Biochemical Engineering Balances and Black Box Models of Cells

Jeffrey D. Varner*

School of Chemical and Biomolecular Engineering*
Cornell University, Ithaca NY 14853

Introduction

The basic building block of any biomolecular process is a cell. Cells are biological agents which consume nutrients, and process these nutrients to make valuable products, waste products and more cells. The chemical transformation of nutrients e.g., sugar to valuable products such as *proteins* or organic molecules uses a vast network of coupled chemical reactions collectively called *metabolic pathways*. Cells differ in their size, shape, behavior and complexity. However, they all share the common thread of requiring nutrients to survive, and many of the pathways that process nutrients are conserved from the simplest bacteria to the most complex cells in our bodies. We can grow cells and use them to make products of interest in special chemical reactors called *bioreactors*. To understand how cells function in bioreactors, and ultimately how to manipulate them for societal benefit, we must first understand how to apply engineering principles such as conservation of mass and energy to biological systems. Toward this goal, we'll start by reviewing basic material balance concepts and use these concepts to write balance equations around *extracellular* nutrients (nutrients outside of the cell). Second, we'll begin to think about how cells process extracellular nutrients, and how to write material balances around cells in bioreactors.

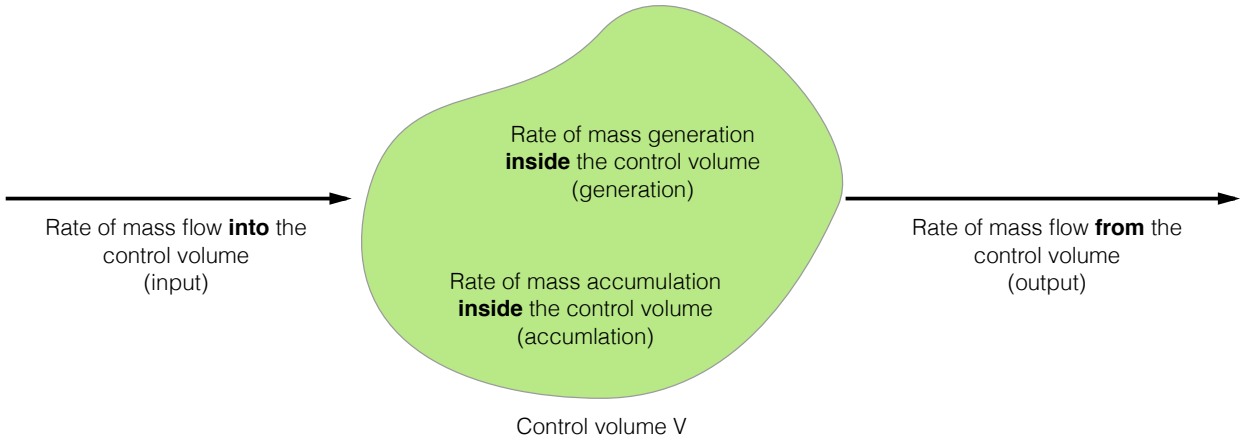

Fig. 1: Schematic of an idealized control volume V. Total macroscopic material balance equations consist of four types of terms. Material flow into and from the control volume, material generation inside the control volume and material accumulation inside the control volume. The abundance of material can be measured using both a mass or mole basis.

Total macroscopic mass and mole balance equations Consider an idealized control volume (Fig. 1). The total macroscopic material balance around the control volume contains four terms, flow in, flow out, generation ($\dot{m}_{gen}$) and accumulation ($\dot{m}_{acc}$):

$$\sum_{s\,=\,1}^{\mathcal{S}} v_s \dot{m}_s + \dot{m}_{gen} = \dot{m}_{acc} \tag{1}$$

The summation term:

$$\sum_{s\,=\,1}^{\mathcal{S}} v_s \dot{m}_s \tag{2}$$

describes the net rate of material flow into and from the control volume, where $\dot{m}_s$ denotes the mass flow *rate* of stream s (typical units of kg/hr) and $\mathcal{S}$ denotes the total number of streams. The quantities v_s are direction parameters. We'll use the convention that streams entering the control volume have positive direction parameters ($v_s = 1$), while streams exiting the control volume (s^*) have negative direction parameters ($v_{s^*} = -1$), The term $\dot{m}_{gen}$ describes the *rate* of total mass generation inside the control volume, while $\dot{m}_{acc}$ describes the *rate* of total mass accumulation inside the control volume. For biochemical processes, total mass is conserved (it is neither created or destroyed), thus, the total mass generation term $\dot{m}_{gen} = 0$.

