada por el sistema fosfotransferasa (PTS). Luego, la glucosa-6-fosfato (g6p) es convertida en fructosa-6-fosfato (f6p) por una reacción de isomerización reversible catalizada por la enzima fosfoglucosa isomerasa (PGI). La siguiente reacción, en la que la enzima fosfofructoquinasa (PFK) se desempeña como catalizador, consiste en la fosforilación de la fructosa-6-fosfato (f6p) para dar fructosa-1,6-difosfato (fdp), siendo el ATP el dador del grupo fosforilo. Luego le sigue la ruptura de la fructosa-1,6-difosfato (fdp) para dar dos triosas, dihidroxiacetona fosfato (dhap) y gliceraldehído-3-fosfato (gap), que es catalizada por la enzima aldolasa (ALDO). La última reacción de la fase preparatoria de la glucólisis es la interconversion de dichas triosas, catalizada por la enzima triosa fosfato isomerasa (TIS). Debido a que solo el gliceraldehído-3-fosfato puede seguir siendo oxidado en los siguientes pasos de la glucólisis, la dihidroxiacetona fosfato es convertida, reversiblemente, en gliceraldehído-3-fosfato.

Como se mencionó anteriormente, la fase de beneficios de la glucólisis incluye las reacciones conservadoras de energía. El primer paso es la oxidación y

fosforilación por fosfato inorgánico del gliceraldehído-3-fosfato (gap) para dar 1,3-difosfoglicerato (pgp), catalizada por la enzima gliceraldehído-3-fosfato deshidrogenasa (GAPDH).

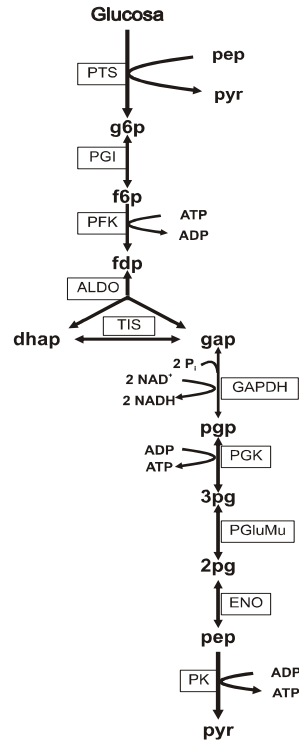


Figura 1.3. Reacciones de la ruta glucolítica en *Escherichia coli*

En la siguiente reacción, la enzima fosfoglicerato quinasa (PGK) transfiere el grupo fosforilo de alta energía del 1,3-difosfoglicerato (pgp) al ADP, lo que lleva a la formación de ATP y 3-fosfoglicerato (3pg). La transferencia de un grupo fosforilo desde un sustrato, en este caso 1,3-difosfoglicerato, para dar ATP se conoce como fosforilación a nivel de sustrato. El 3-fosfoglicerato (3pg) es convertido en 2-fosfoglicerato (2pg), debido a la acción de la enzima fosfoglicerato mutasa (PGM)

que cataliza el desplazamiento reversible del grupo fosforilo entre el carbono 2 y el carbono 3.

Siguiendo la secuencia de reacciones, la enzima enolasa (ENO) cataliza la deshidratación del 2-fosfoglicerato (2pg) para dar fosfoenolpiruvato (pep). El último paso de la glucólisis es la transferencia del grupo fosforilo desde el fosfoenolpiruvato (pep) al ADP originando ATP y piruvato (pyr), esta reacción es catalizada por la enzima piruvato quinasa (PK).

El balance global de la glucólisis es:



El piruvato, luego puede seguir distintos caminos. En condiciones aeróbicas, este compuesto es oxidado, con pérdida de su grupo carboxilo en forma de CO₂, para dar el grupo acetilo del acetyl-coA (AcCoA), que luego es oxidado completamente a CO₂ en el ciclo de los ácidos tricarboxílicos. Por otro lado en condiciones anaeróbicas, el piruvato es catabolizado en las rutas fermentativas para dar compuestos como etanol, ácido láctico, ácido acético, entre otros. En condiciones anaeróbicas también se produce CO₂, en la reacción reversible catalizada por la enzima fosfoenolpiruvato carboxilasa (PEPCxylase) y en la ruta de las pentosas.

1.3.2.2. Ruta de las pentosas fosfato

En esta ruta, la glucosa-6-fosfato es oxidada a pentosas fosfato, compuestos precursores de los nucleótidos, siendo la nicotinamida adenina dinucleótido fosfato en su forma oxidada (NADP^+) el aceptor de electrones en estas reacciones para dar NADPH (nicotinamida adenina dinucleótido fosfato reducido), el cual es necesario en las reacciones de biosíntesis.

La ruta de las pentosas (Figura 1.4) consta de una fase oxidativa, en la cual se producen pentosas fosfato y NADPH, y de una fase no oxidativa, en la que las pentosas fosfato se reciclan a la ruta glucolítica.

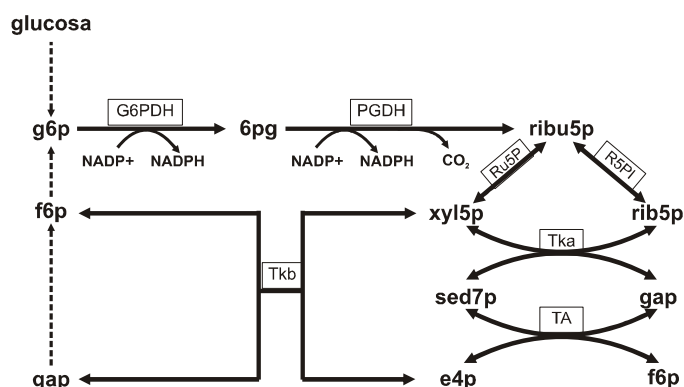


Figura 1.4. Reacciones de la ruta de las pentosas fosfato

La primera reacción de la fase oxidativa de la vía de las pentosas fosfato, es la oxidación de la glucosa-6-fosfato (g6p), catalizada por la enzima glucosa-6-fosfato (G6PDH) deshidrogenasa para dar 6-fosfogluconolactona; el NADP^+ es el aceptor de electrones. Es por esta razón, que en esta reacción hay generación de una molécula de NADPH. Luego, la lactona es hidrolizada para dar 6-fosfogluconato (6pg), el cual luego es oxidado y descarboxilado por la acción de la enzima 6-

fosfogluconato deshidrogenasa (PGDH), obteniéndose como producto la pentosa ribulosa-5-fosfato (ribu5p) con generación de CO₂ y una molécula de NADPH. La siguiente reacción, catalizada por la ribulosa-5-fosfato isomerasa (R5PI), corresponde a la isomeración de la ribulosa-5-fosfato (ribu5p) para dar ribosa-5-fosfato (rib5p). Ésta es la última reacción correspondiente a la fase oxidativa.

En la fase no oxidativa, la ribulosa-5-fosfato, por medio de la enzima ribulosa-5-fosfato epimerasa (Ru5P) se interconvierte en su estereoisómero, xilulosa-5-fosfato (xyl5p). A continuación, en una serie de pasos se produce un reordenamiento de los esqueletos carbonados, obteniéndose a partir de seis azúcares fosfato de cinco carbonos, cinco azúcares fosfato de seis carbonos, completándose el ciclo y permitiendo la oxidación continua de la glucosa-6-fosfato, con producción de NADPH (Nelson y Cox, 2005). En este sentido, la enzima transcetolasa a (TKa), transfiere los carbonos 1 y 2 de la xilulosa-5-fosfato (xyl5p) a la ribosa-5-fosfato (rib5p), formando un compuesto de siete carbonos llamado sedoheptulosa-7-fosfato (sed7p). El fragmento de tres átomos de carbonos que se obtiene de la xilulosa (xyl5p) es el gliceraldehído-3-fosfato (gap). A continuación, la enzima transaldolasa (TA) cataliza una reacción en la cual se elimina un fragmento de tres carbonos de la sedoheptulosa-7-fosfato (sed7p) que se condensa con el gliceraldehído-3-fosfato (gap), formando fructosa-6-fosfato (f6p) y una tetrosa (eritrosa-4-fosfato (e4p)). En la siguiente reacción, catalizada por la enzima transacetolasa b (TKb), se forma fructosa-6-fosfato (f6p) y gliceraldehído-3-fosfato (gap) a partir de la eritrosa-4-fosfato (e4p) y la xilulosa-5-fosfato (xyl5p). Ambos productos de esta reacción (f6p y gap) son re-circulados a la glucólisis.

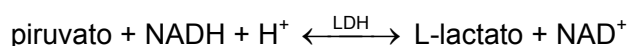
Luego de dos repeticiones de las reacciones descritas anteriormente, las dos moléculas de gliceraldehído-3-fosfato (gap) se convierten en fructosa-1,6-difosfato (fdp) pudiéndose obtener glucosa-6-fosfato (g6p) a partir de dicho metabolito.

1.3.2.3. Rutas fermentativas

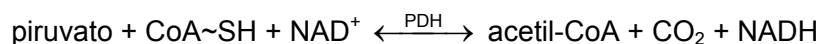
En condiciones de crecimiento micro-aeróbicas o anaeróbicas, el piruvato es catabolizado en las rutas fermentativas. El objetivo primordial de la fermentación es la regeneración de nicotinamida adenina dinucleótido en su forma oxidada (NAD^+), a partir de la oxidación del NADH (nicotinamida adenina dinucleótido reducido), para que siga siendo utilizado en la glucólisis.

Escherichia coli, tiene cinco metabolitos como productos de las vías fermentativas: etanol, acetato, lactato, formiato y succinato. A este tipo de fermentación se la conoce como *fermentación ácida mixta*. El etanol, lactato, acetato y formiato son producidos a partir de piruvato, mientras que el succinato es producido a partir de fosfoenolpiruvato.

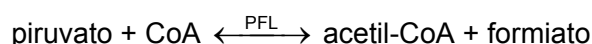
La producción de lactato a partir de la reducción de piruvato es catalizada por la enzima lactato deshidrogenasa (LDH). En esta reacción el agente reductor es el NADH, que se oxida produciendo NAD^+ , según:



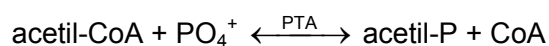
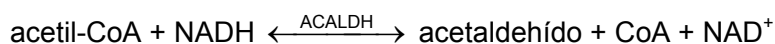
En condiciones aeróbicas la enzima piruvato deshidrogenasa (PDH) cataliza la descarboxilación oxidativa del piruvato obteniéndose acetil-CoA y CO₂ como productos, según la siguiente reacción:



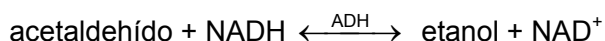
En condiciones anaeróbicas, la enzima piruvato deshidrogenasa es inhibida, siendo la enzima piruvato formiato liasa (PFL) la que cataliza la conversión de piruvato y coenzima A, obteniéndose como productos ácido fórmico y acetil-coenzima A (acetil-CoA) (Knappe *et al.*, 1974), según:



El acetil-CoA sigue dos caminos, uno es la reducción de este compuesto en una reacción catalizada por la enzima acetaldehído deshidrogenasa (ACALDH) para dar acetaldehído y coenzima A, siendo el agente reductor en esta reacción el NADH. El otro destino del acetil-CoA es su conversión en acetil fosfato (acetil-P) por la acción de la enzima fosfotransacetilasa (PTA). Las reacciones son:



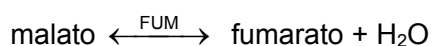
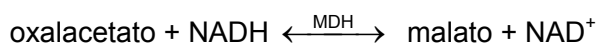
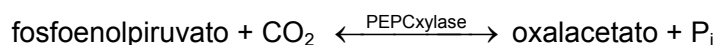
El acetaldehído es reducido a etanol en una reacción catalizada por la enzima alcohol deshidrogenasa (ADH), el dador de electrones en esta reacción de óxido-reducción es el NADH, según:



Por su parte, el acetil-P es convertido en acetato por la acción de la enzima acetato quinasa (ACK), en esta reacción el fosfato es cedido al ADP obteniéndose de esta forma ATP, según la siguiente reacción:



Como se mencionó previamente, otro de los productos de la fermentación ácida mixta en *Escherichia coli* es el succinato. Esta bacteria posee cuatro vías de producción de succinato, sin embargo la reacción más favorecida energéticamente es la que se describe a continuación: el fosfoenolpiruvato es carboxilado obteniéndose oxalacetato en una reacción catalizada por la enzima fosfoenolpiruvato carboxilasa (PEPCxylase), el oxalacetato luego es convertido en malato por la acción de la enzima malato deshidrogenasa (MDH). Luego el malato pasa a fumarato en una reacción catalizada por la enzima fumarato hidratasa (FUM). Este último compuesto es reducido a succinato en una reacción catalizada por la enzima fumarato reductasa (FRD) (Zhang *et al.*, 2009). Las correspondientes reacciones son:



1.4. Objetivos de la tesis

General

Desarrollo de modelos integrados de optimización de bioprocesos a nivel metabólico para identificar estrategias óptimas de operación a nivel celular mediante la aplicación de técnicas de programación matemática y estudios experimentales.

Específicos

- Desarrollo e implementación de modelos de redes metabólicas, de creciente rigurosidad y complejidad, basados en datos disponibles en la literatura y obtenidos en este trabajo.
- Estudio de la regulación dinámica de la ruta glucolítica en *Escherichia coli*
- Calibración y validación de los modelos desarrollados frente a datos experimentales obtenidos en distintas condiciones de trabajo.
- Diseño óptimo de redes metabólicas estables para la producción de metabolitos de interés.

1.5. Estructura de la tesis

La presente tesis está estructurada en seis capítulos, según se detalla a continuación:

En el Capítulo 1 se presenta una breve reseña histórica de los avances en el área de la biotecnología, se introducen conceptos básicos del modelado de rutas metabólicas y se describe el metabolismo celular de *Escherichia coli*.

En el Capítulo 2 se desarrolla un modelo cinético del metabolismo central del carbono y las vías fermentativas de *Escherichia coli* K12. El modelo consiste en un sistema de ecuaciones diferenciales algebraicas no lineales que surge de formular balances de masa para cada metabolito que integra la red. La no linealidad de las ecuaciones está dada por la inclusión de las expresiones cinéticas de las reacciones enzimáticas.

En el Capítulo 3 se desarrolla e implementa una estrategia de análisis de sensibilidad global basado en la varianza para el modelo dinámico, con el fin de identificar los parámetros cinéticos que tienen mayor influencia en las variables de salida del modelo.

El Capítulo 4 presenta la estimación dinámica de parámetros frente a datos experimentales extraídos de bibliografía y obtenidos en el trabajo de tesis. Se describe el trabajo experimental realizado con el objetivo de cuantificar los productos de fermentación en un cultivo de *Escherichia coli*.

En el Capítulo 5, en un marco de optimización, se formulan problemas de diseño óptimo de redes metabólicas que incluyen restricciones de estabilidad del sistema.

En el Capítulo 6 se presentan las conclusiones de la tesis y las líneas de acción futuras.

CAPÍTULO 2

MODELADO DE LA RED METABÓLICA DE *ESCHERICHIA COLI*

2.1. Introducción

La creciente disponibilidad de información a nivel genómico, transcriptómico, así como también a nivel fluxómico y metabolómico, muestran la necesidad de desarrollar herramientas que permitan el estudio y análisis integral de dicha información experimental (Lander, 1999; Schmidt, *et al.*, 1999; Buchholz *et al.*, 2002; Chassagnole *et al.* 2002; Reed & Palsson, 2003; Degenring *et al.* 2004; Schwender, 2008). En este sentido, el modelado matemático de redes metabólicas constituye una estrategia poderosa para interpretar y evaluar los numerosos conjuntos de datos obtenidos experimentalmente.

La metabolómica, tal como se describe en el Capítulo 1, hace referencia al conjunto de datos correspondientes a las concentraciones de metabolitos tanto intra como extracelulares. Estos metabolitos representan los reactivos y productos de las reacciones involucradas en la red metabólica y, si son cuantificados en una escala de tiempo apropiada, permiten estudiar la cinética de las reacciones metabólicas “in vivo”. Este hecho hace posible la formulación y ajuste de modelos

cinéticos de redes metabólicas. Estos modelos describen cuantitativamente la dinámica de las concentraciones de metabolitos y de las velocidades de reacción de las rutas bioquímicas implicadas en la red. Los modelos metabólicos cinéticos están basados en la estructura de la red, contemplan la estequiometría de las reacciones, expresión cinética correspondiente a cada reacción y tienen en cuenta la presencia de inhibidores y activadores de las enzimas, como información regulatoria (Wolgang & Noack, 2011).

Tradicionalmente los modelos cinéticos incluyen ecuaciones cinéticas mecanísticas que son derivadas de ensayos *in vitro* realizados sobre las enzimas aisladas. Sin embargo, debido a la complejidad que exhiben estas expresiones, algunos autores proponen el uso de ecuaciones cinéticas no-mecanísticas de tipo linlog, “cinética conveniente”, expresiones basadas en la ley de acción de masas (Hatzimanikatis *et al.*, 1996; Hatzimanikatis & Bailey, 1997; Nielsen, 1997; Hatzimanikatis *et al.*, 1998; Visser *et al.*, 2000; Degenring *et al.*, 2004; Visser *et al.*, 2004; Liebermeister & Klipp, 2006a,b). En la Tabla 2.1 se muestran los distintos tipos de mecanismos de regulación enzimática junto con su correspondiente expresión cinética (mecanística).

Los modelos cinéticos del metabolismo celular datan de la década del '60. Chance *et al.* (1960) y Garfinkel *et al.* (1964) formulan un modelo de la glucólisis en células tumorales (Snoep, 2005). En la década de los '70, se proponen modelos cinéticos de la glucólisis en eritrocitos, cuya complejidad se va aumentando al incluir fenómenos de transporte y efectos de potencial de membranas (Snoep, 2005).

Tabla 2.1. Tipos de regulación enzimática

Tipo de regulación	Características	Ecuación
Inhibición competitiva	<ul style="list-style-type: none"> El inhibidor compite con el sustrato por los sitios activos de la enzima La presencia del inhibidor competitivo aumenta el valor de K_s aparente El valor de la velocidad máxima se mantiene constante 	$r_i = \frac{r_i^{max} C_S}{K_S \left(1 + \frac{C_I}{K_I} \right) + C_S}$
Inhibición no competitiva	<ul style="list-style-type: none"> El inhibidor no compite por los sitios activos de la enzima La unión del inhibidor a la enzima produce cambios conformacionales que afectan la actividad enzimática El valor de K_s se mantiene constante Disminuye el valor de la velocidad máxima de reacción 	$r_i = \frac{r_i^{max} C_S}{\left(1 + \frac{C_I}{K_I} \right) (K_S + C_S)}$
Inhibición acompetitiva	<ul style="list-style-type: none"> El inhibidor competitivo se une al complejo enzima sustrato, inactivándolo Disminuye el valor de K_s aparente y el de la velocidad máxima 	$r_i = \frac{r_i^{max} C_S}{K_S + C_S \left(1 + \frac{C_I}{K_I} \right)}$
Activación	<ul style="list-style-type: none"> La presencia de un activador aumenta la actividad enzimática. Activación no esencial: la velocidad de reacción no se anula en ausencia del activador Activación esencial: la velocidad de reacción tiende a cero en ausencia de activador 	$r_i = \frac{r_i^{max} C_S (K_A + VC_A)}{(K_S + C_S)(K_A + C_A)}$ <p style="text-align: center;"><i>Activación no esencial</i></p> $r_i = \frac{r_i^{max} C_S C_A}{(K_S + C_S)(K_A + C_A)}$ <p style="text-align: center;"><i>Activación esencial</i></p>
Regulación alostérica	<ul style="list-style-type: none"> Cambios conformacionales o electrónicos en las enzimas por la unión de efectores en sitios distintos a los activos El efector puede ser el mismo sustrato u otro compuesto Respuesta homotrópica: el sustrato regula la actividad enzimática, la regulación es positiva (activación) Respuesta heterotrópica: un compuesto distinto del sustrato regula la actividad enzimática, la regulación puede ser positiva (activación) o negativa (inhibición) 	$r_i = \frac{r_i^{max} C_S^n}{K' + C_S^n}$ <p style="text-align: center;"><i>n = número de sitios activos</i></p>
Inhibición por sustrato	<ul style="list-style-type: none"> Inhibición de la actividad enzimática por exceso de la concentración de sustrato 	$r_i = \frac{r_i^{max}}{1 + \frac{K_1}{C_S} + \frac{C_S}{K_2}}$

C_S , C_I y C_A : concentraciones de sustrato, inhibidor y activador, respectivamente; K_S : contante de saturación media; K_I : constante de inhibición; K_A : constante de activación; V y K_2 : constantes cinéticas; r_i^{max} : velocidad máxima de reacción.

Rizzi *et al.* (1997) formulan un modelo cinético mecanístico de gran tamaño, para la levadura *Saccharomyces cerevisiae*, que incluye la vía glucolítica y el ciclo de los ácidos tricarboxílicos. Vaseghi *et al.* (1999) proponen un modelo cinético basado en expresiones mecanísticas para la vía de las pentosas fosfato, con el objetivo de estudiar la dinámica de dicha vía en *Saccharomyces cerevisiae*.

En lo que respecta al metabolismo de *Escherichia coli*, varios autores han contribuido a la formulación de modelos cinéticos de distintas vías metabólicas, presentes en esta bacteria, que han permitido el estudio de las mismas. Chassagnole *et al.* (2002) formulan un modelo cinético mecanístico que contempla el sistema fosfotransferasa, la glucólisis y la vía de las pentosas. Schmid *et al.* (2004) proponen un modelo para la producción de triptófano incluyendo modificaciones al modelo formulado por Chassagnole (2002). Asimismo, Sevilla *et al.* (2005) incorporan la ruta del metabolismo del trimetilamonio de *Escherichia coli*, al modelo propuesto por Chassagnole (2002). Más recientemente Usuda *et al.* (2010) propone un modelo cinético del metabolismo y del sistema regulatorio para la producción de aminoácidos en *Escherichia coli*.

Por su naturaleza, los modelos cinéticos de redes metabólicas, en comparación con los modelos basados en restricciones, permiten un completo entendimiento del metabolismo celular. Una vez calibrados y validados, estos modelos constituyen una herramienta muy poderosa y cada vez más buscada en la industria biotecnológica, ya que permiten estudiar efectos regulatorios, explorar posibles estados fisiológicos transitorios, presencia de oscilaciones (Wiechert &

Noack, 2011), estabilidad de los estados estacionarios, así como también permiten el re-diseño óptimo del metabolismo celular.

2.2. Modelado de la red metabólica de *Escherichia coli*

En este capítulo se presenta un modelo cinético del metabolismo de la cepa *Escherichia coli* K12 W3110. Se describen las reacciones involucradas en la glucólisis, la vía de las pentosas fosfato, la reacción de incorporación y fosforilación de glucosa extracelular, llevada a cabo por el sistema fosfotransferasa y las rutas fermentativas. El modelo se ha desarrollado con el objetivo de estudiar la respuesta del metabolismo celular a una perturbación de tipo pulso en la concentración del sustrato aplicada a un cultivo continuo, creciendo en condiciones de limitación de glucosa.

En el marco de la presente tesis, se han propuesto modificaciones para ecuaciones cinéticas correspondientes a las enzimas involucradas en un modelo del metabolismo central del carbono propuesto por Chassagnole *et al.* (2002), a la vez que se ha extendido para incluir las reacciones de fermentación en *Escherichia coli*.

La Figura 2.1 muestra la estructura del metabolismo central del carbono de *Escherichia coli*, incluyendo las vías fermentativas ya descritas en el Capítulo 1.

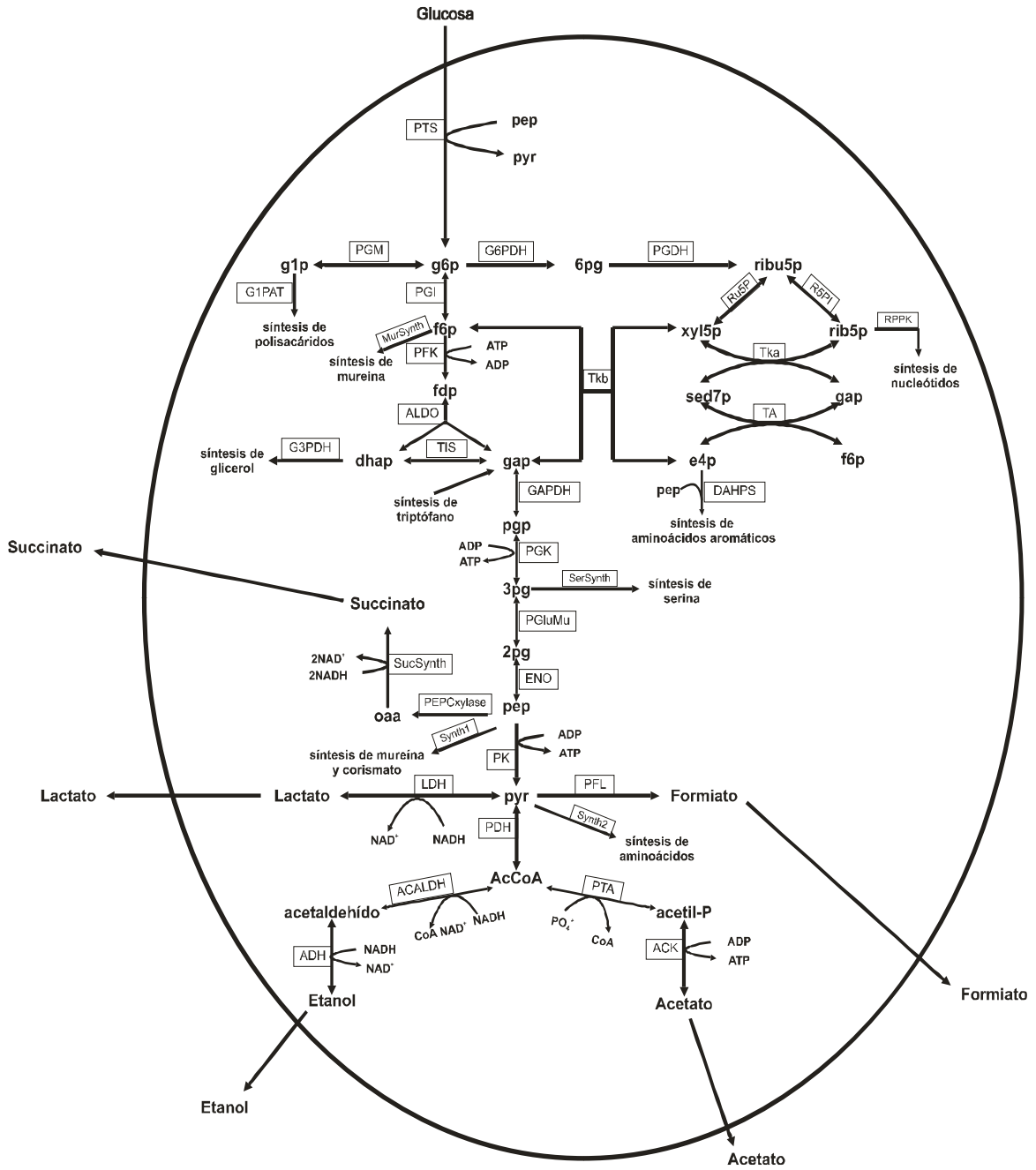


Figura 2.1. Rutas del metabolismo central del carbono y vías fermentativas de *Escherichia coli*

2.2.1 Modelo cinético del metabolismo central del carbono

2.2.1.1 Balances de masa

Los modelos dinámicos de las redes metabólicas comprenden un sistema de ecuaciones diferenciales algebraicas, altamente no lineal, que surge de formular balances de masa para cada nodo de la red, es decir, para cada uno de los metabolitos extra e intracelulares y los co-metabolitos involucrados en las reacciones.

Para el caso de los metabolitos intracelulares, se han formulado balances de masa dinámicos para cada metabolito i de la red de la Figura 2.1, con flujos metabólicos r_j asociados a la enzima j , como se muestra a continuación:

$$\frac{dC_i}{dt} = \sum_j v_{ij} r_j - \mu C_i \quad (2.1)$$

$i = g6p, f6p, fdp, gap, dhap, pgp, 3pg, 2pg, pep, pyr, 6pg, ribu5p, xyl5p, sed7p, rib5p, e4p, g1p.$

$j = PTS, PGI, PFK, ALDO, TIS, GAPDH, PGK, PGluMu, ENO, PK, PDH, PEPCxylase, PGM,$

$G1PAT, RPPK, G3PDH, SerSíntesis, MurSíntesis, DAHPS, TrpSíntesis, MetSíntesis, G6PDH, PGDH,$

$Ru5p, R5PI, TKa, TKb, TA, Síntesis1, Síntesis2.$

donde C_i (mM) es la concentración del metabolito i , r_j (mM/s) corresponde a la expresión cinética de la reacción catalizada por la enzima j , v_{ij} es el coeficiente estequiométrico del metabolito i en la reacción j (< 0 para reactivos, > 0 para productos) y μ (s^{-1}) es la velocidad específica de crecimiento de *Escherichia coli*.

El término μC_i representa la dilución de la concentración del metabolito i por crecimiento de la biomasa (Fredrickson, 1976).

Por su parte, el balance de masa para glucosa extracelular se ha formulado según la siguiente ecuación:

$$\frac{dC_{glc}^{extracelular}}{dt} = D(C_{glc}^{alimentación} - C_{glc}^{extracelular}) + f_{pulso} - \frac{C_x r_{PTS}}{\rho_x} \quad (2.2)$$

donde D (s^{-1}) es la velocidad de dilución, r_{PTS} (mM/s) es la velocidad de reacción del sistema fosfotransferasa, $C_{glc}^{alimentación}$ (mM) representa la concentración de glucosa en la alimentación al reactor, f_{pulso} (mM/s) corresponde a la perturbación aplicada al sistema, C_x (g PS/L volumen de cultivo) y ρ_x (g PS/L volumen celular) corresponden a la concentración y el peso específico de la biomasa, respectivamente (PS: peso seco).

Los co-metabolitos involucrados (AMP, ADP, ATP, NADH, NAD⁺, NADPH y NADPH⁺), participan en todas las reacciones del metabolismo celular, tanto catabólicas como anabólicas, muchas de ellas no incluidas en el modelo propuesto. Por esta razón, no se han formulado balances de masa para co-metabolitos y sus perfiles temporales se han representado con correlaciones en función del tiempo, calculados a partir de datos experimentales (Chassagnole *et al.*, 2002)

Las Tablas 2.2 y 2.3 muestran los balances de masa para los metabolitos intracelulares y los co-metabolitos, respectivamente.

Tabla 2.2. Balances de masa para metabolitos intracelulares

$$\frac{dC_{g6p}}{dt} = r_{PTS} - r_{PGI} - r_{G6PDH} - r_{PGM} - \mu C_{g6p} \quad (2.3)$$

$$\frac{dC_{f6p}}{dt} = r_{PGI} - r_{PFK} + r_{TKb} + r_{TA} - 2r_{MurSintesis} - \mu C_{f6p} \quad (2.4)$$

$$\frac{dC_{fdp}}{dt} = r_{PKF} - r_{ALDO} - \mu C_{fdp} \quad (2.5)$$

$$\frac{dC_{gap}}{dt} = r_{ALDO} + r_{TIS} - r_{GAPDH} + r_{TKa} + r_{TKb} - r_{TA} + r_{TrpSintesis} - \mu C_{gap} \quad (2.6)$$

$$\frac{dC_{dhap}}{dt} = r_{ALDO} - r_{TIS} - r_{G3PDH} - \mu C_{dhap} \quad (2.7)$$

$$\frac{dC_{pgp}}{dt} = r_{GAPDH} - r_{PGK} - \mu C_{pgp} \quad (2.8)$$

$$\frac{dC_{3pg}}{dt} = r_{PGK} - r_{PGGluMu} - r_{SerSintesis} - \mu C_{3pg} \quad (2.9)$$

$$\frac{dC_{2pg}}{dt} = r_{PGluMu} - r_{ENO} - \mu C_{2pg} \quad (2.10)$$

$$\frac{dC_{pep}}{dt} = r_{ENO} - r_{PK} - r_{PTS} - r_{PEPCxylase} - r_{DAHPS} - r_{Sintesis1} - \mu C_{pep} \quad (2.11)$$

$$\frac{dC_{pyr}}{dt} = r_{PK} + r_{PTS} - r_{PDH} - r_{Sintesis2} + r_{MetSintesis} + r_{TrpSintesis} - \mu C_{pyr} \quad (2.12)$$

$$\frac{dC_{6pg}}{dt} = r_{G6PDH} - r_{PGDH} - \mu C_{6pg} \quad (2.13)$$

$$\frac{dC_{ribu5p}}{dt} = r_{PGDH} - r_{Ru5P} - r_{R5PI} - \mu C_{ribu5p} \quad (2.14)$$

$$\frac{dC_{xyl5p}}{dt} = r_{Ru5P} - r_{TKa} - r_{TKb} - \mu C_{xyl5p} \quad (2.15)$$

$$\frac{dC_{sed7p}}{dt} = r_{TKa} - r_{TA} - \mu C_{sed7p} \quad (2.16)$$

$$\frac{dC_{rib5p}}{dt} = r_{R5PI} - r_{TKa} - r_{RPPK} - \mu C_{rib5p} \quad (2.17)$$

$$\frac{dC_{e4p}}{dt} = r_{TA} - r_{TKb} - r_{DAHPS} - \mu C_{e4p} \quad (2.18)$$

$$\frac{dC_{glp}}{dt} = r_{PGM} - r_{GIPAT} - \mu C_{glp} \quad (2.19)$$

Tabla 2.3. Correlaciones para las concentraciones de los co-metabolitos (Chassagnole *et al.*, 2002)

$$C_{atp} = 4,27 - 4,163 \frac{t}{0,657 + 1,43t + 0,364t^2} \quad (2.20)$$

$$C_{adp} = 0,582 + 1,73 (2,731^{-0,15t}) (0,12t + 0,000214 t^3) \quad (2.21)$$

$$C_{amp} = 0,123 + 7,25 \frac{t}{7,25 + 1,47t + 0,17t^2} + 1,073 \frac{t}{1,29 + 8,05t} \quad (2.22)$$

$$C_{nadph} = 0,062 + 0,332 (2,718^{-0,46t}) (0,0166t^{1,58} + 0,000166t^{4,73} + 1,16 \cdot 10^{-10} t^{7,89} + 1,36 \cdot 10^{-13} t^{11} + 1,23 \cdot 10^{-16} t^{14,2}) \quad (2.23)$$

$$C_{nadp} = 0,159 + 0,00554 \frac{t}{2,8 + 0,271t + 0,01t^2} + 0,182 \frac{t}{4,81 + 0,526t} \quad (2.24)$$

$$C_{nadh} = 0,0934 + 0,0011 (2,371^{-0,123t}) (0,844t + 0,104t^3) \quad (2.25)$$

$$C_{nad} = 1,314 + 1,314 (2,73^{(-0,0435t - 0,342)}) - \frac{(t + 7,871)(2,73^{(-0,0218t - 0,171)})}{8,481 + t} \quad (2.26)$$

2.2.1.2. Cinéticas de reacción

Como se mencionó previamente, los modelos cinéticos permiten predecir la respuesta dinámica del metabolismo celular, es decir cómo evolucionan las concentraciones de metabolitos y los niveles de enzimas en el tiempo. Para esto es imprescindible que además de conocer la estequiometría de las reacciones se incorpore información sobre las cinéticas y regulación de las reacciones

enzimáticas. A continuación se presentan las ecuaciones cinéticas (mecanísticas) y se detallan los mecanismos de regulación para las enzimas que participan en el metabolismo central del carbono en *Escherichia coli*: sistema fosfotransferasa, glucólisis y vía de las pentosas fosfato.

2.2.1.2.1. Sistema fosfotransferasa

El sistema fosfotransferasa (PTS) es un sistema de captación de carbohidratos en *Escherichia coli*. Asimismo este sistema actúa como sensor y está involucrado en el control del consumo de algunos carbohidratos; por ejemplo, si hay presente glucosa en el medio de cultivo, la síntesis de proteínas de transporte de otras fuentes carbonadas así como de las correspondientes enzimas catabólicas se ve reprimida (Kremling *et al.*, 2004).

Se trata de una cascada de fosforilación constituida por cinco reacciones en las que un grupo fosforilo es transferido desde el fosfoenolpiruvato (pep) a través de dos intermediarios, la enzima I (EI) y una proteína de transporte (HPr), hasta el complejo EII^{Glc} y finalmente al sustrato específico, que en este caso es la glucosa. El complejo EII^{Glc} está conformado por la proteína $EIIA^{Crr}$ y por el transportador de membrana $EIICB^{Glc}$, específicos para el consumo de glucosa (Kremling *et al.*, 2004; Kremling *et al.*, 2007). En la Figura 2.2 se muestra el mecanismo de reacción de PTS.

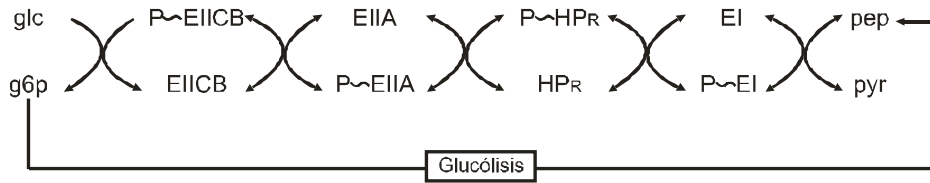


Figura 2.2. Representación esquemática del sistema fosfotransferasa. Fuente: Kremling *et al.* (2004)

Chassagnole *et al.* (2002) proponen una única expresión cinética (Ec. 2.27), desarrollada por Liao *et al.* (1996), para el sistema fosfotransferasa que engloba todos los pasos anteriormente mencionados. Liao *et al.* (1996) proponen esta expresión explicando que el control que ejerce la relación pep:pyr se puede predecir asumiendo que las reacciones que involucran a las enzimas I, HPr y IIA^{Crr} están en equilibrio y que el paso limitante es el de ingreso y fosforilación de la glucosa (Fig. 2.2). La ecuación contempla el efecto inhibitorio que ejerce la glucosa-6-fosfato (g6p) sobre el sistema. La expresión cinética del sistema PTS está dada por la siguiente ecuación:

$$r_{PTS} = \frac{r_{PTS}^{max} C_{glc}^{extracelular} \frac{C_{pep}}{C_{pyr}}}{\left(K_{PTS1} + K_{PTS2} \frac{C_{pep}}{C_{pyr}} + K_{PTS,3} C_{glc}^{extracelular} + C_{glc}^{extracelular} \frac{C_{pep}}{C_{pyr}} \right) \left(1 + \frac{C_{g6p}^{n_{PTS,g6p}}}{K_{PTS,g6p}} \right)} \quad (2.27)$$

En el marco de la presente tesis, se han probado expresiones cinéticas propuestas por otros autores (Rohwer *et al.*, 2000; Kremling *et al.*, 2004), sin obtener mejoras importantes en los resultados.

