

## **INFORMATION TO USERS**

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

**U·M·I**

University Microfilms International  
A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
313/761-4700 800/521-0600



**Order Number 9423336**

**Flux balance analysis of *Escherichia coli* metabolism**

**Varma, Amit, Ph.D.**

**The University of Michigan, 1994**





# **FLUX BALANCE ANALYSIS OF *ESCHERICHIA COLI* METABOLISM**

by  
Amit Varma

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Chemical Engineering)  
in The University of Michigan  
1994

#### **Doctoral Committee:**

Associate Professor Bernhard O. Palsson, Chair  
Professor Robert A. Bender  
Assistant Professor Jennifer J. Linderman  
Professor Frederick C. Neidhardt  
Professor Henry Y. Wang

© Amit Varma 1994  
All Rights Reserved

## **ACKNOWLEDGEMENTS**

I would like to acknowledge the financial support provided by the Rackham Pre-Doctoral Fellowship which was crucial for initiating this work. Through its fellowship program the Horace H. Rackham School of Graduate Studies provides a remarkable catalyst for innovative research at The University of Michigan.

Throughout my graduate studies I was the beneficiary of useful advice from Dr. Srinivas Palanki. I take this opportunity to thank him for his close friendship. I would like to thank Alice Chuck and Sadanand Deshpande for their friendship and availability to lend an ear.

I would like to recognize the love and support of my close and distant family that has been a source of strength to me. In particular, I would like to thank my parents for their patient support and understanding.

The advice I received from my doctoral committee members was extremely useful in setting directions for my research. I would like to thank them all for their support and encouragement. Indeed education goes beyond academics and I appreciate my interactions with my doctoral committee members.

Dr. Palsson for his inspiring pep talks.

Dr. Bender for his anecdotal approach to research education.

Dr. Linderman for being so accessible and easy to talk to.

Dr. Neidhardt for his constant encouragement and guidance.

Dr. Wang for providing an industrial perspective.

The efforts of Dr. Elliot Juni to introduce me to microbial metabolism also deserve recognition.

My work with Dr. Gyun Min Lee was my first research experience at The University of Michigan. It was a pleasure working with him. I also appreciate my interaction with Dr. Stephen Emerson for the insights into hematology research. I would like to thank the undergraduates, Brian Boesch, and Geraldine Terradurro who participated in this research. It was a pleasure to be acquainted with them.

I would also like to thank the support and friendship of all my fellow students. In particular I would like to mention some names: Dr. Minoo Javanmardian, Dr. Jay Keasling, Dr. Anuradha Koratkar, Dr. Sabine Nakouzi, Dr. Sadettin Ozturk, Dr. Joanne Savinell, Dr. Richard Schwartz, Stelios Andreadis, Dimitrios Hatznikolau, Krishnamohan Kosaraju, Choul Gyun Lee, Minette Levee, Ram Mandalam, Ching-an Peng, and Shuang-Ying Yee. I also thank Barbara Dunn for her efficient office organization.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS . . . . .</b>	<b>ii</b>
<b>LIST OF FIGURES . . . . .</b>	<b>vii</b>
<b>LIST OF TABLES . . . . .</b>	<b>xi</b>
<b>NOMENCLATURE . . . . .</b>	<b>xiii</b>
<b>ABSTRACT . . . . .</b>	<b>xiv</b>
<b>CHAPTER</b>	
<b>I. INTRODUCTION . . . . .</b>	<b>1</b>
1.1 Goals . . . . .	3
1.2 Approach . . . . .	3
<b>II. HISTORY OF METABOLIC ANALYSIS . . . . .</b>	<b>7</b>
2.1 Dynamic Analysis of Metabolism . . . . .	8
2.2 Steady State Analysis . . . . .	11
2.3 Linear Optimization . . . . .	14
2.3.1 Shadow Prices . . . . .	16
<b>III. PREMISE FOR A FLUX BALANCE MODEL . . . . .</b>	<b>18</b>
3.1 Network Formulation . . . . .	19
3.2 Metabolic Demands for Growth . . . . .	24
3.3 Applicability of Linear Programming to Analysis of Metabolic Networks . . . . .	26
3.4 Mathematical Algorithms and Computer Implementation . . . . .	27
3.4.1 Simplex Algorithm . . . . .	28
<b>IV. CAPABILITIES OF THE CATABOLIC NETWORK . . . . .</b>	<b>29</b>
4.1 Optimal Production of Cofactors . . . . .	29

4.2	Optimal Production of Biosynthetic Precursors . . . . .	33
4.3	Summary . . . . .	39
<b>V.</b>	<b>STOICHIOMETRIC OPTIMALITY OF GROWTH . . . . .</b>	<b>41</b>
5.1	Optimal Growth Pattern on Glucose . . . . .	41
5.1.1	Which Factors Constrain Growth? . . . . .	43
5.1.2	How Sensitive is the Optimal Solution? . . . . .	48
5.1.3	The use of $\sigma$ to interpret the effects of variation in metabolic loads . . . . .	59
5.2	Growth on Acetate . . . . .	66
5.3	Conclusions . . . . .	68
<b>VI.</b>	<b>BIOCHEMICAL PRODUCTION CAPABILITIES . . . . .</b>	<b>69</b>
6.1	Definition of Metabolic Constraints . . . . .	71
6.2	Maximal Theoretical Performance . . . . .	72
6.3	Optimal Flux Distributions . . . . .	76
6.4	Balanced Growth with Biochemical Production . . . . .	84
6.5	Discussion and Conclusions . . . . .	85
<b>VII.</b>	<b>THE ROLE OF OXYGEN . . . . .</b>	<b>90</b>
7.1	Model Considerations . . . . .	90
7.2	Growth Dependence on Glucose Supply . . . . .	95
7.3	Growth on Glucose with Oxygen Limitations . . . . .	97
7.4	Discussion and Conclusions . . . . .	108
<b>VIII.</b>	<b>STABILITY OF CELL POPULATIONS . . . . .</b>	<b>112</b>
8.1	Analysis of Population Stability . . . . .	114
8.2	Culture Stability . . . . .	117
8.3	Example: Valine Production . . . . .	118
8.3.1	Engineering Optimal Conversion . . . . .	124
8.4	Example: Tryptophan and Lysine . . . . .	127
8.5	General Principles . . . . .	130
8.6	Discussion and Conclusions . . . . .	131
<b>IX.</b>	<b>EXPERIMENTAL OBSERVATIONS . . . . .</b>	<b>137</b>
9.1	Materials and Methods . . . . .	138
9.2	Results . . . . .	142
9.2.1	Strain Characterization . . . . .	142
9.2.2	Stoichiometrically Optimal By-Product Secretion . .	146
9.2.3	Flux Balance Model as a Predictive Tool . . . . .	148

9.3 Discussion . . . . .	157
<b>X. Stoichiometric Optimality: A Perspective . . . . .</b>	<b>160</b>
10.1 Stoichiometric Optimality: A Scientific Principle . . . . .	161
10.2 Applicability of Stoichiometric Optimality . . . . .	165
10.3 Future of Metabolic Analysis . . . . .	167
10.4 Brave New World . . . . .	171
<b>BIBLIOGRAPHY . . . . .</b>	<b>172</b>

## LIST OF FIGURES

### Figure

1.1	The scope of metabolic engineering . . . . .	3
1.2	Overview of the synthesis of an <i>E. coli</i> cell . . . . .	4
2.1	Schematic of flux balances around the fueling and biosynthetic reactions . . . . .	9
2.2	The stoichiometrically feasible domain of steady state reaction fluxes	13
2.3	Schematic representation of the shadow price . . . . .	17
3.1	The catabolic reaction network . . . . .	21
3.2	The biosynthetic reaction network . . . . .	22
4.1	Flux distributions for maximal (a) ATP and (b) NADPH production	30
4.2	Flux distributions for maximal (a) E4P and (b) PEP production .	34
4.3	Alternate flux distribution for maximal E4P production . . . . .	37
5.1	Optimal flux distribution for aerobic growth on glucose . . . . .	42
5.2	Logarithmic sensitivity of maximal growth to biomass composition .	47
5.3	Yield sensitivity to PPS and maintenance . . . . .	51
5.4	Biomass yield sensitivity to the P/O ratio . . . . .	54
5.5	Optimal aerobic growth with maintenance . . . . .	57
5.6	$\sigma$ values for NADPH . . . . .	59
5.7	Flux distribution change over a $\sigma$ discontinuity for NADPH production	60

5.8	$\sigma$ change with PEP drain . . . . .	62
5.9	Flux distribution change over a $\sigma$ discontinuity for PEP production	64
5.10	Optimal biomass yield on acetate as a function of ATP maintenance requirements . . . . .	66
6.1	Flux distributions for maximal cell growth . . . . .	78
6.2	Optimal catabolic flux distributions for maximal biochemical productions . . . . .	80
6.2	Continued . . . . .	81
6.2	Continued . . . . .	82
6.2	Continued . . . . .	83
6.3	The simultaneous production of biomass and specific biochemicals .	84
7.1	Oxygen demand of a cell culture in contrast to the supply rate . . .	91
7.2	Optimal aerobic growth and by-product secretion . . . . .	94
7.3	Optimal by-product secretion and growth under the complete range of oxygenation rates . . . . .	98
7.4	Optimal flux distributions characteristic of the five different phases of oxygenation . . . . .	100
7.5	The five phases of oxygenation shown as a function of the glucose and oxygen supply . . . . .	105
7.6	By-product secretion shown as a function of the cell concentration and glucose and oxygen supply . . . . .	106
7.7	Variation of critical growth rate, $\mu_c$ , for various by-products under different oxygenation rates . . . . .	107
8.1	Culture stability represented by the growth advantage as a function of the oxygen supply . . . . .	117
8.2	Culture stability during valine production as a function of the oxygen supply . . . . .	119

8.3	Variation of the growth rate of non-producers and valine producers as a function of the oxygen supply . . . . .	120
8.4	Flux distributions for the various phases of growth during valine production . . . . .	121
8.5	Total valine production in a culture as a function of the engineered specific productivity . . . . .	125
8.6	Average specific conversion of cultures computed as a function of time	126
8.7	Culture stability represented by $S$ given as a function of the oxygen supply to the cells . . . . .	128
8.8	Total productivity of a culture shown as a function of the specific productivity of the engineered clone . . . . .	129
8.9	Culture size, as the number of producing cells obtainable, for a culture to become 25% non-producing plotted as a function of the dimensionless engineered productivity . . . . .	130
8.10	Total production of a biochemical product as a function of the dimensionless engineered productivity, under different mutational probabilities, and dimensionless economic criteria . . . . .	132
8.11	Stability of an engineered clone in the presence of mutations . . . . .	133
9.1	Schematic of the bacterial metabolism . . . . .	138
9.2	Schematic of bioreactor setup . . . . .	139
9.3	Maximum enzymatic capacity for oxygen uptake . . . . .	143
9.4	Maximum aerobic glucose uptake . . . . .	144
9.5	Maximum anaerobic glucose uptake . . . . .	144
9.6	Determination of biomass and maintenance requirements . . . . .	145
9.7	Aerobic chemostat culture . . . . .	147
9.8	Aerobic chemostat culture . . . . .	148
9.9	Aerobic batch culture . . . . .	149

9.10	Aerobic fed-batch culture . . . . .	151
9.11	Aerobic fed-batch culture . . . . .	153
9.12	Aerobic fed-batch culture . . . . .	154
9.13	Anaerobic batch culture . . . . .	156
10.1	Some evolutionary events leading to the formation of life . . . . .	163

## LIST OF TABLES

### Table

3.1	Requirements of biosynthetic precursors and cofactors for growth . . . . .	24
3.2	Requirements of buiding blocks for growth . . . . .	25
4.1	Maximum energy and redox generation from glucose . . . . .	30
4.2	Maximum biosynthetic precursor yields from glucose . . . . .	33
5.1	Metabolic demands of precursors and cofactors . . . . .	44
5.2	Logarithmic sensitivity of maximal biomass yield . . . . .	52
5.3	Flux distribution for growth on acetate . . . . .	65
6.1	Definition of stoichiometric constraints . . . . .	70
6.2	Maximal theoretical yield of amino acids and nucleotides on a glucose substrate . . . . .	73
6.3	Maximal theoretical yield of amino acids and nucleotides on a glycerol substrate . . . . .	75
6.4	Maximal theoretical yield of amino acids and nucleotides on an acetate substrate . . . . .	77
6.5	Optimal trade-off between growth and biochemical production . . . . .	86
7.1	By-product secretion follows a linear relationship with growth rate . . . . .	95
7.2	Aerobic energy value of different by-products . . . . .	96
7.3	Shadow prices computed from the dual optimization problem . . . . .	99
7.4	Comparison of optimal and experimental anaerobic by-product secretions . . . . .	101

8.1	Shadow prices of redox, energy, and by-products in the various phases of growth . . . . .	123
-----	----------------------------------------------------------------------------------------------	-----

## NOMENCLATURE

- a - Mutational probability *per* cell division.
- A - Ratio of mutational rate to growth rate of producing population.
- b - Vector representing transport flux of metabolites out of the cell.
- C - Engineered conversion (mol product/mol substrate).
- $C_{max}$  - Maximum possible productivity (mol product/mol substrate).
- $d_M$  - Metabolic demands for growth (mmol metabolite/g biomass).
- $k_t a$  - Mass transfer coefficient.
- M - A particular metabolite in the metabolic network.
- $n$  - Subscript  $n$  refers to the non-producing population.
- $p$  - Subscript  $p$  refers to the producing population.
- P - Fractional engineered conversion compared to the maximal conversion ( $C/C_{max}$ ).
- $q_{O_2}$  - Oxygen consumption rate.
- $r_s$  - Rate of substrate uptake (mol/g DW-hr).
- S - Growth advantage of a non-producing population expressed as a ratio ( $\mu_n/\mu_p$ ).
- $S$  - The stoichiometric matrix for the metabolic network. An element  $S_{ij}$  represents the moles of metabolite  $i$  needed for reaction  $j$ .
- t - Time (hr).
- $t_f$  - Culture termination time (hr).
- TP - Total production (mmol product).
- v - The vector of reaction fluxes (mmol flux/g DW-hr). These fluxes are determined using linear optimization of an objective function.
- $v_{gro}$  - Biosynthetic growth flux defined as a drain of biosynthetic precursors in a ratio appropriate to produce cell biomass.
- X - Cell mass (g DW).
- Y - Critical economic productivity (mol product/mol substrate).
- Z - Denotes the linear optimization objective.
- $\alpha$  - Mutational rate (1/hr).
- $\gamma_i$  - Shadow price of the  $i^{th}$  metabolite.
- $\mu$  - Growth rate (1/hr).
- $\nu$  - By-product secretion rate (g by-product/g DW-hr).
- $\sigma$  - Dimensionless economic productivity ( $Y/C_{max}$ ).
- $\tau$  - Dimensionless time ( $\mu_p \cdot t$ ).

## ABSTRACT

### FLUX BALANCE ANALYSIS OF *ESCHERICHIA COLI* METABOLISM

by  
Amit Varma

Chair: Bernhard O. Palsson

The present thesis represents an effort in the emerging area of metabolic engineering. While the physical principles governing the unit operations of bioprocessing are quite well known, the design of the cellular component is poorly understood. Here we place microbial metabolism in an engineering context. We have determined a flux balance metabolic model for the bacterium *Escherichia coli* that includes the catabolic and biosynthetic reactions. Growth of the bacterium is defined in the model as a demand for metabolites based on published composition analyses.

The flux balance model is able to determine metabolic pathway utilization in the bacterium for specific objectives such as biochemical production or cell growth. We propose the principle of stoichiometric optimality, the hypothesis that metabolism functions to achieve an optimal pathway utilization to enhance growth and multiplication of the bacterial cell. Predictions from the model for *E. coli* growth and by-product secretion under various oxygen supply rates are found to provide an elegant interpretation for the observed physiology of *E. coli* metabolism.

We have applied the flux balance model to growth and product formation from clonal populations in a bioreactor. The effect of oxygen supply in the model is investigated for its ability to enhance the productivity and economics of a bioprocess.

We also undertake an experimental confirmation of the model's validity. We show how the flux balance model may be applied to bioprocesses to predict time profiles of cell growth, and nutrient and by-product concentration. Taken together these results provide a quantitative mathematical framework to describe and understand microbial metabolism. The engineering principles developed here should prove of use to bioprocess design and development.

# **CHAPTER I**

## **INTRODUCTION**

Microbial metabolism has been of use to man since ancient times. Domestication of microbial cells for wine making, brewing and baking is graphically depicted on Egyptian tombs. The 20<sup>th</sup> century has seen a particularly large expansion in the use of microbial cells following Pasteur's pioneering work in the late 19<sup>th</sup> century. Bio-processing gained importance as a sector of the chemical industry early this century and formed the basis for the rapidly expanding pharmaceutical industry in the post world war II era.

The ever increasing awareness of the limitations of non-renewable sources of energy and chemical feedstocks has made widespread the use of biotechnology to replace or substitute traditional petrochemical based products. Products from biological sources are also biodegradable and hence considered environmentally safe. Indeed bioprocesses utilizing microbes are being developed for waste remediation, in order to treat both organic and inorganic industrial wastes. The increasing competition in the area of biotechnology as well as the increasing use of biotechnology require a better quantitative understanding of cellular metabolism.

The traditional approach to obtain microbes with desired metabolic characteristics utilizes the techniques of random mutation of a cell population followed by

selection for the desired characteristic. Several examples, notably that of penicillin production, exist that attest to the success of this approach. However, lack of understanding of the cellular component has remained the weakest part of the overall bioprocess in this traditional approach.

The past two decades have seen an explosion of interest in the field of genetic engineering. One can now insert, alter, or delete genes almost at will. Numerous genes have now been cloned and vast arrays of microbial mutants are available. Extensive databases on cellular components and their function, in particular for *Escherichia coli*, are now at hand [32, 53]. The characteristics of the cell and its components need no longer to be considered as fixed. One can change prokaryotic cells almost at will to introduce desired metabolic characteristics. Thus, the cellular component of the biotechnological process can now be considered a subject of rational design as a part of the overall bioprocess design.

However, initial attempts to metabolically engineer cells have often yielded surprising and unpredictable results [2]. *A priori* inspection of the metabolic map as a basis for metabolic manipulation has not given satisfactory results. Functioning of a whole cell involves a set of interwoven and complex set of metabolic and genetic processes. A systemic study that accounts for all these processes is needed as a basis for metabolic engineering and rational construction of bacterial strains with predetermined metabolic characteristics. The present study is the first step in the undertaking of such an ambitious endeavour. As depicted in Figure 1.1, metabolic design of the cell should be an integral part of the overall design of an efficient bioprocess.

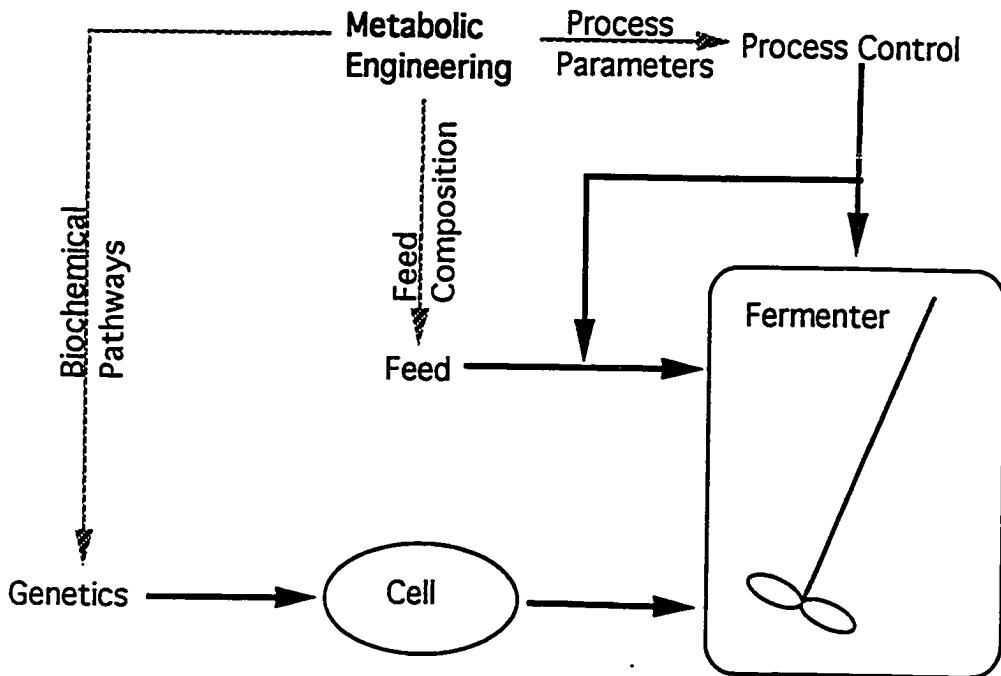


Figure 1.1: The scope of metabolic engineering.

## 1.1 Goals

The overall objective of this study is to determine the metabolic capabilities of microorganisms, and to quantitatively determine the change in metabolism required to achieve a desired goal. The micro-organism chosen for study is the prokaryotic bacterium *Escherichia coli*. The choice of *E. coli* is based primarily on two considerations. First, the wealth of information available on *E. coli* is unsurpassed among all micro-organisms, e.g. [32, 53]. Second, the widespread use of *E. coli* in science as well as industry makes the study very relevant and applicable to current biotechnology.

## 1.2 Approach

In order to better appreciate quantitative metabolic analysis we first survey the background literature on metabolic models in Chapter 2. We consider the applicability and usefulness of various modeling approaches. Linear programming as a tool

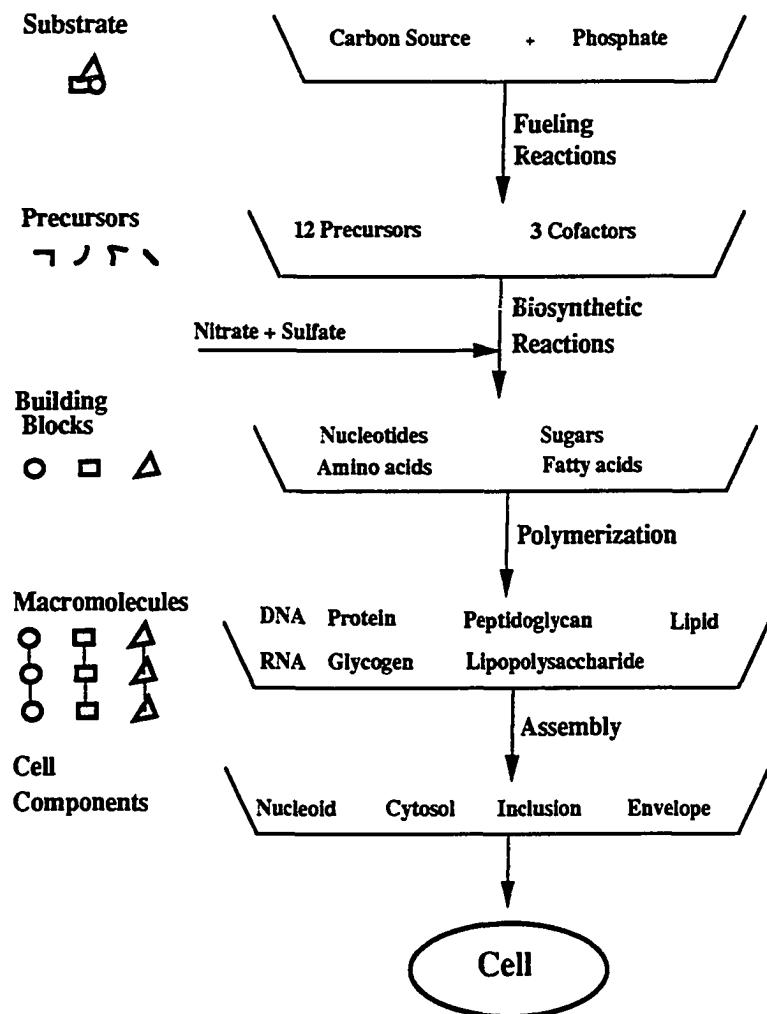


Figure 1.2: Overview of the synthesis of an *E. coli* cell.

to determine steady state metabolism is also described in Chapter 2. As shown in Figure 1.2, synthesis of the *E. coli* cell can be logically broken down into fueling, biosynthesis, polymerization, and assembly processes. Chapter 3 includes the fueling, and biosynthetic reaction network that together with the metabolic demands for growth provide a metabolic model of the *E. coli* cell. We also consider the applicability of linear programming to the metabolic model.

The mathematical analysis begins in Chapter 4 by examining the characteristics and constraints governing the synthesis of the twelve key precursor metabolites that feed into the biosynthetic reactions of metabolism. Chapter 5 then examines the aerobic growth of *E. coli* and the constraints imposed by the metabolic network. We propose stoichiometric optimality as the model of choice both for wild type metabolism as well as for optimal bioprocesses.

Several useful products are also intermediates in the metabolic network. In Chapter 6 we determine the capability of the *E. coli* metabolic network to produce these speciality chemicals or biochemicals. We also determine the metabolic constraints faced while utilizing some typical substrates.

In Chapter 7 we examine the effect of oxygen supply on optimal growth rate and metabolism. By-product secretion and shifts in the utilization of metabolic pathways are observed as the optimal response to changes in the oxygenation conditions.

By influencing the growth rate, oxygen supply can also influence the stability of multiple populations growing in the same culture environment. Chapter 8 quantitates the population stability as a function of growth rate. We are then able to predict the conditions that maximize the stability of an engineered production strain.

Chapter 9 provides some experimental implementation of the flux balance analysis that indicates its applicability. Chapter 10 concludes the present study with some

ideas to stimulate future work in the area of metabolic flux balance analysis and its application. This thesis ends with an overview of future potential in metabolic engineering building on the work presented here.

## CHAPTER II

# HISTORY OF METABOLIC ANALYSIS

The chemistry, stoichiometry and regulation of metabolism are perhaps the most thoroughly studied and best known aspects of the microscopic functioning of living cells. Although detailed knowledge exists about the biochemistry of individual metabolic events, less is known about the systemic nature of metabolic function and how it impacts the overall function of a cell.

**Network of Reactions** The metabolic network consists of the chemical transformations of low molecular weight compounds occurring within the cell. The effect of the metabolic network is to transform the carbon source and minerals into additional cellular material for maintenance as well as growth.

In this chapter we survey some of the mathematical tools that have been used to analyze the network of metabolic reactions. We begin by outlining the purpose of quantitative modeling and then describe linear optimization as an appropriate tool to achieve the goals stated in Chapter 1.

**Why Model?** Mathematical modeling serves many different purposes. At the most basic level it provides a method to compact experimental data. For example, nomographs present in handbooks of physical and chemical data provide a concise

representation of information. Models may also be used to convey a physical law or principle. Thus, for example, the Arrhenius Rate Law conveys the thermodynamic understanding of reaction rates.

In the area of process control one always needs a process model that quantitatively describes the process to be controlled and the interaction of the process with the controlled variables. A process model used for process control needs to be very accurate. However, such a model need not be mechanistic and in fact, correlative models are often quite useful.

Somewhat more sophisticated mathematical models, when used with care, can also serve as predictive models. These models utilize a relatively small experimental database and are able to predict phenomenon outside of that database. Predictive models are often useful for the validation of a theory. A classical example is the validation of Einstein's Theory of Relativity, from the observation of the bending of light by the gravity of the Sun during an eclipse.

Predictive models are also useful to simulate conditions that are interesting but for some reason have not been achieved. For instance, the Grand Unification Theory is often used to predict the existence of subatomic particles which may be verified when technology is sufficiently advanced.

In this work we present a mechanistic model of metabolism. The model is used as a predictive tool to extract engineering principles for bioprocess design. We also provide experimental verification to the model.

## 2.1 Dynamic Analysis of Metabolism

Metabolism consists of a set of reactions that process the various metabolites within the cell. A material balance can be written for each metabolite within the

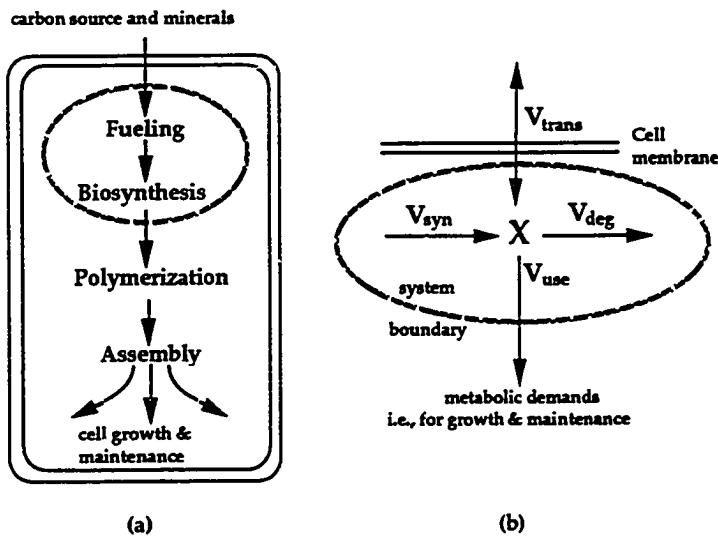


Figure 2.1: Schematic of flux balances around the fueling and biosynthetic reactions.

(a) Fueling and biosynthetic reactions form the system of reactions chosen for metabolic analysis. The products of the metabolic network chosen are processed further and ultimately used for cell growth and maintenance.

(b) The fluxes affecting the concentration of a metabolite,  $X$ , in the cell.

cell to yield the equations interconnecting the various metabolites. The four types of fluxes affecting the concentration of a particular metabolite pool are listed below.

1. Uptake or secretion of the metabolite across the cell membrane,  $V_{trans}$ .
2. Synthesis of the metabolite from other intermediates,  $V_{syn}$ .
3. Consumption of the metabolite by the cell for maintenance or for the production of other cellular components,  $V_{use}$ .
4. Degradation of the metabolite,  $V_{deg}$ .

The fluxes affecting a hypothetical metabolite,  $X$ , are depicted in Figure 2.1b. Equating accumulation to the net rate of production the material balances for the metabolites can be written as:

$$\frac{dX_i}{dt} = V_{syn} - V_{deg} - [V_{use} - V_{trans}] \quad (2.1)$$

Normally, the uptake or secretion flux,  $V_{trans}$ , as well as the growth and maintenance requirements,  $V_{use}$ , are known. Equation (2.1) therefore, reduces to:

$$\frac{dX_i}{dt} = V_{syn} - V_{deg} - b_i \quad (2.2)$$

where  $b_i$  is the net transport into our defined system.

Generally, for a metabolic network that contains  $n$  metabolites all the transient  $n$  material balances can be represented by a single matrix equation.

$$\frac{d\mathbf{X}}{dt} = \mathbf{S} \cdot \mathbf{v} - \mathbf{b} \quad (2.3)$$

where  $\mathbf{X}$  is the vector of metabolite amounts per cell,  $\mathbf{S}$  is the  $n \times m$  stoichiometric matrix,  $n$  is the number of metabolites,  $\mathbf{v}$  is the vector of  $m$  reaction fluxes, and  $\mathbf{b}$  is the vector of known consumption rates, waste production rates, and secretion rates. The element  $S_{ij}$  is the stoichiometric coefficient of the  $i^{th}$  compound in the  $j^{th}$  reaction indicating the amount of the  $i^{th}$  compound produced per unit flux of the  $j^{th}$  reaction. Equation (2.3) can be solved for the complete dynamic response of the metabolite pools,  $\mathbf{X}$ , if the reaction rates,  $\mathbf{v}$ , are known.

Several attempts to systematically model metabolic dynamics have been carried out (e.g., see reviews in [21, 24, 63]) but they have been hampered by the lack of kinetic and regulatory information on the function of all enzymes in a particular cell. The only cell for which a comprehensive kinetic metabolic model currently exists is the mature human red blood cell [34, 41]. The fully differentiated mature red blood cell is somewhat unique in that it has lost its nucleus. The cell therefore

does not have any genetic regulation and a dynamic response can be determined simply from a knowledge of the enzyme kinetics. The red blood cell model thus enables several kinetic studies but these investigations and the conclusions derived from them, although useful, are necessarily limited given the highly differentiated state and specialized function of the red blood cell.

**Other Methods** Much effort has also been devoted towards the development of a theoretical framework for the analysis of metabolic regulation, mostly through the use of logarithmic sensitivity coefficients [25, 35, 66]. Extensive literature exists on this topic and useful overviews are available [11, 79]. The applicability and usefulness of these theories remain to be examined within the context of realistic metabolic models although some recent progress has been made [57].

## 2.2 Steady State Analysis

The time constants of metabolic reactions are typically very fast compared to the time constants of cell growth and culture fermentation. Therefore, one can study the much simpler metabolic steady state as an approximation of the true physical system.

**Elemental Analysis** At the simplest level, steady state analysis involves a mass balance of the major elements present in metabolites. For relatively simple metabolic networks the elemental balance technique can provide useful process information. An excellent review is available in literature [55].

**Stoichiometric Analysis** Much more powerful is stoichiometric steady state analysis for determining the metabolic reaction fluxes. In the stoichiometric analysis

we consider the steady state form of Equation (2.3) thus eliminating the time derivative to obtain Equation (2.4). Note that the steady state assumption also removes metabolite concentration as a variable from the mathematical formulation of the metabolic network. Time as a variable is also not present. Only the stoichiometry of the reactions,  $S$ , and the metabolic demands of the cell,  $b$ , are required in order to solve the system to evaluate the metabolic flux distribution.

$$S \cdot v = b \quad (2.4)$$

To determine the metabolic fluxes in the cell we need to solve Equation (2.4) ( $S$  and  $b$  are known from the metabolic network). For a square  $S$ -matrix, *i.e.* number of rates equal to the number of metabolites or material balances, the solution to Equation (2.4) is unique and is easily obtained by Gaussian elimination.

However, typically the number of rates is greater than the number of mass balances (*i.e.*,  $m > n$ ) resulting in an underdetermined set of equations. Consequently, the stoichiometry of the metabolic network does not uniquely specify the fluxes through the cell's pathways, so that the number of possible flux distributions allowed by the stoichiometry is infinite. The particular flux distribution chosen by the cell is a function of regulatory mechanisms within the cell that determine the kinetic characteristics of cellular enzymes as well as enzyme expression.

Thus, the flux balances, Equation (2.4) determine the wider limits within which metabolic pathways can be utilized. The stoichiometrically set limits are then further narrowed by the kinetic and regulatory function of metabolic enzymes. These two limits placed on metabolic function are illustrated in Figure 2.2.

In the absence of detailed knowledge of enzyme regulation, one can estimate the metabolic flux distribution by specifying an optimization criteria. Given a linear

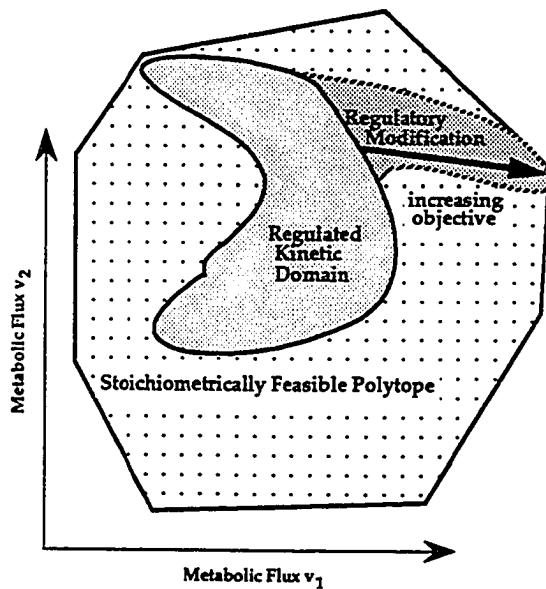


Figure 2.2: The stoichiometrically feasible domain of steady state reaction fluxes within the cell describe a wider limit of metabolic behavior. A two dimensional region is shown to schematically illustrate the stoichiometrically feasible values of two hypothetical metabolic fluxes. The regulated domain of fluxes chosen by the cell forms a subset of the stoichiometrically feasible region.

objective function, an underdetermined set of equations can be solved using linear optimization techniques. The resulting solution is that of a metabolic network biased towards optimizing the postulated objective. Objectives can thus be chosen to explore the optimum response of the metabolic network within the cell.

The history of stoichiometric based metabolic analyses is relatively short. Linear programming has been used to study the stoichiometric constraints on metabolic networks [91]. Where the number of metabolic fluxes exceeds the number of flux balances this approach has been applied to adipocyte metabolism [18]. Acetate secretion from *E. coli* under ATP maximization conditions [46] and the metabolic behavior of hybridoma cells [72] have also been studied.

Linear optimization can also be helpful in determining the capability of the metabolic network of the cell. One can determine the optimum flux distribution for maximizing product in a commercial process. Knowledge of the optimum solution opens up the possibilities of modifying the cell as well as the process in order to approach the optimum flux distribution thus maximizing yield.

### 2.3 Linear Optimization

The solution to Equation (2.4) can be formulated as a linear programming problem, in which one finds the solution that minimizes a particular objective. Mathematically, this minimization is expanded as:

Minimize  $Z$

$$\text{where } Z = \sum c_i \cdot v_i = \langle \mathbf{c} \cdot \mathbf{v} \rangle \quad (2.5)$$

where  $Z$  is the objective which is represented as a linear combination (as defined by the weights in the vector  $\mathbf{c}$ ) of the fluxes  $v_i$ . Appropriate representation of  $Z$  enables

us to formulate physiologically meaningful objective functions. For example, maximizing the production of a particular metabolite,  $X_i$ , could be represented by the objective function as minimizing the negative of the secretion rate for the metabolite, Equation (2.6). The weight  $c_i$  for the synthesis rate for that metabolite would have an arbitrary positive value while all other weights would have a zero value. A function maximizing growth could be formulated as maximizing the growth flux,  $v_{gro}$ , Equation (2.7). The growth flux is defined as drains on the various metabolite pools required for growth, Equation (2.8). The vector  $d_M$  represents the stoichiometric ratios in which metabolites are used for growth.

$$\text{Minimize } Z = -c_i \cdot (V_{i,syn} - V_{i,deg}) \quad (2.6)$$

$$\text{Minimize } Z = -v_{gro} \quad (2.7)$$

$$\text{where } growth = v_{gro} \cdot < d_M > \quad (2.8)$$

The minimization in Equation (2.5) is subject to the constraints:

$$S \cdot v = b \quad (2.9)$$

$$v_i < \alpha_i, i = 1, 2, \dots, m \quad (2.10)$$

The first set of constraints is simply that of the steady state material balances, from Equation (2.4). The second set of constraints introduces a vector of optional parameters,  $\alpha$ , which represents the maximum fluxes allowable through the reactions  $v$ . A maximum limit on fluxes may be introduced due to enzymatic saturation, mass transfer limitations etc.. Fluxes to specific pathways can also be fixed if an experimentally determined value is known. Furthermore, specific drains can also be fixed to provide for growth and maintenance of the cell. Therefore, we have additional

constraints corresponding to the multiple demands of metabolites for growth and maintenance. Any flux in the solution,  $\mathbf{v}$ , of the linear programming problem can also be restricted to be positive, thus permitting us to incorporate irreversible reactions. For the present analysis we have restricted all fluxes to be positive and reversible reactions have been permitted as two opposing fluxes. Thus, the above mathematical formulation has a lot of flexibility that enables us to determine the capabilities of a metabolic network.

### 2.3.1 Shadow Prices

For each linear programming problem defined by Equations (2.5-2.10), there exists an adjoint optimization, called the dual optimization problem [44, 50]. For the above described primal linear optimization problem the dual problem is formulated as another linear optimization.

$$\text{maximize: } \gamma \cdot \mathbf{b} \quad (2.11)$$

$$\text{subject to: } \gamma \cdot \mathbf{S} \leq \mathbf{c} \quad (2.12)$$

The solution to the dual problem yields the values of the dual variables,  $\gamma$ . Interpretation of the dual variables as the shadow prices results in an interesting definition:

$$\text{shadow price } = \gamma_i = -\left. \frac{\partial Z}{\partial b_i} \right|_{Z=Z_{\max}} \quad (2.13)$$

The shadow price is thus the sensitivity of the objective function with respect to each constraint. In other words, if the supply to the cell of compound  $i$  is increased by  $\Delta b_i$  units, then the objective value  $Z$  will decrease by  $\gamma_i \cdot \Delta b_i$  units. The shadow price thus indicates the marginal conversion of a metabolite into the objective.

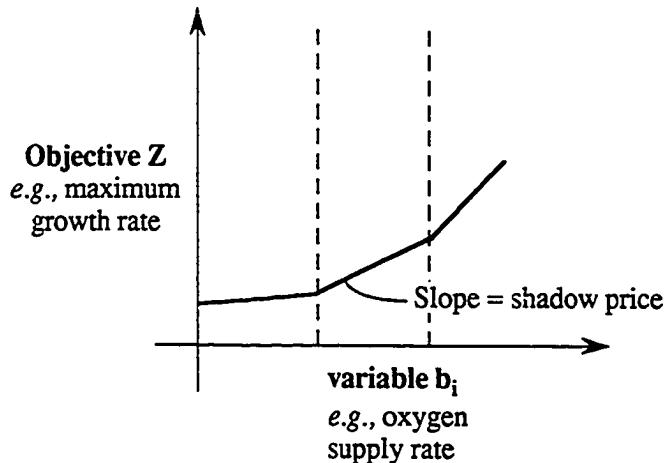


Figure 2.3: Schematic representation of the shadow price which effectively provides the increase in objective possible by providing additional quantities of the  $i^{th}$  metabolite or nutrient.