Just like total mass balances, we can also write a balance equation around the *total* number of moles in a process. Total mole balances have the same four types of terms, input, output, accumulation and generation and take the form:

$$\sum_{s=1}^{\mathcal{S}} v_s \dot{n}_s + \dot{n}_{gen} = \dot{n}_{acc} \tag{3}$$

where $\dot{n}_s$ denotes the rate of transport of total moles in stream s (mol/time), $\dot{n}_{gen}$ denotes the rate of generation of total moles in a process (mol/time), and $\dot{n}_{acc}$ denotes the rate of accumulation of moles inside the process (mol/time), and v_s denotes our direction parameter. However, unlike total mass balances, the generation term in mole balances can be non-zero ($\dot{n}_{gen} \neq 0$) in some cases (especially for reactive systems). We can also write species balances using moles. For example, the balance around species k is given by:

$$\sum_{s=1}^{\mathcal{S}} v_s \dot{n}_s x_{k,s} + \dot{n}_{k,gen} = \dot{n}_{k,acc} \qquad k = 1, 2, \ldots, \mathcal{M} \tag{4}$$

where $x_{k,s}$ denotes the mole fraction of species k in stream s. If stream s is a vapor or gas, we use the symbol $y_{k,s}$ to denote the mole fraction of species k in stream s. We can also express mole-based species balances in terms of the concentration of species k (C_k) in the control volume:

$$\sum_{s=1}^{\mathcal{S}} v_s F_s C_{k,s} + \dot{n}_{k,gen} = \frac{d}{dt}\left(C_k V\right) \qquad k = 1, 2, \ldots, \mathcal{M} \tag{5}$$

where F_s denotes the volumetric flow rate (L/hr) in stream s, $C_{k,s}$ denotes the concentration of species k in stream s, and V denote the volume (L) of the control volume. We can write a balance of the form shown in Eqn. (5) for each of the $\mathcal{M}$ species in the bioreactor.

Example 1 (Volume of culture media in a well stirred bioreactor)

Starting from the general total mass balance equation, derive the volume (V) of culture media in a well mixed bioreactor. Let ρ_s denote the density of liquid in stream s, and ρ denote the density of culture media in the reactor.

Solution: The general total mass balance is given by:

$$\sum_{s=1}^{\mathcal{S}} v_s \dot{m}_s + \dot{m}_{gen} = \dot{m}_{acc} \tag{6}$$

We know there is no mass generation for total mass, thus $\dot{m}_{gen} = 0$. We also know that we can write the mass of liquid in the reactor in terms of the density and volume:

$$m = \rho V \tag{7}$$

which means that we can write the accumulation terms as:

$$\dot{m}_{acc} = \frac{d}{dt}\left(\rho V\right) = V\frac{d\rho}{dt} + \rho\frac{dV}{dt} \tag{8}$$

Substituting the accumation term into the general balance gives:

$$\rho\frac{dV}{dt} = \sum_{s=1}^{\mathcal{S}} v_s \dot{m}_s - V\frac{d\rho}{dt} \tag{9}$$

Often it is more convient to write the mass flow rates in terms of volumetric flow rates and the density of each stream:

$$\dot{m}_s = \rho_s F_s \tag{10}$$

where F_s is the volumtric flow rate for stream s (L/hr). Substituting $\dot{m}_s$ into Eqn (9) gives a final expression governing the Volume of media in the reactor:

$$\frac{dV}{dt} = \sum_{s=1}^{\mathcal{S}} v_s \frac{\rho_s}{\rho} F_s - \frac{V}{\rho}\frac{d\rho}{dt} \tag{11}$$

Black box models of growing cells Metabolic pathways convert extracellular nutrients, for example sugar, oxygen or nitrogen etc into cells, valuable products and waste products. When we

write balances describing the abundance of sugar, nitrogen or other metabolic inputs we should include descriptions of the metabolic pathways in our balances equations. However, exhaustively modeling metabolic pathways is difficult; metabolic pathways consist of thousands of coupled chemical reactions, occurring simultaneously inside growing cells. Thus, it is currently only feasible to completely model metabolism (the process of the breakdown of raw materials and construction of finished products) for very simple organisms (1). Instead, we'll make a simplifying assumption and treat cells as black boxes which consumes nutrients, produce more cells, and metabolic products (later, we'll relax this assumption). Consider the schematic of a simple cell shown in Fig. 2.

