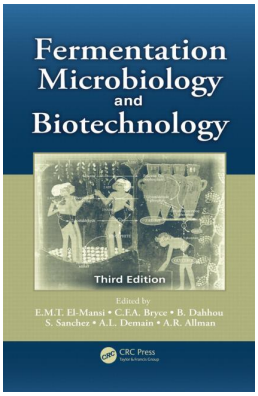

Fermentation Microbiology and Biotechnology

E.M.T. El-Mansi, C.F.A. Bryce, B. Dahhou, S. Sanchez, A.L. Demain, A.R. Allman

Fermentation Kinetics

Publication details
https://www.routledgehandbooks.com/doi/10.1201/b11490-4
Jens Nielsen
Published online on: 12 Dec 2011

3 Fermentation Kinetics: Central and Modern Concepts

Jens Nielsen

CONTENTS

3.1 INTRODUCTION

Growth of microbial cells is the result of many chemical reactions, including fueling reactions, biosynthetic reactions, and assembly reactions (see Figure 2.3). In preparation for cell division, the cells increase in size (or extend their hyphae, in the case of filamentous microorganisms) as the macromolecules are assembled *en route* to biomass formation. Biomass formation can be quantified by measuring the increase in dry weight (see Box 3.1), RNA, DNA, and/or proteins. *In situ* measurements of biomass formation during the course of fermentation can also be monitored by following the increase in turbidity at a given wavelength, as illustrated by Olsson and Nielsen (1997).

Growth of microbial cells is often illustrated with a batch-wise growth of a unicellular organism (either a bacterium or yeast). Here the growth occurs in a constant volume of medium with one growth-limiting substrate component that is used by the cells. Cell growth is generally quantified by the so-called *specific growth rate* μ (h^{-1}), which for such a culture is given by

$$\mu = \frac{1}{x}\frac{dx}{dt} \tag{3.1}$$

BOX 3.1 STANDARD OPERATING PROCEDURE (SOP) FOR DRY WEIGHT DETERMINATION

Biomass is most frequently determined by dry weight measurements. This can be done either using an oven or a microwave oven, with the latter being the fastest procedure. An important prerequisite for the measurement is that the sample is dried completely, and it is therefore important to apply a consistent procedure. A suggested protocol is as follows:

1. Dry the filter (pore size 0.45 μm for yeast or fungi, 0.20 μm for bacteria) on a glass dish in the microwave oven on 150 W for 10 min. Place a tissue paper between the glass and the filter so that the filter does not stick to the glass.
2. Place the filter in a desiccator and allow to cool for 10–15 min. Weigh the filter.
3. Filter the cell suspension through the filter and wash the cells with demineralized water.
4. Place the filter on the glass dish again and dry in the microwave oven for 15 min at 150 W.
5. Put the filter in a desiccator and allow to cool for 10–15 min. Weigh the filter.
6. If more than 30 mg dry weight is present on the filter, the time in the microwave oven may have to be longer.

where x is the biomass concentration (or cell number). The specific growth rate is related to the *doubling time* t_d(h) of the biomass through

$$t_d = \frac{\ln 2}{\mu} \tag{3.2}$$

The doubling time t_d is equal to the generation time for a cell (i.e., the length of a cell cycle for unicellular organisms), which is frequently used by life scientists to quantify the rate of cell growth.

The design and optimization of a given fermentation process require a quantitative description of the process, which, considering the nature of microbial growth, is generally a complex task. Furthermore, often the product is not the cells themselves but a compound synthesized by the cells, and depending on the type of product the kinetics of its formation may vary from one phase of growth to another. Thus, while primary metabolites (see Chapter 4 for more details) are typically formed in conjunction with cellular growth, an inverse relationship between product formation and cell growth is often found in the case of secondary metabolites (see Chapter 5 for more details), and here flux to product formation may be greatest in the stationary phase.

With these differences in mind, it is clear that quantification of product formation kinetics may be a difficult task. However, with the rapid progress in biological sciences, our understanding of cellular function has increased dramatically, and this may form the basis for far more advanced modeling of cellular growth kinetics than seen earlier. Thus, in the literature one may find mathematical models describing events like gene expression, kinetics of individual reactions in central pathways, together with macroscopic models that describe cellular growth and product formation with relatively simple mathematical expressions. These models cannot be compared directly because they serve completely different purposes, and it is therefore important to consider the aim of the modeling exercise in a discussion of mathematical models.

In this chapter, the applications of kinetic modeling to fermentation and cellular processes will be discussed.

3.2 FRAMEWORK FOR KINETIC MODELS

The net result of the many biochemical reactions within a single cell is the conversion of substrates to biomass and metabolic end products (see Figures 2.3 and 3.1). Clearly the number of reactions involved in the conversion of, say, glucose into biomass and desirable end products is very large, and it is therefore convenient to adopt the structure proposed by Neidhardt et al. (1990) for describing cellular metabolism, which can be summarized as follows:

Assembly reactions carry out chemical modifications of macromolecules, their transport to prespecified locations in the cell, and, finally, their assembly to form cellular structures such as cell walls, membranes, the nucleus, and so on.

Polymerization reactions represent directed, sequential linkage of activated molecules into long (branched or unbranched) polymeric chains. These reactions lead to the formation of macromolecules from a set of building blocks such as amino acids, nucleotides, and fatty acids.

Biosynthetic reactions produce the building blocks used in the polymerization reactions. They also produce coenzymes and related metabolic factors, including signal molecules. Furthermore, a large number of biosynthetic reactions occur in functional units called *biosynthetic pathways*, each of which consists of sequential reactions leading to the synthesis of one or more building blocks. Pathways are easily recognized and are often controlled *en bloc*. In some cases their reactions are catalyzed by enzymes made from a polycistronic message of messenger RNA (mRNA) transcribed from a set of 12 genes forming an operon. All biosynthetic pathways begin with one of only 12 precursor metabolites from which all building blocks can be synthesized. Some pathways begin directly with such a precursor metabolite, others indirectly by branching from an intermediate or an end product of a related pathway.

Fueling reactions produce the 12 precursor metabolites needed for biosynthesis. Additionally, they generate Gibbs free energy in the form of adenosine-5'-triphosphate (ATP), which is used for biosynthesis, polymerization, and assembling reactions. Finally, the fueling reactions produce the reducing power needed for biosynthesis. The fueling reactions include all biochemical pathways referred to as *catabolic pathways* (degrading and oxidizing substrates).

Thus, the conversion of glucose into cellular protein, for example, proceeds *via* precursor metabolites formed in the fueling reactions, further *via* building blocks (in this case amino acids) formed in the biosynthetic reactions, and finally through polymerization of the building blocks (or amino acids). In the fueling reactions there are many more intermediates than the precursor metabolites, and similarly a large number of intermediates are also involved in the conversion of precursor

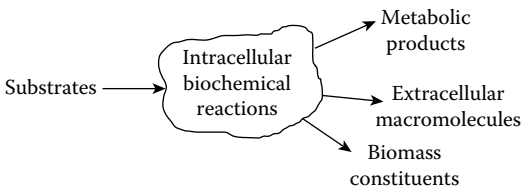

FIGURE 3.1 An overview of the intracellular biochemical reactions in micro-organisms; in addition to the formation of biomass constituents, for example, cellular protein, lipids, RNA, DNA, and carbohydrates, substrates are converted into primary metabolites, for example, ethanol, acetate, lactate; secondary metabolites, for example, penicillin; and/or extracellular macromolecules, for example, enzymes, heterologous proteins, polysaccharides.

TABLE 3.1
Overall Composition of an Average Cell of *Escherichia Coli*

Macromolecule	% of Total Dry Weight	Different Kinds of Molecules
Protein	55.0	1050
RNA	20.5	
rRNA	16.7	3
tRNA	3.0	60
mRNA	0.8	400
DNA	3.1	1
Lipid	9.1	4
Lippolysaccharide	3.4	1
Peptidoglycan	2.5	1
Glycogen	2.5	1
Metabolite pool	3.9	

Source: Data are taken from Ingraham, J.L., Maaloe, O., and Neidhardt, F.C., *Growth of the Bacterial Cell*. Sunderland, MA: Sinauer Associates, 1983.

metabolites into building blocks. The number of cellular metabolites is therefore very large, but still they only account for a small fraction of the total biomass (Table 3.1). The reason for this is the *en bloc* control of the individual reaction rates in the biosynthetic pathways mentioned above. Furthermore, the high affinity of enzymes for the reactants ensures that each metabolite can be maintained at a very low concentration even at a high flux through the pathway (see Box 3.2).

This control of the individual reactions in long pathways is very important for cell function, but it also means that in a quantitative description of cell growth it is not necessary to consider the kinetics of all the individual reactions, and this obviously leads to a significant reduction in the degree of complexity. Consideration of the kinetics of individual enzymes or reactions is therefore necessary only when the aim of the study is to quantify the relative importance of a particular reaction in a pathway.

3.2.1 Stoichiometry

The first step in a quantitative description of cellular growth is to specify the stoichiometry for those reactions to be considered for analysis. For this purpose it is important to distinguish between substrates, metabolic products, intracellular metabolites, and biomass constituents (Stephanopoulos et al. 1998):

A **substrate** is a compound present in the sterile medium, which can be further metabolized or directly incorporated into the cell.

A **metabolic product** is a compound produced by the cells and excreted to the extracellular medium.

Biomass constituents are pools of macromolecules that make up the biomass (e.g., RNA, DNA, protein, lipids, and carbohydrates), but also macromolecular products accumulating inside the cell (e.g., a polysaccharide or a nonsecreted heterologous protein).

Intracellular metabolites are all other compounds within the cell (i.e., glycolytic intermediates, precursor metabolites, and building blocks).

Note that this list distinguishes between biomass constituents and intracellular metabolites, because the timescales of their turnover in cellular reactions are very different: intracellular

BOX 3.2 CONTROL OF METABOLITE LEVELS IN BIOCHEMICAL PATHWAYS

The level of intracellular metabolites is normally very low. This is due to tight regulation of the enzyme levels and to the high affinity most enzymes have towards the reactants. To illustrate this consider two reactions of a pathway – one forming the metabolite X_i and the other consuming this metabolite:

$$....X_{i-1} \overset{v_i}{\rightarrow} X_i \overset{v_{i+1}}{\rightarrow} X_{i+1}....$$

Assuming that there is no allosteric regulation of the two enzyme-catalyzed reactions the kinetics can be described with reversible Michaelis–Menten kinetics:

$$v_i = \frac{v_{i,\max}\left(\frac{c_{i-1}}{K_{i-1}} - \frac{c_i}{K_i}\right)}{1 + \frac{c_{i-1}}{K_{i-1}} + \frac{c_i}{K_i}}$$

where c_i is the metabolite concentration and $v_{i,\max}$ expresses the enzyme activity. If the rate of the first reaction increases drastically, for example, due to an increase in the concentration of the metabolite X_{i-1}, the metabolite X_i will accumulate. This will lead to an increase in the second reaction rate and a decrease in the first reaction rate. Consequently the concentration of metabolite X_i will decrease again. The parameters K_i and K_{i-1} quantify the affinity of the enzyme for the reactant and the product in each reaction, and generally these are in the order of a few μM. Thus, even for low metabolite concentrations (in the order of 10 times K_i) the enzyme will be saturated and the reaction rate will be close to $v_{i,\max}$, but typically the metabolite concentration is of the order of K_i because hereby the enzyme can respond rapidly to changes in the metabolite level, and metabolite accumulations can be avoided.

metabolites have a very fast turnover (typically in the range of seconds) compared with that of macromolecules (typically in the range of hours). This means that on the timescale of growth, the intracellular metabolite pools can be assumed to be in pseudo-steady state.

With the goal of specifying a general stoichiometry for biochemical reactions, we consider a system where N substrates are converted to M metabolic products and Q biomass constituents. The conversions are carried out in J reactions in which K intracellular metabolites participate as pathway intermediates. The substrates are termed S_i, the metabolic products are termed P_i, the biomass constituents are termed $X_{\text{macro,i}}$, and the intracellular metabolites are termed $X_{\text{met,i}}$. With these definitions, the general stoichiometry for the j^{th} reaction can be specified as

$$\sum_{i=1}^{N} \alpha_{ji} S_i + \sum_{i=1}^{M} \beta_{ji} P_i + \sum_{i=1}^{Q} \gamma_{ji} X_{\text{macro,i}} + \sum_{i=1}^{K} g_{ji} X_{\text{met,i}} = 0;\ j = 1,..,J \tag{3.3}$$

Here, α_{ji} is a stoichiometric coefficient for the i^{th} substrate, β_{ji} is a stoichiometric coefficient for the i^{th} metabolic product, γ_{ji} is a stoichiometric coefficient for the i^{th} macromolecular pool, and g_{ji} is a stoichiometric coefficient for the i^{th} intracellular metabolite. All the stoichiometric coefficients are with sign. Thus, all compounds consumed in the j^{th} reaction have negative stoichiometric coefficients, whereas all compounds that are produced have positive stoichiometric coefficients.

Furthermore, compounds that do not participate in the j^{th} reaction have a stoichiometric coefficient of zero.

If there are many cellular reactions (i.e., J is large), it is convenient to write the stoichiometry for all the J cellular reactions in a compact form using matrix notation:

$$\mathbf{AS} + \mathbf{BP} + \mathrm{G}\mathbf{X}_{\mathrm{macro}} + \mathbf{G}\mathbf{X}_{\mathrm{met}} = 0 \tag{3.4}$$

where the matrices **A**, **B**, G, and **G** are stoichiometric matrices containing stoichiometric coefficients in the J reactions for the substrates, metabolic products, biomass constituents, and pathway intermediates, respectively. In these matrices, rows represent reactions and columns metabolites, that is, the element in the j^{th} row and the i^{th} column of **A** specifies the stoichiometric coefficient for the i^{th} substrate in the j^{th} reaction. Formulation of the stoichiometry in matrix form may seem rather complex; however, if the model is simple (i.e., only a few reactions, a few substrates, and a few metabolic products are considered), it is generally more convenient to use the simpler stoichiometric representation in Equation 3.3.

3.2.2 Reaction Rates

The stoichiometry of the individual reaction is the basis of any quantitative analysis. However, of equal importance is specification of the rate of the individual reactions. Normally the rate of a chemical reaction is given as the *forward rate*, which, if termed v_i, specifies that a compound that has a stoichiometric coefficient β in the i^{th} reaction is formed with the rate βv_i. Normally the stoichiometric coefficient for one of the compounds is arbitrarily set to 1, whereby the forward reaction rate becomes equal to the consumption or production of this compound in this particular reaction. For this reason the forward reaction rate is normally specified with the unit moles (or g) h^{-1}, or if the total amount of biomass is taken as reference (so-called specific rates) with the unit moles (or g) (g DW h^{-1}).