The shadow price is schematically depicted in Figure 2.3. In this particular example the growth rate increases in a piece-wise linear manner as the oxygen supply increases. At a particular oxygen supply rate the shadow price for oxygen indicates the marginal or local increase in growth rate that can be achieved by the addition of oxygen. Thus, the shadow price indicates the utility or usefulness of oxygen in accelerating the growth rate.

In the following chapters we present the stoichiometric framework for *E. coli* metabolism and apply the linear optimization tools described above.

## CHAPTER III

### PREMISE FOR A FLUX BALANCE MODEL

The flux balance model for *E. coli* consists of three components: the metabolic network, the metabolic requirements for growth, and the linear optimization analysis. We begin by formulating and presenting the catabolic and anabolic reaction network for the bacterium. The *E. coli* cell is well known in terms of its chemical composition and therefore the metabolic demands that are imposed by cell growth and function can be defined. These metabolic demands have recently been thoroughly documented [32] and are reproduced here for use in the flux balance model. Finally, we address the application of the linear optimization techniques to the flux balance model.

**Catabolism of Substrate** Catabolic reactions in the bacterium serve to degrade a variety of carbon sources into a pool of twelve biosynthetic precursors and three cofactors. The reactions are often classified as: the glycolytic pathway, the pentose phosphate pathway, the tricarboxylic acid TCA cycle with the glyoxylate shunt, the anaplerotic reactions, and the electron transfer system. These reactions are also responsible for providing the interconvertibility between the biosynthetic precursors and cofactors.

Several useful overviews are available [32, 48, 51, 53] that describe the catabolic

network and its ability to degrade several substrates. The catabolic network used for this study is described in literature [87] with a detailed documentation.

**Anabolism to Building Blocks** Anabolic or biosynthetic reactions serve to utilize the biosynthetic precursors and cofactors to synthesize the building blocks of the cell. The reactions involved in biosynthesis alongwith their associated stoichiometry are available from several sources [28, 33, 48, 51, 85].

### 3.1 Network Formulation

The biochemical pathway information described above can be presented in the mathematical form of Equation (2.4). Several reductions of the flux balance equations are possible without altering the mathematical nature of the problem.

**Stoichiometric reduction** The stoichiometric representation of metabolism can be made compact by using simple rules thus reducing the dimensions of  $S$ . The application of these rules does not alter the basic structure of the catabolic network. The rules for network reduction are:

1. Intermediates of reaction pathways, which have only one route for generation and one route for consumption, can be ignored and one flux may denote the flux producing the intermediate and the flux consuming it. For example, the set of reactions producing R5P from G6P can be represented by a single flux converting G6P to R5P.
2. Metabolites that are freely interconvertible without the involvement of a third metabolite can be represented by a single metabolite. An example may be

found in the enzymatically convertible compounds, DHAP and G3P. They can be represented in the metabolic network as T3P.

3. Compounds of no interest here need not be included in the reaction network. They are then implicitly assumed to have an infinite source or sink. Thus, in some situations we may choose not to keep track of O<sub>2</sub> in the reaction network. Cofactor molecules such as coenzyme A and NAD which are carriers for specific molecular species have their flux balances intrinsically balanced, and can therefore be ignored in the network. Inclusion of such metabolites only leads to the generation of a dependent row in the matrix.
4. Although the reactions are not described here, the high energy phosphate bond of the various nucleotides can be traded among the various nucleotides. Therefore the utilization of the high energy bond associated with a non-adenosine nucleotide as well as the second phosphate bond of adenosine phosphate have been considered stoichiometrically equivalent to the third phosphate bond of ATP.

Further, only physiologically occurring reactions are included. Thus, although the reaction catalyzed by pyruvate kinase is reversible *in vitro*, only the physiological forward reaction producing pyruvate is included in the network.

**The *E. coli* metabolic network** Applying the above rules to the metabolic reactions described above we have derived the catabolic network, Figure 3.1, and the biosynthetic network, Figure 3.2. Reaction fluxes are denoted by letter codes which represent the enzymes or their genetic loci. While reaction coupling to cofactors is not shown in these figures, the stoichiometry of coupling is included in the mathematical

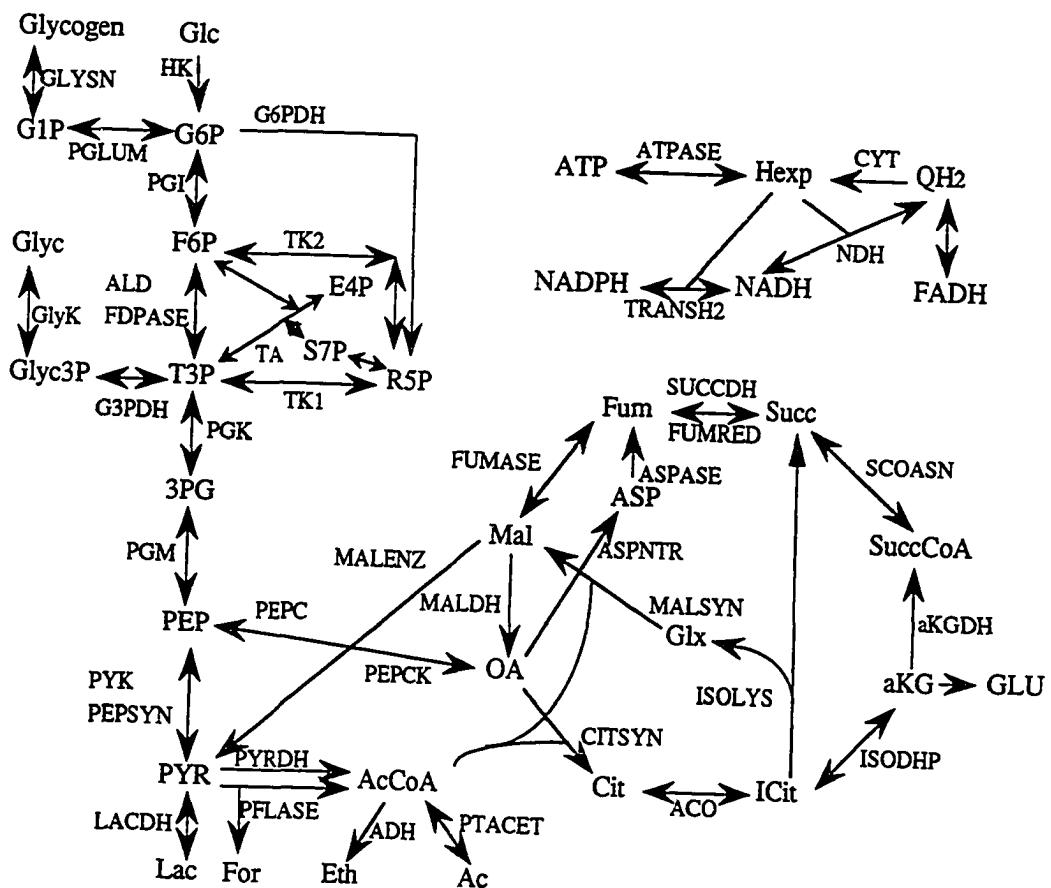


Figure 3.1: The catabolic reaction network.

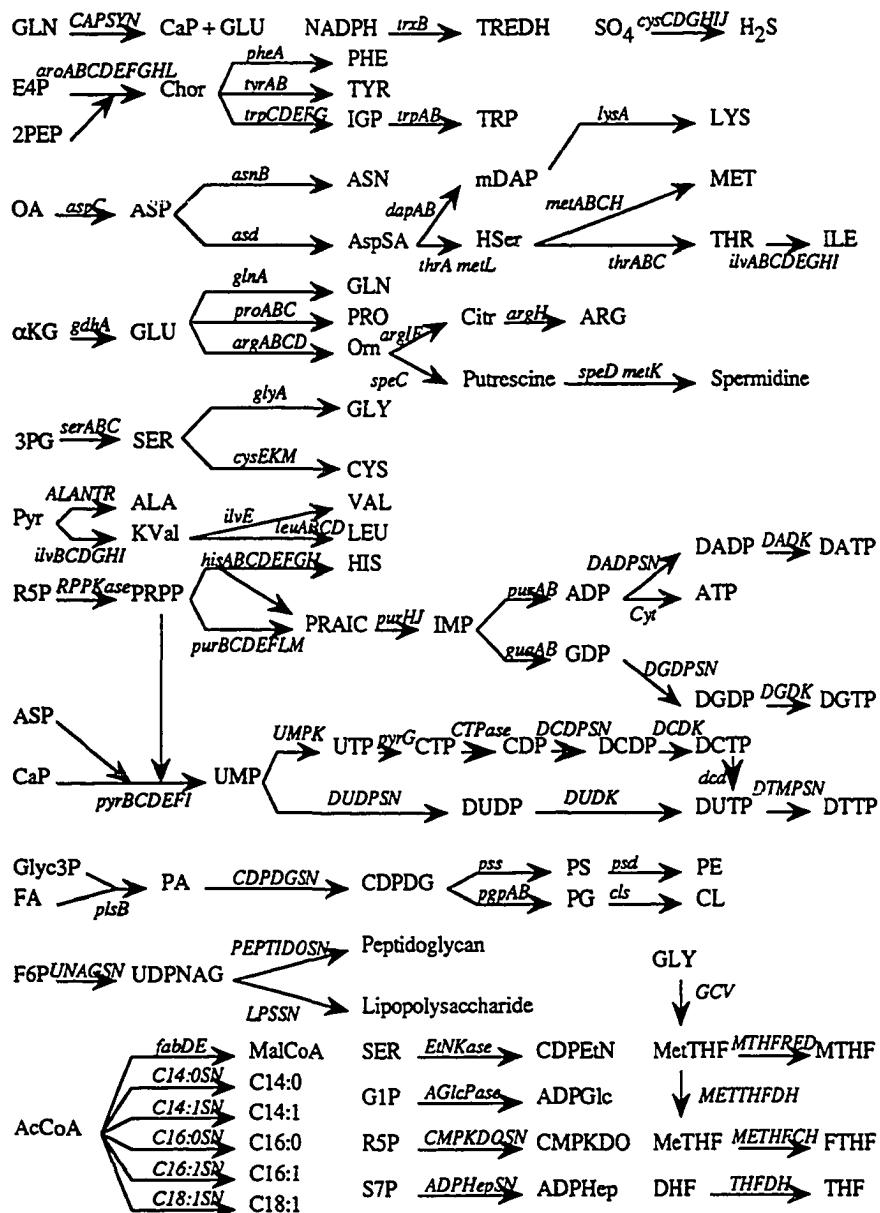


Figure 3.2: The biosynthetic reaction network.

specification of the network. Mathematical representations of the metabolic network in the form of Equation (2.4) are available for the catabolic network [87] as well as for the biosynthetic network [85].

**Network Flexibility** A few observations about the flexibility of the stoichiometric model are in order. A particular solution for steady state pathway fluxes can result in the generation of surplus energy or reductive power. In the network formulated above surplus energy can be eliminated by several metabolic loops. A well known example of a futile cycle may be found in the conversion of F6P into T3P and the reconversion back to F6P which has the net effect of hydrolyzing one high energy phosphate bond.

Similarly surplus energy associated with the proton gradient across the membrane can be transformed into a high energy phosphate bond of ATP by the enzyme ATPase, and then dumped into a futile cycle. Surplus reductive power in the aerobic network can be transformed into the transmembrane proton gradient using oxygen as the electron acceptor. The transmembrane proton gradient is converted into high energy phosphate bonds and dissipated through the futile cycles. Of course, anaerobic conditions would not allow the transformation of surplus reductive power into energy and some other sink for electrons is required.

The network can deal with a surplus of a particular metabolite by two mechanisms. Dissipation can occur through a complete oxidation to  $\text{CO}_2$  and water, or by a secretion pathway that would have to be incorporated in the network. Thus, the basic network that we have formulated accounts for known fueling reactions in *E. coli* and has the metabolic flexibility that the bacterial cell possesses.

Table 3.1: Requirements of biosynthetic precursors and cofactors for growth, from [32, 52]. Metabolic requirements are in mmol *per* 1 g biomass yield.

Metabolite	$d_M$	Metabolite	$d_M$	Metabolite	$d_M$
ATP	41.257	G6P	0.2050	3PG	1.4960
NADH	-3.547	F6P	0.0709	PEP	0.5191
NADPH	18.225	R5P	0.8977	PYR	2.8328
		E4P	0.3610	AcCoA	3.7478
		T3P	0.1290	OA	1.7867
				$\alpha$ KG	1.0789

### 3.2 Metabolic Demands for Growth

As shown in Figure 1.2 all the reactions occurring within the bacterium can be logically divided into the categories of fueling, biosynthetic, polymerization, and assembly reactions. For metabolic analysis we have studied the fueling and biosynthetic reactions. The input to the system therefore consists of the carbon source and minerals and any particular metabolite that we may specify. The output from the metabolic network under consideration consists of the metabolites required for growth, maintenance, and possibly secretion.

The approximate chemical composition for *E. coli* B/r is known in terms of the biosynthetic precursors as well as the building blocks [32, 52], (see Tables 3.1,3.2). The composition is interpreted as a metabolic requirement for growth, Equation (3.1), and thus specifies a growth flux in the flux balance model.

$$\sum_{all M} d_M \cdot M \xrightarrow{v_{gro}} biomass \quad (3.1)$$

Table 3.2: Requirements of building blocks for growth, from [32, 52]. Metabolic requirements are in mmol *per 1 g* biomass yield.

Metabolite	Demand	Metabolite	Demand	Metabolite	Demand
ALA	0.488	PHE	0.176	DGTP	0.0254
ARG	0.281	PRO	0.210	DCTP	0.0254
ASP	0.229	SER	0.205	DTTP	0.0247
ASN	0.229	THR	0.241	Phosphatidyl serine	0.00258
CYS	0.087	TRP	0.054	Phosphatidyl ethanolamine	0.09675
GLU	0.250	TYR	0.131	Cardiolypin	0.00645
GLN	0.250	VAL	0.402	Phosphatidyl glycerol	0.02322
GLY	0.582	~P (Energy)	21.97	Lipopolysaccharide	0.00785
HIS	0.090	ATP	0.165	Peptidoglycan	0.0276
ILE	0.276	GTP	0.203	Glycogen	0.154
LEU	0.428	CTP	0.126	One Carbon	0.0485
LYS	0.326	UTP	0.136	Putrescine	0.0341
MET	0.146	DATP	0.0247	Spermidine	0.007

### 3.3 Applicability of Linear Programming to Analysis of Metabolic Networks

Linear programming as a tool for the analysis of metabolic networks has several advantages as listed below.

1. The information required, the stoichiometry of the cellular reactions, is fairly well-known for most organisms.
2. Limited thermodynamic information is included in the form of reversibility or irreversibility of the reactions. Linkage to energy in the form of the high energy phosphate bond of ATP provides some additional thermodynamic information.
3. The problem formulation allows the incorporation of additional knowledge of maximal activities of specific metabolic reactions. Experimental information can be included by placing limits on the fluxes allowable through specific pathways in the metabolic network.
4. Linear optimization allows the identification of the reaction pathways for optimal metabolism. We are able to determine the optimal pathways used to fulfill specific metabolic needs of the cell such as energy, reductive potential, etc.
5. The relative value of various metabolites and substrates in achieving an objective can be determined. The objective may be growth or production of a commercial intermediate. Identification of optimal substrate compositions as a part of commercial process design is easily amenable to this approach.
6. The optimum flux distribution for a desired goal (such as the commercial production of a metabolite) can be determined. We can thereby identify the

important reactions which are subject to genetic engineering to achieve the production goals.

There are, however, some limitations of this approach to metabolic modeling. We do not consider the regulation of enzymes catalyzing the cellular reactions. The solution obtained may therefore not be acceptable to the regulatory mechanisms of the cells. Of course, genetic engineering can be used to obtain organisms with modified metabolic regulation. The second limitation is the absence of explicit accounting of metabolic concentrations. We are therefore unable to predict the concentrations of the metabolites within the cell. Also we have not incorporated any thermodynamic information in the form of rate expressions. Time as a variable is therefore absent and we are unable to make any predictions as to the time constants of the cellular processes. Lastly, the discovery of new pathways [16] and stoichiometries or a modification of the existing pathways may require further modifications to the network formulated here.

The present literature on the stoichiometric analysis of metabolic networks, however shows that this approach is surprisingly powerful [72, 73, 74, 75] and is very attractive since little and well known information is required. Our goal is to use this biochemically simple yet reliable and thus attractive approach to determine the capabilities of the metabolic network.

### **3.4 Mathematical Algorithms and Computer Implementation**

The metabolic system under consideration for linear optimization is on the order of 100 variables and constraints. For a system of this order the simplex based algorithms give fast and accurate solutions. The computer implementation has therefore

utilized available packages incorporating simplex based algorithms.

### 3.4.1 Simplex Algorithm

Given a basic feasible solution the simplex based algorithms try to find a better solution in the multidimensional space defined by the constraints. A finite number of iterations eventually lead to the optimum solution. The initial solution can be provided manually or generated using additional algorithms. The program used in this study is capable of generating the initial solutions itself.

**Software** Linear optimization has been implemented using the computer code, simple, which is a simplex based FORTRAN routine obtained from the 'genlib' collection of the CERN COMPUTER CENTER PROGRAM LIBRARY [22]. It incorporates anticycling rules using the method of small displacements by Charnes. Graphic displays of the solution were obtained using the program, 'mapper', developed by Tom Groshans at The University of Michigan.

**Hardware** All the computer programs used have a FORTRAN source code and are therefore transportable to an extent. For the present study the programs were implemented on the Apollo and Hewlett Packard workstations.

## CHAPTER IV

# CAPABILITIES OF THE CATABOLIC NETWORK

After gaining an appreciation for metabolic analysis in Chapter 2 and the formulation of the flux balance model in Chapter 3 we are in a position to study the characteristics of the metabolic model. We now use the catabolic network to investigate the ability of *E. coli* to make the three key metabolic cofactors (NADH, NADPH and ATP) as well as the twelve basic biosynthetic precursor molecules during aerobic growth on glucose. We seek to determine optimal pathway utilization, and maximal yields of each metabolite and to assess the systemic constraints that *E. coli* is faced with in their production.

### 4.1 Optimal Production of Cofactors

The cofactors ATP, NADH and NADPH play a central role in bacterial metabolism. We now use linear optimization to determine the maximal production of these cofactors. The type of objective function used has been defined in Equation (2.6). The solutions for maximum ATP and NADPH production are shown in Figures 4.1a and b respectively. These solutions represent the maximal capability of the *E. coli* catabolic network to produce metabolic energy and biosynthetic redox.

Table 4.1: Maximum energy and redox generation on glucose using *E. coli* fueling reactions. The PPS flux is given as the percentage of carbon flow through the oxidative branch of the pentose phosphate pathway with a maximum of six times the glucose input or 600%.

Metabolite	Yield	PPS	ATP shadow	Constraint
		price		
ATP	18.667	0%	-1	-
NADH	11.573	471%	-0.214	Energy
NADPH	11.000	540%	-0.5	

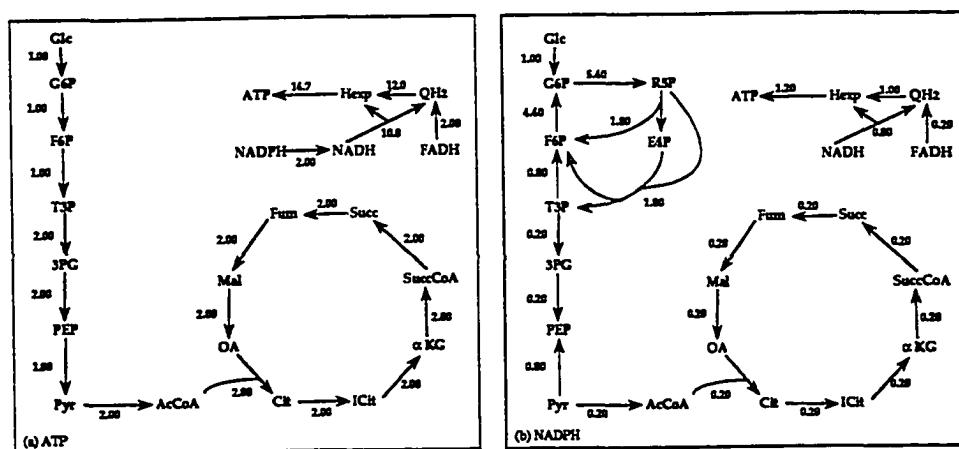


Figure 4.1: Flux distributions for maximal (a) ATP and (b) NADPH production.

The solutions thus represent important and fundamental biological constraints. The key characteristics of these solutions are summarized in Table 4.1.

We find that a maximum of 18.7 moles of ATP can be generated *per* mole of glucose oxidized to CO<sub>2</sub> and water, (see Table 4.1). The optimal solution does not use the pentose phosphate pathway. The use of the TCA cycle for optimal production of ATP is consistent with the general belief that the pentose phosphate pathway functions primarily to generate pentoses and NADPH and is not cycled to produce metabolic energy in the form of ATP. The glyoxalate shunt is believed to have the function of producing TCA cycle intermediates. Consistent with this belief the glyoxalate shunt is not used for the optimal production of ATP. Inactivating the TCA cycle and forcing the flux through the glyoxalate bypass lowers the maximum ATP yield on glucose to 16.7 mol/mol.

The complete oxidation of glucose requires a sink for 24 electrons. Therefore a potential reduction of 12 redox carriers can be obtained. However, in the presence of systemic constraints the base network can produce a maximum of only 11.6 NADH. The difference can be explained based on the ATP shadow price in the optimal NADH solution:

$$\frac{\partial NADH}{\partial ATP} \simeq \frac{\Delta NADH}{\Delta ATP} = 0.214 \quad (4.1)$$

Thus, the additional amount of ATP needed to make the full amount of 12 NADH from glucose is

$$\Delta ATP = \frac{\Delta NADH}{0.214} = \frac{12 - 11.573}{0.214} = 2 \quad (4.2)$$

These two ATPs correspond to the cost of resynthesizing the PEP using PEP synthase from the pyruvate formed during glucose phosphorylation by the phospho-

transferase system. Thus, when the whole network is considered one finds that it is not possible to generate 12 reduced NADH molecules from glucose due to the cost of glucose import.

Similiarly, the theoretical maximum reducing power that can be generated in the form of NADPH is 12 NADPH *per* molecule of glucose. The amount that can actually be generated considering the whole network is lower, or 11 NADPH *per* glucose molecule. As for NADH, the cost of glucose uptake prevents the basic fueling network from generating 12 molecules of NADPH. The additional ATPs required to yield the 12 NADPH molecules can be calculated from the shadow price for ATP:

$$\frac{\partial \text{NADPH}}{\partial \text{ATP}} \simeq \frac{\Delta \text{NADPH}}{\Delta \text{ATP}} \Rightarrow \Delta \text{ATP} = \frac{12 - 11}{0.5} = 2 \quad (4.3)$$

Again the recovery of PEP used for glucose import requires 2 ATPs. The difference in the ATP shadow price for the above two examples is due to the different optimal pathway utilizations for the production of NADPH and NADH.

From the above discussion it is apparent that energy is the constraining factor for maximal generation of reductive power. The function of the enzyme transhydrogenase, is to transfer reductive power between NAD and NADP. From the above optimal solutions we note that the maximal synthesis of NADPH does not utilize transhydrogenation. Biosynthesis requires reductive power in the form of NADPH while NADH is oxidized to produce energy under aerobic conditions [32]. The absence of a requirement of transhydrogenase for maximal NADPH generation in the catabolic network is experimentally indicated by the lack of a phenotype for transhydrogenase mutants [93].

From a study of optimal flux distributions we note that the oxidizing pathway of PPS is used only for redox generation. The optimal generation of the precursors E4P

Table 4.2: Maximum biosynthetic precursor yields from glucose.

Metabolite	Yield	Carbon	ATP shadow	Constraint
		conversion	price	
3PG	2	100%	0	
PEP	2	100%	0	
Pyr	2	100%	0	None
OA	2	133.3%	0	
G6P	0.908	90.8%	-0.046	
F6P	0.908	90.8%	-0.046	
R5P	1.08	90%	-0.055	Energy
E4P	1.33	88.7%	-0.068	
T3P	1.73	86.5%	-0.088	
AcCoA	2	66.7%	0	
$\alpha$ KG	1	83.3%	0	Stoichiometry
SuccCoA	1	66.7%	0	

and R5P, discussed next, occurs through the non-oxidative branch. Experimentally determined pathway utilizations [92] are found to agree with these observations of optimal flux distributions.

## 4.2 Optimal Production of Biosynthetic Precursors

All carbon sources are degraded into a minimum set of 12 biosynthetic precursors [32]. The biosynthetic reactions of the cell utilize these precursors to produce the monomers that go into making the macromolecular constituents of the cell. We

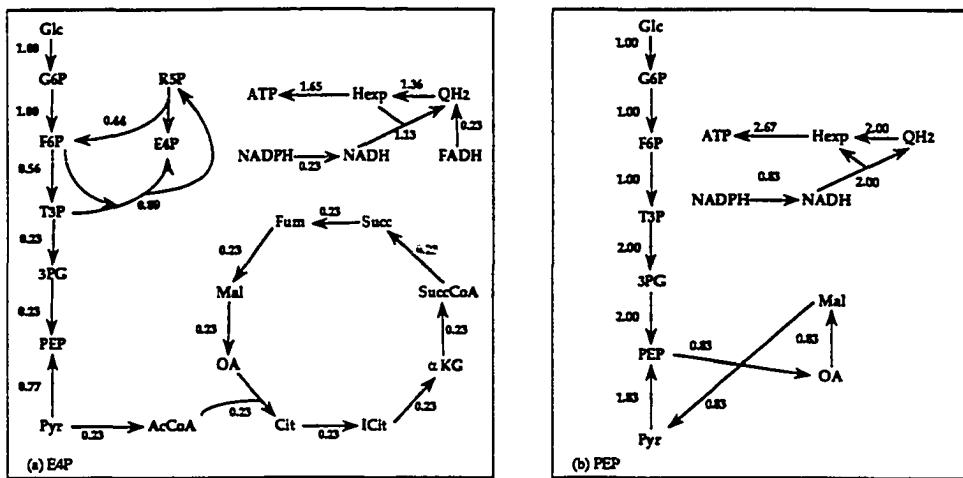


Figure 4.2: Flux distributions for maximal (a) E4P and (b) PEP production.

now determine the capability of the basic *E. coli* catabolic network to produce each of these precursors individually using glucose as the sole carbon source. The results from optimizing the production of each precursor molecule on glucose are summarized in Table 4.2. Full optimal solutions are shown in Figures 4.2a and b for the maximal production of E4P and PEP from glucose as examples of what the optimal solutions look like. The carbon conversion listed in Table 4.2 is the percentage of the carbon in glucose that ends up in the precursor molecule being produced.

The three glycolytic intermediates, 3PG, PEP and Pyr, can be produced with 100% carbon conversion. Their maximum yield from glucose has no energy related constraints (the shadow price for ATP is zero) and, in fact, a surplus of energy is generated which is dissipated through a futile cycle.

The surplus ATP production is readily illustrated by an example. Consider the

generation of PEP from glucose, Figure 4.2b. To make two PEP from glucose, two ATP are required for glucose uptake, one ATP is required for the PFK reaction and two ATP are produced by PGK. Thus, direct stoichiometric coupling of ATP to PEP production results in the consumption of one ATP. However, two NADH are also produced which subsequently yield 2.66 ATP upon oxidation through the electron transport system. Thus, an overall surplus of 1.66 ATP *per glucose* is produced. Note that this amount is dissipated via a PEP→OA→Mal→Pyr→PEP cycle whose net effect is the dissipation of two ATPs, leading to a futile cycle flux of  $1.666/2 = 0.833$ . The network possesses several other equivalent futile cycles which could also be used to generate the same result.

Optimal production of oxaloacetate results in a carbon recovery in excess of 100%. A carbon conversion of 133 % is possible because of the CO<sub>2</sub> fixing reaction catalyzed by PEPC. There are no energy limitations and the TCA cycle is not used. In the base fueling network we do allow unlimited access to CO<sub>2</sub>. Restricting the CO<sub>2</sub> availability reduces the carbon yield to 100%. ATP overproduction is dissipated through a futile cycle as just discussed for PEP.

**Constraints on the production of biosynthetic precursors** In attempting to maximize the production of a metabolite, the metabolic network may be confronted with systemic constraints that prevent a 100% carbon conversion from substrate to metabolite. For the twelve biosynthetic precursors the cell encounters two constraints: energy or network stoichiometry. These constraints are listed in Table 4.2 for the biosynthetic precursors. Note that redox constraints do not appear in the production of precursors. Reductive power is primarily required for the biosynthesis of monomers [32], (see Chapter 6).

The constraint for a particular optimization can be determined by studying the shadow prices, §2.3.1, obtained from the dual optimization problem. Energy constraints are evidenced by non-zero shadow prices of ATP. Stoichiometric limitations are indicated by less than 100% carbon conversion and the absence of an energy constraint.

(i) *Energy.* The monophosphate sugars, G6P, F6P, R5P, E4P and T3P, cannot be produced at a 100% carbon conversion, Table 4.2. Energy is a constraint for all five cases and some carbon must be oxidized fully to provide the required energy. As illustrated above for optimal cofactor production from glucose, the ATP shadow price brings this fact out clearly. For G6P we have that  $\frac{\partial G6P}{\partial ATP} = 0.046$ , therefore 100% carbon conversion requires

$$\Delta ATP = \frac{\Delta G6P}{0.046} = \frac{1 - 0.908}{0.046} = 2 \quad (4.4)$$

additional molecules of ATP which correspond to the cost of importing and phosphorylating glucose. A similar calculation for T3P yields:

$$\Delta ATP = \frac{\Delta T3P}{0.088} = \frac{2 - 1.73}{0.088} = 3 \quad (4.5)$$

Thus not only are 2 ATPs required for PEP recovery but also a third ATP molecule is needed in the PFK reaction. E4P and R5P are produced by the non-oxidative branch of the pentose phosphate pathway in an attempt to reduce loss of carbon through decarboxylation in the oxidative steps. Again the reduced carbon conversion can be readily explained based on the ATP shadow price. For E4P the additional ATP required for 100% carbon conversion is:

$$\Delta ATP = \frac{\Delta E4P}{0.068} = \frac{1.5 - 1.33}{0.068} = 2.5 \quad (4.6)$$

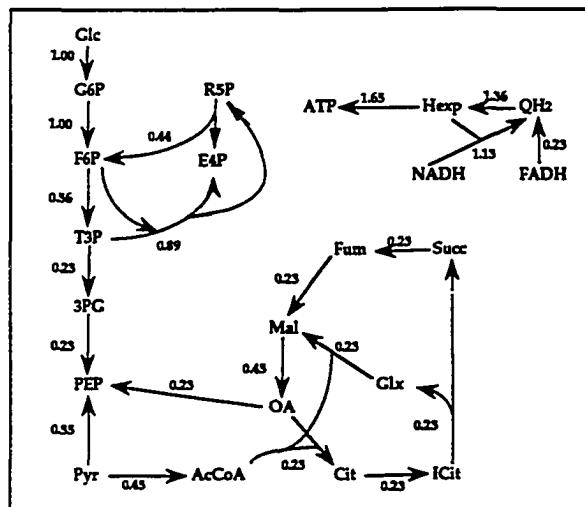
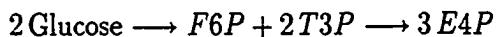


Figure 4.3: Alternate flux distribution for maximal E4P production utilizing the glyoxalate shunt instead of the TCA cycle.

This ATP requirement is consistent with the stoichiometry of E4P synthesis from glucose through the non-oxidative branch.



As discussed above F6P requires 2 ATPs (similar to G6P) and 2 T3Ps require 3 ATPs. Thus, 5 ATPs are required to synthesize 3 E4Ps from two glucose molecules which is equivalent to 2.5 ATP per glucose molecule.

The optimal production of the monophosphate sugars is therefore clearly constrained by energy. A sample flux distribution is displayed in Figure 4.2a, for E4P production. The solution shows the conversion of Pyr to PEP by the enzyme PEP-synthase. The energy required for the reaction is optimally produced by the TCA

cycle coupled with the ETS. As shown earlier, the TCA cycle is the optimal pathway for the generation of ATP. However there is an alternative pathway to convert Pyr to PEP which is energetically equivalent. If we restrict the complete TCA cycle the same yields for sugar monophosphates are obtained by the use of the glyoxalate shunt. A sample solution displaying the use of the glyoxalate shunt for E4P production is shown in Figure 4.3. The flexibility of the fueling network is such that it has more than one equivalent ways to make the ATPs necessary for making sugar phosphates.

This situation is one of multiple optimal solutions and the linear optimization program arbitrarily selects any one of the optimal solutions. The glyoxalate shunt provides an alternate route for the flux from Pyr to PEP using the enzyme PEP carboxykinase. Thus, the diversion of flux from the reaction catalyzed by PEP synthase results in a saving of energy which makes the net use of the glyoxalate shunt, energetically comparable to the TCA cycle for these special cases. However, simply for the production of energy the TCA cycle maintains an edge over the glyoxalate shunt as discussed earlier.

(ii) *Stoichiometry.* Acetyl CoA can only be produced with a 66.7 % carbon conversion and the optimal solution is not constrained by energy. The catabolic network possesses only one route for generation of acetyl CoA and it is by the de-carboxylation of pyruvate. Therefore, maximally a 2/3<sup>rd</sup> carbon recovery is possible, simply due to stoichiometric limitations.

The maximal production of the TCA intermediates  $\alpha$ -ketoglutarate and succinyl CoA is not limited by energy requirements. Yet the yield is less than 100%. Again the stoichiometric structure of the catabolic network forces the loss of carbons in essential reactions needed for the production of  $\alpha$ -ketoglutarate and SuccCoA.

Production of  $\alpha$ -ketoglutarate is associated with decarboxylation in the reactions catalyzed by pyruvate dehydrogenase and isocitrate dehydrogenase. The loss of carbons is mitigated to some extent by the use of the CO<sub>2</sub> fixing anaplerotic reaction catalyzed by PEP-carboxylase. Production of succinyl CoA is associated with a further loss of carbon by decarboxylation in the reaction catalyzed by  $\alpha$ -ketoglutarate dehydrogenase.

### 4.3 Summary

Stoichiometric analysis of metabolic networks is expected to yield rich dividends in terms of systematizing knowledge of metabolic systems, presenting us with the opportunity to explore the complex biochemical process that underlies the function of living cells. In this Chapter we have used linear optimization of the flux balance model to explore the boundaries of achievable metabolic performance. Linear optimization allows the identification of the optimal reaction pathway utilization to fulfill specific metabolic needs of the cell, such as the production of energy, reductive potential, or biosynthetic requirements. The relative value of various metabolites and substrates in achieving an objective can be determined using the shadow prices.

The set of catabolic pathways within the cell serve to degrade all substrates into a common set of biosynthetic precursors and cofactors. We have determined the capability of the bacterial catabolic network to produce these precursors and cofactors aerobically from glucose. All cellular synthesis utilizes the carbon skeletons provided by this pool of precursors while the cofactors provide energy and redox power to the cell. The maximal achievable yields of biosynthetic precursors and cofactors, therefore, represents fundamental determinants of metabolic performance.

Formulation of the flux balance model of *E. coli* catabolism has enabled the

determination of the capabilities of *E. coli* to make the three key cofactors and the twelve biosynthetic precursors. The results from these computations thus represent fundamental systemic constraints on *E. coli* metabolism and therefore important quantities in bacterial physiology. Confirmation of optimal pathway utilization by experimental literature confirms the correctness of the metabolic network formulated and demonstrates the applicability of the flux balance methods for metabolic analysis.

# CHAPTER V

## STOICHIOMETRIC OPTIMALITY OF GROWTH

In the previous Chapter we have determined the stoichiometrically constrained metabolic capabilities of *E. coli* to produce biosynthetic precursors and cofactors. We now examine the ability of the cell to make a balanced mix of cellular constituents corresponding to the biomass composition.

### 5.1 Optimal Growth Pattern on Glucose

Maximal biomass yield can be determined by computing the optimal solution for the objective function of Equation (2.7) that maximizes growth, for a fixed glucose uptake rate,  $V_{glc}$ . Growth is defined based on the biomass composition of Table 3.1. The maximum biomass yield is computed as :

$$Y_{max} = \frac{X}{Glc} = \frac{(V_{gro})_{max}}{V_{glc}} \quad (5.1)$$

The uptake of glucose can either be in mass or molar units. Both units are used here and to convert between them one simply uses the molecular weight of glucose, 180 g/mol or 0.18 g/mmol.

Using only the biosynthetic demands, the maximum yield of biomass computed

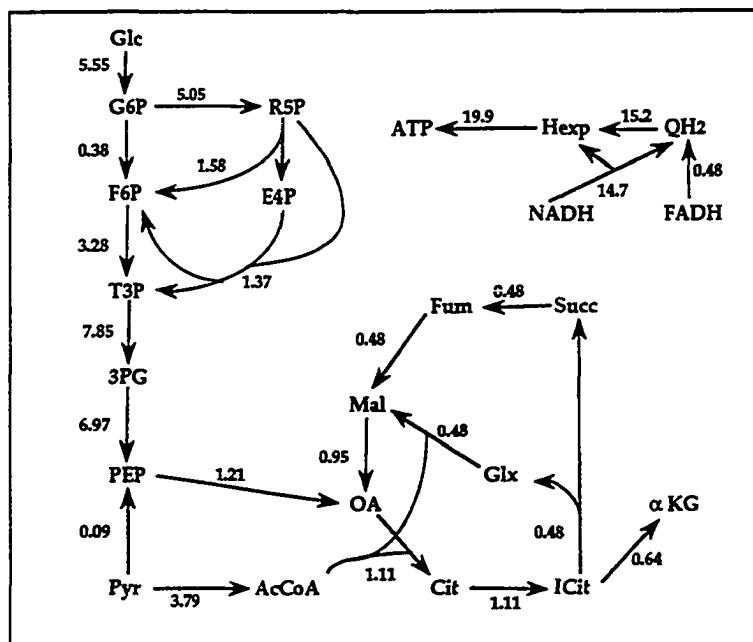


Figure 5.1: Optimal flux distribution for aerobic growth on glucose without any maintenance requirements. The maximal biomass production is 0.588 g DW/g Glc. All fluxes are relative to the input flux of glucose which is 5.55 mmol or 1g glucose.

is 0.588 g DW/g glucose consumed. Under aerobic conditions this yield is higher than typical mass yields of 0.4-0.5 g DW/g glucose and slightly higher than the maximum experimentally observed values of 0.54 g DW/g glucose consumed. This difference is not unexpected since the value of 0.588g DW/g glucose is obtained with no maintenance requirements [76]. The flux distribution that corresponds to the maximum biomass yield is shown in Fig. 5.1.

The calculation of the maximum biomass yield leads to two important questions; first, what limits the yield?, and second, how sensitive is the optimal solution to changes in fluxes, stoichiometry, biosynthetic demands, and the omission of maintenance requirements? We now address these questions.

### 5.1.1 Which Factors Constrain Growth?

Analysis of the shadow prices proved useful when we determined the limiting factors in the production of the biosynthetic precursors [87]. The shadow prices for the optimal solution presented in Fig. 5.1 are listed in Table 5.1. Since all the shadow prices are non zero, the addition of any one of the precursors will improve the optimal yield. This result is not surprising, since a precursor may not only be directly incorporated into biomass but also used to generate other biosynthetic precursor molecules. Therefore, each metabolite possibly has multiple values to the cell, such as energy, redox and the potential to make the various carbon skeletons needed.

Thus, the direct use of shadow prices is not useful when multiple simultaneous demands, *e.g.*  $V_{gro}$ , are imposed on the network. A new metric needs to be developed to identify key constraints for such complex and interconnected networks. We will base our development on glucose as the common reference.

Table 5.1: Metabolic demands of precursors and cofactors required for 1 g biomass yield, as well as shadow prices and the identification of growth constraints. The scaled shadow price is shown without considering maintenance as well as for a maintenance requirement of 4 ATP/Glc as described in the text.