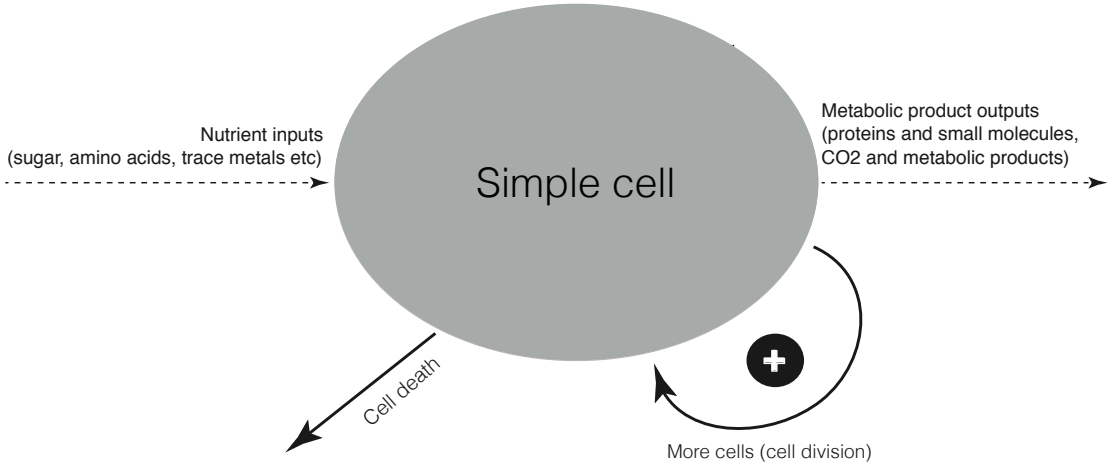

Fig. 2: Schematic of simple cells. Simple cells consume nutrients (carbon, nitrogen, etc), and process them to make metabolic products, waste products and more cells. The dashed arrows denote catalytic behavior that simple cells do while the solid arrows indicate reactions that change the abundance of cells in a bioreactor.

Simple cells consume nutrients $S_1, S_1, \ldots, S_N$ to produce products $P_1, P_2, \ldots, P_M$ and cells X. In reality, this conversion involves many chemical reactions, however for now lets lump all of these reactions into the single reaction:

$$\sum_{i=1}^{N} Y_{XS_i} S_i \xrightarrow{\mu} X + \sum_{j=1}^{M} Y_{XP_j} P_j \tag{12}$$

where the Y_{Xj} are called *yield coefficients*. Yield coefficients are a special type of stoichiometric coefficient that relates the consumption or production of a nutrient or product to the cell growth rate. For example, if species j were a nutrient, Y_{Xj} would quantify how much j was consumed to product a unit of cells. In the general case, yield coefficients are *not constant* (we'll see examples of this later) and must be measured from experimental data, or calculated using theoretical tools

such as flux balance analysis (FBA) (2). Just like other stoichiometric coefficients, yield coefficients have both a direction, and a magnitude; for reactants $Y_{Xi} < 0$, while for products $Y_{Xj} > 0$. The rate at which Eqn. (12) occurs is called the *specific growth rate* and is given the special symbol μ. There are many different models for the functional form of μ (i.e., how the growth rate depends upon the concentration of nutrients), and we'll consider several of these throughout the semester.