2.2.1.2.2. Glucólisis

Como se explica en el Capítulo 1 y se muestra en la Figura 2.1, la primera reacción de la vía glucolítica es catalizada por la enzima fosfoglucona isomerasa (PGI). Esta enzima está presente en dos formas en *Escherichia coli*, la forma más abundante representa el 90% de la actividad total y ambas isoenzimas tienen propiedades cinéticas idénticas (Schreyer & Böck, 1980). Es por esto que la actividad de las dos isoenzimas se engloba en una única reacción (Ec. 2.28). La expresión cinética es del tipo Michaelis-Menten reversible (Richter *et al.*, 1975) y contempla la inhibición ejercida por el 6-fosfogluconato (6pg) (Schreyer & Böck, 1980).

$$r_{PGI} = \frac{r_{PGI}^{max} \left(C_{g6p} - \frac{C_{f6p}}{K_{PGI, eq}} \right)}{K_{PGI, g6p} \left(1 + \frac{C_{f6p}}{K_{PGI, f6p} \left(1 + \frac{C_{6pg}}{K_{PGI, f6p, 6pginh}} \right)} + \frac{C_{6pg}}{K_{PGI, g6p, 6pginh}} \right) + C_{g6p}} \quad (2.28)$$

La transferencia de un grupo fosforilo desde el ATP hacia la fructosa-6-fosfato para dar fructosa-1,6-difosfato (fdp), es catalizada por la enzima fosfofructoquinasa (PFK). *Escherichia coli* también posee dos isoenzimas de la PFK, siendo la PFK-1 la que representa el 90% de la actividad total.

En el modelo formulado en el marco de la presente tesis, se incluye la cinética propuesta por Diaz Ricci (1996) para la enzima fosfofructoquinasa (PFK), alternativa a la cinética propuesta por Chassagnole *et al.* (2002). PFK es una enzima alostérica tetramérica que presenta interacción homotrópica cooperativa e

interacción no-cooperativa con respecto a sus sustratos: fructosa-6-fosfato (f6p) y adenosin-tri-fosfato (ATP), respectivamente. La enzima también presenta activación alostérica heterotrópica por adenosin-di-fosfato (ADP), siendo el fosfoenolpiruvato (pep) un inhibidor alostérico. Diaz Ricci (1996) estudió el efecto del pep en el posible comportamiento oscilatorio de la PFK y desarrolló una detallada deducción de la cinética de dicha enzima teniendo en cuenta todos los mecanismos de regulación y el cambio en el estado conformacional (activa (R) o inactiva (T)) de la enzima en presencia de diferentes efectores. La expresión cinética de la fosfofructoquinasa está dada por las ecuaciones 2.29 y 2.30:

$$r_{PFK} = \frac{r_{PFK}^{max} \left[\frac{eC_{f6p}}{K_{Rf6p}} \left(1 + \frac{eC_{f6p}}{K_{Rf6p}} \right)^{n_{PFK}-1} \left(1 + \frac{C_{adp}}{K_{Radp}} \right)^{n_{PFK}} + L\theta\theta c' \frac{C_{f6p}}{K_{Rf6p}} \left(1 + ce' \frac{C_{f6p}}{K_{Rf6p}} \right)^{n_{PFK}-1} \right]}{\left(1 + \frac{eC_{f6p}}{K_{Rf6p}} \right)^{n_{PFK}} \left(1 + \frac{C_{adp}}{K_{Radp}} \right)^{n_{PFK}} + L \left(1 + ce' \frac{C_{f6p}}{K_{Rf6p}} \right)^{C_{pep}/K_{Tpep}}} \quad (2.29)$$

$$L = L_0 \left[\frac{\left(1 + \frac{C_{pep}}{K_{Tpep}} \right) \left(1 + \frac{C_{adp}}{K_{Tadp}} \right)}{\left(1 + \frac{C_{pep}}{K_{Rpep}} \right) \left(1 + \frac{C_{adp}}{K_{Radp}} \right)} \right]^{n_{PFK}} \quad (2.30)$$

donde L es el coeficiente de equilibrio alostérico, L_0 es el valor de L en ausencia de cualquier efector alostérico, K_{Rf6p} , K_{Radp} y K_{Tadp} son constantes de disociación de los sustratos, K_{Tpep} es la constante de disociación del inhibidor, e y e' son razones de constantes catalíticas, Θ es una razón de constantes de disociación, C es un coeficiente de unión no exclusivo y n_{PFK} es el número de sitios de unión.

La enzima fructosa-1,6-difosfato aldolasa (ALDO) cataliza la ruptura de la fructosa-1,6-difosfato (fdp) para dar dos triosas (Capítulo 1). Esta reacción es reversible con un fuerte desplazamiento del equilibrio a la formación de las triosas (dhap y gap) (Nelson & Cox, 2005). Su cinética (Ec. 2.31) está representada por un mecanismo uni-bi ordenado, el cual establece que luego de la formación del complejo enzima-sustrato (por unión del sustrato a la enzima), se libera un producto y luego el otro en forma secuencial (Richter *et al.*, 1975).

$$r_{ALDO} = \frac{r_{ALDO}^{max} \left(C_{fdp} - \frac{C_{gap} C_{dhap}}{K_{ALDO,eq}} \right)}{K_{ALDO,fdp} + C_{fdp} + \frac{K_{ALDO,gap} C_{dhap}}{K_{ALDO,eq} V_{ALDO,bif}} + \frac{K_{ALDO,dhap} C_{gap}}{K_{ALDO,eq} V_{ALDO,bif}} + \frac{C_{fdp} C_{gap}}{K_{ALDO,gap,inh}} + \frac{C_{dhap} C_{gap}}{K_{ALDO,eq} V_{ALDO,bif}}} \quad (2.31)$$

La siguiente reacción de la glucólisis, correspondiente a la interconversión de las triosas, es catalizada por la enzima triosa fosfato isomerasa (TIS), cuya cinética se representa por una ecuación de tipo Michaelis-Menten reversible (Richter *et al.*, 1975):

$$r_{TIS} = \frac{r_{TIS}^{max} \left(C_{dhap} - \frac{C_{gap}}{K_{TIS,eq}} \right)}{K_{TIS,dhap} \left(1 + \frac{C_{gap}}{K_{TIS,gap}} \right) + C_{dhap}} \quad (2.32)$$

donde $K_{TIS,gap}$ y $K_{TIS,dhap}$ son constantes de saturación media y $K_{TIS,eq}$ es la constante de equilibrio de la reacción.

Las cinéticas de las reacciones catalizadas por las enzimas gliceraldehído-3-fosfato deshidrogenasa (GAPDH) y fosfogliceratoquinasa (PGK), se han modelado según Chassagnole *et al.* (2002), como cinéticas de Michaelis-Menten reversibles de dos sustratos (Ecs. 2.33 y 2.34).

$$r_{GAPDH} = \frac{r_{GAPDH}^{max} \left(C_{gap} C_{nad} - \frac{C_{pgp} C_{nadh}}{K_{GAPDH,eq}} \right)}{\left(K_{GAPDH,gap} \left(1 + \frac{C_{pgp}}{K_{GAPDH,pgp}} \right) + C_{gap} \right) \left(K_{GAPDH,nad} \left(1 + \frac{C_{nadh}}{K_{GAPDH,nadh}} \right) + C_{nad} \right)} \quad (2.33)$$

$$r_{PGK} = \frac{r_{PGK}^{max} \left(C_{adp} C_{pgp} - \frac{C_{atp} C_{3pg}}{K_{PGK,eq}} \right)}{\left(K_{PGK,adp} \left(1 + \frac{C_{atp}}{K_{PGK,atp}} \right) + C_{adp} \right) \left(K_{PGK,pgp} \left(1 + \frac{C_{3pg}}{K_{PGK,3pg}} \right) + C_{pgp} \right)} \quad (2.34)$$

En la ecuación (2.33) $K_{GAPDH,gap}$ y $K_{GAPDH,nad}$ son constantes de saturación media, $K_{GAPDH,pgp}$ y $K_{GAPDH,nadh}$ representan constantes de inhibición competitiva y $K_{GAPDH,eq}$ es la constante de equilibrio de reacción. En la ecuación 2.34 $K_{PGK,adp}$ y $K_{PGK,pgp}$ son las constantes de saturación media, $K_{PGK,eq}$ representa la constante de equilibrio de reacción y $K_{PGK,atp}$ y $K_{PGK,3pg}$ son las constantes de inhibición competitiva.

Los dos pasos siguientes en la glucólisis, correspondientes a la conversión de 3-fosfoglicerato (3pg) a 2-fosfoglicerato (2pg) y la deshidratación del 2pg a fosfoenolpiruvato (pep), son catalizados por las enzimas fosfoglucomutasa y

enolasa, respectivamente. Ambas reacciones son reversibles, representadas por ecuaciones de Michaelis-Menten reversibles (Ecs. 2.35 y 2.36).

$$r_{PGluMu} = \frac{r_{PGluMu}^{max} \left(C_{3pg} - \frac{C_{2pg}}{K_{PGluMu,eq}} \right)}{K_{PGluMu,3pg} \left(1 + \frac{C_{2pg}}{K_{PGluMu,2pg}} \right) + C_{3pg}} \quad (2.35)$$

$$r_{ENO} = \frac{r_{ENO}^{max} \left(C_{2pg} - \frac{C_{pep}}{K_{ENO,eq}} \right)}{K_{ENO,2pg} \left(1 + \frac{C_{pep}}{K_{ENO,pep}} \right) + C_{2pg}} \quad (2.36)$$

En la ecuación 2.35 $K_{PGluMu,3pg}$ y $K_{PGluMu,2pg}$ son las constantes de saturación media, mientras que $K_{PGluMu,eq}$ representa la constante de reacción. Asimismo, en la ecuación 2.36 $K_{ENO,2pg}$ y $K_{ENO,pep}$ corresponden a las constantes de saturación media y $K_{ENO,eq}$ es la constante de equilibrio de reacción.

La última reacción de la glucólisis corresponde a la transferencia del grupo fosforilo desde el fosfoenolpiruvato (pep) al ADP, para dar como productos piruvato (pyr) y ATP y es catalizada por la enzima piruvato quinasa (PK). *Escherichia coli* posee dos isoenzimas, PKI y PKII; la primera es activada por la fructosa-1,6-difosfato ($K_{PK,fdp}$) e inhibida por ATP ($K_{PK,atp}$), mientras que PKII es regulada por un mecanismo de activación por AMP ($K_{PK,amp}$). Ambas isoenzimas muestran propiedades alostéricas respecto del fosfoenolpiruvato (n_{PK}) (Kornberg & Malcovati, 1973; Chassagnole *et al.*, 2002; Muñoz & Ponce, 2003). La ecuación

2.37, corresponde a la expresión cinética de la enzima piruvato quinasa (engloba la actividad de las isoenzimas PKI y PKII). La formulación se basa en la ecuación de Johannes & Hess (1973) con un término adicional que representa la activación por AMP ($K_{PK,amp}$).

$$r_{PK} = \frac{r_{PK}^{max} C_{pep} C_{adp} \left(\frac{C_{pep}}{K_{PK,pep}} + 1 \right)^{(n_{PK}-1)}}{K_{PK,pep} \left(L_{PK} \left(\frac{1 + \frac{C_{atp}}{K_{PK,atp}}}{\frac{C_{fdp}}{K_{PK,fdp}} + \frac{C_{amp}}{K_{PK,amp}} + 1} \right)^{n_{PK}} + \left(\frac{C_{pep}}{K_{PK,pep}} + 1 \right)^{n_{PK}} \right) (C_{adp} + K_{PK,adp})} \quad (2.37)$$

El modelo propuesto para la red metabólica de *Escherichia coli* incluye la descarboxilación del piruvato para dar acetil-coA (accoa), catalizada por la enzima piruvato deshidrogenasa (PDH), cuya cinética es descrita por la ecuación 2.38, correspondiente a la ecuación de Hill que describe en forma simplificada la velocidad de reacción de un enzima alostérica con n (n_{PDH}) sitios activos.

$$r_{PDH} = \frac{r_{PDH}^{max} C_{pyr}^{n_{PDH}}}{K_{PDH,pyr} + C_{pyr}^{n_{PDH}}} \quad (2.38)$$

2.2.1.2.3. Ruta de las pentosas fosfato

Luego de la fosforilación de la glucosa para dar glucosa-6-fosfato, el flujo de carbono se distribuye entre la glucólisis y la vía de las pentosas (Capítulo 1). La enzima glucosa-6-fosfato deshidrogenasa (G6PDH) cataliza la primera reacción

correspondiente a la vía de las pentosas, es por esto que su cinética y regulación han sido ampliamente estudiadas (Sanwal, 1970; Ratushny *et al.*, 2006). Dicha enzima exhibe un mecanismo de regulación complejo; NADPH es un inhibidor competitivo del NADP^+ y un inhibidor no-competitivo de la glucosa-6-fosfato. G6PDH presenta inhibición alostérica heterotrópica por NADH, y además este co-metabolito causa inhibición no-competitiva con respecto a la glucosa-6-fosfato (Sanwal, 1970). Ratushny *et al.* (2006) desarrolló una expresión cinética para la glucosa-6-fosfato deshidrogenasa (Ec. 2.39) a través de la aplicación del método de la función generalizada de Hill. Dicha ecuación contempla todos los mecanismos de regulación descritos anteriormente. En el marco de la presente tesis, se ha incluido dicha ecuación en lugar de la inicialmente propuesta por Chassagnole *et al.* (2002) debido a que representa con mayor rigurosidad los mecanismos descritos y permite un mejor ajuste de los datos experimentales.

$$r_{G6PDH} = \frac{A \cdot B}{C \cdot D} \quad (2.39)$$

donde

$$A = r_{G6PDH}^{max} C_{g6p}$$

$$B = \left(\frac{C_{nadp}}{K_{G6PDH,nadp}} \right)^{1+k_{dtn} \frac{C_{nadh}^{h_{dt}}}{k_{dt}^{h_{dt}} + C_{nadh}^{h_{dt}}}}$$

$$C = \left(K_{G6PDH,g6p} \left(1 + k_{tgn} \frac{C_{nadp}^{h_{tg}}}{k_{tg}^{h_{tg}} + C_{nadp}^{h_{tg}}} \right) + C_{g6p} \right) \left(1 + \frac{C_{nadph}}{K_{G6PDH,nadphin}} + \frac{C_{nadh}}{K_{G6PDH,nadhin}} \right)$$

$$D = \left(1 + \left(\frac{C_{nadp}}{K_{G6PDH,nadp}} \right)^{1+k_{dtn} \frac{C_{nadh}^{h_{dt}}}{k_{dt}^{h_{dt}} + C_{nadh}^{h_{dt}}} + \left(\frac{C_{nadph}}{K_{G6PDH,nadphin2}} \right)^{h_{nadph}} \right) \left(1 + \left(\frac{C_{nadh}}{K_{G6PDH,nadh}} \right)^{h_{nadh}} \right)$$

En la ecuación 2.39 $K_{G6PDH,g6p}$ y $K_{G6PDH,nadp}$ corresponden a las constantes de saturación media de reacción; $K_{G6PDH,nadphin}$ es la constante de inhibición por NADPH; h_{nadph} es una constante que explica la no linealidad del efecto del NADPH en la velocidad de reacción; k_{dtn} , k_{dt} y $K_{G6PDH,nadh}$ son constantes que determinan la eficiencia de efecto del NADH en la velocidad de reacción; las constantes h_{dt} y h_{nadh} explican la no linealidad del efecto del NADH en la velocidad de reacción; k_{tg} y k_{tgn} son constantes que determinan la eficiencia del efecto del NADP en la velocidad de reacción de la enzima G6PDH. Por último, la constante h_{tg} explica la no linealidad del efecto del NADH en la velocidad de reacción.

El siguiente paso en la ruta de las pentosas fosfato es la oxidación y descarboxilación del 6-fosfogluconato (6pg) para dar ribulosa-5-fosfato (ribu5p) según:

$$r_{PGDH} = \frac{r_{PGDH}^{max} C_{6pg} C_{nadp}}{\left(C_{6pg} + K_{PGDH,6pg} \right) \left(C_{nadp} + K_{PGDH,nadp} \left(1 + \frac{C_{nadph}}{K_{PGDH,nadph,inh}} \right) \left(1 + \frac{C_{atp}}{K_{PGDH,atp,inh}} \right) \right)} \quad (2.40)$$

Esta reacción es catalizada por la enzima 6-fosfogluconato deshidrogenasa (PGDH), la cual presenta inhibición por ATP ($K_{PGDH,atp,inh}$) y NADPH ($K_{PGDH,nadph,inh}$) (Orozco de Silva y Fraenkel., 1979; Ratushny *et al.*, 2006). La ecuación 2.40 describe la cinética de reacción de la enzima PGDH teniendo en cuenta los mecanismos de regulación (Vaseghi *et al.* 1999).

Tanto la última reacción de la fase oxidativa de la ruta de las pentosas fosfato como las reacciones correspondientes a la fase no oxidativa (Capítulo 1) son tratadas como reacciones en equilibrio, siendo descritas por ecuaciones cinéticas de primer orden reversibles (Ecs. 2.41 a 2.45), según Vaseghi *et al.* (1999):

$$r_{R5PI} = r_{R5PI}^{max} \left(C_{ribu5p} - \frac{C_{rib5p}}{K_{R5PI,eq}} \right) \quad (2.41)$$

$$r_{Ru5p} = r_{Ru5p}^{max} \left(C_{ribu5p} - \frac{C_{xyl5p}}{K_{Ru5p,eq}} \right) \quad (2.42)$$

$$r_{TKa} = r_{TKa}^{max} \left(C_{rib5p} C_{xyl5p} - \frac{C_{sed7p} C_{gap}}{K_{TKa,eq}} \right) \quad (2.43)$$

$$r_{TKb} = r_{TKb}^{max} \left(C_{xyl5p} C_{e4p} - \frac{C_{f6p} C_{gap}}{K_{TKb,eq}} \right) \quad (2.44)$$

$$r_{TA} = r_{TA}^{max} \left(C_{gap} C_{sed7p} - \frac{C_{e4p} C_{f6p}}{K_{TA,eq}} \right) \quad (2.45)$$

2.2.1.2.4. Reacciones de biosíntesis y anapleróticas

La descarboxilación del fosfoenolpiruvato para dar oxalacetato (ooa) (reacción anaplerótica) es catalizada por la enzima fosfoenolpiruvato carboxilasa (PepCxlase), según:

$$r_{PEPCxlase} = \frac{r_{PEPCxlase}^{max} C_{pep} \left(1 + \left(\frac{C_{fdp}}{K_{PEPCxlase,fdp}} \right)^{n_{PEPCxlase,fdp}} \right)}{K_{PEPCxlase,pep} + C_{pep}} \quad (2.46)$$

Según Kai *et al.* (2003) esta enzima es activada alostéricamente por fructosa-1,6-difosfato. En la ecuación 2.46, el término en el numerador de la expresión, que contiene los parámetros $n_{PEPCxlase,fdp}$ y $K_{PEPCxlase,fdp}$, contempla el mecanismo de regulación alostérica.

Las reacciones de biosíntesis incluyen una serie de reacciones que conducen al producto final (aminoácidos, nucleótidos, entre otros). Sin embargo con el fin de simplificar el modelo, Chassagnole *et al.* (2002) proponen describir cada una de las rutas de biosíntesis con una única ecuación que contemple todas las reacciones o bien con la primera reacción de la ruta. En este sentido, el consumo de pep para la síntesis de mureína y corismato, se describe por la reacción general Síntesis1 (Ec. 2.47); el consumo de pyr destinado a la síntesis de aminoácidos se representa por una única reacción denominada Síntesis2 (Ec. 2.48), así como la síntesis de serina se describe por una única reacción (SerSíntesis) que requiere el consumo de pep (Ec. 2.49). Las tres reacciones globales son modeladas con una cinética de tipo Michaelis-Menten, según:

$$r_{Sintesis1} = \frac{r_{Sintesis1}^{max} C_{pep}}{K_{Sintesis1, pep} + C_{pep}} \quad (2.47)$$

$$r_{Sintesis2} = \frac{r_{Sintesis2}^{max} C_{pyr}}{K_{Sintesis2, pyr} + C_{pyr}} \quad (2.48)$$

$$r_{SerSintesis} = \frac{r_{SerSintesis}^{max} C_{3pg}}{K_{SerSintesis, 3pg} + C_{3pg}} \quad (2.49)$$

Asimismo, las ecuaciones cinéticas de las enzimas ribosa-fosfato pirofosfoquinasa (RPPK, reacción que conduce a las vías de síntesis de nucleótidos), glicerol-3-fosfato deshidrogenasa (G3PDH, enzima que cataliza la síntesis de glicerol requerido para la síntesis de lípidos y fosfolípidos), se describen con cinéticas tipo Michaelis-Menten (Ecs. 2.50 y 2.51).

$$r_{RPPK} = \frac{r_{RPPK}^{max} C_{rib5p}}{K_{RPPK, rib5p} + C_{rib5p}} \quad (2.50)$$

$$r_{G3PDH} = \frac{r_{G3PDH}^{max} C_{dhap}}{K_{G3PDH, dhap} + C_{dhap}} \quad (2.51)$$

La cinética de la enzima DAHPS (3-deoxy-D-arabino-heptulosonato-7-fosfato sintasa) que cataliza la primera reacción en la síntesis de compuestos aromáticos (Akowski y Bauerle, 1997), se describe con la ecuación de Hill para dos sustratos (Ec. 2.52):

$$r_{DAHPS} = \frac{r_{DAHPS}^{max} C_{e4p}^{n_{DAHPS, e4p}} C_{pep}^{n_{DAHPS, pep}}}{\left(K_{DAHPS, e4p} + C_{e4p}^{n_{DAHPS, e4p}} \right) \left(K_{DAHPS, pep} + C_{pep}^{n_{DAHPS, pep}} \right)} \quad (2.52)$$

Por otra parte, también se incluyen en el modelo las reacciones que conducen a la síntesis de polisacáridos, es decir la transferencia del grupo fosfato desde la posición 6, en la glucosa-6-fosfato, a la posición 1, en la glucosa-1-fosfato, catalizada por la enzima fosfoglucomutasa (PGM); y la primera reacción que conduce a la biosíntesis de polisacáridos catalizada por la glucosa-1-fosfato adeniltransferasa (G1PAT). La cinética de la enzima PGM se describe mediante una ecuación de tipo Michaelis-Menten reversible (Ec. 2.53), mientras que la cinética de la G1PAT se modela con una ecuación empírica de dos sustratos que contempla la activación alostérica ($K_{G1PAT,fdp}$ y $n_{G1PAT,fdp}$) por el intermediario glucolítico fdp (Ec. 2.54) (Preiss *et al.*, 1975).

$$r_{PGM} = \frac{r_{PGM}^{max} \left(C_{g6p} - \frac{C_{g1p}}{K_{PGM, eq}} \right)}{K_{PGM, g6p} \left(1 + \frac{C_{g1p}}{K_{PGM, g1p}} \right) + C_{g6p}} \quad (2.53)$$

$$r_{G1PAT} = \frac{r_{G1PAT}^{max} C_{g1p} C_{atp} \left(1 + \left(\frac{C_{fdp}}{K_{G1PAT, fdp}} \right)^{n_{G1PAT, fdp}} \right)}{\left(K_{G1PAT, g1p} + C_{g1p} \right) \left(K_{G1PAT, atp} + C_{atp} \right)} \quad (2.54)$$

Por último las cinéticas de las reacciones de síntesis de mureína, triptófano y metionina se han considerado constantes e iguales a su velocidad máxima (Ecs. 2.55, 2.56 y 2.57, respectivamente) (Chassagnole *et al.*, 2002).

$$r_{MurSintesis} = r_{MurSintesis}^{max} \quad (2.55)$$

$$r_{TrpSintesis} = r_{TrpSintesis}^{max} \quad (2.56)$$

$$r_{MetSintesis} = r_{MetSintesis}^{max} \quad (2.57)$$

Se han analizado y probado numerosas expresiones cinéticas alternativas para diferentes enzimas en el modelo de la red metabólica de *E. coli*, pero sin lograr mejoría en los resultados en la estimación de parámetros. En este sentido, se evaluaron las expresiones cinéticas para las enzimas Aldolasa (Degenring, 2004 (Ec. 2.59); Ceric *et al.*, 2006 (Ec. 2.63)), Fosfoglucoisomerasa (Degenring, 2004 (Ec. 2.58)), Fosfogliceratoquinasa, Fosfoglucomutasa y Enolasa (Ecs. 2.60, 2.61 y 2.62, respectivamente; Usuda *et al.*, 2010).

$$r_{PGI} = k_{f,PGI} C_{g6p}^{n_{PGI,1}} - k_{b,PGI} C_{f6p}^{n_{PGI,2}} \quad (2.58)$$

$$r_{ALDO} = \frac{r_{ALDO}^{max} C_{fdp}^{n_{ALDO,1}}}{K_{m,ALDO}^{n_{ALDO,1}} + C_{fdp}^{n_{ALDO,1}}} \left(\frac{C_{gap}}{K_{gap,ALDO}} \right)^{n_{ALDO,2}} \quad (2.59)$$

$$r_{PGK} = \frac{\left(\frac{r_{PGK,f}^{max} C_{pgp} C_{adp}}{K_{m,pgp} K_{m,adp}} \right) - \left(\frac{r_{PGK,r}^{max} C_{3pg} C_{atp}}{K_{m,3pg} K_{m,atp}} \right)}{\left(1 + \frac{C_{pgp}}{K_{m,pgp}} + \frac{C_{3pg}}{K_{m,3pg}} \right) \left(1 + \frac{C_{adp}}{K_{m,adp}} + \frac{C_{atp}}{K_{m,atp}} \right)} \quad (2.60)$$

$$r_{PGM} = \frac{\left(\frac{r_{PGM,f}^{max} C_{3pg}}{K_{m,3pg}} \right) - \left(\frac{r_{PGM,r}^{max} C_{2pg}}{K_{m,2pg}} \right)}{1 + \frac{C_{3pg}}{K_{m,3pg}} + \frac{C_{2pg}}{K_{m,2pg}}} \quad (2.61)$$

$$r_{ENO} = \frac{\left(\frac{r_{ENO,f}^{max} C_{2pg}}{K_{m,2pg}} \right) - \left(\frac{r_{ENO,r}^{max} C_{pep}}{K_{m,pep}} \right)}{1 + \frac{C_{2pg}}{K_{m,2pg}} + \frac{C_{pep}}{K_{m,pep}}} \quad (2.62)$$

$$r_{ALDO} = \frac{r_{ALDO}^{max} \left(C_{fdp} - \frac{C_{gap} C_{dhap}}{K_{ALDO,eq}} \right) \left(1 + \frac{C_{f6p}}{K_{ALDO,a,f6p}} \right)}{\left(K_{ALDO,m,fdp} + C_{fdp} \left(1 + \frac{C_{gap}}{K_{ALDO,i,gap}} \right) + \frac{\left(1 + \frac{C_{f6p}}{K_{ALDO,a,f6p}} \right)^n}{K_{ALDO,eq}} \left(K_{ALDO,m,dhap} C_{gap} + K_{ALDO,m,dhap} C_{dhap} + C_{gap} C_{dhap} \right) \right)} \quad (2.63)$$

2.2.2. Extensión del modelo metabólico de *Escherichia coli*

2.2.2.1. Balances de masa

El modelo cinético de la red metabólica de *E. coli* descrito en la sección anterior (PTS, glucólisis, PPP) se ha extendido mediante la adición de las reacciones correspondientes a las rutas fermentativas (Capítulo 1). Los balances de masa para los intermediarios de reacción intracelulares, acetil-P y acetaldehído; se han formulado según la ecuación 2.1. Los productos de fermentación se consideran metabolitos extracelulares ya que se asume que su transporte desde el citoplasma

al medio de cultivo es lo suficientemente rápido (Stephanopoulos *et al.*, 1998) y los balances de masa se han formulado según la ecuación general:

$$\frac{dC_i}{dt} = v_{ij} r_j \left(\frac{C_x}{\rho_x} \right) - DC_i \quad (2.64)$$

$i = \text{etanol, formiato, acetato, lactato, succinato}$
 $j = \text{LDH, ACK, ADH, PFL, SucSintesis}$

donde C_i (mM) es la concentración del producto de fermentación i , r_j (mM/s) corresponde a la expresión cinética de la enzima j , v_{ij} es el coeficiente estequiométrico del metabolito i en la reacción j (< 0 para reactivos, > 0 para productos), D (s^{-1}) es la velocidad de dilución y C_x (g PS/L volumen de cultivo) y ρ_x (g PS/L volumen celular) son la concentración y densidad de biomasa, respectivamente (PS: peso seco).

La Tabla 2.4 muestra los balances de masa modificados de los metabolitos intracelulares y los balances de masa los nuevos metabolito intra y extracelulares contemplados en el modelo.

Tabla 2.4. Balances de masa para metabolitos intracelulares y extracelulares en modelo extendido

$$\frac{dC_{pyr}}{dt} = r_{PK} + r_{PTS} - r_{PDH} - r_{Synth2} + r_{MetSintesis} + r_{TrpSintesis} - r_{LDH} - r_{PFL} - \mu C_{pyr} \quad (2.65)$$

$$\frac{dC_{acal}}{dt} = r_{ACALDH} - r_{ADH} - \mu C_{acal} \quad (2.66)$$

$$\frac{dC_{acp}}{dt} = r_{PTA} - r_{ACK} - \mu C_{acp} \quad (2.67)$$

$$\frac{dC_{etanol}}{dt} = r_{ADH} \left(\frac{C_X}{\rho_X} \right) - DC_{etanol} \quad (2.68)$$

$$\frac{dC_{formiato}}{dt} = r_{PFL} \left(\frac{C_X}{\rho_X} \right) - DC_{formiato} \quad (2.69)$$

$$\frac{dC_{lactato}}{dt} = r_{LDH} \left(\frac{C_X}{\rho_X} \right) - DC_{lactato} \quad (2.70)$$

$$\frac{dC_{acetato}}{dt} = r_{ACK} \left(\frac{C_X}{\rho_X} \right) - DC_{acetato} \quad (2.71)$$

$$\frac{dC_{succinato}}{dt} = r_{SucSintesis} \left(\frac{C_X}{\rho_X} \right) - DC_{succinato} \quad (2.72)$$

2.2.2.2. Cinéticas de reacción

Las enzimas fosfotransacetilasa (PTA) y acetato quinasa (ACK) conducen a la síntesis de etanol a partir de acetil-coA (Capítulo 1). La cinética de la PTA (Henkin & Abeles, 1976) muestra la inhibición competitiva de los pares de sustrato-producto de dicha enzima, según:

$$r_{PTA} = \frac{r_{PTA}^{max} \left(\frac{1}{K_{i,AcCoA} K_{mP}} \right) \left(C_{AcCoA} C_P - \frac{C_{acp} C_{CoA}}{K_{eq}} \right)}{\left(1 + \frac{C_{AcCoA}}{K_{i,AcCoA}} + \frac{C_P}{K_{i,P}} + \frac{C_{acp}}{K_{i,acp}} + \frac{C_{CoA}}{K_{i,CoA}} + \frac{C_{AcCoA} C_P}{K_{i,AcCoA} K_{m,P}} + \frac{C_{acp} C_{CoA}}{K_{m,acp} K_{i,CoA}} \right)} \quad (2.73)$$

Asimismo la cinética de la enzima acetato quinasa (ACK) se modela del tipo Michaelis-Menten reversible con pareja de sustrato-producto no competitiva, según (Hoefnagel *et al.*, 2002):

$$r_{ACK} = \frac{r_{ACK}^{max} (K_{m,adp} K_{m,acp})^{-1} \left(C_{acp} C_{adp} - \frac{C_{acetato} C_{atp}}{K_{eq}} \right)}{\left(1 + \frac{C_{acp}}{K_{m,acp}} + \frac{C_{acetato}}{K_{m,acetato}} \right) \left(1 + \frac{C_{adp}}{K_{m,adp}} + \frac{C_{atp}}{K_{m,atp}} \right)} \quad (2.74)$$

La síntesis de etanol a partir de acetil-CoA es catalizada por las enzimas acetaldehído deshidrogenasa (ACALDH) y alcohol deshidrogenasa (ADH). La primera de ellas actúa en la oxidación del acetil-coA para dar acetaldehído, mientras que la ADH cataliza la conversión del acetaldehído a etanol. Ambas cinéticas se describen con una expresión del tipo Michaelis-Menten reversible con pareja de sustrato-producto no competitiva, según las ecuaciones (Hoefnagel *et al.*, 2002):

$$r_{ACALDH} = \frac{r_{ACALDH}^{max} \left(\frac{1}{K_{m,AcCoA} K_{m,nadh}} \right) \left(C_{AcCoA} C_{nadh} - \frac{C_{CoA} C_{nad} C_{acal}}{K_{eq}} \right)}{\left(1 + \frac{C_{nad}}{K_{m,nad}} + \frac{C_{nadh}}{K_{m,nadh}} \right) \left(1 + \frac{C_{AcCoA}}{K_{m,AcCoA}} + \frac{C_{CoA}}{K_{m,CoA}} + \frac{C_{acal}}{K_{m,acal}} + \frac{C_{acal} C_{CoA}}{K_{m,acal} K_{m,CoA}} \right)} \quad (2.75)$$

$$r_{ADH} = \frac{r_{ADH}^{max} (K_{m,acal} K_{m,nadh})^{-1} \left(C_{acal} C_{nadh} - \frac{C_{etanol} C_{nad}}{K_{eq}} \right)}{\left(1 + \frac{C_{nad}}{K_{m,nad}} + \frac{C_{nadh}}{K_{m,nadh}} \right) \left(1 + \frac{C_{acal}}{K_{m,acal}} + \frac{C_{etanol}}{K_{m,etanol}} \right)} \quad (2.76)$$

La enzima piruvato formiato liasa (PFL) cataliza la conversión reversible de piruvato y CoA en formiato y acetil-CoA. Knappe *et al.* (1974) estudiaron las propiedades de la enzima PFL, obteniendo como resultado que el formiato (producto) es un inhibidor competitivo respecto de la coA y un inhibidor no

competitivo respecto del piruvato. Asimismo, Knappe et al. (1974) proponen un mecanismo de tipo ping-pong para describir la cinética de la enzima PFL (Ec. 2.77), este mecanismo se utiliza para explicar las cinéticas de reacciones de dos o más sustratos en el cual uno o más productos deben ser liberados antes que todos los sustratos reaccionen (Marangoni, 2003), según la siguiente ecuación:

$$r_{PFL} = \frac{r_{PFL,for}^{max} C_{pyr} C_{CoA}}{(C_{pyr} C_{CoA} + C_{pyr} K_{mCoA} + C_{CoA} K_{mPyr})} - \frac{r_{PFL,rev}^{max} C_{formiato} C_{AcCoA}}{(C_{formiato} C_{AcCoA} + C_{formiato} K_{mAcCoA} + C_{AcCoA} K_{mformiato})} \quad (2.77)$$

La producción de lactato a partir de piruvato es catalizada por la enzima lactato deshidrogenasa, cuya cinética es del tipo Michaelis-Menten reversible con pareja de sustrato-producto no competitiva (Ec. 2.78) (Hoefnagel et al., 2002):

$$r_{LDH} = \frac{r_{LDH}^{max} (K_{m,pyr} K_{m,nadh})^{-1} \left(C_{pyr} C_{nadh} - \frac{C_{lactato} C_{nad}}{K_{eq}} \right)}{\left(1 + \frac{C_{pyr}}{K_{m,pyr}} + \frac{C_{lac}}{K_{m,lactato}} \right) \left(1 + \frac{C_{nadh}}{K_{m,nadh}} + \frac{C_{nad}}{K_{m,nad}} \right)} \quad (2.78)$$

Por último, la síntesis de succinato a partir de fosfoenolpiruvato (Capítulo 1) se modela según una única reacción global (SucSíntesis), cuya cinética se describe con una expresión del tipo Michaelis-Menten:

$$r_{SucSintesis} = \frac{r_{SucSintesis}^{max} C_{pep}}{K_{m, pep} + C_{pep}} \quad (2.79)$$

La Tabla 2.5 muestra los parámetros cinéticos involucrados en el modelo del metabolismo de *Escherichia coli*, así como su descripción y unidades.