For calculation of the overall production or consumption rate, we have to sum the contributions from the different reactions, that is, the total specific consumption rate of the i^{th} substrate equals the sum of substrate consumptions in all the J reactions:

$$r_{s,i} = -\sum_{j=1}^{J} \alpha_{ji} v_j \tag{3.5}$$

The stoichiometric coefficients for substrates are generally negative, that is, the specific formation rate of the i^{th} substrate in the j^{th} reaction given by $\alpha_{ji} v_j$ is negative, but the specific substrate uptake rate is normally considered as positive, and a minus sign is therefore introduced in Equation 3.5. For the specific formation rate of the i^{th} metabolic product, similarly we have

$$r_{p,i} = \sum_{j=1}^{J} \beta_{ji} v_j \tag{3.6}$$

Equations 3.5 and 3.6 specify some very important relations between what can be directly measured: the specific substrate uptake rates and the specific product formation rates, and the rates of the reactions in the metabolic model. If a compound is consumed or formed in only one reaction, it is quite clear that we can get a direct measurement of this reaction rate. For the biomass constituents

and the intracellular metabolites, we can specify similar expressions for the net formation rate in all the J reactions:

$$r_{\text{macro,i}} = \sum_{j=1}^{J} \gamma_{ji} v_j \tag{3.7}$$

$$r_{\text{met,i}} = \sum_{j=1}^{J} g_{ji} v_j \tag{3.8}$$

These rates are net specific formation rates, because a compound may be formed in one reaction and consumed in another, and the rates specify the net results of consumption and formation in all the J cellular reactions. Thus, if $r_{\text{met,i}}$ is positive there is a net formation of the i^{th} intracellular metabolite, and if it is negative there is a net consumption of this metabolite. Finally, if $r_{\text{met,i}}$ is zero, the rates of formation of the i^{th} metabolite exactly balance its consumption.

If the forward reaction rates for the J cellular reactions are collected in the rate vector **v,** the summations in Equations 3.5 through 3.8 can be formulated in matrix notation as

$$\mathbf{r}_{\text{s}} = -\mathbf{A}^{\text{T}}\mathbf{v} \tag{3.9}$$

$$\mathbf{r}_{\text{p}} = \mathbf{B}^{\text{T}}\mathbf{v} \tag{3.10}$$

$$\mathbf{r}_{\text{macro}} = \text{G}^{\text{T}}\mathbf{v} \tag{3.11}$$

$$\mathbf{r}_{\text{met}} = \mathbf{G}^{\text{T}}\mathbf{v} \tag{3.12}$$

Here $\mathbf{r}_{\text{s}}$ is a rate vector containing the specific uptake rates of the N substrates, $\mathbf{r}_{\text{p}}$ a vector containing the specific formation rates of the M metabolic products, $\mathbf{r}_{\text{macro}}$ a vector containing the net specific formation rate of the Q biomass constituents, and $\mathbf{r}_{\text{met}}$ a vector containing the net specific formation rate of the K intracellular metabolites. Notice that what appears in the matrix equations are the transposed stoichiometric matrices, which are formed from the stoichiometric matrices by converting columns into rows and vice versa (see Example 3a, this chapter). Equations 3.7 and 3.11 give the net specific formation rate of biomass constituents, and because the intracellular metabolites only represent a small fraction of the total biomass, the specific growth rate μ of the total biomass is given as the sum of formation rates for all the macromolecular constituents:

$$\mu = \sum_{i=1}^{Q} r_{\text{macro,i}} = \mathbf{1}_Q^{\text{T}}\mathbf{r}_{\text{macro}} = \mathbf{1}_Q^{\text{T}}\Gamma^{\text{T}}\boldsymbol{\nu} \tag{3.13}$$

where $\mathbf{1}_Q$ is a Q-dimensional row vector with all elements being 1. Equation 3.13 is very fundamental because it links the information supplied by a detailed metabolic model with the macroscopic (and measurable) parameter μ. It clearly specifies that the formation rate of biomass is represented by a sum of formation of many different biomass constituents (or macromolecular pools), a point that will be discussed further in Section 3.4.3.

3.2.3 Yield Coefficients and Linear Rate Equations

The overall yield (e.g., how much carbon in the glucose ends up in the metabolite of interest) is a very important design parameter in many fermentation processes. This overall yield is normally

represented in the form of *yield coefficients*, which can be considered as relative rates (or fluxes) toward the product of interest with a certain compound as reference, often the carbon source or the biomass. These yield coefficients therefore have the units mass per unit mass of the reference (e.g., moles of penicillin formed per mole of glucose consumed or g protein formed per g biomass formed). An often used yield coefficient in the design and operation of aerobic fermentations is the respiratory quotient (RQ), which specifies the moles of carbon dioxide formed per mole of oxygen consumed (see also Example 3a). Several different formulations of the yield coefficients can be found in the literature. Here we will use the formulation of Nielsen et al. (2003), where the yield coefficient is stated with a double subscript Y_{ij}, which states that a mass of j is formed or consumed per mass of i formed or consumed. With the i^{th} substrate as the reference compound, the yield coefficients are given by

$$Y_{s_i s_j} = \frac{r_{s,j}}{r_{s,i}} \tag{3.14}$$

$$Y_{s_i p_j} = \frac{r_{p,j}}{r_{s,i}} \tag{3.15}$$

$$Y_{s_i x} = \frac{\mu}{r_{s,i}} \tag{3.16}$$

In the classical description of cellular growth introduced by Monod (1942) (see Section 3.4.2), the yield coefficient Y_{sx} was taken to be constant, and all the cellular reactions were lumped into a single overall growth reaction where substrate is converted to biomass. However, in the late 1950s it was shown (Herbert 1959) that the yield of biomass with respect to substrate is not constant. In order to describe this, Herbert introduced the concept of *endogenous metabolism* and specified substrate consumption for this process in addition to that for biomass synthesis. At the same time, Luedeking and Piret (1959) found that lactic acid bacteria produce lactic acid at nongrowth conditions, which was consistent with an endogenous metabolism of the cells. Their results indicated a linear correlation between the specific lactic acid production rate and the specific growth rate:

$$r_p = a\mu + b \tag{3.17}$$

In the mid-1960s, Pirt (1965) introduced a similar linear correlation between the specific rate of substrate uptake and the specific growth rate, and suggested the term *maintenance*, which is currently a widely used concept in endogenous metabolism. The linear correlation of Pirt takes the form of

$$r_s = Y_{xs}^{\text{true}}\,\mu + m_s \tag{3.18}$$

where Y_{xs}^{true} is referred to as the true yield coefficient and m_s as the maintenance coefficient. With the introduction of the linear correlations, the yield coefficients can obviously not be constants. Thus, for the biomass yield on the substrate:

$$Y_{sx} = \frac{\mu}{Y_{xs}^{\text{true}}\,\mu + m_s} \tag{3.19}$$

TABLE 3.2
True Yield and Maintenance Coefficients for Different Microbial Species and Growth on Glucose *or Glycerol*

Organism	Substrate	Y_{xs}^{true} [g (g DW)$^{-1}$]	m_s [g (g DW h)$^{-1}$]
Aspergillus nidulans		1.67	0.020
Candida utilis		2.00	0.031
Escherichia coli		2.27	0.057
Klebsiella aerogenes		2.27	0.063
Penicillium chrysogenum		2.17	0.021
Saccharomyces cerevisiae		1.85	0.015
Aerobacter aerogenes	Glycerol	1.79	0.089
Bacillus megatarium		1.67	–
Klebsiella aerogenes		2.13	0.074

Source: Data are taken from Nielsen, J. and Villadsen, J., *Bioreaction Engineering Principles*. New York: Plenum Press, 1994.

which shows that Y_{sx} decreases at low specific growth rates where an increasing fraction of the substrate is used to meet the maintenance requirements of the cell. When the specific growth rate becomes large, the yield coefficient approaches the reciprocal of Y_{xs}^{true}. A compilation of true yield and maintenance coefficients for various microbial species is given in Table 3.2.

The empirically derived linear correlations are very useful for correlating growth data, especially in steady-state continuous cultures where linear correlations similar to Equation 3.18 were found for most of the important specific rates. The remarkable robustness and general validity of the linear correlations indicate that they have a fundamental basis, and this basis is the continuous supply and consumption of ATP, which are tightly coupled in all cells. Thus the role of the energy-producing substrate is to provide ATP to drive biosynthesis and polymerization reactions as well as cell maintenance processes according to the linear relationship:

$$r_{ATP} = Y_{xATP}\mu + m_{ATP} \tag{3.20}$$

which is a formal analog to the linear correlation of Pirt. Equation 3.20 states that ATP produced balances the consumption for growth and for maintenance, and if the ATP yield on the energy-producing substrate is constant (i.e., r_{ATP} is proportional to r_s), it is quite obvious that Equation 3.20 can be used to derive the linear correlation in Equation 3.18, as illustrated in Example 3a. Notice that Y_{xATP} in Equation 3.20 is a true yield coefficient, but it is normally specified without the superscript "true."

The concept of balancing ATP production and consumption can be extended to other cofactors (e.g., NADH and NADPH), and as such it is possible to derive linear rate equations for three different cases (Nielsen and Villadsen 1994):

Anaerobic growth where ATP is supplied by substrate-level phosphorylation
Aerobic growth without metabolite formation
Aerobic growth with metabolite formation

For aerobic growth with metabolite formation, the specific substrate uptake rate takes the form of

$$r_s = Y_{xs}^{true}\mu + Y_{ps}^{true}r_p + m_s \tag{3.21}$$

This linear rate equation can be interpreted as a metabolic model with three reactions:

Conversion of substrate to biomass with a stoichiometric coefficient for the substrate and a forward reaction rate equal to the specific growth rate
Conversion of substrate to the metabolic product with a stoichiometric coefficient for the substrate and a forward reaction rate equal to the specific product formation rate
Metabolism of substrate to meet the maintenance requirements (normally, the substrate is oxidized to carbon dioxide) with the rate m_s

Consequently, the stoichiometry for these three reactions can be specified as

$$-Y_{xs}^{\text{true}} S + X = 0;\ \mu \tag{3.22}$$

$$-Y_{ps}^{\text{true}} S + P = 0;\ r_p \tag{3.23}$$

$$-S = 0;\ m_s \tag{3.24}$$

With this stoichiometry, the linear rate Equation 3.21 can easily be derived using Equation 3.5, that is, the overall specific substrate consumption rate is the sum of substrate consumption for growth, metabolite formation, and maintenance.

Thus, it is important to distinguish between true yield coefficients (which are rather stoichiometric coefficients) and overall yield coefficients, which can be taken to be stoichiometric coefficients in *one lumped reaction* (often referred to as the *black box model*; see Section 3.2.4), which represents all the cellular processes:

$$-Y_{xs} S + Y_{xp} P + X = 0;\ \mu \tag{3.25}$$

Despite the subscript "true," the true yield coefficients are only parameters for a given cellular system as they simply represent overall stoichiometric coefficients in lumped reactions; for example, reaction (3.22) is the sum of all reactions involved in the conversion of substrate into biomass. If, for example, the environmental conditions change, a different set of metabolic routes may be activated, and this may result in a change in the overall recovery of carbon in each of the three processes mentioned above (i.e., the values of the true yield coefficients change). Even the more fundamental Y_{xATP} cannot be taken to be constant, as illustrated in a detailed analysis of lactic acid bacteria (Benthin et al. 1994).

Example 3a: Metabolic Model for Aerobic Growth of *Saccharomyces cerevisiae*

To illustrate the derivation of the linear rate equations for an aerobic process with metabolite formation, we consider a simple metabolic model for the yeast *S. cerevisiae*. For this purpose we set up a stoichiometric model that summarizes the overall cellular metabolism, and based on assumptions of pseudo-steady state for ATP, NADH, and NADPH, linear rate equations can be derived where the specific uptake rates for glucose and oxygen and the specific carbon dioxide formation rate are given as functions of the specific growth rate. Furthermore, by evaluating the parameters in these linear rate equations, which can be done from a comparison with experimental data, information on key energetic parameters may be extracted.

TABLE 3.3
Macromolecular Composition of *Saccharomyces cerevisiae*

Macromolecule	Content [g (g DW)$^{-1}$]
Protein	0.39
Polysaccharides + trehalose	0.39
DNA + RNA	0.11
Phospholipids	0.05
Triacylglycerols	0.02
Sterols	0.01
Ash	0.03

From an analysis of all the biosynthetic reactions, the overall stoichiometry for synthesis of the constituents of a *S. cerevisiae* cell can be specified (Oura 1983) as

$$CH_{1.62}O_{0.53}N_{0.15} + 0.12\ CO_2 + 0.397\ NADH - 1.12\ CH_2O - 0.15\ NH_3 - Y_{xATP}ATP - 0.212\ NADPH = 0 \quad (3a1)$$

The stoichiometry (3a1) holds for a cell with the composition specified in Table 3.3; the substrate is glucose, and inorganic salts (i.e., ammonia) are the nitrogen source. The stoichiometry is given on a C-mole basis (i.e., glucose is specified as CH_2O), and the elemental composition of the biomass was calculated from the macromolecular composition to be $CH_{1.62}O_{0.53}N_{0.15}$ (see Table 3.3). The ATP and NADPH required for biomass synthesis are supplied by the catabolic pathways, and excess NADH formed in the biosynthetic reactions is, together with NADH formed in the catabolic pathways, reoxidized by transfer of electrons to oxygen via the electron transport chain. Reactions (3a2) through (3a5) specify the overall stoichiometry for the catabolic pathways. Reaction (3a2) specifies NADPH formation by the PP pathway, where glucose is completely oxidized to CO_2; reaction (3a3) is the overall stoichiometry for the combined EMP pathway and the TCA cycle; reaction (3a4) is the fermentative glucose metabolism, where glucose is converted to ethanol (this reaction only runs at high glucose uptake rates); and, finally, reaction (3a5) is the overall stoichiometry for the oxidative phosphorylation, where the P/O ratio is the overall (or operational) P/O ratio for the oxidative phosphorylation:

$$CO_2 + 2\ NADPH - CH_2O = 0 \quad (3a2)$$

$$CO_2 + 2\ NADH + 0.667\ ATP - CH_2O = 0 \quad (3a3)$$

$$CH_3O_{0.5} + 0.5\ CO_2 + 0.5\ ATP - 1.5\ CH_2O = 0 \quad (3a4)$$

$$P/O\ ATP - 0.5\ O_2 - NADH = 0 \quad (3a5)$$

Finally, consumption of ATP for maintenance is included simply as a reaction where ATP is used:

$$-ATP = 0 \quad (3a6)$$

Note that with the stoichiometry given on a C-mole basis, the stoichiometric coefficients extracted from the biochemistry (e.g., the formation of 2 moles ATP per mole glucose in the EMP pathway) are divided by six, because glucose contains 6 C moles per mole.