Metabolite <i>M</i>	<i>d<sub>M</sub></i>	Shadow price	Metabolite	Scaled shadow	Scaled shadow
	(mmol)	no maintenance $-\frac{\partial X}{\partial M}$	yield $-\frac{M}{Glc}$	price $\sigma$ 0 ATP/Glc	price $\sigma$ 4 ATP/Glc
ATP	41.2570	-0.0049	-18.700	0.863	0.965
NADH	-3.5470	-0.0065	-11.600	0.714	0.921
NADPH	18.2250	-0.0092	-11.000	0.959	0.947
G6P	0.2050	-0.116	-0.908	0.994	0.961
F6P	0.0709	-0.116	-0.908	0.994	0.961
R5P	0.8977	-0.097	-1.080	0.991	0.962
E4P	0.3610	-0.079	-1.330	0.989	0.963
T3P	0.1290	-0.060	-1.740	0.990	0.964
3PG	1.4960	-0.049	-2.000	0.923	0.947
PEP	0.5191	-0.049	-2.000	0.923	0.947
PYR	2.8328	-0.039	-2.000	0.738	0.956
AcCoA	3.7478	-0.033	-2.000	0.615	0.781
OA	1.7867	-0.049	-2.000	0.923	0.816
$\alpha$ KG	1.0789	-0.072	-1.000	0.682	0.829
SuccCoA	-	-0.064	-1.000	0.600	0.741

**Formulation of a metric to assess biomass yield constraints** The factors constraining the biomass yield can be evaluated using the shadow prices of precursors for growth in combination with precursor yields on glucose. Conceptually, the approach taken here is to evaluate whether the cell is producing biosynthetic precursors and cofactors optimally while trying to achieve maximal biomass yield. If the cell is producing any one of its biosynthetic precursors optimally while maximizing biomass yield then we assume that this resource is limiting. Under those circumstances the optimal production of cell mass and the production of this metabolic resource are parallel, or aligned objectives. Conversely, if a metabolic resource is not being optimally produced from glucose we assume that the cell does not need to optimize the formation of this precursor in order to achieve maximal yield.

The criterion for the identification of yield limiting factors just outlined is comprised of three different processes. These are:

1. The optimal incorporation of glucose into biomass.



The optimal value for  $(X/Glc)$  is 0.588 g/g or 0.106 g/mmol in the absence of maintenance requirements.

2. The optimal yield of a precursor or cofactor from glucose.



These yields  $(M/Glc)$  were calculated in the preceding Chapter 4 and are listed in the fourth column of Table 5.1.

3. The marginal increase in the biomass yield from glucose if a particular precursor or cofactor were added:

$$M \xrightarrow{\partial X / \partial M} X \quad (5.4)$$

The quantity  $(\partial X / \partial M)$  is the shadow price listed in Table 5.1.

The biomass yield on glucose through the intermediate  $M$  is given by:

$$\left( \frac{M}{Glc} \right) \times \left( \frac{\partial X}{\partial M} \right) \quad (5.5)$$

In the criterion stated above, this product is to be compared to the biomass yield:

$$\left( \frac{X}{Glc} \right) \quad (5.6)$$

Since the product of Equation (5.5) is always less than or equal to the biomass yield of Equation (5.6) a dimensionless quantity

$$\sigma = \frac{\left( \frac{M}{Glc} \right) \times \left( \frac{\partial X}{\partial M} \right)}{\left( \frac{X}{Glc} \right)} \quad (5.7)$$

can be defined that assumes numerical values between zero and unity. If  $\sigma$  is unity for a given precursor then we conclude, based on the above line of reasoning, that it is biomass yield limiting. The quantity  $\sigma$  is therefore a shadow price scaled in terms of glucose units and gives a measure of the relative importance of the intermediate for achieving the objective of maximal biomass yield.

After this conceptual development we return to the optimal solution presented in Fig. 5.1. The numerical values for  $\sigma$  for all the precursors and cofactors are calculated and shown in Table 5.1. Interestingly, the sugar monophosphates (G6P, F6P,

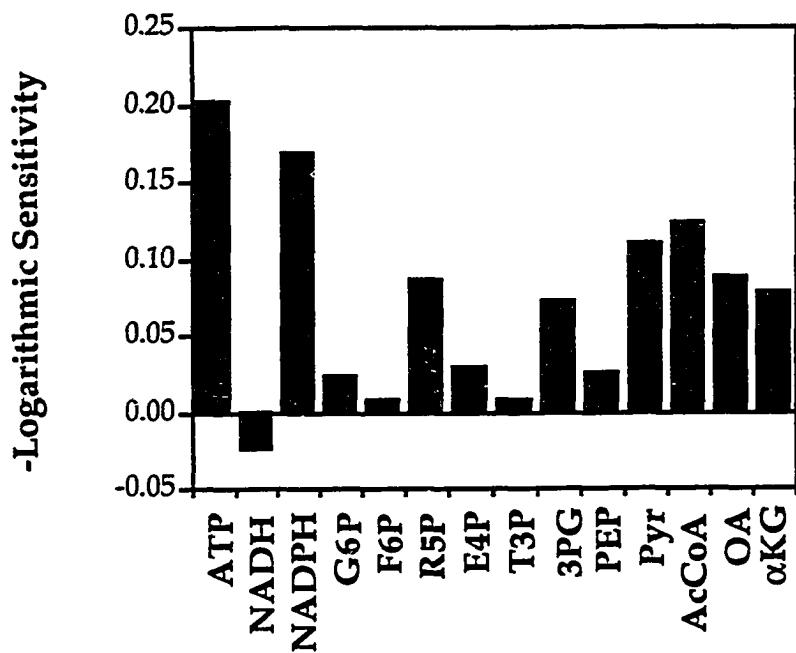


Figure 5.2: Logarithmic sensitivity of maximal biomass yield to the metabolic requirements for growth. The change in sign for NADH reflects a net production of NADH during growth.

T3P, R5P, E4P) have  $\sigma$  values closest to unity and are therefore biomass yield limiting factors under optimal growth on glucose although they are needed in relatively small amounts. In the case of the cofactors, NADPH is very desirable for biomass production as measured by  $\sigma$ , as is ATP. NADH is not an important constraint as measured by  $\sigma$ , which is as expected because the aerobic network uses NADH only to produce ATP, as shown in Fig. 5.1.

Another interpretation of the scaled shadow price,  $\sigma$ , is obtained by considering the usefulness of a metabolite to biomass formation. The shadow price of a metabolite represents the marginal utility of that metabolite towards the process of growth. Scaling of the shadow price as represented by  $\sigma$  gives the maginal utility of a metabolite in glucose equivalents. We are thus able to determine the utility of various metabolites on a common basis. For the maximal biomass yield solution, Fig. 5.1, we find that the sugar monophosphates have the highest utility for cell growth.

### 5.1.2 How Sensitive is the Optimal Solution?

The optimal solution displayed in Fig. 5.1 has several noteworthy characteristics. First, the pentose phosphate shunt (PPS) is extensively used. About 91% of the glucose is metabolized through the PPS, which is far greater than experimentally observed values [92]. Second, the glyoxalate shunt is used as an anaplerotic reaction which does not normally occur during *E. coli* growth on glucose. Third, the conversion of Pyr to PEP occurs via the enzyme PEP synthase which is again not normally observed during growth on glucose.

Thus, the optimal solution for maximum biomass yield in Fig. 5.1 shows some deviations from experimentally observed fluxes. In this section we determine the

sensitivity of the maximal biomass yield to various perturbations and thus try to evaluate the factors that determine the differences between experimentally obtained fluxes and the computed optimal solution.

There is a limited number of factors that can influence the optimal solution given the specifications of the model; basically the metabolic stoichiometry and the metabolic demands. The factors that can affect the optimal solution are:

1. **THE ACTIVE PATHWAYS** Several pathways and reactions that appear in the optimal solution may not be operative in the physiological state considered.
2. **THE METABOLIC DEMANDS** The biosynthetic requirements used for the generation of one gram of dry biomass given in Table 3.1 are only estimates. These requirements may vary.
3. **THE P/O RATIO** The stoichiometry of the metabolic network is well known. The one exception is the P/O ratio, since the stoichiometry of energy transducing membranes is not fixed.
4. **THE MAINTENANCE REQUIREMENT** Living cells have maintenance energy requirements that are not accounted for in the optimal solution presented above.

We now characterize the effect of these factors on the maximal biomass yield.

1. **Constraints on selected pathways** As pointed out above, the optimal solution has three noteworthy deviations from the expected flux distribution corresponding to experimental data on aerobic growth of *E. coli* on glucose. First, there is a high flux through the pentose pathway. Second, the glyoxylate shunt is used. Third, Pyr is converted to PEP by the enzyme PEP synthase.

The optimal biomass yield was examined with respect to these pathways by restricting the flux through them. Forcing both the glyoxylate shunt and PEP synthase fluxes to zero resulted in a negligible drop in the maximum biomass yield, to a value of 0.585 g DW/g glucose. Thus, the optimal biomass yield is not very sensitive to flux variations through these pathways.

The maximal biomass yields are relatively insensitive to variations in the flux through the pentose pathway. Shown in Fig. 5.3a is the maximum biomass yield as a function of the pentose pathway flux. Even if the pentose pathway flux is restricted to zero the maximal biomass yield only drops by 2.5%. Thus, the cell has significant flexibility in determining the metabolic flux distribution around the G6P node without experiencing detrimental effects on the biomass yield. We will discuss later the influence that ATP maintenance energy has on the pentose pathway flux.

**2. Biosynthetic Demands** The metabolic requirements to produce one gram of dry cell weight listed in the second column of Table 5.1 are based on estimates from experimental measurements of cellular composition [32, 51]. The estimates are strain specific and can also change under different environmental conditions. It is therefore pertinent to calculate the sensitivity of the maximal biomass yield to perturbations in the biosynthetic requirements.

Since the biosynthetic demands vary significantly in their absolute values we compare the possible errors in their estimates using a relative, or a logarithmic sensitivity coefficient:

$$\frac{\partial X/X}{\partial d_M/d_M} = \frac{\partial \ln X}{\partial \ln d_M} \quad (5.8)$$

where  $d_M$  is the biosynthetic demand. The logarithmic sensitivity measures the per-

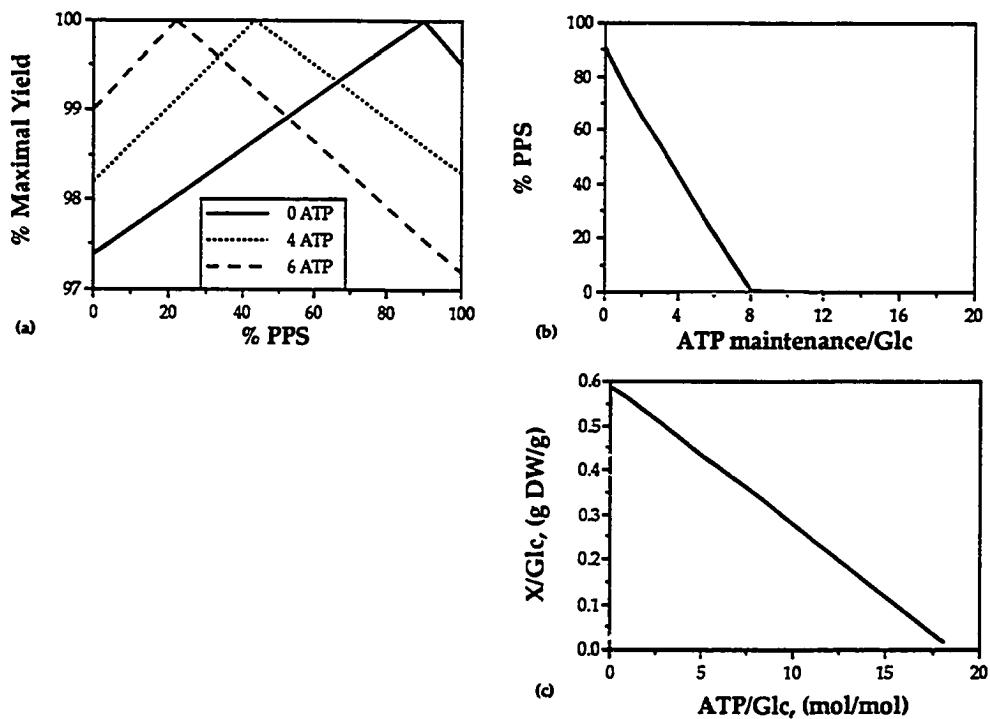


Figure 5.3: (a) Yield sensitivity, expressed as a percentage of the maximal biomass yield, plotted as a function of the percentage PPS flux for three different ATP maintenance requirements: 0, 4, and 6 ATP/Glc. The PPS flux represents the flux through the oxidative branch of the pentose phosphate pathway. (b) Optimal percent PPS flux plotted as a function of ATP maintenance requirements. (c) Variation in maximal biomass yield with changing ATP maintenance requirements.

Table 5.2: Logarithmic sensitivity of maximal biomass yield to the metabolic requirements for growth. The change in sign for NADH reflects a net production of NADH during growth.

Cofactor	Sensitivity $(\frac{\partial X/X}{\partial d_M/d_M})$	Precursor	Sensitivity $(\frac{\partial X/X}{\partial d_M/d_M})$	Precursor	Sensitivity $(\frac{\partial X/X}{\partial d_M/d_M})$
ATP	0.2021	G6P	0.0238	3PG	0.0733
NADH	-0.0231	F6P	0.0082	PEP	0.0254
NADPH	0.1677	R5P	0.0871	Pyr	0.1105
		E4P	0.0285	AcCoA	0.1237
		T3P	0.0077	OA	0.0876
				$\sigma$ KG	0.0777

percentage change in the biomass yield in response to a percentage change in the precursor requirement. Logarithmic sensitivity coefficients have become quite popular for the analysis of kinetic models of metabolism following their introduction [26, 35, 66].

The relative sensitivity coefficients for the biosynthetic precursors and cofactors are calculated as:

$$\frac{\partial X/X}{\partial d_M/d_M} = \frac{d_M}{X} \times \frac{\partial X}{\partial M} \times \frac{\partial M}{\partial d_M} \quad (5.9)$$

$$= d_M \times \frac{\partial X}{\partial M} \quad (5.10)$$

$$\text{since } \frac{\partial M}{\partial d_M} = X \quad (5.11)$$

Thus, only the shadow prices and the metabolic requirements are needed to compute the logarithmic sensitivity coefficients. The numerical values for the logarithmic sensitivity coefficients are shown in Table 5.2.

Examination of the computed logarithmic sensitivity coefficients leads to important observations. First, the biomass yield is not overly sensitive to changes in biosynthetic need of any one of the precursors or cofactors. The highest values for the logarithmic sensitivity coefficient are 0.2 and 0.17 for ATP and NADPH respectively. Quantitatively, these values mean that if the need for these cofactors increases by 10% then the biomass yield will drop 2% and 1.7% respectively. We will discuss further the ATP maintenance requirements below.

Second, changes in the use of biosynthetic precursor molecules do not significantly affect the biomass yield. For the bacteria this feature is highly desirable since biosynthetic demands may vary from one environmental condition to another. Moderate deviations from the values for  $d_M$  listed in Tables 3.1,3.2 thus do not have overly detrimental effects on the biomass yield. From an evolutionary and taxonomic standpoint these results are interesting. This flexibility in meeting metabolic demands provides a basis for the evolution of number of different biomass compositions without losing significantly in biomass yield, and thus without significant loss in competitive advantage. The metabolic network can thus serve the metabolic requirements of many different genomes.

Finally, these results are encouraging for the purpose of redirecting fluxes in bacteria or to “metabolically engineer” them [2, 3, 80]. Moderate diversion of metabolic resources should not lead to engineered strains with vastly inferior growth characteristics. The question of excessive drains is addressed below.

**3. The P/O ratio** The P/O ratio is defined as the number of high energy phosphate bonds formed during the transfer of a pair of electrons to oxygen from NADH via the electron transfer system (ETS). The P/O ratio is determined by two inde-

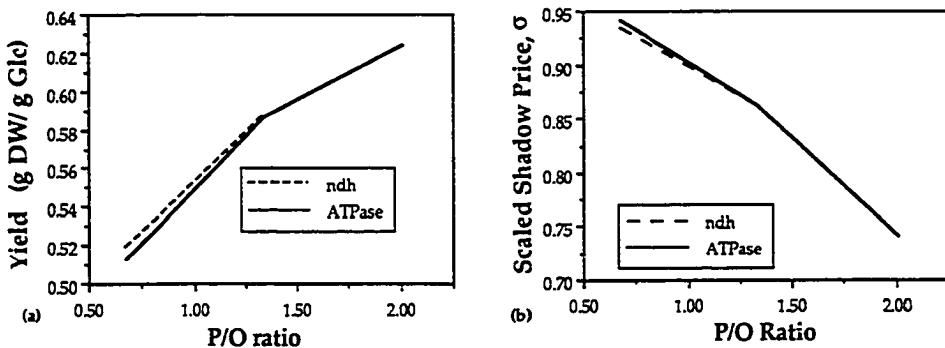


Figure 5.4: (a) The optimal biomass yield as a function of the P/O ratio. The two curves represent changes in the P/O ratio due to different stoichiometries of NADH dehydrogenase (ndh) and ATPase. (b)  $\sigma$  value for ATP as a function of the P/O ratio. The P/O ratio is a combination of the reactions catalyzed by the NADH dehydrogenase and cytochromes and ATPase.

pendent factors; the number of protons translocated by the NADH dehydrogenases and the stoichiometry of ATPase.

The enzyme NADH dehydrogenase (ndh) transfers hydrogen from NADH to quinone. Both energy linked and unlinked activities have been observed [62]. Use of the non-energy linked activity results in a P/O ratio of 0.667 as compared to a P/O ratio of 1.33 for the energy linked dehydrogenase, assuming an ATPase stoichiometry of 3 H<sup>+</sup>/ATP. Fig. 5.4a depicts the maximal biomass yield with either of the NADH dehydrogenases. The use of the non-energy linked NADH dehydrogenase, *i.e.*, reducing the P/O ratio by 50%, results in a significant drop in the biomass yield, to 0.519 g DW/g glucose, a drop of 11.7%.

The energy of the transmembrane proton gradient is converted into high energy phosphate bonds by the enzyme ATPase. The exact stoichiometry of the conversion is not accurately known. A value of  $3 \text{ H}^+/\text{ATP}$  has been used for our analysis, however non-integral values are possible [36, 37, 47]. The effect of deviations in the P/O ratio, due to ATPase stoichiometry, on the optimal biomass yield is also shown in Fig. 5.4a. The use of  $\text{H}^+/\text{ATP} = 2$  which corresponds to a P/O ratio of 2, *i.e.*, an increase of 50%, results in a biomass yield of 0.625 grams dry cells *per* gram glucose an increase of 6.3%. On the other hand, a lower P/O ratio of  $2/3$  (corresponding to  $\text{H}^+/\text{ATP} = 6$ ), a drop of 50% reduces the biomass yield to 0.514 grams biomass per gram glucose, a 12.6% drop. The yield is slightly more sensitive to a P/O ratio change caused by ATPase stoichiometry as compared to ndh.

Thus, we find that within the limits of P/O ratio from  $2/3$  to 2 ( $\pm 50\%$  from our nominal value of 1.33) the maximum biomass yield can vary from 0.514 to 0.625 g DW/g glucose (a 19% range). For the results presented here the P/O ratio of 1.33 recommended in literature [47] has been used. The considerable range of biomass yield argues for the evolution of specialized energy transducing membranes to carry out terminal oxidation. Higher organisms do indeed show specialized structures, *e.g.*, mitochondria and chloroplasts, that are capable of more efficient energy transduction.

Fig. 5.4b shows the changes in the scaled shadow price for ATP with changes in the P/O ratio. As expected with a lower P/O ratio, ATP production assumes a greater importance which is reflected in the relative increase in the  $\sigma$  value for ATP from a value of 0.7 at P/O=2 to a value of 0.9 at P/O=0.67. The increases in the  $\sigma$  value indicates the tendency of the maximal biomass solution to move towards the maximal ATP production solution with a reduction in the P/O ratio.

**4. ATP maintenance requirements** Cells use energy for metabolic functions other than growth. These functions are many, including cellular motility, maintaining cellular osmolarity, macromolecule turnover, transport, and maintenance of transmembrane gradients. Since it has proven difficult, perhaps impossible, to quantify each one of these functions individually, a phenomenological measure known as the maintenance coefficient [49, 59] has gained widespread use. It captures the essence of all such functions and can be determined experimentally [76]. Most maintenance requirements are of energy and may therefore be termed as ATP maintenance requirements. Maintenance represents a significant drain on metabolic resources at low growth rates.

In the stoichiometric framework, maintenance requirements may be incorporated as a specified number of ATP molecules consumed (for activities other than growth) *per* glucose molecule imported into the cell. In this way the effects of the ATP maintenance needs on the optimal biomass yield are readily calculated. The results of such calculation is shown in Fig. 5.3c. The x-axis is given in terms of moles of ATP that are used for maintenance for every mole of glucose consumed. Recall that the ATP shadow price on glucose utilization is 18.7 and thus when 18.7 ATP molecules *per* glucose are used for maintenance the biomass yield drops to zero. The drop in biomass yield with increasing maintenance costs is significant.

Since experimental yields for growth on glucose are on the order of 0.4-0.5 g DW/g glucose, we show the flux distribution for ATP maintenance requirement of 4 ATP *per* glucose consumed in Fig. 5.5. The maximal cell yield with this maintenance demand is 0.467 g DW/g glucose. Note that the fluxes through the glyoxylate shunt and through the PEP synthase have both dropped to zero without imposing any restrictions on these fluxes. Also note that NADPH is produced both by the

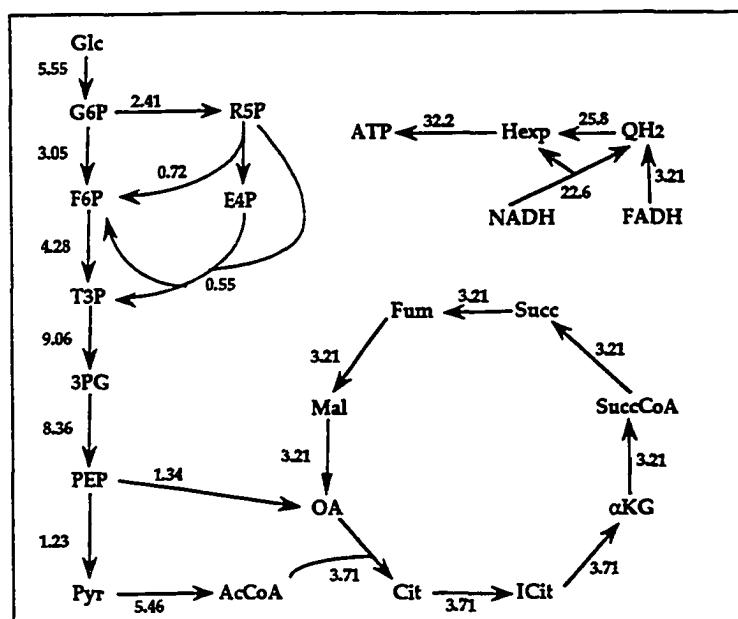


Figure 5.5: The optimal flux distribution for aerobic growth on glucose with a maintenance requirement of 4 ATP/Glc. The maximal biomass production is 0.467 g DW/g Glc.

pentose phosphate pathway as well as from  $\alpha$ KG synthesis. Interestingly, there is no contribution of the transhydrogenase reaction to NADPH production which is consistent with the lack of phenotype for transhydrogenase mutants [93].

Computations of the scaled shadow price,  $\sigma$ , for an ATP maintenance requirement of 4 ATP/Glc are listed in Table 5.1. We observe significant changes in the  $\sigma$  values after the inclusion of ATP maintenance requirements. ATP now has the highest value indicating that energy production has acquired an increased importance.

We next examine the PPS flux as a function of the ATP maintenance requirements for maximal biomass yield. The flux through the PPS does show a stoichiometric optimality after considering ATP maintenance requirements. Fig. 5.3b plots the computed optimal PPS flux for different maintenance requirements. We find that the optimal PPS flux is very sensitive to the maintenance requirements. In fact, maintenance requirements in the range of 4-6 ATP/Glc result in optimal PPS fluxes in the physiological range of 20-40% [92]. Corresponding to these maintenance requirements the maximal biomass yield is in the range of 0.4-0.45 g DW/g glucose which is again a physiologically observed range. The maximum biomass yield sensitivity to the PPS flux under ATP maintenance requirements of 4 and 6 ATP/glucose is shown in Fig. 5.3a. The biomass yield varies over a 2.5% range for the entire range of PPS flux.

Thus, the optimal solution obtained with the inclusion of maintenance costs shows pathway utilizations that correspond more closely to experimental observations than the solution without maintenance costs. We therefore observe that microbial metabolism does function in a manner that is consistent with attainment of stoichiometric optimality.

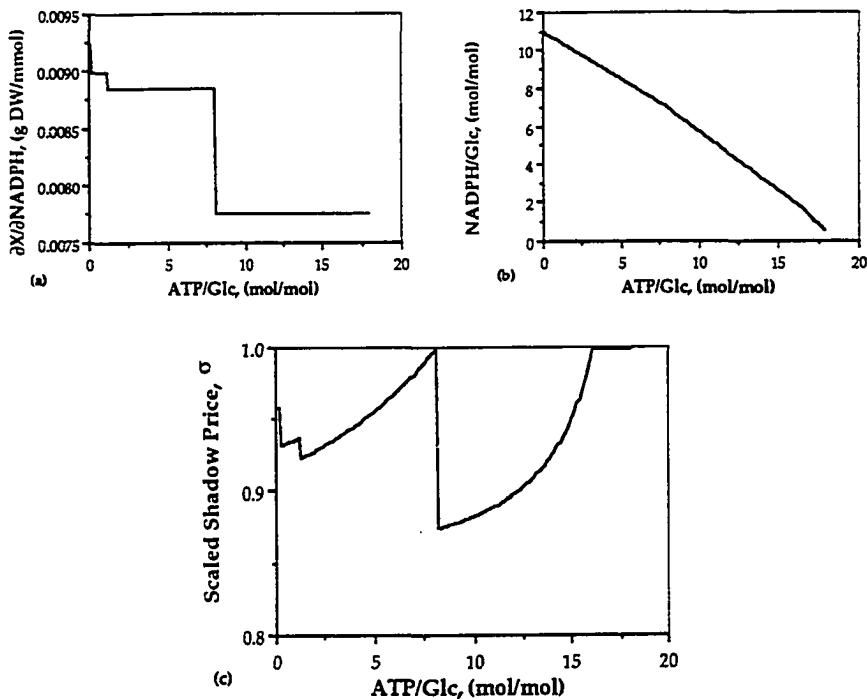


Figure 5.6: The derivation of  $\sigma$  values for NADPH from its defining equation shown as a function of the ATP maintenance demand. The maximal biomass yield is displayed in Fig. 5.3c. (a) The shadow price of NADPH while optimizing yield on glucose with a varying ATP maintenance demand. (b) The maximal production of NADPH from glucose with the same ATP maintenance demands. (c)  $\sigma$ -values.

### 5.1.3 The use of $\sigma$ to interpret the effects of variation in metabolic loads

The two previous subsections illustrate the development of a metric to rank growth constraining factors and the sensitivity of the optimal solution to variations in stoichiometry and metabolic loads. While the catabolic network displays versatility with respect to attaining high biomass yield in face of stoichiometric variations, changing the metabolic loads does change the flux distribution pattern significantly. The parameter  $\sigma$  can be used to interpret such changes. We will illustrate the use of  $\sigma$  with two examples.

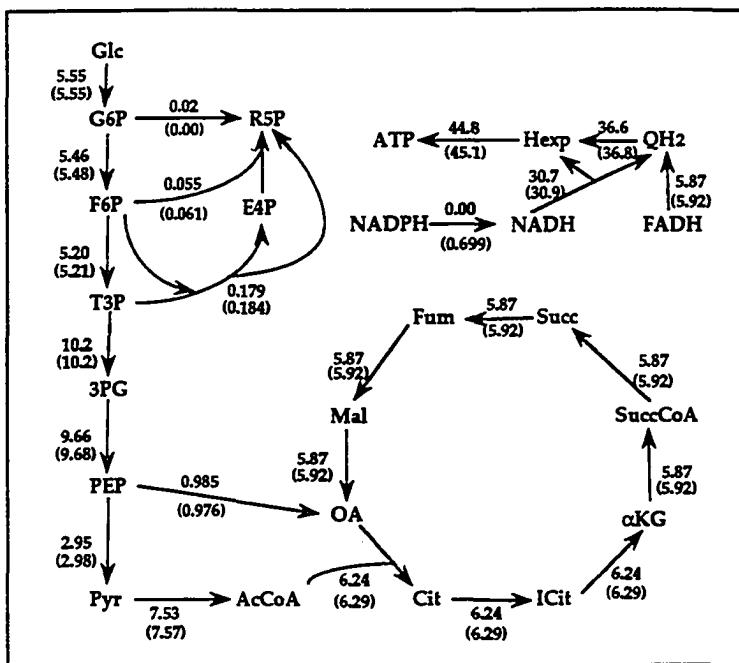


Figure 5.7: Comparison of flux distributions for an ATP maintenance requirement of 8 ATP and 8.1 ATP *per* glucose consumed. Fluxes for 8.1 ATP *per* glucose consumed are shown in brackets. The ratio of shadow prices for NADPH to ATP is 1.60 for 8 ATP and 1.33 for 8.1 ATP *per* glucose consumed.

**ATP maintenance requirements** The variations in the  $\sigma$  value for NADPH and the factors that it is comprised of with changing ATP maintenance requirement are shown in Fig. 5.6. The computed  $\sigma$  values show a pattern with several discontinuities. The continuous changes in  $\sigma$  are due to continually varying biomass yields and NADPH yields on glucose. On the other hand, the discontinuities in  $\sigma$  arise because of discontinuities in the shadow price for NADPH. Discontinuities in shadow prices are a result of changing a constraining boundary. The optimal use of the metabolic pathways changes, resulting in step changes in the shadow price.

Let us now examine a particular discontinuity more closely in order to illustrate the shifting of constraints. The discontinuity chosen is the one occurring between a

maintenance demand of 8 and 8.1 ATP per glucose consumed, Fig. 5.6. The flux distributions for an ATP maintenance demand of 8 and 8.1 are shown in Fig. 5.7. The discontinuity occurs in  $\sigma$  value due to a surplus production of NADPH that occurs at an ATP maintenance of 8.1. The surplus NADPH is converted into ATP via the electron transfer system. The surplus of NADPH produced results in the drop of its shadow price. In fact, for an ATP maintenance of 8.1 (Fig. 5.7b) the ratio of the shadow prices of NADPH to ATP is 1.33 which is the stoichiometric conversion ratio of NADPH to ATP while at a maintenance of 8 ATP/Glc this ratio of shadow prices is 1.60. Therefore, the value of NADPH is reduced to its ATP equivalence and energy becomes the governing yield constraint. This fact is reflected in the  $\sigma$  value for ATP being changed discontinuously to unity (not shown) while the  $\sigma$  value for NADPH falls to 0.875.

Discontinuities in shadow prices occur due to changes in constraining boundaries the consequences of which are a shift in pathway utilization. The increasing demands of ATP for maintenance cause increased utilization of the TCA cycle. Oxidation of ICit to  $\alpha$ KG in the TCA cycle results in the concomitant generation of NADPH. At a maintenance requirement of 8.1 ATP/Glc, a surplus of NADPH is produced that eliminates the need for PPS to produce NADPH. Thus, discontinuity arises and the value of NADPH is reduced to its ATP equivalent as the surplus is funneled down the electron transport system.

**Drain of a Precursor** Often microbial fermentations are carried out to produce specific products of interest. Useful products usually involve the drain of one or more precursors in order to produce the product. We now determine the effects of precursor drain on the biomass yield using the metabolic network formulated.

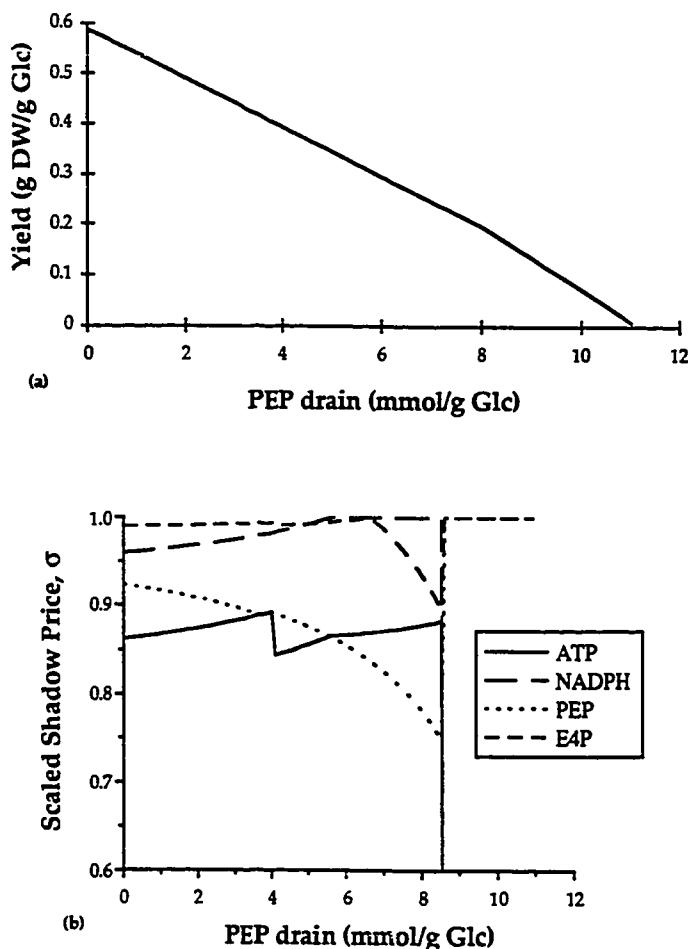


Figure 5.8: (a) Maximal biomass yield on a mass basis with a simultaneous drain of the precursor PEP. The precursor drain is defined as mmol per gram of glucose. (b) Scaled shadow price,  $\sigma$ , for some intermediates as a function of the drain of the precursor PEP.

The intermediate PEP has been chosen as a central metabolite directly or indirectly involved in many biosynthetic fluxes when glucose is used as the carbon source. Fig. 5.8a shows changes in the maximum biomass yield with a specified drain of the precursor PEP. As expected an almost linear decrease in biomass yield is obtained with precursor drain. At a maximal PEP drain of 11.1 mmol/g glucose the biomass yield drops to zero.

To illustrate the constraints on maximal biomass yield on glucose with an associated PEP drain the  $\sigma$  values for some metabolites have been plotted in Fig. 5.8b. For PEP drains below 8.5 mmol/g glucose we note that biosynthetic redox (NADPH) and sugar monophosphates represented by E4P have a high  $\sigma$  values. A PEP drain above 8.5 however results in surplus ATP and redox being produced as noted by their shadow price dropping to zero. PEP now becomes a governing constraint with all other metabolites showing much lower  $\sigma$  values.

Notable changes occur in the flux distribution at the discontinuity occurring at a PEP drain of 8.5 mmol/g glucose. At a PEP drain exceeding 8.5 mmol/g glucose energy and redox stop being constraints. The flux distribution before the discontinuity may be compared to that after the discontinuity, Fig. 5.9. At higher PEP drains an excess of energy and redox are produced. Surplus redox is converted to energy through the electron transport system. The net surplus of energy is dissipated in the network by the F6P/FDP futile cycle, as shown in Fig. 5.9. In the accompanying paper the optimal production of PEP also resulted in the dissipation of surplus ATP through a different futile cycle. Several futile cycles are embedded in the catabolic network, and all can serve to dissipate surplus energy. They all lead to equivalent but different optimal solutions as discussed earlier in Chapter 4. The interpretation of this discontinuity is thus similar to that given above.

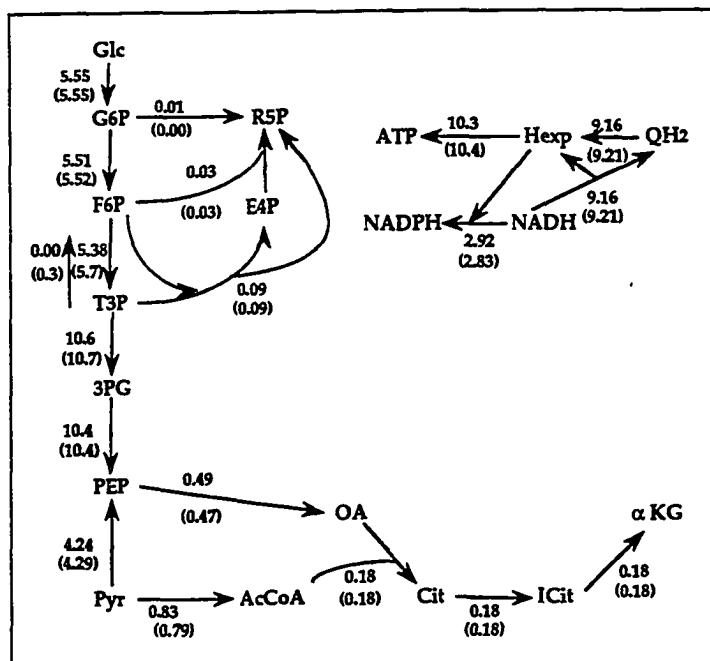


Figure 5.9: Comparison of flux distributions for a PEP drain of 8.5 PEP and 8.6 PEP per gram of glucose utilized. Fluxes for 8.5 PEP per gram glucose consumed are shown in brackets. For 8.6 PEP the shadow price for energy and redox are zero and surpluses are produced. Surpluses are eliminated through the use of the energy dissipative futile cycle  $F6P \rightleftharpoons 2 T3P$ .

Table 5.3: Flux distribution for growth on acetate for maximum biomass yield compared to experimentally observed values reported in literature, [90]. All fluxes are relative to the same input flux of 145 for acetate. The last column shows the flux distribution obtained with the inclusion of an ATP maintenance demand of 1.5 ATP per acetate in the metabolic network.

Enzyme	Experimental Flux	Optimal Fluxes	
		No Maint.	With Maint.
PYRDH	0.0	0.0	0.0
CITSYN	111.0	92.3	109
ISODHP	80.0	53.6	82.7
AKGDH	75.0	49.6	80.0
SUCCDH	106.0	88.3	106
MALDH	127.5	116.0	126
ISOLYS	31.0	38.7	26.4
MALENZ	9.5	10.6	7.2
PEPCK	9.5	17.4	11.9

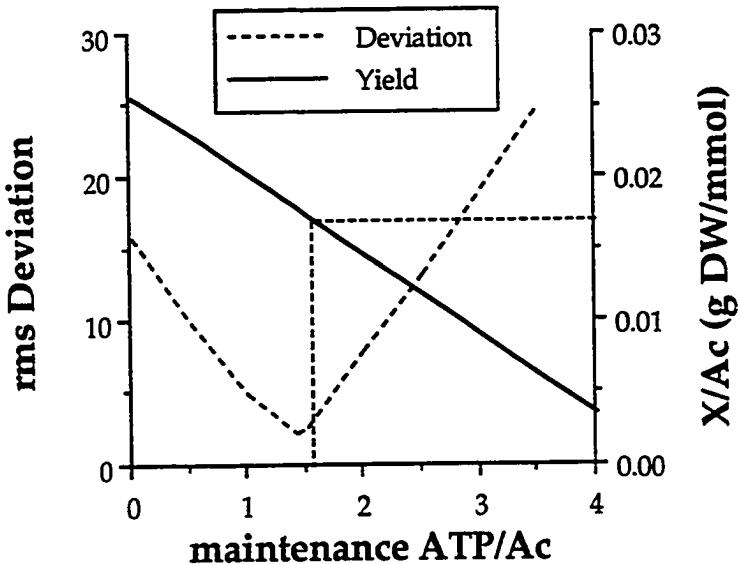


Figure 5.10: Maximal biomass yield on acetate plotted as a function of the ATP maintenance requirements. The root mean square deviation of the experimental flux distribution from the computed maximal biomass yield solution is also shown as a function of the ATP maintenance requirements.

## 5.2 Growth on Acetate

The flux distribution for *E. coli* growth on acetate is reported in literature [90]. These experimentally determined metabolic fluxes are listed in Table 5.3. Here we attempt to compare the experimentally obtained flux distributions to the results computed for maximal biomass yield using the fueling network.

In order to make a suitable comparison we have to determine the maintenance requirement for growth on acetate. We have therefore computed the maximal biomass yield as a function of the ATP maintenance requirements, as shown in Fig. 5.10. The maximum biomass yield on acetate without maintenance is 0.0255 g DW/mmol acetate. Such a high yield is not experimentally observed because of maintenance

demands. Experimentally observed yields are approximately 0.0163 g DW/mmol acetate [12]. Based on these biomass yields we can estimate a maintenance requirement of 1.5 ATP/acetate, see Fig. 5.10. A maintenance of 1.5 ATP/acetate may appear to be high considering that the maximum ATP generation by the fueling network is 4.67 ATP/acetate. However, this maintenance includes the additional costs of acetate uptake by the high energy consuming acetate scavenging pathway [10]. The maximal yield for a maintenance requirement of 1.5 ATP/acetate is 0.0175 g DW/mol acetate.

We have also computed the flux distributions corresponding to the maximal biomass yield under different maintenance requirements. Two such flux distributions, for no maintenance and a maintenance of 1.5 ATP/acetate are listed in Table 5.3. We have evaluated the root mean square deviation for the computed fluxes from the above mentioned experimental fluxes. The results are plotted against the maintenance requirement, see Fig. 5.10. We note the minimum deviation between the computed and measured fluxes occurs around 1.5 ATP/acetate. Remarkably, by defining a physiologically appropriate ATP maintenance requirement we obtain an optimal flux distribution that is similar to the experimentally observed flux distribution.