Extracellular material balance equations. To analyze the behavior of cells in a bioreactor, we need to write material balances around nutrients, cells, metabolic products and the volume of media in the bioreactor. For bioreactors we'll exclusively use mole-based units, however, other mass-based units systems can be found in the biochemical literature. Let's begin writing our balances by expanding the generation term in the general mole-based species balance given in Eqn (5):

$$\dot{n}_{j,gen} = \left(\sum_{r}^{\mathcal{R}} \sigma_{jr}\hat{r}_r\right)V + \left(\sum_{k}^{\mathcal{T}} \tau_{j,k}q_k\right)XV \tag{13}$$

We have two sets of reaction terms, the first describes $\mathcal{R}$ chemical reactions that can occur in the absence of cells (in the bulk fluid in the reactor), while the second describes those $\mathcal{T}$ reactions that occur because of cells (cell-associated reactions). For bulk fluid phase reactions, σ_{jr} denotes the stoichiometric coefficient governing species j in reaction r; if $\sigma_{jr} < 0$ species j is *consumed* by reaction r, if $\sigma_{jr} > 0$ species j is *produced* by reaction r and if $\sigma_{jr} = 0$ species j is not connected with reaction r. The quantity $\hat{r}_r$ denotes the reaction rate per unit volume for reaction r (mmol/L-hr). Similarly, for the cell-associated reaction terms, $\tau_{j,k}$ denotes the stoichiometric coefficient describing how species j is connected with cell-associated reaction k. Reactions associated with cells have a unique unit system called *specific units* which means we write all quantities per unit cell abundance (grams dry weight, or mmol of cells). Thus, q_k, the kth cell-associated reaction rate, has units of mmol/mmol-hr or mmol/gdw-hr etc. Substituting our reaction terms into Eqn (5) gives:

$$\frac{d}{dt}\left(C_jV\right) = \sum_{s}^{N} v_sF_sC_{j,s} + \left(\sum_{r}^{\mathcal{R}} \sigma_{jr}\hat{r}_r\right)V + \left(\sum_{k}^{\mathcal{T}} \tau_{j,k}q_k\right)XV \qquad j = 1, 2, \ldots, \mathcal{M} \tag{14}$$

where $\mathcal{M}$ denotes the number of *metabolites*, X denotes the cellmass level in the reactor (gdw/L or mmol/L), and V denotes the volume (L) of the reactor. Similarly, the cellmass balance is given by:

$$\frac{d}{dt}\left(XV\right) = \sum_{s} v_sF_sX_s + \left(\mu - k_d\right)XV \tag{15}$$

where μ denotes the *specific growth rate* of cells, (hr^{-1}) and k_d denotes the cell death constant (hr^{-1}). Lastly, both the cellmass and metabolite balances involve the volume V which is governed by:

$$\frac{dV}{dt} = \sum_{s\,=\,1}^{\mathcal{S}} v_s\frac{\rho_s}{\rho}F_s - \frac{V}{\rho}\frac{d\rho}{dt} \tag{16}$$

Putting all three types of equations together, expanding all derivatives and dividing both sides of the extracellular metabolite and cellmass balances by the volume gives:

$$\frac{dC_j}{dt} = \sum_{s=1}^{\mathcal{S}} v_s D_s C_{j,s} + \left(\sum_{r=1}^{\mathcal{R}} \sigma_{jr}\hat{r}_r\right) + \left(\sum_{k=1}^{\mathcal{T}} \tau_{j,k} q_k\right) X - \frac{C_j}{V}\frac{dV}{dt} \qquad j = 1, 2, \ldots, \mathcal{M} \quad (17)$$

$$\frac{dX}{dt} = \sum_{s=1}^{\mathcal{S}} v_s D_s X_s + (\mu - k_d)\, X - \frac{X}{V}\frac{dV}{dt} \quad (18)$$

$$\frac{dV}{dt} = \sum_{s=1}^{\mathcal{S}} v_s \frac{\rho_s}{\rho} F_s - \frac{V}{\rho}\frac{d\rho}{dt} \quad (19)$$

where the quantity D_s, called a *dilution rate* (hr^{-1}), is given as:

$$D_s \equiv \frac{F_s}{V} \qquad s = 1, 2, \ldots, \mathcal{S} \quad (20)$$

The cellmass, metabolite and volume balances are a coupled system of $\mathcal{M} + 2$ nonlinear ordinary differential equations, which depending upon the functional forms of the specific growth, death and uptake rates, has no analytic solution. However, these equations can be easily solved numerically using common algorithms included in packages and languages such as MATLAB or Python. To see this, let's consider three special cases of the bioreactor balances, batch (no flow), fed-batch (flow in, no flow out) and continuous (flow in = flow out).