Above, the stoichiometry is written as in Equation 3.3, but we can easily convert it to the more compact matrix notation of Equation 3.4:

$$\begin{pmatrix} -1.120 & 0 \\ -1 & 0 \\ -1 & 0 \\ -1.5 & 0 \\ 0 & -0.5 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} S_{glc} \\ S_{o_2} \end{pmatrix} + \begin{pmatrix} 0 & 0.120 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0.5 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_{eth} \\ P_{CO_2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} X + \begin{pmatrix} -Y_{xATP} & 0.397 & -0.212 \\ 0 & 0 & 2 \\ 0.667 & 2 & 0 \\ 0.5 & 0 & 0 \\ P/O & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} X_{ATP} \\ X_{NADH} \\ X_{NADPH} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{3a7}$$

where X represents the biomass.

We now collect the forward reaction rates for the six reactions in the rate vector $\mathbf{v}$ given by

$$\boldsymbol{v} = \begin{pmatrix} \mu \\ v_{PP} \\ v_{EMP} \\ r_{eth} \\ v_{OP} \\ m_{ATP} \end{pmatrix}$$

In analogy with Equation 3.18, we balance the production and consumption of the three cofactors ATP, NADH, and NADPH. This gives the three equations:

$$-Y_{xATP}\mu + 0.667v_{EMP} + 0.5r_{eth} + P/Ov_{OP} - m_{ATP} = 0 \tag{3a8}$$

$$0.397\mu + 2v_{EMP} - v_{OP} = 0 \tag{3a9}$$

$$-0.212\mu + 2v_{PP} = 0 \tag{3a10}$$

Notice that these balances correspond to zero net specific formation rates for the three cofactors, and the three balances can therefore also be derived using Equation 3.12:

$$r_{met} = G^T v = \begin{pmatrix} -Y_{xATP} & 0 & 0.667 & 0.5 & P/O & -1 \\ 0.397 & 0 & 2 & 0 & -1 & 0 \\ -0.212 & 2 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ v_{PP} \\ v_{EMP} \\ r_{eth} \\ v_{OP} \\ m_{ATP} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{3a11}$$

In addition to the three balances (3a8) through (3a10), we have the relationships between the reaction rates and the specific substrate uptake rates and the specific product formation rate given by Equations 3a5 and 3a6, or, using the matrix notation of Equations 3a9 and 3a10:

$$\begin{pmatrix} r_{glc} \\ r_{O_2} \end{pmatrix} = -\begin{pmatrix} -1.120 & -1 & -1 & -1.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.5 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ v_{PP} \\ v_{EMP} \\ r_{eth} \\ v_{OP} \\ m_{ATP} \end{pmatrix} \tag{3a12}$$

$$= \begin{pmatrix} 1.120\mu + v_{PP} + v_{EMP} + 1.5r_{eth} \\ 0.5v_{OP} \end{pmatrix}$$

$$\begin{pmatrix} r_{eth} \\ r_{CO_2} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0.120 & 1 & 1 & 0.5 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ v_{PP} \\ v_{EMP} \\ r_{eth} \\ v_{OP} \\ m_{ATP} \end{pmatrix} \tag{3a13}$$

$$= \begin{pmatrix} r_{eth} \\ 0.120\mu + v_{PP} + v_{EMP} + 0.5r_{eth} \end{pmatrix}$$

Clearly, the specific ethanol production rate is equal to the rate of reaction (3a5) because the stoichiometric coefficient for ethanol in this reaction is 1, and it is the only reaction where ethanol is involved. Using the combined set of Equations 3a11 through 3a13, the four reaction rates v_{EMP}, v_{PP}, v_{OP}, and m_{ATP} can be eliminated and the linear rate equations 3a14 through 3a16 can be derived:

$$r_{glc} = (a + 1.226)\mu + (1.5 - b)r_{eth} + c = Y_{xs}^{true}\mu + Y_{ps}^{true}r_{eth} + m_s \tag{3a14}$$

$$r_{CO_2} = (a + 0.226)\mu + (0.5 - b)r_{eth} + c = Y_{xc}^{true}\mu + Y_{pc}^{true}r_{eth} + m_c \tag{3a15}$$

$$r_{O_2} = (a + 0.229)\mu - br_{eth} + c = Y_{xo}^{true}\mu + Y_{po}^{true}r_{eth} + m_o \tag{3a16}$$

The three common parameters *a*, *b,* and *c* are functions of the energetic parameters Y_{xATP}, m_{ATP}, and the P/O ratio according to Equations 3a17 through 3a19:

$$a = \frac{Y_{xATP} - 0.458P/O}{0.667 + 2P/O} \tag{3a17}$$

$$b = \frac{0.5}{0.667 + 2P/O} \tag{3a18}$$

$$c = \frac{m_{ATP}}{0.667 + 2P/O} \tag{3a19}$$

If there is no ethanol formation, which is the case at low specific glucose uptake rates, Equation 3a14 reduces to the linear rate Equation 3.18, but the parameters of the correlation are determined by basic energetic parameters of the cells. It is seen that the parameters in the linear correlations are coupled via the balances for ATP, NADH, and NADPH, and the three true yield coefficients cannot take any value. Furthermore, the maintenance coefficients are the same. This is due to the use of the units C-moles per C-mole biomass per hour for the specific rates. If other units are used

for the specific rates, the maintenance coefficients will not take the same values but will remain proportional. This coupling of the parameters shows that there are only three degrees of freedom in the system, and one actually only has to determine two yield coefficients and one maintenance coefficient—the other parameters can be calculated using Equations 3a14 through 3a16.

The derived linear rate equations are certainly useful for correlating experimental data, but they also allow evaluation of the key energetic parameters Y_{xATP}, m_{ATP}, and the operational P/O ratio. Thus, if the true yield coefficients and the maintenance coefficients of Equations 3a14 through 3a16 are estimated, the values of *a*, *b,* and *c* can be found, and these three parameters relate the three energetic parameters through Equations 3a17 through 3a19. Thus, from one of the ethanol yield coefficients, *b* can be found, and, thereafter, the P/O ratio can be determined. Then m_{ATP} can be found from one of the maintenance coefficients, and finally Y_{xATP} can be found from one of the biomass yield coefficients. In practice, however, it is difficult to extract sufficiently precise values of the true yield coefficients from experimental data to estimate the energetic parameters—especially since the three parameters *a*, *b,* and *c* are closely correlated (especially *b* and *c*). However, if either the P/O ratio or Y_{xATP} is known, Equations 3a14 through 3a16 allow an estimation of the two remaining unknown energetic parameters. Consider the situation where there is no ethanol formation; here the true yield coefficient for biomass is 1.48 C-moles glucose (C-mole biomass)$^{-1}$ and the maintenance coefficient (equal to *b*) is 0.012 C-moles glucose (C-mole biomass h)$^{-1}$ (both values taken from Table 3.2). Thus, *a* is equal to 0.254 moles ATP (C-mole^{-1} biomass). If the operational P/O ratio is about 1.5 (which is a reasonable value for *S. cerevisiae*), we find that Y_{xATP} is 1.62 moles ATP (C-mole^{-1} biomass) or about 67 mmoles ATP (g DW^{-1}). Similarly, we find m_{ATP} to be about 2 mmoles (g DW h^{-1}).

In connection with baker's yeast production, it is important to maximize the yield of biomass on glucose:

$$Y_{sx} = \frac{\mu}{(a+1.226)\mu + (1.5-b)r_{eth} + c} \tag{3a20}$$

Clearly, this can best be done if ethanol production is avoided. Thus, the glucose uptake rate is to be controlled below a level where there is respiro-fermentative metabolism. A very good indication of whether there is respiro-fermentative metabolism is the RQ:

$$RQ = \frac{(a+0.226)\mu + (0.5-b)r_{eth} + c}{(a+0.229)\mu - br_{eth} + c} \tag{3a21}$$

If there is no ethanol production, RQ will be close to 1 (independent of the specific growth rate), whereas if there is ethanol production, RQ will be above 1 and will increase with r_{eth} (*b* is always less than 0.5). From measurements of carbon dioxide and oxygen in the exhaust gas, the RQ can be evaluated. If it is above 1, there is respiro-fermentative metabolism resulting in ethanol formation and hence a low yield of biomass on sugar. This is caused by so-called glucose repression of respiration, where a high sugar concentration causes decreased activity of the respiration resulting in ethanol production due to overflow metabolism. Thus, if RQ is larger than 1 it means that the sugar concentration in the reactor must be reduced, and this can be done by reducing the feed rate to the reactor (typically baker's yeast production is operated as a fed-batch process; see Section 3.4).

3.2.4 Black Box Model

In the black box model of cellular growth, all the cellular reactions are lumped into a single reaction. In this overall reaction, the stoichiometric coefficients are identical to the yield coefficients [see also Equation 3.25], and it can therefore be presented as

$$X + \sum_{i=1}^{M} Y_{xp_i} P_i - \sum_{i=1}^{N} Y_{xs_i} S_i = 0 \tag{3.26}$$

Because the stoichiometric coefficient for biomass is 1, the forward reaction rate is given by the specific growth rate of the biomass, which together with the yield coefficients completely specifies the system. As discussed in Section 3.2.3, the yield coefficients are not constants, and the black box model can therefore not be applied to correlate, for instance, the specific substrate uptake rate with the specific growth rate. However, it is very useful for validation of experimental data because it can form the basis for setting up elemental balances. Thus, in the black box model there are $(M + N + 1)$ parameters: M yield coefficients for the metabolic products, N yield coefficients for the substrates, and the forward reaction rate μ. Because mass is conserved in the overall conversion of substrates to metabolic products and biomass, the $(M + N + 1)$ parameters of the black box model are not completely independent but must satisfy several constraints. Thus, the elements flowing into the system must balance the elements flowing out of the system (e.g., the carbon entering the system via the substrates has to be recovered in the metabolic products and biomass). Each element considered in the black box obviously yields one constraint. Thus, a carbon balance gives

$$1 + \sum_{i=1}^{M} f_{\mathrm{p,i}} Y_{xp_i} - \sum_{i=1}^{N} f_{\mathrm{s,i}} Y_{xs_i} = 0 \tag{3.27}$$

where $f_{\mathrm{s,i}}$ and $f_{\mathrm{p,i}}$ represent the carbon content (C-moles mole^{-1}) in the i^{th} substrate and the i^{th} metabolic product, respectively. In the above equation, the elemental composition of biomass is normalized with respect to carbon (i.e., it is represented by the form $CH_aO_bN_c$; see also Example 3a). The elemental composition of biomass depends on its macromolecular content and, therefore, on the growth conditions and the specific growth rate (e.g., the nitrogen content is much lower under nitrogen-limited conditions than under carbon-limited conditions; see Table 3.4). However, except for extreme situations, it is reasonable to use the general composition formula $CH_{1.8}O_{0.5}N_{0.2}$ whenever the biomass composition is not exactly known. Often the elemental composition of substrates and metabolic products is normalized with respect to their carbon content, for example glucose is specified as CH_2O (see also Example 3a). Equation 3.27 is then written on a per C-mole basis as

$$1 + \sum_{i=1}^{M} Y_{xp_i} - \sum_{i=1}^{N} Y_{xs_i} = 0 \tag{3.28}$$

In Equation 3.28, the yield coefficients have units of C-moles per C-mole biomass. Conversion to this unit from other units is illustrated in Box 3.3. Equation 3.28 is very useful for checking the consistency of experimental data. Thus, if the sum of carbon in the biomass and the metabolic products does not equal the sum of carbon in the substrates, there is an inconsistency in the experimental data.

Similar to Equation 3.27, balances can be written for all other elements participating in the conversion (3.26). Thus, the hydrogen balance will read

$$a_x + \sum_{i=1}^{M} a_{\mathrm{p,i}} Y_{xp_i} - \sum_{i=1}^{N} a_{\mathrm{s,i}} Y_{xs_i} = 0 \tag{3.29}$$

where $a_{\mathrm{s,i}}$, $a_{\mathrm{p,i}}$, and a_x represent the hydrogen content (moles C-mole^{-1} if a C-mole basis is used) in the i^{th} substrate, the i^{th} metabolic product, and the biomass, respectively. Similarly, we have for the oxygen and nitrogen balances

$$b_x + \sum_{i=1}^{M} b_{\mathrm{p,i}} Y_{xp_i} - \sum_{i=1}^{N} b_{\mathrm{s,i}} Y_{xs_i} = 0 \tag{3.30}$$

TABLE 3.4
Elemental Composition of Biomass for *Several Microorganisms*

Microorganism	Elemental Composition	Ash Content (w/w %)	Growth Conditions
Candida utilis	$CH_{1.83}O_{0.46}N_{0.19}$	7.0	Glucose limited, $D = 0.05\ h^{-1}$
	$CH_{1.87}O_{0.56}N_{0.20}$	7.0	Glucose limited, $D = 0.45\ h^{-1}$
	$CH_{1.83}O_{0.54}N_{0.10}$	7.0	Ammonia limited, $D = 0.05\ h^{-1}$
	$CH_{1.87}O_{0.56}N_{0.20}$	7.0	Ammonia limited, $D = 0.45\ h^{-1}$
Klebsiella aerogenes	$CH_{1.75}O_{0.43}N_{0.22}$	3.6	Glycerol limited, $D = 0.10\ h^{-1}$
	$CH_{1.73}O_{0.43}N_{0.24}$	3.6	Glycerol limited, $D = 0.85\ h^{-1}$
	$CH_{1.75}O_{0.47}N_{0.17}$	3.6	Ammonia limited, $D = 0.10\ h^{-1}$
	$CH_{1.73}O_{0.43}N_{0.24}$	3.6	Ammonia limited, $D = 0.80\ h^{-1}$
Saccharomyces cerevisiae	$CH_{1.82}O_{0.58}N_{0.16}$	7.3	Glucose limited, $D = 0.080\ h^{-1}$
	$CH_{1.78}O_{0.60}N_{0.19}$	9.7	Glucose limited, $D = 0.255\ h^{-1}$
	$CH_{1.94}O_{0.52}N_{0.25}$	5.5	Unlimited growth
Escherichia coli	$CH_{1.77}O_{0.49}N_{0.24}$	5.5	Unlimited growth
	$CH_{1.83}O_{0.50}N_{0.22}$	5.5	Unlimited growth
	$CH_{1.96}O_{0.55}N_{0.25}$	5.5	Unlimited growth
	$CH_{1.93}O_{0.55}N_{0.25}$	5.5	Unlimited growth
Pseudomonas fluorescens	$CH_{1.83}O_{0.55}N_{0.26}$	5.5	Unlimited growth
Aerobacter aerogenes	$CH_{1.64}O_{0.52}N_{0.16}$	7.9	Unlimited growth
Penicillium chrysogenum	$CH_{1.70}O_{0.58}N_{0.15}$		Glucose limited, $D = 0.038\ h^{-1}$
	$CH_{1.68}O_{0.53}N_{0.17}$		Glucose limited, $D = 0.098\ h^{-1}$
Aspergillus niger	$CH_{1.72}O_{0.55}N_{0.17}$	7.5	Unlimited growth
Average	$CH_{1.81}O_{0.52}N_{0.21}$	6.0	

Source: Compositions for *P. chrysogenum* are taken from Christensen, L.H., et al., *J. Biotechnol*. 42:95–107, 1995; other data are taken *from* Roels, J.A., *Energetics and Kinetics in Biotechnology*. Amsterdam: Elsevier Biomedical Press, 1983.