Interesting questions have arisen with respect to the use of alternate pathways during growth on acetate. It has been suggested that the glyoxalate shunt is the evolutionary predecessor to the complete TCA cycle [30]. Restricting the TCA cycle and thus forcing all the flux through the glyoxalate shunt we obtain the maximal biomass yield of 0.014 g DW/mmol acetate. The 20% drop in biomass yield would then correspond to the evolutionary pressure for the development of the TCA cycle. Another alternative for energy generation is the use of the pentose phosphate path-

way. However use of the pentose phosphate pathway for energy generation leads to a maximal computed biomass yield of 0.0109 g DW/mmol acetate. This yield is substantially lower and provides an argument against the use of the pentose phosphate pathway for energy generation. In conclusion, as it does for growth on glucose, the use of acetate as a substrate follows stoichiometric optimality principles remarkably closely.

### 5.3 Conclusions

The influence of metabolic stoichiometry on cellular function and behavior has been studied within the context of flux balancing and linear optimization. The metabolic capabilities and behavior of *E. coli* have been analyzed within this framework and the results compared to experimental data. Remarkably close agreement between computed optimality and a range of different experimental observations is obtained. Taken together all the results, obtained in this and the preceding Chapter support the hypothesis that the metabolic and growth behavior of *E. coli* is consistent with stoichiometric optimality and the underlying regulation strives to attain behavioral patterns that can be described by stoichiometric optimality principles. Finally, the results presented show that analysis through stoichiometric flux balancing can potentially be used both to guide metabolic engineering of bacteria and to provide teleological arguments for evolutionary pressures on metabolic network design and behavior.

## CHAPTER VI

# BIOCHEMICAL PRODUCTION CAPABILITIES

The commercial production of biochemicals often utilizes the metabolic reactions of a microbe in order to achieve the conversion of substrate into the desired product. The subversion of microbial metabolism to over-produce the product is usually based on random mutagenesis in a selective environment along with the addition of external genetic material. In the present Chapter we use flux balance analysis with the combined catabolic and anabolic network to determine the capabilities of *E. coli* to convert glucose, glycerol, and acetate as substrates into amino acids and nucleotides. The optimal metabolic pathway utilization for biochemical production represents the ultimate goal of flux redistribution in a commercial bioprocess. Criteria are established that determine the metabolic constraints limiting a particular optimal solution. Metabolic constraints during the production of amino acids and nucleotides from a glucose, glycerol, or acetate carbon source are determined. We also address the industrially relevant trade-off between growth and product formation.

**Maintenance Requirements** In addition to the composition-based metabolic demands for growth there are also maintenance requirements in viable cells. Activities such as gradient maintenance, regulatory functions, protein turnover are

Table 6.1: Constraints faced in the production of amino acids and nucleotides along-with their defining criteria. Decoupling the metabolic network for a specific cofactor is essentially the same as externally supplying a surplus of the cofactor. A ‘-’ indicates any possible value. Hexp refers to the energy of the proton gradient.

Constraint	CO <sub>2</sub> Production	Shadow Prices		Decoupled
		Hexp	NADPH	
None	$\leq 0$	0	0	
Stoichiometry	$> 0$	0	0	
Redox	-	0	$> 0$	
Energy	-	$> 0$	$> 0$	
Redox + Energy	-	$> 0$	$> 0$	
Redox + Stoichiometry	$> 0$	0	-	Redox
Energy + Stoichiometry	$> 0$	$> 0$	$> 0$	
Energy + Redox + Stoichiometry	$> 0$	-	$> 0$	Energy
	$> 0$	-	-	Energy + Redox

accounted for by including a maintenance energy loss in the metabolic network. A fit of the model to experimental data [49, 76] yields a requirement of 23 mmol ATP/g biomass for growth associated and 5.87 mmol ATP/g DW-hr for non-growth associated maintenance, also see §7.1. These two maintenance terms are also included in the flux balance model.

## 6.1 Definition of Metabolic Constraints

Optimization of an objective can be limited by metabolic constraints. The constraints identified here are: stoichiometry, redox, energy, and combinations thereof. We need to establish a criteria for determining the constraints in a particular optimal solution. Stoichiometric constraints are defined as the use of necessary decarboxylation steps resulting in a net CO<sub>2</sub> evolution. Redox and energy constraints are characterized by the utility of these cofactors in increasing production. Constraints and the criteria defining them are listed in Table 6.1.

The presence of a constraint is indicated by a specific combination of the CO<sub>2</sub> production and the shadow prices of energy and biosynthetic redox. For example, a stoichiometric constraint is indicated by a net CO<sub>2</sub> evolution (i.e., less than 100% carbon conversion) and zero energy and redox shadow prices. A redox constraint is indicated by a positive NADPH shadow price with a zero energy shadow price, indicating only the utility of redox for the objective of biochemical production. An energy constraint would result in a positive shadow price for both energy and redox due to the convertibility of redox to energy by the electron transfer system.

Combinations of constraints can be determined by decoupling the metabolic network from appropriate cofactors. Such decoupling implies not considering the particular cofactor in any of the reactions in the metabolic network. Decoupling can also be interpreted as providing an unlimited surplus of the cofactor to the metabolic network. For example, the combined energy and redox constraint can be distinguished from the energy constraint by a positive shadow price for redox in the energy decoupled metabolic network. Other combinations of constraints are similarly determined by the appropriate decoupling of the metabolic network for energy and redox as

shown in Table 6.1. Thus, the results from the flux balance based approach, such as the net CO<sub>2</sub> evolution and the utility of energy and biosynthetic redox, is interpreted in the form of metabolic constraints. The criteria for metabolic constraints given in Table 6.1 is hence in the nature of a logical truth table.

In addition, for the combined energy and redox constraint it is possible to determine the relative importance of the two constraints. Since the shadow prices of the proton gradient (Hexp) and NADPH represent the utility of energy and biosynthetic redox respectively, the ratio of the two shadow prices would be indicative of the relative importance of redox compared to energy. A higher value indicates a stronger redox constraint.

In comparison, the stoichiometric conversion of biosynthetic redox to energy has a ratio of 4 Hexp/NADPH. Since redox can be converted to energy, the ratio of the shadow prices must have a value above 4. The reverse conversion of energy to redox is not biochemically possible. However, by allowing a mathematical reversal of fluxes we obtain a ratio of 6 Hexp/NADPH. Comparison of the ratio of the shadow prices to these numbers provides a good estimate of the importance of the energy and redox constraints.

## 6.2 Maximal Theoretical Performance

We have determined the production capabilities of the *E. coli* metabolic network by incorporating a drain for specific biochemicals in the metabolic network and maximizing them using linear programming. We compute maximal yields of amino acids and nucleotides from three substrates; glucose, glycerol and acetate.

**GLUCOSE:** The maximal conversion of glucose into amino acids and nucleotides is listed in Table 6.2. In addition, the maximal biomass yield under fully aerobic condi-

Table 6.2: Maximal theoretical yield of amino acids and nucleotides on a glucose substrate in the absence of constant maintenance energy requirement. Network constraints are determined according to the truth table shown in Table 6.1. The redox to energy value represents the ratio of the shadow prices of the biosynthetic redox (NADPH) to those of the proton gradient.

Product	Maximum Yield (mol/mol Glc)	CO <sub>2</sub> Evolved** (mol/mol Glc)	Constraint‡	Redox Energy
Biomass	0.097*	1.910	E + S	
ALA	2.000	0.000	N	
ARG	0.774	1.360	E	
ASN	1.560	-0.240	E	
ASP	1.820	-1.260	E	
CYS†	0.975	3.080	E + R	5.4
GLU	1.000	1.000	S	
GLN	1.000	1.000	S	
GLY††	2.000	0.000	N	
HIS	0.730	1.620	E + S	
ILE	0.734	1.600	E + R	6.0
LEU	0.667	2.000	S	
LYS	0.784	1.300	E + R	6.0
MET†	0.574	3.130	E + R	5.7
PHE	0.529	1.240	E + S	
PRO	1.000	1.000	S	
SER	2.000	0.000	N	
THR	1.230	1.090	E + R	5.7
TRP	0.414	1.450	E + S	
TYR	0.548	1.070	E + S	
VAL	1.000	1.000	S	
AMP	0.500	0.996	E	
CDP	0.540	1.140	E	
GMP	0.498	1.020	E	
UMP	0.600	0.600	E	

\* Biomass yield is in g DW/mmol Glc

\*\* Carbon conversion can be calculated as  $\frac{(6 - \text{CO}_2 \text{ Evolved})}{6}$

† Included are the redox requirements of sulfate reduction.

†† A 1-carbon drain has been included.

‡ The constraints are indicated as E - energy, R- biosynthetic redox,  
S - stoichiometry, N - none.

tions is shown. These values determined without including the constant maintenance energy represent the maximal theoretical stoichiometric production capability of the metabolic network. The goal of metabolic engineering is to move from the normal state of biomass production towards over-producing a specific biochemical. The net CO<sub>2</sub> production for the various conversions is also listed in Table 6.2 and is indicative of the carbon conversion. The carbon conversion is a function of the stoichiometry, the redox requirements, and the energy requirements of the specific product.

Applying the truth table shown in Table 6.1 to the maximal production of various amino acids and nucleotides from a glucose substrate we observe the constraints listed in Table 6.2. A variety of constraints are observed in the production of the different biochemicals depending on the biosynthetic requirements. In comparison, biomass generation is shown to have a combined stoichiometric and energy constraint. It is interesting to note that redox is always associated with energy as a constraint in the cases considered. The ratio of redox and energy shadow prices is also listed in Table 6.2 for the combined redox and energy constraints.

**GLYCEROL:** Glycerol can also be used as a substrate for the production of biochemicals. The results from computations of the maximum theoretical yields and the corresponding constraints are shown in Table 6.3. In contrast to a glucose substrate we note that lysine production on glycerol does not have a redox constraint. Since glycerol is a more reduced substrate as compared to glucose we would expect redox to be less of a constraint with glycerol as the carbon source. Also, since redox can also be converted to energy we expect the energy constraints to be reduced as well. Thus, we note that aspartate, phenylalanine, and tyrosine do not have energy constraints on glycerol, in contrast to glucose.

The redox to energy value for the glycerol substrate is also observed to be higher

Table 6.3: Maximal theoretical yield of amino acids and nucleotides on a glycerol substrate in the absence of the constant maintenance energy requirement. Network constraints are determined according to the truth table shown in Table 6.1.

Product	Maximum Yield (mol/mol Glyc)	CO <sub>2</sub> Evolved** (mol/mol Glyc)	Constraint‡	<u>Redox</u> Energy
Biomass	0.054234*	0.711	E + S	
ALA	1.000	0.000	N	
ARG	0.430	0.419	E	
ASN	0.902	-0.610	E	
ASP	1.000	-1.000	N	
CYS†	0.852	1.340	E + R	6.0
GLU	0.500	0.500	S	
GLN	0.500	0.500	S	
GLY††	1.000	0.000	N	
HIS	0.413	0.520	E + S	
ILE	0.407	0.560	E + R	6.0
LEU	0.333	1.000	S	
LYS	0.435	0.388	E	
MET†	0.325	1.380	E + R	6.0
PHE	0.300	0.300	S	
PRO	0.500	0.500	S	
SER	1.000	0.000	N	
THR	0.698	0.207	E + R	6.0
TRP	0.237	0.391	E + S	
TYR	0.300	0.300	S	
VAL	0.500	0.500	S	
AMP	0.286	0.143	E	
CDP	0.304	0.262	E	
GMP	0.286	0.143	E	
UMP	0.342	-0.070	E	

\* Biomass yield is in g DW/mmol glycerol

\*\* Carbon conversion can be calculated as  $\frac{(3 - \text{CO}_2 \text{ Evolved})}{3}$

† Included are the redox requirements of sulfate reduction.

†† A 1-carbon drain has been included.

‡ The constraints are indicated as E - energy, R- biosynthetic redox,  
S - stoichiometry, N - none.

than that for glucose which is peculiar considering that glycerol is a more reduced substrate. The explanation for this anomaly lies in the stoichiometry of redox coupling. Glycerol incorporation into the central catabolic pathways produces redox in the form of NADH. NADH is easily converted to energy by the electron transfer system. More difficult is the conversion of NADH to biosynthetic redox (NADPH) by transhydrogenation which requires a net input of energy. Thus the extra reducing power of glycerol is more easily converted to energy rather than biosynthetic redox. Therefore the redox to energy ratio when metabolizing glycerol is higher than for glucose metabolism.

**ACETATE:** Computations of the maximal theoretical yields and corresponding constraints using acetate as a substrate are presented in Table 6.4. Energy is observed to be a constraint for all the products considered which is indicative of the poor energetic value of acetate. Stoichiometry is also seen to be a constraint for many of the conversions of acetate into products. Acetate enters the catabolic network through the TCA cycle and must therefore pass through several decarboxylation steps in order to produce the required biosynthetic precursors. The net CO<sub>2</sub> from these steps results in stoichiometric constraints during biochemical production.

Taken together, these results show that constraints on the production of biochemicals from a given substrate are a function of the energy and redox content of the substrate as well as its point of entry into the metabolic network.

### 6.3 Optimal Flux Distributions

The efficient production of biochemicals requires redirection of metabolic fluxes so that formation of the desired product is favoured. In Figure 6.1 we present the optimal catabolic and biosynthetic flux distributions for maximal growth using a

Table 6.4: Maximal theoretical yield of amino acids and nucleotides on an acetate substrate in the absence of the constant maintenance energy requirement. Network constraints are determined according to the truth table shown in Table 6.1. Acetate uptake is assumed to utilize the scavenging AcCoA synthase pathway.

Product	Maximum Yield (mol/mol Ac)	CO <sub>2</sub> Evolved** (mol/mol Ac)	Constraint‡	Redox Energy
Biomass	0.018438*	1.220	E + S	
ALA	0.393	0.821	E + S	
ARG	0.151	1.100	E	
ASN	0.324	0.706	E	
ASP	0.382	0.471	E	
CYS†	0.181	1.460	E + R + S	6.0
GLU	0.268	0.658	E + S	
GLN	0.250	0.750	E + S	
GLY††	0.394	0.818	E + S	
HIS	0.137	1.180	E + S	
ILE	0.144	1.130	E + R + S	6.0
LEU	0.159	1.040	E + S	
LYS	0.155	1.070	E + S	
MET†	0.111	1.450	E + R + S	6.0
PHE	0.100	1.100	E + S	
PRO	0.210	0.952	E + S	
SER	0.394	0.818	E + S	
THR	0.250	1.000	E + R	6.0
TRP	0.076	1.160	E + S	
TYR	0.103	1.070	E + S	
VAL	0.196	1.020	E + S	
AMP	0.095	1.050	E + S	
CDP	0.100	1.100	E + S	
GMP	0.093	1.070	E + S	
UMP	0.113	0.986	E + S	

\* Biomass yield is in g DW/mmol Acetate

\*\* Carbon conversion can be calculated as  $\frac{(2 - \text{CO}_2 \text{ Evolved})}{2}$

† Included are the redox requirements of sulfate reduction.

†† A 1-carbon drain has been included.

‡ The constraints are indicated as E - energy, R- biosynthetic redox,  
S - stoichiometry, N - none.

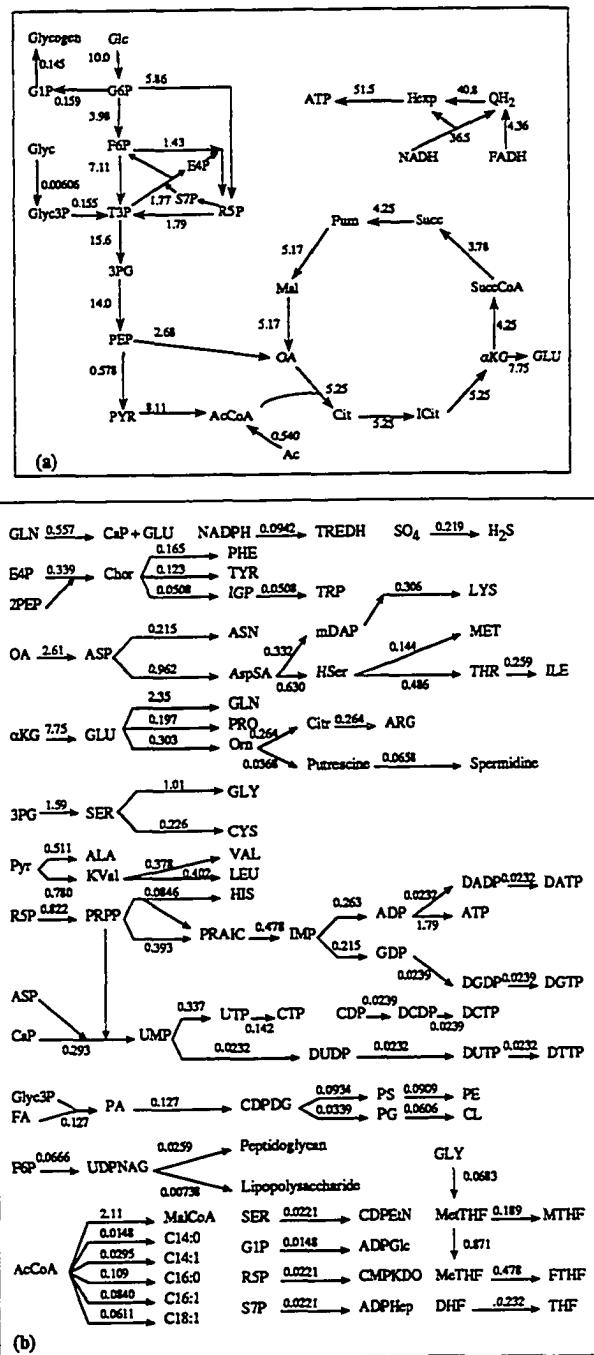


Figure 6.1: Flux distributions for maximal cell growth with an input of 10 mmol Glc/g DW-hr. (a) Catabolic flux distribution. (b) Biosynthetic flux distribution.

glucose supply of 10 mmol Glc/g DW-hr, resulting in a growth rate of  $0.94 \text{ hr}^{-1}$ . Under these conditions the energy yield is 70 mmol ATP/g biomass. The flux distribution shows the well-accepted use of metabolic pathways in *E. coli* during aerobic growth. An interesting feature of the catabolic flux distribution is the utilization of the acetate formed during the biosynthetic reactions.

Figure 6.2 displays the optimal catabolic flux distributions corresponding to the production of the various amino acids and nucleotides with a glucose supply of 10 mmol Glc/g DW-hr. The maximal yields are listed with the flux distributions. Note the dissipation of surplus energy shown as a drain of the proton gradient in some flux distributions. Surplus energy is observed in the flux distributions of ALA, GLU, GLN, and VAL which have either no constraints or stoichiometric constraints, Table 6.2. On the other hand, GLY, PRO, and SER which also have no constraints or stoichiometric constraints (Table 6.2) do not show a surplus of energy. The lack of surplus energy in these cases is due to the constant maintenance energy requirements being larger than the surpluses. The constant maintenance energy requirements are included in the computations for Figure 6.2.

Some interesting observations pertaining to the optimal utilization of pathways follow from the results given in Figure 6.2. The optimal production of ALA and VAL does not consume any oxygen as indicated by the absence of a flux through the cytochromes. Thus these amino acids are optimally produced by fermentation. Utilization of the glyoxalate shunt is observed during the optimal production of ARG, ASN, MET, THR, CDP, and UMP. Since the glyoxalate shunt is not observed to be operative while glucose is a substrate, this shunt may represent a prime target for metabolic engineering for the production of these biochemicals. Similarly the complete TCA cycle is seen to be operative during the optimal production of CYS,

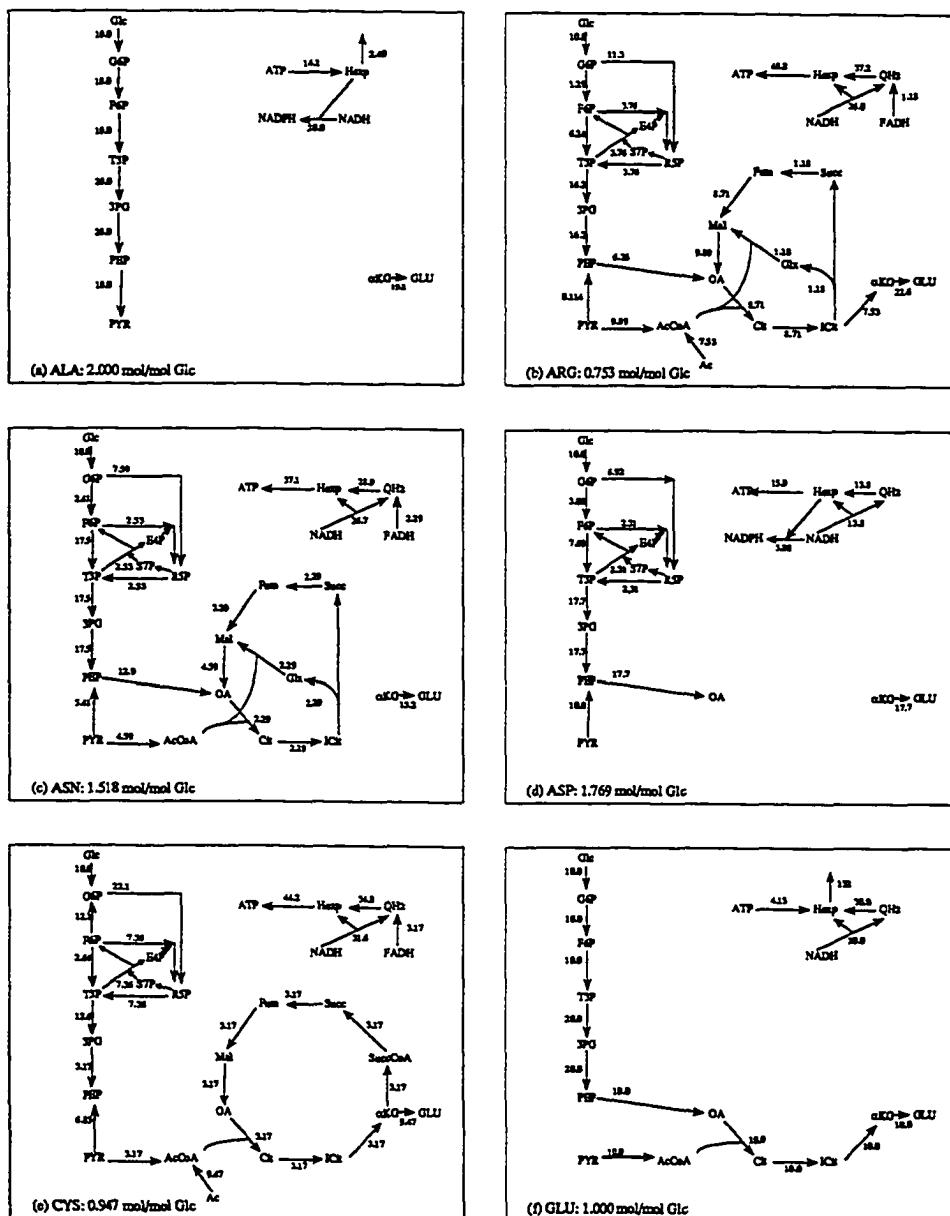


Figure 6.2: Optimal catabolic flux distributions for maximal biochemical productions. A substrate input of 10 mmol Glc/g DW-hr has been provided and maintenance requirements have been included. The maximal productions are listed in the individual flux distributions in mmol product/g DW-hr.

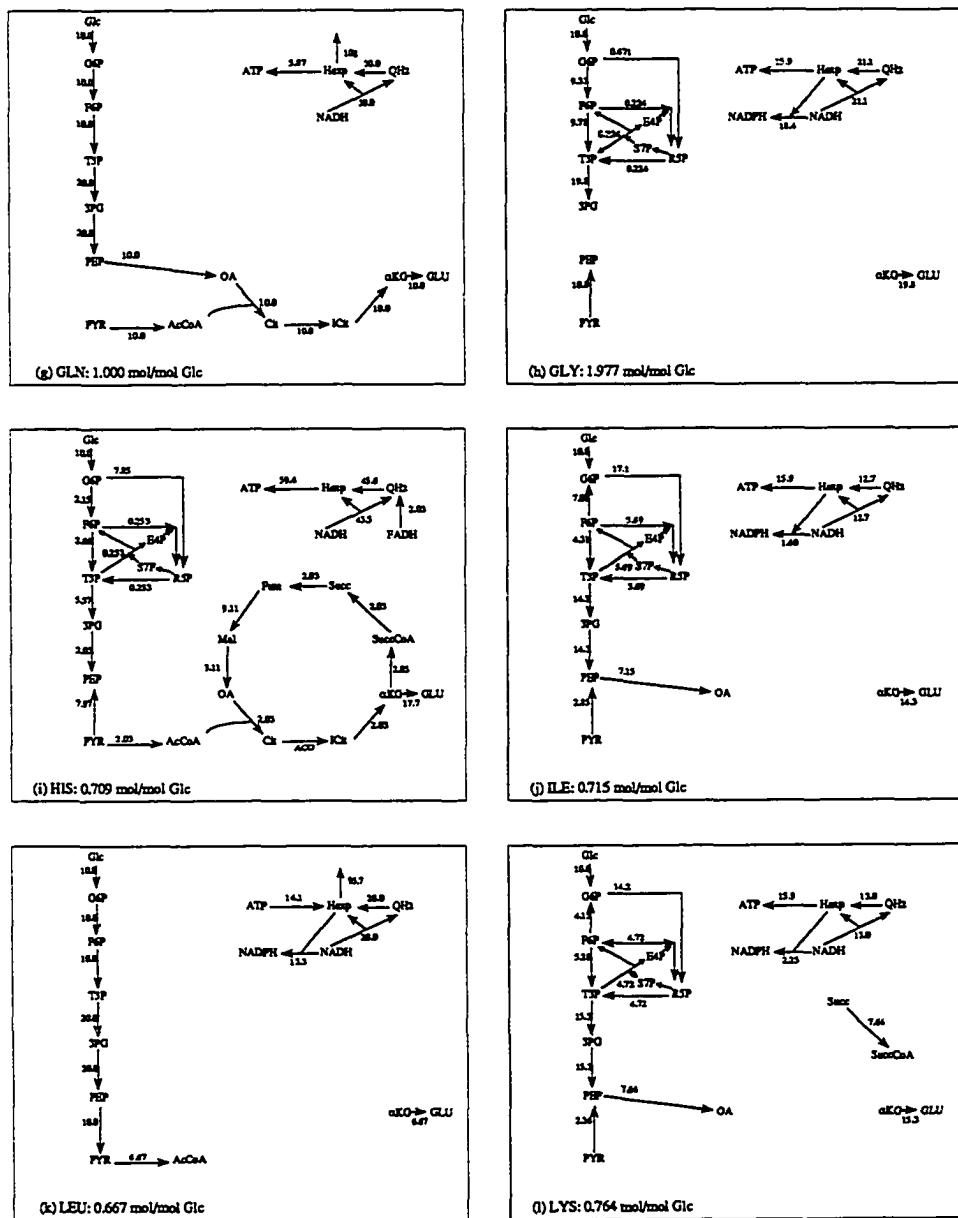


Figure 6.2: Continued.

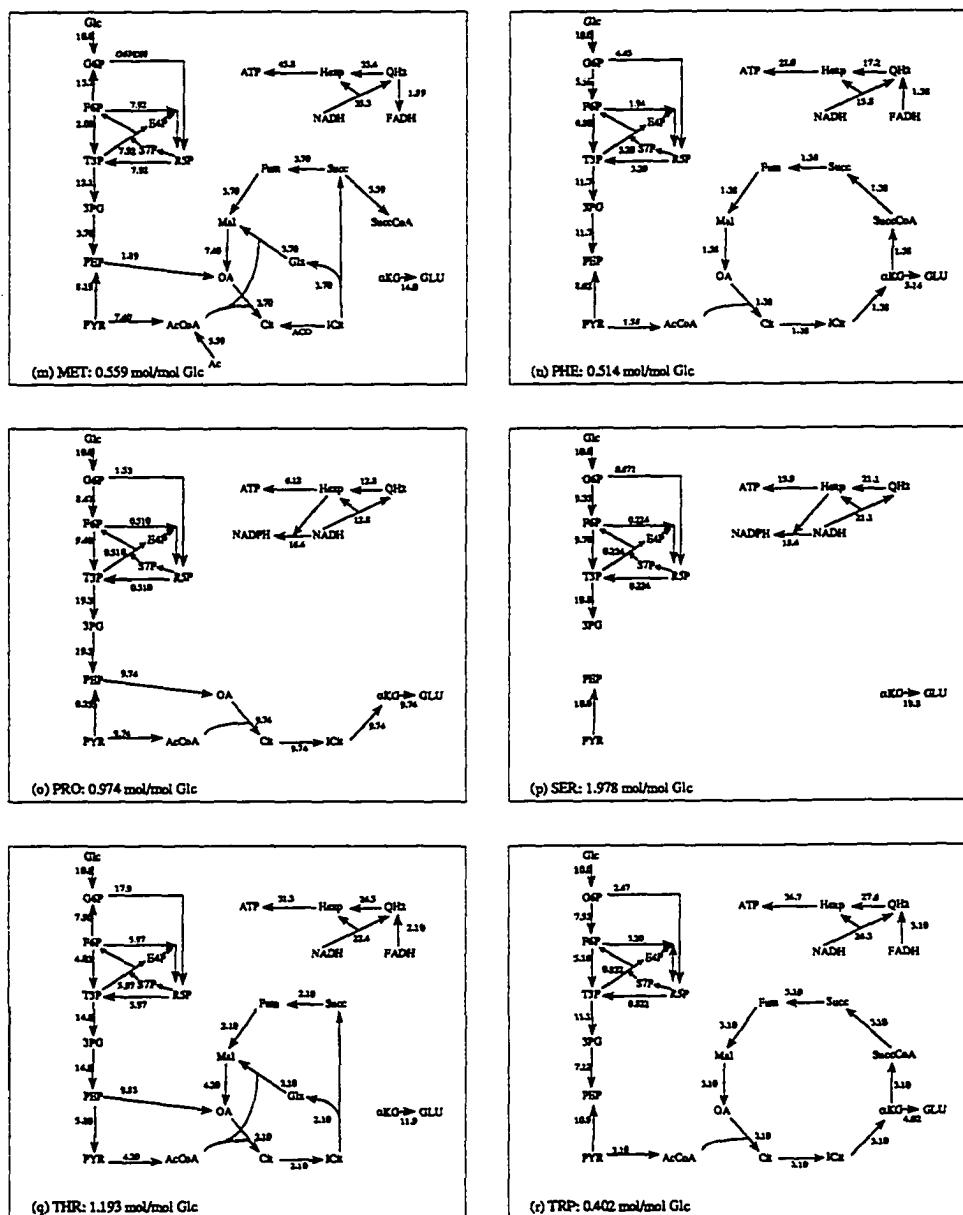


Figure 6.2: Continued.

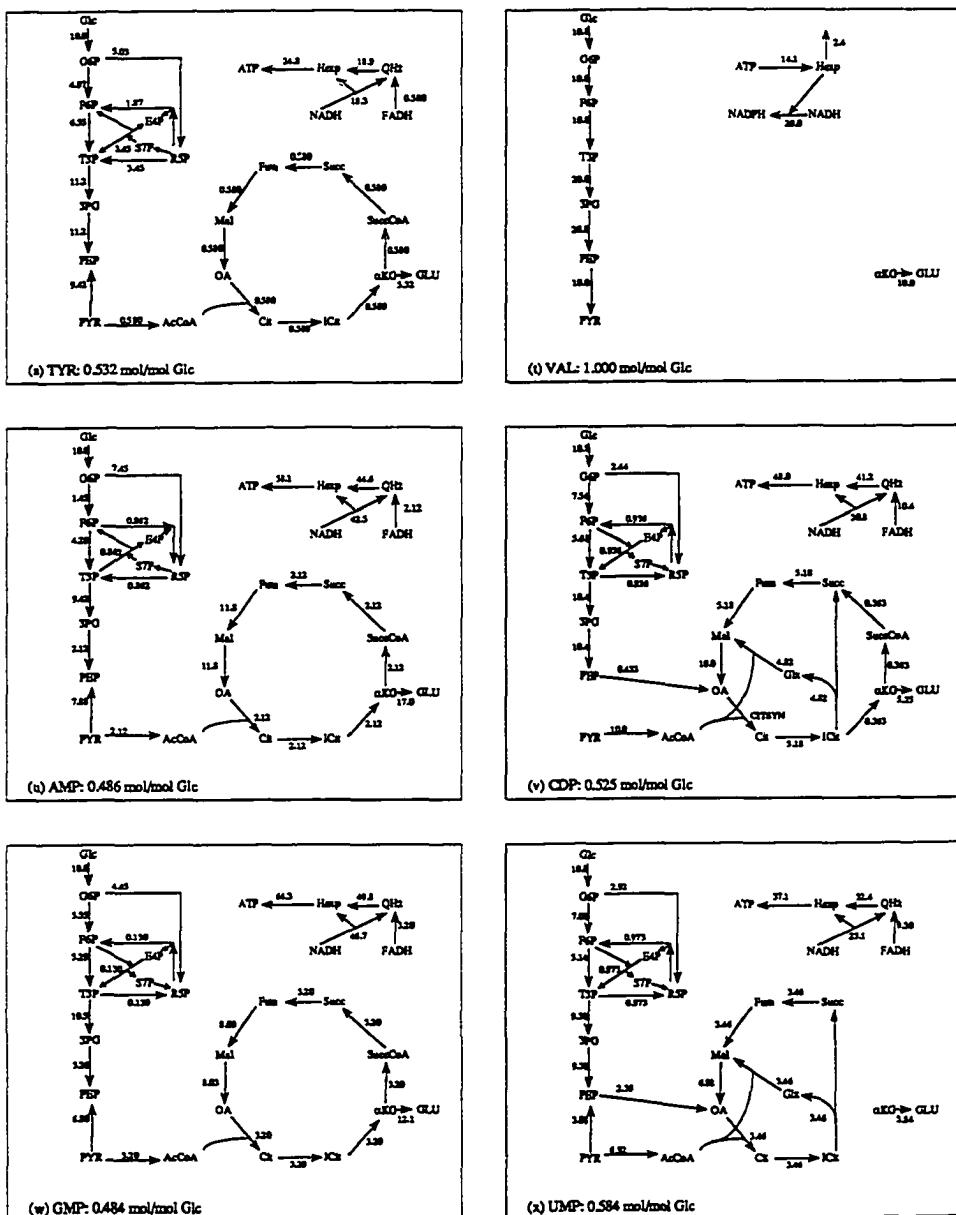


Figure 6.2: Continued.

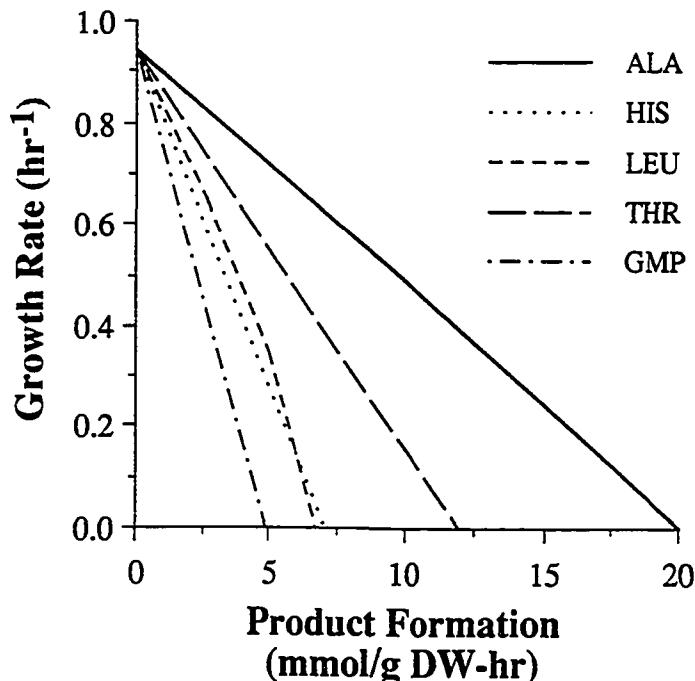


Figure 6.3: The simultaneous production of biomass and specific biochemicals. The optimal trade-off between biomass generation and biochemical production shows a piece-wise linear negative correlation.

HIS, PHE, TRP, TYR, AMP, and GMP. Again, repression of the TCA cycle in the presence of glucose is a suitable area for study in these cases.

In general, the flux distributions for maximal product formation should be compared to the flux distribution for maximal growth, Figure 6.1. Efficient biochemical production requires the redistribution of metabolic fluxes from producing biomass towards producing the specific product.

#### 6.4 Balanced Growth with Biochemical Production

It is often desirable to produce the product while simultaneously generating biomass. An optimal trade-off between growth and biochemical production can be assessed by choosing a production rate for a particular product between zero and

the maximum production rate, and then maximizing the growth rate. Figure 6.3 shows the optimal trade-off between growth and the production rate for a few select biochemicals. A negative correlation is observed between growth of the cell and biochemical production. The trade-off is piece-wise linear and encloses a convex space.

The production of leucine provides a good example of a piece-wise linear relationship between biomass generation and product formation. The production of leucine results in a surplus of energy generation, Figure 6.2k. Biomass generation on the other hand is constrained for energy, Table 6.2. Thus, the combined biomass and leucine production is able to utilize the energy surplus of leucine production. The increase in efficiency of the combined solution as compared to the addition of the individual solutions is the reason for the non-linear trade-off enclosing a convex space.

However, the deviation from absolute linearity is not significant for most of the metabolic products considered here. We have therefore tabulated the initial slopes of the trade-off lines for all the amino acids and nucleotides for maximal growth in Table 6.5. These slopes are indeed the shadow prices of the corresponding biochemicals computed from the dual solution. The shadow prices represent the marginal decrease in growth due to product formation. This trade-off is an important determinant of balanced growth and product formation.

## 6.5 Discussion and Conclusions

The primary goal in producing metabolic products with microbial cells is to obtain a high conversion of the substrate into the desired product. Determining the limits of substrate to product conversion is a key concern. The maximum theoretical yield is constrained by the stoichiometry of the reaction pathways in the metabolic network

Table 6.5: Optimal trade-off between growth and biochemical production represented as the shadow prices of the biochemicals in the maximal growth solution. The shadow price units are g biomass/100 mmol product. An input supply of 10 mmol Glc/g DW-hr has been used for the computations resulting in the growth rate of 0.94 g biomass/g DW-hr.

Product	$\partial\mu/\partial\text{Product}$	Product	$\partial\mu/\partial\text{Product}$	Product	$\partial\mu/\partial\text{Product}$
ALA	4.48	HIS	12.7	THR	7.61
ARG	12.5	ILE	12.9	TRP	22
ASN	5.99	LEU	11.3	TYR	16.8
ASP	4.98	LYS	12.1	VAL	8.95
CYS	9.56	MET	16	AMP	18.5
GLU	7.17	PHE	17.4	CDP	17.3
GLN	7.68	PRO	9.29	GMP	18.7
GLY	2.15	SER	4.31	UMP	15.8

which includes balancing the consumption and generation of metabolic cofactors. We have determined these theoretical limits on microbial performance by applying flux balance analysis to the metabolic network of the bacterium *E. coli*.

In the natural state, metabolism of microbes is directed towards growth. It has been suggested that metabolic regulation in microbial cells has evolved to maximize growth within stoichiometric constraints [88]. Over-production of a desired product thus requires the redirection of metabolic fluxes from generating biomass towards producing the desired biochemical product. We have determined the flux distributions that correspond to the maximal production of various amino acids and nucleotides, as illustrative biochemical products. The goal of engineering the strain can therefore be defined as the redirection of metabolic fluxes from the optimal growth to the optimal biochemical production.

Thus, engineering metabolism in order to produce a specific biochemical product raises two questions: what production level can be achieved, and how it may be attained? We choose lysine production as an illustrative example to answer these two questions. The maximal theoretical yield of lysine on a molar basis on various substrates is computed as glucose = 0.784, glycerol = 0.435, and acetate = 0.155. The corresponding carbon conversion is computed as glucose = 78%, glycerol = 87%, and acetate = 46%. Thus, it would appear that glycerol is the best substrate with the highest carbon conversion. However, we also note that although maximal lysine production has energy constraints with all the substrates considered, it has additional constraints of redox on glucose and stoichiometry on acetate. Therefore, there exists the potential for a synergistic effect in the presence of multiple substrates that could further enhance the carbon conversion. Some precedent for the co-metabolization of glucose and acetate, at least during a transitional growth phase, does exist [90].

The actual optimal conversion of substrate into the desired product requires the manipulation of metabolic fluxes. For the case of lysine production the optimal flux distribution converting glucose into lysine is shown in Figure 6.2l. Achieving this flux distribution represents the goal of metabolic engineering a strain for lysine production. The use of alternate or mixed substrates also needs investigation as they can potentially ease the task of flux redirection for optimal performance.