Tabla 2.5. Nomenclatura, descripción y unidades de los parámetros del modelo del metabolismo de *Escherichia coli*

Enzima	Parámetros	Valor	Descripción
Sistema fosfotransferasa (PTS)	K_{PTS1}	0,02	Constante de saturación media (mM)
	K_{PTS2}	0,03	Constante (mM)
	K_{PTS3}	0,09	Constante
	$K_{PTS,g6p}$	0,3	Constante de inhibición (mM)
	$n_{PTS,g6p}$	4	Constante
	r_{PTS}^{max}	4000	Velocidad máxima de reacción (mM s ⁻¹)
Fosfoglucoisomerasa (PGI) (EC: 5.3.1.9)	$K_{PGI,g6p}$	4,46	Constante de saturación media (mM)
	$K_{PGI,f6p}$	0,266	Constante de inhibición (mM)
	$K_{PGI,eq}$	0,43	Constante de equilibrio
	$K_{PGI,g6p,6pg,inh}$	0,2	Constante de inhibición (mM)
	$K_{PGI,f6p,6pg,inh}$	0,2	Constante de inhibición (mM)
	r_{PGI}^{max}	8,694	Velocidad maxima de reacción (mM s ⁻¹)
Fosfofructoquinasa (PFK) (EC: 2.7.1.11)	e	0,99	Razón de constantes catalítica
	θ	1	Razón de constantes de disociación
	C	0,0005	Coefficiente de unión no-exclusivo
	e'	0,999	Razón de constantes catalítica
	K_{Rf6p}	0,054	Coefficiente de disociación del sustrato (mM)
	K_{Radp}	0,0025	Coefficiente de disociación del sustrato (mM)
	K_{Tpep}	0,75	Coefficiente de disociación del inhibidor (mM)
	K_{Tadp}	1,3	Coefficiente de disociación del sustrato (mM)
	L_0		Constante alostérica
	n_{PFK}	4	Número de sitios de unión
	r_{PFK}^{max}	0,5	Velocidad máxima de reacción (mM s ⁻¹)
Aldolasa (ALDO) (EC: 4.1.2.13)	$K_{ALDO,fdp}$	0,075	Constante de saturación media (mM)
	$K_{ALDO,dhap}$	0,088	Constante de saturación media (mM)
	$K_{ALDO,gap}$	0,088	Constante de saturación media (mM)
	$K_{ALDO,gap,inh}$	1,2	Constante de inhibición (mM)

	$V_{ALDO,bif}$	0,2	Relación de velocidad de reacción directa/reversa
	$K_{ALDO,eq}$	0,144	Constante de equilibrio (mM)
	r_{ALDO}^{max}	1,145	Velocidad máxima de reacción (mM s ⁻¹)
Triosafosfato isomerasa (TIS) (EC: 5.3.1.1)	$K_{TIS,dhap}$	2,8	Constante de saturación media(mM)
	$K_{TIS,gap}$	0,3	Constante de saturación media (mM)
	$K_{TIS,eq}$	1,39	Constante de equilibrio
	r_{TIS}^{max}	102,056	Velocidad máxima de reacción (mM s ⁻¹)
Gliceraldehído-3-fosfato deshidrogenasa (GAPDH) (EC: 1.2.1.12)	$K_{GAPDH,gap}$	0,19	Constante de saturación media (mM)
	$K_{GAPDH,pgp}$	1,1E-4	Constante de inhibición (mM)
	$K_{GAPDH,nad}$	0,252	Constante de saturación media (mM)
	$K_{GAPDH,nadh}$	1,09	Constante de inhibición (mM)
	$K_{GAPDH,eq}$	0,063	Constante de equilibrio
	r_{GAPDH}^{max}	1000	Velocidad máxima de reacción (mM s ⁻¹)
Fosfoglicerato quinasa (PGK) (EC: 2.7.2.3)	$K_{PGK,pgp}$	0,0468	Constante de saturación media (mM)
	$K_{PGK,3pg}$	0,473	Constante de inhibición (mM)
	$K_{PGK,adp}$	0,185	Constante de saturación media (mM)
	$K_{PGK,atp}$	0,653	Constante de inhibición (mM)
	$K_{PGK,eq}$	180	Constante de equilibrio
	r_{PGK}^{max}	2226,032	Velocidad máxima de reacción (mM s ⁻¹)
Fosfoglicerato mutasa (PGluMu) (EC: 5.4.2.1)	$K_{PGluMu,3pg}$	0,2	Constante de saturación media (mM)
	$K_{PGluMu,2pg}$	0,369	Constante de inhibición (mM)
	$K_{PGluMu,eq}$	0,187	Constante de equilibrio
	r_{PGluMu}^{max}	76,972	Velocidad máxima de reacción (mM s ⁻¹)
Enolasa (ENO) (EC: 4.2.1.11)	$K_{ENO,2pg}$	0,1	Constante de saturación media (mM)
	$K_{ENO,pep}$	0,135	Constante de inhibición (mM)
	$K_{ENO,eq}$	6,7	Constante de equilibrio
	r_{ENO}^{max}	452,742	Velocidad máxima de reacción (mM s ⁻¹)
Piruvato quinasa (PK) (EC: 2.7.1.40)	$K_{PK,pep}$	0,1523	Constante de saturación media (mM)
	$K_{PK,adp}$	0,26	Constante de saturación media (mM)
	$K_{PK,atp}$	22,5	Constante de inhibición (mM)
	$K_{PK,fdp}$	0,2591	Constante de activación (mM)
	$K_{PK,amp}$	0,2	Constante de activación (mM)
	L_{PK}	1000	Constante alostérica
	n_{PK}	4	Número de sitios de unión
	r_{PK}^{max}	0,2673	Velocidad máxima de reacción (mM s ⁻¹)
Piruvato deshidrogenasa (PDH) (EC: 1.2.4.1)	$K_{PDH,pyr}$	1,3357	Constante de saturación media (mM)
	n_{PDH}	1	Número de sitios activos

	r_{PDH}^{max}	4,1363	Velocidad máxima de reacción (mM s ⁻¹)
Fosfoenolpiruvato carboxilasa (PEP Cxylase) (EC: 4.1.1.31)	$K_{PEPCxylase,pep}$	4,07	Constante de saturación media (mM)
	$K_{PEPCxylase,fdp}$	0,7	Constante de activación (mM)
	$n_{PEPCxylase,fdp}$	4	Número de sitios activos
	$r_{PEPCxylase}^{max}$	0,095	Velocidad máxima de reacción (mM s ⁻¹)
Fosfoglucomutasa (PGM) (EC: 5.4.2.2)	$K_{PGM,g6p}$	1,038	Constante de saturación media (mM)
	$K_{PGM,g1p}$	0,0136	Constante de inhibición (mM)
	$K_{PGM,eq}$	0,196	Constante de equilibrio
	r_{PGM}^{max}	0,5618	Velocidad máxima de reacción (mM s ⁻¹)
Glucosa-1-fosfato adeniltransferasa (G1PAT) (EC: 2.7.7.27)	$K_{G1PAT,g1p}$	3,2	Constante de saturación media (mM)
	$K_{G1PAT,atp}$	4,42	Constante de saturación media (mM)
	$K_{G1PAT,fdp}$	0,119	Constante de activación (mM)
	$n_{G1PAT,fdp}$	2	Número de sitios activos
	r_{G1PAT}^{max}	0,005	Velocidad máxima de reacción (mM s ⁻¹)
Ribosa fosfato pirofosfoquinasa (RPPK) (EC: 2.7.6.1)	$K_{RPPK,rib5p}$	0,1	Constante de saturación media (mM)
	r_{RPPK}^{max}	0,0096	Velocidad máxima de reacción (mM s ⁻¹)
Glicerol-3-fosfato deshidrogenasa (G3PDH) (EC: 1.1.1.8)	$K_{G3PDH,dhap}$	1,0	Constante de saturación media (mM)
	r_{G3PDH}^{max}	0,0107	Velocidad máxima de reacción (mM s ⁻¹)
Síntesis de serina (SerSíntesis)	$K_{SerSíntesis,3pg}$	1,0	Constante de saturación media (mM)
	$r_{SerSíntesis}^{max}$	0,0148	Velocidad máxima de reacción (mM s ⁻¹)
DAHPS Sintasa (DAHPS) (EC: 2.5.1.54)	$K_{DAHPS,e4p}$	0,035	Constante de saturación media (mM)
	$K_{DAHPS,pep}$	0,0053	Constante de saturación media (mM)
	$n_{DAHPS,e4p}$	2,6	Número de sitios activos
	$n_{DAHPS,pep}$	2,2	Número de sitios activos
	r_{DAHPS}^{max}	0,0722	Velocidad máxima de reacción (mM s ⁻¹)
Glucosa-6-fosfato deshidrogenasa (G6PDH) (EC: 1.1.1.49)	$K_{G6PDH,g6p}$	0,4915	Constante de saturación media (mM)
	$K_{G6PDH,nadp}$	0,02	Constante de saturación media (mM)
	$K_{G6PDH,nadph,inh}$	0,18	Constante de inhibición (mM)
	h_{NADPH}	1,4	Constante de no linealidad (NADPH)
	k_{dtn}	0,3	Constante de eficiencia de NADH
	k_{dt}	0,1	Constante de eficiencia de NADH
	k_{NADH}	0,3	Constante de eficiencia de NADH
	h_{hdt}	2	Constante de no linealidad (NADH)
	h_{NADH}	4	Constante de no linealidad (NADH)
	k_{tgn}	1	Constante de eficiencia de NADP
k_{tg}	0,4	Constante de eficiencia de NADP	

	h_{ign}	2	Constante de nolienalidad (NADP)
	$r_{\text{G6PDH}}^{\text{max}}$	1,2028	Velocidad máxima de reacción (mM s^{-1})
6-fosfogluconato deshidrogenasa (PGDH) (EC: 1.1.1.44)	$K_{\text{PGDH},6\text{pg}}$	25,0	Constante de saturación media (mM)
	$K_{\text{PGDH},\text{nadp}}$	0,0506	Constante de saturación media (mM)
	$K_{\text{PGDH},\text{nadph},\text{inh}}$	0,0138	Constante de inhibición (mM)
	$K_{\text{PGDH},\text{atp},\text{inh}}$	3	Constante de inhibición (mM)
	$r_{\text{PGDH}}^{\text{max}}$	2,84	Velocidad máxima de reacción (mM s^{-1})
Ribulosa fosfato epimerasa (Ru5p) (EC: 5.1.3.1)	$K_{\text{Ru5p},\text{eq}}$	1,4	Constante de equilibrio (mM)
	$r_{\text{Ru5p}}^{\text{max}}$	5,1816	Velocidad máxima de reacción (mM s^{-1})
Ribosa fosfato isomerasa (R5PI) (EC: 5.3.1.6)	$K_{\text{R5PI},\text{eq}}$	4	Constante de equilibrio (mM)
	$r_{\text{R5PI}}^{\text{max}}$	3,733	Velocidad máxima de reacción (mM s^{-1})
Transcetolasa a (TKa) (EC: 2.2.1.1)	$K_{\text{TKa},\text{eq}}$	1,2	Constante de equilibrio (mM)
	$r_{\text{TKa}}^{\text{max}}$	9,372	Velocidad máxima de reacción (mM s^{-1})
Transcetolasa b (TKb) (EC: 2.2.1.1)	$K_{\text{TKb},\text{eq}}$	10	Constante de equilibrio (mM)
	$r_{\text{TKb}}^{\text{max}}$	45,616	Velocidad máxima de reacción (mM s^{-1})
Transaldolasa (TA) (EC: 2.2.1.2)	$K_{\text{TA},\text{eq}}$	1,05	Constante de equilibrio (mM)
	$r_{\text{TA}}^{\text{max}}$	8,461	Velocidad máxima de reacción (mM s^{-1})
Síntesis1	$K_{\text{Síntesis1},\text{pep}}$	1	Constante de saturación media (mM)
	$r_{\text{Síntesis1}}^{\text{max}}$	0,015	Velocidad máxima de reacción (mM s^{-1})
Síntesis2	$K_{\text{Síntesis2},\text{pyr}}$	1	Constante de saturación media (mM)
	$r_{\text{Síntesis2}}^{\text{max}}$	0,057	Velocidad máxima de reacción (mM s^{-1})
Síntesis de mureina	$r_{\text{MurSíntesis}}^{\text{max}}$	0,0003	Velocidad máxima de reacción (mM s^{-1})
Síntesis de triptófano	$r_{\text{TrpSíntesis}}^{\text{max}}$	2	Velocidad máxima de reacción (mM s^{-1})
Síntesis de metionina	$r_{\text{MetSíntesis}}^{\text{max}}$	0,00003	Velocidad máxima de reacción (mM s^{-1})
Lactato deshidrogenasa (LDH) (EC: 1.1.1.27)	$K_{\text{m},\text{pyr}}$	7,2	Constante de saturación media (mM)
	$K_{\text{m},\text{nadh}}$	0,5	Constante de saturación media (mM)
	K_{eq}	1000	Constante de equilibrio
	$K_{\text{m},\text{lactato}}$	100	Constante de saturación media (mM)
	$r_{\text{LDH}}^{\text{max}}$	35,3	Velocidad máxima de reacción (mM s^{-1})
Fosfoacetilasa (PTA) (EC: 2.3.1.8)	$K_{\text{m},\text{AcCoA}}$	0,2	Constante de inhibición (mM)
	K_{eq}	0,0281	Constante de equilibrio
	$K_{\text{i},\text{P}}$	2,6	Constante de inhibición (mM)
	$K_{\text{i},\text{acp}}$	0,2	Constante de inhibición (mM)
	$K_{\text{i},\text{CoA}}$	0,029	Constante de inhibición (mM)
	$K_{\text{m},\text{P}}$	2,6	Constante de saturación media (mM)

	$K_{m,acp}$	0,7	Constante de saturación media (mM)
	r_{PTA}^{max}	0,5	Velocidad máxima de reacción (mM s ⁻¹)
Acetato quinasa (ACK) (EC: 2.7.2.1)	$K_{m,adp}$	0,5	Constante de saturación media (mM)
	$K_{m,acp}$	0,16	Constante de saturación media (mM)
	$K_{m,ac}$	7	Constante de saturación media (mM)
	$K_{m,atp}$	0,07	Constante de saturación media (mM)
	K_{eq}	100	Constante de equilibrio
	r_{ACK}^{max}	20	Velocidad máxima de reacción (mM s ⁻¹)
Acetaldehído deshidrogenasa (ACALDH) (EC: 1.2.1.10)	$K_{m,AcCoA}$	0,007	Constante de saturación media (mM)
	$K_{m,nadh}$	0,05	Constante de saturación media (mM)
	$K_{m,nad}$	0,08	Constante de saturación media (mM)
	$K_{m,CoA}$	0,008	Constante de saturación media (mM)
	$K_{m,acal}$	10	Constante de saturación media (mM)
	K_{eq}	1	Constante de equilibrio
	r_{ACALDH}^{max}	0,9	Velocidad máxima de reacción (mM s ⁻¹)
Alcohol deshidrogenasa (ADH) (EC: 1.1.1.1)	$K_{m,acal}$	0,03	Constante de saturación media (mM)
	$K_{m,nadh}$	0,05	Constante de saturación media (mM)
	$K_{m,nad}$	0,08	Constante de saturación media (mM)
	$K_{m,etanol}$	1	Constante de saturación media (mM)
	K_{eq}	1000	Constante de equilibrio
	r_{ADH}^{max}	0,5	Velocidad máxima de reacción (mM s ⁻¹)
Piruvato formiato liasa (PFL)	$K_{m,pyr}$	2,05	Constante de saturación media (mM)
	$K_{m,CoA}$	0,0068	Constante de saturación media (mM)
	$K_{m,formiato}$	24,5	Constante de saturación media (mM)
	$K_{m,AcCoA}$	0,051	Constante de saturación media (mM)
	$r_{PFL,for}^{max}$	8,2E-4	Velocidad máxima de reacción (mM s ⁻¹)
	$r_{PFL,rev}^{max}$	7,7E-7	Velocidad máxima de reacción (mM s ⁻¹)
SucSíntesis	$K_{m,pep}$	20	Constante de saturación media (mM)
	$r_{SucSíntesis}^{max}$	1,5	Velocidad máxima de reacción (mM s ⁻¹)

De esta manera, el modelo resultante cuenta con 25 variables de estado, correspondientes a las concentraciones de metabolitos intra y extracelulares; 44 variables algebraicas, que representan las velocidades de reacción y las concentraciones de los co-metabolitos, y 164 parámetros cinéticos.

2.3. Conclusiones

En este capítulo se presenta un modelo cinético mecanístico que describe la glucólisis, la ruta de las pentosas fosfato, el sistema fosfotransferasa, reacciones anapleróticas y de biosíntesis y reacciones fermentativas del metabolismo de *Escherichia coli* K12. El modelo permite describir la respuesta dinámica de las concentraciones de los metabolitos y los flujos de las enzimas ante perturbaciones aplicadas al sistema. Los perfiles de concentración de los co-metabolitos se representan por correlaciones basadas en datos experimentales, debido a que el modelo no contempla todas las reacciones en la que los co-metabolitos participan. Las cinéticas de las reacciones enzimáticas se han obtenido, en general, a partir de estudios realizados *in vitro*. El desarrollo de modelos cinéticos de redes metabólicas, como el descrito en el presente capítulo, permite la determinación de las principales cinéticas *in vivo*, mediante la aplicación de una perturbación tipo pulso en un cultivo continuo y el ajuste de las cinéticas frente a la respuesta dinámica del sistema. Previo a la etapa de estimación de parámetros y debido al alto número de parámetros de modelo, se lleva a cabo un análisis de sensibilidad global en el siguiente capítulo para determinar los parámetros más significativos.

CAPÍTULO 3

ANÁLISIS DE SENSIBILIDAD GLOBAL

3.1. Introducción

Durante la última década, se ha producido un notable aumento en el interés por el desarrollo de nuevos productos y procesos utilizando recursos renovables como materia prima, a través del diseño y optimización de nuevas cepas microbianas. La Ingeniería Metabólica juega un rol importante en este punto, ya que esta disciplina se encarga del mejoramiento de las cepas, evaluando la posibilidad de introducir nuevos caminos metabólicos, eliminar y/o modificar los ya existentes en el metabolismo celular, utilizando herramientas de la Ingeniería Genética y de esta forma alcanzar un objetivo determinado, como la maximización en la producción de un metabolito de interés, la minimización en el gasto energético, entre otros (Bailey, 1991; Stephanopoulos & Stafford, 2002).

Actualmente es posible obtener datos experimentales que permiten lograr una mejor caracterización del estado fisiológico de los microorganismos a distintos niveles, tales como genómico, transcriptómico, proteómico y metabolómico. En la literatura hay disponibilidad de una cantidad creciente de mediciones de

concentraciones de metabolitos intra y extracelulares así como también de niveles y actividad enzimática, que permiten evaluar el funcionamiento dinámico de la células y que pueden ser utilizados para el desarrollo de modelos cinéticos (Majewski & Domach, 1990; Varma & Palsson, 1994; Lee, Palakornkule, Domach, & Grossmann, 2000; Burgard & Maranas, 2001; Ghosh, Zhu, Grossmann, Ataii, & Domach, 2005; Namjoshi & Ramkrishna, 2005).

Los modelos cinéticos, una vez calibrados, constituyen una herramienta de gran utilidad para la Ingeniería Metabólica, ya que permiten, en principio, guiar las manipulaciones genéticas y predecir las consecuencias de las mismas a nivel celular. Estos modelos describen el metabolismo celular, brindando información dinámica de la evolución de los metabolitos involucrados en la red metabólica, a través de ecuaciones diferenciales y algebraicas que surgen de los balances de masa planteados para los nodos de la red (Capítulo 2). Por otra parte, los modelos dinámicos, al incorporar las expresiones cinéticas, cuentan con un gran número de parámetros, tales como velocidades máximas de reacción, constantes de saturación media, constantes de inhibición, entre otros, que deben ser estimados para una dada condición de crecimiento del cultivo de los microorganismos bajo estudio. Sin embargo, en la mayoría de los casos, no todos los parámetros presentes en el modelo pueden ser calibrados y resulta apropiado identificar y establecer un ranking de los parámetros más importantes del modelo, llevando a cabo un análisis de sensibilidad.

Existen dos categorías principales de métodos de análisis de sensibilidad: local y global. En el enfoque local, los índices de sensibilidad se calculan como la

derivada primera parcial de las variables de estado respecto de los parámetros del modelo; es decir se estudia el efecto que ejercen pequeñas perturbaciones de los parámetros, variándolos uno a uno, sobre los estados del modelo, asumiendo linealidad de los mismos alrededor de la trayectoria nominal. Los métodos que emplean el enfoque local son altamente dependientes del modelo y pueden arrojar resultados erróneos cuando el modelo es altamente no lineal, como es el caso de los modelos que describen sistemas biológicos.

Por su parte, los métodos de análisis de sensibilidad global se basan en explorar el rango completo de variación de los parámetros mediante el uso de técnicas de muestreo como las simulaciones Monte Carlo (SMC) o Hipercubo Latino (HL) (Sobol', 1990; Saltelli, *et al.*, 1999; Sobol' & Levitan, 1999; Sobol', 2001; Saltelli *et al.*, 2004). Una desventaja de estos métodos es su alto costo computacional, pero el creciente poder de las computadoras ha permitido la aplicación de estas técnicas aún en modelos altamente complejos. Además, estos métodos proveen resultados más realistas, debido a que se pueden identificar las interacciones entre los parámetros. Estos métodos también resultan independientes del modelo debido a que no se asume linealidad o aditividad del modelo (Saltelli *et al.*, 2004).

Mauch *et al.* (1997) proponen un método de sensibilidad local con el objetivo de determinar los coeficientes de control de flujos y de concentración, tanto en estado estacionario como dinámico, para una red metabólica genérica, y lo aplican a un modelo de red metabólica conformada por dos ecuaciones diferenciales, tres expresiones cinéticas y doce parámetros cinéticos. Por otra parte, Noack *et al.* (2008) aplican una técnica de análisis de sensibilidad local a un modelo dinámico

de red metabólica de *Escherichia coli* y desarrollan animaciones y técnicas de visualización para mostrar la variación temporal de las sensibilidades a partir del estado estacionario inicial hasta la inyección de un pulso de glucosa.

En el presente capítulo, se describe la metodología y se presentan los resultados obtenidos al aplicar análisis de sensibilidad global a un modelo dinámico de una red metabólica compleja correspondiente a *Escherichia coli* K12 W3110 (Di Maggio *et al.*, 2009) considerando la vía Embden-Meyerhof-Parnas, el sistema fosfotransferasa y la vía de las pentosas (Di Maggio *et al.*, 2010). El modelo comprende 18 ecuaciones diferenciales, 37 ecuaciones algebraicas, con 123 parámetros, de los cuales 20 se han considerado como parámetros inciertos.

Se ha implementado el método propuesto por Sobol', basado en el empleo de varianzas condicionales e incondiciones para el cálculo de índices de sensibilidad que permiten identificar los parámetros más influyentes en el modelo de la red metabólica de *Escherichia coli*. La determinación de los índices de sensibilidad de Sobol'-Saltelli (Saltelli *et al.*, 1999; Sobol', 2001) permite establecer un orden de los parámetros según su efecto sobre las principales variables de estado.

3.2. Análisis de Sensibilidad Global. Metodología

El análisis de sensibilidad global se basa en explorar el rango de variación completo de los parámetros del modelo, muestreando de la función de distribución de probabilidad asociada a cada parámetro del modelo y realizando repetidas simulaciones. Estos métodos tienen un mayor costo computacional que las

técnicas locales, pero proveen resultados más realistas debido a que se pueden identificar las interacciones entre los parámetros (Saltelli *et al.*, 2004; Pastres *et al.* 2004). Además, los métodos globales no requieren asumir linealidad y/o aditividad del modelo. Las técnicas de análisis de sensibilidad global incluyen el método de Morris (Morris, 1991), el Test de Amplitud de Sensibilidad de Fourier (FAST) (Cukier *et al.*, 1973; Cukier, *et al.*, 1978; Koda *et al.*, 1979 a,b; Koda & Seinfeld, 1982; McRae *et al.*, 1982; Liepmann & Stephanopoulos, 1985) y el método de Sobol' (Sobol', 1991). Este último se basa en realizar la misma descomposición de la varianza que FAST, pero a través de la aplicación de métodos Monte Carlo en lugar de análisis espectral (Sobol', 1990; Saltelli & Sobol', 1995; Homma & Saltelli, 1996; Liu & Chen, 1998; Saltelli *et al.*, 1999; Sobol' & Levitan, 1999; Sobol', 2001; Saltelli & Tarantola, 2002; Saltelli *et al.*, 2004).

Antes de introducir la metodología propuesta por Sobol', se presentan conceptos básicos de estadística. Sea una variable aleatoria continua C (que en una red metabólica podría ser una concentración de metabolito intracelular), que es función de k variables aleatorias x_i (que podrían ser parámetros del modelo de la red metabólica):

$$C = f(x_1, x_2, \dots, x_k) \quad (3.1)$$

Si todas las x_i varían según su función de densidad de probabilidad, la correspondiente incertidumbre en la variable C se puede cuantificar a través de la varianza incondicional $V(C)$. Para ordenar los x_i según la disminución de la varianza que se produce cuando se conoce el verdadero valor de x_i , se emplea el concepto de varianza condicional.

La fórmula general para la descomposición de la varianza o ley de varianza total es:

$$V(C) = V(E(C|x_i)) + E(V(C|x_i)) \quad (3.2)$$

o, alternativamente

$$V(C) = V(E(C|x_{-i})) + E(V(C|x_{-i})) \quad (3.3)$$

donde V y E corresponden a los operadores varianza y valor esperado, respectivamente. En la ecuación (3.2), $V(E(C|x_i))=V_i$ es la varianza (sobre todas las realizaciones posibles del parámetro x_i) del valor esperado condicional de la variable C , bajo la variación de todos los parámetros, excepto x_i . Esto representa la reducción esperada en la varianza de la variable C que podría obtenerse si el parámetro x_i pudiese conocerse y de este modo fijarse y se denomina efecto de primer orden asociado al parámetro x_i .

El segundo término, $E(V(C|x_i))$, es el valor esperado (sobre todas las realizaciones del parámetro x_i) de la varianza condicional de la variable C bajo la variación de todos los parámetros, excepto x_i . Esto representa el valor medio de la varianza de la variable C que se obtendría si x_i pudiese conocerse o fijarse.

Como en la ecuación (3.2), la suma es constante (e igual a la varianza incondicional), buscar el parámetro x_i que, una vez fijado en su valor verdadero produce menor varianza media en C ($E(V(C|x_i))$), es equivalente al que produce mayor reducción en la varianza media ($V(E(C|x_i))$).

Por su parte, la ecuación (3.3) se puede describir en forma análoga, reemplazando x_i por “todos los parámetros excepto x_i ” (x_{-i}). De esta manera, el término $E(V(C|x_{-i}))=V_i^{TOT}$ calcula el valor esperado de la varianza de la variable C que se obtendría si todos los parámetros excepto x_i pudiesen conocerse o fijarse.

Si las ecuaciones (3.2) y (3.3) se dividen por la varianza incondicional, se obtienen las siguientes expresiones:

$$1 = \frac{V(E(C|x_i))}{V(C)} + \frac{E(V(C|x_i))}{V(C)} \quad (3.4)$$

$$1 = \frac{V(E(C|x_{-i}))}{V(C)} + \frac{E(V(C|x_{-i}))}{V(C)} \quad (3.5)$$

El índice de sensibilidad de primer orden, S_i , y el índice de sensibilidad total S_i^{TOT} se definen según:

$$S_i = \frac{V(E(C|x_i))}{V(C)} = \frac{V_i}{V(C)} \quad (3.6)$$

$$S_i^{TOT} = \frac{E(V(C|x_{-i}))}{V(C)} = \frac{V_i^{TOT}}{V(C)} \quad (3.7)$$

El cálculo de las varianzas incondicionales y condicionales de variables aleatorias continuas implica la resolución de integrales múltiples. Sobol' (1990) propone una metodología simplificada para el cálculo de las varianzas, que se describe a continuación.

La idea básica es que dada la función $C=f(x,t)$, definida en $\Omega^k = \{x / 0 \leq x_i \leq 1, i=1,2,\dots,k\}$, donde C es una variable de estado diferencial o algebraica, x es un vector de k parámetros de entrada del modelo y t es el tiempo, esta función puede descomponerse en términos de dimensionalidad creciente, según la ecuación (3.8) (Sobol', 1990). Para facilitar la descripción del método, en la siguiente discusión omitimos el índice t , asumiendo que el valor esperado y la varianza son calculados en cada instante de tiempo para la variable de estado:

$$f = f_0 + \sum_i f_i(x_i) + \sum_i \sum_{j>i} f_{ij}(x_i, x_j) + \dots + f_{1,2,\dots,k}(x_1, x_2, \dots, x_k) \tag{3.8}$$

La ecuación (3.8) es una representación ANOVA de la función $C=f(x)$ si entre cada par de parámetros x_i y x_j se cumple (ortogonalidad):

$$\int_0^1 \int_0^1 f(x_i) f(x_j) dx_i dx_j = 0 \tag{3.9}$$

Por su parte, el valor esperado incondicional de C (f_0), puede calcularse como:

$$E(C) = \int_0^1 \dots \int_0^1 f(x) p(x_{i1}) \dots p(x_{in}) dx_{i1} dx_{in} = f_0 \tag{3.10}$$

donde $p(x_{in})$ es la función de distribución de probabilidad.

Los respectivos valores esperados condicionales, se calculan según:

$$E(C|x_i) = \int_0^1 \dots \int_0^1 f(x) p(x_{i1}) \dots p(x_{in}) \prod_{k \neq i} dx_k = f_0 + f_i(x_i) \tag{3.11}$$

$$E(C|x_i, x_j) = \int_0^1 \dots \int_0^1 f(x) p(x_{i1}) \dots p(x_{in}) \prod_{k \neq i,j} dx_k = f_0 + f_i(x_i) + f_j(x_j) + f_{ij}(x_i, x_j) \tag{3.12}$$

Los sucesivos valores esperados condicionales se calculan de manera análoga.

La ecuación (3.10) corresponde a la definición de valor esperado o esperanza

matemática de una variable aleatoria continua (C) que es función de variables aleatorias (x_1, x_2, \dots, x_n) , la ecuación (3.11) es la definición de valor esperado condicional de la variable aleatoria continua C dado x_i , $E(C|x_i)$. La ecuación (3.12) corresponde a la definición de valor esperado condicional de la variable C dados x_i y x_j , $E(C|x_i, x_j)$.

Asumiendo que el cuadrado de $f(x)$ es integrable, entonces todos los términos de la ecuación (3.8) también lo son. Luego elevando al cuadrado la ecuación 3.8 e integrando sobre el dominio Ω , se obtiene:

$$\int_0^1 \dots \int_0^1 f^2(x) dx_{i_1} \dots dx_{i_n} - f_0^2 = \sum_{s=1}^n \sum_{i_1 < \dots < i_s} \int_0^1 \dots \int_0^1 f_{i_1, i_2, \dots, i_s}^2 dx_{i_1} \dots dx_{i_s} \tag{3.13}$$

donde,

$$V = \int_0^1 \int_0^1 \dots \int_0^1 f^2(x) dx_{i_1} \dots dx_{i_n} - f_0^2 \tag{3.14}$$

$$V_{i_1, i_2, \dots, i_s} = \int_0^1 \int_0^1 \dots \int_0^1 f_{i_1, i_2, \dots, i_s}^2 dx_{i_1} \dots dx_{i_s} \tag{3.15}$$

La ecuación (3.14) corresponde a la definición de la varianza incondicional de una variable aleatoria continua C, mientras que la ecuación (3.15) es la definición de la varianza condicional de la variable aleatoria continua C.

De este modo se obtiene la siguiente descomposición ANOVA de la varianza incondicional (Sobol', 1990):

$$V = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \dots + V_{12\dots k} \tag{3.16}$$

Sobol' (2001) propuso una metodología para estimar los índices de sensibilidad, según la deducción detallada en esta sección, basada en simulaciones Monte Carlo, y que requiere una cantidad mínima de evaluaciones de las funciones.

Los principales pasos para el caso de modelos dinámicos son los siguientes:

1. Generación de dos conjuntos aleatorios diferentes de parámetros del modelo: $\xi = (\eta, \zeta)$ y $\xi' = (\eta', \zeta')$. Cada matriz tiene dimensión $N \times k$, donde N es el número de muestras para las simulaciones Monte Carlo y k es el número de parámetros; η es un vector de dimensiones $N \times 1$, el cual contiene N valores aleatorios del parámetro x_i y ζ es una sub-matriz de dimensión $N \times (k-1)$ que contiene valores aleatorios para los $k-1$ parámetros de entrada restantes.
2. Creación de dos nuevas matrices por combinación de ξ y ξ' , las cuales son requeridas para el cálculo de los perfiles de las varianzas condicionales de las variables de estado, C ($C_0(t)$ representa $E(C(t) | x_i)$), del modelo dinámico

$$\frac{1}{N} \sum_{i=1}^N C(t, \xi_i) \xrightarrow{P} C_0(t) \quad (3.17)$$

$$\frac{1}{N} \sum_{i=1}^N C^2(t, \xi_i) \xrightarrow{P} V(t) + C_0^2(t) \quad (3.18)$$

$$\frac{1}{N} \sum_{i=1}^N C(t, \xi_i) C(t, \eta_i, \zeta_i') \xrightarrow{P} V(t)_i + C(t, \xi_i) C(t, \xi_i') \quad (3.19)$$

$$\frac{1}{N} \sum_{i=1}^N C(t, \xi_i) C(t, \eta_i', \zeta_i) \xrightarrow{P} V(t)_{-i} + C_0^2(t) \quad (3.20)$$

3. Los perfiles de los índices de sensibilidad son calculados por las correspondientes definiciones dadas anteriormente, ecuaciones (3.5) y (3.6) a lo largo del horizonte de tiempo.

Tanto S_i como S_i^{TOT} constituyen buenas medidas de los efectos de la variación de los parámetros sobre las variables de estado del modelo. S_i representa la reducción sobre la varianza incondicional de la variable de estado que puede ser obtenida si x_i se fija a su valor verdadero. Por otro lado, S_i^{TOT} toma en cuenta las interacciones entre parámetros, por lo tanto proporciona información de la parte no aditiva del modelo.

Usualmente, $\sum_{i=1}^k S_i < 1$ y $S_i < S_i^{TOT}$. Sin embargo, para modelos puramente aditivos y parámetros ortogonales $\sum_{i=1}^k S_i = 1$, lo cual puede observarse si la ecuación (3.16) es dividida por $V(C)$ y los términos de interacción son cancelados

$$1 = \sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \dots + S_{12\dots k} \quad (3.21)$$

Un índice adicional, el S_i^{int} , tiene en cuenta los efectos de interacciones entre los parámetros del modelo y puede calcularse como:

$$S_i^{int} = S_i^{TOT} - S_i \quad (3.22)$$

En resumen, cuando S_i es alto, x_i es un parámetro influyente; cuando S_i^{int} es importante, es decir S_i y S_i^{TOT} son muy diferentes, las interacciones entre x_i y

otros parámetros cobran importancia; cuando S_i y S_i^{TOT} alcanzan valores similares (S_i^{int} cercano a cero), no hay interacciones entre x_i y otros parámetros o en su defecto son pequeñas, en el caso de que tanto S_i como S_i^{TOT} toman valores pequeños, esto implica que x_i no es un parámetro influyente (ni solo ni en interacciones con otros parámetros).

3.3. Resultados y discusión

En este trabajo se ha implementado la metodología análisis de sensibilidad global (ASG) descrita en la sección anterior, a un modelo cinético de una red metabólica de gran complejidad de *Escherichia coli*.

El análisis de sensibilidad global en un sistema de ecuaciones diferenciales algebraicas requirió un esfuerzo computacional especial. Como primer paso, se ha asociado una distribución de probabilidad normal a cada parámetro con media en su valor nominal y una desviación estándar del 20% (Tarmy & Kaplan, 1968; Blangy *et al.*, 1968; Sanwal, 1970; Knappe *et al.*, 1974; Ingram *et al.*, 1987; Diaz Ricci, 1996; Rohwer *et al.*, 2000; Chassagnole *et al.*, 2002; Hoefnagel *et al.*, 2002; Kremling *et al.*, 2004; Ceric & Kurtanjek, 2006; Ratushny *et al.*, 2006). En segunda instancia, se llevaron a cabo simulaciones Monte Carlo para las tres matrices aleatorias, según fue descrito en la sección 3.2. El número de escenarios requerido, N , es función de la complejidad del modelo (Baklouti *et al.*, 2006). Dicho número fue estimado realizando simulaciones estocásticas aumentando el número

de escenarios (N) y comparando los valores medios de los perfiles para las principales variables de estado del modelo. Las Figuras 3.1 (a) y (b) muestran los perfiles de la concentración media de glucosa-6-fosfato (C_{g6p}) y piruvato (C_{pyr}) como función de N , que toma valores entre 100 y 2500.

Se puede observar (Figs. 3.1 a y b) que los perfiles de las concentraciones medias se mantienen sin cambios para 2000 y 2500 escenarios. Este resultado sugiere que un número de escenarios igual a 2500 sería adecuado para la implementación del ASG.

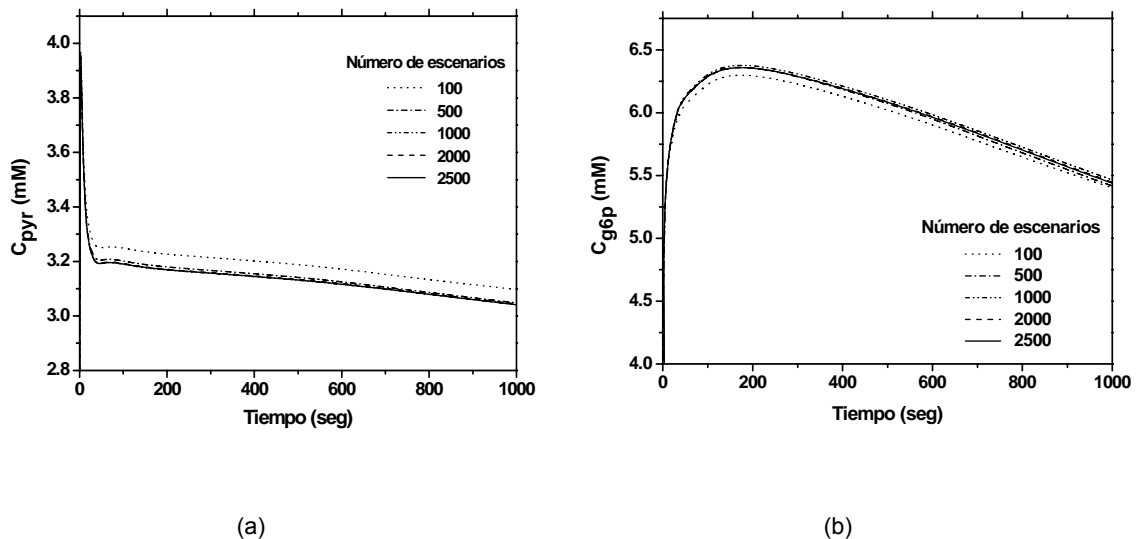


Figura 3.1. Perfiles de concentración media de piruvato (a) y de glucosa-6-fosfato (b) en función del número de escenarios, luego de la aplicación de un pulso de glucosa de 0,15 g/mL

En una primera etapa, se seleccionó un conjunto de 20 parámetros (Tabla 3.1), realizando simulaciones estocásticas variando los parámetros uno por uno sobre su rango de variación, y calculando la varianza incondicional de las principales variables de estado, de manera tal que se seleccionaron los parámetros factibles de estimar bajo consideraciones biológicas y con valores de varianza mayores a

0,5. El análisis de sensibilidad global se llevó a cabo sobre el conjunto de esos 20 parámetros.

Tabla 3.1. Valor nominal de los parámetros inciertos (Chassagnole *et al.*, 2002)

Enzima	Parámetro	Valor nominal
Sistema Fosfotransferasa (PTS)	$K_{PTS,1}$	3082,3 mM
	$K_{PTS,2}$	0,01 mM
	$K_{PTS,3}$	245,3 mM
	$K_{PTS,g6p}$	2,15 mM
	r_{PTS}^{max}	389696,03 mMseg ⁻¹
Fosfofructoquinasa (PFK)	$K_{PFK,f6ps}$	0,325 mM
	$K_{PFK,pep}$	0,00179 mM
	r_{PFK}^{max}	0,406 mM seg ⁻¹
Gliceraldehído-3-fosfato deshidrogenasa (GAPDH)	$K_{GAPDH,gap}$	0,683 mM
	$K_{GAPDH,pgp}$	0,0000104 mM
	r_{GAPDH}^{max}	705,917 mM seg ⁻¹
Piruvato quinasa (PK)	$K_{PK,pep}$	0,31 mM
	$K_{PK,fdp}$	0,19 mM
	r_{PK}^{max}	0,046 mM seg ⁻¹
Piruvato deshidrogenasa (PDH)	$K_{PDH,pyr}$	1159 mM
	r_{PDH}^{max}	4,596 mM seg ⁻¹
Glucosa-6-fosfato deshidrogenasa (G6PDH)	$K_{G6PDH,g6p}$	14,4 mM
	r_{G6PDH}^{max}	1,063 mM seg ⁻¹
6-fosfogluconato deshidrogenasa (PGDH)	$K_{PGDH,6pg}$	37,5 mM
	r_{PGDH}^{max}	12,491 mM seg ⁻¹

Las simulaciones Monte Carlo fueron implementadas en g-PROMS (PSE Enterprise, 2007), el sistema de ecuaciones diferenciales algebraicas se resolvió con DASOLV (Brenan *et al.*, 1996). Los resultados numéricos obtenidos de las simulaciones, es decir los perfiles temporales de las variables de estado diferenciales y algebraicas, se exportaron como archivos de datos a un ambiente de Fortran 90 para el cálculo de las varianzas y los índices de sensibilidad de

primer orden, totales y de interacción, según las ecuaciones (3.17) a (3.20) y (3.6), (3.7) y (3.22). El cálculo de los índices en un entorno de Fortran 90 se realizó debido a limitaciones de memoria en el uso de gPROMS, con un modelo de gran tamaño, alto número de parámetros inciertos (20) y de escenarios (2500). Los cálculos fueron realizados para un horizonte de tiempo de 1000 segundos.

Los resultados del análisis de sensibilidad global aplicado al modelo cinético de la red metabólica de *Escherichia coli* descrito en el Capítulo 2, se muestran en las Figuras 3.2 a 3.8 y en las Tablas 3.2 y 3.3. Las Figuras 3.2, 3.3, 3.5, 3.7 y 3.8 muestran los perfiles temporales acumulados de los índices de sensibilidad de primer orden (S_i) para las concentraciones de piruvato, glucosa extracelular, 6-fosfogluconato, para la velocidad de reacción del sistema fosfotransferasa y para la concentración de fosfoenolpiruvato, respectivamente. En dichas figuras se puede observar que la sumatoria de S_i no alcanza, en la mayor parte de los casos, un valor igual a 1. Sin embargo, alcanza valores superiores a 0,5. Debido a que la sumatoria de los índices de primer orden es una medida de la aditividad del modelo, se podría concluir que no existen fuertes interacciones entre los parámetros para las variables de estado analizadas. Las Figuras 3.4 y 3.6 muestran los perfiles del índice de sensibilidad de interacción (S_i^{int}) para las concentraciones de glucosa extracelular y 6-fosfogluconato, respectivamente. De las gráficas de S_i^{int} también puede observarse que los efectos de interacción entre parámetros son pequeños. Como se mencionó en la sección 3.2 los índices de interacción se calculan según la ecuación 3.22.