BOX 3.3 CALCULATION OF YIELDS WITH RESPECT TO C-MOLE BASIS

Yield coefficients are typically described as moles (g DW)$^{-1}$ or g (g DW)$^{-1}$. To convert the yield coefficients to a C-mole basis, information on the elemental composition and the ash content of biomass is needed. To illustrate the conversion, we calculate the yield of 0.5 g DW biomass (g glucose)$^{-1}$ on a C-mole basis. First, we convert the g DW biomass to an ash-free basis, that is determine the amount of biomass that is made up of carbon, nitrogen, oxygen and hydrogen (and, in some cases, also phosphorus and sulfur). With an ash content of 8% we have 0.92 g ash-free biomass (g DW biomass)$^{-1}$, which gives a yield of 0.46 g ash-free biomass (g glucose)$^{-1}$. This yield can now be directly converted to a C-mole basis using the molecular weights in g C-mole^{-1} for ash-free biomass and glucose. With the standard elemental composition for biomass of $CH_{1.8}O_{0.5}N_{0.2}$ we have a molecular weight of 24.6 g ash-free biomass C-mole^{-1}, and therefore find a yield of 0.46/24.6 = 0.0187 C-moles biomass (g glucose)$^{-1}$. Finally, by multiplication with the molecular weight of glucose on a C-mole basis (30 g C-mole^{-1}), a yield of 0.56 C-moles biomass (C-mole glucose)$^{-1}$ is found.

$$c_x + \sum_{i=1}^{M} c_{\mathrm{p,i}} Y_{xp_i} - \sum_{i=1}^{N} c_{\mathrm{s,i}} Y_{xs_i} = 0 \tag{3.31}$$

where $b_{\mathrm{s,i}}$, $b_{\mathrm{p,i}}$, and b_{x} represent the oxygen content (moles C-mole^{-1}) in the i^{th} substrate, the i^{th} metabolic product, and the biomass, respectively; and $c_{\mathrm{s,i}}$, $c_{\mathrm{p,i}}$, and c_{x} represent the nitrogen content (moles C-mole^{-1}) in the i^{th} substrate, the i^{th} metabolic product, and the biomass, respectively. Normally, only these four balances are considered; balances for phosphate and sulfate may also be set up, but generally these elements are of minor importance. The four elemental balances (3.28) through (3.31) can be conveniently written by collecting the elemental composition of biomass, substrates, and metabolic products in the columns of a matrix **E**, where the first column contains the elemental composition of biomass, columns 2 through $M + 1$ contain the elemental composition of the M metabolic products, and columns $M + 2$ to $M + N + 1$ contain the elemental composition of the N substrates. With the introduction of this matrix, the four elemental balances can be expressed as

$$\mathbf{E\,Y = 0} \tag{3.32}$$

where **Y** is a vector containing the yield coefficients (the substrate yield coefficients are given with a minus sign). With $N + M + 1$ variables, $N + M$ yield coefficients and the forward rate of reaction (3.26) and four constraints, the degree of freedom is $F = M + N + 1 - 4$. If exactly F variables are measured, it may be possible to calculate the other rates by using the four algebraic equations given by (3.32), but, in this case, there are no redundancies left to check the consistency of the data. For this reason, it is advisable to strive for more measurements than the degrees of freedom of the system.

Example 3b: Elemental Balances in a Simple Black Box Model

Consider the aerobic cultivation of the yeast *S. cerevisiae* on a defined, minimal medium (i.e., glucose is the carbon and energy source and ammonia is the nitrogen source). During aerobic growth, the yeast oxidizes glucose completely to carbon dioxide. However, as mentioned in Example 3a, high glucose concentrations cause repression of the respiratory system, resulting in ethanol formation. Thus, at these conditions both ethanol and carbon dioxide should be considered as metabolic products. Finally, water is formed in the cellular pathways. This is also included as a product in the overall reaction. Thus, the black box model for this system is

$$X + Y_{\mathrm{xe}}\ \mathrm{ethanol} + Y_{xc}\ \mathrm{CO_2} + Y_{xw}\ \mathrm{H_2O} - Y_{\mathrm{xs}}\ \mathrm{glucose} - Y_{xo}\ \mathrm{O_2} - Y_{xN}\ \mathrm{NH_3} = 0 \tag{3b1}$$

which can be represented with the yield coefficient vector

$$\mathbf{Y} = (1\ Y_{xe}\ Y_{xc}\ Y_{xw} - Y_{xs} - Y_{xo} - Y_{xN})\mathrm{T} \tag{3b2}$$

We now rewrite the conversion using the elemental composition of the substrates and metabolic products. For biomass we use the elemental composition of $\mathrm{CH_{1.83}O_{0.56}N_{0.17}}$, and therefore we have

$$\mathrm{CH_{1.83}O_{0.56}N_{0.17}} + Y_{xe}\ \mathrm{CH_3O_{0.5}} + Y_{xc}\ \mathrm{CO_2} + Y_{xw}\ \mathrm{H_2O} - Y_{xs}\ \mathrm{CH_2O} - Y_{xo}\ \mathrm{O_2} - Y_{xN}\ \mathrm{NH_3} = 0 \tag{3b3}$$

Some may find it difficult to identify $\mathrm{CH_3O_{0.5}}$ as ethanol, but the advantage of using the C-mole basis becomes apparent immediately when we look at the carbon balance:

$$1 + Y_{xe} + Y_{xc} - Y_{xs} = 0 \tag{3b4}$$

This simple equation is very useful for checking the consistency of experimental data. Thus, using the classical data of von Meyenburg (1969), we find $Y_{xe} = 0.713$, $Y_{xc} = 1.313$, and $Y_{xs} = 3.636$ at a dilution rate of $D = 0.3$ h^{-1} in a glucose-limited continuous culture. Obviously the data are not consistent as the carbon balance is not closed. A more elaborate data analysis (Nielsen and Villadsen 1994) suggests that the missing carbon may be accounted for by ethanol evaporation or stripping due to intensive aeration of the bioreactor.

Similarly, using Equation 3.31, we find that a nitrogen balance gives

$$Y_{xN} = 0.17 \tag{3b5}$$

If the yield coefficients for ammonia uptake and biomass formation do not conform to Equation 3b5, an inconsistency is identified in one of these two measurements, or the nitrogen content of the biomass is different from that specified.

We now write all four elemental balances in terms of the matrix equation 3.32:

$$\mathbf{E} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1.83 & 3 & 0 & 2 & 2 & 0 & 3 \\ 0.56 & 0.5 & 2 & 1 & 1 & 2 & 0 \\ 0.17 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} \leftarrow \text{carbon} \\ \leftarrow \text{hydrogen} \\ \leftarrow \text{oxygen} \\ \leftarrow \text{nitrogen} \end{matrix} \tag{3b6}$$

where the rows indicate, respectively, the content of carbon, hydrogen, oxygen, and nitrogen and the columns give the elemental composition of biomass, ethanol, carbon dioxide, water, glucose, oxygen, and ammonia, respectively. Using Equation 3.32, we find

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1.83 & 3 & 0 & 2 & 2 & 0 & 3 \\ 0.56 & 0.5 & 2 & 1 & 1 & 2 & 0 \\ 0.17 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ Y_{xe} \\ Y_{xc} \\ Y_{xw} \\ -Y_{xs} \\ -Y_{xo} \\ -Y_{xN} \end{pmatrix} = \begin{pmatrix} 1 + Y_{xe} + Y_{xc} - Y_{xs} \\ 1.83 + 3Y_{xe} + 2Y_{xw} - 2Y_{xs} - 3Y_{xN} \\ 0.56 + 0.5Y_{xe} + 2Y_{xc} + Y_{xw} - Y_{xs} - 2Y_{xo} \\ 0.17 - Y_{xN} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{3b7}$$

The first and last rows are identical to the balances derived in Equations 3b4 and 3b5 for carbon and nitrogen, respectively. The balances for hydrogen and oxygen introduce two additional constraints. However, because the rate of water formation is impossible to measure, one of these equations must be used to calculate this rate (or yield). This leaves only one additional constraint from these two balances.

3.3 MASS BALANCES FOR BIOREACTORS

In Section 3.2, we derived equations that relate the rates of the intracellular reaction with the rates of substrate uptake, metabolic product formation, and biomass formation. These rates are the key elements in the dynamic mass balances for the substrates, the metabolic products, and the biomass, which describe the change in time of the concentration of these state variables in a bioreactor. The bioreactor may be any type of device, ranging from a shake flask to a well-instrumented bioreactor. Figure 3.2 is a general representation of a bioreactor. It has a volume V (unit: L), and it is fed with a stream of fresh, sterile medium with a flow rate F (unit: L h^{-1}). Spent medium is removed with a flow rate of F_{out} (unit: L h^{-1}). The medium in the bioreactor is assumed to be completely (or ideally) mixed, that is, there is no spatial variation in the concentration of the different medium compounds. For small-volume bioreactors (<1 L) (including shake flasks), this can generally be achieved through

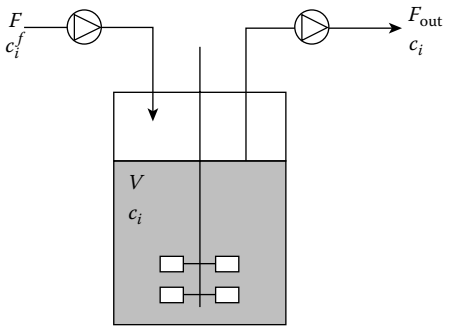

FIGURE 3.2 Bioreactor with addition of fresh, sterile medium and removal of spent medium where c^f_i is the concentration of the i^{th} compound in the feed and c_i is the concentration of the i^{th} compound in the spent medium. The bioreactor is assumed to be very well mixed (or ideal), so that the concentration of each compound in the spent medium becomes identical to its concentration in the bioreactor.

aeration and some agitation, whereas for laboratory stirred-tank bioreactors (1–10 L) special designs may have to be introduced in order to ensure a homogeneous medium (Sonnleitner and Fiechter 1988; Nielsen and Villadsen 1993). The bioreactor may be operated in many different modes of which we will only consider the three most common:

Batch, where $F = F_{out} = 0$ (i.e., the volume is constant)
Continuous, where $F = F_{out}$ 0 (i.e., the volume is constant)
Fed-batch (or semibatch), where F 0 and $F_{out} = 0$ (i.e., the volume increases)

These three different modes of reactor operation are discussed in this section, but first we derive general dynamic mass balances for the substrates, metabolic products, biomass constituents, and biomass.

3.3.1 Dynamic Mass Balances

The basis for derivation of the general dynamic mass balances is the mass balance equation

$$\text{Accumulated} = \text{Net formation rate} + \text{In} - \text{Out} \tag{3.33}$$

where the first term on the RHS is given by Equations 3.5 through 3.8 for substrate, metabolic product, biomass constituents, and intracellular metabolites, respectively. The term "In" represents the flow of the compound into the bioreactor, and the term "Out" the flow of the compound out from the bioreactor. In the following we consider substrates, metabolic products, biomass constituents, intracellular metabolites, and the total biomass separately.

We consider the i^{th} substrate, which is added to the bioreactor via the feed and is consumed by the cells present in the bioreactor. The mass balance for this compound is

$$\frac{d(c_{s,i}V)}{dt} = -r_{s,i}xV + Fc^f_{s,i} - F_{out}c_{s,i} \tag{3.34}$$

where r_i is the specific consumption rate of the compound (unit: moles (g DW h^{-1}); $c_{s,i}$ is the concentration in the bioreactor, which is assumed to be the same as the concentration in the outlet (unit: moles L^{-1}), $c^f_{s,i}$ (i) is the concentration in the feed (unit: moles L^{-1}); and x is the biomass concentration in the bioreactor (unit: g DW L^{-1}). The first term in Equation 3.34 is the accumulation

term, the second term is the consumption (or reaction) term, the third term is accounting for the inlet, and the last term is accounting for the outlet. Rearrangement of this equation gives

$$\frac{dc_{s,i}}{dt} = -r_{s,i}x + \frac{F}{V}c_{s,i}^{f} - \left(\frac{F_{out}}{V} + \frac{1}{V}\frac{dV}{dt}\right)c_{s,i} \tag{3.35}$$

For a fed-batch reactor,

$$F = \frac{dV}{dt} \tag{3.36}$$

and $F_{out} = 0$. So the term within the parentheses becomes equal to the so-called **dilution rate** given by

$$D = \frac{F}{V} \tag{3.37}$$

For a continuous and a batch reactor, the volume is constant (i.e., $dV/dt = 0$, and $F = F_{out}$), and so for these bioreactor modes also the term within the parentheses becomes equal to the dilution rate. Equation 3.35 therefore reduces to the mass balance (3.38) for any type of operation:

$$\frac{dc_{s,i}}{dt} = -r_{s,i}x + D(c_{s,i}^{f} - c_{s,i}) \tag{3.38}$$

The first term on the right-hand side of Equation 3.38 is the volumetric rate of substrate consumption, which is given as the product of the specific rate of substrate consumption and the biomass concentration. The second term accounts for the addition and removal of substrate from the bioreactor. The term on the left-hand side of Equation 3.38 is the accumulation term, which accounts for the change in time of the substrate, which in a batch reactor (where $D = 0$) equals the volumetric rate of substrate consumption.