For a deeper understanding of optimal biochemical production we need to identify metabolic constraints in an optimal solution. Biosynthesis of the product requires inputs of redox, energy, carbon units, and minerals. Redox and energy can be produced as a by-product during substrate conversion to the carbon skeleton. An inadequate production of redox and energy present a constraint that would require the oxidation of substrate in order provide the required redox and energy. In addition, specific metabolic reactions can result in the release or uptake of CO<sub>2</sub>, thus affecting the net carbon conversion. In the presence of adequate minerals, metabolic constraints during the production of specific biochemicals can be categorized in terms of stoichiometry, redox, and energy. These constraints determined according to the criteria in the truth table shown in Table 6.1 represent the demands or stress placed on cellular metabolism during biochemical over-production. Computations for glucose, glycerol and acetate substrates demonstrate that constraints on metabolism depend on the nature of the substrate in terms of its redox and energy content as well as its point of entry into the metabolic network. Of the various combinations of constraints observed it is interesting that redox is always associated with energy as a constraint. A large fraction of cellular metabolic energy requirements are met by oxidative phosphorylation which essentially converts redox into energy. Therefore, generally a redox constraint would result in a simultaneous energy constraint.

A balance between growth and biochemical production is often important for a successful bioprocess. The generation of biomass in the bioprocess is necessary to provide the backbone of metabolism used to achieve substrate conversion into the desired biochemical. Although the growth phase (trophophase) is often considered as separate from the product formation phase (idiophase), it has been demonstrated that simultaneous growth and product formation can indeed result in a trophophase-idiophase separation [23]. We have therefore computed the optimal trade-off between simultaneous growth and biochemical production. As expected, a negative correlation is observed in the trade-off. For most of the biochemicals considered here the trade-off is practically linear. However, some instances of non-linearity do exist that demonstrate a synergistic effect between biomass and biochemical production. Thus, the balanced growth and biochemical solution shows a higher efficiency compared to a simple addition of the individual solutions. An appropriate balance between growth and biochemical production may have particular importance for non-secreted products, such as for polyhydroxybutyrate accumulation in cells [61].

The present Chapter demonstrates an application of the flux balance analysis of *E. coli* metabolism towards the study of biochemical production. The limits of substrate conversion to biochemicals and the distribution of flux in the metabolic network represent the goals of metabolic engineering. Representing the limits of metabolic behavior the analysis is useful for evaluating and analyzing performance. The general conceptual framework presented here can be used to obtain a detailed analysis for a particular product as a guide to the development of a bioprocess.

# CHAPTER VII

## THE ROLE OF OXYGEN

Microbial cultures are often oxygen limited. Restrictions in oxygen supply can occur due to either capacity limits of the respiratory system or external supply restrictions. When oxygen is limited the carbon substrate is only partially oxidized which leads to by-product secretion. Secretion of metabolic by-products allows for the generation of energy while balancing cellular redox metabolism.

Stoichiometric analysis is well suited to address questions associated with the optimal secretion pattern of metabolic by-products under oxygen limitations. Here the answers to these questions are obtained for *E. coli* based on the principle of stoichiometric optimality. By-product secretion patterns due to the enzymatic limits of the respiratory chain are evaluated based on stoichiometric optimality. We also examine the effects of externally imposed oxygen limitations from completely anaerobic to aerobic conditions.

### 7.1 Model Considerations

During steady state cell culture the rate of oxygen consumption has to balance with the rate of oxygen supply. Mathematically this balance is expressed as:

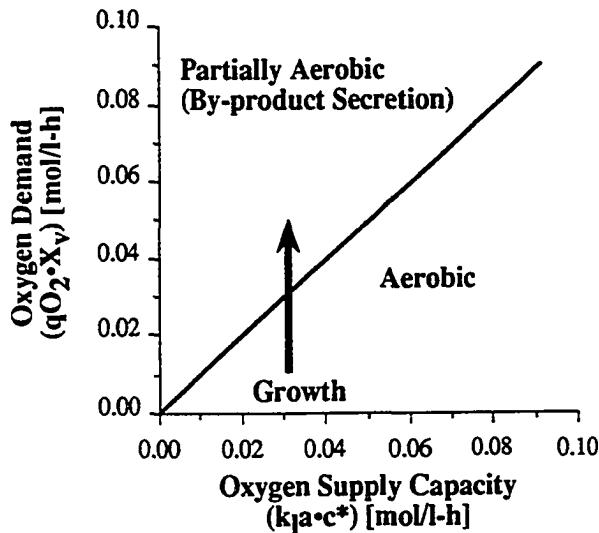


Figure 7.1: Oxygen demand of a cell culture in contrast to the supply rate. Typical cell culture apparatus use surface aeration, agitation and air sparging depending on the scale of operation. Cell growth results in oxygen demand outstripping supply capacity resulting in by-product secretion.

$$q_{O_2} \cdot X_v = k_l a \cdot c^* \quad (7.1)$$

where  $q_{O_2}$  is the per cell oxygen consumption,  $X_v$  is the viable cell density,  $c^*$  is the dissolved oxygen concentration, and  $k_l a$  is the mass transfer coefficient that is a characteristic of the particular cell culture equipment used.

Figure 7.1 depicts graphically the balance between oxygen demand and supply. The x-axis represents the oxygen mass transfer capacity of the equipment while the oxygen demand of the culture is shown on the y-axis. Typical cultivation is initiated at low cell concentrations, usually in the fully aerobic region. Cell growth results in increasing oxygen demand as indicated by the arrow in Figure 7.1. At a certain cell concentration the oxygen demand of the culture exceeds the supply capacity resulting

in a partially aerobic culture. Here we examine the effect of oxygen limitations on optimal by-product secretion and metabolic behavior.

**Basis** To introduce uptake and growth rates into the flux balance analysis, we need to introduce a basis. As a basis, consider a certain mass of cells in culture over some period of time such that the product of cell mass and time is 1 g DW-hr. The basis of 1 g DW-hr may represent 1 g of cells over a period of 1 hr or 10 g of cells over 6 minutes, etc. It is analogous to units such as man-hours and kilowatt-hours. This choice of basis allows all fluxes within the metabolic network to be expressed as mmol *per* g DW-hr. Thus, uptake limits can be specified in the metabolic network and hence define the metabolic capacity. As before, growth of the bacterium is defined as a flux draining biosynthetic precursors and cofactors in an appropriate ratio. In the context of our choice of basis the growth flux now becomes g biomass generated *per* g DW-hr which is the growth rate.

**Uptake Limits** The maximum oxygen uptake rate of *E. coli* cells growing on a variety of carbon sources in an aerobic culture was reported to be about 20 mmol O<sub>2</sub>/g DW-hr [1]. Since a surplus of nutrients was provided in this study one may consider this limit to represent the maximal respiratory capacity of the particular *E. coli* strain. Saturation type kinetic expressions of aerobic glucose consumption have been applied by several workers [15, 34, 77] with a maximum uptake of approximately 20 mmol Glc/g DW-hr. However, this limit can be exceeded under several growth conditions such as anaerobiosis. Although the above maximum uptake rates may vary for different strains as well as conditions of growth, for the present analysis it is only necessary to recognize that some upper limit on uptake exists.

**Maintenance Requirements** Having introduced time we can now account for both growth associated maintenance as well as the non-growth associated maintenance [49, 54, 59, 60, 82]. Substrate utilization for maintenance and growth is mathematically expressed as:

$$q = m_s + \frac{1}{Y_{max}} \times \mu \quad (7.2)$$

where  $m_s$  is the substrate requirement for non-growth associated maintenance activities,  $Y_{max}$  is the maximum biomass yield that includes the growth associated maintenance, and  $\mu$  is the growth rate. At 30 °C, under aerobic conditions,  $m_s$  has been observed to have numerical values in the range of 0.055-0.07 g Glc/g DW-hr [49, 59, 76]. The parameter  $Y_{max}$  represents the maximum possible yield and has the value of 0.524 g DW/g glucose [76]. These parameter values have been determined at 30°C and temperature has been found to have a major impact on the maintenance requirement [17, 45].

For the analysis presented here, non-growth associated maintenance requirements are specified as an equivalent ATP maintenance demand in mmol ATP/g DW-hr. Growth associated maintenance requirements are specified as an additional ATP requirement *per* unit biomass produced. These ATP requirements can be determined from a plot of glucose uptake versus optimal computed growth rate given the experimentally determined values for  $m_s$  and  $Y_{max}$ . Figure 7.2 shows glucose uptake as a function of the computed optimal growth rate. A linear relationship is obtained showing an agreement between Equation (7.2) and the flux balance model. The corresponding growth associated ATP requirements are 23 mmol ATP/g DW while the non-growth associated maintenance requirements are 5.87 mmol ATP/g DW-hr. In the metabolic network the growth associated ATP requirements are included along

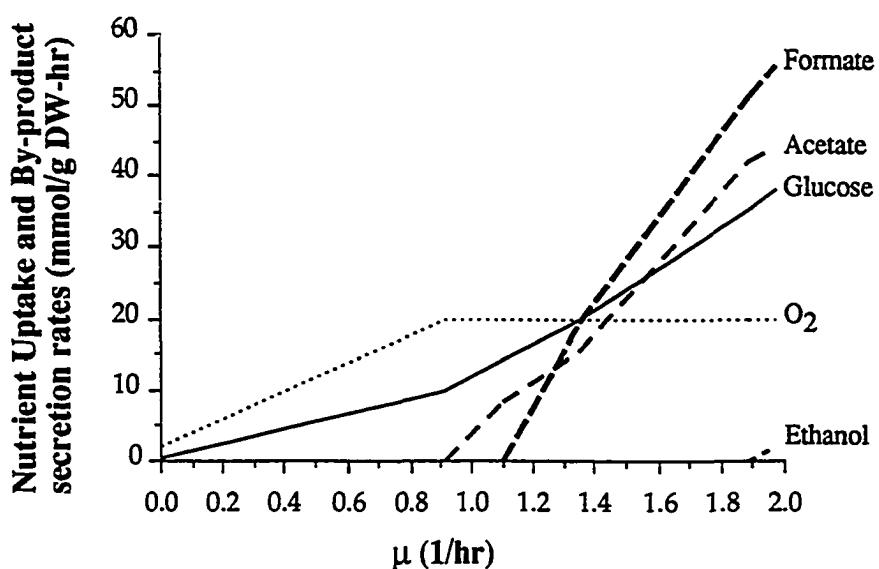


Figure 7.2: Optimal aerobic growth and by-product secretion of *E. coli* using glucose as a substrate. A maximal oxygen uptake rate of 20 mmol  $O_2$ /g DW-hr was used for these computations.

Table 7.1: By-product secretion follows a linear relationship with growth rate, Equation (7.3). Secretion parameters under the maximal aerobic oxygen uptake of 20 mmol O<sub>2</sub>/g DW-hr (Figure 7.2) are listed here.

By-Product	$\mu_c$ (hr <sup>-1</sup> )	$\nu_1$ (g/g DW)
Acetate	0.91	2.6
Formate	1.11	3.7
Ethanol	1.89	1.1

with the metabolic requirements for growth while the non-growth associated maintenance is specified as a drain flux of 5.87 mmol on the ATP pool since the basis corresponds to 1 g DW-hr.

## 7.2 Growth Dependence on Glucose Supply

The computed optimal growth rate of *E. coli* on glucose increases linearly with glucose utilization up to a growth rate of 0.9 hr<sup>-1</sup>, Figure 7.2, when oxygen uptake hits the respiratory constraint of 20 mmol O<sub>2</sub>/g DW-hr [1]. At higher growth rates oxygen supply is limiting and glucose cannot be fully oxidized to CO<sub>2</sub>. By-product secretion is observed as a means of eliminating surplus redox from the extra glucose. Acetate is the first by-product secreted as oxygen becomes limiting. At slightly higher growth rates formate is also secreted, Figure 7.2. At very high glucose uptake rates ethanol is also secreted in the optimal growth solution.

A piece-wise linear relationship between by-product secretion and growth rate is found, Figure 7.2. This relationship can be expressed in mathematical form as:

$$\nu = \begin{cases} 0 & \text{if } \mu < \mu_c \\ \nu_1(\mu - \mu_c) & \text{if } \mu > \mu_c \end{cases} \quad (7.3)$$

where  $\nu$  is the by-product formation rate in units of g by-product/g DW-hr. The

Table 7.2: Aerobic energy value of different by-products *per* unit oxygen consumed. Also listed are the energy and redox generated by the anaerobic catabolism of glucose into the by-products.

By-product	Energy content <i>per</i> mol by-product			Anaerobic production <i>per</i> mol glucose	
	ATP	O <sub>2</sub>	ATP/O <sub>2</sub>	ATP	NADH*
Acetate	4.67	2	2.33	4.00	2.00
Formate	1.33	0.5	2.67	-	-
Ethanol	8.33	3	2.78	2.00	-2.00
Lactate	8.33	3	2.78	2.00	0.00
Succinate	9.00	3.5	2.57	-2.33	-2.00
Glucose	18.67	6.0	3.11	0.00	0.00

\* The by-products acetate and ethanol also involve concomitant production of 2 moles of formate *per* mol glucose.

values for the parameters in Equation 7.3 for the various by-products are given in Table 7.1. The parameters in Table 7.1 correspond to a maximum oxygen uptake rate of 20 mmol O<sub>2</sub>/g DW-hr. An inability to supply this oxygen requirement to the culture would result in acetate secretion at lower growth rates resulting in different parameter values as discussed below.

A linear relationship between acetate secretion and growth rate have been observed under a variety of experimental conditions [4, 9, 19]. The numerical values reported for the two parameters vary;  $\mu_c = 0.14 - 0.52 \text{ hr}^{-1}$  and  $\nu_1 = 0.645 - 1.6 \text{ g acetate/g DW}$  [4, 19, 46]. The variety of parameter values is perhaps an indication of different oxygen uptake limitations as discussed below.

The sequence of optimal by-product secretion can perhaps be best explained by considering the energy and redox value of the by-products. Column 4 of Table 7.2 lists the energy value per unit redox of various by-products. Acetate has the

lowest energy to redox value and should therefore be secreted based solely on this criterion during oxygen limitations. As oxygen limitations become more severe we also have to consider the redox balance. Formate secretion in the degradation of pyruvate to acetyl CoA is a direct and simple means for eliminating surplus redox. Columns 5 & 6 of Table 7.2 list the energy and redox generated while producing the listed by-products. Although acetate secretion provides the maximum energy from substrate level phosphorylation, it also generates surplus redox. Ethanol and succinate are the only by-products that provide a suitable redox sink. While ethanol secretion generates net energy, succinate secretion requires energy. It is therefore not surprising that the optimal growth solution shows ethanol secretion as a means for eliminating surplus redox.

### 7.3 Growth on Glucose with Oxygen Limitations

The optimal growth rate is computed to be a piece-wise linearly increasing function with increasing the oxygen uptake rate, Figure 7.3. A glucose uptake rate of 10 mmol Glc/g DW-hr was arbitrarily chosen for these computations, however, the results do scale with other glucose uptake rates as is described later. There are four breakpoints in the curve dividing metabolic behavior into five phases. Table 7.3 lists the shadow prices for several metabolites in the five phases. We now discuss the five phases individually and discuss changes in metabolic behavior.

**PHASE I.** Under completely anaerobic conditions an optimal growth rate of  $0.26 \text{ hr}^{-1}$  is obtained, assuming the glucose uptake rate of 10 mmol Glc/g DW-hr, resulting in a biomass yield of 0.026 g biomass/mmol glucose. The flux distribution shows the well accepted anaerobic pattern of pathway utilization including the use of the anaerobic pyruvate formate lyase, Figure 7.4a. The non-use of the

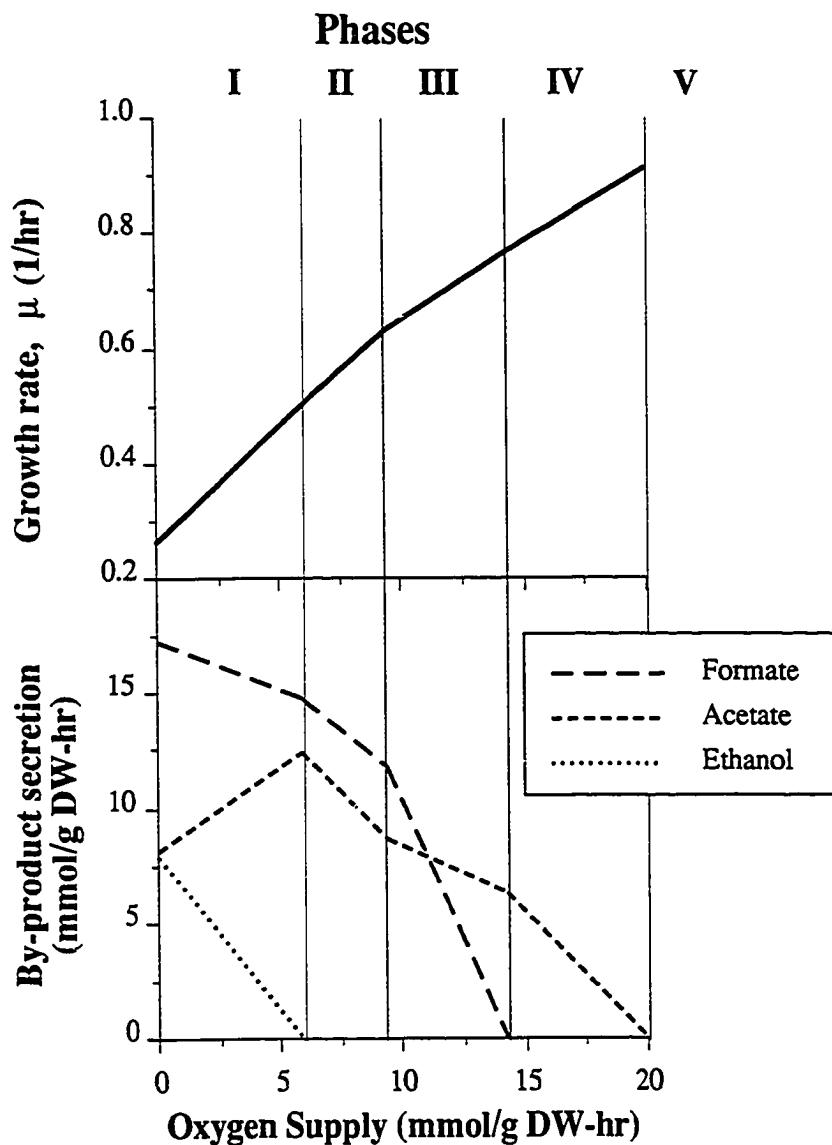


Figure 7.3: Optimal by-product secretion and growth under the complete range of oxygenation rates. Five distinct phases of oxygenation are seen based on the growth characteristics and by-product secretion patterns. A glucose uptake rate of 10 mmol Glc/g DW-hr was used for these computations.

Table 7.3: Shadow prices computed from the dual optimization problem for selected metabolites under different oxygenation conditions. Shadow prices measure the marginal importance of a metabolite towards accelerating growth. The columns list the shadow prices under the oxygenation conditions of Figure 7.4.

Metabolite	Oxygen uptake (mmol/g DW-hr)				
	0	7	12	16	20.05
$\mu$ (hr <sup>-1</sup> )	0.262	0.538	0.701	0.809	0.913
O <sub>2</sub>	0.0399	0.0370	0.0283	0.0257	0.0000
ATP	0.0109	0.0107	0.0106	0.0104	0.0049
Hexp*	0.0036	0.0036	0.0035	0.0035	0.0016
NADPH	0.0018	0.0029	0.0035	0.0038	0.0079
NADH	-0.0054	-0.0043	0.000	0.0010	0.0065
FADH	-0.0127	-0.0114	-0.0071	-0.0059	0.0033
QH <sub>2</sub>	-0.0127	-0.0114	-0.0071	-0.0059	0.0033
Acetate	0.0000	0.0000	0.0000	0.0000	0.0242
Ethanol	0.0000	0.0021	0.0106	0.0125	0.0422
Formate	0.0000	0.0000	0.0000	0.0010	0.0065
Lactate	0.0054	0.0064	0.0106	0.0125	0.0422
Succinate	0.0109	0.0128	0.0177	0.0170	0.0504

\* Hexp represents the energy of the transmembrane proton gradient as the translocation (export) of one proton across the membrane.

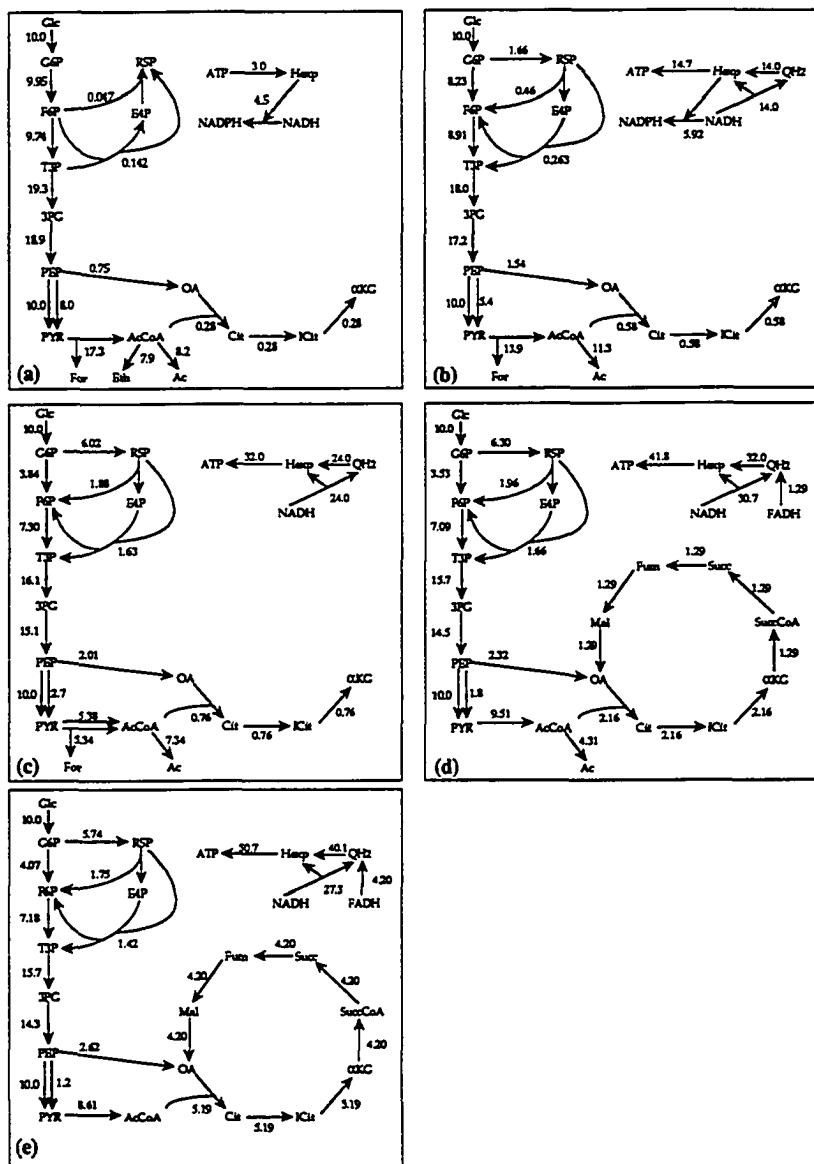


Figure 7.4: Optimal flux distributions characteristic of the five different phases of oxygenation shown in Figure 7.3. Biosynthetic precursors and cofactors are also used for growth as evidenced by the net generation of these metabolites. A glucose uptake of 10 mmol Glc/g DW-hr has been used for these computations. The various oxygenation rates in mmol O<sub>2</sub>/g DW-hr are: a) 0 b) 7 c) 12 d) 16 e) 20.5.

Table 7.4: Comparison of experimental and optimal by-product secretion rates and biomass yield for anaerobic growth of *E. coli* on glucose.

Product (mmol/mmol Glc)	Optimal	Experimental [6]
Acetate	0.815	0.75
Ethanol	0.786	0.87
Formate	1.73	1.13
Lactate	0.00	traces
Succinate	0.00	0.12
Yield (g DW/mmol Glc)	0.026	0.03

TCA cycle in the optimal anaerobic solution has also been experimentally observed. We also observe that the majority of the biosynthetic redox potential is produced by transhydrogenation—from NADH to NADPH. Acetate, ethanol and formate are secreted. Experimentally obtained by-product secretion patterns vary considerably depending on the culture conditions [6, 7]. Table 7.4 lists an experimentally observed by-product secretion rates as well as the optimal by-product secretion pattern along with the corresponding biomass yield. A close similarity is seen between the optimal by-product secretions and the experimental observation indicating that anaerobically growing *E. coli* may follow stoichiometric optimality. The observed variation in by-product secretion is probably a function of the ecological niche of the organism as well as a response to specific culture conditions.

An interpretation of optimal metabolic behavior under anaerobic conditions can be obtained using the shadow prices listed in Table 7.3. The three secreted by-products; acetate, ethanol, and formate are found to have a zero shadow price indicating that they are useless to improve growth. Succinate and lactate have a positive

shadow price and are therefore not secreted in the optimal growth solution. The three redox carriers; NADH, FADH, and QH<sub>2</sub> have a negative shadow price indicating a desire to eliminate surplus redox.

It is interesting, however, that biosynthetic redox, NADPH, is desirable as indicated by its positive shadow price. In the metabolic network biosynthetic redox can be produced from NADH by transhydrogenation, a process that utilizes the energy of the proton gradient. Since transhydrogenation is indeed used in the optimal solution, Figure 7.4a, we find the shadow price of NADPH to be the sum of the shadow prices of NADH and two protons translocated (2-Hexp), Equation (7.4). As the energetic value of NADPH is greater than its redox undesirability we find that NADPH has a net utility to the cell.

$$\begin{aligned}\gamma_{NADPH} &= \gamma_{NADH} + 2.0 \times \gamma_{Hexp} \\ 0.0018 &= -0.0054 + 2.0 \times 0.0036\end{aligned}\tag{7.4}$$

**PHASE II.** The next phase is typified by the cessation of ethanol secretion, Figure 7.3. Acetate and formate are the only by-products secreted. The corresponding flux distribution is shown in Figure 7.4b. The cytochromes are used for the terminal transfer of electrons to oxygen. In this phase we observe the contribution of the pentose phosphate pathway to biosynthetic redox generation, though transhydrogenation is still the major pathway for biosynthetic redox generation.

ATP shadow prices listed in Table 7.3 show a decrease with increasing aerobiosis indicating that ATP has a greater importance under anaerobic conditions. Except for biosynthetic redox potential other redox is still undesirable for growth as indicated by their negative shadow price. Of the secretable products in Table 7.3 we

observe that now ethanol has value to the cell and is therefore not secreted. It is surprising that ethanol has value to the cell since ethanol secretion provides the means for eliminating NADH which has a negative value to the cell. This observation is explained by considering acetate secretion (producing one ATP) as an alternative to ethanol secretion. Thus, the positive shadow price of ethanol is the sum of the shadow prices of one ATP and two NADH, Equation (7.5). In other words, the usefulness of generating energy from acetate secretion outweighs the usefulness of ethanol secretion to eliminate redox.

$$\begin{aligned}\gamma_{Ethanol} &= \gamma_{ATP} + 2.0 \times \gamma_{NADH} \\ 0.0021 &= 0.0107 + 2.0 \times -0.0043\end{aligned}\tag{7.5}$$

**PHASE III.** At an oxygen supply rate of 12 mmol O<sub>2</sub>/g DW-hr we observe another phase of oxygenation, although secretion of acetate and formate still occurs. This phase is characterized by a zero shadow price of NADH, shown in Table 7.3, indicating that the cell is no longer forced to eliminate NADH. The energy poor redox carriers, FADH and QH<sub>2</sub> are however still undesirable for growth. Inspecting the corresponding flux distribution shown in Figure 7.4c we observe the utilization of both pyruvate dehydrogenase and pyruvate formate lyase for the conversion of pyruvate to AcCoA. Thus, the cell now deliberately produces NADH in this conversion rather than eliminating redox as formate. This phase therefore depicts the first utilization of the aerobic pyruvate dehydrogenase enzyme. We also note that transhydrogenation is no longer used for biosynthetic redox generation and almost all the biosynthetic redox requirements are met by the pentose phosphate pathway.

**PHASE IV.** Increasing the oxygen supply to 16 mmol O<sub>2</sub>/g DW-hr we observe the

'aerobic' phase. The optimal flux distribution corresponding to this oxygen supply is shown in Figure 7.4d. We now observe the utilization of the complete TCA cycle for energy generation. The anaerobic pyruvate formate lyase is no longer used. Acetate is now the only by-product secreted. The corresponding shadow prices shown in Table 7.3 indicate that acetate with a zero shadow price is the only non-desirable by-product. Formate and NADH now show a positive utility for growth. The low energy redox carriers; FADH and QH<sub>2</sub> are still undesirable with a negative shadow price.

**PHASE V.** The last phase of oxygenation is that of adequate oxygen supply. Figure 7.4e shows the optimal flux distribution corresponding to an oxygen uptake of 20.05 mmol O<sub>2</sub>/g DW-hr. No by-product is formed and the entire substrate is utilized for growth. The TCA cycle is the major source for energy through oxidative phosphorylation. Almost equal amounts of biosynthetic redox are produced by isocitrate dehydrogenase and the pentose phosphate pathway. The corresponding shadow prices shown in Table 7.3 are positive for all the metabolites indicating no surpluses. Oxygen has a zero shadow price indicating an adequate supply. All redox molecules now show a positive shadow price indicating their utility for growth in proportion to their conversion to energy by the respiratory chain.

Although the above computations of the five oxygenation states utilized a glucose supply of 10 mmol Glc/g DW-hr the same qualitative results hold for other glucose supply rates. Figure 7.5 shows the different phases of oxygenation as a function of the glucose supply rate. The oxygenation condition is determined by the combination of the glucose and oxygen supply rates.

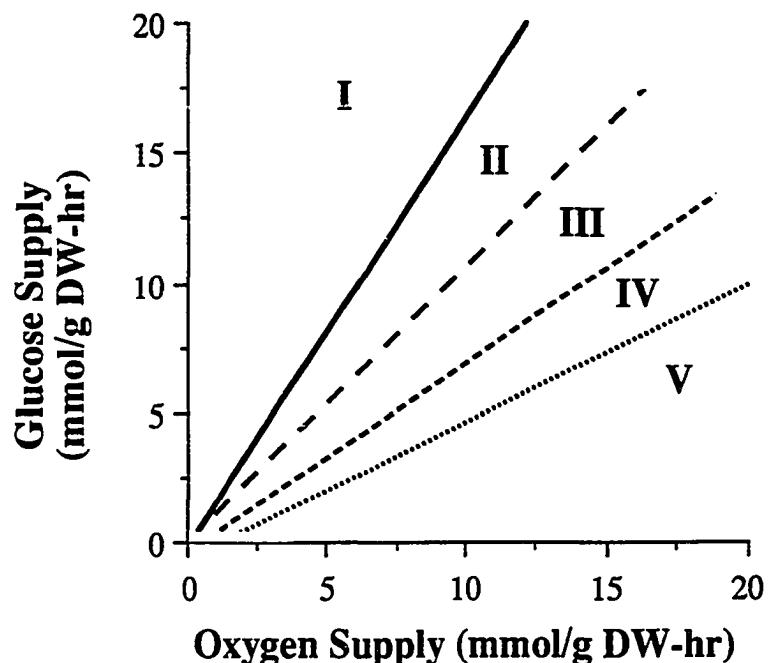


Figure 7.5: The five phases of oxygenation shown as a function of the glucose and oxygen supply. A minimum glucose supply of about 0.31 mmol Glc/g DW-hr is required to fulfill the non-growth associated maintenance energy requirements.

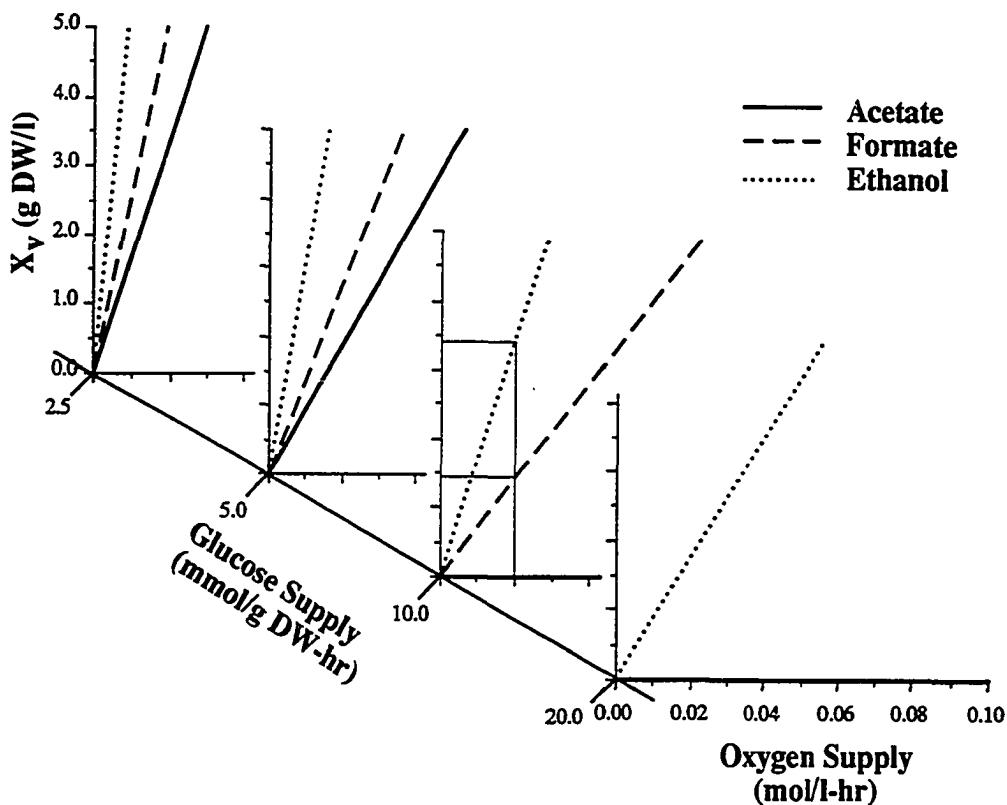


Figure 7.6: By-product secretion shown as a function of the cell concentration and glucose and oxygen supply. Cell concentrations that exceed the by-product secretion lines result in the production of by-product. Fully aerobic cultures may be found only below the acetate secretion line. At high glucose supply rates enzymatic limitations prevent a sufficient oxygen uptake rate and by-products are secreted at all cell concentrations as illustrated by the acetate and formate lines merging with the x-axis at a glucose level of 20 mmol Glc/g DW-hr.

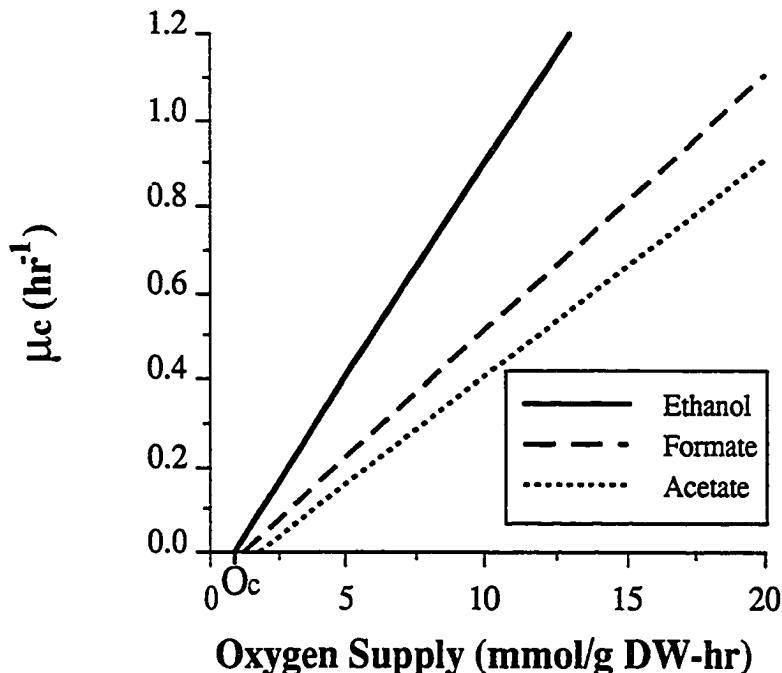


Figure 7.7: Variation of critical growth rate,  $\mu_c$ , for various by-products under different oxygenation rates.

**The relationship of cell concentration, glucose and oxygen supply to by-product secretion** We return now to the issue raised earlier; namely, what is the balance between glucose and oxygen supply rates and the cell concentration at which by-product secretion occurs. Oxygenation in typical cell culture equipment is subject to constraints as determined by the mass transfer coefficient ( $k_{la}$ ). On the other hand the glucose supply is a variable that can easily be controlled. For a fixed glucose supply, when oxygen demand increases due to cell growth and exceeds the oxygen supply, then by-product secretion is observed.

Figure 7.6 summarizes the relationship between the three quantities: oxygen supply, glucose supply, and biomass concentration. It shows the conditions at which by-product secretion occurs. In general, acetate is the first by-product to be secreted, followed by formate, and finally ethanol. At high glucose uptake rates, such as 10 and

20 mmol Glc/g DW-hr the oxygen consumption is enzymatically limited. Therefore by-product secretion occurs at all oxygen supply constraints as indicated by the by-product secretion lines merging with the x-axis in Figure 7.6. To give a specific example; if the glucose supply rate is 10 mmol Glc/g DW-hr and the oxygenation is limited to 0.02 mol/liter/hr then acetate is always secreted, formate is secreted at a cell concentration of 1.5 g DW/liter and ethanol is secreted at 3.5 g DW/liter.

**What are the limits of anaerobic growth?** This interesting question arises upon the examination of the above results. The answer can be found by returning to Figure 7.2. The computations in this figure are for the maximal uptake rate of oxygen and we observe critical growth rates ( $\mu_c$ ) at which by-product secretion occurs. These critical growth rates do vary with the oxygen uptake rate.

Computation of the critical growth rates with varying oxygen uptake rates demonstrate a linear relationship for all three by-products, Figure 7.7. Interestingly, the critical growth rate for ethanol reaches zero at finite oxygen uptake rates,  $O_c = 0.88 \text{ mmol O}_2/\text{g DW-hr}$ . This result means that at any growth rate at or below  $O_c$  all three by-products are secreted. Thus, this condition can be defined as the limit of anaerobic growth.

## 7.4 Discussion and Conclusions

With growth defined as composition-based metabolite requirements, the maintenance requirements are determined from the established relationship between glucose uptake rate and growth rate. We obtain a growth associated maintenance of 23 mmol ATP/g DW and a non-growth associated maintenance of 5.87 mmol ATP/g DW-hr. These values for maintenance are based on growth data obtained at 30 °C and likely to vary with temperature [17, 45].

The growth associated maintenance costs can be compared to the 41.26 mmol of ATP that are stoichiometrically required to synthesize 1 g biomass from biosynthetic precursors [32]. Thus, ATP maintenance requirements for growth are approximately 50% of those required for biosynthesis and polymerization. Under normal growth conditions the non-growth associated maintenance of 5.87 mmol ATP/g DW-hr represents a small fraction of the energy requirements for growth. These ATP maintenance requirements compare well with the reported requirements of 6.8 mmol ATP/g DW-hr for non-growth associated maintenance and 71 mmol ATP/g biomass for growth associated total ATP requirements for the anaerobically growing organism *Aerobacter aerogenes* [81].

With the inclusion of both growth associated and non-growth associated maintenance requirements we can compute the stoichiometrically optimal patterns of metabolic by-product secretion under oxygen limitations. Under varying levels of glucose uptake we observe an optimal secretion of by-products. The secretion is linear with growth rate, occurring above a critical growth rate. Within experimental ranges of cell culture similar experimental observations of acetate secretion have been reported [4, 9, 19] and the concept of a restriction in either the respiratory system or a key TCA cycle enzyme has been suggested [46]. An increasing oxygen limitation results in the optimal secretion of acetate, formate, and ethanol in that order. The sequence of secretion is explained on the basis of the energy to redox value of the by-products as well as their capability to provide a sink for redox potential.

The critical growth rate determined for the by-products acetate, formate, and ethanol shows a linear increase with the oxygen supply rate. Above the critical growth rate for ethanol we observe the secretion of all the by-products typical of an anaerobic culture. Interestingly, below a critical oxygen supply of  $O_c = 0.88$  mmol

$\text{O}_2/\text{g DW-hr}$  we observe all by-products secreted at all growth rates. This critical oxygen supply corresponds to the non-growth associated maintenance energy requirements of the bacterium.

Oxygenation in a typical cell culture equipment is often limited by mass transfer constraints. Typical mass transfer limits of oxygenation fall in the range of 0.05-0.1 mol  $\text{O}_2/\text{l-hr}$ . Thus, a growing culture provided with sufficient glucose faces external limits on oxygen supply. Above a critical cell concentration the oxygen demand of the culture outstrips supply and by-product secretion results. Therefore, depending on the cell concentration a glucose limited culture may face limits on the oxygen mass transfer and thus become oxygen limited. Using the flux balance based approach we are able to predict oxygen limitations and the secretion of by-products.