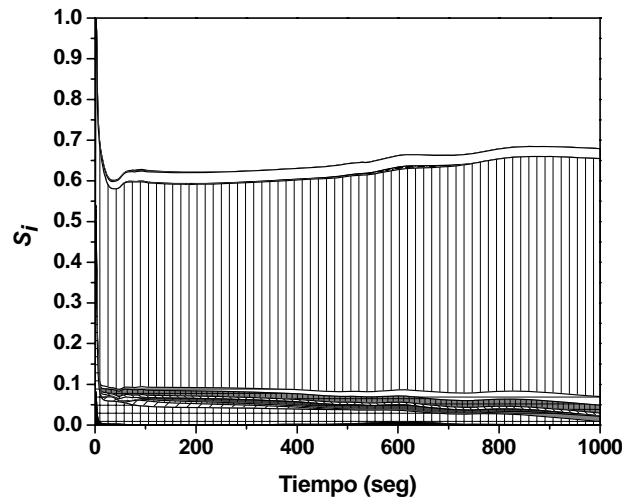
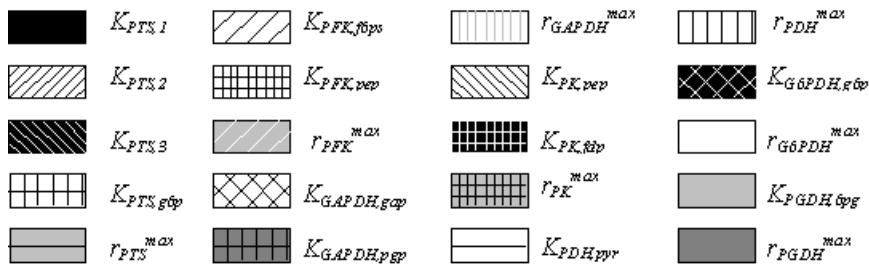


Figura 3.2. Perfil de S_i para la concentración de piruvato, luego de la aplicación de un pulso de glucosa de 0,15 g/mL



Referencias para las Figuras 3.2 a 3.8

La Figura 3.2 muestra el perfil de los índices de primer orden para la concentración de piruvato. El parámetro que mayor influencia tiene en el comportamiento de esta variable durante el horizonte de tiempo completo, es la velocidad máxima de reacción de la enzima piruvato deshidrogenasa (r_{PDH}^{max}). Esta influencia puede ser determinada con una visualización detallada de la red

metabólica ya que el piruvato es sustrato de la reacción catalizada por la enzima PDH, la aplicación del análisis de sensibilidad confirma este hecho.

Las Figuras 3.3 y 3.4 presentan la evolución temporal de los índices de sensibilidad de primer orden (S_i) y de interacciones (S_i^{int}) para la concentración de glucosa extracelular. Ambos gráficos muestran que los mayores efectos se producen al comienzo del horizonte de tiempo, inmediatamente después de la aplicación del pulso de glucosa, cuando la concentración de dicho sustrato aumenta respecto del valor de estado estacionario. Los tres parámetros identificados como los más influyentes para esta variable de estado son r_{PFK}^{max} , (velocidad máxima de reacción de la enzima fosfofructoquinasa, PFK), $K_{PFK,fbps}$ (constante de saturación media de la enzima PFK para la fructosa-6-fosfato) y r_{PTS}^{max} (velocidad máxima de reacción del sistema fosfotransferasa). Es de destacar, que r_{PFK}^{max} tiene un importante efecto sobre el comportamiento de la concentración de glucosa extracelular durante todo el horizonte de tiempo. Se puede observar que como la glucosa extracelular no es sustrato, producto, ni inhibidor de la reacción catalizada por dicha enzima dicha enzima, este metabolito no tiene influencia directa sobre la actividad de la fosfofructoquinasa (Figura 2.1). Sin embargo, el fosfoenolpiruvato es un metabolito común entre las enzimas PFK y PTS siendo inhibidor de la actividad de la PFK y sustrato de la reacción catalizada por el sistema PTS segunda junto con la glucosa extracelular. La relación entre el fosfoenolpiruvato y la glucosa extracelular podría explicar la influencia de la velocidad máxima de la enzima PFK sobre la glucosa extracelular, como fue determinado por el análisis de sensibilidad global. Basado en

consideraciones similares se puede inferir el efecto de $K_{PFK, f6ps}$ (constante de saturación media de la enzima PFK para la fructosa-6-fosfato) sobre la concentración de glucosa extracelular. El tercer parámetro identificado como influyente, r_{PTS}^{max} , cuyos mayores valores en el S_i se detectan al comienzo del horizonte de tiempo, durante los primeros 50 segundos. De acuerdo con los principios de la cinética enzimática cuando la concentración de sustrato es elevada la velocidad de reacción (r_j) se aproxima a la velocidad máxima (r_j^{max}). En la ecuación de Michaelis-Menten (Ec. 2.27), cuando la concentración de sustrato (C_s) es elevada, superando el valor de K_m (concentración de sustrato para la cual se alcanza el valor medio de la velocidad máxima de reacción), en el denominador de la ecuación (2.27) se puede depreciar el valor de K_m obteniéndose, de esta manera, que la velocidad de reacción (r_j) se hace igual a su velocidad máxima (r_j^{max}). En el modelo desarrollado en esta tesis, como se mencionó anteriormente, a $t=0$ se aplica un pulso de glucosa lo que produce un aumento en la concentración de dicho azúcar, que es sustrato de la reacción en la que el sistema fosfotransferasa actúa como catalizador. Por esta razón puede explicarse el efecto que tiene el parámetro r_{PTS}^{max} sobre el comportamiento dinámico de la concentración de la glucosa extracelular, durante los primeros 50 segundos.

En la Figura 3.4 se muestra la evolución de los perfiles de los índices de interacción para la concentración de glucosa extracelular, donde se puede verificar que la interacción del parámetro r_{PFK}^{max} con los restantes del modelo no es despreciable. Se concluye que dicho parámetro contribuye a la varianza incondicional de la glucosa con su efecto total, esto es S_i^{TOT} es mayor de S_i .

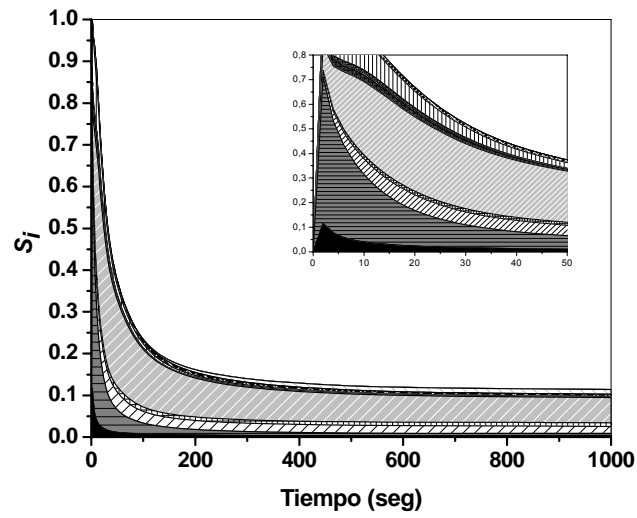


Figura 3.3. Perfil de S_i para la concentración de glucosa extracelular, luego de la aplicación de un pulso de glucosa de 0,15 g/mL

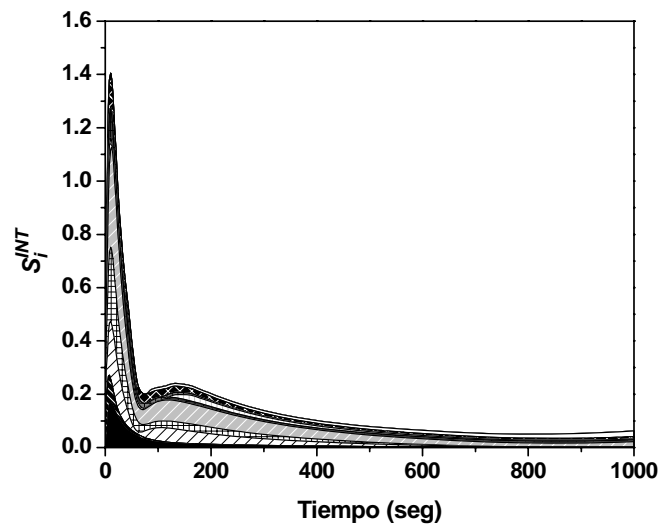


Figura 3.4. Perfil de S_i^{int} para la concentración de glucosa extracelular, luego de la aplicación de un pulso de glucosa de 0,15 g/mL

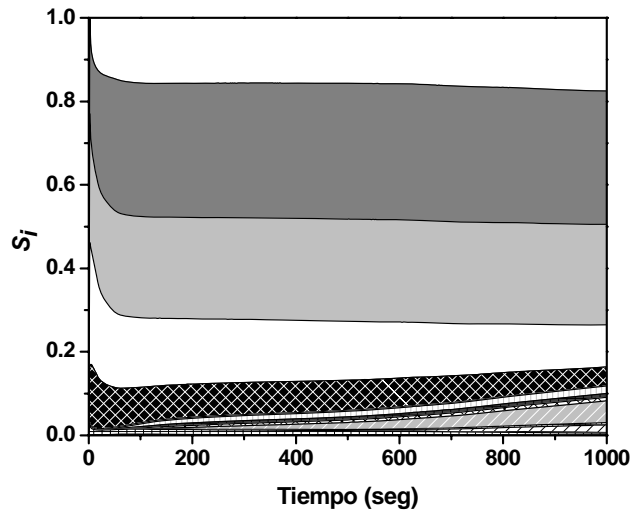


Figura 3.5. Perfil de S_i para la concentración de 6-fosfogluconato, luego de la aplicación de un pulso de glucosa de 0,15 g/mL

El 6-fosfogluconato (6pg) es un metabolito intracelular que se obtiene por deshidrogenación de la glucosa, en la primera reacción involucrada en la vía de las pentosas; es sustrato de la reacción catalizada por la enzima 6-fosfogluconato deshidrogenasa (PGDH) y producto de la reacción en la que participa la enzima glucosa-6-fosfato deshidrogenasa (G6PDH) (Fig. 2.1). En la Figura 3.5 se muestran los perfiles de S_i para la concentración de 6pg. Se puede observar que los mayores aportes a la varianza están dados por la constante de saturación media y la velocidad máxima de reacción de la enzima PGDH ($K_{PGDH,6pg}$ y r_{PGDH}^{max} , respectivamente) y por la constante de saturación media y la velocidad máxima de reacción de la enzima G6PDH ($K_{G6PDH,g6p}$ y r_{G6PDH}^{max}). Por su parte, los efectos de interacciones entre parámetros no son despreciables tanto para $K_{PGDH,6pg}$ como para r_{PGDH}^{max} , como puede observarse en la Figura 3.6.

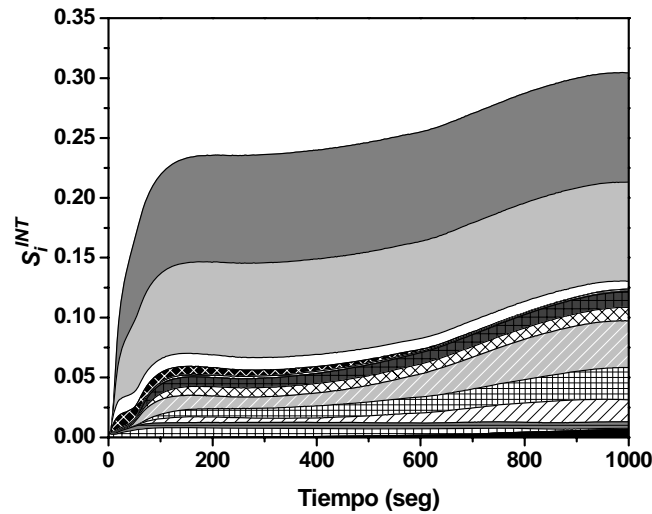


Figura 3.6. Perfil de S_i^{int} para la concentración de 6-fosfogluconato, luego de la aplicación de un pulso de glucosa de 0,15 g/mL

En este capítulo también se han analizado los índices de sensibilidad para las variables algebraicas del modelo de la red metabólica de *E. coli*. La Figura 3.7 muestra los perfiles de los índices de primer orden (S_i), para la velocidad de reacción del sistema fosfotransferasa (r_{PTS}). En este caso, los parámetros que realizan los mayores aportes de primer orden son $K_{PTS,g6p}$ (constante de inhibición por producto del sistema PTS), $K_{PFK,f6ps}$ (constante de saturación media de la enzima PFK), r_{PFK}^{max} (velocidad máxima de reacción de la enzima PFK) y r_{G6PDH}^{max} (velocidad máxima de reacción de la enzima G6PDH). Estos resultados sugieren que modificando la afinidad de las enzimas hacia los sustratos, en este caso modificando la constante de saturación media de la PFK hacia fructosa-6-fosfato podrían originarse cambios en la actividad del sistema fosfotransferasa. Resultados similares podrían obtenerse si se pudiera modificar las

concentraciones de las enzimas PFK y G6PDH, lo que produciría una variación en los valores de r_{PFK}^{max} y r_{G6PDH}^{max} .

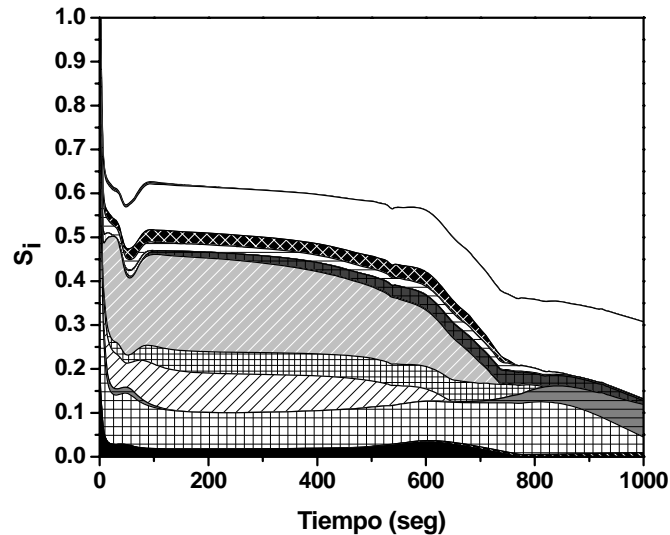


Figura 3.7. Perfil de S_i para la velocidad de reacción del sistema fosfotransferasa, luego de la aplicación de un pulso de glucosa de 0,15 g/mL

Los mayores efectos ocurrirían a través de la modificación de la concentración de la enzima glucosa-6-fosfato deshidrogenasa, ya que presenta los mayores valores de S_i . Un hecho llamativo que surge de la aplicación del análisis de sensibilidad global es que la velocidad de reacción del sistema fosfotransferasa, no se ve afectada por la velocidad máxima de reacción correspondiente a esta enzima (Fig. 3.7).

En las Figuras 3.8 a y b se observa la evolución de los índices de sensibilidad de primer orden para la concentración de fosfoenolpiruvato. Los parámetros cinéticos de la enzima gliceraldehído-3-fosfato deshidrogenasa, r_{GAPDH}^{max} (velocidad máxima de reacción), $K_{GAPDH,gap}$ (constante de saturación media) y $K_{GAPDH,pgp}$ (constante de inhibición), dominan el comportamiento dinámico de este metabolito

a lo largo de todo el horizonte de tiempo (Fig. 3.8 a). Durante los primeros 15 segundos de análisis, r_{PTS}^{max} (velocidad máxima de reacción del sistema fosfotransferasa) es el parámetro más influyente sobre la concentración de fosfoenolpiruvato, como se puede observar en la Figura 3.8 b.

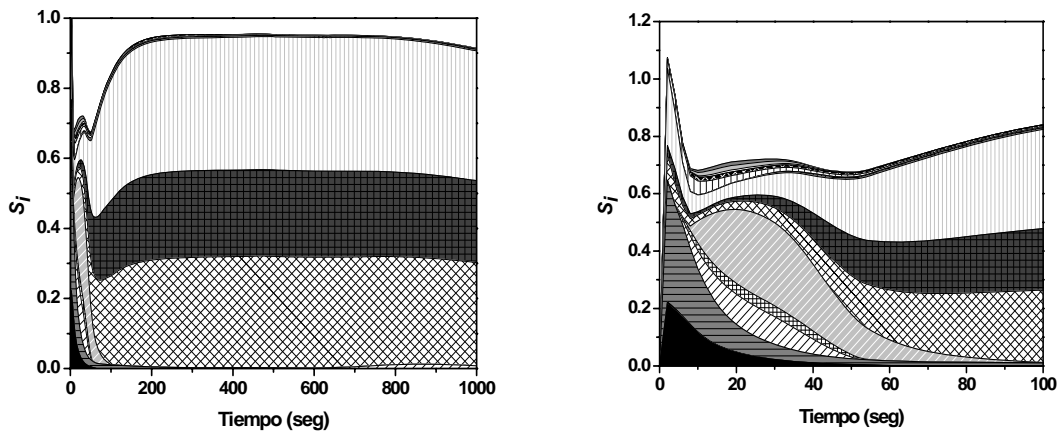


Figura 3.8. Perfil de S_i para la concentración de fosfoenolpiruvato, luego de la aplicación de un pulso de glucosa de 0,15 g/mL
(a) Horizonte de tiempo completo y (b) Primeros 100 segundos de análisis

Este metabolito, junto con la glucosa extracelular, es sustrato de la reacción en la que interviene el sistema fosfotransferasa. Cuando la concentración de glucosa se incrementa debido a la aplicación del pulso de este azúcar en el medio de cultivo, la velocidad de reacción de la enzima PTS se iguala a su velocidad máxima de reacción (r_{PTS}^{max}), provocando este dominio en el comportamiento dinámico del pep, tal como se explicó anteriormente para el caso de la glucosa extracelular. Por otra parte, cuando se analiza el intervalo de tiempo comprendido entre los 15 y los 35 segundos, se observa que la velocidad máxima de reacción de la fosfofructoquinasa (r_{PFK}^{max}) pasa a ser un parámetro influyente para el

fosfoenolpiruvato (Fig. 3.8 b). Este metabolito es un inhibidor alostérico de la actividad de la enzima PFK (Diaz Ricci, 1996) cuya actividad se ve incrementada, hasta cerca de su valor máximo, debido al pulso de glucosa. Este aumento brusco en la actividad podría ser la explicación del efecto del parámetro r_{PFK}^{max} en la evolución dinámica del pep.

En las Tablas 3.2 y 3.3 se muestran valores de S_i y S_i^{int} para las variables concentración de glucosa-6-fosfato y velocidad de reacción de la enzima piruvato quinasa, respectivamente, en tres instantes de tiempo ($t = 50, 300$ y 800 seg).

Para el caso de la concentración de glucosa-6-fosfato (Tabla 3.2) se puede

observar que para los tres instantes de tiempo la $\sum_{i=1}^{20} S_i$ es cercana a 1, lo cual nos

estaría indicando que la varianza incondicional de la variable C_{g6p} podría ser explicada por el efecto de primer orden. Los parámetros que mayor influencia tienen sobre esta variable en los primeros instantes de tiempo son $K_{PTS,g6p}$, r_{PTS}^{max} y r_{GAPDH}^{max} , siendo estos tres parámetros los que explican el 46% de la varianza de la concentración de glucosa-6-fosfato.

En la Tabla 3.2 se puede observar que para los restantes instantes de tiempo, es decir $t=300$ seg y $t=800$ seg, la velocidad máxima de reacción del sistema fosfotransferasa (r_{PTS}^{max}) deja ser un parámetro influyente, volviéndose más importante la velocidad de reacción de la enzima fosfofructoquinasa (r_{PFK}^{max}). Si se observan los valores de S_i^{int} , se puede destacar el efecto de la interacción del parámetro r_{PFK}^{max} con los restantes.

Tabla 3.2. Valores de S_i y S_i^{int} en tres instantes de tiempo para la concentración de glucosa-6-fosfato

	S_i			S_i^{int}		
	t=50	t=300	t=800	t=50	t=300	t=800
$K_{PTS,1}$	0,09	0,0235	0,0	0,0001	0,0302	0,0234
$K_{PTS,2}$	0,0	0,0	0,0	0,0	0,0	0,0
$K_{PTS,3}$	0,00369	0,00244	0,00109	0,0	0,0	0,0
$K_{PTS,g6p}$	0,3135	0,1088	0,0295	0,0	0,0206	0,0207
r_{PTS}^{max}	0,0625	0,0309	0,0113	0,0103	0,0	0,0
$K_{PFK,f6ps}$	0,0	0,0526	0,0882	0,1173	0,0479	0,0677
$K_{PFK,pep}$	0,0	0,0104	0,0270	0,0338	0,0531	0,0570
r_{PFK}^{max}	0,0	0,1216	0,2124	0,1591	0,1188	0,1163
$K_{GAPDH,gap}$	0,0195	0,0732	0,0489	0,0	0,0177	0,0113
$K_{GAPDH,pgp}$	0,0205	0,0577	0,0381	0,0	0,0239	0,0188
r_{GAPDH}^{max}	0,0858	0,1533	0,0947	0,0742	0,0206	0,0154
$K_{PK,pep}$	0,0	0,0	0,0	0,0	0,0	0,0
$K_{PK,fdp}$	0,0	0,0	0,0	0,0	0,0	0,0
r_{PK}^{max}	0,0	0,0	0,0	0,0	0,0	0,0
$K_{PDH,pyr}$	0,0543	0,0205	0,0057	0,0	0,0	0,0
r_{PDH}^{max}	0,0	0,0	0,0	0,0	0,0	0,0
$K_{G6PDH,g6p}$	0,0403	0,0230	0,0320	0,0126	0,0094	0,0167
r_{G6PDH}^{max}	0,0200	0,0223	0,0482	0,0229	0,0182	0,0037
$K_{PGDH,6pg}$	0,0	0,0	0,0	0,0	0,0	0,0
r_{PGDH}^{max}	0,0	0,0	0,0	0,0	0,0	0,0
$\sum_{i=1}^{20} S_i$	0,71009	0,70024	0,63709			

En cuanto a la velocidad de reacción de la enzima fosfofructoquinasa (r_{PFK}), los valores de los índices de sensibilidad de primer orden (Tabla 3.3), permiten inferir que los parámetros que mayor aporte realizan a la varianza de esta variable son $K_{PTS,g6p}$, $K_{PFK,f6ps}$, $K_{PFK,pep}$ y r_{PFK}^{max} . Estos cuatro factores que explican entre el 47% y el 76% de la varianza de la variable r_{PFK} . Los efectos de interacción entre parámetros se consideran despreciables para esta variable algebraica.

El análisis de sensibilidad global basado en análisis de varianza tiene como objetivo establecer la importancia relativa de los parámetros involucrados en la dinámica de un modelo cinético del metabolismo central del carbono de *Escherichia coli*, apuntando a dos cuestiones principales:

1. Establecimiento de parámetros a fijar: este análisis tiene como finalidad identificar los parámetros que pueden ser fijados en cualquier valor dado de sus dominios sin reducir significativamente la varianza de las variables de estado (Homma & Saltelli, 1996). En este caso se evalúan los índices de efecto total (S_i^{TOT})
2. Establecimiento de priorización de parámetros: para determinar los factores más importantes los cuales, si son fijados a sus valores verdaderos, podrían llevar a mayores reducciones en las varianzas de las variables de estado. En este caso se estudian los efectos de primer orden (S_i).

Basado en los criterios anteriormente mencionados y analizando las siguientes variables algebraicas, que corresponden a velocidades de reacción de enzimas que son puntos de regulación en la red metabólica (r_{PTS} , r_{PFK} , r_{PK} , r_{G6PDH}) y diferenciales, que corresponden a las concentraciones de los sustratos y productos de las reacciones catalizadas por las enzimas mencionadas anteriormente (C_{glc}^{ext} , C_{g6p} , C_{f6p} , C_{fdp} , C_{pyr} , C_{pep} , C_{6pg}), se identificaron los parámetros más influyentes en las variables de estado del modelo.

Tabla 3.3. Valores de S_i y S_i^{int} en tres instantes de tiempo para la velocidad de reacción de la enzima piruvato quinasa

	S_i			S_i^{int}		
	t=50	t=300	t=800	t=50	t=300	t=800
$K_{PTS,1}$	0,0134	0,0125	0,0056	0,0	0,0	0,0
$K_{PTS,2}$	0,0	0,0	0,0	0,0	0,0	0,0
$K_{PTS,3}$	0,0	0,0001	0,0019	0,0	0,0	0,0
$K_{PTS,g6p}$	0,0652	0,0507	0,0505	0,0	0,0	0,0
r_{PTS}^{max}	0,0026	0,0	0,0117	0,0023	0,0018	0,0
$K_{PFK,f6ps}$	0,1689	0,1626	0,0618	0,0	0,0	0,1015
$K_{PFK,pep}$	0,0639	0,0864	0,0830	0,0	0,0	0,0376
r_{PFK}^{max}	0,4628	0,4363	0,2801	0,0	0,0	0,0675
$K_{GAPDH,gap}$	0,0	0,0174	0,0251	0,0185	0,0065	0,0
$K_{GAPDH,pgp}$	0,0130	0,0357	0,0457	0,0004	0,0	0,0
r_{GAPDH}^{max}	0,0	0,0	0,0	0,0	0,0	0,0
$K_{PK,pep}$	0,0016	0,0012	0,0026	0,0	0,0	0,0
$K_{PK,fdp}$	0,0	0,0	0,0	0,0	0,0	0,0
r_{PK}^{max}	0,0016	0,0017	0,0037	0,0007	0,0027	0,0066
$K_{PDH,pyr}$	0,0131	0,0095	0,0043	0,0	0,0007	0,0
r_{PDH}^{max}	0,0	0,0	0,0	0,0	0,0	0,0
$K_{G6PDH,g6p}$	0,0022	0,0028	0,0289	0,0	0,0009	0,0166
r_{G6PDH}^{max}	0,0	0,0	0,0384	0,0	0,0	0,0
$K_{PGDH,6pg}$	0,0017	0,0	0,0	0,0012	0,0	0,0
r_{PGDH}^{max}	0,0019	0,0001	0,0	0,0018	0,0	0,0
$\sum_{i=1}^{20} S_i$	0,8119	0,817	0,6433			

Analizando los valores de S_i^{TOT} , bajo la consideración del establecimiento de los parámetros a fijar, se eliminaron ocho parámetros de los veinte iniciales. Entre éstos se pueden destacar, constantes cinéticas del sistema fosfotransferasa (K_{PTS1} , K_{PTS2} , K_{PTS3}), constantes cinéticas y velocidad máxima de reacción de la enzima piruvato quinasa ($K_{PK,pep}$, $K_{PK,fdp}$, r_{PK}^{max}), constante cinética de la enzima

fosfofructoquinasa ($K_{PFK_{pep}}$) y la constante de saturación media de la enzima piruvato deshidrogenasa ($K_{PDH_{pyr}}$).

Del análisis de establecimiento de priorización de parámetros, es decir teniendo en cuenta el efecto de primer orden, r_{PFK}^{max} , K_{f6p} , K_{PTSg6p} , $K_{GAPDHgap}$, $K_{GAPDHpgp}$ y r_{GAPDH}^{max} se identificaron como los seis parámetros más influyentes para la glucosa extracelular y para algunos de los intermediarios glucolíticos como son glucosa-6-fosfato (g6p), fructosa-6-fosfato (f6p), fructosa-1,6-difosfato (fdp) y fosfoenolpiruvato (pep). Otros parámetros del modelo afectan solo a una variable de estado del modelo. En este sentido se puede destacar a la velocidad máxima de reacción del sistema fosfotransferasa (r_{PTS}^{max}) que influye fuertemente en la dinámica de la concentración de glucosa extracelular. Asimismo, la velocidad máxima de reacción de la enzima piruvato deshidrogenasa (r_{PDH}^{max}) afecta exclusivamente a la concentración de piruvato. Por otra parte, el comportamiento de la concentración de 6-fosfogluconato está dominado por las velocidades máximas de reacción y constantes de saturación media de las enzimas glucosa-6-fosfato deshidrogenasa (r_{G6PDH}^{max} y $K_{G6PDH,g6p}$) y 6-fosfogluconato deshidrogenasa (r_{PGDH}^{max} y $K_{PGDH,6pg}$).

Los parámetros seleccionados según el segundo criterio son aquellos para los que el valor de S_i es igual o mayor a 0,05 a lo largo de todo el horizonte de tiempo.

En resumen el análisis de sensibilidad global aplicado al modelo de la red metabólica de *Escherichia coli*, permitió identificar doce parámetros como los más influyentes en el comportamiento dinámico tanto de las variables diferenciales como algebraicas analizadas. Es importante destacar que estos parámetros están

involucrados en cinéticas de enzimas que son puntos regulatorios de la red metabólica, tales como PTS, PFK, G6PDH; así como también en expresiones cinéticas de enzimas que no lo son, como es el caso de GAPDH, PGDH, PDH. Los parámetros seleccionados corresponden a velocidades máximas de reacción, constantes de saturación media y constantes de inhibición de algunas enzimas involucradas en la red metabólica bajo estudio, y están relacionados con propiedades específicas de estos catalizadores biológicos. Por ejemplo, la velocidad máxima de reacción es directamente proporcional a la concentración de enzima, las constantes de saturación media, de inhibición o activación representan la afinidad de una enzima por su sustrato o algún efector. La identificación y posterior calibración de estos parámetros proveen importante información que puede tomarse como punto de partida para la modificación de propiedades específicas de las enzimas.

3.4. Conclusiones

En este capítulo se ha realizado la implementación de una técnica de análisis de sensibilidad global, aplicada a un modelo cinético del metabolismo central del carbono de *Escherichia coli*, que comprende un sistema de ecuaciones diferenciales algebraicas no lineales. Se calcularon los índices de sensibilidad de primer orden y totales, como así también los índices de interacción para las

variables de estado diferenciales y algebraicas involucradas en el modelo mediante la aplicación de un método basado en la varianza (Sobol', 1993).

Mediante la observación de los índices de primer orden (S_i) se identificaron doce parámetros que controlan el comportamiento dinámico del modelo. El parámetro que mayor influencia tiene en las variables de salida del modelo es r_{PFK}^{max} , velocidad máxima de reacción de la enzima fosfofructoquinasa, la cual es un punto de regulación de la vía glucolítica. Por otra parte, mediante la aplicación del ASG, se identificaron parámetros que describen propiedades específicas de enzimas que no son regulatorias de las vías bajo estudio y que tienen gran influencia en el comportamiento dinámico de variables algebraicas y diferenciales del modelo. Asimismo, se detectaron aquellos parámetros que no producen impacto en la evolución de las variables de estado del modelo a lo largo de todo el horizonte de tiempo de estudio. Finalmente, los valores de S_i^{int} , que cuantifican las interacciones entre parámetros, son pequeños para la mayoría de los parámetros. Los resultados obtenidos del análisis de sensibilidad global llevado a cabo son de fundamental importancia para determinar los parámetros más influyentes, que se estimarán en la calibración del modelo, en el Capítulo 4.

CAPÍTULO 4

ESTIMACIÓN DINÁMICA DE PARÁMETROS

4.1. Introducción

Los modelos dinámicos de redes metabólicas, basados en primeros principios, permiten determinar los perfiles temporales tanto de las concentraciones de los metabolitos intra y extracelulares como de las velocidades de reacción. Estos modelos están formados por un sistema de ecuaciones diferenciales algebraicas, que corresponden a los balances de masa para los metabolitos; y expresiones cinéticas de las enzimas involucradas en la red metabólica, respectivamente. Dichas cinéticas enzimáticas involucran un elevado número de parámetros que deben ser calibrados para una condición de cultivo.

La estimación de parámetros en modelos cinéticos del metabolismo celular tiene como objetivo determinar los valores de dichos parámetros para una condición *in vivo*. Normalmente se dispone de datos cinéticos de enzimas *in vitro*. La estimación *in vivo* permite caracterizar el funcionamiento de las enzimas dentro de la red metabólica, en presencia de las demás enzimas que constituyen dicha red. Los avances en las técnicas experimentales y el consecuente incremento de la

disponibilidad de datos, tanto intracelulares como extracelulares, del comportamiento dinámico del metabolismo celular, ha facilitado el desarrollo y calibración de modelos dinámicos de redes metabólicas, que permiten predecir el comportamiento microbiano y que, como se mencionó anteriormente, constituyen una herramienta de gran valor para la Ingeniería Metabólica.

Chassagnole *et al.* (2002) estiman el conjunto completo de parámetros cinéticos (86 parámetros) del modelo dinámico propuesto para el metabolismo central del carbono de *Escherichia coli* K12 W3110, aplicando *Simulated Annealing*, considerando los datos experimentales obtenidos en el mismo trabajo. En un trabajo más reciente, Chassagnole *et al.* (2006) ajustan el modelo dinámico a las mediciones de concentraciones de metabolitos intracelulares, por cuadrados mínimos, utilizando evolución diferencial (*Differential Evolution*). Rizzi *et al.* (1997) utilizan dos herramientas que permiten la reducción del número de parámetros a estimar para un modelo cinético del metabolismo de *Saccharomyces cerevisiae*. Estas herramientas incluyen, la descomposición estructural del modelo y la aplicación de análisis de sensibilidad, explorando la matriz de información de Fisher. Los autores proponen una función objetivo de cuadrados mínimos, resolviendo el problema de estimación de parámetros aplicando el método Simplex modificado según Nelder & Mead (1965), un algoritmo de optimización no restringida que no emplea información de gradientes.

Por otra parte Ceric & Kurtanjek (2006) comparan tres estrategias de optimización: el mencionado algoritmo de optimización no restringida basado en evaluaciones de la función de Nelder & Mead (1965), *Simulated Annealing* y Evolución

diferencial (*Differential Evolution*), en la resolución del problema de estimación de parámetros aplicado a un modelo dinámico del metabolismo central de *Escherichia coli*. Los autores reportan que los mejores resultados fueron obtenidos con el método de Evolución diferencial, ya que permitió explorar un área más amplia en el espacio de los parámetros y cuya velocidad de convergencia no disminuye al acercarse al mínimo de la función objetivo, como en el caso del método Simplex (Nelder & Mead, 1965).

Recientemente, Won *et al.* (2011) proponen modificaciones al modelo del metabolismo central del carbono de *Escherichia coli* formulado por Chassagnole *et al.* (2002) y estiman parámetros cinéticos correspondientes a las enzimas de la vía glucolítica. La función objetivo propuesta por los autores es cuadrados mínimos ponderados por una matriz de pesos, Q , cuyos valores varían entre 1 y 10 dependiendo de la confiabilidad de los datos experimentales. Estos autores resuelven el problema en MATLAB.

En este capítulo, se presenta la formulación y resolución del problema de estimación de parámetros como un problema de optimización dinámica sujeto a un sistema de ecuaciones diferenciales algebraico que representa el modelo dinámico cinético correspondiente al metabolismo central del carbono para *Escherichia coli* formulado en el Capítulo 2 (Di Maggio *et al.*, 2010; Chassagnole *et al.*, 2002). La estimación de parámetros se lleva a cabo basada en datos experimentales de bibliografía (Chassagnole *et al.*, 2002; Degenring *et al.*, 2004). Asimismo, se presentan los resultados de experiencias de fermentación realizadas en un bio-reactor continuo frente a un pulso de glucosa, en las condiciones de

crecimiento del cultivo empleado en la estimación de parámetros del metabolismo celular. Finalmente, se emplean dichos datos experimentales para el ajuste de un modelo cinético no estructurado del quimiostato, formulado también en este capítulo.

En ambos casos el problema de estimación dinámica de parámetros se abordó mediante estrategias de optimización dinámica restringida secuencial (gPROMS, PSE Enterprise, 2007).

4.2. Métodos

4.2.1. Problema de estimación dinámica de parámetros

La necesidad del desarrollo de modelos cinéticos que describan apropiadamente el comportamiento celular, requiere de un gran conocimiento de la dinámica de los componentes de la red metabólica. Las cinéticas de las enzimas se han obtenido, generalmente, de ensayos *in vitro*. Como se mencionó anteriormente, con el objetivo de mejorar la respuesta del modelo, se deben estimar los parámetros cinéticos más influyentes en condiciones *in vivo*; es decir, con las enzimas funcionando dentro de la red metabólica. Para tal fin se realizan determinaciones mediante la inserción de un pulso de sustrato en el bio-reactor y efectuando mediciones de las concentraciones de los metabolitos tanto intra como extracelulares en los tiempos transitorios.

Los datos de entrada necesarios para abordar el problema de estimación de parámetros son: a) datos descriptivos del cultivo (velocidad de dilución, densidad

celular, concentración celular, entre otros); b) valores de los parámetros cinéticos conocidos; c) perfiles de concentración de los metabolitos intracelulares y extracelulares involucrados en la red metabólica; d) condiciones iniciales para las variables de estado del modelo y e) cotas inferiores y superiores para los parámetros a estimar. Como se indicó en el Capítulo 3, se ha realizado una búsqueda bibliográfica minuciosa para obtener el rango de variación de los parámetros del modelo que serán ajustados (Tarmy & Kaplan, 1968; Blangy *et al.*, 1968; Sanwal, 1970; Knappe *et al.*, 1974; Ingram *et al.*, 1987; Diaz Ricci, 1996; Rohwer *et al.*, 2000; Chassagnole *et al.*, 2002; Hoefnagel *et al.*, 2002; Kremling *et al.*, 2004; Ceric & Kurtanjek, 2006; Ratushny *et al.*, 2006).

El problema de estimación dinámica de parámetros restringido por el sistema diferencial algebraico no lineal y que describe el metabolismo de *Escherichia coli*, se ha formulado en el ambiente g-PROMS (g-PROMS, 2007) con una función objetivo de máxima verosimilitud (*Maximum likelihood*). El algoritmo de optimización determina los valores del vector de parámetros cinéticos inciertos y los correspondientes al modelo de la varianza, tales que maximicen la probabilidad de que el modelo de la red metabólica prediga los valores de las mediciones obtenidas experimentalmente. Asumiendo que los errores de las mediciones son independientes y normalmente distribuidos, con media cero y desviación estándar σ_{ij} , la siguiente función objetivo corresponde a un estimador de máxima verosimilitud:

$$\varphi = \frac{N}{2} \ln(2\pi) + \frac{1}{2} \min_p \sum_{i=1}^{NM} \sum_{j=1}^{NT} \left[\ln(\sigma_{ij}^2) + \frac{(C_{ij}^M - C_{ij})^2}{\sigma_{ij}^2} \right] \quad (4.1)$$

donde la suma en la función objetivo es sobre NM variables de estado medidas y NT puntos experimentales para cada variable medida C_{ij}^M ; σ_{ij} es la varianza de la medición j de la variable i la cual es determinada por el modelo de varianza de las variables medidas, en este caso se tomó como modelo el de varianza constante; \mathbf{p} corresponde al vector de parámetros a ser estimados.

El problema de estimación de parámetros para la red metabólica de *E. coli* es el siguiente:

$$\varphi = \frac{N}{2} \ln(2\pi) + \frac{1}{2} \min_p \sum_{i=1}^{NM} \sum_{j=1}^{NT} \left[\ln(\sigma_{ij}^2) + \frac{(C_{ij}^M - C_{ij})^2}{\sigma_{ij}^2} \right] \quad (4.2)$$

s.t.