Dynamic mass balances for the metabolic products are derived in analogy with those for the substrates and take the form

$$\frac{dc_{p,i}}{dt} = r_{p,i}x + D(c_{p,i}^{f} - c_{p,i}) \tag{3.39}$$

where the first term on the right-hand side is the volumetric formation rate of the i^{th} metabolic product. Normally, the metabolic products are not present in the sterile feed to the bioreactor, and $c_{p,i}^{f}$ is therefore often zero. In these cases the volumetric rate of product formation in a steady-state continuous reactor is equal to the dilution rate multiplied by the concentration of the metabolic product in the bioreactor (equal to that in the outlet).

With sterile feed, the mass balance for the total biomass is derived directly:

$$\frac{dx}{dt} = (\mu - D)x \tag{3.40}$$

where μ (unit: h^{-1}) is the specific growth rate of the biomass given by Equation 3.13.

For the biomass constituents, we normally use the biomass as the reference (i.e., their concentrations are given with the biomass as the basis). In this case, the mass balance for the i^{th} biomass constituent is derived from (sterile feed is assumed)

$$\frac{d(X_{\text{macro,i}}x\text{V})}{dt} = r_{\text{macro,i}}x\text{V} - F_{\text{out}}X_{\text{macro,i}}x \tag{3.41}$$

where $X_{\text{macro,i}}x$ is the concentration of the i^{th} biomass component in the bioreactor (unit: g L^{-1}) and $r_{\text{macro,i}}$ is the specific net rate of formation of the i^{th} biomass constituent. Rearrangement of Equation 3.41 gives

$$\frac{dX_{\text{macro,i}}}{dt} = r_{\text{macro,i}} - \left(\frac{F_{\text{out}}}{\text{V}} + \frac{1}{x}\frac{dx}{dt} + \frac{1}{\text{V}}\frac{d\text{V}}{dt}\right)X_{\text{macro,i}} \tag{3.42}$$

Again, we have that for any mode of bioreactor operation:

$$D = \frac{F_{\text{out}}}{\text{V}} + \frac{1}{\text{V}}\frac{d\text{V}}{dt} \tag{3.43}$$

which, together with the mass balance (3.40) for the total biomass concentration, gives the mass balance:

$$\frac{dX_{\text{macro,i}}}{dt} = r_{\text{macro,i}} - \mu X_{\text{macro,i}} \tag{3.44}$$

where $X_{\text{macro,i}}$ is the concentration of the i^{th} biomass constituent within the biomass. Different units may be applied for the concentrations of the biomass constituents, but they are normally given as g (g DW)$^{-1}$, because then the sum of all the concentrations equals 1, that is,

$$\sum_{i=1}^{Q} X_{\text{macro,i}} = 1 \tag{3.45}$$

Furthermore, this unit corresponds with the experimentally determined macromolecular composition of cells, where weight fractions are generally used. In Equation 3.44, it is observed that the mass balance for the biomass constituents is completely independent of the mode of operation of the bioreactor (i.e., the dilution rate does not appear in the mass balance). However, there is indirectly a coupling via the last term, which accounts for dilution of the biomass constituents when the biomass expands due to growth. Thus, if there is no net synthesis of a macromolecular pool, but the biomass still grows, the intracellular level decreases.

For intracellular metabolites it is not convenient to use the same unit for their concentrations as for the biomass constituents. These metabolites are dissolved in the matrix of the cell; therefore, it is more appropriate to use the unit moles per liquid cell volume for the concentrations. The intracellular concentration can then be compared directly with the affinities of enzymes, typically quantified by their *K*m values, which are normally given with the unit moles per liter. If the concentration is known in one unit, it is, however, easily converted to another unit if the density of the biomass (in the range of 1 g cell per mL cell) and the water content (in the range of 0.67 ml water per ml cell) is

known. Even though a different unit is applied, the biomass is still the basis, and the mass balance for the intracellular metabolites therefore takes the same form:

$$\frac{dX_{\text{met,i}}}{dt} = r_{\text{met,i}} - \mu X_{\text{met,i}} \tag{3.46}$$

where X_{met},i is the concentration of the i^{th} intracellular metabolite. It is important to distinguish between concentrations of intracellular metabolites given in moles per liquid reactor volume and in moles per liquid cell volume. If concentrations are given in the former unit, the mass balance will be completely different.

3.3.2 Batch Reactor

This is the classical operation of the bioreactor, and many life scientists use it because it can be carried out in a relatively simple experimental setup. Batch experiments have the advantage of being easy to perform, and by using shake flasks a large number of parallel experiments can be carried out. The disadvantage is that the experimental data are difficult to interpret because there are dynamic conditions throughout the experiment (i.e., the environmental conditions experienced by the cells vary with time). However, by using well-instrumented bioreactors, at least some variables (e.g., pH and dissolved oxygen tension) may be kept constant.

As mentioned in the previous section, the dilution rate is zero for a batch reactor, and the mass balances for the biomass and the limiting substrate therefore take the form

$$\frac{dx}{dt} = \mu x \quad x(t=0) = x_0 \tag{3.47}$$

$$\frac{dc_s}{dt} = -r_s x \quad c_s(t=0) = c_{s,0} \tag{3.48}$$

where x_0 indicates the initial biomass concentration, which is obtained immediately after inoculation, and c_s,0 is the initial substrate concentration. According to the mass balance, the biomass concentration will increase as indicated in Figure 3.3 and the substrate concentration will decrease until its concentration reaches zero and growth stops. Because the substrate concentration is zero at the end of the cultivation, the overall yield of biomass on the substrate can be found from

$$Y_{sx}^{\text{overall}} = \frac{x_{\text{final}} - x_0}{c_{s,0}} \tag{3.49}$$

where x final is the biomass concentration at the end of the cultivation. Normally $X_0 \ll x_{\text{final}}$, and the overall yield coefficient can therefore be estimated from the final biomass concentration and the initial substrate concentration alone. Notice that the yield coefficient determined from the batch experiment is the overall yield coefficient and not Y_{sx} or $(Y_{xs}^{\text{true}})^{-1}$. The yield coefficient Y_{sx} may well be time dependent as it is the ratio between the specific growth rate and the substrate uptake rate; see Equation 3.16. However, if there is little variation in these rates during the batch culture (e.g., if there is a long exponential growth phase and only a very short declining growth phase), the overall yield coefficient may be very similar to the yield coefficient. The true yield coefficient, on the other hand, is difficult to determine from batch cultivation because it requires information about the maintenance coefficients, which can hardly be determined from a batch experiment. However, in batch cultivation the specific growth rate is close to its maximum throughout most of the growth

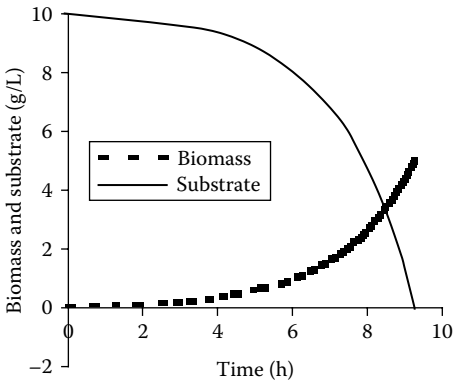

FIGURE 3.3 Batch fermentation described with Monod kinetics. The biomass concentration is found using Equation (c2) and the corresponding substrate concentration is found from Equation (c1). μ_{max} is 0.5 h^{-1}, K_s is 50 mg l^{-1} (a quite high value), and Y_{sx} is 0.5 g g^{-1}. The initial substrate concentration c_s, 0 is 10 g l^{-1}. The substrate concentration decreases from 0.5 g l^{-1} to 0 in less than 5 min, and this is the interesting substrate concentration range for estimation of K_s.

phase, and the substrate consumption due to maintenance is therefore negligible. According to Equation 3.17, the true yield coefficient is close to the observed yield coefficient determined from the final biomass concentration.

3.3.3 Chemostat

A typical operation of the continuous bioreactor is the so-called *chemostat*, where the added medium is designed such that there is a *single limiting substrate*. This allows for controlled variation in the specific growth rate of the biomass. The advantage of the continuous bioreactor is that a steady state can be achieved, which allows for precise experimental determination of specific rates. Furthermore, by varying the feed flow rate to the bioreactor the environmental conditions can be varied, and valuable information concerning the influence of the environmental conditions on the cellular physiology can be obtained. The continuous bioreactor is attractive for industrial applications because the productivity can be high. However, often the titer (i.e., the product concentration) is lower than can be obtained in the fed-batch reactor, and it is therefore a trade-off between productivity and titer. Furthermore, it is rarely used in industrial processes because it is sensitive to contamination (e.g., via the feed stream) and to the appearance of spontaneously formed mutants that may outcompete the production strain. Other examples of continuous operation besides the chemostat are the *pH-stat*, where the feed flow is adjusted to maintain the pH constant in the bioreactor, and the *turbidostat*, where the feed flow is adjusted to maintain the biomass concentration at a constant level.

From the biomass mass balance (3.40), it can easily be seen that in a steady-state continuous reactor, the specific growth rate equals the dilution rate:

$$\mathrm{m} = D \tag{3.50}$$

Thus, by varying the dilution rate (or the feed-flow rate) in a continuous culture, different specific growth rates can be obtained. This allows detailed physiological studies of the cells when they are grown at a predetermined specific growth rate (corresponding to a certain environment experienced by the cells). At steady state, the substrate mass balance (3.38) gives

$$0 = -r_s x + D\left(c_s^f - c_s\right) \tag{3.51}$$

which, upon combination with Equation 3.50 and the definition of the yield coefficient, directly gives

$$x = Y_{sx}\left(c_s^f - c_s\right) \tag{3.52}$$

Thus, the yield coefficient can be determined from measurement of the biomass and the substrate concentrations in the bioreactor (the substrate concentration in the feed flow should generally be known as it is determined in the setup of the experiment).

Besides the advantage for obtaining steady-state measurements, the chemostat is well suited to study dynamic conditions because it is possible to perform well-controlled transients. Thus, it is possible to study the cellular response to a sudden increase in the substrate concentration by adding a pulse of the limiting substrate to the reactor or to a sudden change in the dilution rate. These experiments both start and end with a steady state, so the initial and end conditions are well characterized, and this facilitates the interpretation of the cellular response. One type of transient experiment is especially suited to determining an important kinetic parameter, namely, the maximum specific growth rate. By increasing the dilution rate to a value above μ_{max}, the cells will wash out from the bioreactor and the substrate concentration will increase (and eventually reach the same value as in the feed). After adaptation of the cells to the new conditions, they will attain their maximum specific growth rate and the dynamic mass balance for the biomass becomes

$$\frac{dx}{dt} = (\mu_{max} - D)x \tag{3.53}$$

or

$$\frac{x(t-t_0)}{x(t_0)} = \exp((\mu_{max} - D)(t-t_0)) \tag{3.54}$$

where t_0 is the time at which the cells have become adapted to the new conditions and grow at their maximum specific growth rate. Thus, the maximum specific growth rate is easily determined from a plot of the biomass concentration versus time on a semilog plot.

3.3.4 Fed-Batch Reactor

This operation is probably the most common in industrial processes, because it allows for control of the environmental conditions, for example maintaining the glucose concentration at a certain level, as well as enabling the formation of very high titers (up to several hundred grams per liters of some metabolites), which is important for subsequent downstream processing. For a fed-batch reactor, the mass balances for biomass and substrate are given by Equations 3.38 and 3.40. Normally, the feed concentration is very high (i.e., the feed is a very concentrated solution) and the feed flow is low, giving a low dilution rate.

For the fed-batch reactor, the dilution rate is given by

$$D = \frac{1}{V}\frac{dV}{dt} \tag{3.55}$$

To keep D constant, there needs to be an exponentially increasing feed flow to the bioreactor, which is normally practically impossible as it may lead to oxygen limitations. The feed flow is therefore adjusted or increased until limitations in the oxygen supply set in, at which point the feed flow is kept constant. This will give a decreasing specific growth rate. However, because the biomass

concentration usually increases, the volumetric uptake rate of substrates (including oxygen) may be kept approximately constant. From the above it is quite clear that there may be many different feeding strategies in a fed-batch process, and optimization of the operation is a complex problem that is difficult to solve empirically. Even when a very good process model is available, calculation of the optimal feeding strategy is a complex optimization problem. In an empirical search for the optimal feeding policy, the two most obvious criteria are

Keep the concentration of the limiting substrate constant.
Keep the volumetric growth rate of the biomass (or uptake of a given substrate) constant.

A constant concentration of the limiting substrate is often applied if the substrate inhibits product formation, and the chosen concentration is therefore dependent on the degree of inhibition and the desire to maintain a certain growth of the cells. A constant volumetric growth rate (or uptake of a given substrate) is applied if there are limitations in the supply of oxygen or in heat removal.

Fed-batch cultures were used in the production of baker's yeast as early as 1915. The method was introduced by Dansk Gæringsindustri and is therefore sometimes referred to as the Danish method. It was recognized that an excess of malt in the medium would lead to a higher growth rate, resulting in an oxygen demand in excess of what could be met in the fermentors. This resulted in the development of respiratory catabolism of the yeast, leading to ethanol formation at the expense of biomass production. The yeast was allowed to grow in an initially weak medium to which additional medium was added at a rate less than the maximum rate at which the organism could use it. In modern fed-batch processes for yeast production, the feed of molasses is under strict control, based on the automatic measurement of traces of ethanol in the exhaust gas of the bioreactor. Although such systems may result in low specific growth rates, the biomass yield is generally close to the maximum obtainable, and this is especially important in the production of baker's yeast, where there is much focus on the yield. Apart from the production of baker's yeast, the fed-batch process is used today for the production of secondary metabolites (where penicillin is a prominent group), industrial enzymes, and many other products derived from cultivation processes.

3.4 KINETIC MODELS

Kinetic modeling expresses the verbally or mathematically expressed correlation between rates and reactant or product concentrations that, when inserted into the mass balances derived in Section 3.3, permit a prediction of the degree of conversion of substrates and the yield of individual products at other operating conditions. If the rate expressions are correctly set up, it may be possible to express the course of an entire fermentation experiment based on initial values for the components of the state vector (e.g., concentration of substrates). This leads to simulations, which may finally result in an optimal design of the equipment or an optimal mode of operation for a given system. The basis of kinetic modeling is to express functional relationships between the forward reaction rates **v** of the reactions considered in the model and the concentrations of the substrates, metabolic products, biomass constituents, intracellular metabolites, and/or biomass:

$$v_i = f_i(c_s, c_p, X_{macro}, X_{met}, x) \tag{3.56}$$

If, during the cultivation, the biomass composition remains constant, then the rates of the internal reactions must necessarily be proportional. This is referred to as **balanced growth**. In this case the growth process can be described in terms of a single variable that defines the state of the biomass. This variable is quite naturally chosen as the biomass concentration x (g DW l^{-1}). This is the basis of the so-called **unstructured models** that have proved adequate during 50 years of practical application to design cultivation processes (especially steady-state or batch cultivations), to install suitable

control devices, and to estimate which process conditions are likely to give the best return on the investment in process equipment. However, these unstructured models generally have poor predictive strength and as such are of little value in fundamental studies of cellular function.