What are the τ_* and q_*? The τ_* and q_* terms that appear in the extracellular species balances describe the stoichiometry and specific rate of consumption or production (mmol/gdw-L) of extracellular metabolites by simple cells. Lets assume that $\tau_* = -1$ for nutrients, while $\tau_* = 1$ for products. We can then relate the specific uptake rate of nutrients to the specific growth rate of simple cells using the Pirt equation (3):

$$q_k = \frac{1}{Y^*_{Xk}}\mu + m_k \qquad k \in \{nutrients\} \quad (21)$$

where Y^*_{Xk} is the *true* biomass yield on nutrient k, and m_k denotes the maintenance utilization of nutrient k. Similarly, we can use the Luedeking and Piret equation (4) to relate the specific rate of product formation to the specific growth rate:

$$q_j = \frac{1}{Y^*_{Xj}}\mu + \theta_j \qquad j \in \{products\} \quad (22)$$

where Y^*_{Xj} denotes the true product yield, and θ_j denotes the non-growth associated production of product j. For now, we'll only consider growth associated product formation, thus θ_j = 0.

Batch cultures $D_{in} = D_{out} = 0$. Lets consider growth of simple cells on a limiting nutrient S e.g., glucose in an idealized well-mixed stirred tank bioreactor. Denote the fluid-phase concentration of simple cells as X. Batch operation ($D_{in} = D_{out} = 0$) reduces our system of extracellular balances

to:

$$\frac{dS}{dt} = -\left(\frac{\mu}{Y_{XS}} + m_S\right) X \quad (23)$$

$$\frac{dX}{dt} = (\mu - k_d)\, X \quad (24)$$

$$\frac{dV}{dt} = 0 \quad (25)$$

The cellmass and substrate balances can be easily solved numerically using common algorithms included in packages such as MATLAB (Fig. 3). Cell mass growth occurs in three phases: lag, exponential and stationary/death phases. During the lag-phase, cells are manufacturing the proper systems internally to process the substrate. Once this adaptation phase is complete, cells begin to process substrate as quickly as possible which leads to a phase of maximum growth called the exponential phase. Once the substrate is exhausted, cells no longer divide, and growth stops leading to the stationary phase.

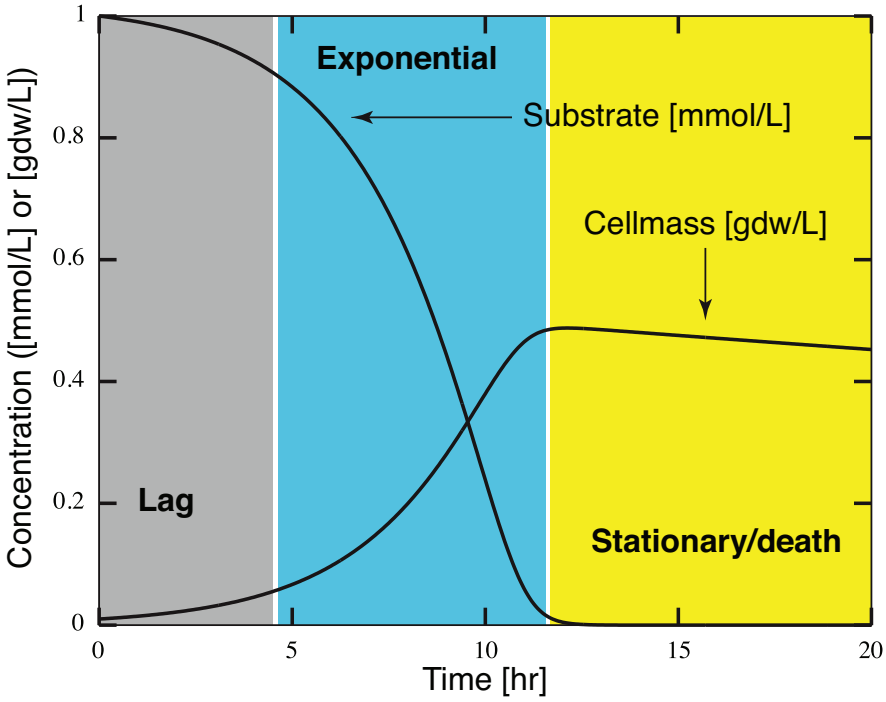

Fig. 3: Batch growth of simple cells in a well-mixed bioreactor as a function of time. The x-axis denotes the fermentation tile while the y-axis denotes the cellmass [gdw/L] or substrate [mmol/L] concentration.