Modelo DAE del metabolismo de *E. coli* (Ecs.2.2 - 2.57)

$$C_i(0) = C_i^0$$

$$p^L \leq p \leq p^U$$

El desempeño del modelo se ha evaluado cuantitativamente mediante dos medidas diagnósticas calculadas en base a los valores medios de las principales variables de estado del modelo. El primer estadístico aplicado es una medida del sesgo del modelo, el error medio (EM), el cual debe ser cercano a cero y da información de la tendencia del modelo a sobre o subestimar una variable. La segunda medida empleada caracteriza la exactitud del modelo y es el error relativo (ER).

$$EM = \frac{\sum (\text{Valores observados} - \text{Valores predichos})}{n} \quad (4.3)$$

$$ER = \frac{\sum |\text{Valores observados} - \text{Valores predichos}|}{\sum \text{Valores observados}} \quad (4.4)$$

donde n es el número de observaciones.

4.2.2. Algoritmo de optimización

El problema de máxima verosimilitud se ha formulado y resuelto en gPROMS haciendo uso de la herramienta de estimación de parámetros gEST. El solver utilizado fue MXLKHD, que utiliza una estrategia de BDF (*Backward Differentiation Formulae*), con técnicas de matriz rala, en la rutina DASOLV para la resolución del sistema de ecuaciones diferencial algebraico y para el cálculo de las ecuaciones de sensibilidad de las ecuaciones del modelo y de la función objetivo respecto de los parámetros a ser estimados. La información de las ecuaciones de sensibilidad, es decir los gradientes de las restricciones y de la función objetivo respecto de los parámetros en cada instante de tiempo, es empleada en el problema de Programación no Lineal (NLP) del lazo exterior para determinar la dirección de búsqueda en el espacio de los parámetros a determinar. El problema NLP es resuelto con un método de Programación Cuadrática Sucesiva, implementado en el solver SRQPD.

4.2.3. Estimación dinámica de parámetros: Modelo del metabolismo central del carbono

El problema de estimación de parámetros se ha resuelto en base a datos experimentales reportados por Chassagnole *et al.* (2002) para un cultivo continuo de *Escherichia coli* K12, perturbado con un pulso de glucosa. La Tabla 4.1 muestra las condiciones de cultivo.

Se incluyeron datos experimentales correspondientes a los perfiles de concentraciones de glucosa extracelular (C_{glc}^{ext}), glucosa-6-fosfato (C_{g6p}), fructosa-6-fosfato (C_{f6p}), fructosa-1,6-difosfato (C_{fdp}), piruvato (C_{pyr}), fosfoenolpiruvato (C_{pep}), 6-fosfogluconato (C_{6pg}) y gliceraldehído-3-fosfato (C_{gap}). Los datos fueron tomados durante los 20 segundos siguientes a la inserción del pulso de glucosa (Chassagnole *et al.*, 2002). La pequeña escala de tiempo en la que se realizan las mediciones de las concentraciones de los metabolitos permite despreciar los efectos de la regulación genética y, de esta forma, considerar constantes la concentración y actividad de las enzimas (Chassagnole *et al.* 2002, 2006). Esto implica que, con dichas mediciones, solo se podrían observar los efectos de la regulación enzimática ante una perturbación impuesta al sistema (Fig. 4.1). En base a lo explicado anteriormente, las velocidades máximas de reacción (r_i^{max}), que representan la concentración y actividad de las enzimas, se consideran constantes, es decir invariantes en el tiempo. Las velocidades máximas de reacción fueron calculadas según Rizzi *et al.* (1997), como se detalla en el Capítulo 3.

Tabla 4.1. Condiciones de cultivo (Chassagnole *et al.* 2002)

$\text{pH} = 7$
 $T = 35^\circ\text{C}$
 $D = 0,1 \text{ h}^{-1}$
 $\rho_x = 564 \text{ gPS L}^{-1}$ volumen celular
 $C_x = 8,7 \text{ gPS L}^{-1}$ volumen de cultivo
 $C_{glc}^{alimentación} = 111,1 \text{ mM}$
 $\mu = 0,1 \text{ h}^{-1}$
 condiciones aeróbicas

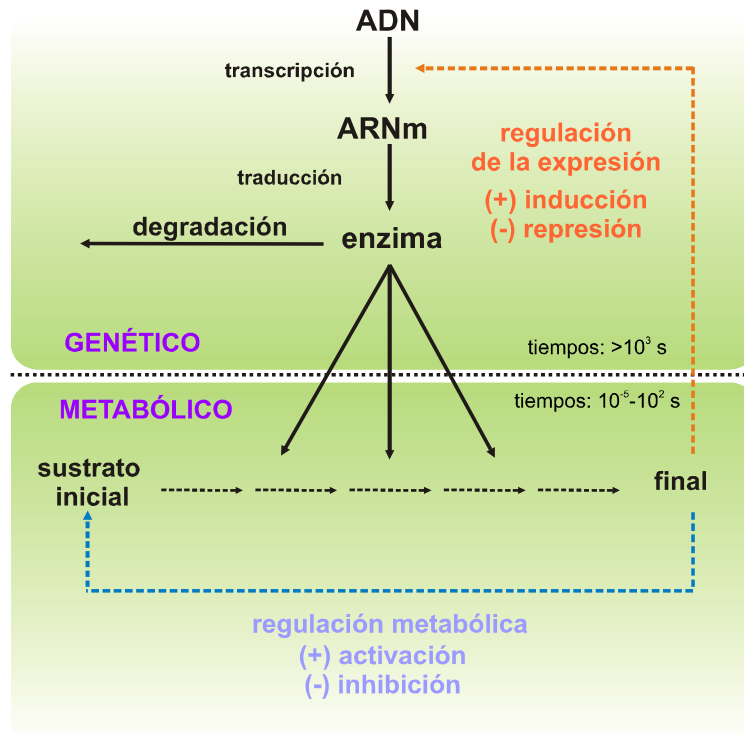


Figura 4.1. Lazo de control genético y metabólico. Las líneas punteadas indican interacciones regulatorias, mientras que las flechas sólidas indican transformaciones químicas primarias. Los diferentes cuadros separan el dominio de la regulación metabólica y la transcripcional.

Basado en el análisis de sensibilidad (Capítulo 3), se ha seleccionado un subconjunto de seis parámetros cinéticos para ser estimados, como son: constantes de saturación media de las enzimas fosfofructoquinasa, gliceraldehído-3-fosfato deshidrogenasa, glucosa-6-fosfato deshidrogenasa y 6-fosfogluconato deshidrogenasa (K_{Rf6p} , $K_{GAPDH,gap}$, $K_{G6PDH,g6p}$ y $K_{PGDH,6pg}$, respectivamente); y constantes de inhibición para el sistema fosfotransferasa y la enzima gliceraldehído-3-fosfato deshidrogenasa ($K_{PTS,g6p}$ y $K_{GAPDH,pgp}$, respectivamente). Asimismo, se consideraron necesarios estimar los siguientes parámetros: constante de saturación media de la piruvato deshidrogenasa ($K_{PDH,pyr}$), constantes de saturación media y de activación de la enzima piruvato quinasa ($K_{PK,fdp}$ y $K_{PK,pep}$, respectivamente)

4.2.4. Estimación dinámica de parámetros: Modelo cinético de bio-reactor

En esta tesis se han llevado a cabo cultivos continuos de *Escherichia coli* en las condiciones descritas en Chassagnole *et al.* (2002) con el objeto de formular y ajustar modelos de producción de compuestos de fermentación en bio-reactores continuos.

Los procesos pueden ser caracterizados por su tiempo de relajación o tiempo característico. Cuando un sistema en estado estacionario es perturbado, se denomina tiempo característico al que dicho sistema necesita para alcanzar el 63% de la distancia entre los valores del estado estacionario de partida y los nuevos. Así, por ejemplo las reacciones enzimáticas que ocurren en el metabolismo celular poseen un tiempo característico del orden de los

milisegundos. Por su parte, la dinámica de cultivos batch, fed-batch y las transiciones de cultivos continuos presentan un tiempo de relajación del orden de horas. De esta manera, los procesos con menor tiempo de relajación que el del sistema bajo estudio, pueden ser considerados en pseudo-estado estacionario (Stephanopoulos *et al.*, 1998). Por lo expuesto anteriormente, en la presente tesis se formuló un modelo no segregado no estructurado (de “caja negra”), de un bio-reactor continuo, con el objetivo de estudiar la dinámica de los productos de fermentación cuando el sistema es perturbado con un pulso de glucosa, en las mismas condiciones que se efectuaron las experiencias para el modelo metabólico. Los balances de masa para el sustrato y los productos de fermentación se formulan de la siguiente manera:

$$\frac{dC_{glc}^{ext}}{dt} = D(C_{glc}^{alimentación} - C_{glc}^{extracelular}) - Y_{XS} \mu C_X \quad (4.5)$$

$$\frac{dC_i}{dt} = r_i - DC_i, \quad i=etanol, acetato, lactato, formiato, succinato \quad (4.6)$$

$$r_i = K_i C_X C_{glc}^{ext} \left(\frac{\mu}{\mu^{max}} \right) \quad i=etanol, acetato, lactato, formiato, succinato \quad (4.7)$$

donde, C_{glc}^{ext} y $C_{glc}^{alimentación}$ (g/l) corresponden a la concentración de glucosa en el bio-reactor y en la alimentación, C_X (g/l) es la concentración de biomasa, respectivamente; D (h^{-1}) es la velocidad de dilución, Y_{XS} ($g_{sustrato}/g_{biomasa}$) es el rendimiento sustrato-biomasa, μ y μ^{max} (h^{-1}) corresponden a la velocidad específica de crecimiento y la velocidad específica de crecimiento máxima, C_i (g/l) es la

concentración de los productos de fermentación (etanol, acetato, formiato, lactato y succinato) y K_i ($\text{g}_{\text{producto}}/\text{g}_{\text{sustrato}} \text{g}_{\text{biomasa}} \text{h}$) corresponde a la constante cinética de formación de cada producto de fermentación. La Ecuación 4.7 corresponde a una expresión cinética de producción de acetato, para *E. coli* utilizando glucosa como fuente de carbono (Alba & Calvo, 2000). Esta expresión cinética fue utilizada para la producción de los restantes productos de fermentación.

4.2.4.1. Ensayos de cultivos continuos

Como parte del trabajo de tesis, se han llevado a cabo ensayos de fermentación en bio-reactores continuos, en las condiciones para las que se ajustó el modelo cinético de redes metabólicas. Los ensayos de cultivos continuos se llevaron a cabo en PROIMI (Planta Piloto de Procesos Industriales Microbiológicos) en San Miguel de Tucumán, bajo la supervisión de la Ingeniera Nora Perotti y la colaboración de los señores Martín Díaz Alfaro y Daniel Etchegorry.

4.2.4.1.1. Condiciones de cultivo

En esta tesis, se ha cultivado *Escherichia coli* K12 W3110 bajo limitación de glucosa, en medio sintético, en condiciones aeróbicas, a 35°C en un bio-reactor tanque agitado continuo (New Brunswick Scientific, Discovery Serie 100), cuyo volumen de trabajo es de 1L. La composición del medio de cultivo fue calculada de forma tal de obtener una concentración de biomasa de 10 gDW/L (Chassagnole *et al.*, 2002); la misma se detalla a continuación: 4 g/L Na_2SO_4 , 5,36 g/L $(\text{NH}_4)_2\text{SO}_4$, 1 g/L NH_4Cl , 7,3 g/L K_2HPO_4 , 1,8 g/L $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$, 120 g/L $(\text{NH}_4)_2\text{-H-citrato}$, 4 ml/L

MgSO₄ (1M), 6 ml/L solución de elementos traza, 0,02 g/L tiamina y 20 g/L glucosa. La solución de tiamina se esterilizó por filtración y se adicionó al medio estéril. Las soluciones de elementos traza y de MgSO₄ fueron esterilizadas en autoclave y luego adicionadas al medio estéril. Asimismo, la solución de glucosa fue preparada y esterilizada en autoclave en forma separada para evitar la caramelización del azúcar, y luego se incorporó al medio de cultivo estéril. El pH del medio de cultivo se mantuvo en 7, con control automático utilizando soluciones de NaOH y HCl de concentración 2 M. El flujo de aire se mantuvo a 1 vvm haciendo circular aire seco generado por un compresor y la velocidad de agitación se mantuvo en 1500 rpm. El nivel del cultivo se controló automáticamente, manteniéndose mediante la adición de antiespumante. La Figura 4.2 muestra el bio-reactor continuo donde se han llevado a cabo las experiencias de adición de pulsos de glucosa a cultivos de *E. coli*.

4.2.4.1.2. Ensayos de pulso de glucosa y toma de muestras

Una vez alcanzado el estado estacionario, a una velocidad de dilución (D) de 0,1 h⁻¹, para el cual la $C_{glc}^{ext} = 0,028$ g/L, se introdujo un pulso de glucosa. Se consideró que luego de tres tiempos de residencia ($\tau=1/D$) el cultivo continuo alcanza nuevamente al estado estacionario.

Se realizaron dos perturbaciones al sistema. En el primer caso, se inyectaron 4 ml de una solución estéril de glucosa (15% w/w) con una jeringa, para obtener una concentración final en el bio-reactor de 0,6 g/L. Luego de realizada la perturbación al sistema, se comenzó la toma de muestras según la siguiente frecuencia: en los

primeros 30 minutos, luego de aplicada la perturbación, se extrajo una muestra cada 5 minutos, en la segunda hora, una muestra cada 10 minutos, en la tercera hora, una muestra cada 20 minutos, en la cuarta hora, una muestra cada media hora y luego se continuó extrayendo una muestra por hora durante las 6 horas siguientes.



Figura 4.2. Cultivos de *E.coli* en bio-reactor continuo

En el segundo caso, se inyectaron 3,5 ml de una solución estéril de glucosa (30% w/w), a partir del estado estacionario ($C_{glc}^{ext} = 0,041$ g/L), para dar una concentración final en el bio-reactor de 1,0 g/L. En esta oportunidad, la frecuencia de muestreo fue la que se detalla a continuación: la primera muestra se extrajo a los 5 minutos de haber realizado el pulso de glucosa, en la hora siguiente se extrajeron dos muestras, una cada 30 minutos y por último se continuó extrayendo 1 muestra por hora durante 8 horas. En ambos casos, el volumen de solución de

glucosa inyectado al perturbar el sistema fue menor que el 1% del volumen de trabajo del bio-reactor, hecho que permite despreciar cualquier efecto en la velocidad de dilución. Las muestras obtenidas fueron analizadas según se detalla en la sección 4.2.4.2.

4.2.4.2. Métodos analíticos

4.2.4.2.1. Peso seco

Para la determinación de peso seco se centrifugó 1 ml de la muestra original (3 ml) a 12.000 g, durante 3 minutos, se descartó el sobrenadante y el pellet obtenido se secó en estufa a 105°C hasta obtener peso constante. Esta determinación se realizó en las muestras tomadas cada hora.

4.2.4.2.2. Glucosa y productos de fermentación

Para la determinación de glucosa y productos de fermentación (etanol, ácido acético, ácido láctico, ácido fórmico y ácido succínico) se centrifugaron 2 ml de la muestra original (3 ml) a 12.000 g, durante 3 minutos, se recuperó el sobrenadante, el cual se mantuvo a -18°C hasta el momento del análisis.

La determinación de los compuestos se realizó en PLAPIQUI mediante cromatografía líquida de alta presión (HPLC) (Waters e2695). Se empleó una columna de intercambio iónico Phenomenex Rezex ROA - Organic Acid (300 x 7,8 mm, 8 µm de tamaño de partícula). Como fase móvil se utilizó una solución de H₂SO₄ 0,005N, el volumen de inyección de muestra fue de 10 µL, para el caso de

los estándares utilizados para realizar las curvas de calibrado el volumen de inyección fue de 25 μL . El flujo de la fase móvil fue de 0,4 mL/min, mientras que la temperatura de la columna se mantuvo en 60°C. Se utilizó un detector de índice de refracción (Waters 2414), cuya temperatura de operación se mantuvo en 40°C. Para la determinación cuantitativa de la glucosa y productos de fermentación se realizaron curvas de calibrado para cada uno de los compuestos (Tabla 4.2).

Tabla 4. 2. Curvas de calibrado para glucosa, etanol, ácido láctico, ácido acético, ácido succínico y ácido fórmico.

Compuesto	Rango [g/L]	Ecuación de la curva	R ²
Glucosa	0,028-1,44	$y = 0,0000002x + 0,017$	0,999
Etanol	0,0011-1,103	$y = 0,0000007x - 0,0007787$	0,999
Acido láctico	0,0106-2,6938	$y = 0,0000006x + 0,0012077$	0,999
Acido acético	0,0125-3,125	$y = 0,0000006x + 0,004$	0,999
Acido succínico	0,0012-1,164	$y = 0,0000004x$	0,999
Acido fórmico	0,0305-3,054	$y = 0,0000008x - 0,007$	0,999

x = área del pico, y = concentración [g/L]

4.3. Resultados y discusión

4.3.1. Modelo del metabolismo central del carbono

El problema de estimación de parámetros propuesto (Ec. 4.1) se ha formulado en g-PROMS como un problema de máxima verosimilitud.

Basado en resultados del análisis de sensibilidad implementado en el Capítulo 3 y considerando que sólo las constantes cinéticas pueden ser calibradas, 9 parámetros cinéticos fueron estimados. Los datos experimentales fueron extraídos de Chassagnole *et al.* (2002), en la Tabla 4.3 se muestran las condiciones iniciales

de cada variable de estado, que equivalen a la concentración del estado estacionario inicial antes de la aplicación del pulso de glucosa; respectivamente.

En la Tabla 4.4 se observan los valores estimados de dichos parámetros así como también su valor inicial y sus cotas, tomados de bibliografía (Tarmy & Kaplan, 1968; Blangy *et al.*, 1968; Sanwal, 1970; Knappe *et al.*, 1974; Ingram *et al.*, 1987; Diaz Ricci, 1996; Rohwer *et al.*, 2000; Chassagnole *et al.*, 2002; Hoefnagel *et al.*, 2002; Kremling *et al.*, 2004; Ceric & Kurtanjek, 2006; Ratushny *et al.*, 2006).

Tabla 4.3. Condiciones iniciales para las variables de estado (Chassagnole *et al.*, 2002; Degenring 2004)

Degenring (2004)		Chassagnole <i>et al.</i> (2002)	
Metabolito	Concentración (mM)	Metabolito	Concentración (mM)
Glucosa	0,27	Glucosa	0,056
g6p	0,16	g6p	3,48
f6p	0,17	f6p	0,6
fdp	2,71	fdp	0,272
gap	5,58	gap	0,218
dhap	5,02	dhap	0,167
3pg	2,37	pgp	0,008
2pg	2,36	2pg	2,13
pep	0,92	3pg	0,399
pyr	1,26	pep	2,67
6pg	3,66	pyr	2,67
5p	1,37	6pg	0,808
		ribu5p	0,111
		rib5p	0,398
		xyl5p	0,138
		sed7p	0,276
		e4p	0,098
		g1p	0,653

5p=rib5p+ribu5p+xyl5p

En la solución óptima los residuos ponderados toman un valor igual a 38,86 y menor que el valor de χ^2 (44,98), esto es indicador de un buen ajuste.

Las Figuras 4.2(a) a 4.2(g) muestran los perfiles obtenidos para las principales variables de estado comparados con los datos experimentales (Chassagnole *et al.*, 2002). Se puede observar que el modelo reproduce razonablemente la respuesta dinámica de cada metabolito al pulso de glucosa aplicado.

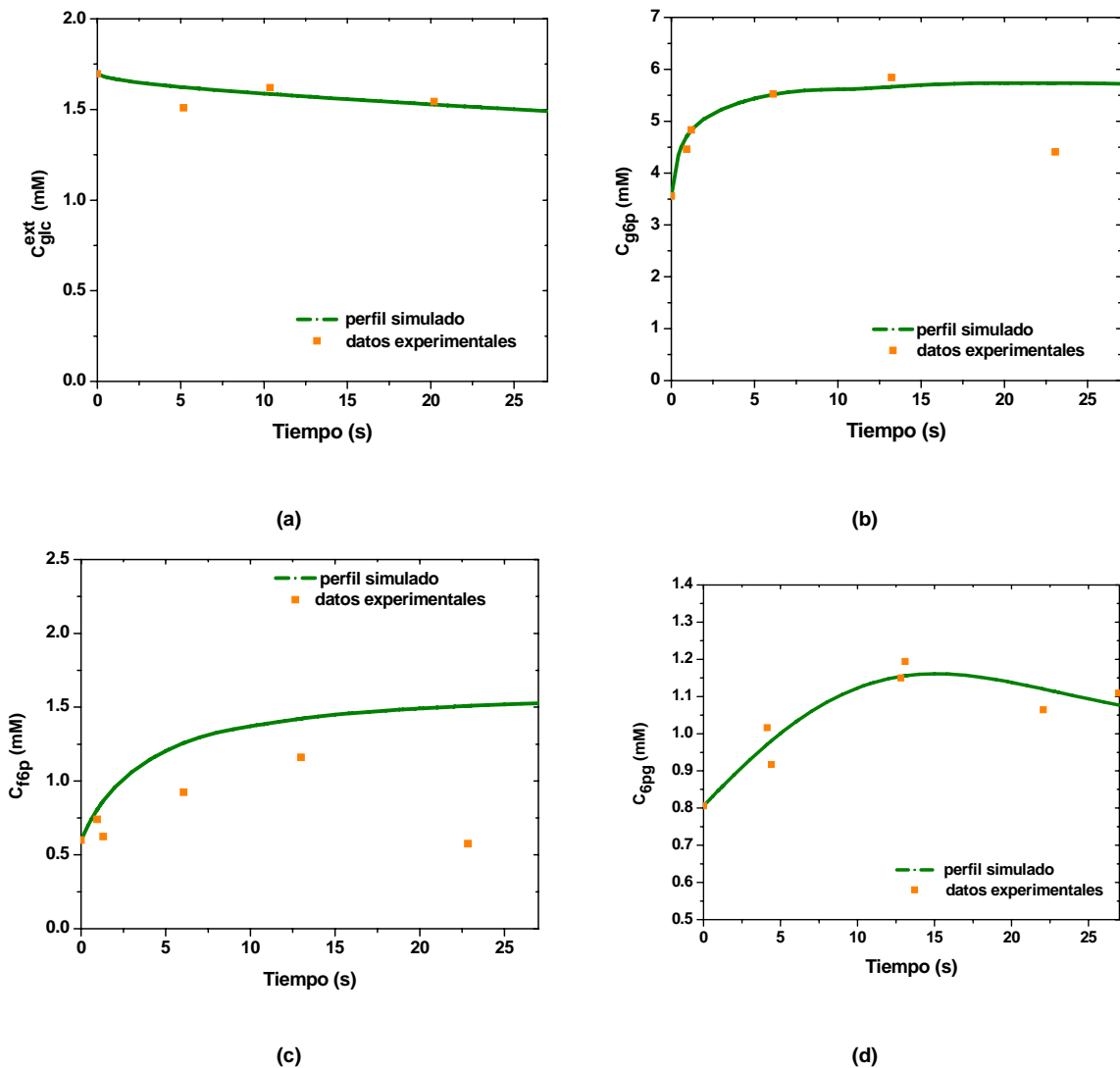


Figura 4.3. Perfil simulado y datos observados (Chassagnole *et al.*, 2002) para la concentración de glucosa extracelular (a), glucosa-6-fosfato (b), fructosa-6-fosfato (c) y 6-fosfogluconato (d).

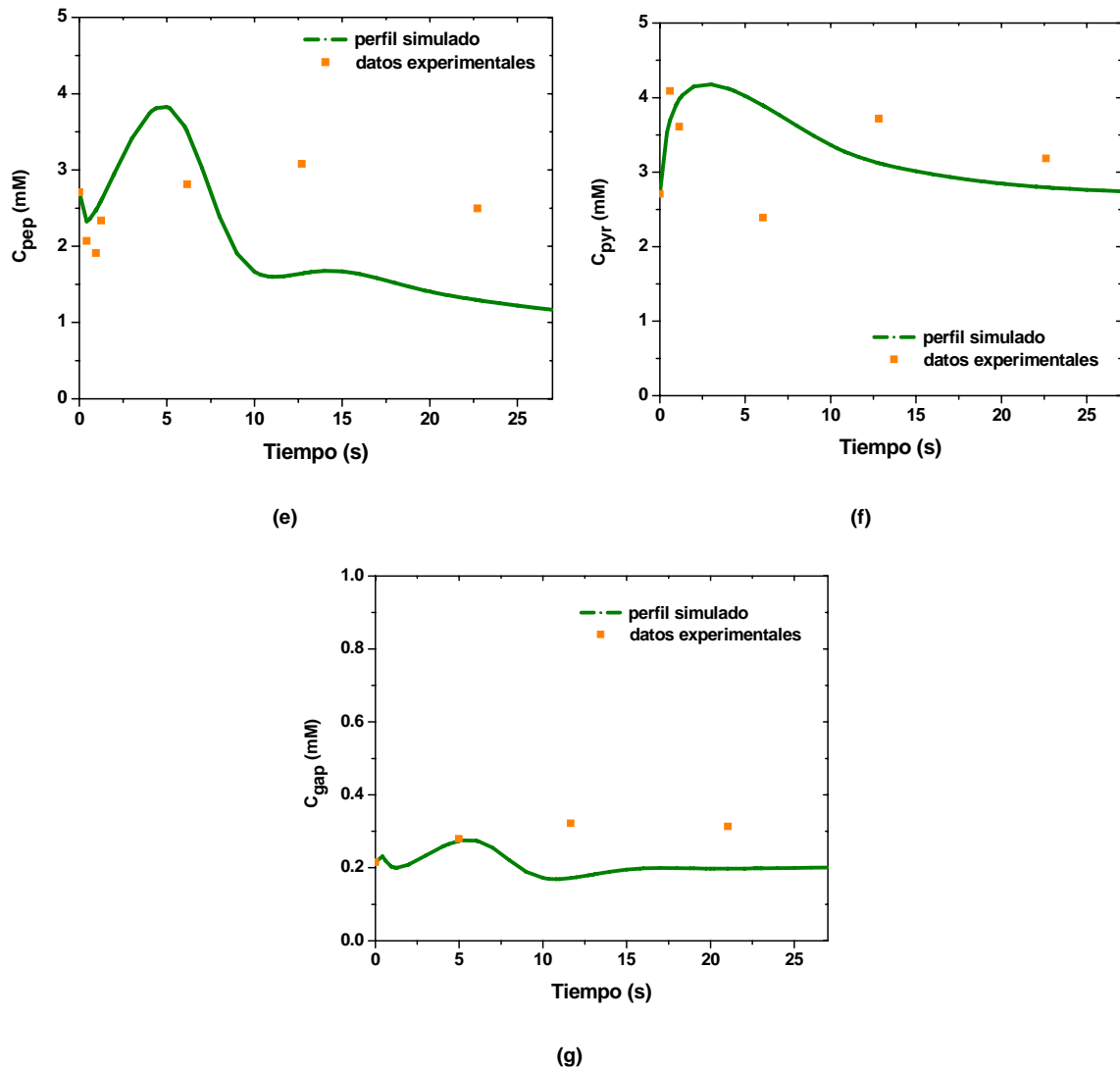


Figura 4.4. Perfil simulado y datos observados (Chassagnole *et al.*, 2002) para la concentración fosfoenolpiruvato (e), piruvato (f) y gliceraldehído-3-fosfato (g).

En la Tabla 4.5 se muestran los valores del error relativo (ER) y el error medio (EM) calculados a partir de los resultados obtenidos en la estimación de parámetros. Los valores del error medio muestran un buen compromiso entre los datos observados y las salidas del modelo para todas las variables analizadas. El EM en el caso de fructosa-6-fosfato tiene el mayor valor, igual a -0,306 (Tabla 4.4).

Tabla 4.4. Valor óptimo, nominal y cotas superior e inferior de los parámetros estimados del modelo de la red metabólica de *E. coli*

Parámetro	Valor inicial	Cota inferior	Cota superior	Valor óptimo
$K_{GAPDH,gap}$	0,19	0,04	0,68	0,381
$K_{GAPDH,pgp}$	0,0001	0,00001	0,015	0,0002
$K_{mG6PDH,g6p}$	0,491	0,005	3,51	1,127
$K_{PDH,pyr}$	1,336	0,5	1,8	1,5
$K_{PGDH,6pg}$	25	18,5	37,6	20
$K_{PFK,fdp}$	0,259	0,12	0,9	0,9
$K_{PK,pep}$	0,152	0,031	0,2	0,1
$K_{PTS,g6p}$	0,3	0,2	0,4	0,2
K_{f6p}	0,00125	0,0006	0,15	0,0007

Con respecto al valor del error relativo las variables que mejor ajusta el modelo son la glucosa extracelular (ER = 2,61%), 6-fosfogluconato (ER = 3,30%) y glucosa-6-fosfato (ER = 6,17%). Los valores de ER para las concentraciones de piruvato (ER = 16,52%), gliceraldehído-3-fosfato (ER = 24,13%) y fosfoenolpiruvato (ER = 25,31%) muestran que el ajuste es razonable. La mayor discrepancia entre los datos observados y la predicción del modelo ocurre nuevamente para la concentración de fructosa-6-fosfato, cuyo valor de ER asciende al 39,71% (Tabla 4.5).

Tabla 4.5. Mediciones estadísticas para la bondad de ajuste del modelo dinámico del metabolismo celular de *Escherichia coli*

	C_{glc}^{ext}	C_{g6p}	C_{f6p}	C_{6pg}	C_{gap}	C_{pep}	C_{pyr}
ER	2.61%	6.17%	39.71%	3.30%	24.13%	25.31%	16.52%
EM	-0.015	-0.230	-0.306	-0.001	0.068	0.124	-0.081

La dinámica de los flujos alrededor de fosfoenolpiruvato se muestra en la Figura 4.3(a). Se puede observar que cuando el sistema es perturbado con un pulso de glucosa en $t=0$, la actividad del sistema fosfotransferasa es lo suficientemente elevada, logrando reproducir el decrecimiento observado en la concentración de pep, que es metabolito de dicho sistema PTS (Fig. 4.2(e)). Asimismo, los datos experimentales muestran que la concentración de pep comienza a aumentar a partir de los 3 s (Fig. 4.2(e)), el modelo logra reproducir dicho incremento debido al aumento en la actividad de la enzima enolasa (ENO) (Fig. 4.3(a)), sin embargo a partir de los 5 s el modelo predice una disminución en la concentración de pep, que no se corresponde con la observación experimental, debido a que la actividad de la enzima fosfoenolpiruvato carboxilasa (PEPCxylase) sumada a la actividad del sistema fosfotransferasa (PTS) son superiores a la actividad de la enzima enolasa (Fig. 4.3(a)).

Como se observa en la Figura 4.2(c) el modelo sobreestima la concentración de fructosa-6-fosfato (f6p), hecho que puede atribuirse a que el flujo de la enzima fosfoglucosa isomerasa (PGI), que cataliza la conversión de g6p a f6p, alcanza valores mayores que el resto de los flujos en el nodo f6p (Figura 4.3(b)). En $t=20$ s, la velocidad de la PGI se iguala a la velocidad de reacción de la enzima fosfofructoquinasa (PFK).

Es importante destacar que la certeza de los datos experimentales es altamente dependiente del método de recolección de las muestras. Un procedimiento típico de muestreo para la cuantificación de metabolitos tanto intracelulares como extracelulares comprende dos etapas: una es la recolección de la muestra y otra

es el enfriamiento rápido de las mismas para detener las reacciones metabólicas y de esta manera evitar cambios en las concentraciones de los metabolitos (Hiller *et al.*, 2007). Según Schaefer *et al.* (1999) las técnicas de muestreo para medición de metabolitos intracelulares solo son exitosas si (i) se logra tomar una muestra representativa del medio de cultivo sin perturbar las reacciones; (ii) la inactivación del metabolismo es rápida comparada con las velocidades de las reacciones metabólicas; (iii) el muestreo, la inactivación y los procedimientos de extracción no afectan la estabilidad de los metabolitos y (iv) la velocidad de muestreo es lo suficientemente rápida como para estudiar la dinámica de las reacciones metabólicas.

La Figura 4.4 muestra los perfiles temporales simulados de las concentraciones de ribulosa-5-fosfato (Figura 4.4(a)), xilulosa-5-fosfato (Fig. 4.4(b)) y ribosa-5-fosfato (Fig. 4.4(c)), que son metabolitos intermediarios de la vía de las pentosas.

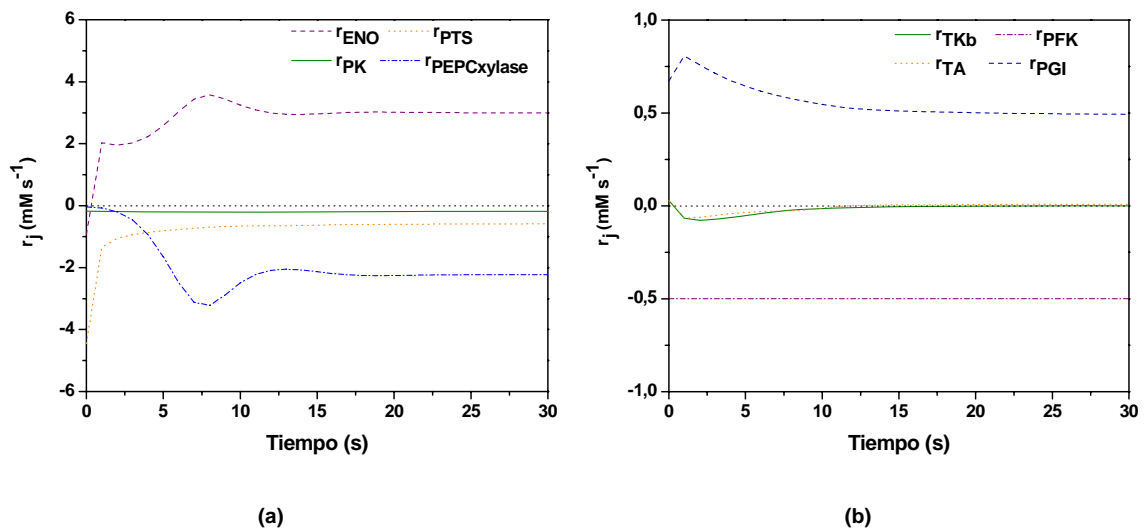


Figura 4.5. Perfil de flujos simulados en los nodos fosfoenolpiruvato (a) y fructosa-6-fosfato (b)

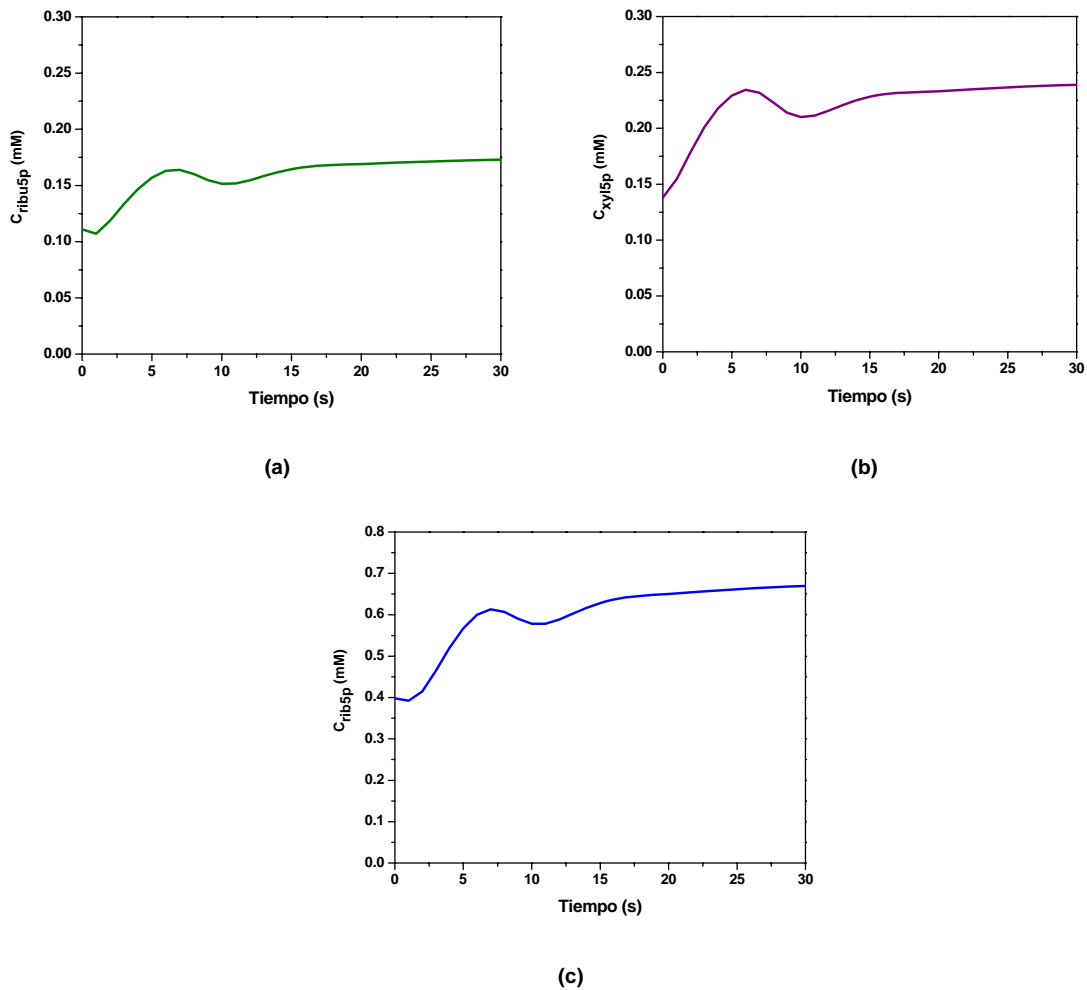


Figura 4.6. Perfiles simulados para concentraciones de tres intermediarios de la ruta de las pentosas: ribulosa-5-fosfato (a), xilulosa-5-fosfato (b) y ribosa-5-fosfato (c)

Una vez calibrado el modelo, se evaluó la performance del mismo comparando los resultados obtenidos por simulación con un conjunto de datos experimentales reportado por Degenring (2004). En la Tabla 4.3 se muestran las concentraciones de los metabolitos en el estado estacionario, que corresponden a las condiciones iniciales del sistema. La Tabla 4.6 muestra las condiciones a las que fue realizado el cultivo. En la Figura 4.5 se observan los perfiles predichos y los datos observados de fructosa-1,6-difosfato (a), fosfoenolpiruvato (b), glucosa-6-fosfato

(c) y piruvato (d). Se puede observar que el modelo calibrado reproduce las tendencias de los datos experimentales, siendo la fructosa-1,6-difosfato, la variable mejor representada.

Tabla 4.6. Condiciones de cultivo (Degenring 2004)

pH = 7
T = 37°C
D = 0,125 h ⁻¹
ρ_X = no reportada
C _X = 10 gPS L ⁻¹ volumen de cultivo
C _{glc} ^{alimentación} = no reportada
μ = 0,125 h ⁻¹
condiciones aeróbicas

Las discrepancias observadas entre los perfiles predichos y los datos observados pueden atribuirse a que, Degenring (2004) no reporta un conjunto completo para las condiciones iniciales de cada metabolito involucrado en la red. Por otra parte, para realizar las simulaciones, se han utilizado las correlaciones para las concentraciones de los co-metabolitos propuestas por Chassagnole *et al.* (2002), ya que Degenring (2004) no reporta las concentraciones del conjunto completo de co-metabolitos involucrado en la red metabólica bajo estudio. Esto puede constituir una fuente importante de error ya que las condiciones de cultivo del trabajo de Chassagnole *et al.* (2002) difieren de las condiciones de trabajo de Degenring (2004).