3.4.1 Degree of Model Complexity

A typical discussion on the complexity of mathematical modeling of biochemical systems may be initiated by asking the question of whether a mechanistic model or an empirical model should be applied instead. To illustrate this, consider the fractional saturation y of a protein at a ligand concentration c_l. This may be described by the Hill equation (Hill 1910):

$$y = \frac{c_l^h}{c_l^h + K} \tag{3.57}$$

where h and K are empirical parameters. Alternatively, the fractional saturation may be described by the equation of Monod et al. (1965):

$$y = \frac{\left(La\left(1+\frac{ac_l}{K_R}\right)^3 + \left(1+\frac{c_l}{K_R}\right)^3\right)\frac{c_l}{K_R}}{L\left(1+\frac{ac_l}{K_R}\right)^4 + \left(1+\frac{c_l}{K_R}\right)^4} \tag{3.58}$$

where L, a, and K_R are parameters. Both equations address the same experimental problem, but whereas Equation 3.57 is completely empirical with h and K as fitted parameters, Equation 3.58 is derived from a hypothesis for the mechanism; the parameters therefore have a direct physical interpretation. If the aim of the modeling is to understand the underlying mechanism of the process, Equation 3.57 can obviously not be applied because the kinetic parameters are completely empirical and give no (or little) information about the ligand binding to the protein. In this case, Equation 3.58 should be applied, because by estimating the kinetic parameter the investigator is supplied with valuable information about the system and the parameters can be directly interpreted.

If, on the other hand, the aim of the modeling is to simulate the ligand binding to the protein, Equation 3.57 may be as good as Equation 3.58—and this equation may even be preferable because it is simpler in structure, has fewer parameters, and actually often gives a better fit to experimental data than Equation 3.57. Thus, the answer to which model is preferred depends on the aim of the modeling exercise. The same can be said about the unstructured growth models (Section 3.4.2), which are completely empirical but are valuable for extracting key kinetic parameters for growth. Furthermore, they are well suited to simple design problems and for teaching.

If the aim is to simulate dynamic growth conditions, one may turn to simple structured models (Section 4.4.3), for example the compartment models, which are also useful for an illustration of structured modeling in the classroom. However, if the aim is to analyze a given system in further detail, it is necessary to include far more structure in the model. In this case one often describes only individual processes within the cell, such as a certain pathway or gene transcription from a certain promoter. Similarly, if the aim is to investigate the interaction between different cellular processes (e.g., the influence of a plasmid copy number on chromosomal DNA replication), a single-cell model (Section 3.4.4) has to be applied.

Finally, if the aim is to look into population distributions, which in some cases may have an influence on growth or production kinetics, either a segregated or a morphologically structured model has to be applied (Section 3.5).

3.4.2 Unstructured Models

Even when there are many substrates, one of these substrates is usually limiting (i.e., the rate of biomass production depends exclusively on the concentration of this substrate). At low concentrations c_s of this substrate μ is proportional to c_s, but for increasing values of c_s an upper value μ_{max} for the specific growth rate is gradually reached. This is the verbal formulation of the Monod (1942) model:

$$\mu = \mu_{max} \frac{c_s}{c_s + K_s} \tag{3.59}$$

which has been shown to correlate fermentation data for many different microorganisms. In the Monod model, K_s is that value of the limiting substrate concentration at which the specific growth rate is half its maximum value. Roughly speaking, it divides the μ versus c_s plot into a low substrate concentration range where the specific growth rate is strongly (almost linearly) dependent on c_s, and a high substrate concentration range where μ is independent of c_s.

When glucose is the limiting substrate, the value of K_s is normally in the micromolar range (corresponding to the mg l^{-1} range), and it is therefore experimentally difficult to determine. Some of the K_s values reported in the literature are compiled in Table 3.5. It should be stressed that the K_s value in the Monod model does not represent the saturation constant for substrate uptake but an overall saturation constant for the entire growth process.

Some of the most characteristic features of microbial growth are represented quite well by the Monod model:

The constant specific growth rate at high substrate concentration
The first-order dependence of the specific growth rate on substrate concentration at low substrate concentrations

In fact, one may argue that the two features that make the Monod model work so well in fitting experimental data are deeply rooted in any naturally occurring conversion process: the size of the machinery that converts substrate must have an upper value, and all chemical reactions

TABLE 3.5
Compilation of K_s values *for Sugars*

Species	Substrate	K_s (mg l^{-1})
Aerobacter aerogenes	Glucose	8
Escherichia coli	Glucose	4
Klebsiella aerogenes	Glucose	9
	Glycerol	9
Klebsiella oxytoca	Glucose	10
	Arabinose	50
	Fructose	10
Lactococcus cremoris	Glucose	2
	Lactose	10
	Fructose	3

Source: Values are taken from Nielsen, J. and Villadsen, J., *Bioreaction Engineering Principles*. New York: Plenum Press, 1994.

will end up as first-order processes when the reactant concentration tends to zero. The satisfactory fit of the Monod model to many experimental data should never be misconstrued to mean that Equation 3.59 is a mechanism of fermentation processes. The Langmuir rate expression of heterogeneous catalysis and the Michaelis–Menten rate expression in enzymatic catalysis are formally identical to Equation 3.59, but the denominator constant has a direct physical interpretation in both cases (the equilibrium constant for dissociation of a catalytic site–reactant complex) whereas K_s in Equation 3.59 is no more than an empirical parameter used to fit the average substrate influence on all cellular reactions pooled into the single reaction by which substrate is converted to biomass.

In the Monod model, it is assumed that the yield of biomass from the limiting substrate is constant; in other words, there is proportionality between the specific growth rate and the specific substrate uptake rate. The Monod model is, however, normally used together with a maintenance consumption of substrate, that is, the specific substrate uptake is described by the linear relation; see Equation 3.18. The Monod model including maintenance is probably the most widely accepted model for microbial growth, and it is well suited for analysis of steady-state data from a chemostat (see Example 3c). Often the model is combined with the Luedeking and Piret (1959) model for metabolite production in which the specific rate of product formation is given by Equation 3.17. The Luedeking and Piret model was derived on the basis of an analysis of lactic acid fermentation and is in principle only valid for metabolic products formed as a direct consequence of the growth process (i.e., metabolites of primary metabolism). However, the model may in some cases be applied to other products (e.g., secondary metabolites), but this should not be done automatically.

Example 3c: The Monod Model

Despite its simplicity, the Monod model is very useful for extracting key growth parameters, and it generally fits simple batch fermentations with one exponential growth phase and steady-state chemostat cultures (but rarely with the same parameters). We first consider a batch process, where substrate consumption due to maintenance can usually be neglected. In this case, there is an analytical solution to the mass balances for the concentrations of substrate and the limiting substrate (Nielsen and Villadsen 1994):

$$c_s = c_{s,0} - Y_{xs}\left(x - x_0\right) \tag{3c1}$$

$$\mu_{max} t = \left(1 + \frac{K_s}{c_{s,0} + Y_{xs} x_0}\right) \ln\left(\frac{x}{x_0}\right) - \frac{K_s}{c_{s,0} + Y_{xs} x_0} \ln\left(1 + \frac{Y_{sx}\left(x_0 - x\right)}{c_{s,0}}\right) \tag{3c2}$$

Using this analytical solution, it is in principle possible to estimate the two kinetic parameters in the Monod model, but since K_s generally is very low it is in practice not possible to estimate this parameter from a batch cultivation (see Figure 3.3).

For a steady-state, continuous culture, the mass balance for the biomass, together with the Monod model, gives

$$D = \mu_{max} \frac{c_s}{c_s + K_s} \tag{3c3}$$

or

$$c_s = \frac{D K_s}{\mu_{max} - D} \tag{3c4}$$

Thus, the concentration of the limiting substrate increases with the dilution rate. When substrate concentration becomes equal to the substrate concentration in the feed, the dilution rate attains its maximum value, which is often called the *critical dilution rate*:

$$D_{crit} = \mu_{max} \frac{c_s^f}{c_s^f + K_s} \tag{3c5}$$

When the dilution rate becomes equal to or larger than this value, the biomass is washed out of the bioreactor. Equation 3b4 clearly shows that the steady-state chemostat is well suited to studying the influence of the substrate concentration on the cellular function (e.g., product formation), because by changing the dilution rate it is possible to change the substrate concentration as the only variable. Furthermore, it is possible to study the influence of different limiting substrates on the cellular physiology (e.g., glucose and ammonia).

Besides quantification of the Monod parameters, the chemostat is well suited to determine the maintenance coefficient. Because the dilution rate equals the specific growth rate, the yield coefficient is given by

$$Y_{sx} = \frac{D}{Y_{xs}^{true} D + m_s} \tag{3c6}$$

or, if we use Equation 3.52,

$$x = \frac{D}{Y_{xs}^{true} D + m_s} \left(c_s^f - c_s \right) \approx \frac{D}{Y_{xs}^{true} D + m_s} c_s^f \tag{3c7}$$

because $c_s^f >> c_s$ except for dilution rates close to the critical dilution rate. Equation 3c7 shows that the biomass concentration decreases at low specific growth rates, where the substrate consumption for maintenance is significant compared with that for growth. At high specific growth rates (high dilution rates), maintenance is negligible and the yield coefficient becomes equal to the true yield coefficient (Figure 3.4). By rearrangement of Equation 3c7, a linear relationship between the specific substrate uptake rate and the dilution rate is found, and when using this, the true yield coefficient and the maintenance coefficient can easily be estimated using linear regression.

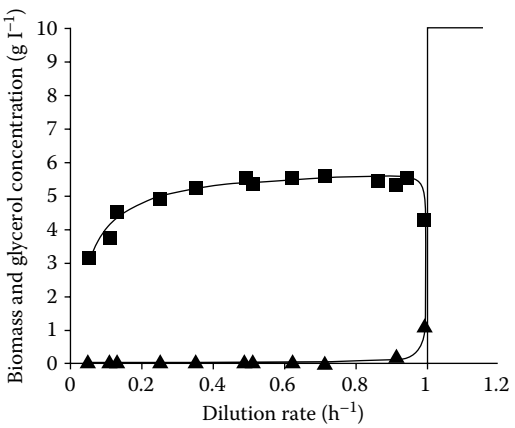

FIGURE 3.4 Growth of *Aerobacter aerogenes* in a chemostat with glycerol as the limiting substrate where the biomass concentration (■) decreases for increasing dilution rate due to the maintenance metabolism, and when the dilution rate approaches the critical value the biomass concentration decreases rapidly. The glycerol concentration (▲) increases slowly at low dilution rates, but when the dilution rate approaches the critical value it increases rapidly. The lines are model simulations using the Monod model with maintenance, and with the parameter values: $c_s^f = 10$ g L^{-1}; $\mu_{max} = 1.0$ h^{-1}; $K_S = 0.01$ g L^{-1}; $ms = 0.08$ g (g DW h)$^{-1}$; $Y_{xs}^{true} = 1.70$ g (g DW)$^{-1}$.

It is unlikely that the Monod model can be used to fit all kinds of fermentation data. Many authors have tried to improve on the Monod model, but generally these empirical models are of little value. However, in some cases growth is limited either by substrate concentration or by the presence of a metabolic product, which acts as an inhibitor. In order to account for this, the Monod model is often extended with additional terms. Thus, for inhibition by high concentrations of the limiting substrate

$$\mu = \mu_{max} \frac{c_s}{c_s^2/K_i + c_s + K_s} \tag{3.60}$$

and for inhibition by a metabolic product

$$\mu = \mu_{max} \frac{c_s}{c_s + K_s} \frac{1}{1 + p/K_i} \tag{3.61}$$

Equations 3.60 and 3.61 may be a useful way of including product or substrate inhibition in a simple model, and often these expressions are also applied in connection with structured models. Extension of the Monod model with additional terms or factors should, however, be carried out with some restraint because the result may be a model with a large number of parameters but of little value outside the range in which the experiments were made.

3.4.3 Compartment Models

Simple structured models are in one sense improvements to the unstructured models, because some basic mechanisms of the cellular behavior are at least qualitatively incorporated. Thus, the structured models may have some predictive strength, that is, they may describe the growth process at different operating conditions with the same set of parameters. But one should bear in mind that "true" mechanisms of the metabolic processes are of course not considered in simple structured models even if the number of parameters is quite large.

In structured models all the biomass components are lumped into a few key variables (i.e., the vectors $\mathbf{X}_{macro}$ and $\mathbf{X}_{met}$), which are hopefully representative of the state of the cell. The microbial activity thus becomes a function of not only the biotic variables, which may change with very small time constants, but also the cellular composition, and consequently the "history" of the cells (i.e., the environmental conditions the cells have experienced in the past).

The biomass can be structured in a number of ways. For example, in simple structured models only a few cellular components are considered, whereas in highly structured models up to 20 intracellular components are considered (Nielsen and Villadsen 1992). As discussed in Section 3.4.1, the choice of a particular structure depends on the aim of the modeling exercise, but often one starts with a simple structured model onto which more and more structures are added as new experiments are added to the database. Even in highly structured models many of the cellular components included in the model represent "pools" of different enzymes, metabolites, or other cellular components. The cellular reactions considered in structured models are therefore empirical in nature because they do not represent the conversion between "true" components. Consequently, it is permissible to write the kinetics for the individual reactions in terms of reasonable empirical expressions, with a form judged to fit the experimental data with a small number of parameters. Thus, Monod-type expressions are often used because they summarize some fundamental features of most cellular reactions (i.e., being first order at low substrate concentration and zero order at high substrate concentration). Despite their empirical nature, structured models are normally based on some well-known cell mechanisms, and they therefore have the ability to simulate certain features of experiments quite well.