Fed-batch cultures $D_{in} \neq 0, D_{out} = 0$. Lets consider growth of simple cells on a limiting nutrient S e.g., glucose in an idealized well-mixed stirred tank bioreactor. Denote the fluid-phase concentration of simple cells as X. Feb-batch (or sometimes also called semi-batch) operation is the typical mode that bioreactors are operated in industrially (however there are a few exceptions). Feb-batch requires $F_{in} \neq 0$ and $F_{out} = 0$, where the F_{in} trajectory, called the feeding profile, is

adjusted to maximize some target objectives e.g., the final product *titer* (the concentration of the desired product at the end of the run) or the *productivity* (the rate of product formation). In a fed-batch bioreactor the cellmass and the nutrient balances are given by:

$$\frac{dS}{dt} = D\left(S_{in} - S\right) - \left(\frac{\mu}{Y_{XS}} + m_S\right) X \tag{26}$$

$$\frac{dX}{dt} = \left(\mu - k_d\right) X - DX \tag{27}$$

$$\frac{dV}{dt} = F_{in} \tag{28}$$

where we have assumed sterile feed. The value of the feeding profile F_{in} can be a *difficult* function of time. Similar to batch growth, the cellmass/nutrient balances are a coupled nonlinear system of equations with no analytical solution (especially if F_{in} is time-dependent). However, we can also easily solve this system numerically, where a typical solution profile is given in Fig. 4.

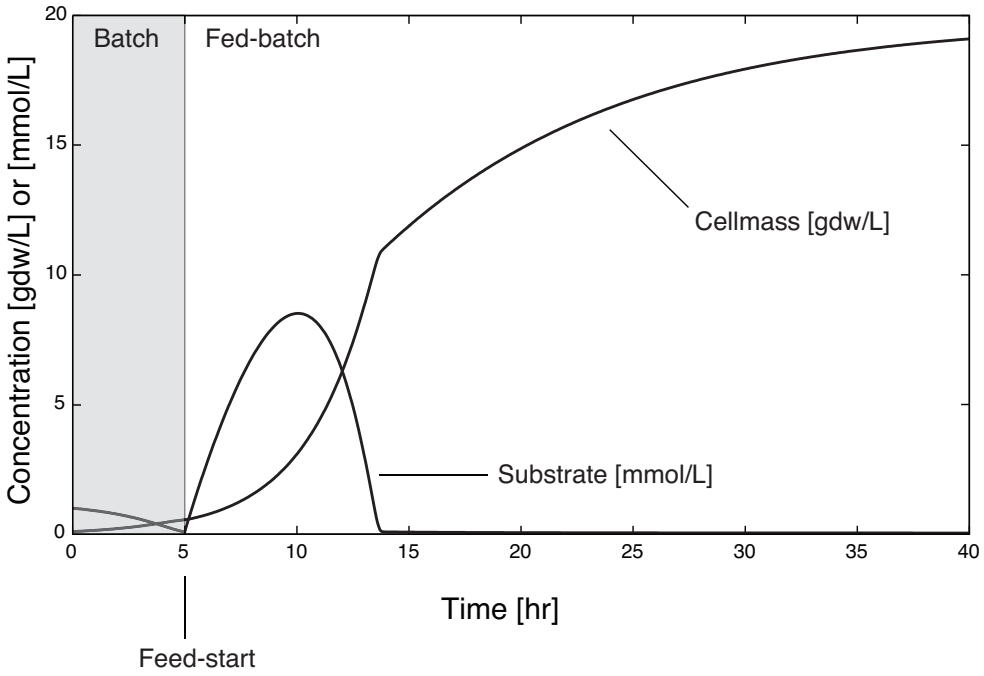

Fig. 4: Fed-batch growth of simple cells in a well-mixed bioreactor as a function of time. The x-axis denotes the fermentation tile while the y-axis denotes the cellmass [gdw/L] or substrate [mmol/L] concentration. The reactor is run in batch mode until the substrate is nearly exhausted, then the feed is started in the reactor. The feed-profile for this simulation is given by a simple exponential ramp $F_{in}\left(t\right) = F_o \exp\left(\alpha t\right)$.