The optimal growth solution under a variety of oxygenation conditions shows several shifts in metabolic behavior. These phases of oxygenation are defined by changes in the value of different redox carriers. Since the redox balance is coupled to several pathways in the metabolic network, we also observe shifts in metabolic pathway utilization. Several physiological observations are also related to the supply of carbon source and oxygen. In the presence of plentiful glucose, flux redistribution is seen to occur, resulting in the inhibition of oxidative phosphorylation. It is probable that changes in the value of redox carriers, as illustrated by the flux balance based approach, are responsible for actuating the regulatory mechanisms that cause the observed changes in pathway utilization. Some effects on metabolic regulation have been observed during the transition from fully aerobic to partially aerobic fermentation under different glucose uptake rates [14, 29]. The shifts of metabolic behavior are externally observable in the form of changing by-product secretion patterns.

In conclusion, we find that physiological observations of metabolic behavior are

consistent with optimization of growth rate within stoichiometric constraints. We therefore expect the general principle of stoichiometric optimality to provide a physiological basis for the objectives of metabolic regulatory mechanisms. Stoichiometrically optimal metabolism provides a common basis for the interpretation of observed metabolic by-product secretion patterns under a variety of culture conditions and limiting nutrients.

## CHAPTER VIII

# STABILITY OF CELL POPULATIONS

The biological route for the production of chemicals utilizes cellular metabolism to convert a substrate into the desired product. An important limitation of industrial biochemical production is the population instability of an engineered strain that results from over-growth of a non-producing mutant population. As reported in literature, even a small growth advantage of non-producers can result in their rapid outgrowth of the engineered strain, e.g. [31, 40, 65]. Engineering of cells to redirect metabolic fluxes towards the desired product [2, 3, 80] produces an unnatural condition within the cell which may compromise stability of the engineered clone. It is thus desirable to provide culture conditions that favor growth of the engineered production strain.

Stability of recombinant cells in large-scale culture depends on strain, or genetic stability, as well as population stability [31]. Strain stability as represented by the mutation of the engineered cells is determined by strain properties, the nature of engineered changes, and the genetic techniques used. Strain stability is therefore affected by choices of vector, copy number, *etc.*, as well as process parameters such as temperature, pH, nutrient concentration, and so forth. Although several methods can be used to reduce the mutational frequency it is not generally possible to completely

eliminate mutations in a large scale industrial culture [20].

On the other hand, population instability refers to the outgrowth of the engineered strain by a mutant strain, resulting from the growth advantage of a mutant population over the engineered strain. Population stability may be affected by the environmental conditions, such as selective media.

In this Chapter we extend the population balance analysis to determine the total productivity of a culture based on economic considerations of production. We determine the impact of engineered productivity (with the resulting loss in growth) on population stability and total production. Optimal metabolism and growth rates of both producing and non-producing populations are determined by a flux balance model. Flux balance analysis has been used to determine the capabilities of microbial metabolism [18, 46, 73, 87, 91]. This approach has been successfully applied to *Escherichia coli* metabolism to study optimal growth patterns [88] as well as the production of biochemicals [85]. Here we use flux balance analysis to investigate simultaneous growth and biochemical production to determine population stability under various oxygenation conditions.

Our hypothesis is that the influence of oxygenation on the growth rate would be different for the engineered producing strain and the non-producing population of mutant strains. The rationale for this hypothesis is that different amounts of reductive power are required for generating biomass than for product formation. Therefore, a change in the oxygen supply which is the redox sink would differentially affect the growth rate of producer and non-producer cells. Computed productivity of valine, lysine, and tryptophan are used here to illustrate the enhancement in culture stability that can be achieved through a suitable control of the oxygen supply rate.

## 8.1 Analysis of Population Stability

Established differential equations, Equations (8.1-8.2), for describing the dynamics of population stability of producer and non-producer clones [31, 40, 65] were used in the present analysis:

$$\frac{dx_p}{dt} = \mu_p x_p - \alpha x_p \quad (8.1)$$

$$\frac{dx_n}{dt} = \mu_n x_n + \alpha x_p \quad (8.2)$$

where  $x$  is the cell number, the subscript ‘ $p$ ’ refers to the population of producing cells and ‘ $n$ ’ to the population of non-producing cells,  $\mu$  is the growth rate, and  $\alpha$  is the mutational rate of conversion of producers to non-producers. The dimensionless solution for the time varying population dynamics is also available [31, 40, 65]:

$$X_p(\tau) = e^{(1-A)\tau} \quad (8.3)$$

$$X_n(\tau) = \left[ \frac{A}{1 - A - S} \right] [e^{(1-A)\tau} - e^{S\tau}] + X_{no} e^{S\tau} \quad (8.4)$$

where  $\tau = \mu_p t$  is dimensionless time,  $A = \alpha/\mu_p = a/\ln(2)$ ,  $a$  is the mutational probability per cell division,  $S = \mu_n/\mu_p$  is the growth advantage of the non-producing strain, and  $X_{no}$  is the dimensionless initial non-producer cell number.

**Biochemical Production** The flux balance analysis with a growth rate maximization objective has been shown to simulate the growth and metabolic behavior of *E. coli* [46, 86, 87, 88]. We therefore use the flux balance model to determine the growth advantage of non-producers over the producing population ( $S = \mu_n/\mu_p$ ).

The by-product secretion pathways for acetate, formate, ethanol, lactate, and succinate have been included in the metabolic network to allow for by-product secre-

tion under partially aerobic conditions [46, 86, 78]. The choice to secrete any of these by-products is determined by the optimal growth criteria. In addition, although a glucose supply of 10 mmol Glc/g DW-hr has been specified for all the computations, the results can be approximately scaled for other glucose supply rates based on the oxygen to glucose supply ratio.

Starting a culture with an engineered conversion,  $C$  (mol product/mol substrate), the generation of non-producers results in a reduction of net conversion. The net productivity at any period of time can be estimated as  $(\frac{C X_p}{X_p + X_n})$ . An economic criteria for optimal culture termination in an industrial biochemical production can be formulated as a critical economic conversion. Thus, the decision to terminate the culture should occur when the net conversion falls below the predetermined threshold value based on the cost economics of production. Mathematically, the critical economic criteria may be stated as:

$$Y < \frac{C X_p}{X_p + X_n} \quad (8.5)$$

where  $Y$  is the critical conversion. The above result can be expressed in a non-dimensional form, Equation (8.6), by defining a dimensionless engineered conversion as  $P = C/C_{max}$  and a dimensionless economic conversion as  $\sigma = Y/C_{max}$ .

$$\frac{\sigma}{P} < \frac{X_p}{X_p + X_n} \quad (8.6)$$

$C_{max}$  represents the maximal theoretical conversion of substrate into the desired product. Since all the substrate is utilized to make the product the growth rate is zero during a  $C_{max}$  product formation rate. For the products considered here the value of  $C_{max}$  is: valine = 1, lysine = 0.78, tryptophan = 0.41 (mol/mol) [85].

**Assuming Linear Trade-Off** Some general results for designing specific productivity in strains can be derived by assuming a linear reduction in growth rate due to biochemical production in order to estimate the growth advantage [85], Equation (8.7).

$$\mu_p = \mu_n \left( 1 - \frac{C}{C_{max}} \right) \quad (8.7)$$

or in a dimensionless form,

$$S = \frac{1}{1 - P} \quad (8.8)$$

Based on the above economic considerations one can estimate the time for culture termination and hence the final producer cell number from Equations (8.3-8.4). A time integral of product formation upto the termination of the culture yields an estimate of the total production ( $TP$ ), Equations (8.9-8.10). The total production in terms of dimensionless parameters (assuming  $A \ll 1$ ) is given by Equation (8.11).

$$TP = \int_0^{t_f} r_s \cdot C \cdot X_p \, dt \quad (8.9)$$

$$= \frac{r_s \cdot C}{(1 - A) \mu_p} [e^{(1-A)\mu_p t_f} - 1] \quad (8.10)$$

$$= \frac{r_s \cdot C_{max}}{\mu_n} \frac{P}{1 - P} (X_{pf} - 1) \text{ if } A \ll 1 \quad (8.11)$$

where the culture termination time,  $t_f$ , and final cell number,  $X_{pf}$ , are determined from the above economic considerations, Equation (8.6). From an operational standpoint, the total production determined here is based on the assumption of operation in an exponential phase. Physical capacity limits of the equipment that result in an earlier culture termination are not considered here.

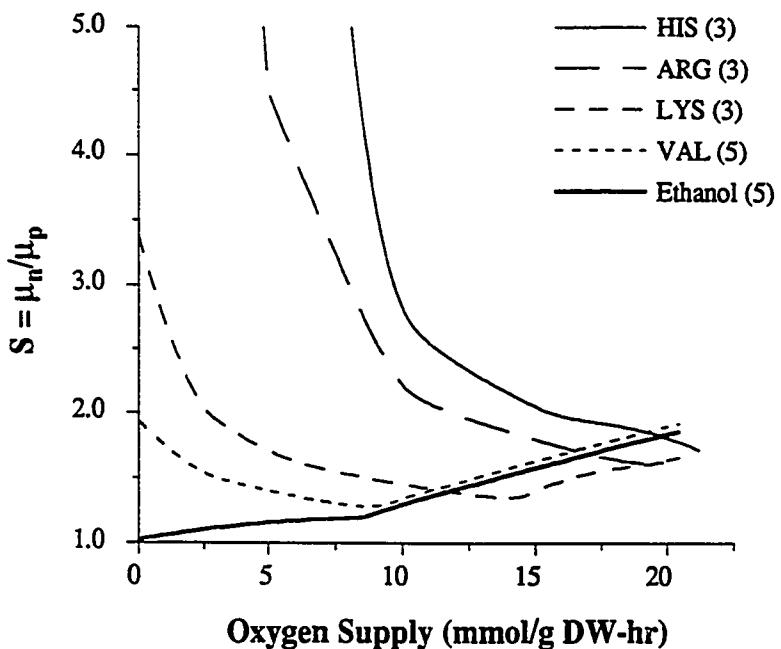


Figure 8.1: Culture stability represented by the growth advantage as a function of the oxygen supply for several biochemical products. A glucose supply rate of 10 mmol Glc/g DW-hr has been specified. The biochemical production (mmol/g DW-hr) listed for the respective plots.

## 8.2 Culture Stability

Flux balance analysis allows the computation of the optimal growth rate which is a good approximation [46, 86, 87, 88] to experimental observations. Reduction in the optimal growth rate due to product secretion can also be estimated from the flux balance model. Thus, we can determine culture stability as measured by the growth advantage of non-producers.

Let us first consider changes in culture stability for some illustrative producers of; histidine, arginine, lysine, valine, and ethanol as a function of the oxygen supply. The growth advantage ( $S$ ), determined using the flux balance model, as a function of the oxygen supply for a fixed glucose supply rate of 10 mmol Glc/g DW-hr is

shown in Figure 8.1. For some secreted products, the growth advantage has a minimum or optimal value at a particular oxygen supply rate (Figure 8.1). A lower growth advantage is desirable since it results in an increased longevity of the culture before outgrowth by a non-producing population. Operating at an oxygen supply corresponding to the minimum growth advantage can therefore result in an increased productivity from the culture over time.

The optimal oxygen supply rate is determined by the nature of the specific product and in particular with the redox requirements for product formation. To illustrate one may consider the simple example of ethanol production. Since ethanol formation has the largest redox requirements an ethanol producing population is most stable at low oxygen supply rates, as evidenced by the minimum growth advantage (Figure 8.1). However, most products have a complex relationship between the growth advantage and the oxygen supply rate as defined by the metabolic network.

### 8.3 Example: Valine Production

To further define the effects of variation in the oxygen supply rate, we now focus our attention on valine production. Figure 8.2 shows the variation in the growth advantage  $S$  as a function of the oxygen supply during valine production. A minimum growth advantage is found for all production rates of valine. For oxygen supply below this optimum, the growth rate of both producing and non-producing cells is reduced. However, since the growth rate of producer cells is lower the percentage drop in their growth rate with reduced oxygen supply is much greater leading to high numerical values of the growth advantage  $S$  at low oxygen supply rates. If oxygen supply is greater than the optimal value we find that the producer cells are not affected since they have an adequate oxygen supply. They are unable to utilize the

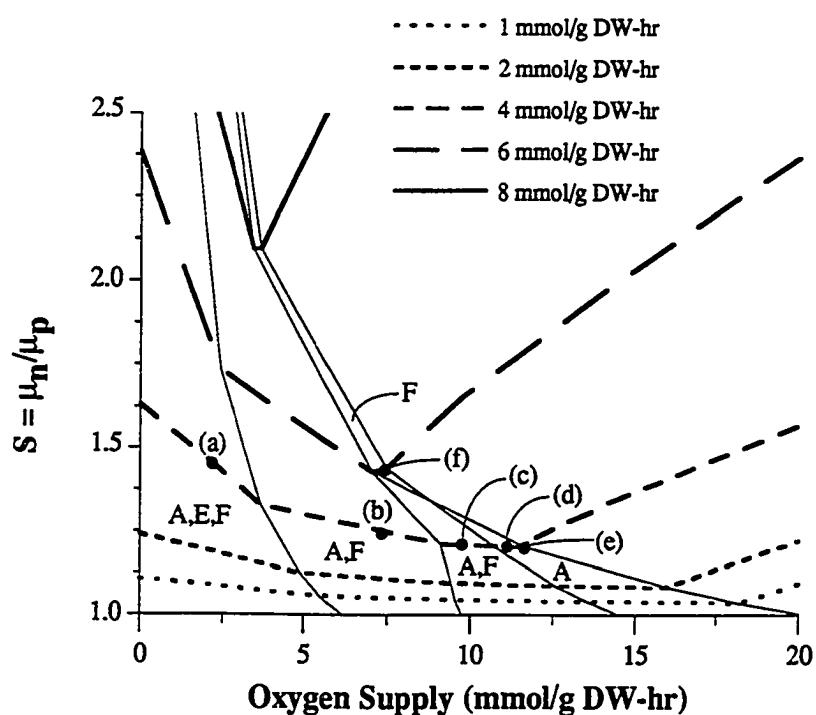


Figure 8.2: Culture stability during valine production as a function of the oxygen supply. The theoretical maximum valine production is 10 mmol VAL/g DW-hr. Production rates are indicated in units of mmol/g DW-hr. By-product secretion from the producer cells is shown in regions bounded by thin lines. The notation used for the secreted by-products is: A-acetate, E-ethanol, and F-formate. Flux distributions of Figure 8.4 are determined at the conditions indicated by a numbered dot.

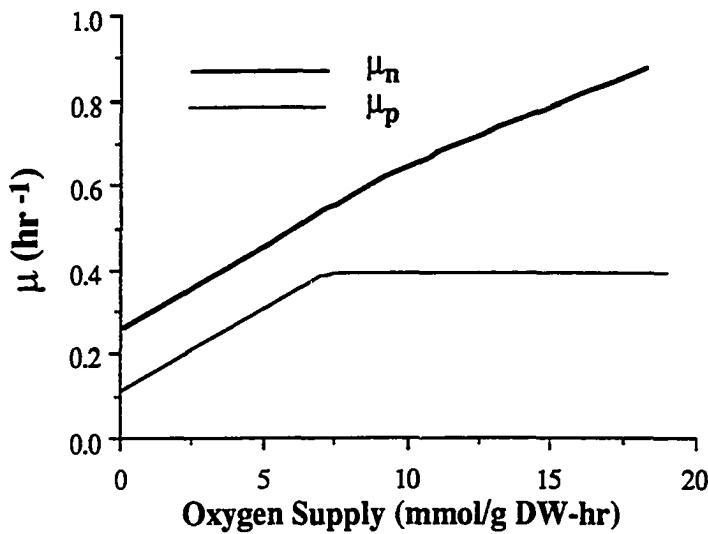


Figure 8.3: Variation of the growth rate of non-producers and valine producers as a function of the oxygen supply. A valine production of 6 mmol/g DW-hr was specified.

additional oxygen availability to enhance their growth rate. However, non-producer cells can still utilize oxygen to accelerate their growth rate. Hence we observe a higher numerical value of the growth advantage at an oxygen supply higher than the optimal value.

These changes in the growth rates of producer and non-producer populations with the oxygen supply are shown in Figure 8.3. While an increased oxygen supply results in a continuously increasing non-producer growth rate, the producer population reaches a maximum growth rate at an oxygen supply of 12.5 mmol O<sub>2</sub>/g DW-hr. Thus, the growth advantage increases with oxygen supply at high oxygenation. At low oxygenation levels while both growth rates increase with oxygen supply the percentage effect is greater for the producing population. Thus, the growth advantage falls with increasing oxygen supply at lower oxygen supply rates.

Different phases based on the particular by-products secreted during optimal

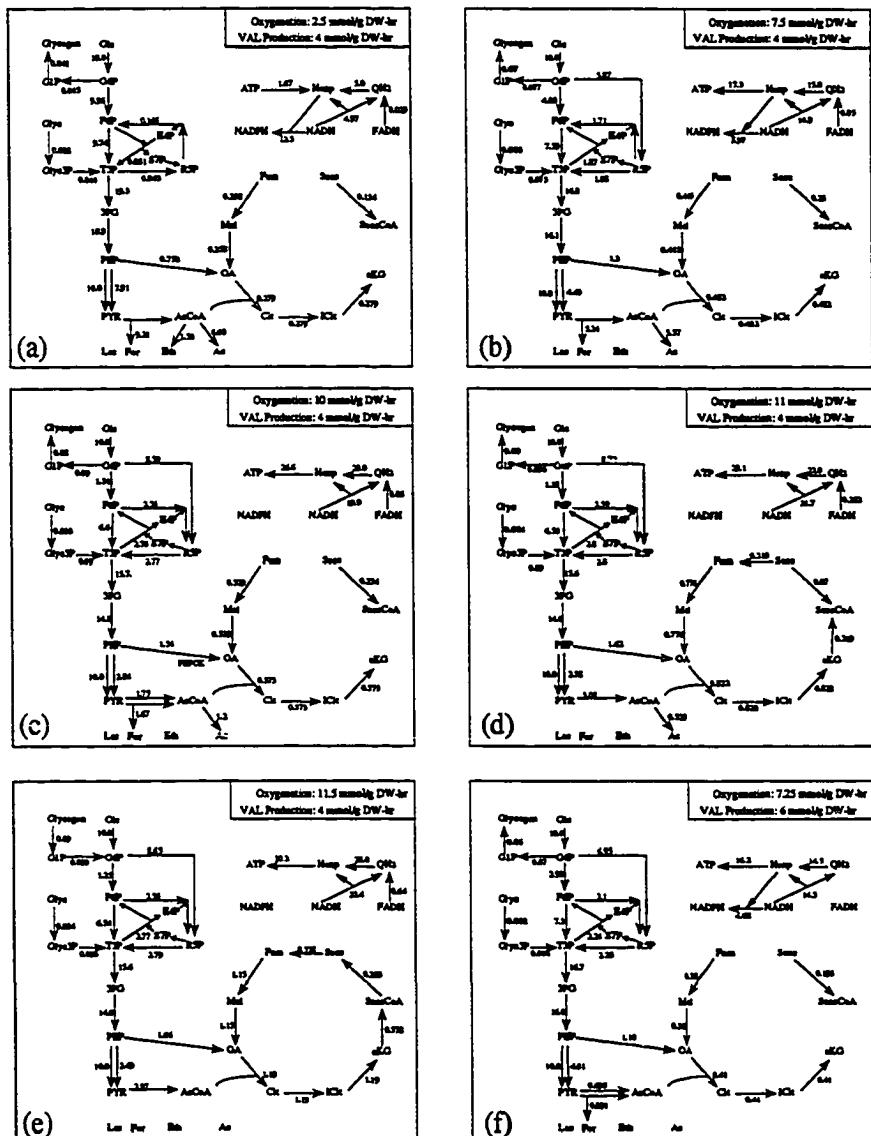


Figure 8.4: Flux distributions for the various phases of growth during valine production observed in Figure 8.2. A glucose supply of 10 mmol Glc/g DW-hr is specified and other parameters are listed in Table 8.1.

growth and valine production are marked in Figure 8.2. Catabolic flux distributions characteristic of these phases are shown in Figure 8.4. Under oxygen limitations we observe a change in the optimal metabolic pathways and the secretion of acetate, formate and ethanol as a means to eliminate surplus redox. While the TCA cycle is active at high oxygenation, Figure 8.4d,e, it is not operative during oxygen limitations. Also, as oxygenation increases, Figure 8.4a,b,c,d,e biosynthetic redox (NADPH) generation shifts from transhydrogenation to the pentose phosphate cycle. Thus, the different phases of by-product secretion, Figure 8.2, are characterized by changes in the optimal metabolic pathways.

By-product secretion under different oxygenations can also be explained by examining shadow prices for the various metabolites. Table 8.1 lists the shadow prices for various redox carriers, energy, and by-products typical of the different phases, Figure 8.2. Once again the shadow price reflects the utility of a metabolite to accelerate growth. Thus, a negative value for the shadow price in Table 8.1 represents the undesirability of a metabolite. By-products that are undesirable are secreted and hence have a zero shadow price.

The different redox carriers are positioned in the respiratory chain in a hierarchical manner such that they have different energy values. This fact is clearly brought out by the differences in their shadow prices which represents their value to the cell. Thus, the value of redox is ordered as:  $NADPH > NADH > QH_2$ . The various phases are characterized by the value of the redox shadow prices as well as the by-product shadow prices. As oxygen supply becomes limiting, Table 8.1(e)→(a) we observe a reduction in the value of redox (lower shadow price). As a result by-products are secreted in order to eliminating surplus redox. Under extreme oxygen limitations, Table 8.1(a), the three by-products: acetate, formate, and ethanol are secreted.

Table 8.1: Shadow prices of redox, energy, and by-products in the various phases of growth and by-product secretion observed in Figure 8.2.

	Flux Distributions in Fig. 8.4					
	(a)	(b)	(c)	(d)	(e)	(f)
Production rate mmol/g DW-hr	4	4	4	4	4	6
Oxygen Supply mmol/g DW-hr	2.5	7.5	10	11	11.5	7.25
$\mu$ (1/hr)	0.26	0.45	0.54	0.57	0.58	0.39
Metabolite	Shadow Prices - $\frac{\partial \mu}{\partial \text{metabolite}} \times 100$					
NADPH	0.183	0.289	0.359	0.389	0.808	0.522
NADH	-0.55	-0.433	0	0.106	0.673	0
QH <sub>2</sub>	-1.28	-1.15	-0.719	-0.602	0.337	-0.522
O <sub>2</sub>	4.03	3.75	2.88	2.62	0	2.09
Hexp*	0.367	0.361	0.359	0.354	0.168	0.261
Acetate	0	0	0	0	2.49	1.57
Formate	0	0	0	0.106	0.673	0
Ethanol	0	0.216	1.08	1.27	4.34	2.35
Lactate	0.55	0.649	1.08	1.27	4.34	2.35
Succinate	0.367	0.577	1.44	1.73	5.18	2.61

\* Hexp represents the energy of exporting one proton across the membrane.

A glucose supply of 10 mmol Glc/g DW-hr is specified;  
valine production rate, and oxygen supply rate for the different phases are as listed.

Thus, shadow prices as well as the optimal flux distributions show a metabolic pathway utilization characteristic of the various phases. These optimal flux distributions represent the goals of metabolic engineering to redirect catabolic fluxes towards the desired product.

### 8.3.1 Engineering Optimal Conversion

During production of a biochemical by cells in culture an over-growth of producing cells by non-producing cells is often found to occur over time. As a result there is a reduction in productivity with lower product yields as the culture progresses. Thus, the number of useful culture generations is limited by the over-growth of non-producing cells. An optimal economic threshold for culture termination occurs when the net productivity of the culture falls below a predetermined marginal conversion based on the cost economics of production, Equation (8.6).

As an example consider the total valine production possible, under the economic constraints outlined above, using cells of different engineered productivities, Figure 8.5. The arbitrarily chosen economic threshold conversion of 0.4 mol valine/mol glucose denotes the conversion above which production is profitable. Thus, all strains engineered for a conversion greater than 0.4 mol/mol are capable of an economically viable production.

A comparison of full oxygen supply to the optimal oxygen supply for culture stability shows that the total production can be increased by several orders of magnitude by a suitable control of oxygen supply, Figure 8.5. Thus, optimal strain engineering as well as the selection of optimal culture conditions can have a tremendous impact on the production attainable in cultures.

Figure 8.6 displays the average productivity of a culture as a function of time.

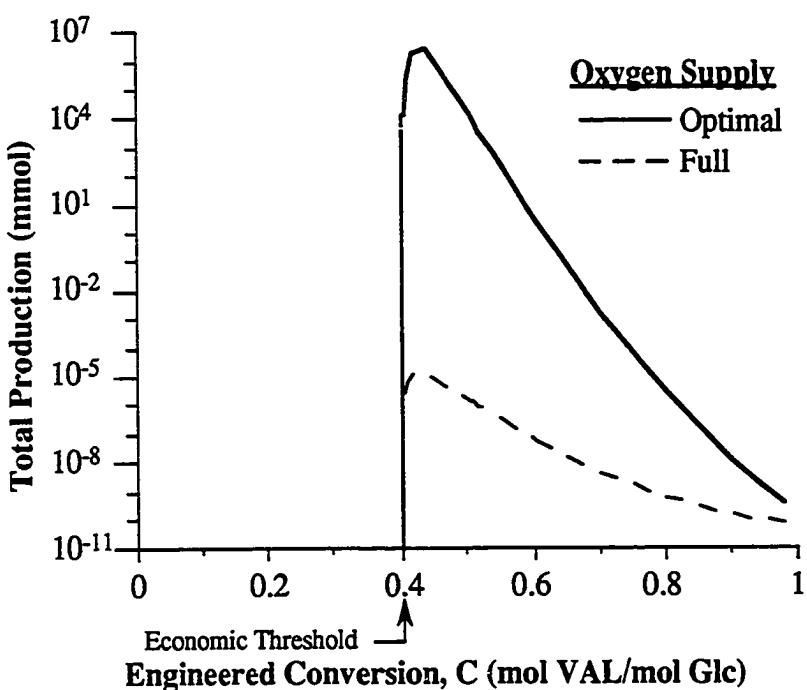


Figure 8.5: Total valine production in a culture as a function of the engineered specific productivity under the economic constraints that the marginal conversion should not fall below 0.4 mol VAL/mol Glc. A mutational rate of  $a = 10^{-6}$  per cell division was assumed. The total production was calculated from Equation (8.11).

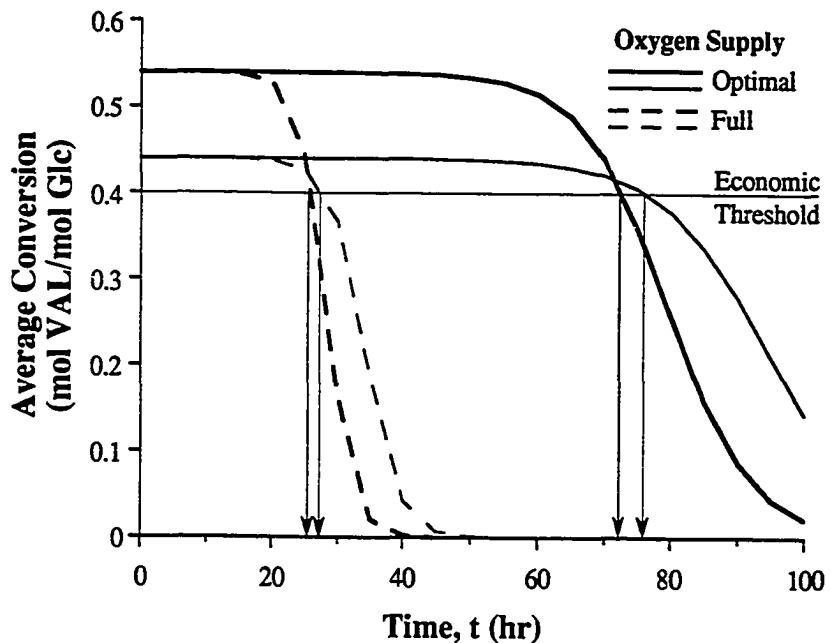


Figure 8.6: Average specific conversion of cultures computed as a function of time. The cultures start with single engineered clones producing 0.44 mol VAL/mmol Glc and 0.54 mol VAL/mol Glc. Optimal oxygenation conditions are shown to have a large impact on the production stability of the culture. A mutational rate of  $a = 10^{-6}$  per cell division has been assumed. The termination times are indicated by arrows on the x-axis.

It is evident from Figure 8.6 that control of oxygenation can have a tremendous impact on stabilizing the productivity. Thus, optimal oxygenation provides a longer economic process time for both cell growth, and valine production resulting in the larger total production.

In contrast the engineered productivity has a smaller impact on the time for culture termination, Figure 8.6. However, cell concentration reaches large values only towards the end of the culture at which time most of the production occurs. Therefore, even the small difference in time causes a large change in the total production achievable as evidenced in Figure 8.5. Maximal total production is observed to occur at an engineered productivity very close to the economic threshold conversion. Engineering productivity above this optimal value is counter-productive since there is an extremely rapid decay of the total production.

#### 8.4 Example: Tryptophan and Lysine

To further demonstrate the principle of oxygenation control as a means to enhance culture productivity we also consider the production of the commercially important amino acids; tryptophan and lysine. The culture stability parameter,  $S$ , is shown at various production levels for these amino acids under different oxygenation conditions, Figure 8.7. Once again we note that a reduced oxygen supply can result in an increased population stability of the producer cells.

Production levels achievable for the two amino acids are shown in Figure 8.8 based on the arbitrary economic criteria shown. Again optimal oxygenation is capable of enhancing productivity over several orders of magnitude. Furthermore, an optimal strain engineering strategy requires that specific productivity be engineered close to the economic threshold conversion in order to maximize total production.

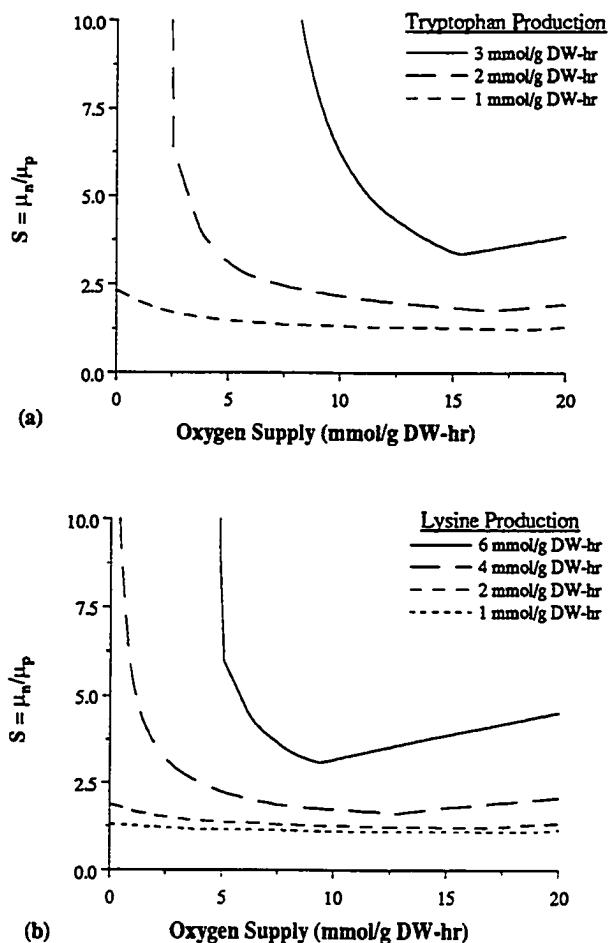


Figure 8.7: Culture stability represented by  $S$  given as a function of the oxygen supply to the cells for (a) tryptophan, and (b) lysine production. A glucose supply of 10 mmol Glc/g DW-hr was used for the computations.

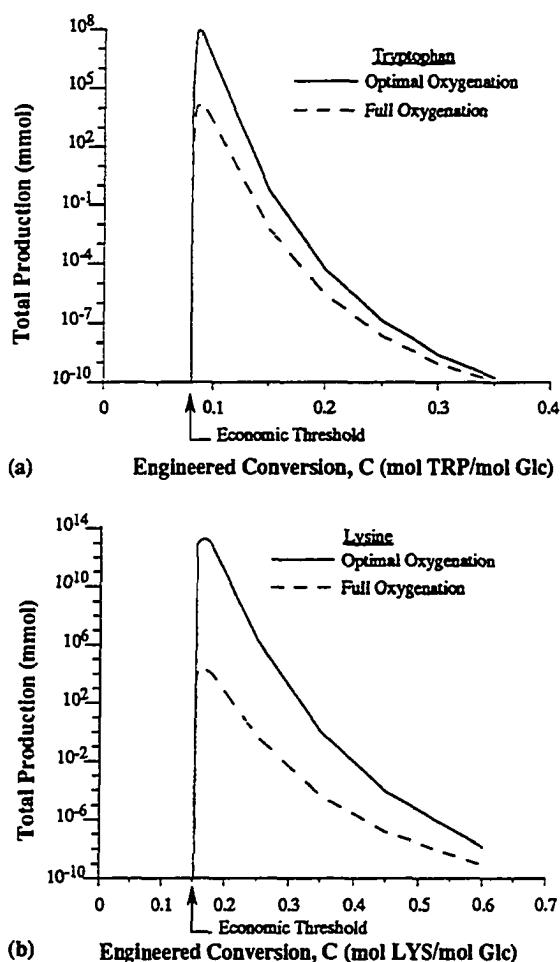


Figure 8.8: Total productivity of a culture shown as a function of the specific productivity of the engineered clone with a glucose supply of 10 mmol Glc/g DW-hr for (a) tryptophan, and (b) lysine production. The criteria for termination of the culture is that the marginal productivity falls below 0.8 mmol TRP/g DW-hr for tryptophan and 1.5 mmol LYS/g DW-hr for lysine. A mutational rate of  $a = 10^{-6}$  per cell division was assumed.

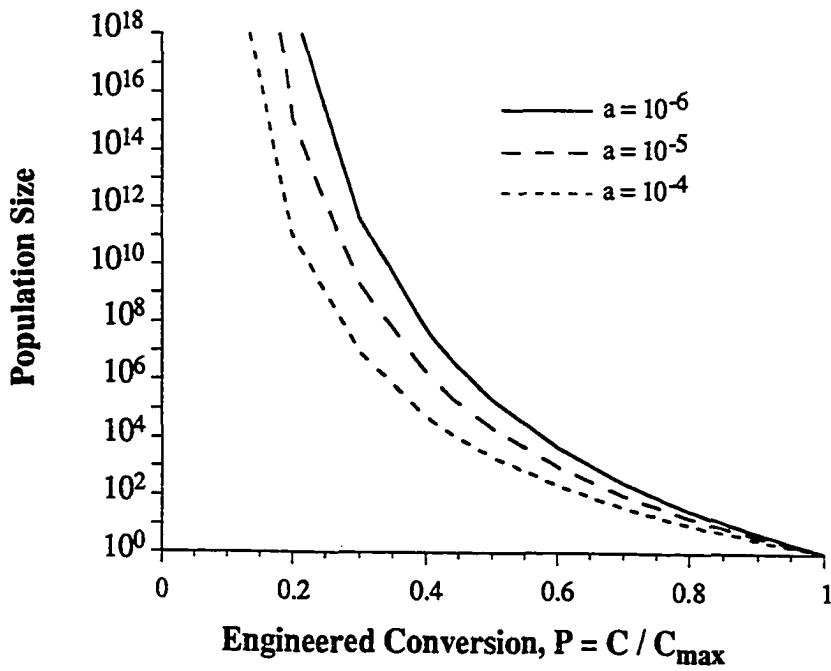


Figure 8.9: Culture size, as the number of producing cells obtainable, for a culture to become 25% non-producing plotted as a function of the dimensionless engineered productivity. The mutation probability ( $a$ ) used in the computations is indicated.

## 8.5 General Principles

Independent of the choice of product, some general principles of culture stability may be derived from the analysis just presented. Synthesis of any product results in the diversion of substrate from producing growth. Hence it is reasonable to assume that a practically linear reduction in growth rate can be obtained during product secretion [85], as expressed by Equation (8.7). Thus, the reduction in growth rate or the growth advantage can be estimated from the engineered productivity.

With the above estimation of the growth advantage we use population dynamics to estimate the culture size. Figure 8.9 shows the producer population size achievable before 25% of the total population convert to non-producers. The probability of

mutation,  $a$ , is seen to have a large influence on the culture size. Furthermore, the culture size decays rapidly with increasing engineered productivity. Thus, at a high productivity a reasonable producer population size can only be obtained by the use of selective conditions.

We next consider the total production possible under various conditions of engineered productivity and economic criteria. Figure 8.10 displays this relationship using the dimensionless parameters described in the methods section. Given that production costs, and the economics of the market specify the minimum threshold conversion, we find that the maximal total production is very sensitive to the mutational probability,  $a$ . We also find that optimal strain engineering calls for an engineered production rate close to the economic criteria. Indeed over-engineering productivity to any extent may be highly undesirable as seen in the rapid drop in total production.

The total production is observed to be infinite at the maximal dimensionless engineered productivity of unity ( $P = 1$ ), Figure 8.10. At this maximal production the growth rate is reduced to zero which eliminates the problems of culture stability which are growth related. Thus, the infinite production possible without growth corresponds to an alternative production technology of separating the growth and production phases.

## 8.6 Discussion and Conclusions

An important limitation of large-scale biochemical production is the population instability of the engineered cells due to mutations and overgrowth by non-producing populations. Loss of productivity in the culture sets an economic limit on culture longevity or culture size. Laboratory strain screening with non-selective or poorly

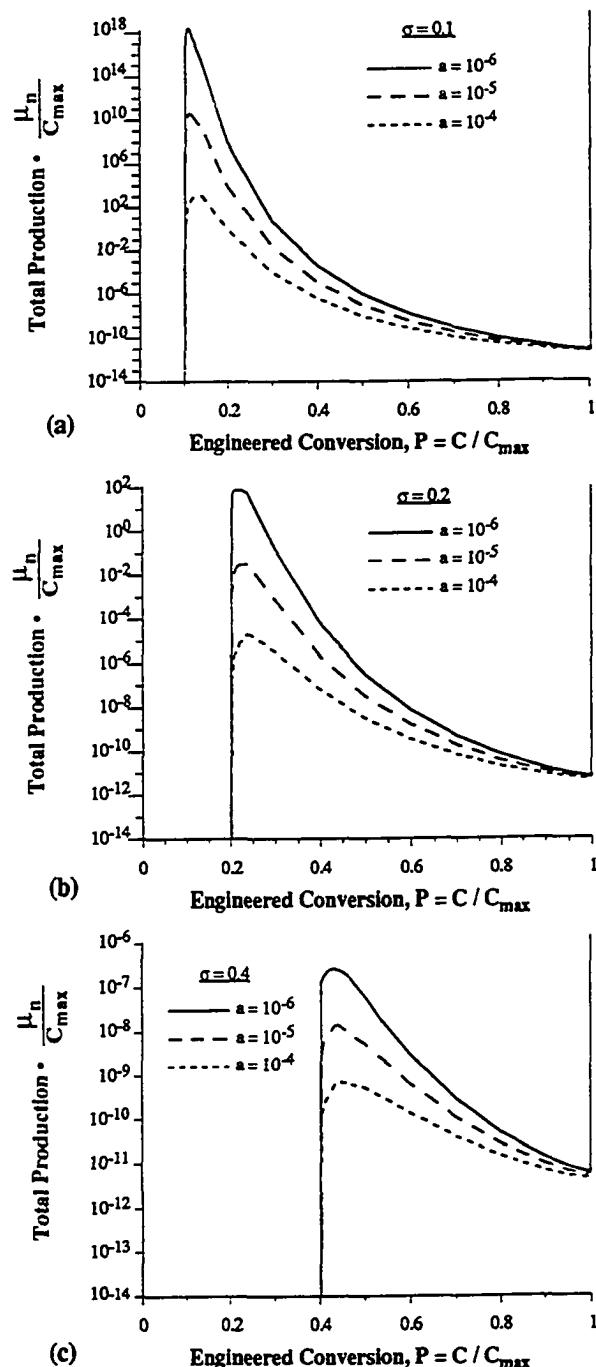


Figure 8.10: Total production of a biochemical product as a function of the dimensionless engineered productivity, under different mutational probabilities, and dimensionless economic criteria.