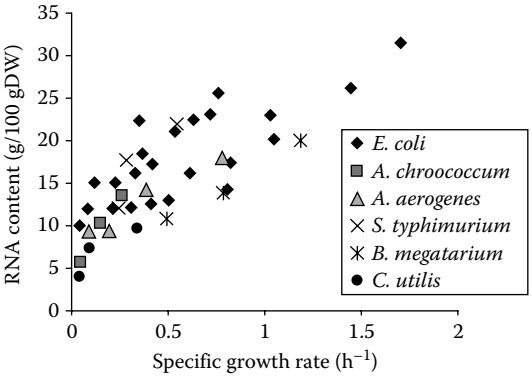

FIGURE 3.5 The level of stable RNA as a function of specific growth rate for different micro-organisms. (Data are taken from Nielsen, J. and J. Villadsen., *Bioreaction Engineering Principles*, New York: Plenum Press, 1994.)

The first structured models appeared in the late 1960s from the group of Fredrickson and Tsuchiya at the University of Minnesota (Ramkrishna et al. 1966, 1967; Williams 1967), who also were the first to formulate microbial models within a general mathematical framework similar to that used to describe reaction networks in classical catalytic processes (Tsuchiya et al. 1966; Fredrickson et al. 1967; Fredrickson 1976). Since this pioneering work, many other simple structured models have been presented (Harder and Roels 1982; Nielsen and Villadsen 1992).

In these simple structured models, the biomass is divided into a few compartments. These compartments must be chosen with care, and cell components with similar functions should be placed in the same compartment, for example, all membrane material and otherwise rather inactive components in one compartment, and all active material in another compartment. With the central role of the protein-synthesizing system (PSS) in cellular metabolism, this is often a key component in compartment models. Besides a few enzymes, the PSS consists of ribosomes, which contain approximately 60% ribosomal RNA (rRNA) and 40% ribosomal protein. Because rRNA makes up more than 80% of the total stable RNA in the cell, the level of the ribosomes is easily identified through measurements of the RNA concentration in the biomass. As seen in Figure 3.5, the RNA content of many different microorganisms increases linearly with the specific growth rate. Thus, the level of the PSS is well correlated with the specific growth rate. It is therefore a good representative of the activity of the cell, and this is the basis of most simple structured models (see Example 3d).

Example 3d: Two-Compartment Model

Nielsen *et al.* (1991a, 1991b) presented a two-compartment model for the lactic acid bacterium *Lactococcus cremoris*. The model is a direct descendant of the model created by Williams (1967) with similar definitions for the following two compartments:

The active (A) compartment contains the PSS and small building blocks.
The structural and genetic (G) compartment contains the rest of the cell material.

The model considers both glucose and a complex nitrogen source (peptone and yeast extract), but in the following presentation we discuss the model with only one limiting substrate (glucose). The model considers two reactions for which the stoichiometry is

$$\gamma_{11} X_A - s = 0 \tag{3d1}$$

$$\gamma_{22} X_G - X_A = 0 \tag{3d2}$$

In the first reaction, glucose is converted into small building blocks in the A compartment, and these are further converted into ribosomes. The stoichiometric coefficient γ_{11} can be considered as a yield coefficient because metabolic products (lactic acid, carbon dioxide, etc.) are not included in the stoichiometry. In the second reaction, building blocks present in the A compartment are converted into macromolecular components of the G compartment. In this process, some by-products may be formed and the stoichiometric coefficient γ_{22} is therefore slightly less than 1. The kinetics of the two reactions have the same form:

$$v_i = k_i \frac{c_s}{c_s + K_{s,i}} X_A; \quad i = 1,2 \tag{3d3}$$

From Equation 3.13, the specific growth rate for the biomass is found to be

$$\mu = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \gamma_{11} & -1 \\ 0 & \gamma_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \gamma_{11} v_1 - (1 - \gamma_{22}) v_2 \tag{3d4}$$

or, with the kinetic expression for v_1 and v_2 inserted,

$$\mu = \left(\gamma_{11} k_1 \frac{c_s}{c_s + K_{s,1}} - (1 - \gamma_{22}) k_2 \frac{c_s}{c_s + K_{s,2}} \right) X_A \tag{3d5}$$

Thus, the specific growth rate is proportional to the size of the active compartment. The substrate concentration c_s influences the specific growth rate both directly and indirectly by determining the size of the active compartment. The influence of the substrate concentration on the synthesis of the active compartment can be evaluated through the ratio r_1/r_2:

$$\frac{r_1}{r_2} = \frac{k_1}{k_2} \frac{c_s + K_{s,2}}{c_s + K_{s,1}} \tag{3d6}$$

If $K_{s,1}$ is larger than $K_{s,2}$, the formation of X_A is favored at high substrate concentration, and it is thus possible to explain the increase in the active compartment with the specific growth rate. Consequently, when the substrate concentration increases rapidly, there are two effects on the specific growth rate:

- First a rapid increase in the specific growth rate, which is a result of mobilization of excess capacity in the cellular synthesis machinery.
- Thereafter, there is a slow increase in the specific growth rate, which is a result of a slow buildup of the active part of the cell (i.e., additional cellular synthesis machinery has to be formed in order for the cells to grow faster).

This is illustrated in Figure 3.6, which shows the biomass concentration in two independent wash-out experiments. In both cases the dilution rate was shifted to a value (0.99 h^{-1}) above the critical dilution rate (0.55 h^{-1}), but in one experiment the dilution rate before the shift was low (0.1 h^{-1}) and in the other experiment it was high (0.5 h^{-1}). The wash-out profile is seen to be very different, with a much faster wash-out when there was a shift from a low dilution rate. When the dilution rate is shifted to 0.99 h^{-1}, the glucose concentration increases rapidly to a value much higher than $K_{s,1}$ and $K_{s,2}$, and this allows growth at the maximum rate. However, when the cells come from a low dilution rate, the size of the active compartment is not sufficiently large to allow rapid growth, and XA therefore has to be built up before the maximum specific growth rate is attained. On the other hand, if the cells come from a high dilution rate, XA is already high and the cells immediately attain their maximum specific growth rate. It is observed that the model is

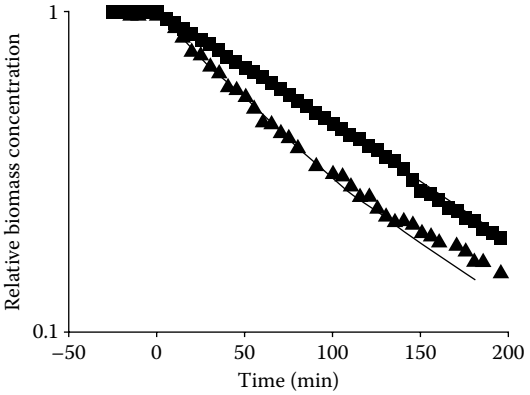

FIGURE 3.6 Measurement of biomass concentration, two transient experiments, of *L. cremoris* growing in glucose-limited chemostat; the dilution rate was shifted from an initial value of 0.10 h^{-1} (▲) or 0.50 h^{-1} (■) to 0.99 h^{-1}, respectively. The biomass concentration is normalized by the steady-state biomass concentration before the step change, which was made at time zero. The lines are model simulations. (The data are taken from Nielsen, J., Nikolajsen, K., and Villadsen, J., *Biotechnol. Bioeng.*, 38:11–23, 1991b.)

able to correctly describe the two experiments (all parameters were estimated from steady-state experiments), and the model correctly incorporates information about the previous history of the cells.

The model also includes the formation of lactic acid; the kinetics was described using a rate equation similar to Equation 3d3. Thus, the lactic acid formation increases when the activity of the cells increases, and so it is ensured that there is a close coupling between the formation of this primary metabolite and the growth of the cells.

It is interesting to note that even though the model does not include a specific maintenance reaction, it can actually describe a decrease in the yield coefficient of biomass on glucose at low specific growth rates. The yield coefficient is given by

$$Y_{sx} = \gamma_{21}\left(1-(1-\gamma_{22})\frac{k_2}{k_1}\frac{c_s+K_{s,1}}{c_s+K_{s,2}}\right) \tag{3d7}$$

Because $K_{s,1}$ is larger than $K_{s,2}$, the last term within the parentheses decreases for increasing specific growth rates, and the yield coefficient will therefore also increase for increasing substrate concentration.

3.4.4 Single-Cell Models

Single-cell models are in principle an extension of the compartment models, but with the description of many different cellular functions. Furthermore, these models depart from the description of a population and focus on the description of single cells. This allows consideration of characteristic features of the cell, and it is therefore possible to study different aspects of cell function:

- Cell geometry can be accounted for explicitly, and so it is possible to examine its potential effects on nutrient transport.
- Temporal events during the cell cycle can be included in the model, and the effect of these events on the overall cell growth can be studied.
- Spatial arrangements of intracellular events can be considered, even though this would lead to significant model complexity.

To set up single-cell models, it is necessary to have a detailed knowledge of the cell, and single-cell models have therefore only been described for well-studied cellular systems such as *Escherichia coli*, *S. cerevisiae*, *Bacillus subtilis*, and human erythrocytes. The most comprehensive single-cell model is the so-called Cornell model set up by Shuler and coworkers (Shuler et al. 1979), which contains 20 intracellular components. This model predicts a number of observations made with *E. coli,* and it has formed the basis for setting up several other models (Nielsen and Villadsen 1992). Thus, Peretti and Bailey (1987) extended the model to describe plasmid replication and gene expression from a plasmid inserted into a host cell. This allowed study of host–plasmid interactions, especially the effects of copy number, promoter strength, and ribosome binding site strength on the metabolic activity of the host cell and on the plasmid gene expression.

3.4.5 Molecular Mechanistic Models

Despite the level of detail, the single-cell models are normally based on an empirical description of different cellular events (e.g., gene transcription and translation). This is a necessity because the complexity of the model would become very high if all these individual events were to be described with detailed models that include mechanistic information. In many cases, however, it is interesting to study these events separately, and for models where mechanistic information is included, they have to be used. These models are normally set up at the molecular level, and they can therefore be referred to as *molecular mechanistic models*. Many different models of this type can be found in the literature, but most fall in one of two categories:

Gene transcription models
Pathway models

Gene transcription models aim at quantifying gene transcription based on knowledge of the promoter function. The *lac*-promoter of *E. coli* is one of the best studied promoter systems of all, and so this system has been modeled most extensively. Furthermore, this promoter (or its derivatives) is often used in connection with the production of heterologous proteins by this bacterium, because it is an inducible promoter. In a series of papers, Lee and Bailey (1984a, 1984b, 1984c, 1984d) presented an elegant piece of modeling of this system, and through combination with a model for plasmid replication they could investigate, for example, the role of point mutations in the promoter on gene transcription. This promoter system is quite complex, with both activator and repressor proteins, and empirical investigation can therefore be laborious; the detailed mathematical model is a valuable tool to guide the experimental work.

In pathway models the individual enzymatic reactions of a given pathway are described with enzyme kinetic models, and it is therefore possible to simulate the metabolite pool levels and the fluxes through different branches of the pathway. These models have mainly concentrated on glycolysis in *S. cerevisiae* (Galazzo and Bailey 1990; Rizzi et al. 1997), because much information about enzyme regulation is available for this pathway. However, complete models are also available for other pathways, such as the penicillin biosynthetic pathway (Pissarra et al. 1996). These pathway models are experimentally verified by comparing modeling simulations with measurements of intracellular metabolite pool levels – something that has only been possible with sufficient precision in the last couple of years, because it requires rapid quenching of the cellular activity and sensitive measurement techniques.

Pathway models are very useful in studies of metabolic fluxes, because they allow quantification of the control of flux by the individual enzymes in the pathway. This can be done by calculation of sensitivity coefficients (or the so-called *flux control coefficients*; see Box 3.4), which quantify the relative importance of the individual enzymes in the control of flux through the pathway (Stephanopoulos et al. 1998).

BOX 3.4 QUANTIFICATION OF FLUX CONTROL

In a study of flux control in a biochemical pathway the concept of metabolic control analysis is very useful (Stephanopoulos et al., 1998). Here the flux control of the individual enzymatic reactions on the steady state flux J through the pathway is quantified by the so-called flux control coefficients (FCC):

$$C_i^J = \frac{v_i}{J}\frac{dJ}{dv_i}$$

where v_i is the rate of the i^{th} reaction. If the enzyme concentration of the i^{th} enzyme is increased its rate will normally increase, and a higher flux through the pathway may be the result. However, it is likely that due to allosteric regulation (or other regulation phenomena) there may be a very small effect of increasing the enzyme concentration. This is exactly what is quantified by the FCC, that is the relative increase in the steady state flux upon a relative increase in the enzyme activity. Clearly a step with a high FCC has a large control of the flux through the pathway. If a kinetic model is available for the pathway the flux control coefficients for each step can easily be calculated using model simulations. The FCCs can also be determined experimentally by changing the enzyme concentration (or activity) genetically, by titration with the individual enzymes, or by adding specific enzyme inhibitors (Stephanopoulos et al. 1998).

A general criticism of the application of kinetic models for complete pathways is that despite the level of detail included, they cannot possibly include all possible interactions in the system and therefore only represent one model of the system. The robustness of the model is extremely important especially if the kinetic model is to be used for predictions, and unfortunately most biochemical models, even very detailed models, are only valid at operating conditions close to those where the parameters have been estimated (i.e., the predictive strength is limited). For analysis of complex systems it is, however, not necessary that the model gives a quantitatively correct description of all the variables, because even models that give a qualitatively correct description of the most important interactions in the system may be valuable in studies of flux control.

3.5 POPULATION MODELS

Normally it is assumed that the population of cells is homogeneous (i.e., all cells behave identically). Although this assumption is certainly crude if a small number of cells is considered, it gives a very good picture of certain properties of the cell population because there are billions of cells per ml medium (see Box 3.5).

Furthermore, the kinetics is often linear in the cellular properties (e.g., in the concentration of a certain enzyme), and the overall population kinetics can therefore be described as a function of the average property of the cells (Nielsen et al. 2003). There are, however, situations where cell property distributions influence the overall culture performance, and here it is necessary to consider the cellular property distribution, and this is done in the so-called *segregated models*. In the following we will discuss two approaches to segregated modeling.