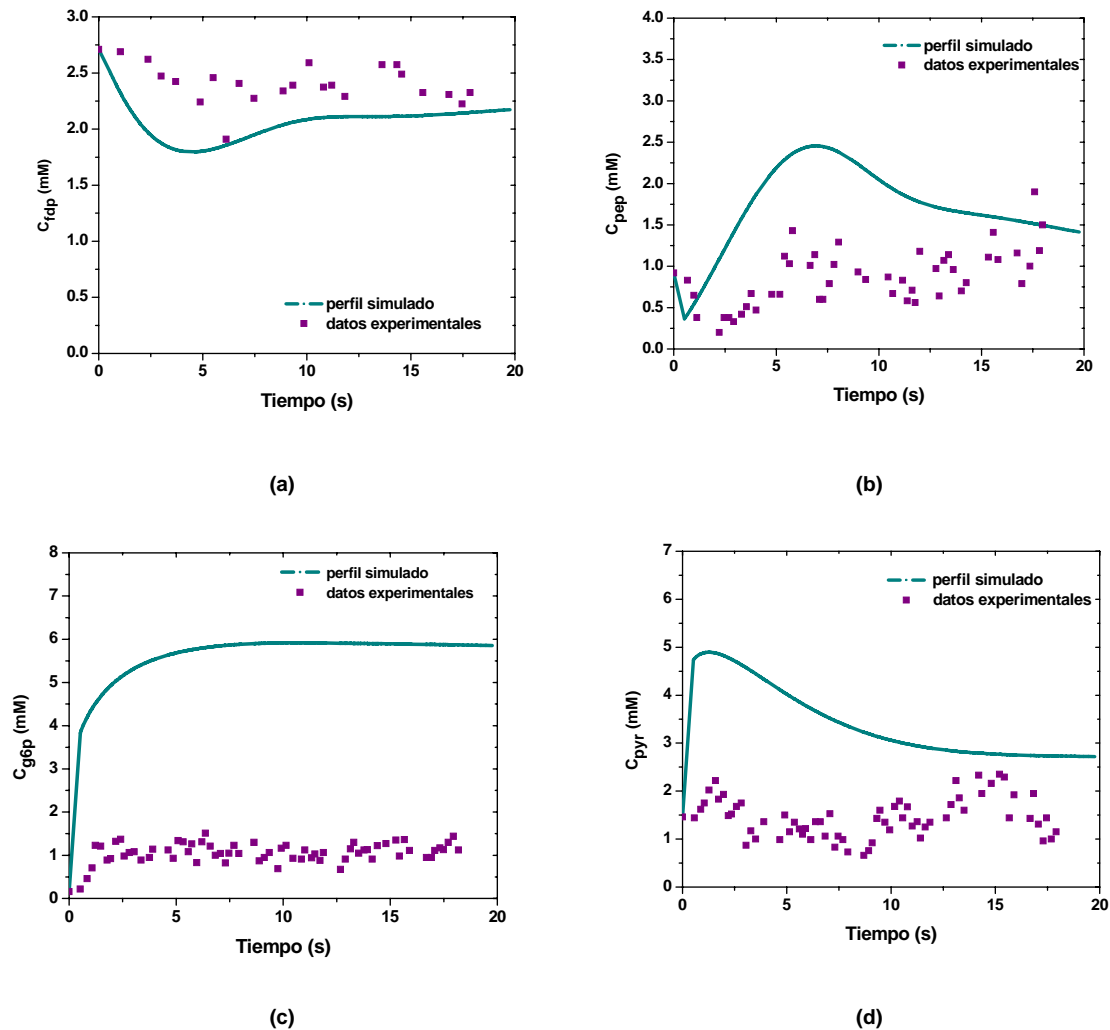


Figura 4.7. Perfil simulado y datos observados para las concentraciones de fructosa-1,6-difosfato (a), fosfoenolpiruvato (b), glucosa-6-fosfato (c) y piruvato (d).

4.3.2. Ensayos de cultivo continuo y calibración del modelo del bio-reactor

El problema de estimación de parámetros propuesto (Ec. 4.1) ha sido formulado en g-PROMS como un problema de máxima verosimilitud.

Utilizando los datos correspondientes a la segunda experiencia, llevada a cabo en la presente tesis (pulso de glucosa=1,0 g/l), se estimaron las cinco constantes

cinéticas de formación de los productos de fermentación, el rendimiento sustrato-biomasa y la velocidad específica de crecimiento máxima. La determinación de peso seco mostró que la concentración de biomasa se mantuvo constante e igual a 10 g/l durante el tiempo completo de ensayo. Por esta razón no se incluyó el balance de masa para la biomasa en la formulación del modelo del bio-reactor (Ecs. 4.5 a 4.7).

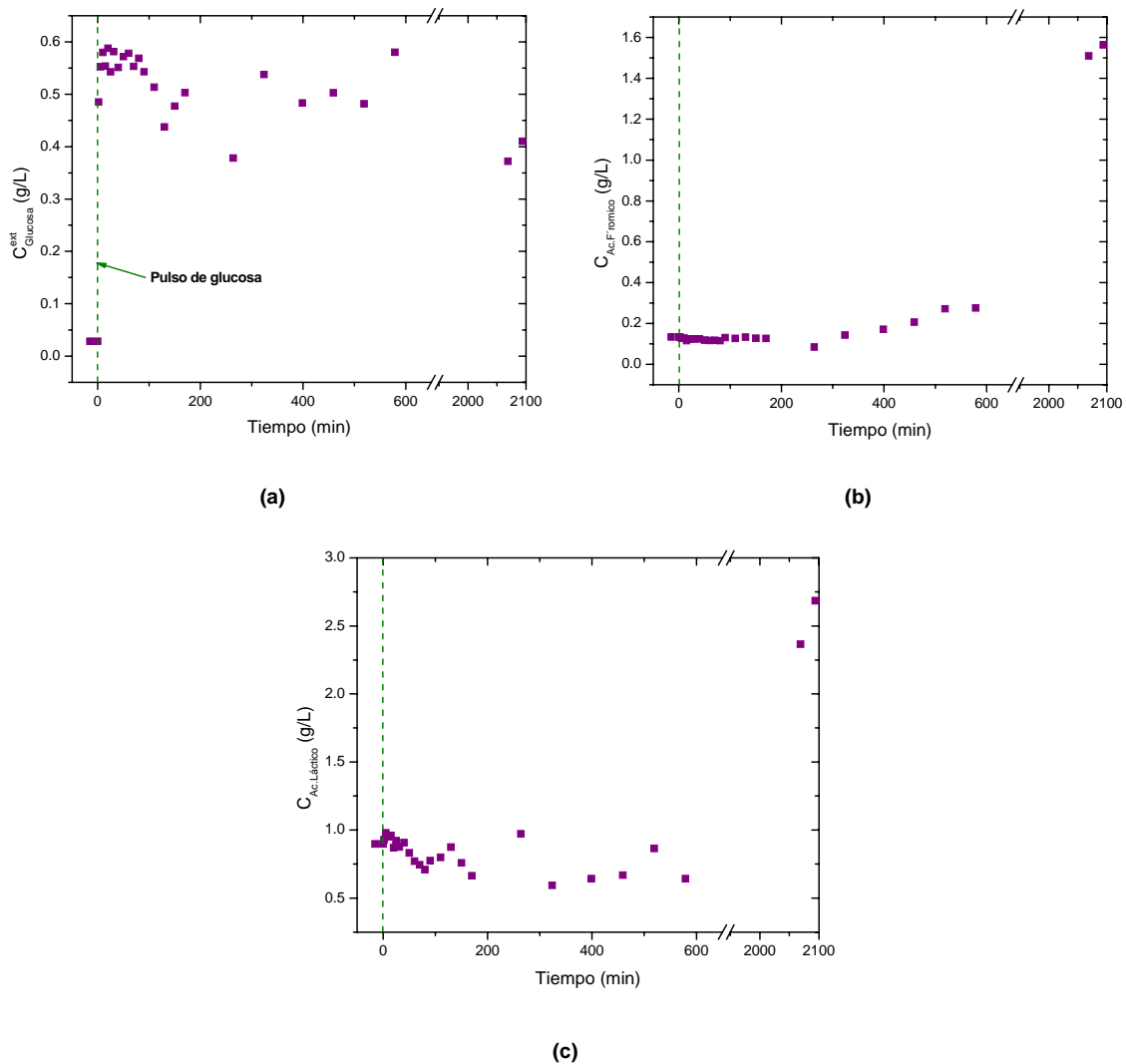


Figura 4.8. Datos observados para la concentración de glucosa extracelular (a), formiato (b) y lactato (c).

En la Tabla 4.7 se observan los valores estimados de los parámetros así como también su valor inicial. En la solución óptima los residuos ponderados toman un valor igual a 51,54 y menor que el valor de χ^2 (56,94), indicador de un buen ajuste. La Figura 4.6 muestra los datos observados para la concentración de glucosa extracelular (a), de formiato (b) y de lactato (c), obtenidos en la primera experiencia (pulso 1).

En la Figura 4.6(a), se observa que la concentración de glucosa extracelular residual es de 0,028 g/l. Luego de aplicado el pulso de glucosa, dicha concentración alcanza un valor de 0,57 g/l. A las 35 h de ensayo, la concentración de glucosa se mantenía en 0,41 g/l, resultado que se obtuvo tanto con una medición espectrofotométrica como cromatográfica (Sección 4.2.4.2.2).

Tabla 4.7. Valores nominal y optimizado para los parámetros estimados del modelo del bio-reactor

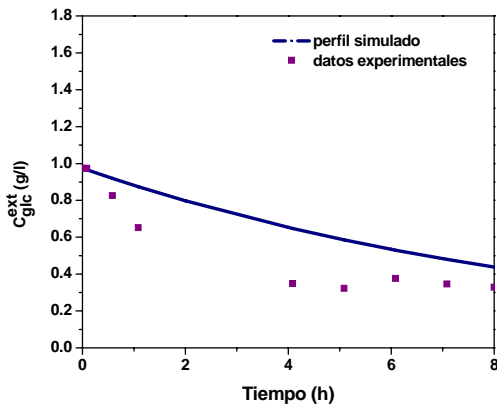
Parámetro	Valor nominal	Valor optimizado
μ^{max}	1,4	2,3
Y_{XS}	1,0	2,0
K_{etoh}	6,0	0,237
K_{ac}	6,0	1,636
K_{lac}	6,0	1,271
K_{suc}	6,0	0,124
K_{form}	6,0	0,411

En las Figuras 4.6 (b) y (c) se puede observar que luego de las 35 h de ensayo, tanto la concentración de formiato como de lactato no logran estabilizarse. Esto nos permite inferir que el sistema bajo estudio no se estabiliza luego de transcurridos los tres tiempos de residencia.

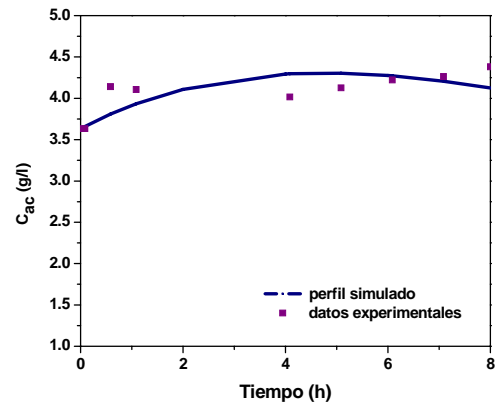
La Figura 4.7 muestra los datos observados, obtenidos cuando se perturbó por segunda vez al sistema, junto con los perfiles simulados de las concentraciones de glucosa extracelular (a), acetato (b), formiato (c), etanol (d), lactato (e) y succinato (f). La Figura 4.7 muestra que *Escherichia coli* K12 aún creciendo en condiciones aeróbicas produce los cinco productos de fermentación, siendo los mayoritarios el acetato y el lactato. Las concentraciones alcanzadas fueron de 4,5 y 3,5 g/l de acetato y lactato, respectivamente. Según Lin *et al.* (2005), en condiciones aeróbicas el acetato es el principal sub-producto del metabolismo de *Escherichia coli*, coincidiendo con los resultados obtenidos en la presente tesis. Por otra parte, Lin *et al.* (2005) sostienen que el succinato es un producto minoritario del metabolismo anaeróbico de *Escherichia coli*, no produciéndose en condiciones aeróbicas. Sin embargo Diaz Ricci *et al.* (1991a) logran cuantificar succinato en un cultivo de *Escherichia coli* creciendo en condiciones aeróbicas. En la presente tesis se detectó la producción de succinato, tal como se muestra en la Figura 4.7(f), siendo uno de los productos minoritarios.

La Figura 4.7 muestra que modelo reproduce razonablemente la respuesta dinámica del sustrato y de los productos al pulso de glucosa aplicado.

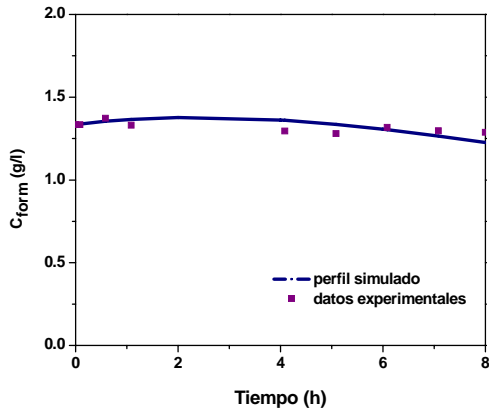
En la Tabla 4.8 se muestran los valores del error relativo (ER) y el error medio (EM) calculados a partir de los resultados obtenidos en la estimación de parámetros. Los valores del error medio muestran un buen compromiso entre los datos observados y las salidas del modelo para todas las variables analizadas, siendo la glucosa extracelular la variable que presenta el mayor EM (0,128). Con respecto a los valores de ER se puede observar que las mejores predicciones del



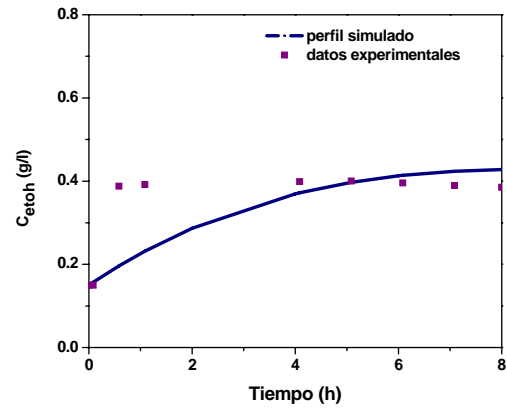
(a)



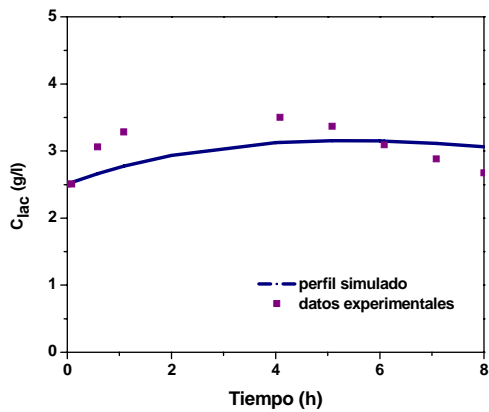
(b)



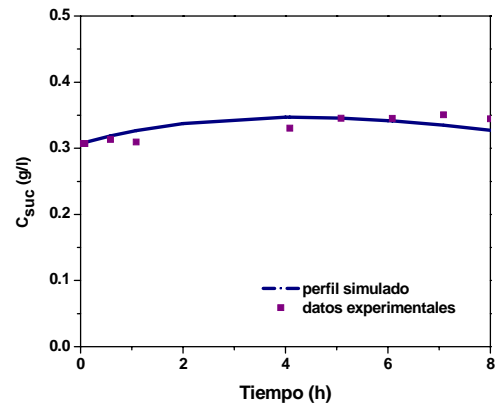
(c)



(d)



(e)



(f)

Figura 4.9. Perfiles y datos observados para conc. de glucosa extracelular (a), acetato (b), formiato (c), etanol (d), lactato (e) y succinato (f)

modelo se logran para las concentraciones de lactato (ER=8,17%), acetato (ER=3,70%), formiato (ER=2,40%) y succinato (ER=2,32%). Por otra parte las mayores discrepancias se obtienen para la concentración de glucosa extracelular, cuyo valor de ER asciende al 24,30% y para la concentración de etanol, con ER=15,85%.

Tabla 4.8. Mediciones estadísticas para la bondad de ajuste del modelo del bio-reactor

	C_{glc}^{ext}	C_{lac}	C_{ac}	C_{form}	C_{suc}	C_{etoh}
ER	24,30%	8,17%	3,70%	2,40%	2,32%	15,85%
EM	-0,128	0,088	0,031	-0,004	-0,002	0,031

4.4. Conclusiones

En este capítulo se ha formulado un problema de estimación dinámica de parámetros tanto para el modelo del metabolismo central del carbono de *Escherichia coli* como para el modelo de un bio-reactor. El problema de estimación dinámica de parámetros restringido por el sistema diferencial algebraico no lineal se ha formulado en el ambiente g-PROMS (g-PROMS, 2007) con una función objetivo de máxima verosimilitud (*Maximum likelihood*).

En el caso del modelo del metabolismo central del carbono, se han estimado nueve parámetros en base a datos experimentales de concentraciones de metabolitos intracelulares y de glucosa extracelular obtenidos de bibliografía. Los resultados de la estimación de parámetros muestran que el modelo predice

razonablemente la respuesta dinámica del metabolismo celular ante la aplicación de una perturbación de tipo pulso.

Finalmente, se han estimado siete parámetros del modelo cinético de bio-reactor continuo. En este caso se utilizaron los perfiles de concentración de productos de fermentación y de glucosa extracelular obtenidos en este trabajo como datos de entrada para la estimación de parámetros. Los resultados muestran que la cinética de producción propuesta para todos los productos de fermentación logra reproducir razonablemente el comportamiento de los datos observados.

CAPÍTULO 5

DISEÑO ÓPTIMO DE REDES METABÓLICAS ESTABLES

5.1. Introducción

En la naturaleza existe una gran diversidad de microorganismos que se han adaptado a diferentes ambientes, a lo largo del tiempo. Es sabido que la vida microbiana puede desarrollarse en condiciones extremas, como por ejemplo temperaturas por debajo de los 0°C o superiores a los 100°C, pHs cercanos a 1 o a 13, elevadas concentraciones de sal, entre otras. Este proceso de adaptación ha posibilitado simultáneamente la manifestación de una amplia gama de rutas metabólicas.

En ciertos casos, es posible introducir una vía metabólica específica a un microorganismo o bien modificar una existente, mediante el uso de técnicas provistas por la Ingeniería Genética, con el objeto de optimizar los rendimientos de un producto de interés o mejorar propiedades celulares que harían que dichos microorganismos sean más apropiados para una aplicación comercial (Stephanopoulos *et al.*, 1998). Estas técnicas de modificación dirigida demostraron

ser potencialmente mejores que las técnicas clásicas de mutagénesis al azar, siempre y cuando se conozca previamente el efecto de la modificación.

La complejidad que exhibe el metabolismo celular debido a la interacción de rutas metabólicas y a la regulación alostérica, hace indispensable el uso de modelos matemáticos, en particular modelos cinéticos del metabolismo celular, para evaluar los efectos que se producirían en la red al modificar las concentraciones de las enzimas, sus propiedades o bien, introducir una nueva ruta.

Por otra parte, los modelos cinéticos permitirían predecir la distribución óptima de los niveles de las enzimas, por ejemplo para maximizar la productividad, mediante la aplicación de técnicas de optimización (Visser *et al.*, 2004), dando lugar a lo que se conoce como “problema de diseño”.

Es importante mencionar que las manipulaciones genéticas pueden llevar a la red metabólica a un nuevo estado estacionario diferente del original (Nikolaev, 2010), y potencialmente vulnerable a perturbaciones externas.

Los modelos cinéticos permiten predecir no solo la existencia de diferentes estados estacionarios sino también caracterizarlos dinámicamente. Esto es de suma importancia ya que los sistemas biológicos pueden exhibir multiestabilidad, oscilaciones y comportamiento caótico (Vital-Lopez *et al.*, 2006).

Es por esto que se considera de gran importancia caracterizar el estado estacionario alcanzado desde el punto de vista de la estabilidad cuando se modifican los niveles de las enzimas o bien incluir en la formulación del problema de diseño restricciones que aseguren la estabilidad del nuevo estado estacionario.

Diaz Ricci (1996) estudia la estabilidad local del comportamiento de la enzima fosfofructoquinasa (PFK), analizando los autovalores de un modelo cinético del comportamiento dinámico del sustrato y producto de la enzima PFK que contempla la expresión cinética de dicha enzima. Los resultados muestran que el comportamiento dinámico de la fosfofructoquinasa puede exhibir nodos y espirales tanto estables como inestables dependiendo de la concentración de fosfoenolpiruvato.

Visser *et al.* (2004) formulan un problema de diseño óptimo para una red metabólica de *Escherichia coli*. En este trabajo se calculan los niveles óptimos de las enzimas con el objetivo de maximizar la producción de serina y el flujo a través del sistema fosfotransferasa. Para abordar el problema de diseño se comparan dos modelos cinéticos del metabolismo celular: un modelo mecanístico (Chassagnole *et al.*, 2002) y un modelo en el cual se utilizan expresiones cinéticas aproximadas del tipo “linlog”. Vital-López, *et al.* (2006) estudian la estabilidad del metabolismo central del carbono de *Escherichia coli* (Chassagnole *et al.*, 2002) para los niveles óptimos de las enzimas obtenidos al maximizar la producción de serina (Nikolaev *et al.*, 2005). Para tal fin, los autores presentan diagramas de bifurcación, considerando los niveles de diferentes enzimas como parámetros de bifurcación. Los resultados muestran que para ciertos valores de los parámetros el metabolismo de *Escherichia coli* presenta un comportamiento oscilatorio. Nikolaev *et al.* (2005) formulan un problema de programación no lineal mixto entero (MINLP), para el modelo del metabolismo central del carbono propuesto por Chassagnole *et al.* (2002), que permite identificar conjuntos pequeños de enzimas

que podrían ser modificadas para alcanzar un dado objetivo, así como también la distribución óptima de los niveles de enzimas en el nuevo estado estacionario. El problema de optimización fue resuelto utilizando un método híbrido (determinístico-estocástico), teniendo como objetivos la maximización de la síntesis de serina y del flujo a través del sistema fosfotransferasa (PTS).

Chang & Sahinidis (2005) desarrollan un algoritmo de optimización global para el diseño de redes metabólicas, incorporando restricciones de estabilidad al problema haciendo uso del test de Routh-Hurwitz y del teorema de Vieta para determinar la estabilidad de la solución óptima. El algoritmo fue aplicado a una red metabólica simple que describe la síntesis bacteriana de triptófano (tres variables de estado) y a un modelo de fermentación anaeróbica en *Saccharomyces cerevisiae*, (catorce metabolitos, cinco variables de estado). Por otro lado, Gerhard *et al.* (2008) abordan el problema de optimización de estado estacionario que asegure la estabilidad de una red de biosíntesis de triptófano, utilizando el enfoque “Constructive nonlinear dynamics” (CNLD) propuesto por Mönningmann & Marquardt (2003). Más recientemente, Nikolaev (2010) reformula el problema de diseño (Nikolaev *et al.*, 2005) incluyendo restricciones que aseguran que la parte real de los autovalores del Jacobiano del sistema es negativa, para tal fin los autores calculan los autovalores del sistema linealizado. Este autor resuelve el problema con *Simulated Annealing*.

En este capítulo se presenta la formulación del problema de diseño óptimo de redes metabólicas estables y su resolución como un problema de optimización de autovalores (Blanco & Bandoni, 2007). El objetivo es diseñar redes metabólicas

que maximicen un objetivo biotecnológico, asegurando la estabilidad del estado estacionario resultante (Di Maggio *et al.*, 2009; 2010). El modelo estudiado incluye la vía Embden-Meyerhof-Parnas, la vía de las pentosas fosfato y el sistema fosfotransferasa de *Escherichia coli* K12 (Chassagnole *et al.*, 2002) y de las reacciones de fermentación.

El problema de optimización no lineal, correspondiente a la maximización de la producción de serina, en un primer caso, y etanol, en el segundo caso, sujeto a las ecuaciones de balance de masa en estado estacionario junto con las restricciones de estabilidad, se ha resuelto en un marco de optimización, con un algoritmo de Programación Cuadrática Sucesiva (Cuthrell & Biegler, 1985).

5.2. Métodos

5.2.1. Problema de diseño óptimo de redes metabólicas

El problema de diseño de redes metabólicas se formula como un problema de optimización donde las velocidades máximas de reacción (r_j^{max}) son las variables de optimización, ya que están asociados a las concentraciones de las enzimas correspondientes (e_j), según:

$$r_j^{max} = k_{cat,j} e_j \quad (5.1)$$

donde $k_{cat,j}$ es la constante catalítica.

En la formulación del problema de diseño óptimo de la red metabólica, se plantean los balances de masa para los metabolitos extra e intracelulares, en estado

estacionario (Ecs. 5.2 a 5.26), basados en el modelo dinámico descrito en el Capítulo 2:

$$D(C_{glc}^{alimentación} - C_{glc}^{extracelular}) - \frac{C_x r_{PTS}}{\rho_x} = 0 \quad (5.2)$$

$$r_{PTS} - r_{PGI} - r_{G6PDH} - r_{PGM} - \mu C_{g6p} = 0 \quad (5.3)$$

$$r_{PGI} - r_{PKF} + r_{TKb} + r_{TA} - 2r_{MurSynth} - \mu C_{f6p} = 0 \quad (5.4)$$

$$r_{PKF} - r_{ALDO} - \mu C_{f6p} = 0 \quad (5.5)$$

$$r_{ALDO} + r_{TIS} - r_{GAPDH} + r_{TKa} + r_{TKb} - r_{TA} + r_{TrpSynth} - \mu C_{gap} = 0 \quad (5.6)$$

$$r_{ALDO} - r_{TIS} - r_{G3PDH} - \mu C_{dhap} = 0 \quad (5.7)$$

$$r_{GAPDH} - r_{PGK} - \mu C_{pgp} = 0 \quad (5.8)$$

$$r_{PGK} - r_{PGGluMu} - r_{SerSynth} - \mu C_{3pg} = 0 \quad (5.9)$$

$$r_{PGluMu} - r_{ENO} - \mu C_{2pg} = 0 \quad (5.10)$$

$$r_{ENO} - r_{PK} - r_{PTS} - r_{PEPCxylase} - r_{DAHPS} - r_{Synth1} - \mu C_{pep} = 0 \quad (5.11)$$

$$r_{PK} + r_{PTS} - r_{PDH} - r_{Synth2} + r_{MetSynth} + r_{TrpSynth} - \mu C_{pyr} = 0 \quad (5.12)$$

$$r_{G6PDH} - r_{PGDH} - \mu C_{6pg} = 0 \quad (5.13)$$

$$r_{PGDH} - r_{Ru5P} - r_{R5PI} - \mu C_{ribu5p} = 0 \quad (5.14)$$

$$r_{Ru5P} - r_{TKa} - r_{TKb} - \mu C_{xyl5p} = 0 \quad (5.15)$$

$$r_{TKa} - r_{TA} - \mu C_{sed7p} = 0 \quad (5.16)$$

$$r_{R5PI} - r_{TKa} - r_{RPPK} - \mu C_{rib5p} = 0 \quad (5.17)$$

$$r_{TA} - r_{TKb} - r_{DAHPS} - \mu C_{e4p} = 0 \quad (5.18)$$

$$r_{PGM} - r_{GIPAT} - \mu C_{glp} = 0 \quad (5.19)$$

$$r_{ACALDH} - r_{ADH} - \mu C_{acal} = 0 \quad (5.20)$$

$$r_{PTA} - r_{ACK} - \mu C_{acp} = 0 \quad (5.21)$$

$$r_{ADH} \left(\frac{C_x}{\rho} \right) - DC_{etanol} = 0 \quad (5.22)$$

$$r_{PFL} \left(\frac{C_x}{\rho} \right) - DC_{formiato} = 0 \quad (5.23)$$

$$r_{LDH} \left(\frac{C_x}{\rho} \right) - DC_{lactato} = 0 \quad (5.24)$$

$$r_{ACK} \left(\frac{C_x}{\rho} \right) - DC_{acetato} = 0 \quad (5.25)$$

$$r_{SucSynth} \left(\frac{C_x}{\rho} \right) - DC_{succinato} = 0 \quad (5.26)$$

Según la ecuación (5.1) cambios en los valores de r_j^{max} inducen cambios en los niveles o concentraciones de enzimas, lo cual produce una modificación en la concentración total de proteínas.

Las proteínas ocupan entre 20-30 % del volumen celular. Para el caso de *Escherichia coli* la concentración de proteínas totales ronda los 230 mg/ml de volumen celular (Brown, 1991). El incremento en los niveles de proteínas puede ocasionar problemas de difusión de metabolitos como consecuencia del aumento de viscosidad del citoplasma, comprometiéndose la viabilidad celular. Por otra parte, si se exceden los límites de solubilidad de ciertas proteínas se podría ocasionar la precipitación de dichos polímeros, dando lugar a lo que se conoce como “cuerpos de inclusión” (Brown, 1991). Finalmente, un incremento en los niveles de proteínas requiere de un mayor consumo y/o una mayor tasa de síntesis de aminoácidos y de hidrólisis de ATP, lo que podría ocasionar estrés al

cultivo (Brown, 1991; Visser *et al.*, 2004). Los modelos cinéticos de redes metabólicas no contemplan los cambios fisiológicos anteriormente mencionados. Por esta razón, se considera de gran importancia incluir en el problema de diseño restricciones adicionales que aseguren que la concentración total de enzimas en el nuevo estado estacionario se mantenga cercana a la concentración del estado estacionario original. Esto puede lograrse incorporando la siguiente restricción:

$$\frac{1}{M} \sum_{j=1}^M \frac{r_j^{max}}{r_j^{max,0}} = 1 \quad (5.27)$$

donde M es el número total de enzimas involucradas en la red metabólica y $r_j^{max,0}$ es la velocidad máxima de reacción en el estado estacionario inicial. De esta manera, la ecuación (5.27) impone que un incremento en la concentración de una dada enzima es compensado por el decrecimiento en la concentración de las restantes. Esta hipótesis se confirmó con observaciones experimentales que han mostrado que la sobreexpresión de una dada enzima puede ocasionar un decrecimiento en la concentración de las restantes (Nikolaev, 2010). La ecuación (5.27) no contempla por sí misma los cambios coordinados en todas las enzimas, cuando se modifican solo L de las M enzimas ($L < M$). Para tener en cuenta estos cambios se introduce la siguiente ecuación:

$$\frac{r_{j_1}^{max}}{r_{j_1}^{max,0}} = \dots = \frac{r_{j_K}^{max}}{r_{j_K}^{max,0}} = \gamma \quad (5.28)$$

donde K representa el número de enzimas no modificadas ($K=M-L$); $j_1 \dots j_K$ son los índices correspondientes a las enzimas no modificadas y γ es un coeficiente de proporcionalidad.

Por otra parte, las células deben conservar la homeostasis, para lo cual tienden a mantener la concentración de sus metabolitos intracelulares en un nivel prácticamente constante. Así, si una red metabólica optimizada alcanza concentraciones de estado estacionario muy diferentes a la del estado estacionario original, es de esperarse que la expresión génica se vea influenciada de manera tal que el sistema vuelva al estado estacionario en que se encontraba. Estos efectos no pueden ser previstos por los modelos cinéticos ya que, usualmente, no describen los mecanismos de regulación genética (Visser *et al.*, 2004; Nikolaev *et al.*, 2005). Además, el incremento en la concentración de los metabolitos intracelulares puede ocasionar la inhibición de la actividad de ciertas enzimas o bien la degradación de los propios metabolitos (Visser *et al.*, 2004). Para determinar la región factible de la red metabólica de forma tal que se satisfaga la condición de homeostasis, se incluye la siguiente ecuación en la formulación del problema de optimización:

$$\frac{1}{N} \sum_{i=1}^N \frac{|C_i - C_i^0|}{C_i^0} \leq \delta \quad (5.29)$$

donde, δ , corresponde a la fracción máxima de cambio permitida para la concentración media de metabolitos intracelulares; C_i^0 es la concentración del metabolito i en el estado estacionario inicial y N es el número de metabolitos involucrados en la red metabólica.

El problema de diseño óptimo de la red metabólica se ha formulado en un primer caso en base al modelo del metabolismo central del carbono de *Escherichia coli* K12 (Ecs. 5.2 a 5.19) y luego al modelo extendido formulado para la misma cepa

(Ecs. 5.2 a 5.26), ambos descritos en el Capítulo 2 de esta tesis, teniendo como objetivo la maximización de la producción de serina y de etanol, respectivamente. Las concentraciones de los co-metabolitos involucrados en la red se han considerado constantes.

Las ecuaciones (5.27) y (5.28) se re-escribieron como las ecuaciones (5.30) y (5.31), respectivamente, para una mejor formulación del problema de optimización.

$$\frac{r_{j_1}^{max}}{r_{j_1}^{max,0}} + \dots + \frac{r_{j_L}^{max}}{r_{j_L}^{max,0}} + KY = M \quad (5.30)$$

$j_1 \dots j_L =$ índices de las enzimas modificadas

$$r_{j_s}^{max} = \gamma r_{j_s}^{max0} \quad s = 1, \dots, K \quad (5.31)$$

La ecuación (5.31) es válida para las s enzimas no modificadas.

El problema de diseño óptimo de la red metabólica queda finalmente formulado como:

$$\begin{aligned} \max_{r_{j_1}^{max}, \dots, r_{j_L}^{max}} \quad & r_m = r_m^{max} f(c, q) \quad m = \text{SerSynth o ADH} \\ \text{sujeto a} \quad & \sum_{j=1}^M S_{ij} r_j = 0 \quad i = 1, \dots, N \quad \text{a)} \\ & r_j = r_j^{max} f(c, q) \quad j = 1, \dots, M \quad \text{b)} \\ & \frac{r_{j_1}^{max}}{r_{j_1}^{max,0}} + \dots + \frac{r_{j_L}^{max}}{r_{j_L}^{max,0}} + KY = M \quad \text{c)} \\ & r_{j_s}^{max} = \gamma r_{j_s}^{max0} \quad s = 1, \dots, K \quad \text{d)} \\ & \frac{1}{N} \sum_{i=1}^N \frac{|C_i - C_i^0|}{C_i^0} \leq \delta \quad \text{e)} \end{aligned} \quad (5.32)$$

donde \mathbf{c} y \mathbf{q} son los vectores de concentración y parámetros del modelo, respectivamente; el primer conjunto de restricciones (5.32 a) corresponde a las ecuaciones (5.2) a (5.26) dependiendo el caso de estudio; el segundo set de restricciones (5.32 b) representa las ecuaciones (2.27) a (2.57) y (2.73) a (2.79); mientras que las restantes restricciones (5.32 c, d, e) corresponden a las ecuaciones (5.30), (5.31) y (5.29), respectivamente.

El problema presentado en la ecuación (5.32) es formulado para el diseño óptimo de la red metabólica de *Escherichia coli*, maximizando la producción de serina (aminoácido) y de etanol, respectivamente.

5.2.2. Problema de diseño óptimo de redes metabólicas bajo restricciones de estabilidad

Los modelos cinéticos de redes metabólicas están representados por un sistema de ecuaciones diferenciales-algebraicas no lineales (Capítulo 2), en los que las expresiones cinéticas son funciones altamente no lineales. Debido a ello, estos sistemas pueden exhibir multiplicidad de estados estacionarios, oscilaciones, bifurcación paramétrica de los estados estacionarios y comportamiento caótico, entre otros (Vital-Lopez *et al.*, 2006; Matallana *et al.*, 2009; Blanco & Bandoni, 2007).

En este capítulo se propone una metodología de optimización de autovalores (Blanco & Bandoni, 2007) para asegurar la estabilidad asintótica del punto óptimo obtenido en el problema de diseño. Se estudia el modelo del metabolismo central de *Escherichia coli* (Di Maggio *et al.*, 2009), así como también su versión

extendida que incluye las reacciones de fermentación (Di Maggio *et al.*, 2010), ambos formulados en el Capítulo 2.

5.2.2.1. Estabilidad de sistemas no lineales. Fundamentos

En esta sección se describen conceptos básicos de estabilidad de sistemas no lineales. Para una descripción más detallada se recomiendan los trabajos de Vidyasagar (1993), Kahlil (1998), Blanco & Bandoni (2003, 2007) y Matallana *et al.* (2009).

Sea el siguiente sistema dinámico no lineal (realizado un cambio de variables para examinar la estabilidad del sistema en el origen):

$$\begin{aligned} \frac{d\mathbf{c}}{dt} &= \mathbf{f}(\mathbf{c}, \mathbf{q}, \mathbf{d}) \\ \mathbf{c}(t_0) &= \mathbf{c}_0 \end{aligned} \tag{5.33}$$

donde \mathbf{c} es el vector de variables de estado, de dimensión n ; \mathbf{d} corresponde al vector de variables de diseño, de dimensión m . Generalmente, el vector de parámetros (\mathbf{q}) es conocido. Para un cierto valor de los parámetros \mathbf{q} (\mathbf{q}^*), las soluciones del sistema de ecuaciones (5.33) en estado estacionario (\mathbf{c}^*) son los puntos de equilibrio del sistema físico. Es importante mencionar que para los sistemas dinámicos no lineales las propiedades de estabilidad de los puntos de equilibrio son locales. Esto implica que para cierto punto de equilibrio asintóticamente estable, existe una región en el espacio de las variables de estado donde se originan las trayectorias que convergen a dicho punto. Dicha región se conoce como dominio de atracción.

El sistema (5.33) se puede aproximar por un modelo lineal alrededor de un punto de equilibrio del mismo:

$$\frac{dc}{dt} \approx A(c - c^*) \tag{5.34}$$

donde **A** es el Jacobiano del sistema (5.33) evaluado en c^* , (Ec. 5.35).

$$A = \frac{df}{dc} \tag{5.35}$$

El comportamiento dinámico del sistema (5.33) para un punto de equilibrio particular depende localmente del espectro de la matriz **A**, cuyos autovalores y autovectores se definen como:

$$Av = \lambda v \tag{5.36}$$

donde λ (escalar) y **v** (vector) son los autovalores y autovectores de **A**, respectivamente. La naturaleza de los puntos de equilibrio del sistema (5.33), en cuanto a sus propiedades de estabilidad y convergencia dinámica, es localmente determinada por el tipo (real o imaginario) y el signo de los autovalores de **A** (Tabla 5.1).

Tabla 5.1. Características de los autovalores y comportamiento del sistema

Autovalores		Comportamiento
Parte real	Parte imaginaria	
< 0	= 0	Decrecimiento exponencial a cero
> 0	= 0	Crecimiento exponencial
< 0	≠ 0	Oscilaciones amortiguadas
> 0	≠ 0	Oscilaciones no amortiguadas
= 0	≠ 0	Oscilaciones sostenidas

Un tipo importante de equilibrio es el localmente asintóticamente estable. Para que un punto de equilibrio sea localmente asintóticamente estable, todos los autovalores de la matriz **A** deben poseer partes reales negativas cuando son evaluados en dicho punto de equilibrio. Para asegurar estabilidad en el marco de un problema de diseño óptimo, dicha condición debería incluirse como restricción en el problema de optimización. Sin embargo, se puede demostrar que aún cuando los elementos de una matriz sean diferenciables, la expresión analítica de sus autovalores puede no serlo (Overton, 1992).

En este caso, la inclusión de las expresiones analíticas de los autovalores como restricciones del problema de optimización no permitiría la aplicación de algoritmos basados en gradiente.