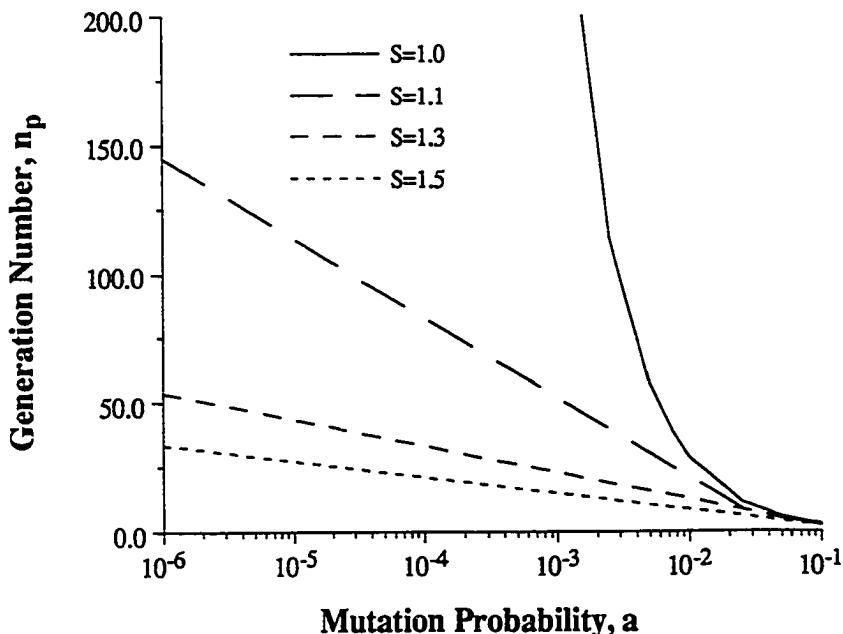


Figure 8.11: Stability of an engineered clone in the presence of mutations. The number of generations of the engineered producer strain, ' $n_p$ ', for the culture conversion to 25% mutants is plotted as a function of the mutation probability, 'a'. Culture stability as measured by the number of generations is extremely sensitive to the growth advantage of the mutant shown as  $S = (\mu_n / \mu_p)$ .

selective media can potentially suffer from outgrowth by low production mutant populations. In order to alleviate such problems of population stability it is useful to identify the conditions favoring stability of a culture.

Several studies [31, 40, 65] have utilized population balances to study the dynamics of different populations in a culture. Non-producing populations are assumed to arise from the original clone by random mutations. Stability of the producer cells is found to be critically dependent on the growth advantage of non-producers over the producer cells. Figure 8.11, derived from standard population balance models in literature [31, 40, 65] shows the extreme sensitivity of culture size to the growth

advantage of non-producers over the original clone. An increase in the number of generations ( $n_p$ ) results in an exponential increase in the culture size as well as productivity. Thus, there is a significant incentive to determine conditions favoring culture stability for large-scale industrial cultures.

On an individual single cell level instability arises from a loss or mutation of genetic material that leads to a loss in productivity. This process has been characterized as strain or genetic instability and may arise from the segregational loss of extrachromosomal material, mutational loss of genes, or reversion of the genetic changes. Segregational losses can be substantially reduced by the use of selective techniques or a suitable choice of vector. However, random mutations in *E. coli* genetic material are estimated to occur at the rate of  $10^{-6}$ - $10^{-7}$  per gene replication which can change several fold under special conditions [84]. Thus, host cell modifications of a random nature can result in the loss of productivity [20].

Loss of productivity in a single cell assumes importance if the cell proliferates at a higher rate than the original strain. A higher proliferative status of non-producing cells is likely due to the reduced productivity burden on cellular metabolism. Thus, the growth advantage of a non-producing mutant represented by the ratio of growth rates determines the culture stability. In the present study we have investigated the burden that productivity places on cellular metabolism and the conditions of oxygenation that may help reduce this stress. Such conditions, favoring stability of the genetically engineered population, can be useful to obtain a large culture size for biochemical production.

The flux balance model coupled to a maximal growth objective provides a useful method for estimating culture stability. We have previously demonstrated, Chapters 5, 7, that the maximal growth objective provides a suitable basis for simulation

of metabolic behavior in *E. coli* [86, 88]. As a model for oxygen consumption by a cell we permit consumption up to an arbitrarily specified supply rate. Thus, by suitably specifying the oxygen supply rate we are able to determine metabolic behavior over the entire range of oxygen supply rates. It should be noted that although we only consider a variation in the oxygen supply the key variable changed is the oxygen to glucose supply ratio.

The complete range of oxygenation has been investigated to determine the growth characteristics during biochemical production. We observe that depending on the product an appropriate choice of oxygen supply can lead to a reduced growth advantage of non-producer cells over the engineered high-producing strain. Since even a small reduction in the growth advantage can result in many more useful generations before the eventual outgrowth by non-producers a dramatic increase in the stability of the engineered strain can be obtained. Thus, the secretion of products from a cell can be stabilized through an appropriate selection of culture parameters.

For the illustrative examples of valine, tryptophan, and lysine production we demonstrate that the stability of an engineered strain can be markedly improved by an appropriate choice of oxygenation conditions. Based on economic criteria for production these improvements in culture stability can translate into large differences in the longevity of an economically viable culture. Thus, a culture stabilized by optimal oxygenation shows an improvement in the total production over several orders of magnitude.

Limitations of culture stability are an important consideration in the overall design of a bioprocesses. In general, we find that the stability of a biochemical producing culture is very sensitive to the probability of mutation as well as the engineered productivity. However, significant improvements in stability can be obtained by the

appropriate choice of oxygenation parameters. An investigation of culture parameters as a means for relieving production stress and thereby enhancing culture stability is therefore useful for the optimal design of a bioprocess.

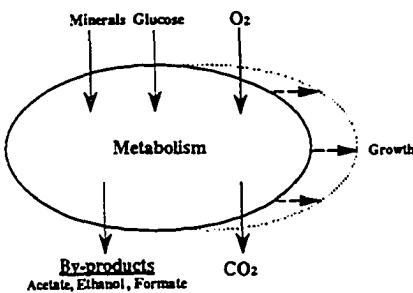
The present Chapter illustrates the important concept of controlling process parameters for improving the culture stability in a culture of genetically modified cells. Control of oxygen supply is an engineering technique that can be easily used to supplement other genetic techniques in order to enhance the stability of microbial cultures in bioprocesses. Thus, an integrated approach towards bioprocess design, using both strain development and process development to achieve the goals of high productivity and stability are expected to improve performance.

## CHAPTER IX

# EXPERIMENTAL OBSERVATIONS

Metabolism of the facultative organism *Escherichia coli* is known to result in the formation of several by-products under certain conditions. Acetate production in aerobic cultures has often been described as an overflow phenomenon [9, 83]. Under anaerobic conditions additional by-products such as formate, hydrogen, ethanol, etc. are also observable. Here we address the secretion of by-products by a nearly wild-type *E. coli* strain under various culture conditions and relate the observations to the flux balance model. The phenomenon is schematically represented in Fig. 9.1 with the carbon source (glucose) metabolized inside the bacterial cell to produce additional biomass. Oxygen is used by the cell as a terminal electron acceptor and several potential by-products can be secreted.

The fairly well known stoichiometry of the catabolic reactions has provided a useful basis to formulate quantitative models to predict metabolism in the bacterium. A balance between metabolic fluxes has been suggested to explain the phenomenon of acetate secretion in *E. coli* [46] and ethanol production in yeast [78]. We have formulated a comprehensive model of catabolism for *E. coli* that has proven useful to explain metabolism in the bacterium under aerobic and anaerobic conditions [86], Chapter 7. Here we show that experimental data on growth and by-product secretion



**Figure 9.1:** Schematic of the bacterial metabolism under consideration. The carbon source (glucose) and an electron acceptor (oxygen) are metabolized by the cell to produce biomass. Several by-products may also be produced in the process.

by a wild-type *E. coli* can be predicted using the flux balance model.

We have acquired data on *E. coli* growth on glucose under aerobic chemostat, batch, and fed-batch conditions as well as anaerobic batch conditions. We show that the flux balance model with an objective of maximizing growth is an appropriate representation of metabolism under these conditions, and furthermore show that with known initial conditions of the culture the flux balance model is able to predict the time profiles of growth and by-product concentration in the culture.

## 9.1 Materials and Methods

**Culture** An *E. coli* K-12 strain W3110 (ATCC# 27325) was used for all the experiments. The strain has been described as a nearly wild type and is able to grow on glucose mineral medium. Defined M9 medium ( $\text{Na}_2\text{HPO}_4$  6g/l,  $\text{KH}_2\text{PO}_4$  3g/l, NaCl 0.5g/l,  $\text{NH}_4\text{Cl}$  1g/l,  $\text{MgSO}_4$  2mM,  $\text{CaCl}_2$  0.1mM,  $\text{FeCl}_3$  0.01mM) was used for all the experiments with 2g/l glucose except for the fed-batch experiments where glucose was continuously added as described. A temperature of 38° was maintained for the culture in the bioreactor as well as the incubator.

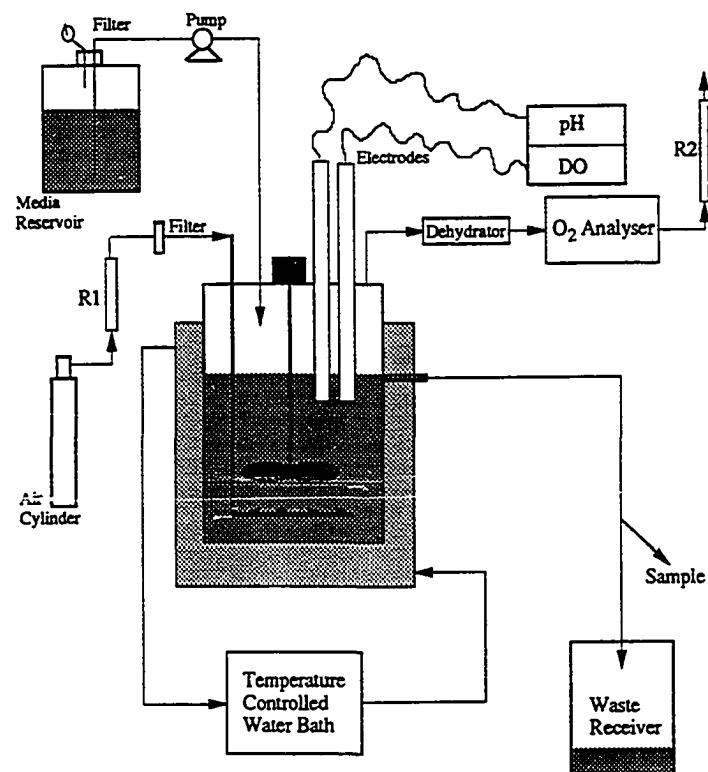


Figure 9.2: Schematic of bioreactor setup.

**Bioreactor Setup** Batch, fed-batch, and chemostat experiments were performed in a water jacketed Pegasus bioreactor (Pegasus, Ontario). The bioreactor setup is shown in Fig. 9.2. Temperature was controlled at 38° using a Haake water bath (Haake, Berlin, Germany). Medium inflow in chemostat experiments was achieved using a peristaltic pump and measured volumetrically at the outlet. Glucose solution was injected using a syringe pump in fed-batch experiments. Cell density was measured either with a coulter counter Model ZM (Coulter Electronics Inc, Hialeah, Florida) or by optical density at 600 nm using a Shimadzu UV 160 spectrophotometer (Shimadzu, Kyoto, Japan) and calibrated with the dry weight. A dry weight calibration of 0.32g DW/OD<sub>600</sub> was obtained.

Gas sparging was achieved using 3 Type D sintered glass spargers (10-20 $\mu$ m, Ace Glass Inc., Louisville). A small bubble size was obtained that helped keep dissolved oxygen above 50% in aerobic experiments. Gas flow rate was measured using Gilmont rotameters in the inlet and outlet. Oxygen percentage in the inlet and outlet gases was measured using fuel cell based oxygen analyzers, Systech Model 2000 (Illinois Instruments Company Inc.; McHenry, Illinois) and Micro-oxymax (Columbus Instruments International Corporation, Columbus, Ohio). Both instruments gave similar results.

**Analytical Assays** Glucose and lactate concentrations were measured using a YSI Glucose Lactate analyzer Model 2000 (Yellow Springs Instruments Inc., Yellow Springs, Ohio). Acetate was measured by enzymatic conversion to acetyl phosphate and reaction with hydroxylamine [42, 43, 64]. Ethanol, formate, and succinate were measured using available enzymatic kits (Boehringer Mannheim Corporation, Indianapolis).

**Flux Balance Model** In order to fulfill the objective of maximizing the growth rate we allow the model to secrete any amount of the by-products; acetate, ethanol, formate, lactate, and succinate in all simulations. The extent of by-product secretion is determined by the criterion of maximizing the growth rate. We then compare model predictions of growth rates and by-products to experimental data.

**Predictive Algorithm** For the prediction of the time profile of metabolism we have divided the experimental time into small steps of  $\Delta t$ . For the first time step the initial concentration values are specified for the model. Starting with the first time step the flux balance model is used to predict concentrations for the next step using the following algorithm:

1. Substrate concentration is given by the substrate concentration predicted from the last step plus any additional substrate provided in a fed-batch mode.

$$S_c = S_{co} + \frac{\text{supply} \cdot \Delta t}{\text{volume}} \quad (\text{mmol/l}) \quad (9.1)$$

2. The substrate concentration is appropriately scaled to define the substrate availability to the flux balance model.

$$\text{Substrate available} = \frac{S_c}{X \cdot \Delta t} \quad (\text{mmol/g DW - hr}) \quad (9.2)$$

3. The flux balance model is used to evaluate the actual substrate uptake ( $S_u$ ) (may be less than the substrate available) and the growth rate ( $\mu$ ).
4. Concentrations for the next time step are calculated by integrating the standard differential equations:

$$\frac{dX}{dt} = -\mu X \quad \Rightarrow \quad X = X_o \cdot e^{\mu \Delta t} \quad (9.3)$$

$$\frac{\partial S_c}{\partial t} = -S_u \cdot X \quad \Rightarrow \quad S_c = S_{co} + \frac{S_u}{\mu} X_o [1 - e^{\mu \Delta t}] \quad (9.4)$$

In the above algorithm we denote both glucose and by-products as substrates that can be used by cells. Thus, as a result of implementing the above algorithm we predict the concentration time profiles of cells, glucose, and by-products.

## 9.2 Results

The flux balance model is based solely on the stoichiometry of the metabolic reactions in the bacterium. Utilization of specific pathways is determined by the goal of optimizing the growth rate. The only parameter specifications to the model are the maintenance and biomass requirements as well as the enzymatic capacity of substrate uptake which are strain specific. Therefore, we first characterize the strain and then apply the predictive algorithm to several situations.

### 9.2.1 Strain Characterization

As parameters for the flux balance model we determine the maximum enzymatic limits of oxygen and glucose uptake from batch experiments for our particular strain. We also derive the biomass maintenance requirements which are strain and temperature specific.

**Oxygen Uptake Rate** The enzymatic capacity of the cell to consume oxygen was determined by batch experiments as shown in Fig. 9.3. The logic of the experimental approach is that during the exponential phase of a batch culture the glucose substrate

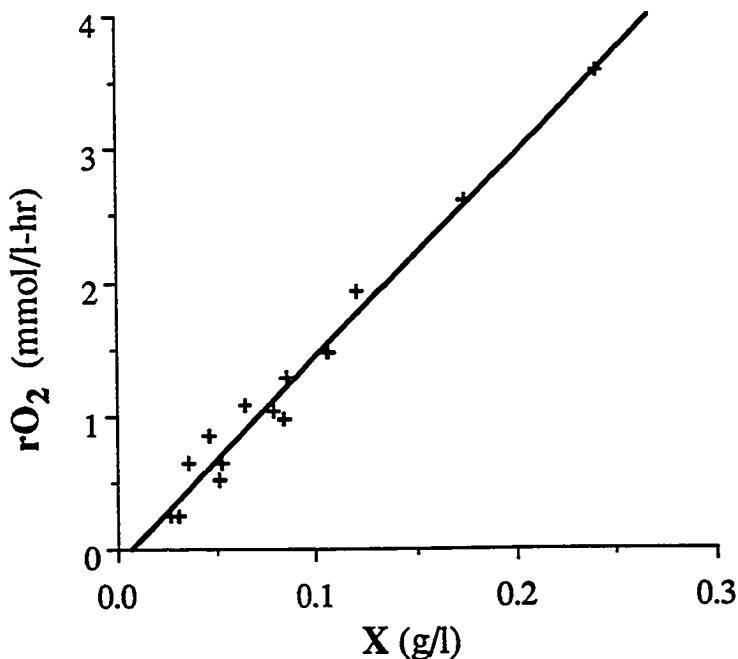


Figure 9.3: Maximum enzymatic capacity for oxygen uptake determined from batch experiments as the slope of the best fit, 15 mmol O<sub>2</sub>/g DW-hr. The data are a composite of several batch experiments where oxygen consumption was measured in the gas phase.

is in plentiful supply and growth rate is at a maximum. In such a situation the cell is expected to consume the maximum amount of oxygen possible [1].

From a modeling perspective the flux balance model also predicts a maximal oxygen consumption during exponential batch culture. Hence we obtain this model parameter from the batch experiments shown in Fig. 9.3 as 15 mmol O<sub>2</sub>/g DW-hr.

**Glucose Uptake Rate** The maximum glucose uptake rate is required for metabolic simulations of batch and fed-batch cultures in order to specify a finite upper limit on the metabolic capacity of a single cell. The enzymatic capacity for glucose uptake is determined from the biomass yield and growth rate in batch experiments. From aerobic batch cultures we have determined this value to be 10.5 mmol Glc/g DW-hr,

Fig. 9.4.

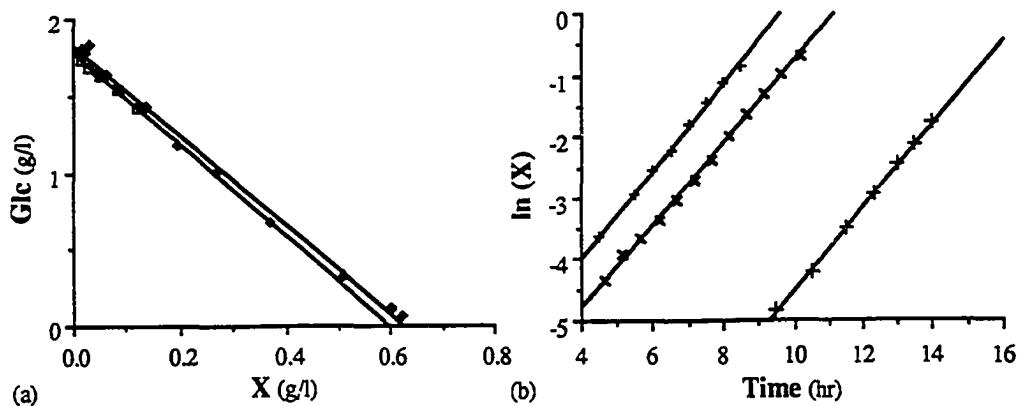


Figure 9.4: Maximum aerobic glucose uptake rate ( $10.5 \text{ mmol Glc/g DW-hr}$ ) determined from batch experiments as the growth rate ( $0.68 \text{ hr}^{-1}$ ) divided by the biomass yield ( $0.064 \text{ g DW/mmol Glc}$ ). The plots are composites of batch experiments.

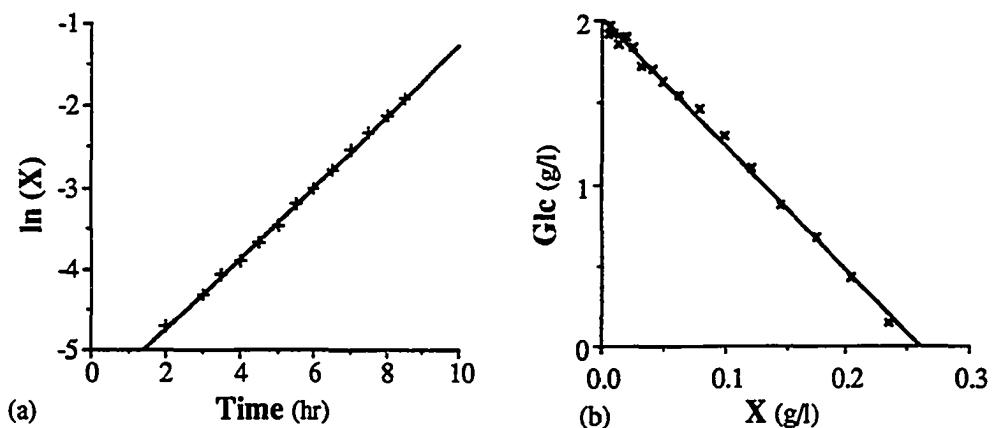


Figure 9.5: Maximum anaerobic glucose uptake rate ( $18.5 \text{ mmol Glc/g DW-hr}$ ) determined from batch experiments as the growth rate ( $0.43 \text{ hr}^{-1}$ ) divided by the biomass yield ( $0.023 \text{ g DW/mmol Glc}$ ).

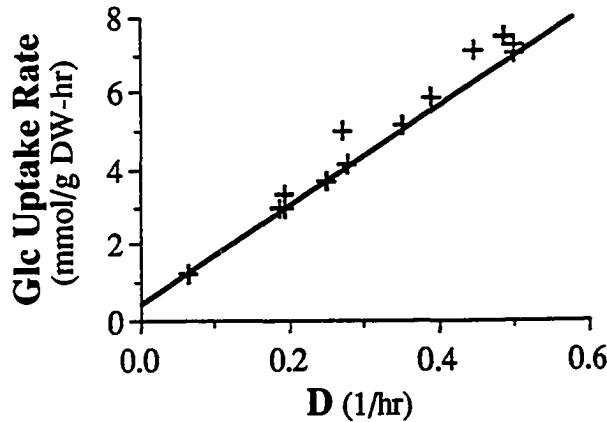


Figure 9.6: Determination of biomass and maintenance requirements from chemostat experiments. There were no mineral limitations in the chemostat. The solid line denotes the best fit of the flux balance model to the experimental datapoints shown. The non-growth associated maintenance requirements determined from the y-intercept are 7.6 mmol ATP/g DW-hr while the growth associated maintenance is defined as 13 mmol ATP/g DW. In addition biomass requirements are scaled to be 30% higher than the definition from literature [32] to account for strain specific differences.

It is a well recognized fact that the glucose uptake rate is higher under anaerobic conditions (the pasteur effect). We have determined the glucose uptake capacity under anaerobic conditions to be around 18.5 mmol/g DW-hr for our strain and experimental conditions, Fig. 9.5.

**Maintenance and Biomass Requirements** Synthesis of biomass in the model has been defined based on composition analysis in literature [32, 88]. In order to fully describe biomass generation in the model we have to include the maintenance requirements of the bacterium as well as account for strain specific differences in biomass composition.

Maintenance requirements are determined by the standard approach [76] of plotting glucose uptake as a function of the growth rate in a chemostat. As shown in

Fig. 9.6 a fit of the model to the plot gives us the non-growth associated maintenance [86] as the y-intercept which is equal to 7.6 mmol ATP/g DW-hr. Growth associated maintenance [86] is determined by an appropriate fit of the model slope to the experimental data of Fig. 9.6. For our system we have determined the growth associated maintenance to be 13 mmol ATP/g biomass. This maintenance adds on to the ATP energy requirements for biomass synthesis.

In addition to maintenance requirements the biomass requirements are also scaled in order to reflect strain specific differences. Thus, a scaling factor of 1.3 (ie., 30% higher biomass requirements) is used to obtain the appropriate fit of the flux balance model to data, Fig. 9.6. It should be pointed out that the biomass composition in terms of specific biosynthetic precursors may be different for our strain compared to the composition determined for a particular strain in literature [32]. However, it has been shown that the flux balance model is not highly sensitive to any particular biosynthetic precursor [88]. Thus, the non-specific scaling factor of 1.3 is expected to take into account any discrepancy in the biomass composition. In addition, the scaling factor also accounts for any differences among laboratories in measuring biomass.

From the above model specifications we have a total ATP requirement for producing biomass equal to 69 mmol ATP/g biomass, a value that compares favourably with 71 mmol ATP/g biomass determined for the closely related *Aerobacter aerogenes* [81].

### 9.2.2 Stoichiometrically Optimal By-Product Secretion

The flux balance model is based on the presumption of stoichiometrically optimal metabolism. In other words the model requires the optimal utilization of metabolic pathways in order to produce cell growth. We now consider the applicability of this

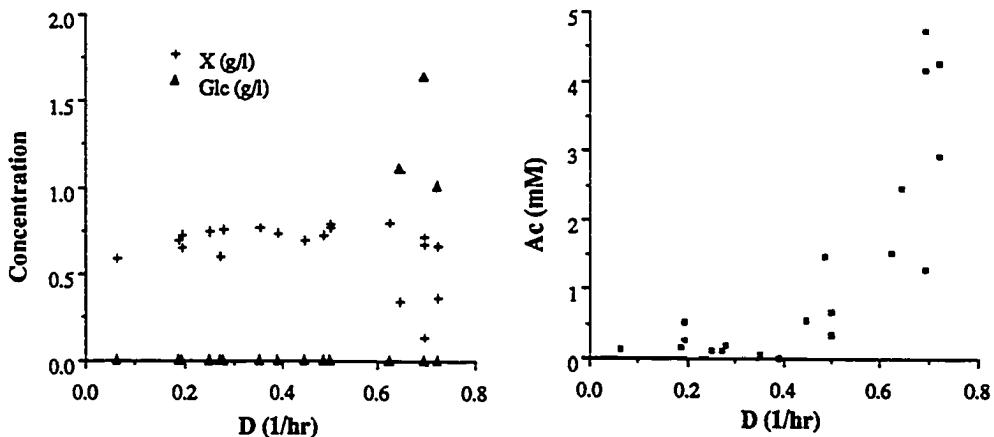


Figure 9.7: Aerobic chemostat culture showing cell density, glucose, and acetate concentrations as a function of the dilution or growth rate. The chemostat was not limited for minerals.

assumption to the aerobic culture of cells at different growth rates.

*E. coli* cells were grown in an aerobic chemostat at several different dilution rates. The chemostat was not limited for minerals. Cell density, glucose, and by-product concentrations were measured in the outlet from the culture vessel and are plotted in Fig. 9.7. At low dilution rates we observe that all the glucose is consumed without the secretion of any by-product. Cell density is also high at low dilution rates. At high dilution rates we observe that all the glucose is not consumed and there is a corresponding low cell density. Increasing dilution rates also result in the secretion of acetate as a by-product.

Fig. 9.8 plots the glucose uptake and acetate secretion rate at different dilution rates. Simulation in the flux balance model is carried out by specifying a glucose uptake rate and determining optimal growth and by-product secretion. We note in Fig. 9.8 that the flux balance model is able to simulate growth and acetate secretion.

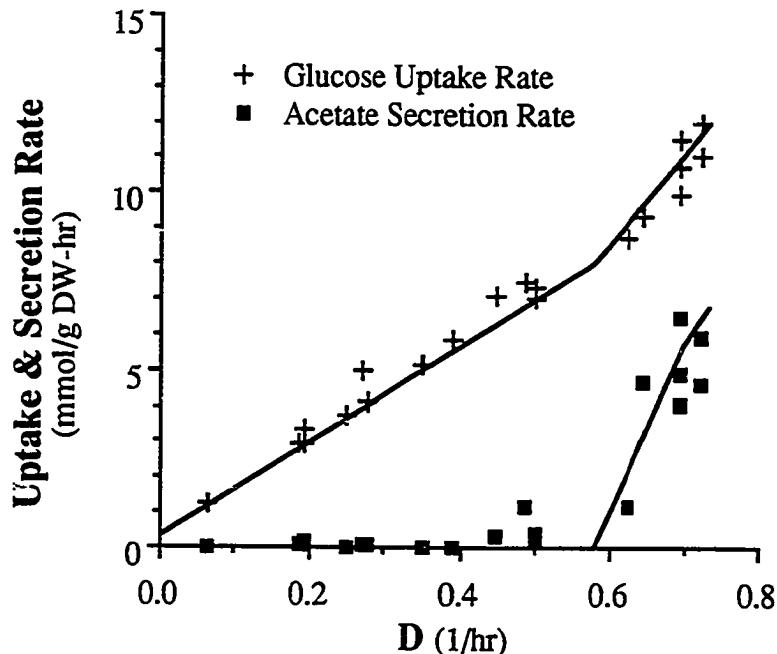


Figure 9.8: Aerobic chemostat culture showing the glucose uptake rate and acetate secretion rate as a function of the dilution or growth rate. The chemostat was not limited for minerals. The solid lines represent the flux balance model simulations.

It is interesting that the model is able to properly simulate both the onset of acetate secretion as well as the slope of the acetate secretion line. Another interesting point is that neither the model nor the experiments show the secretion of any other by-product. An excellent explanation for by-product secretion may be found in Chapter 7.

### 9.2.3 Flux Balance Model as a Predictive Tool

We now consider the applicability of the flux balance model as a predictive tool in batch and fed-batch cultures. It should be pointed out that the flux balance model is strictly applicable in a metabolic steady state. However, one may stretch the applicability to an unsteady state if we divide the time period into small segments where each segment does not deviate much from the steady state. While assessing

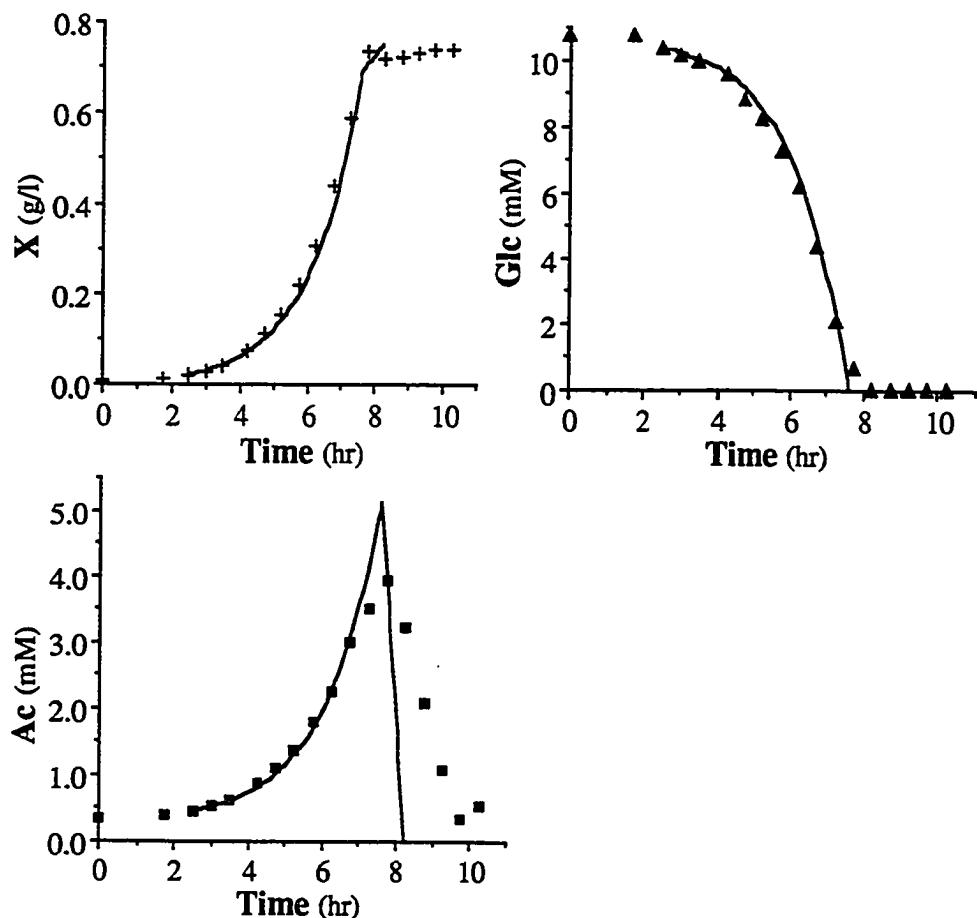


Figure 9.9: Aerobic batch culture showing time profiles of cell density, glucose concentration, and acetate concentration. The culture was not limited for minerals. The solid lines are the flux balance model predictions for the time profiles in the culture.

the results one must also keep in mind the reservation that the flux balance model cannot properly address the lag phase where the cellular metabolic machinery is not fully operational.

**Aerobic Batch Culture** Fig. 9.9 shows experimental data from a typical aerobic batch culture. Also shown are predictive simulations from the flux balance model. The predictions were carried out as discussed in the methods by specifying the initial conditions to the model.

As is evident from Fig. 9.9 the flux balance model is able to accurately predict time profiles for the cell density as well as the glucose uptake rate. It is also able to predict the acetate secretion and accumulation in the medium. We note that model predicts the reutilization of acetate, but is unable to account for the time delays in transition from glucose consumption to acetate consumption. Transitions between pathway utilization are a function of the regulatory process and not metabolism. Since the flux balance model includes only metabolism and not regulation the discrepancy in predicting the time delay is understandable.

Furthermore, an enzymatic limit on acetate uptake may have significance in limiting the acetate reconsumption rate. We have not determined the enzymatic limits on acetate uptake. Thus, the current specification of the flux balance model allows unlimited acetate consumption. However, there is a broader enzymatic limit of oxygen uptake that restricts acetate consumption to 11.3mmol Ac/g DW-hr.

**Aerobic Fed-Batch Culture** We now turn our attention to a predictive application of the flux balance model in fed-batch cultures. We initiate the fed-batch by inoculating the bacteria in defined mineral media without a carbon source. Glucose as the carbon source is continuously fed to the culture using a syringe pump. Once

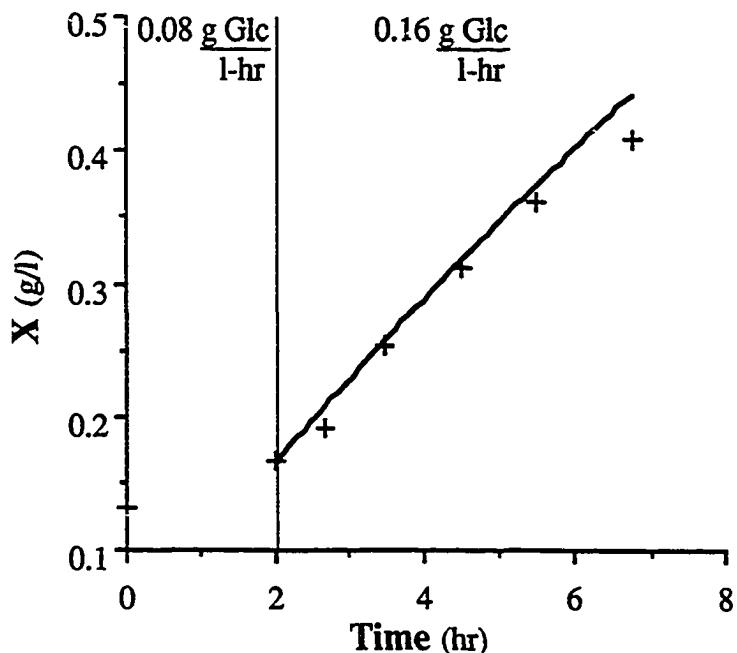


Figure 9.10: Aerobic fed-batch culture with continuous glucose injection at the rate noted in the figure. The culture was not limited for minerals. The time profile of cell density is shown with model predictions shown as the solid line. Glucose and acetate concentrations were zero for both experiments and model predictions.

again we ignore any lag phase for the simulations. We specify the initial concentrations and the glucose feed rate to the predictive algorithm and obtain the time profiles of cell density, glucose, and by-products.

We first consider the case of a culture inoculated with a relatively high cell density and a relatively low feed rate of glucose. Fig. 9.10 shows the experimental and predicted time profiles of such a culture. We note a high degree of correlation between the experimental and predicted time profiles of cell concentration. Also neither experiments nor model predictions show any accumulation of glucose and acetate.

Next we consider a fed-batch culture at a somewhat higher glucose feed rate relative to the cell density shown in Fig. 9.11. Once again we note the ability of the model to predict the time profile of cell density. We also note that glucose does not accumulate in the culture medium which is also predicted by the flux balance model.

We observe an interesting phenomenon of accumulation and reconsumption of acetate in the culture. One may explain the phenomenon in simple terms by recognizing that in the early part of the culture the cell density is low compared to the glucose supply rate. Thus growth rate is high and acetate is secreted similar to the observations for the chemostat cultures, Fig. 9.8. In the later part of the culture the cell density is high enough to consume all glucose supplied as well as to consume some of the acetate present in the medium. Thus, we observe a depletion of acetate in the medium. It is interesting to note that in the later part of the culture glucose and acetate are metabolized simultaneously.

Finally, we turn our attention to a fed-batch culture where the inoculum density is quite low. Fig. 9.12 shows experimental and predicted time profiles of cell density, glucose, and acetate concentration in the culture medium. Once again we observe an accurate prediction of the culture variables by the flux balance model.

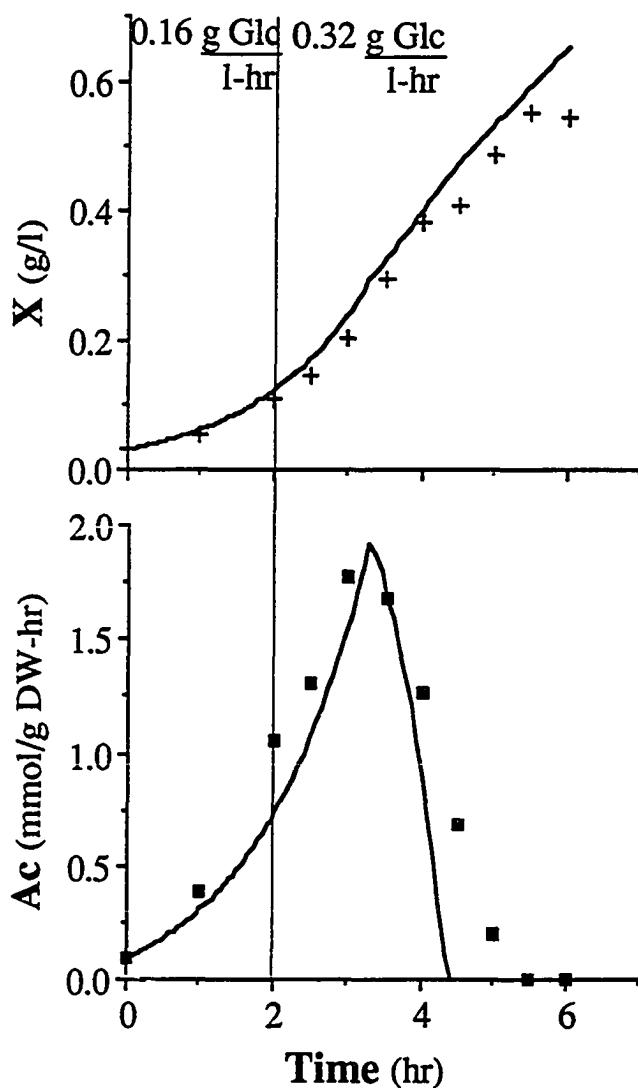


Figure 9.11: Aerobic fed-batch culture with continuous glucose injection at the rate noted in the figure. The culture was not limited for minerals. The time profile of cell density and acetate concentration is shown with model predictions shown as the solid line. Glucose concentration was zero for both experiments and model predictions.

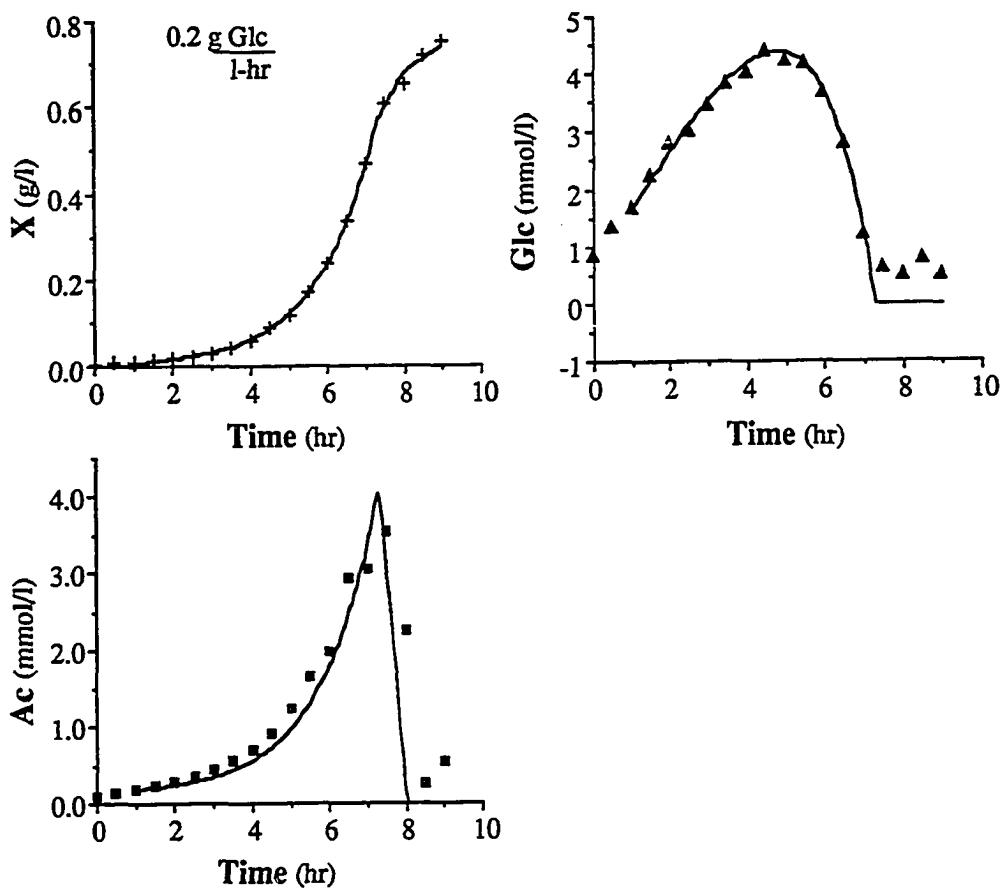


Figure 9.12: Aerobic fed-batch culture with continuous glucose injection at the rate noted in the figure. The culture was not limited for minerals. The time profile of cell density, glucose, and acetate concentrations is shown with model predictions shown as the solid line.