3.5.1 Morphologically Structured Models

The simplest approach to model distribution in the cellular property is by the so-called *morphologically structured models* (Nielsen and Villadsen 1992; Nielsen 1993). Here the cells are divided into a finite number Q of cell states Z (or morphological forms), and conversion between the different

BOX 3.5 DETERMINISTIC VERSUS STOCHASTIC MODELING

In a description of cellular kinetics macroscopic balances are normally used, that is the rates of the cellular reactions are functions of average concentrations of the intracellular components. However, living cells are extremely small systems with only a few molecules of certain key components, and it does not really make sense to talk about 'the DNA concentration in the cell' for example, because the number of macromolecules in a cell is always small compared with Avogadro's number. Many cellular processes are therefore stochastic in nature and the deterministic description often applied is in principle not correct. However, the application of a macroscopic (or deterministic) description is convenient and it represents a typical engineering approximation for describing the kinetics in an average cell in a population of cells. This approximation is reasonable for large populations because the standard deviation from the average 'behavior' in a population with elements is related to the standard deviation for an individual cell through

$$\sigma_{pop} = \frac{\sigma}{\sqrt{e}}$$

Thus with a population of 10^9 cells mL^{-1}, which is a typical cell concentration during a cultivation process, it can be seen that the standard deviation for the population is very small. There are, however, systems where small populations occur, for example at dilution rates close to the maximum in a chemostat, and here one may have to apply a stochastic model.

cell states is described by a sequence of empirical metamorphosis reactions. Ideally these metamorphosis reactions can be described as a set of intracellular reactions, but the mechanisms behind most morphological conversions are largely unknown. Thus, it is not known why filamentous fungi differentiate to cells with a completely different morphology from that of their origin. It is therefore not possible to set up detailed mechanistic models describing these changes in morphology; empirical metamorphosis reactions are therefore introduced. The stoichiometry of the metamorphosis reactions is given by analogy with Equation 3.4:

$$\Delta \mathbf{Z} = \mathbf{0} \tag{3.62}$$

where Δ is a stoichiometric matrix. Z_q represents both the q^{th} morphological form and the fractional concentration (g q^{th}, morphological form (g DW^{-1}). With the metamorphosis reactions, one morphological form is spontaneously converted to other forms. This is of course an extreme simplification because the conversion between morphological forms is the sum of many small changes in the intracellular composition of the cell. With the stoichiometry in Equation 3.62, it is assumed that the metamorphosis reactions do not involve any change in the total mass, and the sum of all stoichiometric coefficients in each reaction is therefore taken to be zero. The forward reaction rates of the metamorphosis reactions are collected in the vector **u**. Each morphological form may convert substrates to biomass components and metabolic products. These reactions may be described by an intracellularly structured model, but in order to reduce the model complexity a simple unstructured model is used for description of the growth and product formation of each cell type (e.g., the specific growth rate of the q^{th} morphological form is described by the Monod model). The specific growth rate of the total biomass is given as a weighted sum of the specific growth rates of the different morphological forms:

$$\mu = \sum_{i=1}^{Q} \mu_i Z_i \tag{3.63}$$

The rate of formation of each morphological form is determined both by the metamorphosis reactions and by the growth-associated reactions for each form (for derivation of mass balances for the morphological forms, see Nielsen and Villadsen 1994). The concept of morphologically structured models is well suited to describing the growth and differentiation of filamentous microorganisms (Nielsen 1993), but it may also be used to describe other microbial systems where a cellular differentiation has an impact on the overall culture performance.

3.5.2 Population Balance Equations

The first example of a heterogeneous description of cellular populations was presented in 1963 by Fredrickson and Tsuchiya. In their model, single-cell growth kinetics was combined with a set of stochastic functions describing cell division and cell death. The model represents the first application of a completely segregated description of a cell population. In the model, the cell population is described by a number density function $f(X,t)$, where $f(X,t)\mathrm{d}X$ is the number of cells with property X being in the interval X to $X + \mathrm{d}X$. The dynamic balance for $f(X,t)$ is given by

$$\frac{\partial f(X,t)}{\partial t} + \frac{\partial}{\partial X}(f(X,t)v(X,t)) = 2\int_X^\infty b(X^*,t)\,p(X^*,X,t)\,f(X^*,t)\,dX^* - b(X,t)f(X,t) - Df(X,t) \tag{3.64}$$

where v is the net rate of formation of the cell property, X. $b(X,t)$ is the breakage function (i.e., the rate of cell division for cells with property X), and $p(X^*,X,t)$ is the partitioning function (i.e., the probability that a cell with property X is formed upon division of a cell with property X^*). Through the functions p and b, a stochastic element can be introduced into the model, but these functions can also be completely descriptive. The balance equation 3.64 was applied in the original work of Fredrickson and Tsuchiya, but in a later paper a general framework for segregated population models was presented (Fredrickson *et al.* 1967). Segregated models represent the complete description of a cell population, and they take into account that all cells in a population are not identical. However, complete cellular segregation is rarely applied in cell culture models for two main reasons:

- For large populations, the average properties will normally represent the overall population kinetics quite well.
- The mathematical complexity of Equation 3.64 is quite substantial, especially if more than a single-cell property is considered (i.e., the number density function becomes multidimensional).

If the kinetics for product formation is not zero or first order in a given cell property, application of an average property model will, however, not give the same result as a segregated description. This is the case for production of a heterologous protein in plasmid-containing cells of *E. coli*, where the product formation kinetics is not first order in the plasmid copy number. A segregated model therefore has to be applied to give a good description of the product formation kinetics (Seo and Bailey 1985). The simplest segregated models are when the cellular property is described by a single variable (e.g., cell age), and in Example 3e, the age distribution of an exponentially growing culture is derived from the general balance (Equation 3.64).

Example 3e: Age Distribution Model

The simplest segregated population models are those where the cellular property is taken to be described solely by the cell age a. In this case, the rate of increase in the cellular property $v(a,t)$ is

equal to 1. Furthermore, if it is assumed that cell division occurs only at a certain cell age $a = t_d$, the two first terms on the right-hand side of Equation 3.64 become equal to zero. At steady state, the balance therefore becomes

$$\frac{d\phi(a)}{da} = -D\phi(a) \tag{3e1}$$

where ϕ is a normalized distribution function:

$$\phi(a) = \frac{f(a)}{n} \tag{3e2}$$

with n being the total cell number, given as the zero moment of the number density function $f(a)$. The solution to Equation 3e1 is

$$\phi(a) = \phi(0)e^{-Da} \tag{3e3}$$

Due to the normalization, the 0th moment of $f(a)$ is 1, that is,

$$\int_o^{t_d} \phi(a)da = 1 \tag{3e4}$$

which leads to

$$\phi(a) = \frac{D}{1-e^{-Dt_d}}e^{-Da} \tag{3e5}$$

The cell balance relating to cell division (the so-called renewal equation) is given by

$$\phi(0) = 2\phi(t_d) \tag{3e6}$$

which, together with Equation 3e3, directly gives Equation 3.2. Furthermore, when Equation 3.2 is inserted in Equation 3e5, we have the simpler expression

$$\phi(a) = 2De^{-Da} \tag{3e7}$$

Thus, the fraction of cells with a given age decreases exponentially with age, and the decrease is determined by the specific growth rate of the culture (equal to the dilution rate at steady state). The average cell age is given as the first moment of $f(a)$:

$$\langle a \rangle = \int_0^{t_d} a\phi(a)da = \frac{1-\ln 2}{D} \tag{3e8}$$

Consequently, the average age of the cells decreases for increasing specific growth rates.

SUMMARY

Understanding microbial growth kinetics is an essential requirement for the design and successful operation of industrial fermentation processes and for obtaining quantitative information about the function of microbial cells.

The primary objective of this chapter is to introduce the reader to the basic principles of the wide-ranging aspects of microbial growth kinetics and dynamics, from the basic principles to the more advanced concept of modeling.

Based on the information given in this chapter, the reader should be able to design fermentation processes for the production of biomass and microbially derived products.

Furthermore, the reader should be able to set up simple mathematical models describing microbial growth as well as evaluate more complex mathematical models.

REFERENCES

Benthin, S., U. Schulze, J. Nielsen, and J. Villadsen. 1994. Growth energetics of *Lactococcus cremoris* FDI during energy, carbon and nitrogen limitation in steady state and transient cultures. *Chem. Eng. Sci*. 49:589–609.

Christensen, L.H., C.M. Henriksen, J. Nielsen, J. Villadsen, and M. Egel-Mitani. 1995. Continuous cultivation of *P. chrysogenum*: Growth on glucose and penicillin production. *J. Biotechnol*. 42:95–107.

Fredrickson, A.G. 1976. Formulation of structured growth models. *Biotechnol. Bioeng*. 18:1481–6.

Fredrickson, A.G., D. Ramkrishna, and H.M. Tsuchiya. 1967. Statistics and dynamics of procaryotic cell populations. *Math. Biosci*. 1:327–74.

Fredrickson, A.G. and H.M. Tsuchiya. 1963. Continuous propagation of micro-organisms. *AIChE J*. 9:459–68.

Galazzo, J.L. and J.E. Bailey. 1990. Fermentation pathway kinetics and metabolic ßux control in suspended and immobilized *Saccharomyces cerevisiae*. *Enzym. Microb. Technol*. 12: 162–72.

Harder, A. and J.A. Roels. 1982. Application of simple structured models in bioengineering. *Adv. Biochem. Eng*. 21:55–107.

Herbert, D. 1959. Some principles of continuous culture. *Recent Prog. Microbiol.* 7:381–96.

Hill, A.V. 1910. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J. Physiol. Lond*. 40:4–7.

Ingraham, J.L., O. Maaloe, and F.C. Neidhardt. 1983. *Growth of the Bacterial Cell*. Sunderland, MA: Sinauer Associates.

Lee, S.B. and J.E. Bailey. 1994a. A mathematical model for λdv plasmid replication: analysis of wild-type plasmid. *Gene* 11:151–65.

Lee, S.B. and J.E. Bailey. 1994b. A mathematical model for λdv plasmid replication: Analysis of copy number mutants. *Gene* 11:166–77.

Lee, S.B. and J.E. Bailey. 1994c. Genetically structured models for *lac* promoter-operator function in the *Escherichia coli* chromosome and in multicopy plasmids: *lac* operator function. *Biotechnol. Bioeng*. 26:1372–82.

Lee, S.B. and J.E. Bailey. 1994d. Genetically structured models for *lac* promoter-operator function in the *Escherichia coli* chromosome and in multicopy plasmids: *lac* promoter function. *Biotechnol. Bioeng*. 26:1383–9.

Luedeking, R. and E.L. Piret. 1959. A kinetic study of the lactic acid fermentation: Batch process at controlled pH. *J. Biochem, Microbiol. Technol. Eng*. 1:393–412.

Meyenburg, K. von. 1969. Katabolit-Repression und der Sprossungszyklus von *Saccharomyces cerevisiae*. PhD dissertation, ETH, Zürich.

Monod, J. 1942. *Recherches sur la Croissance des Cultures Bacteriennes*. Paris: Hermann and Cie.

Monod, J., J. Wyman, and J-P. Changeux. 1965. On the nature of allosteric transitions: A plausible model. *J. Molec. Biol*. 12:88–118.

Neidhardt, F.C., J.L. Ingraham, and M. Schaechter. 1990. *Physiology of the Bacterial Cell: A Molecular Approach*. Sunderland, MA: Sinauer Associates.

Nielsen, J. 1993. A simple morphologically structured model describing the growth of Þlamentous microorganisms. *Biotechnol, Bioeng*. 41:715–27.

Nielsen, J., K. Nikolajsen, and J. Villadsen. 1991a. Structured modeling of a microbial system 1: A theoretical study of the lactic acid fermentation. *Biotechnol. Bioeng*. 38:1–10.

Nielsen, J., K. Nikolajsen, and J. Villadsen. 1991b. Structured modelling of a microbial system 2: Verification of a structured lactic acid fermentation model. *Biotechnol. Bioeng*. 38:11–23.

Nielsen, J. and J. Villadsen. 1992. Modeling of microbial kinetics. *Chem. Eng. Sci*. 47:4225–70.

Nielsen, J. and J. Villadsen. 1993. Bioreactors: Description and modelling. In Rehm, H.J., and Reed, G., eds., *Biotechnology*, vol. 3, 2nd ed., pp. 77–104. Weinheim: VCR Verlag.

Nielsen, J. and J. Villadsen. 1994. *Bioreaction Engineering Principles*. New York: Plenum Press.

Nielsen, J. J. Villadsen, and G. Lidén. 2003. *Bioreaction Engineering Principles*, 2nd ed. New York: Kluwer Academic/Plenum.

Olsson, L. and J. Nielsen. 1997. On-line and in situ monitoring of biomass in submerged cultivations. *TIBTECH* 15:517–22.

Oura, E. 1983. Biomass from carbohydrates. In Rehm, H-J., and Reed, G., eds., *Biotechnology*, vol. 3, 2nd ed., pp. 3–42. Weinheim: VCR Verlag.

Peretti, S.W. and J.E. Bailey. 1987. Simulations of host–plasmid interactions in *Escherichia coli*: Copy number, promoter strength, and ribosome binding site strength effects on metabolic activity and plasmid gene expression. *Biotechnol. Bioeng*. 29:316–28.

Pirt, S.J. 1965. The maintenance energy of bacteria in growing cultures. *Proc. Roy. Soc. London*, *Ser. B*. 163:224–31.

Pissarra, P.N., J. Nielsen, and M.J. Bazin. 1996. Pathway kinetics and metabolic control analysis of a high-yielding strain of *Penicillium chrysogenum* during fed-batch cultivations. *Biotechnol. Bioeng*. 51:168–76.

Ramkrishna, D., A.G. Fredrickson, and H.M. Tsuchiya. 1966. Dynamics of microbial propagation: Models considering endogenous metabolism. *J. Gen. Appl. Microbiol*. 12:311–27.

Ramkrishna, D., A.G. Fredrickson, and H.M. Tsuchiya. 1967. Dynamics of microbial propagation: Models considering inhibitors and variable cell composition. *Biotechnol. Bioeng*. 9:129–70.

Rizzi, M., M. Baltes, U. Theobald, and M. Reuss. 1997. In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol. Bioeng*. 55:592–608.

Roels, J.A. 1983. *Energetics and Kinetics in Biotechnology*. Amsterdam: Elsevier Biomedical Press.

Seo, J.H. and J.E. Bailey. 1985. A segregated model for plasmid content and product synthesis in unstable binary fission recombinant organisms. *Biotechnol. Bioeng*. 27:156–65.

Shuler, M.L., S.K. Leung, and C.C. Dick. 1979. A mathematical model for the growth of a single bacterial cell. *Ann. NY Acad. Sci*. 326:35–55.

Sonnleitner, B. and A. Fiechter. 1988. High performance bioreactors: A new generation. *Anal. Chim. Acta* 213:199–205.

Stephanopoulos, G., J. Nielsen, and A. Aristodou. 1998. *Metabolic Engineering.* San Diego, CA: Academic Press.

Tsuchiya, H.M., A.G. Fredrickson, and R. Aris. 1966. Dynamics of microbial cell populations. *Adv. Chem. Eng*. 6:125–206.

Williams, F.M. 1967. Av model of cell growth dynamics. *J. Theor. Biol*. 15:190–207.