Por otra parte, tampoco es posible determinar expresiones analíticas para el cálculo de los autovalores en matrices de dimensiones mayores a cuatro (Kokossis & Floudas, 1994), lo que impide su inclusión directa como restricciones en problemas de optimización de escala mediana a grande.

Para un tratamiento general del problema de diseño bajo condiciones de estabilidad, es necesario desarrollar formulaciones alternativas, equivalentes a la condición de negatividad sobre la parte real de los autovalores de la matriz **A**.

Blanco & Bandoni (2007), por ejemplo, aplican conceptos derivados de la teoría de estabilidad de Lyapunov (Kahlil, 1998).

Según Lyapunov, la estabilidad de un sistema se puede estudiar por medio de una función de energía **V(c)**, denominada función de Lyapunov, definida positiva en un entorno del punto de equilibrio (5.37 a). Si la derivada temporal de dicha función

de energía es definida negativa en el entorno estudiado (5.37 b), entonces es posible asegurar que el punto de equilibrio analizado es asintóticamente estable.

$$\begin{aligned} \mathbf{V}(\mathbf{c}) > 0, \mathbf{V}(\mathbf{0}) = 0 & \quad \text{a)} \\ \frac{d\mathbf{V}}{dt} = \frac{\partial \mathbf{V}}{\partial \mathbf{c}} \frac{d\mathbf{c}}{dt} \leq 0 & \quad \text{b)} \end{aligned} \quad (5.37)$$

No hay un procedimiento general para determinar funciones de Lyapunov para sistemas no lineales en general. Para sistemas lineales (o linealizados), autónomos, es posible obtener una función de Lyapunov por medio de la resolución de un sistema de ecuaciones lineales algebraicas. El procedimiento se deriva a continuación.

Se puede demostrar que, para un sistema lineal, la siguiente función cuadrática constituye una función de Lyapunov:

$$\mathbf{V}(\mathbf{c}) = \mathbf{c}^T \mathbf{P} \mathbf{c} \quad (5.38)$$

donde \mathbf{P} es una matriz simétrica y definida positiva. Reemplazando (5.34) en la ecuación (5.38):

$$\mathbf{V}(\mathbf{c}) = \mathbf{c}^T (\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A}) \mathbf{c} \quad (5.39)$$

De la condición (5.37b) es posible asegurar que el equilibrio correspondiente al sistema (5.34) es asintóticamente estable si se satisface la siguiente condición:

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} < 0 \quad (5.40)$$

o, en forma equivalente:

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{Q} \quad (5.41)$$

donde \mathbf{Q} es simétrica y definida positiva, como por ejemplo, la matriz identidad:

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} + \mathbf{I} = \mathbf{0} \tag{5.42}$$

La ecuación (5.42) se conoce como ecuación de Lyapunov.

Blanco & Bandoni (2003, 2007) proponen formular el problema de diseño bajo condiciones de estabilidad incorporando la ecuación 5.42 y junto con la condición de definición positiva sobre \mathbf{P} en el problema original. Asimismo, estos autores formulan la condición de definida positiva en base al criterio de Sylvester de los menores de una matriz. Dicho criterio establece que la condición necesaria y suficiente para que una matriz real simétrica sea definida positiva es que sus menores principales sean positivos. Blanco & Bandoni (2007) aplican esta condición sobre la inversa de la matriz \mathbf{P} (y no directamente sobre \mathbf{P}), obteniendo una formulación más robusta del problema de optimización. Di Maggio *et al.* (2009, 2010) emplean este enfoque para formular el problema de diseño óptimo de redes metabólicas estables.

5.2.2.2 Formulación del problema de diseño óptimo de redes metabólicas bajo restricciones de estabilidad

En esta sección se presenta una formulación del problema de diseño óptimo de la red metabólica de *Escherichia coli* que incluye restricciones para asegurar la estabilidad asintótica del punto de equilibrio. La metodología aplicada, basada en la teoría de estabilidad de Lyapunov y la formulación propuesta por Blanco & Bandoni (2007), incorpora a la formulación 5.32 la ecuación de Lyapunov (5.42) y

la condición de definición positiva sobre la matriz de Lyapunov (\mathbf{P}), a través de la condición de Sylvester sobre sus menores principales:

$$\begin{aligned}
 & \max_{r_{j_1}^{\max} \dots r_{j_L}^{\max}} \quad r_m = r_m^{\max} f(\mathbf{c}, \mathbf{q}) \\
 & \text{sujeto a} \\
 & \mathbf{A}^T(\mathbf{c})\mathbf{P} + \mathbf{P}\mathbf{A}(\mathbf{c}) + \mathbf{I} = \mathbf{0} \quad \text{e)} \\
 & \det(\mathbf{P}_i^{-1}) > \xi \quad i = 1, \dots, N \quad \text{f)} \\
 & \sum_{j=1}^M \mathbf{S}_{ij} \mathbf{r}_j = \mathbf{0} \quad i = 1, \dots, N \\
 & r_j = r_j^{\max} f(\mathbf{c}, \mathbf{q}) \quad j = 1, \dots, M \\
 & \frac{r_{j_1}^{\max}}{r_{j_1}^{\max,0}} + \dots + \frac{r_{j_L}^{\max}}{r_{j_L}^{\max,0}} + K\gamma = M \\
 & r_{j_s}^{\max} = \gamma r_{j_s}^{\max,0} \quad s = 1, \dots, K \\
 & \frac{1}{N} \sum_{i=1}^N \frac{|C_i - C_i^0|}{C_i^0} \leq \delta
 \end{aligned} \tag{5.43}$$

donde \mathbf{c} y \mathbf{q} son los vectores de concentración y parámetros del modelo, respectivamente; \mathbf{A} es el Jacobiano del sistema dinámico definido por los balances de masa; la restricción (5.43e) corresponde a la ecuación de Lyapunov. La restricción (5.43f) corresponde a la condición de Sylvester (menores principales de \mathbf{A} positivos), siendo ξ una constante positiva pequeña. El valor de dicha constante permite modificar en forma indirecta el espectro del Jacobiano y consecuentemente la evolución dinámica del sistema hacia el equilibrio. Las restantes restricciones corresponden al problema de diseño óptimo de la red metabólica, descrito en la Sección 5.2.1 (Ec. 5.32).

La solución del problema formulado según la ecuación (5.43) determina una red metabólica de *Escherichia coli* para maximizar la producción de un objetivo biotecnológico (producción de serina o etanol, en los casos de estudio), asegurando la estabilidad asintótica del estado estacionario óptimo.

5.3. Resultados y discusión

En esta sección se presentan los resultados obtenidos para el problema de diseño óptimo de una red metabólica de *E. coli*. En primer lugar (Caso 1), el objetivo es maximizar la producción de un aminoácido de interés, serina. En segundo lugar (Caso 2), se maximiza la producción de etanol. Se resuelven los respectivos problemas sin restricciones de estabilidad (Ec. 5.32) y considerando restricciones que aseguran estabilidad de las rutas metabólicas (Ec. 5.43) y se comparan los resultados obtenidos. Los problemas de Programación no Lineal fueron implementados en Fortran 90 y resueltos con un algoritmo de Programación Cuadrática Sucesiva (Cuthrell & Biegler, 1985).

Caso 1. Maximización de la producción de serina

En base a los resultados del análisis de sensibilidad global sobre el modelo dinámico de *Escherichia coli* (Capítulo 3), se seleccionaron las siguientes velocidades máximas de reacción como variables de diseño ($L=3$): velocidad máxima de reacción de síntesis de serina ($r_{SerSynth}^{max}$), velocidad máxima de

reacción de la enzima glucosa-6-fosfato isomerasa (r_{PGI}^{max}), y velocidad máxima de producción de corismato y mureina (r_{synth1}^{max}) (ver Figura 2.1).

En este caso el problema de diseño, tanto con restricciones de estabilidad como sin ellas, se ha formulado para el modelo cinético en estado estacionario de la vía Embden-Meyerhof-Parnas, el sistema fosfotransferasa y la vía de las pentosas fosfato (Di Maggio *et al.*, 2009) y comprende dieciocho balances de masa en estado estacionario (Ecs. 5.2 a 5.19), treinta ecuaciones cinéticas asociadas a cada una de las enzimas (Ecs. 2.27 a 2.57) y veintinueve restricciones adicionales correspondientes a las ecuaciones (5.29), (5.30) y (5.31), descritas en la Sección 5.2.1. Cuando se formula el problema de diseño bajo restricciones de estabilidad (Ec. 5.43) se incorporan además, dieciocho restricciones de igualdad y dieciocho de desigualdad, que corresponden a la ecuación de Lyapunov (Ec. 5.42) y a las ecuaciones que aseguran la positividad de los menores principales de \mathbf{P} , respectivamente.

Los resultados obtenidos para el diseño de la red metabólica de *Escherichia coli* cuando el objetivo es la maximización de la producción de serina se muestran en las Figuras 5.1, 5.2 y 5.3. En ambas formulaciones (Ecs. 5.32 y 5.43) la restricción correspondiente a la condición de mantenimiento de la homeostasis es acotada con un valor de $\delta=0,3$. La Tabla 5.2 muestra los valores de las concentraciones y de los flujos de las reacciones enzimáticas para el estado estacionario de referencia.

En la Tabla 5.3 se presentan los valores de las variables de optimización y la función objetivo en el estado estacionario de referencia y en los estados estacionarios alcanzados luego de la optimización.

Tabla 5.2. Flujos y concentraciones de metabolitos en el estado estacionario de referencia

Metabolito	Concentración (mM)	Reacción	Flujo (mM/seg)
Glucosa	0,056	PTS	0,153
g6p	3,48	PGI	0,044
f6p	0,6	PFK	0,109
fdp	0,272	ALDO	0,109
gap	0,218	TIS	0,106
dhap	0,167	G3PDH	0,002
pgp	0,008	GAPDH	0,245
2pg	2,13	PGK	0,245
3pg	0,399	PGM	0,002
pep	2,67	ENO	0,232
pyr	2,67	PepCxylase	0,038
6pg	0,808	PK	0,029
ribu5p	0,111	PDH	0,143
rib5p	0,398	Synth1	0,011
xyl5p	0,138	Synth2	0,041
sed7p	0,276	G6PDH	0,107
e4p	0,098	PGDH	0,107
g1p	0,653	Ru5p	0,064
atp	4,27	R5PI	0,043
amp	0,595	Tka	0,035
adp	0,955	TKb	0,029
nadp ⁺	0,195	TA	0,035
nadph	0,062	RPPK	0,008
nad ⁺	1,47	DAHPS	0,005
nadh	0,1	PGlumu	0,232
		G1PAT	0,002
		MurSynth	0,002
		TrpSynth	0,005
		SerSynth	0,014
		MetSynth	0,002
		PFL	7E-06
		LDH	0,002
		PTA	0,006
		ACALDH	0,003
		ADH	0,005
		ACK	0,011

Los resultados numéricos, obtenidos al resolver el problema formulado en la ecuación (5.32), muestran que la producción de serina podría ser incrementada desde 0,014 mM/s, que corresponde a su valor en el estado estacionario de referencia, hasta 0,094 mM/s cuando la velocidad máxima de síntesis de serina ($r_{SerSynth}^{max}$) llega a su cota superior ($=10 \cdot r_{SerSynth}^{max,0}$). Analizando los autovalores de la matriz $\mathbf{A}(\mathbf{c})$, éstos muestran que la red metabólica para *Escherichia coli* es estable, siendo la parte real del mayor autovalor: $-3,7E-4$. Si bien la red es estable, se considera que al estar el autovalor tan cercano al eje imaginario, cualquier perturbación lo podría sacar de la zona de estabilidad asintótica.

Esta observación está en concordancia con Vital-Lopez *et al.* (2006), quienes determinaron que existe un punto de bifurcación en el comportamiento del sistema cuando la velocidad máxima de síntesis de serina alcanza un valor igual a $9,8 r_{SerSynth}^{max,0}$, volviéndose inestable el sistema cuando $r_{SerSynth}^{max} > 9,8 r_{SerSynth}^{max,0}$.

Por esa razón, se ha resuelto el problema con la formulación (5.43) que, además de asegurar la estabilidad asintótica del equilibrio, permite modificar la posición de los autovalores de la matriz $\mathbf{A}(\mathbf{c})$ a través del parámetro ξ . Con este objetivo se han realizado diferentes pruebas con el valor del parámetro ξ . Para $\xi = 1E-4$, la parte real del mayor autovalor es $-1,5E-3$. En este caso, la producción de serina se pudo incrementar en un 330%, desde su valor de estado estacionario de referencia (0,014 mM/s) a 0,059 mM/s. Se puede observar que se logra mejorar el margen de estabilidad del sistema a expensas de una disminución en la función objetivo. Asimismo, en la Tabla 5.3 se muestra que el valor optimizado de la

velocidad máxima de producción de serina no se encuentra en su cota máxima, como en el caso anterior.

En la Figura 5.1 se observa que los mayores incrementos en las concentraciones respecto de su estado de referencia, se producen para g6p, f6p y g1p, siendo de alrededor del 49% para el caso del diseño sin restricciones de estabilidad y de alrededor del 23% para el caso del diseño estable. Asimismo, la mayor variación ocurre para la concentración de fosfoenolpiruvato (pep).

Tabla 5.3. Valores de las variables de optimización y de la función objetivo (Caso 1)

Variables de optimización (mM/seg)	Estado de referencia	Diseño óptimo	Diseño óptimo bajo restricciones de estabilidad
r_{PGI}^{max}	495,870	442,316	495,842
$r_{SerSynth}^{max}$	0,020	0,200	0,099
r_{Synth1}^{max}	0,015	0,000	0,000
$r_{SerSynth}$	0,014	0,094	0,059

Es importante destacar que para el estado estacionario alcanzado cuando se incorporan las restricciones de estabilidad al problema diseño, los aumentos y disminuciones de las concentraciones (respecto del estado estacionario de referencia) son menores que para el caso de diseño sin restricciones de estabilidad (Figura 5.1). Este fenómeno constituye una manifestación del conflicto entre los objetivos “económico” y “operativo” que verifica la mayoría de los procesos y sistemas (Blanco & Bandoni, 2003).

El cambio en los flujos optimizados respecto de los del estado de referencia se puede observar en la Figura 5.2. La velocidad de producción de serina ($r_{SerSynth}$) aumenta con un factor de 6,8 para el caso del diseño sin restricciones de

estabilidad, aumentando con un factor de 4,3 cuando se formula el problema de diseño estable.

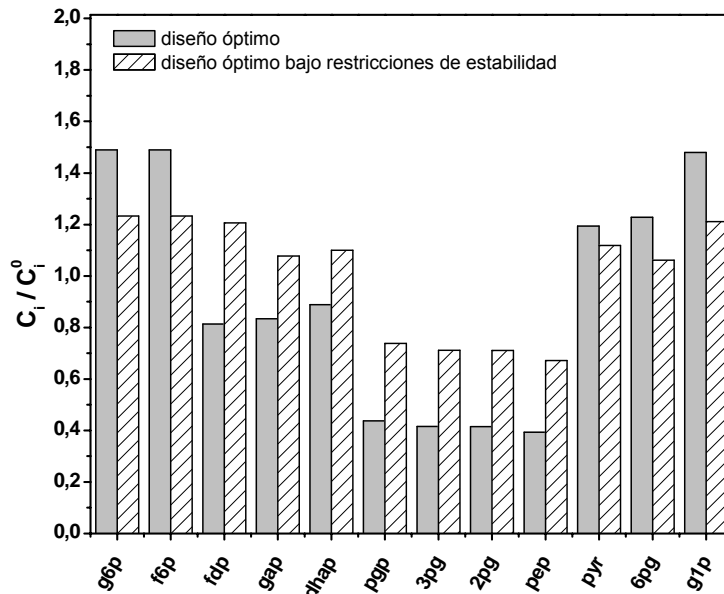


Figura 5.1. Concentraciones óptimas para los diseños con y sin restricciones de estabilidad para la maximización de la producción de serina

La maximización de la velocidad de producción de serina hace que los flujos de las enzimas glucolíticas corriente abajo de ésta, se vean incrementados respecto de los flujos alcanzados en el estado estacionario inicial (Figura 5.2).

Los resultados muestran que para ambos casos el óptimo en la producción de serina implica que la velocidad de producción de corismato y mureina disminuya al mínimo o se anule, ya que $r_{Synth1}^{max} = 0$. Esto implicaría, en términos de Ingeniería Genética, la delección de dicha reacción.

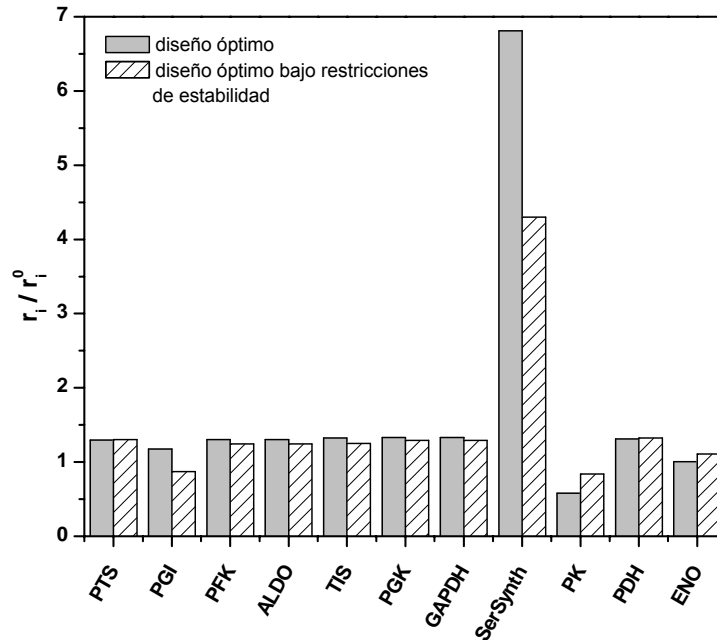


Figura 5.2. Flujos metabólicos óptimos para los diseños con y sin restricciones de estabilidad en la maximización de la producción de serina

La Figura 5.3 muestra la distribución óptima de flujos en la red metabólica de *Escherichia coli* para el estado estacionario de referencia y el estado estacionario alcanzado cuando se formula el problema de diseño bajo restricciones de estabilidad. Se puede notar que el flujo a través de la vía de las pentosas es superior en la red metabólica optimizada, sin embargo el flujo de carbono destinado a la síntesis de aminoácidos aromáticos y de nucleótidos en la red optimizada se mantiene cercano a los valores de referencia. Esto implica que la mayor parte del flujo de carbono es re-circulado a la vía glucolítica produciéndose un aumento del pool de fructosa-6-fosfato (f6p) y gliceraldehído-3-fosfato (gap) por la reacción catalizada por la enzima transcetolasa b (TKb) (Figura 2.1). El

incremento en las concentraciones de f6p y gap favorece el flujo en la vía glucolítica, siendo destinado a la síntesis de serina.

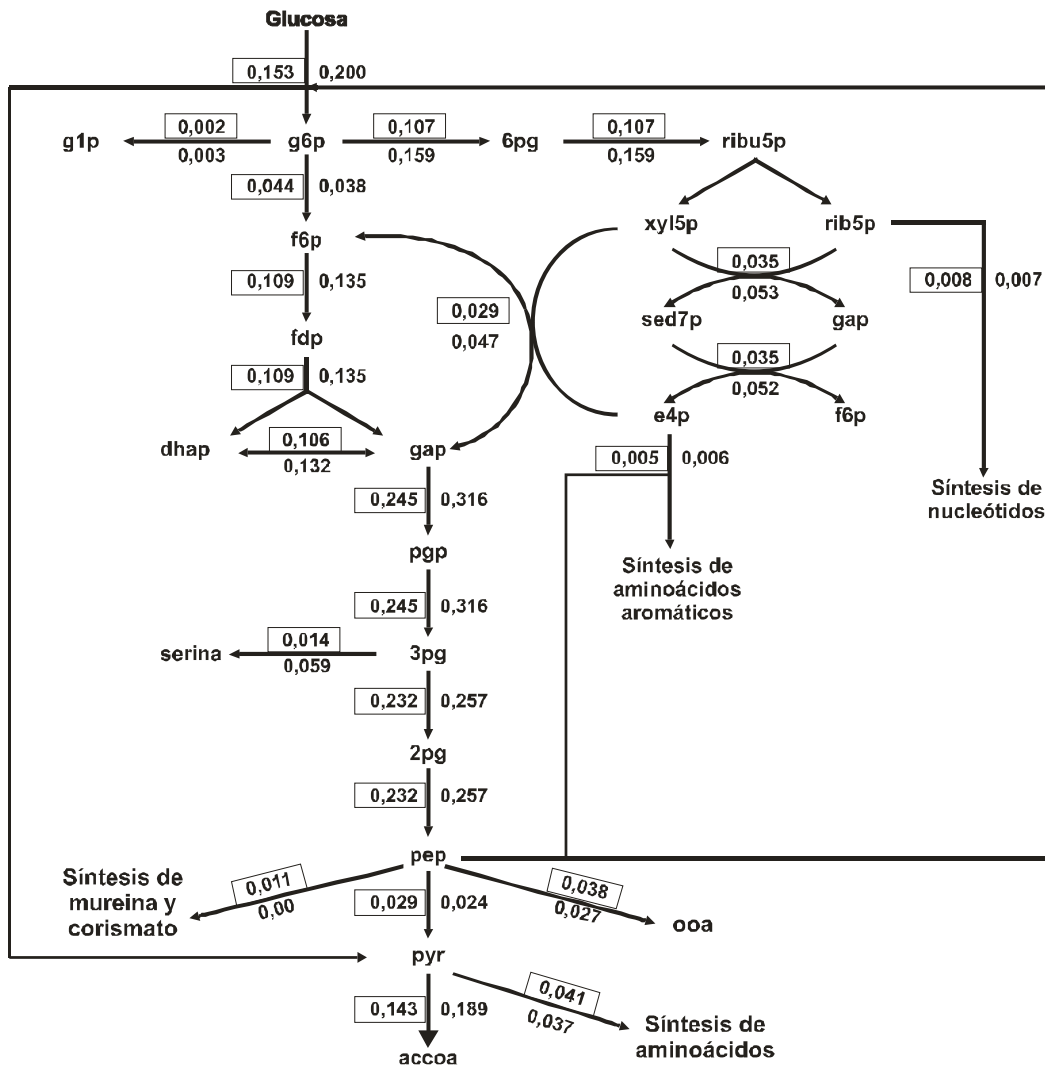


Figura 5.3. Distribución óptima de flujos para la maximización de la producción de serina en estado de referencia (□) y el estado estacionario estable

Caso 2. Maximización de la producción de etanol

Las variables de diseño en el problema de diseño óptimo de la red metabólica para maximizar la producción de etanol son las velocidades máximas de reacción corespondientes a la enzima fosfofructoquinasa (r_{PFK}^{max}), piruvatoquinasa (r_{PK}^{max}), alcohol deshidrogenasa (r_{ADH}^{max}) y velocidad máxima de producción de corismato y mureina (r_{Synth1}^{max}). Las enzimas PFK y PK corresponden a la vía glucolítica, mientras que la enzima ADH es una reacción de fermentación (Figura 2.1).

El problema de diseño, tanto con restricciones de estabilidad como sin ellas, se ha formulado para el modelo cinético extendido en estado estacionario de la vía Embden-Meyerhof-Parnas, el sistema fosfotransferasa, la vía de las pentosas fosfato (Di Maggio *et al.*, 2009) y las reacciones de fermentación de *Escherichia coli* (Di Maggio *et al.*, 2010). Las restricciones del problema de optimización incluyen veinticinco balances de masa (Ecs. 5.2 a 5.26), treinta y ocho ecuaciones cinéticas asociadas a cada una de las enzimas (Ecs. 2.27 a 2.57 y 2.73 a 2.79) y treinta y seis restricciones adicionales ($K+2$) correspondientes a las ecuaciones (5.29), (5.30) y (5.31); en este caso la ecuación (5.12) que corresponde al balance de masa de piruvato es reemplazada por la ecuación (2.65), formulada en estado estacionario. Cuando se resuelve el problema formulado según la ecuación (5.43), se incorporan al set de restricciones veinticinco restricciones de igualdad y veinticinco de desigualdad, correspondientes a la ecuación de Lyapunov (Ec. 5.42) y a las ecuaciones que imponen que los menores principales de \mathbf{P} sean positivos, respectivamente.

Los resultados obtenidos para el diseño de la red metabólica de *Escherichia coli* cuando se maximiza la velocidad de producción de etanol se muestran en las Figuras 5.4, 5.5 y 5.6. En ambas formulaciones (Ecs. 5.32 y 5.43), la variación permitida en el cambio medio de las concentraciones de los metabolitos es del 30% ($\delta=0,3$). La Tabla 5.4 muestra los valores de las concentraciones y de los flujos de las reacciones enzimáticas para el estado estacionario de referencia y en los diseños óptimos determinados. Los resultados numéricos muestran que la velocidad de producción de etanol se puede incrementar en un 38% respecto del estado estacionario de referencia (Tabla 5.4 y Figura 5.5). Al analizar los autovalores de la matriz $A(c)$ se observa que la parte real de algunos autovalores es positiva, por lo que se puede concluir que el nuevo estado estacionario alcanzado para la red metabólica de *Escherichia coli*, no es estable.

Tabla 5.4. Valores de las variables de optimización y de la función objetivo (Caso 2)

Variables de optimización (mM/seg)	Estado de referencia	Diseño óptimo	Diseño óptimo bajo restricciones de estabilidad
r_{PFK}^{max}	0,12	0,177	0,155
r_{PK}^{max}	0,047	0,278	0,117
r_{Synth1}^{max}	0,015	0,023	0,001
r_{ADH}^{max}	0,5	3,934	1,316
r_{ADH}	0,005	0,007	0,006

Al formular el problema de diseño bajo restricciones de estabilidad (Ec. 5.43), se observa que, para un valor de $\xi=1E-5$ la parte real del mayor autovalor de la matriz $A(c)$ es $-2,1E-3$. En este caso, la producción de etanol se pudo incrementar un 36%, desde su valor de estado estacionario de referencia, 0,005 mM/s a 0,006

mM/s (Tabla 5.4). Se puede observar que se logra mejorar la estabilidad del sistema a expensas de una disminución de la producción de etanol.

En cuanto a las concentraciones de los metabolitos en los nuevos estados estacionarios (Figura 5.4), se puede mencionar que los mayores cambios se presentan para las concentraciones de piruvato (pyr), ácido láctico, ácido fórmico y etanol (156%, 125%, 57% y 21%, respectivamente) para el primer caso. Por otra parte, cuando se resuelve el problema de diseño bajo restricciones de estabilidad, el mayor incremento (84%) se observa en la concentración de fdp (fructosa-1,6-difosfato). En este caso, la concentración de los intermediarios glucolíticos es mayor en el estado estacionario estable, excepto para el piruvato (pyr) (Figura 5.4) que es el precursor de la mayor parte de los productos de fermentación. En la Figura 5.5 se observa que el mayor incremento en los flujos se produce para la reacción de la enzima piruvato quinasa (PK), tanto para el diseño estable como inestable. Esto puede atribuirse, en el caso del estado estacionario estable, al efecto conjunto del aumento en el valor de la velocidad máxima de dicha reacción (Tabla 5.4) y del aumento en la concentración de fructosa-1,6-difosfato, que es un activador de la actividad de la enzima piruvato quinasa (Capítulo 2).

Los resultados del problema de diseño bajo consideraciones de estabilidad muestran que los flujos de las enzimas glucolíticas así como los correspondientes a las reacciones que conducen a la producción de etanol se ven incrementados respecto del estado estacionario de referencia, lo que conduce a un mayor flujo de carbono destinado a la síntesis del alcohol de interés (Figura 5.5).

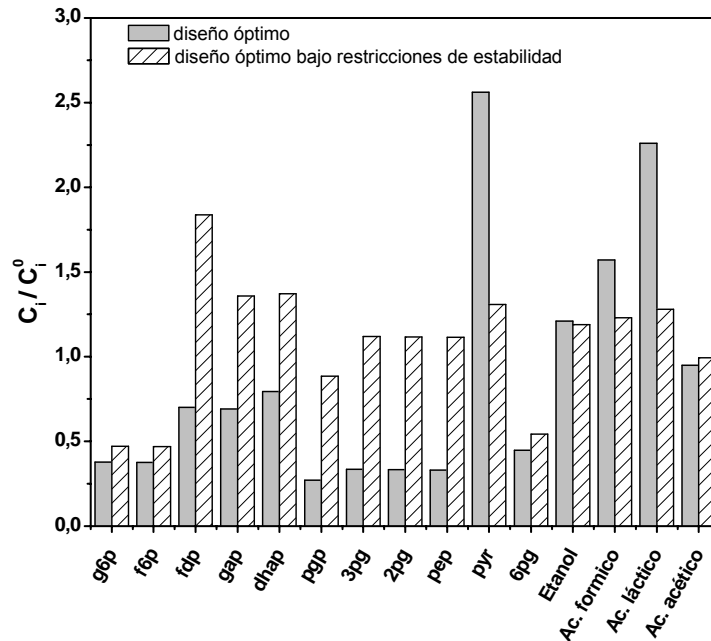


Figura 5.4. Concentraciones óptimas para los diseños con y sin restricciones de estabilidad para la maximización de la producción de etanol

La Figura 5.6 muestra la distribución óptima de flujos en el estado estacionario estable y de referencia. Se observa, como se mencionó previamente, que el flujo de carbono en la vía glucolítica aumenta cuando se maximiza la producción de etanol, no ocurriendo lo mismo con el flujo de carbono a través de la vía de las pentosas fosfato. Con respecto a los flujos alrededor de acetil-coenzimaA (accoa), en el estado estacionario estable, el flujo de la enzima ACALDH (acetaldehído deshidrogenasa) que conduce a la síntesis de etanol se incrementa, mientras que el flujo de la enzima PTA (fosfotransacetilasa) que conduce a la producción de ácido acético disminuye. Por otra parte se observa que solo una fracción del flujo de carbono aportado por la enzima PDH (que cataliza la reacción de piruvato a accoa) es destinado por las células a las reacciones de fermentación. Esto puede

atribuirse a que debido a que el modelo de la red metabólica está formulado y calibrado para condiciones de crecimiento aeróbica, parte del flujo a través de la enzima piruvato deshidrogenasa (PDH) es dirigido hacia el ciclo de Krebs.

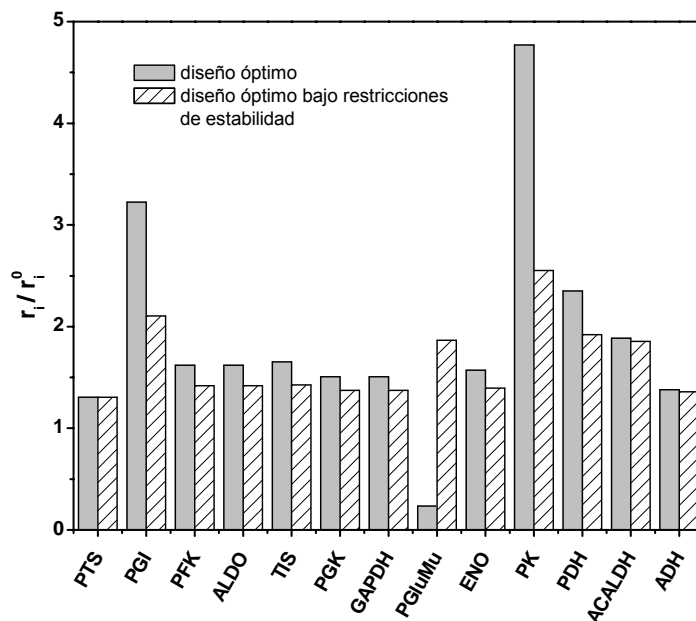


Figura 5.5. Flujos metabólicos óptimos para los diseños con y sin restricciones de estabilidad para la maximización de la producción de etanol

Como se mencionó previamente, la Figura 5.4 muestra que se produce un incremento en la concentración de etanol en ambos casos. Partiendo de una concentración inicial igual a la obtenida experimentalmente en la presente tesis (Capítulo 4), la concentración de etanol asciende a un valor 3,63 mM cuando se resuelve el problema de optimización sin restricciones de estabilidad (Ec. 5.32), mientras que cuando se incluyen dichas restricciones al problema de optimización (Ec. 5.43) la concentración de etanol alcanza un valor de 3,57 mM. Estos niveles de etanol se obtienen cuando se amplifican los flujos de las enzimas acetaldehído

deshidrogenasa (ACALDH) y alcohol deshidrogenasa (ADH), respecto del estado de referencia, como se observa en la Figura 5.5.

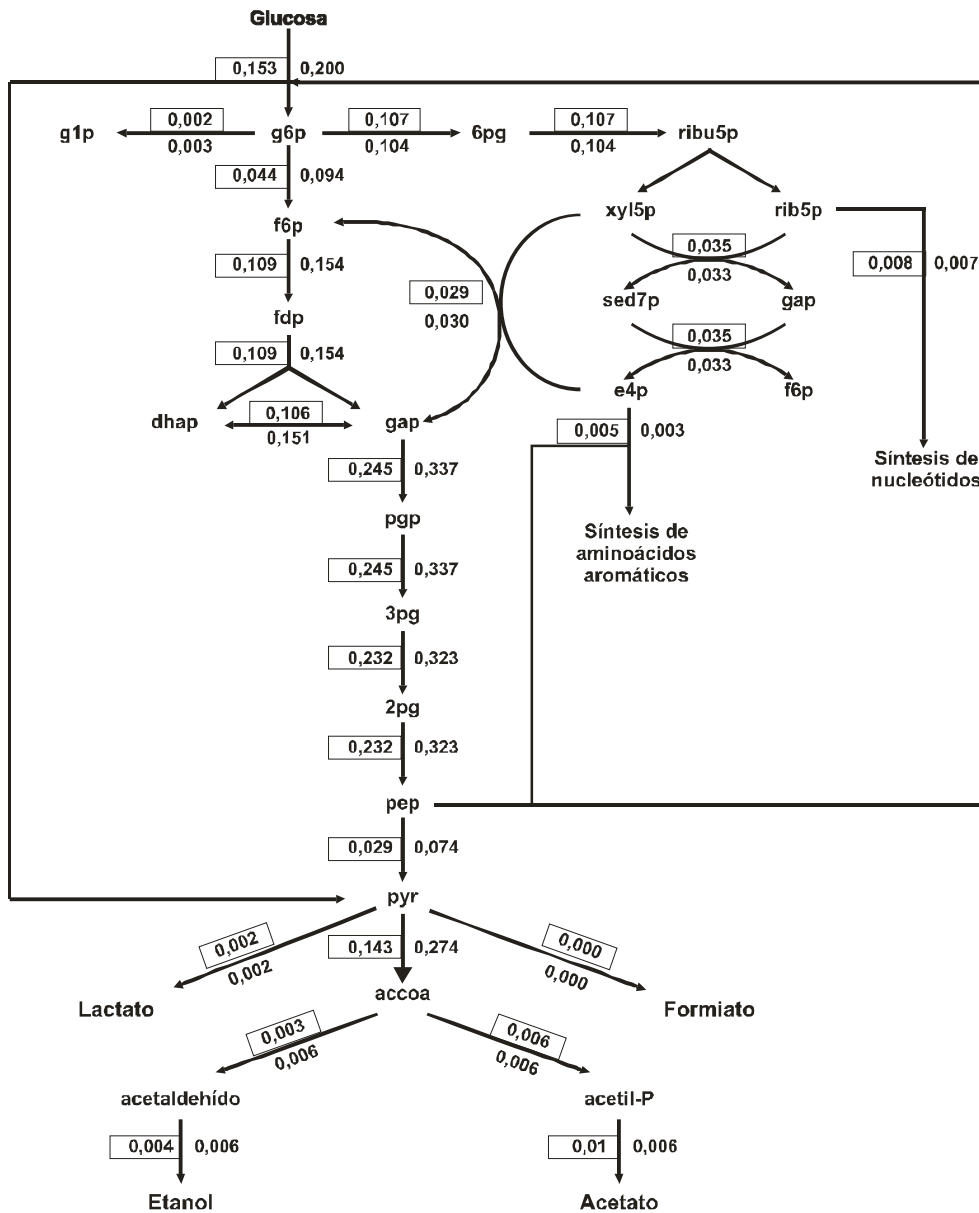


Figura 5.6. Distribución óptima de flujos para la maximización de la producción de etanol en el estado de referencia (□) y en el estado estacionario estable

Resultados reportados en Ingram & Conway (1988), muestran que la cepa recombinante de *Escherichia coli*, obtenida por inserción de los genes heterólogos (provenientes de *Z. mobilis*) que codifican las enzimas piruvato descarboxilasa (PDC) y alcohol deshidrogenasa (ADH), logra una producción de etanol de 337 mM, notablemente mayor.

Los resultados obtenidos en este capítulo confirman que aún amplificando los flujos de las enzimas naturales de formación de etanol (ACALDH y ADH) en *Escherichia coli*, no se obtienen importantes aumentos en la producción de etanol, lo cual justifica la inserción de nuevas rutas metabólicas (PDC y ADH).

5.4. Conclusiones

En este capítulo se han formulado problemas de programación no lineal como problemas de diseño óptimo para modelos cinéticos mecanísticos de gran escala de la red metabólica de *Escherichia coli*, considerando la incorporación de restricciones de estabilidad. Se ha mostrado, que dicho enfoque permite diseñar una red metabólica que maximiza un objetivo biotecnológico (producción de un aminoácido o de etanol), asegurando la estabilidad del sistema en los niveles de diseño tanto de las enzimas como de las concentraciones de los metabolitos.

Los resultados numéricos indican la necesidad de incorporar restricciones de estabilidad al problema de diseño, ya que de otra manera se podrían determinar estados estacionarios inestables que lleven a las células a una situación de vulnerabilidad a perturbaciones externas. Para el caso de la formulación del

problema de diseño cuando se pretende maximizar la velocidad de producción de serina, los resultados han mostrado un amplio incremento en la producción de dicho aminoácido, sugiriéndose la delección de la reacción de síntesis de corismato y mureína para alcanzar el objetivo. En este caso, se ha observado que el flujo a través de la vía de las pentosas fosfato es dirigido a la glucólisis incrementando las concentraciones de fructosa-6-fosfato y gliceraldehído-3-fosfato, lo cual produce un aumento del flujo de carbono en la vía glucolítica que, a su vez, contribuye a incrementar la producción de serina.

Por otra parte cuando se maximiza la velocidad de producción de etanol, los resultados han mostrado que se requiere una amplificación en los flujos de las enzimas glucolíticas para alcanzar dicho objetivo, aunque se está lejos aún de la producción de las cepas recombinantes a las que se insertan genes heterólogos de bacterias etanologénicas (por ejemplo *Z. mobilis*).

Finalmente, se puede concluir que el enfoque de diseño bajo restricciones de estabilidad empleado en esta tesis constituye una herramienta eficiente y confiable para el diseño de redes metabólicas estables y de gran tamaño.

CAPÍTULO 6

CONCLUSIONES Y TRABAJOS FUTUROS

6.1. Conclusiones generales

En esta tesis se han realizado contribuciones para la optimización de redes metabólicas, mediante el desarrollo de modelos matemáticos basados en primeros principios y técnicas avanzadas de optimización. Se trata de modelos cinéticos y dinámicos, altamente no lineales, con elevado número de parámetros, que describen las principales reacciones enzimáticas que se llevan a cabo en el interior de un microorganismo. Los problemas de estimación de parámetros para estos modelos se han formulado en un marco de optimización con función objetivo de máxima verosimilitud, sujeto a un sistema de ecuaciones diferenciales algebraicas. Los modelos desarrollados, ajustados y validados con datos de bibliografía, han permitido en primer lugar la determinación de parámetros de cinéticas enzimáticas *in vivo*, lo cual constituye un importante aporte ya que se dispone normalmente de datos *in vitro* de la cinética enzimática y no de la enzima funcionando dentro de la red metabólica. Por otro lado, los modelos desarrollados en esta tesis se han empleado en el diseño óptimo de cepas para la sobreproducción de productos de interés biotecnológico, como pueden ser

aminoácidos o bioetanol. En este sentido, se ha presentado una metodología, en un marco de optimización restringida, para el diseño óptimo para estabilidad de las cepas resultantes.