At low cell densities we also observe an accumulation of glucose in the culture medium, Fig. 9.12. Glucose accumulation is simply explained as a result of feeding more glucose than the cells can consume. Later in the culture as the cell density increases the cells are able to consume all the glucose supplied as well as deplete the acetate accumulated in the culture medium.

Similar to Fig. 9.11 we again observe the accumulation of acetate in the culture medium in Fig. 9.12. Once again we note that as cell density increases and glucose is depleted from the culture the cells are able to metabolize the acetate that was accumulated in the culture medium.

**Anaerobic Batch Culture** Since *E. coli* is a facultative organism we also consider the ability of the flux balance model to predict time profiles in anaerobic cultures. Experiments were run similar to the aerobic batch experiments with the exception that a nitrogen atmosphere was provided.

Fig. 9.13 shows the typical time profile of cell density and by-product concentrations. Once again the solid lines represent the model predictions for the culture. Three major by-products; acetate, ethanol and formate are found to be secreted and accumulate in the medium. We observe a high degree of correlation between the flux balance model and the anaerobic batch culture.

The by-product succinate was also experimentally found to be secreted at a marginal rate. Accumulation of succinate was quite small in the culture medium (less than 1mM at the end of culture). The flux balance model however does not predict any succinate secretion. Thus, there appears to be a small deviation from optimal metabolism under anaerobic conditions.

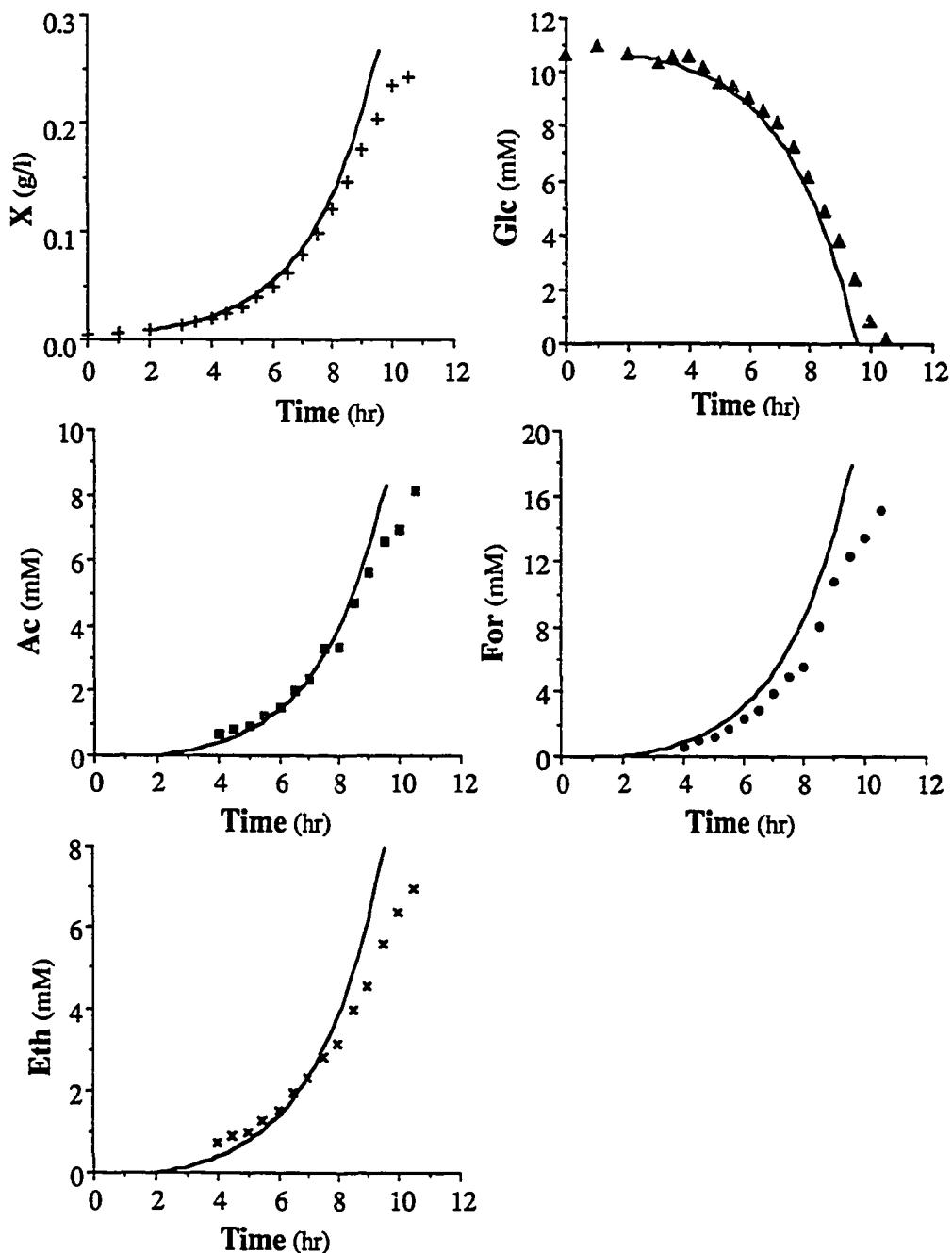


Figure 9.13: Anaerobic batch culture showing the time profiles of cell density and various by-product concentrations. Solid lines represent the model predictions.

### 9.3 Discussion

The flux balance model is a mechanistic description of metabolism in the cell. Based on the presumption of optimal metabolic performance it has been used to explain metabolic physiology of *E. coli* [46, 86]. Usefulness of the model arises from the large amount of information on metabolic pathway utilization provided by the model which makes it a useful quantitative tool for studying the metabolic engineering of pathways. It has been suggested to be a useful tool for bioprocess design and development [89].

Here we have experimentally addressed the applicability of the model. The two related issues of model verification and model prediction are addressed. The model is fully defined by the enzymatic capacity limits of oxygen and glucose uptake. In a larger sense specification of such limits prevents a single cell from displaying an infinite metabolic capacity, a physical impossibility.

Cells in the aerobic chemostat are shown to secrete acetate above a specific growth rate in a manner similar to that reported in literature [4]. We have shown here that the flux balance model can explain this secretion of acetate following a logic reported in literature [46].

We have provided a predictive algorithm for time profiles of metabolism in the unsteady state. The algorithm makes the assumption of a steady state in each small time step. The ability of the predictive algorithm to predict time profiles for cell density, glucose, and by-product concentrations in the culture medium provides a verification for the flux balance model as well as demonstrates its usefulness for bioprocess engineering.

Some deviations from model predictions are observed during transitions in the

utilization of metabolic pathways, such as the change from glucose metabolism to acetate metabolism. It is observed that the transition is less sudden for fed-batch cultures resulting in somewhat better model predictions for fed-batch transitions.

It should be recognized that the flux balance model is a model for metabolism and does not incorporate the time dynamics of regulation. Thus, the flux balance is able to accurately predict the occurrence of metabolic transitions but has trouble in circumstances where regulatory changes require time. Thus, for example, the flux balance model is inappropriate for simulating the lag phase in cultures. Regulatory transitions have been modeled in literature using a cybernetic approach [5, 38].

An interesting phenomenon observed in the aerobic fed-batch cultures is that of acetate reconsumption in the presence of a glucose feed. It is generally assumed that the presence of glucose represses the utilization of other substrates. In contrast we observe in both experiments and the model predictions that a sufficiently high cell density can result in the simultaneous consumption of glucose and acetate. Co-metabolization of glucose and acetate has also been previously reported [90].

An important question in the observation of co-metabolism is that of clone-specific metabolism. A glucose limited chemostat has been shown to result in the generation of polymorphisms, starting with a single clone of *E. coli* [27]. Clones isolated from the chemostat were found to have a different metabolism rate for glucose and acetate. Co-metabolism of glucose and acetate observed in the fed-batch experiment, Fig. 9.11 and Fig. 9.12 could either result from single cell metabolism or clone-specific metabolism. However, since the experiments only last a few hours one may reasonably assume that genetic mutants cannot be significant unless present in the inoculum. One can also conceptualize arguments for clone specific metabolism as a result of regulation with the same genetic clone. These questions form an interesting

area for future research.

With the above discourse on the usefulness and applicability of the flux balance model we would like to end with a brief mention of some mathematical and computational aspects. While the flux balance model may appear complicated, its structure and formulation is really rather simple. For large metabolic networks that include both catabolism as well as biosynthesis [85] the model can become computationally intensive. However, it is computationally solvable on current day microcomputers in real time. It is also interesting to note that microcomputers are expected to increase in computational speed by a factor of 10 by 1995 [13].

## CHAPTER X

### Stoichiometric Optimality: A Perspective

Ever since the elucidation of the chemical transformations that underlie intermediary metabolism in the 1930s and 1940s, the need has existed to analyse the holistic nature of metabolic physiology. This need has represented a formidable challenge since a full dynamic description of the metabolic network requires a detailed kinetic description of each enzyme. The extensive work on enzyme kinetics, in the 1960's and into the 1970's, made clear that such information was very difficult if not impossible to obtain for most cells. In addition, problems with the correlation between the *in vivo* as opposed to the *in vitro* conditions left nagging doubts about the physiological relevance of the data obtained. This apprehension has been further exacerbated by the discoveries, within the past decade or so, that most enzymes work in multi-enzyme complexes *in vivo*.

Much work exists on developing a reliable quantitative framework for the analysis of metabolic dynamics and control (with early attempts beginning in the late 1960's [66], intensifying in the 1970's [26, 35], and extending into the early 1980's [58, 63]). However, the many difficulties encountered in applying this early mathematical framework to a single cell has resulted in the inability to provide a systemic description of metabolism. In contrast, the relatively recent work on a stoichiomet-

ric interpretation of metabolism [18, 46, 73, 72, 91] lead us to recognize the potential of the newer flux balance methods of metabolic analysis.

The present thesis deals with the development and application of a flux balance model for *E. coli* metabolism. The model is based solely on the stoichiometry of the chemical reactions occurring within the bacterium. Study of this model is instructive in outlining the functional nature of metabolism.

The metabolic model for *E. coli* has a direct applicability to bioprocess design. The model is able to predict the optimal pathway utilization as well as the interaction of the cell with its environment. Genetic pathway modification can then be carried out to modify the metabolic fluxes as desired. The model also has potential use in the control of the bioprocess to maintain optimal environmental conditions.

### 10.1 Stoichiometric Optimality: A Scientific Principle

The living process is characterized by self-replication and metabolism. The present day complexity of life is presumed to have arisen by a sequential process of evolution. Several evolutionary events have been identified in the process of evolution of life on earth. Some of these steps are shown in Fig. 10.1. The earliest evolutionary events are thought to be the generation of complex organic compounds in the pre-biotic reducing environment of the earth. These organics are assumed to have provided the raw material for the generation of the first replicative machineries. Several lines of experimental evidence exist to suggest that the first self-replicative entity could have been either clay, protein, or nucleic acids.

Orderly chemical transformations lead to the development of metabolism. Many such schemes are likely to have developed, and through selection and adaptation, life-enabling metabolic networks were established. What characteristic of these early

networks leads to survival? At present the selection mechanisms are unknown.

To break down the problem of evolution of a metabolic pathway in present day cells, one can propose three distinct steps. At the primary level a suitable enzyme must be generated to convert the reactant into products. The enzymatic scheme must be coupled to redox and energy carrying cofactors. At the second level, a regulatory mechanism needs to be established to permit a suitable response of the pathway. Finally, a system of regulatory control has to be established to enhance survivability.

Fig. 10.1 depicts the three elements involved in a metabolic pathway; the stoichiometry, the regulatory mechanism, and the regulatory control. Much theory and experimental evidence is available [67, 68, 69, 70, 71] to support the "Demand Theory of Genetic Regulation" as an evolutionary principle guiding the choice of regulatory mechanisms for a particular metabolic function. One may also address the two related issues; stoichiometry of metabolic pathways, and functional nature of regulatory control.

Our work on the stoichiometric interpretation of microbial metabolism suggests an interesting proposition, namely; the function of regulatory control in a cell is to enable the cell to utilize its metabolic pathways in a stoichiometrically optimal fashion. Thus, the cell is able to maximize its survivability under the varying conditions of its habitats.

Consider the case of *Escherichia coli*. The life of the bacteria is divided between two remarkably different environments. The gut provides the bacterium with a nutritionally rich anaerobic environment while the soil is a nutritionally lean aerobic environment. One expects a different set of pathways to be operational under the different conditions. The flexibility of metabolic pathways to adapt to different

The Planetary Environment	Evolutionary Events	Theory and Experiments
Reduced Atmosphere in Chemical Equilibrium	Formation of Organic Chemicals	S. Miller (Spark Apparatus)
High Energy Solar Radiation	Simple Replicative Machinery • Clays • Proteins • Nucleic Acids	A.G. Cairns-Smith A.I. Oparin M. Eigen, L. Orgel
Changes in Temperature	Single Cell Primitive Organisms	
Oxidized Atmosphere in Chemical Disequilibrium	Metabolism as a Force of Change	
Reduced High Energy Solar Radiation	Evolution of Metabolic Pathways • Reaction Stoichiometry  • Regulatory Mechanisms  • Regulatory Control	<b><i>Development of Metabolic Pathways</i></b>  <b><i>Demand Theory of Genetic Regulation</i></b> M. Savageau
	Cellular Organelles • Chloroplasts • Mitochondria • Nucleus	<b><i>Stoichiometric Optimality of Cellular Metabolism</i></b>  <b><i>Driving force for Organelle Development</i></b>  <b>Synergism as a driving force</b> L. Margulis
	Evolution of Multicellular Organisms	

Figure 10.1: Some evolutionary events leading to the formation of life.

environments is what we determine as stoichiometric optimality. Such a flexibility also provides a microbe with a survival advantage.

A contrasting example may be developed in the form of a specialized microbe living in a hot thermal oceanic vent. Such an organism may not have the necessary flexibility of pathways to adapt to any other environment. The question arises: how do we square the existence of such an organism with a stoichiometrically optimal organism living in the same environment? The answer we propose may lie in the Burden of Regulation. While a flexibility of pathways is an advantage in a changing environment it does carry with it the necessary regulatory machinery which is an additional burden on metabolism. In contrast, a specialized microbe without such regulatory machinery would grow very well in a static environment. Indeed there may even be an incentive to loose unnecessary pathways and regulatory mechanisms in a static environment.

Thus, survival through the optimal use of metabolism logically relates to the structure of the metabolic transformations (the stoichiometry) and its optimal use. Such optimal use is likely to have been important in the evolution of life and the survival of primitive cells. As biological behavior increases in complexity, into multi-cellular aggregates and so forth, the importance of optimal metabolism may diminish as a critical determinant in survival. However, a large fraction of living cells today are primitive cells and the principle of stoichiometric optimality is still operative in many environments. The physiology of prokaryotic metabolism, particularly in nutritionally rich environments, is probably still governed by stoichiometric optimality. Stoichiometric optimality thus constitutes an important scientific principle that has basic scientific value, is of industrial importance for strain development and bioprocess optimization, may have medical importance with respect to microbial pathology,

and may have expanded medical importance in the future as the era of gene therapy arrives.

## 10.2 Applicability of Stoichiometric Optimality

**Metabolism: The decision Maker** It is well known that metabolism has the ability to make “decisions”, that is change its flux distribution in response to external and internal stimuli. Elaborate kinetic schemes have been synthesized to explain metabolic switches and describe them as “creative functions” using bifurcation theory applied to simple kinetic models [8]. Such explanations try to correlate kinetic behavior, which, as stated above is not that well known, to physiological function, normally assumed to be optimal in some way or another. On the other hand, stoichiometric optimality allows us to approach this issue in a much more straightforward manner. Further, this approach is only based on relatively unambiguous information and the basis for the decisions is clear.

Using the flux balance analysis we have identified several shifts in pathway utilization that occur in response to changes in the environment. Thus, for example, moving from an aerobic to an anaerobic environment we note that several catabolic pathways shut down while other dormant pathways are stimulated. It is of interest to know the trigger for these physiological decisions to change pathway utilization.

The mathematical structure of the flux balance analysis allows us to formulate the mathematical dual of the linear optimization problem. The solution to the dual optimization provides the shadow prices which allow an interpretation of the physiological decisions mentioned above. Shadow prices in our formulation of the flux balance model represent the value of specific metabolites to accelerate growth. Thus we interpret the physiological decision of by-product secretion to occur when

the specific by-product has no value (shadow price is zero) to the cell.

**Bioprocesses and Engineering** The mechanistic flux balance model presented here provides a useful tool for bioprocess design. By providing predictions of metabolism the model allows us to study the interaction of the cell with its environment. Thus, one can undertake integrated bioprocess design of both the cellular and processing components.

Availability of a predictive metabolic model has direct applicability as a component of a process model that can be used for process control. Metabolic models are also useful for rational medium design for cell culture as well as optimizing utilization of multiple substrates. Industrial applications arise in the areas of biochemical production and waste degradation.

**Strain Development** The engineering of micro-organisms to produce therapeutic proteins, biochemicals, etc., requires the addition of new metabolic pathways. The introduction of new pathways causes a deviation from the normal metabolic state of the organism. One may use the tools of stoichiometric and thermodynamic analysis to investigate the effect of the changed metabolic pools on the feasibility of metabolic pathways.

We have already illustrated the use of the model for enhancing population stability of an engineered strain. One may also consider altering metabolism within the cell in order to influence the stability of an engineered strain. Selection of strains showing a particular phenotype is frequently carried out in the presence of selective media. A systemic analysis of metabolism is able to identify further environmental conditions favorable to the selection of a particular phenotype.

**Metabolic Imbalance of Disease** It has long been known that many infections and diseases derive from imbalances in metabolism. For example, storage diseases, such as glycogenosis, are characterized by an intracellular accumulation of unmetabolized substrate. These diseases result from specific enzymic deficiencies that are targets for gene therapy [39].

Recent developments in the area of gene therapy place in our hands the capability to introduce genetic change as a therapeutic measure. Gene therapy can be used to correct metabolic deficiencies by providing the necessary enzymes. One can design strategies for redressing the metabolic imbalance of a diseased state through engineered metabolic change.

While gene therapy provides us with a method to introduce specific enzyme catalyzed pathways the systemic effect of such pathways is not known. A systemic analysis of the metabolic network may prove an aid to the analysis of metabolic imbalances.

Our key understanding is that metabolic imbalance occurs due to a change in the normal metabolic pathways. Abnormal metabolism causes changes in the size of metabolite pools. Using a systemic approach one may investigate the effects of changes in metabolite pools on reaction pathways as well as the effect of reaction pathways on the metabolic state of the cell.

### 10.3 Future of Metabolic Analysis

The principles of flux balance analysis developed in this work provide a material balance dimension to the metabolic network. Another dimension that may be explored is that of thermodynamic changes during the metabolic transformations. Thus one may account for both the material and energy flows in the metabolic net-

work. At this point a further development and inclusion of the regulatory framework would provide a sense of completeness to understanding metabolism in a living cell.

**Introduction of Thermodynamics** Metabolism as a network of reactions is governed by several fundamental physical laws. Flux balance analysis described in this thesis essentially implements the law of conservation of mass. As a further development to the flux balance analysis one may incorporate thermodynamic information in the model.

While the steady state flux balance model is a very useful tool for determining the optimal metabolic flux distribution in a cell, it is incapable of determining whether the reaction pathways are thermodynamically feasible. The conceptualization of a cell as a tiny reactor carries the implication that metabolite concentrations must determine the direction of individual reactions. While the action of suitably regulated enzymes can modulate the reaction rates, enzymes cannot by themselves dictate the direction of reactions.

Several lines of evidence suggest that thermodynamics plays an important role in determining the structure of cellular metabolism. Cells have evolved several innovative methods to bypass the limitations of thermodynamics [56]. Thus, channelling and multienzyme complexes provide a means by which several reactions may occur with a particular reactant before it is able to diffuse away into the cytoplasmic soup. One may conceptualize channelling as increasing the local concentration of intermediates of the reaction sequence. Alternate mechanisms of multicompartmentalization such as in chloroplasts and mitochondria have also evolved as a means to improve overall efficiency. Thus, the P/O ratio in mitochondria is significantly higher than in simple prokaryotes.

We now show a mathematical formalism that describes the feasible metabolite concentrations for a given set of active reaction pathways. The analysis extends the current linear programming/flux balance approach to the interactions between metabolite pools. The formalism is based on the requirement that the free energy change in an active reaction must be negative. Mathematically one may write:

$$\Delta G = \Delta G^\circ + RT \ln \Pi C_i^{S_{ij}} < 0 \quad (10.1)$$

$$\Delta G = \Delta G^\circ + RT \sum S_{ij} \cdot \ln C_i < 0 \quad (10.2)$$

where  $\Delta G$  represents the free energy change of the reaction,  $\Delta G^\circ$  represents the standard free energy change,  $C_i$  is the concentration of the  $i^{th}$  metabolite and  $S_{ij}$  is the stoichiometric coefficient of the  $i^{th}$  metabolite in the  $j^{th}$  reaction.

Eqn. (10.1) may be written for all the reactions that are active in the cell under a particular set of conditions. These set of equations represent constraints on the feasible metabolite concentrations in the cytoplasm. Of course, the concentrations must also be constrained to lie within the physical limits of the cell.

The thermodynamic information of  $\Delta G^\circ$  required to solve Eqn. (10.1) is obtainable either as the standard free energy of formation for the metabolites, or from the equilibrium constants for the metabolic reactions. These values are easily available from standard literature sources.

Similar to the flux balance analysis the log metabolite concentrations are now restricted to a linear feasible domain determined by the requirements of Eqn. (10.2). Thus, the mathematical structure of Eqn. (10.2) is accessible to the powerful techniques of linear programming and determines the thermodynamic limits of metabolism. One may directly examine the range of feasible metabolite concentrations for a

particular metabolic flux distribution. Mathematical duality of the problem might also prove useful in determining the thermodynamic relationship between metabolic pools.

**Identification of Regulators** The metabolic network carrying stoichiometric as well as thermodynamic information can potentially be used to identify regulatory mechanisms as well as candidate regulators. Correlations between pathway utilization and metabolite concentrations can be used to narrow the range of regulatory molecules. Model directed experimental verification can then be carried out to establish cause and effect for the proposed regulatory mechanism.

**The Driving Force for Evolution** Improvements in metabolism that increase efficiency should result in the evolution of newer cells and sub-cellular organelles. One may investigate the limits of metabolism due to thermodynamic restrictions. We expect that it is these limits that require the development of compartmentalization, channeling, and multienzyme complexes as evolutionary steps. It is possible to determine the limits of metabolic efficiency in these structures and compare it to that in the simple cell.

An interesting issue is the comparison of metabolic regulation to thermodynamic control. The hypothesis is that while regulation controls the level of reaction fluxes by controlling enzyme expression and activation, thermodynamics can also determine whether reactions may proceed. It is useful to know whether pathways are controlled by thermodynamics with a redundant regulatory mechanism. The significance of this determination is that such pathways will show a rigidity to change. Thus, genetic engineering of such pathways would acquire an added dimension.

#### 10.4 Brave New World

History is often defined by the significant scientific and technological advances that have taken place. Thus, the industrial revolution was defined by the use of coal for power. The petrochemical industry brought about the next advance making the world appear much smaller. The taming of the atom in the 1940's has influenced much of mankinds direction in the following decades. In a similar vein, the taming of the genome proposes to usher a brave new era of biotechnology that will potentially influence the next few decades.

## **BIBLIOGRAPHY**

## BIBLIOGRAPHY

- [1] K. B. Anderson and K. von Meyenburg. Are growth rates of *Escherichia coli* in batch cultures limited by respiration. *J. Bacteriol.*, 144:114–23, 1980.
- [2] J. E. Bailey. Toward a science of metabolic engineering. *Science*, 252:1668–75, 1991.
- [3] J. E. Bailey, S. Birnbaum, J. L. Galazzo, C. Khosla, and J. V. Shanks. Strategies and challenges in metabolic engineering. *Ann. N. Y. Acad. Sci.*, 589:1–15, 1990.
- [4] R. Bajpai. Control of bacterial fermentations. *Ann. N.Y. Acad. Sci.*, 506:446–58, 1987.
- [5] S. Baloo and D. Ramkrishna. Metabolic regulation in bacterial continuous cultures: I. *Biotech. and Bioeng.*, 38:1337–52, 1991.
- [6] A. Belaich and J. P. Belaich. Microcalorimetric study of the anaerobic growth of *Escherichia coli*: Growth thermograms in a synthetic medium. *J. Bacteriol.*, 125:14–18, 1976.
- [7] A. C. Blackwood, A. C. Neish, and G. A. Ledingham. Dissimilation of glucose at controlled pH values by pigmented and non-pigmented strains of *Escherichia coli*. *J. Bacteriol.*, 72:497–9, 1956.
- [8] A. Boiteux, B. Hess, and E. E. Sel'kov. Creative functions of instability and oscillations in metabolic systems. *Curr. Top. Cell Regul.*, 17:171–203, 1980.
- [9] R. J. Brittan. Extracellular Metabolic Products of *Escherichia coli* during rapid growth. *Science*, 119:578, 1954.
- [10] T. D. K. Brown, M. C. Jones-Mortimer, and H. L. Kornberg. The enzymatic interconversion of acetate and acetyl-coenzyme A in *Escherichia coli*. *J. Gen Microbiol.*, 102:327–36, 1977.
- [11] Control of metabolic processes. In Athel Cornish-Bowden and Maria Luz Cardenas, editors, *NATO ASI Series A: Lifesciences Vol. 190*. Plenum Press, New York, 1990.
- [12] A. P. Damoglou and E. A. Dawes. Studies on the lipid content and phosphate requirement of glucose- and acetate- grown *Escherichia coli*. *Biochem. J.*, 110:775–81, 1968.

- [13] P. J. Denning. An end and a begining. *American Scientist*, 81:416–8, 1993.
- [14] H. W. Doelle and N. W. Hollywood. Bioenergetic aspects of aerobic glucose metabolism of *Escherichia coli* K-12 under varying specific growth rates and glucose concentrations. *Microbios*, 21:47–60, 1978.
- [15] M. M. Domach, S. K. Leung, G. G. Cahn, and M. L. Shuler. Computer model for glucose-limited growth of a single cell of *Escherichia coli* Br-A. *Biotech. Bioeng.*, 26:203–16, 1984.
- [16] K. M. Draths and J. W. Frost. Conversion of D-glucose into catechol: The not-so-common pathway of aromatic biosynthesis. *J. Am. Chem. Soc.*, 113:9361–3, 1991.
- [17] I. S. Farmer and C. W. Jones. The effect of temperature on the molar growth yield and maintenance requirement of *Escherichia coli* w during aerobic growth in continuous culture. *FEBS Lett.*, 67:359–63, 1976.
- [18] D. A. Fell and J. A. Small. Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem. J.*, 238:781–6, 1986.
- [19] J. Fieschko and T. Ritch. Production of human alpha consensus interferon in recombinant *Escherichia coli*. *Chem. Eng. Commun.*, 45:229–40, 1986.
- [20] J. Fu, A. P. Togna, M. L. Shuler, and D. B. Wilson. *Escherichia coli* host cell modifications in continuous culture affecting heterologous protein overproduction: A population dynamics study. *Biotechnol. Prog.*, 8:340–6, 1992.
- [21] D. Garfinkel, L. Garfinkel, M. Pring, S. B. Green, and B. Chance. Computer applications to biochemical kinetics. *Ann. Rev. Biochem.*, 39:473–98, 1970.
- [22] M. Gyr. Linear optimization using the simplex algorithm (simple). In *CERN computer center program library*. CERN, Geneva, Switzerland, 1978.
- [23] J. J. Heijnen, J. A. Roels, and A. H. Stouthamer. Application of balancing methods in modeling the penicillin fermentation. *Biotechnol. Bioeng.*, 21:2175–201, 1979.
- [24] R. Heinrich, S. M. Rapoport, and T. A. Rapoport. Metabolic regulation and mathematical models. *Prog. Biophys. Mol. Biol.*, 32:1–82, 1977.
- [25] R. Heinrich and T. A. Rapoport. A linear steady state treatment of enzymatic chains. Critique of the crossover theorem and a general procedure to identify interaction sites with an effector. *Eur. J. Biochem.*, 42:97–105, 1974.
- [26] R. Heinrich and T. A. Rapoport. A linear steady state treatment of enzymatic chains. General properties, control and effector strength. *Eur. J. Biochem.*, 42:89–95, 1974.

- [27] R. B. Helling, C. N. Vargas, and J. Adams. Evolution of *Escherichia coli* during growth in a constant environment. *Genetics*, 116:349–58, 1987.
- [28] K. M. Herrmann and R. L. Somerville. *Amino acids: Biosynthesis and genetic regulation*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1983.
- [29] N. Hollywood and H. W. Doelle. Effect of specific growth rate and glucose concentration on growth and glucose metabolism of *Escherichia coli* K-12. *Microbios*, 17:23–33, 1976.
- [30] W. H. Holms. The central metabolic pathways of *Escherichia coli*: Relationship between flux and control at a branch point, efficiency of conversion to biomass, and excretion of acetate. *Curr. Top. Cell. Regul.*, 28:69–105, 1986.
- [31] T. Imanaka and S. Aiba. A perspective on the application of genetic engineering: Stability of recombinant plasmid. *Ann. NY Acad. Sci.*, 369:1–14, 1981.
- [32] J. L. Ingraham, O. Maaløe, and F. C. Neidhardt. *Growth of the bacterial cell*. Sinauer associates Inc., Sunderland, Massachusetts, 1983.
- [33] M. Inouye. *Bacterial Outer Membranes*. John Wiley & Sons, New York, 1979.
- [34] A. Joshi and B. O. Palsson. Metabolic dynamics in the human red cell. Part I. A comprehensive model. *J. Theor. Biol.*, 141:515–28, 1989.
- [35] H. Kacser and J. A. Burns. The control of flux. *Symp. Soc. Exp. Biol.*, 27:65–104, 1973.
- [36] E. R. Kashket. Stoichiometry of the H<sup>+</sup>-ATPase of growing and resting, aerobic *Escherichia Coli*. *Biochemistry*, 21:5534–8, 1982.
- [37] E. R. Kashket. Stoichiometry of the H<sup>+</sup>-ATPase of *Escherichia Coli* cells during anaerobic growth. *FEBS Letters*, 154:343–6, 1983.
- [38] D. S. Kompala, D. Ramakrishna, N. B. Jansen, and G. T. Tsao. Investigation of bacterial growth on mixed substrates: Experimental evaluation of cybernetic models. *Biotech. Bioeng.*, 28:1044–55, 1986.
- [39] J. W. Lerrick and K. L. Burck. *Gene therapy: Application of molecular biology*. Elsevier, New York, 1991.
- [40] G. M. Lee, A. Varma, and B. O. Palsson. Application of population balance model to the loss of hybridoma antibody productivity. *Biotechnol. Prog.*, 7:72–5, 1991.
- [41] I-Der Lee and B.O. Palsson. A comprehensive model of human erythrocyte metabolism: Extensions to include pH effects. *Biomed. Biochim. Acta*, 49:771–789, 1990.

- [42] F. Lipmann and L. C. Tuttle. Acetyl phosphate: Chemistry, determination, and synthesis. *J. Biol. Chem.*, 153:571–82, 1944.
- [43] F. Lipmann and L. C. Tuttle. A specific micromethod for the determination of acyl phosphates. *J. Biol. Chem.*, 159:21–8, 1945.
- [44] D. G. Luenberger. *Linear and nonlinear programming*. Addison-Wesley Publishing, Reading, Massachusetts, 1984.
- [45] S. E. Mainzer and W. P. Hempfling. Effects of growth temperature on yield and maintenance during glucose-limited continuous culture of *Escherichia coli*. *J. Bacteriol.*, 126:251–6, 1976.
- [46] R. A. Majewski and M. M. Domach. Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotech. and Bioeng.*, 35:732–8, 1990.
- [47] P. C. Maloney. Coupling to an energized membrane: Role of ion-motive gradients in the transduction of metabolic energy. In F. C. Neidhardt, editor, *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology*, pages 222–43. American society for microbiology, 1987.
- [48] J. Mandelstam and K. McQuillen. *Biochemistry of bacterial growth*. John Wiley & Sons, New York, 1973.
- [49] A. G. Marr, E. H. Nilson, and D. J. Clark. The maintenance requirement of *Escherichia coli*. *Ann. N. Y. Acad. Sci.*, 102:536–48, 1963.
- [50] K. G. Murty. *Linear programming*. John Wiley & sons, New York, 1983.
- [51] In F. C. Neidhardt, editor, *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology*. American society for microbiology, 1987.
- [52] F. C. Neidhardt. Chemical composition of *Escherichia coli*. In F. C. Neidhardt, editor, *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology*, pages 3–6. American society for microbiology, 1987.
- [53] F. C. Neidhardt, J. L. Ingraham, and M. Schaechter. *Physiology of the bacterial cell. A molecular approach*. Sinauer Associates, Sunderland, Massachusetts, 1990.
- [54] O. M. Neijssel and D. W. Tempest. Bio-energetic aspects of aerobic growth of *Klebsiella aerogenes* NCTC 418 in carbon-limited and carbon sufficient chemostat culture. *Arch. Microbiol.*, 107:215–21, 1976.
- [55] H. J. Noorman, J. J. Heijnen, and K. Ch. A. M. Luyben. Linear relations in microbial reaction systems: A general overview of their origin, form, and use. *Biotech. and Bioeng.*, 38:603–18, 1991.
- [56] J. Ovadi. Physiological significance of metabolite channelling. *J. Theor. Biol.*, 152:1–22, 1991.

- [57] B. O. Palsson and I. D. Lee. Model complexity has a significant effect on the numerical value and interpretation of metabolic sensitivity coefficients. *J. Theor. Biol.*, accepted, 1992.
- [58] B. O. Palsson and E. N. Lightfoot. Mathematical modelling of dynamics and control in metabolic networks. I. On michaelis-menten kinetics. *J. Theor. Biol.*, 111:273–302, 1984.
- [59] S. J. Pirt. The mamintenance energy of bacteria in growing cultures. *Proc. Roy. Soc. B*, 163:224–31, 1965.
- [60] S. J. Pirt. Maintanence energy: A general model for energy-limited and energy sufficient growth. *Arch. Microbiol.*, 133:300–2, 1982.
- [61] Y. Poirer, D. E. Dennis, K. Klomparens, and C. Somerville. Polyhydroxybutyrate, a biodegradable thermoplastic, produced in transgenic plants. *Science*, 256:520–3, 1992.
- [62] R. K. Poole and B. A. Haddock. Energy-linked reduction of nicotinamide-adenine dinucleotide in membranes derived from normal and various respiratory-deficient mutant strains of *Escherichia coli* K12. *Biochem. J.*, 144:77–85, 1974.
- [63] J. G. Reich and E. E. Sel'kov. *Energy metabolism of the cell*. Academic Press, New York, 1981.
- [64] I. A. Rose, M. Grunberg-Manago, S. R. Korey, and S. Ochoa. Enzymatic phosphorylation of acetate. *J. Biol. Chem.*, 211:737–56, 1954.
- [65] S. I. Rubinow. *Introduction to Mathematical Biology*. John Wiley & Sons, New York, 1975.
- [66] M. A. Savageau. Biochemical systems analysis. II. The steady state solutions for an n-pool system using a power-law approximation. *J. Theor. Biol.*, 25:370–9, 1969.
- [67] M. A. Savageau. Genetic regulatory mechanisms and the ecological niche of *escherichia coli*. *Proc. Natl. Acad. Sci.*, 71:2453–2455, 1974.
- [68] M. A. Savageau. Design of molecular control mechanisms and the demand for gene expression. *Proc. Natl. Acad. Sci.*, 74:5647–51, 1977.
- [69] M. A. Savageau. Regulation of differentiated cell-specific functions. *Proc. Natl. Acad. Sci.*, 80:1411–1415, 1983.
- [70] M. A. Savageau. Mathematics of organizationally complex systems. *Biomed. Biochim. Acta*, 44:839–44, 1985.
- [71] M. A. Savageau. A theory of alternate designs for biochemical control systems. *Biomed. Biochim. Acta*, 44:875–880, 1985.

- [72] J. M. Savinell and B. O. Palsson. Network analysis of intermediary metabolism using linear optimization: II. Interpretation of hybridoma cell metabolism. *J. Theor. Biol.*, 154:455–73, 1992.
- [73] J. M. Savinell and B. O. Palsson. Network analysis of intermediary metabolism using linear optimization: I. Development of mathematical formalism. *J. Theor. Biol.*, 154:421–54, 1992.
- [74] J. M. Savinell and B. O. Palsson. Optimal selection of metabolic fluxes for *in vivo* experimental determination: I. Development of mathematical methods. *J. Theor. Biol.*, 155:201–14, 1992.
- [75] J. M. Savinell and B. O. Palsson. Optimal selection of metabolic fluxes for *in vivo* experimental determination: II. Application to *Escherichia coli* and hybridoma cell metabolism. *J. Theor. Biol.*, 155:215–42, 1992.
- [76] K. L. Schulze and R. S. Lipe. Relationship between substrate concentration, growth rate, and respiration rate of *Escherichia coli* in continuous culture. *Arch. Mikrobiol.*, 48:1–20, 1964.
- [77] M. L. Shuler, S. Leung, and C. C. Dick. A mathematical model for the growth of a single bacterial cell. *Ann. N. Y. Acad. Sci.*, 326:35–55, 1979.
- [78] B. Sonnleitner and O. Kappeli. Growth of *Saccharomyces cerevisiae* is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis. *Biotech. Bioeng.*, 28:927–37, 1986.
- [79] Structural and organizational aspects of metabolic regulation. In P. A. Srere, M. E. Jones, and C. K. Matthews, editors, *UCLA Symposia on molecular and cellular Biology, New Series, Vol. 133*. Wiley-Liss, New York, 1990.
- [80] G. Stephanopoulos and J. J. Vallino. Network rigidity and metabolic engineering in metabolite overproduction. *Science*, 252:1675–81, 1991.
- [81] A. H. Stouthamer and C. W. Bettenhaussen. Determination of the efficiency of oxidative phosphorylation in continuous cultures of *Aerobacter aerogenes*. *Arch. Microbiol.*, 102:187–92, 1975.
- [82] A. H. Stouthamer and C. W. Bettenhaussen. Energetic aspects of anaerobic growth of *Aerobacter aerogenes* in complex medium. *Arch. Microbiol.*, 111:21–3, 1976.
- [83] D. W. Tempest and O. M. Neijssel. Overflow metabolism in aerobic micro-organisms. *Biochem. Soc. Trans.*, 7:82–5, 1979.
- [84] W. Trobner and R. Piechocki. Selective advantage of *polA1* mutator over *polA+* strains of *Escherichia coli* in a chemostat. *Naturwissenschaften*, 72:377–8, 1985.

- [85] A. Varma, B. W. Boesch, and B. O. Palsson. Biochemical production capabilities of *Escherichia coli*. *Biotech. Bioeng.*, 42:59–73, 1993.
- [86] A. Varma, B. W. Boesch, and B. O. Palsson. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under varying oxygenation rates. *Applied and Environmental Microbiology*, 59:2465–73, 1993.
- [87] A. Varma and B. O. Palsson. Metabolic capabilities of *Escherichia coli*: I. Synthesis of biosynthetic precursors and cofactors. *J. Theor. Biol.*, to appear, 1993.
- [88] A. Varma and B. O. Palsson. Metabolic capabilities of *Escherichia coli*: II. Optimal growth patterns. *J. Theor. Biol.*, to appear, 1993.
- [89] A. Varma and B. O. Palsson. Predictions for oxygen supply control to enhance population stability of engineered production strains. *Biotech. Bioeng.*, to appear, 1993.
- [90] K. Walsh and D. E. Jr Koshland. Branch point control by the phosphorylation state of isocitrate dehydrogenase. A quantitative examination of fluxes during a regulatory transition. *J. Biol. Chem.*, 260:8430–7, 1985.
- [91] M. R. Watson. A discrete model of bacterial metabolism. *Computer Applications in the Biosciences*, 2:23–7, 1986.
- [92] T. Wood. *The pentose phosphate pathway*. Academic Press, Inc., Orlando, Florida, 1985.
- [93] K. J. Zahl, C. Rose, and R. L. Hanson. Isolation and partial characterization of a mutant of *Escherichia coli* lacking pyridine nucleotide transhydrogenase. *Arch. Biochem. Biophys.*, 190:598–602, 1978.