Continuous cultures $D_{in} = D_{out} \equiv D$. Lastly, we consider the growth of simple cells on a limiting nutrient S e.g., glucose in a continuous idealized well-mixed stirred tank reactor. Denote the fluid-phase concentration of simple cells as X. The material balance governing the level of cellmass in

the reactor as a function of time is given by:

$$\frac{dS}{dt} = D\left(S_{in} - S_{out}\right) - \left(\frac{\mu}{Y_{SX}} + m_S\right) X - \frac{S}{V}\frac{dV}{dt} \quad (29)$$

$$\frac{dX}{dt} = D\left(X_{in} - X_{out}\right) + \left(\mu - k_d\right) X - \frac{X}{V}\frac{dV}{dt} \quad (30)$$

$$\frac{dV}{dt} = F_{in} - F_{out} \quad (31)$$

where V denotes the fluid-phase volume in the reactor, $F_j, j = in, out$ denotes the volumetric flow into and from the reactor, X_{in} denotes the cellmass abundance in the reactor feed, and μ and k_d denote the specific rate of growth and death, respectively. For continuous operation, $F_{in} = F_{out} \equiv F$, which implies $dV/dt = 0$. After an initial transient period directly following reactor startup, the bioreactor operates at a steady-state defined by the culture and reactor parameters (Fig. 5).

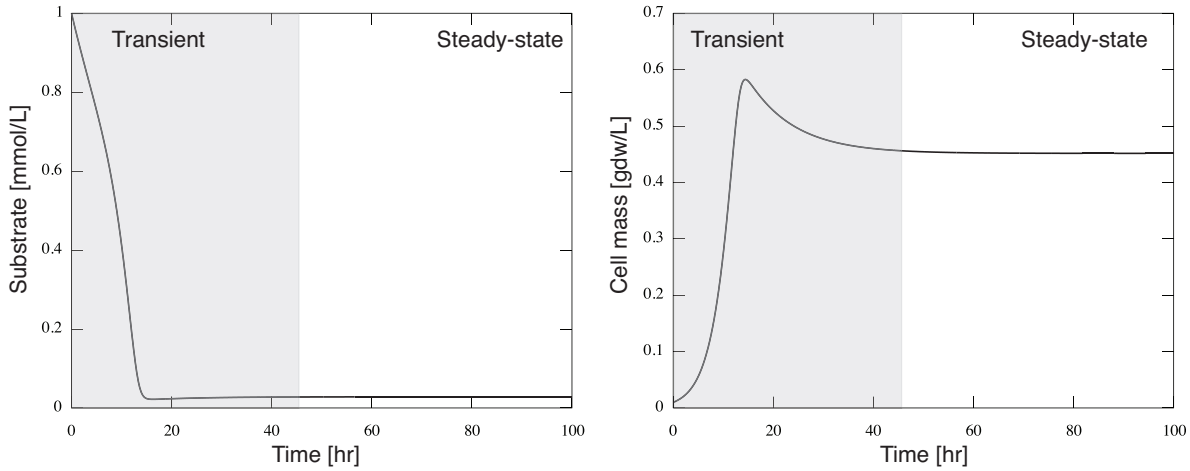

Fig. 5: Typical continuous culture time trajectories in a well mixed bioreactor. After an initial transient period, the substrate (left) and cell mass (right) trajectories come to a steady-state.

References

1. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, et al. (2012) A whole-cell computational model predicts phenotype from genotype. Cell 150: 389-401.
2. Orth JD, Thiele I, Palsson BØ (2010) What is flux balance analysis? Nat Biotechnol 28: 245-8.
3. Pirt SJ (1965) The maintenance energy of bacteria in growing cultures. Proc R Soc Lond B Biol Sci 163: 224-31.
4. Luedeking R, Piret EL (2000) A kinetic study of the lactic acid fermentation. batch process at controlled ph. reprinted from journal of biochemical and microbiological technology engineering vol. i, no. 4. pages 393-412 (1959). Biotechnol Bioeng 67: 636-44.