6.2 Conclusiones particulares

A continuación se enumeran las principales contribuciones realizadas en la presente tesis.

En el capítulo 2 se formula un modelo dinámico cinético del metabolismo central del carbono, describiendo la glucólisis, la ruta de las pentosas fosfato, el sistema fosfotranferasa, reacciones anapleróticas y de biosíntesis en *Escherichia coli* K12. Asimismo, se extiende el modelo para incluir las reacciones de fermentación del metabolismo. El modelo extendido posee 25 variables de estado, correspondientes a las concentraciones de metabolitos intra y extracelulares; 44 variables algebraicas, que representan las velocidades de reacción y las concentraciones de los co-metabolitos, y 166 parámetros cinéticos. Es importante destacar que las cinéticas de las reacciones enzimáticas se han obtenido, en general, a partir de estudios realizados *in vitro*. El desarrollo de modelos cinéticos de redes metabólicas, como el descrito en esta tesis, permite la determinación de las principales cinéticas *in vivo*.

En el capítulo 3, se implementa una técnica de análisis de sensibilidad global (ASG) para estudiar la sensibilidad paramétrica del modelo cinético del

metabolismo central del carbono de *Escherichia coli*. La metodología ASG implementada se basa en la descomposición de la varianza de las variables de estado y el empleo de varianzas condicionales para determinar índices de sensibilidad en cada instante de tiempo. Las varianzas incondicionales y condicionales se calculan mediante simulaciones Monte Carlo en un entorno gPROMS. Se determinan y analizan perfiles de los índices de sensibilidad para las variables de estado diferenciales y algebraicas. Mediante la observación de los índices de primer orden (Si) se identifican parámetros (12) que controlan el comportamiento dinámico del modelo. Asimismo, la metodología ASG empleada permite la determinación de parámetros no significativos a lo largo de todo el horizonte de tiempo. Los resultados obtenidos del análisis de sensibilidad global llevado a cabo son de fundamental importancia para determinar los parámetros más influyentes.

En el capítulo 4, se lleva a cabo estimación de parámetros empleando técnicas avanzadas de optimización dinámica y se presenta el trabajo experimental realizado en esta tesis. En primer lugar, y como se indicó anteriormente, los modelos cinéticos de redes metabólicas, como el desarrollado en esta tesis, permiten la determinación de las principales cinéticas *in vivo*, mediante la aplicación de una perturbación tipo pulso en un cultivo continuo y el ajuste de las cinéticas frente a la respuesta dinámica del sistema. En este capítulo, se formula un problema de optimización sujeto a restricciones diferenciales y algebraicas que representan el modelo de la red metabólica de *E coli*, donde los principales parámetros a estimar son los grados de libertad del sistema. Se seleccionaron dos

conjuntos de datos de bibliografía para estimar y validar el modelo. Esto ha permitido contar con un modelo con cinéticas ajustadas *in vivo*, con estrategias de programación matemática avanzada y mejores ajustes que los reportados en la bibliografía. En este capítulo se presenta también estimación de parámetros para un modelo no estructurado, no segregado de un quimiostato, en las condiciones que se consideraron para el ajuste del modelo metabólico. Los datos experimentales de las concentraciones extracelulares que se reportan en este capítulo se obtuvieron como parte del trabajo de tesis.

El capítulo 5 presenta la formulación de problemas de programación no lineal como problemas de diseño óptimo para modelos cinéticos de la red metabólica de *Escherichia coli*. Se considera también la incorporación de restricciones de estabilidad en el problema de optimización. Este enfoque permite diseñar una red metabólica que maximiza la producción de un objetivo biotecnológico, asegurando la estabilidad del sistema en los niveles de diseño de las enzimas y de las concentraciones de los metabolitos. Los resultados numéricos indican la necesidad de incorporar restricciones de estabilidad al problema de diseño, para evitar la determinación de estados estacionarios inestables que lleven a las células a una situación de vulnerabilidad a perturbaciones externas. La estrategia propuesta para la inclusión de las condiciones de estabilidad no ha sido antes reportada para una red metabólica compleja. Se puede concluir que el enfoque de diseño bajo restricciones de estabilidad empleado en esta tesis constituye una herramienta eficiente y confiable para el diseño de redes metabólicas estables y de gran tamaño

Finalmente, se es de destacar que las técnicas de ingeniería de sistemas de procesos constituyen herramientas poderosas para el modelado, diseño y optimización de caminos metabólicos, de fundamental importancia en el desarrollo de la biotecnología actual.

6.3. Trabajos futuros

La metodología aplicada en esta tesis y los resultados obtenidos contribuirán a avanzar hacia estrategias más robustas para el diseño de redes metabólicas estables, así como hacia la formulación de modelos integrados de redes metabólicas y bio-reactores. A continuación se describen posibles líneas de investigación referidas al trabajo desarrollado en esta tesis.

- ✓ Inclusión de potenciales caminos metabólicos en cepas, asociados a variables binarias, en el marco de problemas de programación matemática no lineal mixto entera. Los resultados obtenidos muestran la conveniencia de incluir en el diseño óptimo la potencial inserción de genes heterólogos en las cepas bajo estudio.
- ✓ Análisis de estrategias alternativas de inclusión de condiciones de estabilidad, basado en la teoría de estabilidad de sistemas no lineales.
- ✓ Formulación de modelos integrados de bio-reactor y redes metabólicas, como problemas de optimización dinámica en dos niveles. Se propone incluir, en el nivel externo de optimización, ecuaciones diferenciales que describen el comportamiento del bio-reactor. En el nivel interno se incluiría

un problema de optimización de la red metabólica, en estado estacionario. De esta manera los modelos integrados quedan representados por un sistema diferencial algebraico. La relación entre el medio interno y externo a la célula está dada por las reacciones de transporte de los metabolitos extracelulares.

BIBLIOGRAFÍA CITADA

A

Akowski, J. y Bauerle, R. (1997). Steady-State kinetics and inhibitor binding of 3-Deoxy-D-arabino-heptulosonate-7-phosphate synthase (Tryptophan sensitive) from *Escherichia coli*. *Biochemistry*, 36, 15817-15822.

Alba, M. & Calvo, E. (2000). Characterization of bioreaction processes: Aerobic *Escherichia coli* cultures. *Journal of Biotechnology*, 84, 107-118.

Asenjo, J., Ramirez, P., Rapaport, I., Aracena, J., Goles, E. & Andrews, B. (2007). A discrete mathematical model applied to genetic regulation and metabolic networks. *Journal of Microbiology and Biotechnology*, 17 (3), 496-510.

B

Bailey, J. (1991). Toward a science of metabolic engineering. *Science*, 252 (5013), 1668-1675.

- Bailey, J. (2001). Reflections on the scope and the future of metabolic engineering and its connections to functional genomics and drug discovery. *Metabolic Engineering*, 3, 111-114.
- Bairoch, A. (2000). The ENZYME database in 2000. *Nucleic Acid Research*, 28 (1), 304-305.
- Biegler, J. & Cuthrell E. (1985). Improved infeasible path optimization for sequential modular simulators II. The optimization algorithm. *Computers and Chemical Engineering*, 9, 257-265.
- Biegler, L. T., Cervantes, A. M. & Waechter, A. (2002). Advances in simultaneous strategies for dynamic process optimization. *Chemical Engineering Science*, 57, 575-593.
- Blanco, A.M. & Bandoni, J.A. (2003). Interaction between process design and process operability of chemical processes: An eigenvalue optimization approach. *Computers and Chemical Engineering*, 27, 1291.
- Blanco, A. M. & Bandoni, J. A. (2007). Eigenvalue optimization-based formulation for nonlinear dynamics and control problems. *Chemical Engineering and Processing: Process Intensification*, 46, 1192-1199.
- Blattner, F., Plunkett, G., Bloch, C., Perna, N., Burland, V., Riley, M., Collado-Vides, J., Glasner, J., Rode, C., Mayhew, G., Gregor, J., Davis, N., Kirkpatrick, H., Goeden, M., Rose, D., Mau, B. & Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science*, 277, 1453-1462.
- Bolten, C., Kiefer, P., Letisse, F., Portais, J., & Wittmann C. (2007). Sampling for metabolome analysis of microorganisms. *Analytical Chemistry*, 79 (10), 3843-3849.
- Bonarius, H., Schmid, G., & Tramper, J. (1997). Flux analysis of underdetermined metabolic networks: The quest for the missing constraints. *Trends in Biotechnology*, 15, 308-314.
- Brenan, K. E., Campbell, S. L., & Petzold, L. (1996). The numerical solution of initial value problems in differential-algebraic equations (2nd ed.). SIAM Classics Series Elsevier Science Publishing Co.

Buchholz, A., Hurlebaus, J., Wandrey, C. & Takors, R. (2002). Metabolomics: quantification of intracellular metabolite dynamics. *Biomolecular Engineering*, 19, 5-15.

Bud, R. (2009). History of Biotechnology, en Bowler, P. & Pickstone, J. The modern biological and earth sciences. *The Cambridge History of Science* (6), 524-538. Ed.: Cambridge University Press.

Burgard, A. P., & Maranas, C. D. (2001). Probing the performance limits of the *Escherichia coli* metabolic network subject to gene additions or deletions. *Biotechnology and Bioengineering*, 74, 364–375.

C

Ceric, S. & Kurtanjek, Z. (2006). Model identification, parameter estimation and dynamic flux analysis of *E. coli* central metabolism. *Chemical and Biochemical Engineering Quarterly*, 20, 3, 243-253.

Cervantes, A. M., Tonelli, S., Brandolin, A., Bandoni, J. A. & Biegler, L. T. (2002). Large-scale dynamic optimization for grade transitions in a low density polyethylene plant. *Computers & Chemical Engineering*, 26 (2), 227-237.

Cervantes, A. M., Waechter, A., Tutuncu, R. & Biegler, L. T. (2000). A reduce space interior point strategy for optimization of differential algebraic systems. *Computers & Chemical Engineering*, 24, 39-51.

Chance, B., Garfinkel, D., Higgins, J. & Hess, B. (1960). Metabolic control mechanisms V. A solution for the equations representing interaction between glycolysis and respiration in ascites tumor cells. *Journal of Biological Chemistry*, 235, 2426-2439.

Chang, Y. & Sahinidis, N. (2005). Optimization of metabolic pathways under stability considerations. *Computers & Chemical Engineering*, 29, 467-479.

- Chassagnole, C., Noisommit-Rizzi, N., Schmid, J. W., Mauch, K., & Reuss, M. (2002). Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnology and Bioengineering*, 79, 53–72.
- Chassagnole, C., Doncescu, A., Manyri, L. & Yang, L. (2006). Parameter estimation by differential evolutionary algorithms for simulation of metabolic pathways in *Escherichia coli*. *Proceedings of the 20th International Conference on Advanced Information Networking and Applications (AINA'06)*
- Chung, B., Selvarasu, S., Camattari, A., Ryu, J., Lee, H., Ahn, J., Lee, H. & Lee, D. (2010). Genome-scale metabolic reconstruction and in silico analysis of methylotrophic yeast *Pichia pastoris* for strain improvement. *Microbial Cell Factories*, 9:50.
- Cukier, R. I., Fortuin, C.M., Shuler, K. E., Petschek, A. G., & Schaibly, J. H. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. *Journal of Chemical Physics*, 59, 3873–3878.
- Cukier, R. I., Levine, H. B., & Shuler, K. E. (1978). Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26(1), 1-42.
- Cuthrell, J. E. & Biegler, L.T. (1987). On the optimization of differential-algebraic process systems. *AIChE Journal*, 33 (8), 1257-1270.

D

- Dauner, M., Bailey, J. & Sauer, U. (2001). Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*. *Biotechnology and Bioengineering*, 76, 144-156.
- de Figueiredo, L., Podhorski, A., Rubio, A., Kaleta, C., Beasley, J., Schuster, S. & Planes, F. (2009). Computing the shortest elementary flux modes in genome-scale metabolic networks. *Bioinformatics*, 25 (23), 3158-3165.
- Degenring, D., Froemel, C., Dikta, G. & Takors, R. (2004). Sensitivity analysis for the reduction of complex metabolism models. *Journal of Process Control*, 14, 729-745.

- Degenring, D. (2004). Erstellung und validierung mechanistischer modelle für den mikrobiellen stoffwechsel zur auswertung von substrat-puls-experimenten. Dissertation. University of Rostock.
- Di Maggio, J., Diaz Ricci, J.C. & Diaz, M.S. (2008), Global sensitivity analysis in metabolic networks, *Mecánica Computacional*, 27, 2531-2540.
- Di Maggio, J., Díaz Ricci J.C. & Díaz, M.S. (2009a). Global Sensitivity Analysis in dynamic metabolic networks. *Computer Aided Chemical Engineering*, 26, 1075-1080.
- Di Maggio, J., Blanco, A.M., Bandoni, J.A. & Díaz, M.S. (2009b). Design of stable large-scale metabolic network. *Computer Aided Chemical Engineering*, 27, 1755-1760.
- Di Maggio, J., Diaz Ricci, J. C. & Diaz, M. S. (2010a). Global sensitivity analysis in dynamic metabolic networks. *Computers & Chemical Engineering*, 34 (5), 770-781.
- Di Maggio, J., Diaz Ricci, J.C. & Diaz, M.S. (2010b). Parameter estimation in kinetic models for large scale metabolic networks with advanced mathematical programming techniques. *Computer Aided Chemical Engineering*, 28, 355-360.
- Di Maggio, J., Blanco, A. M., Bandoni, J. A. & Diaz, M.S. (2010c). Large-scale metabolic network design under stability constraints for sustainable biofuel production. ALIO/INFORMS 2010 Joint International Meeting, 6-9 junio 2010, Buenos Aires, Argentina.
- Diaz Ricci, J. C., Regan, L. & Bailey, J.E. (1991a). Effect of alteration of the acetic acid synthesis pathway on the fermentation pattern of *Escherichia coli*. *Biotechnology & Bioengineering*, 38, 1318-1324.
- Diaz Ricci, J. C., Hitzmann, B., & Bailey, J. E. (1991b). In vivo NMR analysis of the influence of pyruvate decarboxylase and alcohol dehydrogenase of *Zymomonas mobilis* on the anaerobic metabolism of *Escherichia coli*. *Biotechnology Progress*, 7, 305–310.
- Diaz Ricci, J. C. (1996). Influence of phosphoenolpyruvate on the dynamic behavior of phosphofructokinase of *Escherichia coli*. *Journal of Theoretical Biology*, 178, 145–150.

Diaz Ricci, J. C. (2000). ADP modulates the dynamic behavior of the glycolytic pathway of *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 271, 244–249.

Duarte, N., Palsson, B. & Fu, P. (2004). Integrated analysis of metabolic phenotypes in *Saccharomyces cerevisiae*. *BMC Genomics*, 5:63.

E

Edwards, J. & Palsson, B. (2000). The *Escherichia coli* MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (10), 5528-5533.

Edwards, J., Covert, M. & Palsson, B. (2002). Metabolic modelling of microbes: The flux-balance approach. *Environmental Microbiology*, 4 (3), 133-140.

F

Fredrickson, A. G. (1976). Formulation of structured growth models. *Biotechnology and Bioengineering*, 18(10), 1481–1486.

G

Garfinkel, D. & Hess, B. (1964). Metabolic control mechanisms VII. A detailed computer model of the glycolytic pathways in ascites cells. *Journal of Biological Chemistry*, 239, 971-983.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acid Research*, 31 (13), 3784-3788.

Gerhard, J., Mönnigmann, M., & Marquardt, W. (2008). Steady state optimization with guaranteed stability of a tryptophan biosynthesis model. *Computers & Chemical Engineering*, 32, 2914-2919.

Ghosh, S., Zhu, T., Grossmann, I. E., Ataai, M.M., & Domach, M.M. (2005). Closing the loop between feasible flux scenario identification for construct evaluation and resolution of realized fluxes via NMR. *Computers & Chemical Engineering*, 29 (3), 459–466.

Goto, S., Nishioka, T. & Kanehisa, M. (2000). LIGAND: Chemical database of enzyme reactions. *Nucleic Acid Research*, 28 (1), 380-382.

Goto, S., Okuno, Y., Hattori, M., Nishioka, T. & Kanehisa, M. (2002). LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acid Research*, 30 (1), 402-404.

g-PROMS, <http://www.psenterprise.com>.

H

Han, M. & Lee, S. (2006). The *Escherichia coli* proteome: past, present, and future prospects. *Microbiology and Molecular Biology Reviews*, 70 (2), 362-439.

Hardiman, T., Meinhold, H., Hofmann, J., Ewald, J. Sieman-Herzberg, M. & Reuss, M. (2010). Prediction of kinetic parameters from DNA-binding site sequences for modeling global transcription dynamics in *Escherichia coli*. *Metabolic Engineering*, 12, 196-211.

Hatzimanikatis, V., Floudas, C. & Bailey, J. (1996). Analysis and design of metabolic reaction networks via mixed-integer linear optimization. *AIChE Journal*, 42, 1277-1292.

Hatzimanikatis, V. & Bailey, J. (1997). Effects of spatiotemporal variations of metabolic control, approximate analysis using (log) linear kinetic models. *Biotechnology and Bioengineering*, 1, 75-87.

- Hatzimanikatis, V., Emmerling, M., Sauer, U. & Bailey, J. (1998). Application of mathematical tools for metabolic design of microbial ethanol production. *Biotechnology and Bioengineering*, 58, 154-161.
- Henkin, J. & Abeles, R. (1976). Evidence against an Acyl-enzyme intermediate in the reaction catalyzed by clostridial phosphotransacetylase. *Biochemistry*, 15 (16), 3472-3479.
- Hiller, J., Franco-Lara, E., Papaioannou, V., Weuster-Botz, D. (2007). Fast sampling and quenching procedures for microbial metabolic profiling. *Biotechnology Letters*, 29, 1161-1167.
- Hoefnagel, M., Starrenburg, M., Martens, D., Hugenholtz, J., Kleerebezem, M., Van Swam, I., Bongers, R., Westerhoff, H. & Snoep, J. (2002). Metabolic engineering of lactic acid bacteria, the combined approach: kinetic modelling, metabolic control and experimental analysis. *Microbiology*, 148, 1003-1013.
- Homma, T., & Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52, 1–17.
- Hulse, J. (2004). Biotechnologies: past history, present state and future prospects. *Trends in Food Science & Technology*, 15, 3-18.

I

- Ingram, L., Conway, T., Clark, D., Sewell, G. & Preston, J. (1987). Genetic engineering of ethanol production in *Escherichia coli*. *Applied and Environmental Microbiology*, 53, 2420-2425.

K

- Kai, Y., Matsumura, H. & Izui, K. (2003). Phosphoenolpyruvate carboxylase: three-dimensional structure and molecular mechanisms. *Archives of Biochemistry and Biophysics*, 414, 170-179.
- Kai, Y., Matsumura, H. y Izui, K. (2003). Phosphoenolpyruvate carboxylase: three-dimensional structure and molecular mechanisms. *Archives of Biochemistry and Biophysics*, 414, 170-179.
- Kahlil, H.K., *Nonlinear Systems*, Prentice Hall, 1996.
- Kanehisa, M. & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28 (1), 27-30.
- Karp, P., Riley, M., Paley, S., Pellegrini-Toole, A. & Krummenacker, M. (1998). EcoCyc: Encyclopedia of *Escherichia coli* genes and metabolism. *Nucleic Acids Research*, 26 (1), 50-53.
- Kauffman, K., Prakash, P. & Edwards, J. (2003). Advances in flux balance analysis. *Current opinion in Biotechnology*, 14, 491-496.
- Keseler, I., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muñoz-Rascado, L., Bonavides-Martinez, C., Paley, S., Krummenacker, M., Altman, T., Kaipa, P., Spaulding, A., Pacheco, J., Latendresse, M., Fulcher, C., Sarker, M., Shearer, A., Mackie, A., Paulsen, I., Gunsalus, R. & Karp, P. (2011). EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Research*, 39, 583-590.
- Klamt, S & Stelling, J. (2003). Two approaches for metabolic pathway analysis?. *Trends in Biotechnology*, 21 (2), 64-69.
- Knappe, J., Blaschkowski, H., Gröbner, P. & Schmitt, T. (1974). Pyruvate Formate-Lyase of *Escherichia coli*: the Acetyl-Enzyme Intermediate. *European Journal of Biochemistry*, 50 (1), 253-263.

- Koda, M., & Seinfeld, J. H. (1982). Sensitivity analysis of distributed parameter systems. *IEEE Transactions on Automatic Control*, 27, 951–955.
- Koda, M., Dogru, A. H., & Seinfeld, J. H. (1979a). Sensitivity analysis of partial differential equations with application to reaction and diffusion processes. *Journal of Computational Physics*, 30, 259–282.
- Koda, M., McRae, G. J., & Seinfeld, J.H. (1979b). Automatic sensitivity analysis of kinetic mechanisms. *International Journal of Chemical Kinetics*, 11, 427–444.
- Kokossis, A. C. & Floudas, C. A. (1994). Stability in optimal design: Synthesis of complex reactor networks. *AIChE Journal*, 40 (5), 849-861.
- Kornberg, H. & Malcovati, M. (1973). Control in situ of the pyruvate kinase activity of *Escherichia coli*. *FEBS Letters*, 32 (2), 257-259.
- Kotera, M., Hirakawa, M., Tokimatsu, T., Goto, S. & Kanehisa, M. (2012). The KEGG databases and tools facilitating omics analysis: Latest developments involving human diseases and pharmaceuticals. *Methods in Molecular Biology*, 802, 19-39.
- Kremling A., Bettenbrock, K. & Gilles, E. (2007). Analysis of global control of *Escherichia coli* carbohydrate uptake. *BMC Systems Biology*, 1:42.
- Kremling, A., Fischer, S., Sauter, T., Bettenbrock, K. & Gilles, E. (2004). Time hierarchies in the *Escherichia coli* carbohydrate uptake and metabolism. *BioSystems*, 73, 57-71.

L

- Lander, E. (1999). Array of hope. *Nature genetics*, 21, 3-4.
- Lee, S., Palakornkule, C., Domach, M. M., & Grossmann, I. E. (2000). Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. *Computers & Chemical Engineering*, 24(2–7), 711–716.

- Liao, J. & Oh, M. (1999). Toward predicting metabolic fluxes in metabolically engineered strains. *Metabolic Engineering*, 1, 214-223.
- Liao, J., Hou, S. & Chao, Y. (1996). Pathway analysis, engineering, and physiological considerations for redirecting central metabolism. *Biotechnology and Bioengineering*, 52, 129-140.
- Liebermeister, W. & Klipp, E. (2006a). Bringing metabolic network to life: convenience rate law and thermodynamic constraints. *Theoretical Biology and Medical Modelling*, 3:41.
- Liebermeister, W. & Klipp, E. (2006b). Bringing metabolic network to life: integration of kinetic, metabolic and proteomic data. *Theoretical Biology and Medical Modelling*, 3:42.
- Liepmann, D., & Stephanopoulos, G. (1985). A dynamic model of a closed ecosystem: Development and global sensitivity analysis. *Ecological Modelling*, 30, 13–30.
- Lin, H., Bennett, G. & San, K. (2005). Metabolic engineering of aerobic succinate production systems in *Escherichia coli* to improve process productivity and achieve the maximum theoretical succinate yield. *Metabolic Engineering*, 7, 116-127.
- Lin, X., Brandes, A., Zucker, J. & Church, G.M. (2005). A new MILP based approach for in silico reconstruction of metabolic networks and its application to marine cyanobacterium *prochlorococcus marinus*. *AIChE Annual Meeting, Conference Proceedings 2005*, 9308.
- Liu, J., & Chen, R. (1998). Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93 (443), 1032-1044.
- Lorenz, P. & Zinke, H. (2005). White biotechnology: differences in US and EU approaches?. *Trends in Biotechnology*, 23 (12), 570-57

M

- Majewski, R. A., & Domach, M. M. (1990). Simple constrained optimization view of acetate overflow in *Escherichia coli*. *Biotechnology and Bioengineering*, 35, 732-738.

- Marangoni, A. (2003). Enzyme kinetics. A modern approach. Wiley Interscience.
- Masoudi-Nejad, A., Goto, S., Endo, T.R. & Kanehisa, M. (2007). KEGG bioinformatics resource for plant genomics research. *Methods in Molecular Biology*, 406, 437-458.
- Matallana, L., Blanco, A. & Bandoni, J. (2008). Eigenvalue optimization techniques for nonlinear dynamic analysis and design. En Ferruh Erdogdu, *Optimization in Food Engineering*. Ed.; Taylor and Francis Group.
- Mauch, K., Arnold, S., & Reuss, M. (1997). Dynamic sensitivity analysis for metabolic systems. *Chemical Engineering Science*, 52(15), 2589–2598.
- Mc Rae, G. J., Tilden, J. W., & Seinfeld, J. H. (1982). Global sensitivity analysis - A computational implementation of the Fourier Amplitude Sensitivity Test (FAST). *Computers & Chemical Engineering*, 6, 15.
- Mönnigmann, M., & Marquardt, W. (2003). Steady state process optimization with guaranteed robust stability and robust feasibility. *American Institute Of Chemical Engineers Journal*, 49, 3110-3126.
- Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33, 161–174.
- Motamediam, E. & Naeimpoor, F. (2011). Flux distribution in *Bacillus subtilis*: Inspection on plurality of optimal solutions. *Iranian Journal of Biotechnology*, 9 (4), 260-266.
- Muñoz, M. & Ponce, E. (2003). Pyruvate kinase: current status of regulatory and functional properties. *Comparative Biochemistry and Physiology*, part. B, 135, 197-218.

N

- Namjoshi, A. A., & Ramkrishna, D. (2005). A cybernetic modeling framework for analysis of metabolic systems. *Computers & Chemical Engineering*, 29(3), 487–498.

- Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7, 308-313.
- Nelson, D. & Cox, M. (2005). *Lehninger. Principios de Bioquímica*. Cuarta edición. Editorial: Omega.
- Nielsen, J. (1997). Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. *Biochemical Journal*, 321, 1, 133-138.
- Nikel, P., de Almeida, A., Melillo, E., Galvagno, M. & Pettinari, M. (2006). New recombinant *Escherichia coli* strain tailored for the production of Poly(3-Hydroxybutyrate) from agroindustrial by-products. *Applied and Environmental Microbiology*, 72 (6), 3949-3954.
- Nikolaev, E. (2010). The elucidation of metabolic pathways and their improvements using stable optimization of large-scale kinetic models of cellular systems. *Metabolic Engineering*, 12 (1), 26-38.
- Nikolaev, E., Pharkya, P., Maranas, C. & Armaou, A. (2005). Optimal selection of enzyme levels using large-scale kinetic models. *Proceedings of 16th I.F.A.C. World Congress*, Prague, Czech Republic.
- Noack, S., Wahl, A., Haunschild, M., Qeli, E., Freisleben, B., & Wiechert, W. (2008). Visualizing regulatory interdependencies and parameter sensitivities in biochemical models. *Mathematics and Computers in Simulation*, 79(4), 991–998.

O

- Oldiges, M. & Takors, R. (2005). Applying metabolic profiling techniques for stimulus-response experiments: chances and pitfalls. *Advances in Biochemical Engineering/Biotechnology*, 92, 173-196.
- Orozco de Silva, A. & Fraenkel, D. (1979). The 6-phosphogluconate dehydrogenase reaction in *Escherichia coli*. *The Journal of Biological Chemistry*, 254 (20), 10237-10242.

P

Papin, J., Stelling, J., Price, N., Klamt, S., Schuster, S. & Palsson, B. (2004). Comparison of network-based pathway analysis methods. *Trends in Biotechnology*, 22 (8), 400-405.

Pastres, R., & Ciavatta, S. (2005). A comparison between the uncertainties in model parameters and in forcing functions: Its application to a 3D water-quality model. *Environmental Modelling & Software*, 20, 981–989.

Patil, K., Akesson, M. & Nielsen, J. (2004). Use of genome-scale microbial models for metabolic engineering. *Current Opinion in Biotechnology*, 15, 64-69.

Phalakornkule, C., Lee, S., Zhu, T., Koepsel, R., Atai, I. & Grossmann, I. (2001). A MILP-based flux alternative generation and NMR experimental design strategy for metabolic engineering. *Metabolic Engineering*, 3, 124-137.

Pramanik, J. & Keasling, J. (1997). Stoichiometric model of *Escherichia coli* metabolism: incorporation of growth rate dependent biomass composition and mechanistic energy requirements. *Biotechnology and Bioengineering*, 56, 398-421.

Preiss, J., Greenberg, E. & Sabraw, A. (1975). Biosynthesis of bacterial glycogen. Kinetic studies of a glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) from a glycogen deficient mutant of *Escherichia coli* B. *The Journal of Biological Chemistry*, 250 (19), 7631-7638.

Price, N., Papin, J., Schilling, C. & Palsson, B. (2003). Genome-scale microbial in silico models: the constraints-based approach. *Trends in Biotechnology*, 21 (4), 162-169.

R

Ramakrishna, R., Edwards, J., McCulloch, A. & Palsson, B. (2001). Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 280, 695-704.

- Ratushny A., Smirnova O., Usuda Y. & Matsui K. (2006). Regulation of the pentose phosphate pathway in *Escherichia coli*: Gene network reconstruction and mathematical modeling of metabolic reactions. Modelling of molecular genetic systems in bacterial cell, 40-44.
- Reed, J. & Palsson, B. (2003). Thirteen years of building constraint-based in silico models fo *Escherichia coli*. Journal of Bacteriology, 185 (9), 2692-2699.
- Ritcher, O., Betz, A. & Giersch, C. (1975). The response of oscillating glycolysis in the NADH/NAD system: A comparison between experiment and a computer model. BioSystems, 7, 137-146.
- Rizzi, M., Baltés, M., Theobald, U., & Reuss, M. (1997). In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*. II. Mathematical model. Biotechnology and Bioengineering, 55(4), 592–607.
- Rodríguez, A. & Infante, D. (2009). Network models in the study of metabolism. Electronic Journal of Biotechnology, 12 (3).
- Rodríguez, M. A. (2009). Modelamiento y Optimización Dinámica de Procesos Criogénicos de Separación. (Tesis de Doctor en Ingeniería Química, Universidad Nacional del Sur, Bahía Blanca, Argentina)
- Rohwer, J., Meadow, N., Roseman, S., Westerhoff, H. & Postma, P. (2000). Understanding glucose transport by the bacterial phosphoenolpyruvate:glycose phosphotransferase system on the basis of kinetic measurements in vitro. The Journal of Biological Chemistry, 275 (45), 34909-34921.

S

- Saltelli, A., & Sobol', I. M. (1995). About the use of rank transformation in sensitivity analysis of model output. Reliability Engineering & System Safety, 50, 225–239.

- Saltelli, A., & Tarantola, S. (2002). On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal. *Journal of the American Statistical Association*, 97(459), 702–709.
- Saltelli, A., Tarantola, S., & Chan, K. (1999). Quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1), 39–56.
- Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity analysis in practice: A guide to assessing scientific models*. Wiley.
- Sanwal, B. (1970). Regulatory mechanism involving nicotinamide adenine nucleotides as allosteric effectors. *The Journal of Biological Chemistry*, 245 (7), 1626-1631.
- Sauer, U., Cameron, D. y Bailey, J. (1998). Metabolic capacity of *Bacillus subtilis* for the production of purine nucleosides, riboflavin, and folic acid. *Biotechnology and Bioengineering*, 59, 227-238.
- Sauer, U., Lasko, D., Fiaux, J., Hochuli, M., Glaser, R., Szyperski, T., Wüthrich, K. & Bailey, J.(1999). Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. *Journal of Bacteriology*. 181, 6679-6688.
- Schaefer, U., Boos, W., Takors, R. & Weuster-Botz, D. (1999). Automated sampling device for monitoring intracellular metabolite dynamics. *Analytical Chemistry*, 270, 88-96.
- Schilling, C., Letscher, D. & Palsson, B. (2000). Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of Theoretical Biology*, 203 (3), 229-248.
- Schmid, J., Mauch, K., Reuss, M., Gilles, E. & Kremling, A. (2004). Metabolic design based on a coupled gene expression-metabolic network model of tryptophan production in *Escherichia coli*. *Metabolic Engineering*, 6, 364-377.
- Schmidt, K., Carlsen M, Nielsen J. & Villadsen J. (1997). Modeling isotopomer distributions in biochemical networks using isotopomer mapping matrices. *Biotechnology and Bioengineering*, 55, 831-840.

- Schmidt, K., Nielsen, J. & Villadsen, J. (1999). Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two dimensional NMR spectroscopy and complete isotopomer models. *Journal of Biotechnology*, 71, 175-190.
- Schomburg I, Chang A, Hofmann O, Ebeling C, Ehrentreich F & Schomburg D. (2002). BRENDA: a resource for enzyme data and metabolic information. *Trends in Biochemical Sciences*, 27 (1), 54-56.
- Schreyer, R. & Böck, A. (1980). Phosphoglucose isomerase from *Escherichia coli* K 10: Purification, properties and formation under aerobic and anaerobic condition. *Archives of Microbiology*, 127, 289-298.
- Schuster, S., Dandekar, T. & Fell, D. (1999). Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, 17 (2), 53-60.
- Schwender, J. (2008). Metabolic flux analysis as a tool in metabolic engineering of plants. *Current opinion in Biotechnology*, 19, 131-137.
- Sevilla, A., Schmid, J., Mauch, K., Iborra, J., Reuss, M. & Cánovas, M. (2005). Model of central and trimethylammonium metabolism for optimizing L-carnitine production by *E. coli*. *Metabolic Engineering*, 7, 401-425.
- Sigrist, C., Cerutti, L., de Castro, E., Langendijk-Genevaux, P., Bulliard, V., Bairoch, A. & Hulo, N. (2009). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Research*, 38, 161-166.
- Snoep, J. (2005). The Silicon Cell initiative: working towards a detailed kinetic description at the cellular level. *Current Opinion in Biotechnology*, 16, 336-343.
- Sobol', I.M. (1990). Sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 2, 112-118 in Russian, translated in English in Sobol' (1993). *Mathematical Modelling and Computational Experiment*, 1, 407-414.
- Sobol', I. M. & Levitan, Y. L. (1999). On the use of variance reducing multipliers in Monte Carlo computations of a global sensitivity index. *Computer Physics Communications*, 117(1999), 52-61.

- Sobol', I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55, 271-280.
- Soetaert, W. & Vandamme, E. (2010). *Industrial biotechnology. Sustainable Growth and Economic Success*. Ed.: Wiley-VCH.
- Sohn, S., Graf, A., Kim, T., Gasser, B., Maurer, M., Ferrer, P., Mattanovich, D. & Lee, S. (2010). Genome-scale metabolic model of methylotrophic yeast *Pichia pastoris* and its use for in silico analysis of heterologous protein production. *Biotechnology Journal*, 5, 705-715.
- Stephanopoulos, G., Aristidou, A. & Nielsen, J. (1998). *Metabolic engineering: Principles and methodology*. Elsevier.
- Stephanopoulos, G., & Stafford, D. (2002). Metabolic engineering: A new frontier of chemical reaction engineering. *Chemical Engineering Science*, 57, 2595-2602.
- Steuer, R. (2006). On the analysis and interpretation of correlations in metabolomic data. *Briefings in Bioinformatics*, 7 (2), 151-158.
- Suthers, P., Burgard, A., Dasika, M., Nowroozi, F., Van Dien, S., Keasling, J. & Maranas, C. (2007). Metabolic flux elucidation for large-scale models using ¹³C labeled isotopes. *Metabolic Engineering*, 9 (5-6), 387-405.
- Szyperski, T. (1995). Biosynthetically directed fractional ¹³C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. *European Journal of Biochemistry*. 232, 433-448.

T

- Tarmy, E. & Kaplan, N. (1968). Chemical characterization of D-Lactate dehydrogenase from *Escherichia coli* B. *The Journal of Biological Chemistry*, 243 (10), 2579-2586.

Teixeira, A., Santos, S., Carinhas, N., Oliveira, R. & Alves, P. (2008). Combining metabolic flux analysis tools and ¹³C NMR to estimate intracellular fluxes of cultured astrocytes. *Neurochemistry International*, 52, 478-486.

U

Uhlenbusch, I., Sahm, H. & Sprenger, G. (1991). Expression of an L-alanine dehydrogenase gene in *Zymomonas mobilis* and excretion of L-alanine. *Applied and Environmental Microbiology*, 57, 1360-1366.

Usudaa, Y., Nishio, Y., Iwatani, S., Van Dien, S., Imaizumi, A., Shimbo, K., Kageyama, N., Iwahata, D., Miyano, H. & Matsui, K. (2010). Dynamic modeling of *Escherichia coli* metabolic and regulatory systems for amino-acid production. *Journal of Biotechnology*, 147, 17-30.

V

Varma, A. & Palsson, B. O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild type *Escherichia coli* W3110. *Applied and Environmental Microbiology*, 60(10), 3724–3731.

Varma, A. & Palsson, B. (1994). Metabolic flux balancing: Basic concepts, scientific and practical use. *Nature Biotechnology*, 12, 994-998.

Varner, J. & Ramkrishna, D. (1999). Mathematical models of metabolic pathways. *Current opinion in Biotechnology*, 10, 146-150.

Vaseghi, S., Baumeister, A., Rizzi, M. & Reuss, M. (1999). In vivo dynamics of the pentose Phosphate pathway in *Saccharomyces cerevisiae*. *Metabolic Engineering*, 1, 128-140.

Vidyasagar, M. (1993). *Nonlinear systems analysis*. Prentice Hall.

Visser, D., Schmid, J., Mauch, K., Reuss, M. & Heijnen, J. (2004). Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metabolic Engineering*, 6, 378-390.

Visser, D., van der Heijden, R., Mauch, K., Reuss, M. & Heijnen, S. (2000). Tendency modeling: A new approach to obtain simplified kinetic models of metabolism applied to *Saccharomyces cerevisiae*. *Metabolic Engineering*, 2, 252-275.

Vital-Lopez, F., Maranas, C. & Armaou, A. (2006). Bifurcation analysis of the metabolism of *E. coli* at optimal enzyme levels. Proceedings of the 2006 American Control Conference, Minneapolis, Minnesota, USA.

W

Waechter, A. & Biegler, L. (2006). On the implementation of an interior point filter line search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106, 1, 25-57.

Won, W., Park, C., Park, C., Lee, S., Lee, K. & Lee, J. (2011). Parameter estimation and dynamic control analysis of central carbon metabolism in *Escherichia coli*. *Biotechnology and Bioprocess Engineering*, 16, 216-228.

Z

Zabala, V. & Biegler, L. T. (2006). Large-scale parameter estimation in low-density polyethylene tubular reactors. *Industrial and Engineering Chemistry Research*, 45 (23), 7867-7881.

Zhang, X., Jantama, K., Moore, J., Jarboe, L., Shanmugam, K., & Ingram, L. (2009). Metabolic evolution of energy-conserving pathways for succinate production in

Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 106 (48), 20180-20185.

Zupke, C. & Stephanopoulos, G. (1994). Modeling of isotope distributions and intracellular fluxes in metabolic networks using atom mapping matrices. *Biotechnology Progress*, 10, 489-498.